languages

Special Issue

Romance Languages at the Forefront of Language Acquisition Research

Guest Editors:

Prof. Juana M. Liceras
Department of Modern Languages and Literatures and Department of Linguistics, University of Ottawa, Canada

jliceras@uottawa.ca

Prof. Raquel Fernández Fuertes
Departamento de Filología Inglesa, Universidad de Valladolid, Spain

raquelff@lia.uva.es

Editor in Chief:
Prof. Usha Lakshmanan

Associate Editor:
Dr. Osmer Balam

Deadline for manuscript submissions:
31 December 2017

Message from the Guest Editors

Dear colleagues,

The goal of this Special Issue is to showcase state of the art work on the L1, bilingual and non-native acquisition of Romance languages from a theoretical and empirical perspective. The volume will examine how recent learnability issues are approached using acquisition data from different Romance languages. We particularly encourage contributions dealing with different populations, including but not limited to L1 acquisition, L2 acquisition, bi/multilingual and heritage language acquisition, language processing and language disorders.

The overall target length of this special volume would be 90,000 to 135,000 words with contributions that are between 8000 and 10,000 words in length; that is, 10 to 12 articles plus a solid introductory content article.

The Abstract submission deadline is 5 June 2017 (400-450 words including bibliography).

Please send your abstracts to languages@mdpi.com.

Prof. Juana M. Liceras
Prof. Raquel Fernández Fuertes
Guest Editors

Author Benefits

Open Access: free for readers, free publication for well-prepared manuscripts submitted in 2017
Rapid publication: accepted papers are immediately published online.

Special Issue "Romance Languages at the Forefront of Language Acquisition Research—Volume 1"

- Print Special Issue Flyer
- Special Issue Editors
- Special Issue Information
- Keywords
- Published Papers

A special issue of *Languages* (ISSN 2226-471X).

Deadline for manuscript submissions: **closed (31 December 2017)**.

Share This Special Issue

Special Issue Editors

Guest Editor
Prof. Dr. Juana M. Liceras
1) Department of Modern Languages and Literatures and Department of Linguistics, University of Ottawa, Canada
2) Departamento de Lenguas Aplicadas, Universidad Nebrija, Madrid, Spain
Website | E-Mail
Phone: 1-613-5625800 Ext. 3742

Interests: linguistic theory and language acquisition; Romance linguistics; bilingualism; migrant and refugee language; contact linguistics; code-switching; pidging and creole grammars; pedagogical grammar; bilingualism and non-typical language development

Guest Editor
Prof. Dr. Raquel Fernández Fuertes
Departamento de Filología Inglesa, Universidad de Valladolid, Spain
Website | E-Mail
Phone: 629864746

Interests: linguistic theory, comparative grammar and bilingual acquisition; language-in-contact situations (e.g., code-switching, crosslinguistic influence, natural translation, word order issues); spontaneous and experimental data analyses from simultaneous and sequential bilinguals

Show export options Displaying articles 1-13

Open Access Article
Non-Native Dialect Matters: The Perception of European and Brazilian Portuguese Vowels by Californian English Monolinguals and Spanish–English Bilinguals
Languages **2018**, *3*(3), 37; https://doi.org/10.3390/languages3030037
Received: 10 April 2018 / Revised: 30 August 2018 / Accepted: 3 September 2018 / Published: 12 September 2018
PDF Full-text (1356 KB) | HTML Full-text | XML Full-text

Open Access Feature Paper Article
The Lexical Development of Canadian-Born Romanian L1 Bilingual Kindergarteners
Languages **2018**, *3*(3), 33; https://doi.org/10.3390/languages3030033
Received: 1 February 2018 / Revised: 6 August 2018 / Accepted: 7 August 2018 / Published: 16 August 2018
PDF Full-text (1452 KB) | HTML Full-text | XML Full-text

Open Access Article
Brazilian Bimodal Bilinguals as Heritage Signers
Languages **2018**, *3*(3), 32; https://doi.org/10.3390/languages3030032
Received: 6 April 2018 / Revised: 5 August 2018 / Accepted: 6 August 2018 / Published: 10 August 2018
PDF Full-text (785 KB) | HTML Full-text | XML Full-text

Open Access Article
Mood Selection in Relative Clauses by French–Spanish Bilinguals: Contrasts and Similarities between L2 and Heritage Speakers
Languages **2018**, *3*(3), 31; https://doi.org/10.3390/languages3030031
Received: 20 February 2018 / Revised: 27 July 2018 / Accepted: 31 July 2018 / Published: 6 August 2018
PDF Full-text (1893 KB) | HTML Full-text | XML Full-text

Open Access Article
French Postverbal Subjects: A Comparison of Monolingual, Bilingual, Trilingual, and Multilingual French
Languages **2018**, *3*(3), 29; https://doi.org/10.3390/languages3030029
Received: 13 November 2017 / Revised: 23 July 2018 / Accepted: 24 July 2018 / Published: 28 July 2018
Cited by 1 | PDF Full-text (5418 KB) | HTML Full-text | XML Full-text

Open Access Article
Continuity in the Adult and Children's Comprehension of Subject and Object Relative Clauses in French and Italian
Languages **2018**, *3*(3), 24; https://doi.org/10.3390/languages3030024
Received: 5 January 2018 / Revised: 26 June 2018 / Accepted: 29 June 2018 / Published: 5 July 2018
PDF Full-text (816 KB) | HTML Full-text | XML Full-text

Open Access Article
Acquisition of French Causatives: Parallels to English Passives
Languages **2018**, *3*(2), 18; https://doi.org/10.3390/languages3020018
Received: 3 December 2017 / Revised: 28 March 2018 / Accepted: 22 May 2018 / Published: 31 May 2018
PDF Full-text (3401 KB) | HTML Full-text | XML Full-text

Open Access Article
Null Subject Occurrence in Monolingual Spanish SLI: A Discriminant Function Analysis
Languages **2018**, *3*(2), 17; https://doi.org/10.3390/languages3020017
Received: 14 November 2017 / Revised: 10 May 2018 / Accepted: 17 May 2018 / Published: 23 May 2018
PDF Full-text (793 KB) | HTML Full-text | XML Full-text

Open Access Article
Acquisition of L2 French Object Pronouns by Advanced Anglophone Learners
Languages **2018**, *3*(2), 15; https://doi.org/10.3390/languages3020015
Received: 22 December 2017 / Revised: 19 April 2018 / Accepted: 19 April 2018 / Published: 3 May 2018
PDF Full-text (862 KB) | HTML Full-text | XML Full-text

Open Access Article
Does Typological Proximity Really Matter? Evidence from Mandarin and Brazilian Portuguese-Speaking Learners of Spanish
Languages **2018**, *3*(2), 13; https://doi.org/10.3390/languages3020013
Received: 30 December 2017 / Revised: 12 April 2018 / Accepted: 12 April 2018 / Published: 25 April 2018
PDF Full-text (9502 KB) | HTML Full-text | XML Full-text

Open Access Article
The Mixed Effects of Phonetic Input Variability on Relative Ease of L2 Learning: Evidence from English Learners' Production of French and Spanish Stop-Rhotic Clusters
Languages **2018**, *3*(2), 12; https://doi.org/10.3390/languages3020012
Received: 9 January 2018 / Revised: 9 April 2018 / Accepted: 13 April 2018 / Published: 24 April 2018
PDF Full-text (2970 KB) | HTML Full-text | XML Full-text

Open Access Article
The Interpretation of Pronouns across Spanish-Speaking Populations
Languages **2018**, *3*(2), 11; https://doi.org/10.3390/languages3020011
Received: 24 December 2017 / Revised: 10 April 2018 / Accepted: 11 April 2018 / Published: 20 April 2018
PDF Full-text (1076 KB) | HTML Full-text | XML Full-text

Open Access Article
Language Mixing in the Nominal Phrase: Implications of a Distributed Morphology Perspective
Languages **2018**, *3*(2), 10; https://doi.org/10.3390/languages3020010
Received: 14 January 2018 / Revised: 6 April 2018 / Accepted: 9 April 2018 / Published: 12 April 2018
PDF Full-text (4897 KB) | HTML Full-text | XML Full-text |
Supplementary Files

languages

MDPI

Article

Non-Native Dialect Matters: The Perception of European and Brazilian Portuguese Vowels by Californian English Monolinguals and Spanish–English Bilinguals

Jaydene Elvin [1,2,*], Alba Tuninetti [2,3] and Paola Escudero [2,3]

1 Department of Linguistics, California State University, Fresno, 5245 North Backer Ave., M/S PB92, Fresno, CA 9374, USA
2 Australian Research Council (ARC) Centre of Excellence for the Dynamics of Language, The Australian National University, Building 9, H.C. Coombs Building, Canberra, ACT 2601, Australia; A.Tuninetti@westernsydney.edu.au (A.T.); paola.escudero@westernsydney.edu.au (P.E.)
3 The MARCS Institute for Brain, Behaviour & Development, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia
* Correspondence: jaydene@mail.fresnostate.edu; Tel.: +1-559-278-2349

Received: 10 April 2018; Accepted: 3 September 2018; Published: 12 September 2018

Abstract: Studies show that second language (L2) learners' perceptual patterns differ depending on their native dialect (e.g., Chládková and Podlipský 2011; Escudero and Williams 2012). Likewise, speakers from the same native language background show different perceptual patterns depending on the dialect to which they are exposed (e.g., Escudero and Boersma 2004; Escudero and Chládková 2010). The Second Language Linguistic Perception model (L2LP; Escudero 2005) accounts for these differences, explicitly stating that the acoustic similarity between the native and target dialect affects L2 perception. This study investigated whether Californian English monolingual and Spanish–English bilingual listeners differ in their perception of European Portuguese (EP) and Brazilian Portuguese (BP) vowels. Escudero et al. (2009a) showed that there were differences in the acoustic realization of vowels in BP and EP. Stressed vowels were longer in BP than in EP, with differences in vowel height observed for some vowels (e.g., /ɛ/ is higher in EP than in BP). According to the L2LP model, these acoustic differences between dialects will affect vowel perception; therefore, we predicted that there would be differences in the listeners' perception of certain vowel contrasts in BP and EP. Participants completed a non-native categorization task and a discrimination task presented in the XAB format. The results from the non-native categorization task predicted differential vowel perception depending on both the dialect and vowel contrast that listeners heard, which were mostly confirmed with an interaction between dialect and contrast in the discrimination results. We contextualize these results with respect to models of L2 speech perception, highlighting that dialectal differences impact language perception and may influence later language learning.

Keywords: non-native speech perception; dialectal differences; bilingual advantage; Brazilian Portuguese; European Portuguese

1. Introduction

It is well known that second language (L2) learners struggle to produce the sounds of their L2 in a native-like manner and research shows that this is largely due to the influence of their native language (L1) on the L2. In fact, a number of theoretical models of speech perception such as the Perceptual Assimilation Model (PAM; Best 1995) and the Second Language Linguistic Perception model (L2LP;

Escudero 2005) claim that a learner's ability to produce L2 sounds faithfully is largely dependent on their ability to perceive differences between the target sounds. In other words, a listener speaks with a foreign accent because they are listening with a foreign accent.

Given that learners rely on their L1 knowledge in order to perceive and acquire L2 sounds, cross-linguistic difficulty is not uniform across speakers of different language and dialect backgrounds. Thus, speakers with differing L1 backgrounds will differ in their perception of the same L2 speech sounds. In fact, according to the L2LP model (Escudero 2005; van Leussen and Escudero 2015), a learner's L2 perceptual patterns and difficulties are dependent on the cross-linguistic acoustic similarity between both the L1 dialect and the target L2. This suggests that difficulties in L2 speech perception are not necessarily language specific, but rather, dialectally specific. This claim that L2 perceptual patterns differ according to the native dialect is supported in a number of studies. For example, Moravian and Bohemian Czech listeners showed differential perceptual assimilation patterns of Dutch which seemed to correspond to differences in the acoustic realizations of their vowels in their native dialect (Chládková and Podlipský 2011). Likewise, Escudero and Williams (2012) showed that differences in the acoustic realizations of Iberian and Peruvian Spanish vowels led to differences in their discrimination accuracy of some Dutch vowel contrasts. In particular, Iberian Spanish learners had higher accuracy scores when perceiving differences in the Dutch /a/–/ɑ/ contrast. The authors explain that this finding is due to the fact that both vowels in the Dutch contrast are acoustically similar to one single native category in Peruvian Spanish, making them difficult to distinguish. However, these same two Dutch vowels were acoustically similar to two separate vowel categories in Iberian Spanish, which is likely to have influenced their overall higher accuracy when discriminating the vowels in this contrast.

Since differences in L2 speech perception are found in speakers of the same language but from different dialects, it should likewise be expected that speakers will differ in their perception of a language spoken in two different dialects. For example, Escudero and Boersma (2004) showed that Spanish listeners acquire the English /i/–/ɪ/ in different ways, depending on the target dialect they were learning, specifically, Scottish English and Southern-British English. Furthermore, Escudero and Chládková (2010) investigated Peruvian Spanish listeners' categorization of American English (AmE) and Standard Southern-British English (SSBE) vowels. The authors successfully predicted that the Peruvian listeners would categorize some AmE and SSBE vowels differently based on how those vowels compared acoustically to Peruvian Spanish. The authors concluded that based on the perceptual differences between the two English dialects, their perceptual development in each variety would likely differ. In another study, Baker and Smith (2010) investigated the differences between the perception and production of French vowels by learners of European (EF) and Quebec French (QF). The authors noted that there was an additional acoustic cue in QF for the distinction of the /i/–/y/–/u/ vowel contrast and they were interested to see whether learners of QF would better distinguish between these vowels than EF learners for both dialects. The authors indeed found that the QF learners distinguished these three vowels better than EF learners and concluded that the L2 dialect that one is exposed to does affect how well a learner will perceive and produce certain vowels.

To contribute to the emerging body of findings showing that learning of an L2[1] is affected by the dialect to which one is exposed, we investigated whether Californian English (CE) monolinguals and Spanish–English (SpE) bilinguals showed differences in their perception of Brazilian (BP) and European (EP) Portuguese vowels. BP and EP share a vowel system that contains seven oral vowels produced in a stressed position, namely /i, e, ɛ, a, ɔ, o, u/, which are described acoustically in Escudero et al. (2009a). Californian English is said to have 10 steady-state vowels, namely, /i, ɪ, e, ɛ, æ, ʌ, ɑ, o ʊ, u/ (Ladefoged 1999), while Mexican Spanish consists of only five steady-state vowels, /i, e, a, o, u/ (Díaz Granado 2011). We chose Portuguese as the target language, because it is relatively

[1] Throughout our study we refer to our participants' perception of BP and EP as non-native perception rather than L2 perception given that Portuguese is a non-native language for our participants as it is they do not have it as an L2 and they are not currently learning it.

understudied in the field of cross-linguistic speech perception. Although the authors of the study previously contributed to this field with regards to the investigation of Brazilian Portuguese, there remains a gap in the investigation of cross-linguistic influences on the perception of European Portuguese. Our choice to include bilingual participants is motivated by the fact that many studies show differences in bilingual versus monolingual language processing, including speech processing (see (Kroll and Bialystok 2013) for a review). Therefore, examining the differences in bilingual vs. monolingual speech perception and processing in the scope of the L2LP model not only adds information to the model itself, but also to the research on the bilingual advantage. We also chose these specific languages because Portuguese presents an interesting scenario for our language groups in that its vowel inventory falls in between Californian English and Mexican Spanish in regards to its size. Studies such as Iverson and Iverson and Evans (2007, 2009) claim that learners with a larger vowel inventory than the target language should be better able to perceive the L2 contrasts than those whose native vowel inventory is smaller than the target language. However, some studies (e.g., Elvin et al. 2014; Elvin 2016; Alispahic et al. 2017) found that this is not always the case. Therefore, we are interested in investigating the effect of the two vowel systems of differing sizes on bilinguals' vowel perception in addition to determining whether their results are comparable to those of monolinguals.

In order to predict CE and SpE listeners' perception of BP and EP, the L2LP model recommends a thorough cross-linguistic acoustic analysis of how the vowels in each L2 dialect compare to the vowels produced in the listeners' native dialect. All of the participants in this study were either Californian English monolingual or Mexican Spanish–English bilingual speakers from the Central Valley (California, USA). Ideally, acoustic comparisons should be based on the listener's own native productions to ensure the most accurate predictions (Escudero 2005; Elvin et al. 2016a); however, in the absence of such data, researchers can turn to acoustic studies of the native and target dialects with published formant values to make preliminary predictions. Given that there are no published studies with formant values for Californian English from the Central Valley, we used the published formant values of Southern-Californian English vowels presented in Hagiwara (1997). We also refer to a study conducted by Grijalva et al. (2013) which, in addition to describing the acoustic space of Southern-Californian English and Mexican Spanish, also provides formant values for Mexican Spanish spoken by bilingual Spanish–English speaker[2]. A visual plot of vowels produced in Southern-Californian English (Hagiwara 1997), Mexican Spanish produced by Spanish–English bilinguals (Grijalva et al. 2013), and Brazilian and European Portuguese (Escudero et al. 2009a) is shown in Figure 1. The speakers in Hagiwara (1997) study were 15 (six male) Southern-Californian English speaking monolinguals aged between 18 and 26. The vowels were extracted from words produced in the /bVt/, /tVk/, and /hVd/ consonantal contexts produced in a carrier phrase. The Spanish vowels from the Grijalva et al. (2013) study were extracted from words in a pVso context produced in a carrier phrase by 11 (three male) Mexican Spanish–English bilingual students from the University of California, San Diego. Finally, we refer to the formant values of BP and EP reported in (Escudero et al. 2009a), from which the stimuli for the present perception study were selected. The vowels were produced in a disyllabic CVCV, surrounded by two identical voiceless stops or fricatives (/p, t, k, f, s/), in a carrier phrase. These tokens were produced by 20 BP speakers from São Paulo and 20 EP speakers from Lisbon.

---

[2] Grijalva et al. (2013) only report the formant values for /i/, /e/, /a/, /o/, and /u/ by Southern-Californian English, Mexican Spanish, and English–Spanish bilingual speakers. For this reason, we report the formant values for all Southern-Californian English vowels from Hagiwara (1997).

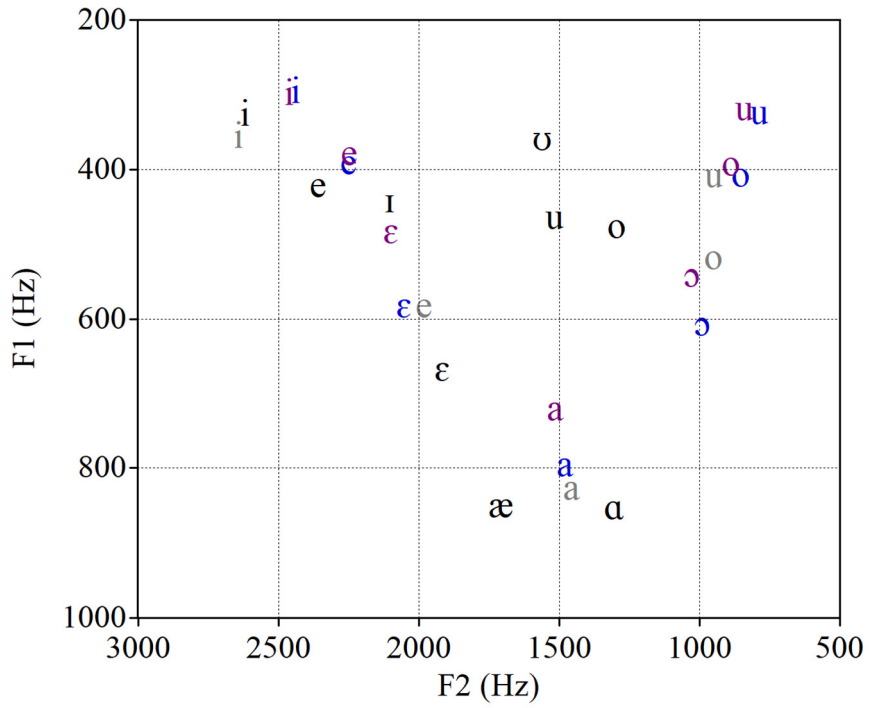


Figure 1. Southern-Californian English monolinguals' production of English (black; Hagiwara 1997) together with Californian English bilinguals' production of Mexican Spanish vowels (gray; Grijalva et al. 2013) compared against Brazilian Portuguese (BP; blue) and European Portuguese (EP; purple) vowels (Escudero et al. 2009a). Note: All formant values were averaged across male and female speakers.

Visual inspection of the vowels plotted in Figure 1 shows that, as reported in Escudero et al. (2009a), many of the vowels in BP and EP are acoustically similar to each other. However, we do observe that EP /a/ and /ɔ/ are acoustically higher than their BP counterparts. We further observe that the lower-mid vowel /ɛ/ is higher in EP than it is in BP, and it is, in fact, closer to EP /e/ than what was observed in BP as previously reported by Escudero et al. (2009a). The authors suggested that this may indicate a future merger of EP /e/ and /ɛ/. The fact that the vowels in the EP /e/–/ɛ/ contrast are acoustically closer than their BP counterpart may make this contrast more difficult to perceive in EP. A visual inspection of the acoustic plotting of the vowel space in BP and EP also shows that /o/ and /u/ in both dialects are produced in close proximity to one another which may make this contrast difficult to discriminate in both dialects. Escudero and colleagues further claim that, on average, BP vowels were produced with longer duration than EP vowels and this may make BP contrasts easier than EP contrasts to discriminate.

Turning to the cross-linguistic comparison of Southern-Californian-English vowels and Mexican-Spanish vowels to BP and EP, we can visually observe that, for the most part, Mexican-Spanish vowels are acoustically closer to both BP and EP than the Southern-Californian-English vowels. Our cross-linguistic comparison of the vowels seems to indicate that /i/ in both BP and EP is acoustically similar to the Southern-Californian-English /i/, whereas BP and EP /e/ seems to be acoustically similar to the Southern-Californian-English /ɪ/ and /e/ vowels. We can also observe that the Mexican-Spanish /i/ seems to fall between the BP and EP /i/ and /e/ vowels, which may make this contrast difficult for bilinguals to perceive if Spanish is activated. The same seems to be true for BP and EP /o/ and /u/ where there is a corresponding vowel that is acoustically similar in English, but the Mexican-Spanish /u/ is acoustically similar to BP and EP /o/ and is also the closest vowel to BP and EP /u/. For the BP and EP /ɛ/, the closest Southern-Californian-English vowel is /ɛ/ and the closest Mexican-Spanish vowel is /e/. However, it is worth noting that the Southern-Californian /ɛ/ and Mexican-Spanish /e/ vowels are acoustically closer to BP /ɛ/ than EP, and that may result in

more consistent categorization of this vowel in BP than in EP. Furthermore, the Mexican-Spanish /a/ is acoustically closer to BP /a/ than it is to EP, and we also observe that the Mexican-Spanish /o/ is acoustically closer to EP /ɔ/ than BP. Interestingly, there appear to be no Southern-Californian-English vowels that are acoustically similar to either BP or EP /ɔ/, and therefore, monolinguals may be more inconsistent with their categorization of this vowel in both dialects.

In sum, our preliminary cross-linguistic acoustic comparison seems to indicate that the Mexican-Spanish vowels as produced by Spanish–English bilinguals are acoustically more similar to the target BP and EP vowels. If the Spanish–English bilinguals in our study activate Spanish to perceive the BP and EP vowels, they may be more consistent in their categorization of vowels in both dialects compared to the English monolinguals. As a result, they may also outperform the monolinguals in their discrimination of the two dialects. However, activation of Spanish may also lead to more difficulty, in particular with the Portuguese /i/–/e/ and /o/–/u/ contrasts, given that both vowels in each contrast appear acoustically similar to a single Mexican-Spanish vowel. Furthermore, some vowel contrasts may be easier to perceive in BP than EP (e.g., /e/–/ɛ/) because BP vowels are longer than EP and some Southern-Californian-English and Mexican-Spanish vowels are acoustically more similar to BP than EP vowels.

Differences in acoustic measurements across dialects of the same language should impact how learners perceive and produce words in those languages. However, another factor which may influence how listeners perceive differences in speech may be linked to their language background. Specifically, bilinguals were shown to display advantages in the linguistic domain, including better learning of novel word forms (Kaushanskaya and Marian 2009; Kaushanskaya 2012; Escudero et al. 2016). This may be the result of enhanced phonological short-term memory, allowing bilinguals greater flexibility in accessing and manipulating sound forms (e.g., Kaushanskaya 2012). However, bilinguals are also required to manipulate two languages, which may lead to competition between them and slower processing as that competition is resolved (e.g., Norris and McQueen 2008).

In phonemic perception, there is conflicting evidence about how bilinguals perform relative to monolinguals. Some studies show that bilinguals perform similarly to monolinguals in non-native vowel perception and production. French–English bilinguals performed similarly to English monolinguals in a vowel perception and production task, where they were asked to discriminate the English vowels /i/–/ɪ/ in an 11-step continuum (Mack 1989). Because isolated vowels require less memory and have no semantic associations, they may behave similarly between monolinguals and bilinguals, even though bilinguals show advantages in linguistic processing tasks. Werker (1986) showed no difference between monolingual, bilingual, and trilingual listeners when they were tasked with discriminating non-native phonetic contrasts, suggesting that broad linguistic experience (e.g., being bilingual/trilingual vs. monolingual) does not confer an advantage when it comes to discriminating novel language contrasts. Instead, specific language experience helps; that is, participants with exposure and experience to the tested language performed better than bilinguals and trilinguals (Werker 1986). Similarly, school-age French–English bilingual and English monolingual children showed no difference in discriminating non-native Russian phonemes (Lambert and MacNamara 1969) even though this same task was used in Rabinovitch and Parver (1966) and their results showed that bilingual university students performed better than monolingual university students. Gallardo del Puerto (2007) showed that the degree of bilingualism (high vs. low use of both languages) does not affect non-native phoneme discrimination, and the author concludes that bilingualism does not affect non-native phonetic discrimination even though there was no comparison to monolingual data.

Other studies also show that broad linguistic experience does seem to play an important role in perceiving differences in non-native contrasts. For example, English monolingual adults compared with multilingual speakers on their discrimination of Japanese geminate consonants showed poorer performance Enomoto (1994). Additionally, there were two groups of multilingual speakers tested—one group had language-specific experience (i.e., spoke a language with gemination, such as Italian) and the other group did not. No differences in performance were found between these groups, suggesting that broad

language experience helped more than specific-language experience. Similarly, French–English bilinguals performed better overall compared to English monolinguals when asked to discriminate contrasts native to French and non-native contrasts (not present in French or English) (Cohen et al. 1967). A similar study examined school-aged children learning French and found that bilingually-instructed children tended to perform better than monolingually-instructed children on discrimination of native and non-native phonetic sequences, but this was not at the level seen in Cohen et al. 1967) and Davine et al. (1971).

A recent study examined the difference in non-native phonetic perception and learning between bilingual and monolingual participants, specifically focusing on the ideas that bilinguals may have an easier time learning a third language due to cognitive advantages (see Cenoz 2003) and that similarity between the already-known languages and the language to be learned may influence the perception and acquisition of non-native contrasts (Antoniou et al. 2015). Mandarin–English bilingual and English monolinguals learned non-native contrasts that were similar to contrasts in either Mandarin or English (though none of them actually existed in either of those languages). The bilingual group was found to be more accurate in their learning compared to the monolinguals, giving support for the idea that bilinguals have a cognitive advantage in processing new language information. Importantly, this held up even when testing another bilingual group (Korean–English), but these results are also qualified by the similarity of the learned contrasts to the experimental groups' native language (e.g., Korean–English bilinguals were better at learning the Korean-like contrasts) and by the overall difficulty of the contrast itself regardless of language background. Therefore, the bilingual advantage interacts with phonetic similarity and universal difficulty (Antoniou et al. 2015).

It is commonly accepted that bilinguals and L2 speakers will process their second language or a new language through the lens of their L1 (e.g., Best and Tyler 2007; Escudero 2005). The mechanisms and factors that underlie and influence this processing are outlined and debated in various models, including the PAM and L2LP. Here, we focus on the L2LP which explicitly outlines how the acoustic and perceptual similarities between the first and second (or foreign) languages shape how listeners will perceive phones in those languages. Additionally, the model claims that listeners' perception of those phones will differ depending on the dialect to which they are exposed. This model is particularly useful for the current study because we expect that these naïve listeners' perceptual assimilation of the non-native phones to their native language(s) will be influenced by the acoustic distance between the sounds in both target dialects and their native inventory. We expect to observe differences in the perceptual assimilation patterns of each dialect and will use these results to predict our listeners' performance in the non-native discrimination task. Given that we are examining SpE bilinguals and CE monolinguals, we expect that the bilinguals may be better able to discriminate Portuguese phonemes because, according to the L2LP model, they have two separate vowel systems that could be activated to aid vowel perception. In particular, Spanish and Portuguese show a smaller acoustic distance between their vowel spaces (see Elvin et al. 2014), and the Spanish–English bilinguals may rely on the Spanish vowel system when discriminating BP and EP vowels. However, we are also specifically interested in a perceptual difference between dialects, which has not yet been tested with bilinguals and monolinguals.

Thus, our study investigates the perception of BP and EP vowels by CE monolinguals and SpE bilinguals. Based on the acoustic differences shown in Escudero et al. (2009a) and, in particular, their finding that, on average, BP vowels show longer duration than EP vowels, we may see differences in categorization and discrimination between the dialects, such that participants are better at BP than EP vowels. Alternatively, if duration differences between the dialects do not affect categorization and discrimination, no dialect difference is expected. However, we do expect an interaction between dialect and vowel contrast because we also see some differences in vowel quality across the two dialects. This was most apparent with certain vowel contrasts (e.g., /e/–/ɛ/), and therefore, certain vowel contrasts may be perceived more accurately in BP compared to EP. Finally, we expect a bilingual advantage for the bilinguals given that Spanish vowels are acoustically closer to Portuguese vowels than English, as observed in Figure 1, which may facilitate their discrimination of Portuguese vowels.

2. Materials and Methods

2.1. Participants

The listeners in this study were 17 SpE bilingual and nine CE monolingual females aged between 20 and 27 years old.[3] All of the CE monolinguals were born and raised in the Central Valley of California. They further reported that their parents were native American English speakers born and raised in California. The SpE bilingual participants reported either Spanish or English as their native language, and were born to parents whose native language was Mexican Spanish. All participants were university students enrolled in an introductory Linguistics course. Participation was voluntary and all participants provided informed consent and were free to withdraw at any time without consequence. A demographical description of our participants' language background is provided in Table 1.

Table 1. A summary of the demographic details of our Californian English monolingual and Spanish–English bilingual population.

	Californian-English Participants	Spanish–English Bilingual Participants
Native language	American English (9) [a]	Spanish (14) Spanish/English (2) American English (1; age of Spanish onset = 13)
Place of birth	Central Valley, California, USA (9)	Central Valley, California, USA (11) Mexico (6; age of arrival to USA = 3–13 years of age)
Mean self-rated proficiency [b]	English: 7 Spanish: 2 (2)	English: 6.625 Spanish: 6.375 [c]
Mother's place of birth	California, USA (9)	Mexico (15) California (2)
Mother's native language	American English (9)	Spanish (14) Spanish/English (2) Mixtec (1)
Father's place of birth	California, USA (9)	Mexico (17)
Father's native language	American English (9)	Spanish (16) Mixtec (1)

[a] The number of participants for each factor is shown in parentheses. [b] Self-rated proficiency (speaking/comprehension) is on a 1-to-7-point Likert-type scale with 10 indicating the highest level of ability and 1 indicating the lowest level of ability. [c] The proficiency rating is missing for one participant who did not fill in that section correctly. USA—United States of America.

2.2. Stimuli

The stimuli used in this study were naturally produced BP[4] and EP isolated vowel tokens (and, therefore, shorter in duration) selected from the Escudero et al. (2009a) corpus. The vowels were produced by 10 male and 10 female speakers from São Paulo, Brazil, and by 10 male and 10 female speakers from Lisbon, Portugal. All vowel tokens were produced in the fVfe context in a carrier sentence, e.g., *Fêfe. Em fêfe e fêfo temos ê* which translates to: '*Fêfe. In fêfe and fêfo we have ê*' (Escudero et al. 2009a). The vowel in the first syllable was always stressed and corresponded to one of the seven Portuguese vowels /i, e, ɛ, a, o, ɔ, u/. Our choice to present vowels in isolation that were extracted from a voiceless fricative context was motivated by previous studies that used similar methods (e.g., Escudero and Chládková 2010; Vasiliev 2013; Elvin et al. 2014). In Vasiliev (2013), the author explains that voiceless stops differ in their voice onset time (VOT) and formant transitions in English, Spanish, and Portuguese, and suggests that it is

[3] Although we would have liked to include Spanish-speaking monolinguals, as well as more American-English-speaking monolinguals, in this study, we were unable to do so due to the fact that Latin-American Spanish–English bilinguals make up the majority of the population in the Central Valley of California. However, we conducted a mixed-effects logistic regression model in our analyses which controlled for differences in effect sizes that may be due to differences in sample sizes, such that a sample size difference, as seen here, does not influence the final statistical results.

[4] The BP vowels were the same vowels previously used in Vasiliev (2013); Elvin et al. (2014); and Elvin (2016).

better to use vowels extracted from a voiceless fricative context to avoid these differences. Vasiliev (2013) further explains that the use of isolated vowels extracted from their context reduces the processing load while still preserving the information contained in the formant transitions. Additionally, Escudero (2005) states that, if vowels are presented in isolation, listeners must then rely on their abstract representations of the target vowels, which helps avoid the contextual effects that may influence vowel categorization and discrimination. There were a total of 140 vowel tokens used in both the non-native categorization and XAB categorical discrimination task (10 speakers × 2 dialects × 7 vowels). The A and B stimuli in the XAB task were synthetic tokens that represented the seven BP and EP vowels, previously used and described in Vasiliev (2013).

2.3. Procedure

Participants were tested in a quiet room at the university. Participants first completed a non-native auditory discrimination task in the XAB format using the same procedure as in previous studies (e.g., Escudero and Williams 2012; Vasiliev 2013; Elvin et al. 2014; Elvin 2016). The XAB task targeted six BP and EP vowel contrasts, namely, /a/–/ɔ/, /a/–/ɛ/, /i/–/e/, /o/–/u/, /e/–/ɛ/, and /o/–/ɔ/. For each trial, participants were instructed to listen to the three vowel tokens, which were played one after the other, and then decided whether the first token they heard (X) sounded more like the second (A) or third (B), and to select their response by clicking the corresponding option (either "2" or "3") on the screen. There were 44 trials for each BP and EP contrast, and each trial was presented in a random order. As in Vasiliev (2013) and Elvin et al. (2014), the X stimuli were always natural tokens and the A and B tokens were always the two synthetic vowel tokens for that particular contrast. Additionally, a synthetic token of each of the vowels in the contrast was used as the X stimulus to ensure that participants understood the task and could match acoustically equal tokens (Elvin et al. 2014). The reasoning for the use of synthetic tokens for the A and B responses is provided in detail in Vasiliev (2013) and Elvin et al. (2014); specifically, by having listeners make a comparison between natural and synthetic tokens, they are completing a categorical discrimination task and their decision is based on phonemic rather than acoustic differences. The inter-stimulus interval of 1.2 seconds was the same as previous studies to ensure phonological processing (Escudero et al. 2009b; Vasiliev 2013; Elvin et al. 2014). A practice session was conducted using a fairly easy contrast, namely /i/–/u/. The experiment took approximately 50 min to complete.

Following the non-native discrimination task, participants then completed a non-native categorization task, where they heard the 70 BP and 70 EP tokens one by one. In the categorization task, participants were instructed to match the vowel sound that they heard with one of the 10 CE vowels (i, ɪ, e, ɛ, æ, ʌ, ɑ, o ʊ, u) on the screen. Given that English is not orthographically transparent, the 10 vowels were presented in one of the 10 key words in the bVt consonantal context, namely, *beat*, *bit*, *bait*, *bet*, *bat*, *but*, *bought*, *boat*, *butcher*, and *boot*, which correspond to the CE phonemes /i, ɪ, e, ɛ, æ, ʌ, ɑ, o ʊ, u/. The non-native categorization task was a multiple-forced choice task as participants were required to choose one of the response options, even when unsure. Participants received a short practice session before beginning the task and the task took about 15 min to complete. Both tasks were completed on a laptop computer using the Praat software (Boersma and Weenink 2018) and the stimuli were presented via headphones.

3. Results

3.1. Non-Native Categorization

In order to investigate whether listeners perceptually assimilate BP and EP vowels to different L1 categories, we conducted a non-native discrimination task. We used the results of the non-native categorization task to predict performance in the non-native discrimination task. The percentage of times each BP and EP vowel was categorized as a CE vowel is presented in Table 2 for the monolinguals and Table 3 for the bilinguals.

Table 2. Californian English monolingual listeners' classification percentages for Brazilian Portuguese (BP, left) and European Portuguese (EP, right). The native vowel category with the highest classification percentage appears in bold and those percentages below chance (i.e., 10%) are not shown.

Monolinguals	Brazilian Portuguese								European Portuguese						
	i	**e**	**ɛ**	**a**	**o**	**ɔ**	**u**		**i**	**e**	**ɛ**	**a**	**o**	**ɔ**	**u**
i	**60**	17						**i**	**66**	22					
ɪ	23	**34**						**ɪ**	20	**33**					
e		16	19					**e**		12	16				
ɛ		31	**67**					**ɛ**		26	**63**				
æ				**54**				**æ**				**34**			
ʌ				13				**ʌ**				27		13	
ɑ				26	30	**72**		**ɑ**				26		**53**	
o					**31**	12		**o**					27	23	
ʊ					11			**ʊ**					**12**		
u					23		**73**	**u**					**39**		**81**

As can be observed in Table 2, the CE monolingual participants categorized the majority of BP and EP vowels as more than one native CE vowel. In particular, monolingual listeners categorized both BP and EP /i/ as CE /i/ 60% and 66% of the times, with CE /ɪ/ as the second most commonly selected vowel (23% and 20%). Their categorization of BP and EP /e/ was spread across four CE vowels namely, /i/, /ɪ/, /e/, and /ɛ/. BP and EP /ɛ/ was predominately categorized as CE /ɛ/ (67% in BP and 63% in EP), with a smaller percentage categorized as CE /e/ (19% in BP and 16% in EP). Categorization of BP and EP /a/ was spread across CE /æ/, /ʌ/, and /ɑ/, with /æ/ being the most popular choice (54% in BP and 34% in EP). EP /o/ was most commonly categorized as CE / u/ (39%), followed by CE /o/ and / ʊ/. For BP /o/, categorization was spread across four categories, namely, /ɑ/, /o/, /ʊ/, and /u/, with CE /o/ having the higher percentage of categorization. BP and EP /ɔ/ were largely categorized as CE /ɑ/ (72% in BP and 53% in EP), followed by CE /o/, as well as CE /ʌ/ for EP. Finally, BP and EP /u/ was consistently categorized as CE /u/ (73% in BP and 81% in EP).

Turning to the bilingual participants, in Table 3, we also observe similar patterns of categorization across the two dialects. Specifically, the categorization of BP /i/ equally spread across CE /i/ and /ɪ/ (each with 41% categorization). EP /i/ was also categorized to the same two vowels; however, CE /ɪ/ was chosen more often than CE /i/ (44% vs. 36%). Similar to the monolinguals, SpE bilinguals categorized BP and EP /e/ across four categories, namely, /i/, /ɪ/, /e/, and /ɛ/, with CE /ɪ/ as the vowel most commonly chosen, followed by CE /ɛ/ (24% categorization in BP and 16% in EP). The SpE bilinguals consistently chose CE /ɛ/ when they heard tokens of BP and EP /ɛ/ (64% in BP and 49% in EP); however, they also chose CE /e/ and, in the case of EP, /ɛ/ was also perceived as CE /ɪ/. The SpE bilinguals did not differ from the monolinguals in that their categorization of BP and EP /a/ was spread across CE /æ/, /ʌ/, and /ɑ/, with /æ/ being the most popular choice (62% for BP and 52% for EP). The majority of the BP /o/ tokens were categorized as CE /ɑ/ (54% of the time), followed by CE /o/ (23% of the time). The categorization of EP /o/ was spread across four vowel categories, specifically, CE /ɑ/ (26%), /u/ (25%), /o/ (23%) and /ʌ/ (14%). Additionally, most of the BP and EP /ɔ/ tokens were categorized as CE /ɑ/ (58% in BP vs. 53% in EP), with a smaller percentage categorized as CE /ʌ/ and /o/ and, in the case of BP, /æ/. Finally, BP and EP /u/ was consistently categorized as CE /u/ (56% in BP and 60% in EP), with a small percentage of tokens categorized as /ʌ/ and /o/.

Table 3. SpE bilingual listeners' classification percentages for BP (left) and EP (right). The native vowel category with the highest classification percentage appears in bold and those percentages below chance (i.e., 10%) are not shown.

Bilinguals	Brazilian Portuguese								European Portuguese						
	i	**e**	**ɛ**	**a**	**o**	**ɔ**	**u**		**i**	**e**	**ɛ**	**a**	**o**	**ɔ**	**u**
i	**41**	11						**i**	36	12					
ɪ	**41**	**45**						**ɪ**	**44**	**62**	19				
e		14	11					**e**		5	15				
ɛ		24	**64**					**ɛ**		16	**49**				
æ				**62**		12		**æ**				**52**			
ʌ				21		14	12	**ʌ**				31	14	12	12
ɑ				10	**54**	**58**		**ɑ**				10	**26**	**53**	
o					23	10	12	**o**					23	21	12
ʊ								**ʊ**							
u							**56**	**u**					**25**		**60**

Based on these categorization percentages, we would not predict any overall differences across the two dialects, which suggests that the overall difference in vowel duration does not play a role in non-native perception. However, there are some instances where the amount of categorization per vowel differs across the two dialects. For example, CE monolinguals categorized BP /a/ as CE /æ/ 54% of the time, yet only 34% of the time in EP. Likewise, they classified BP /ɔ/ as CE /ɑ/ 72% of the time, but only 53% of the time in EP. Additionally, EP /u/ was categorized 81% of the time as CE /u/ and 73% of the time for BP. Similar differences are observed among the SpE bilinguals. Specifically, EP /e/ was categorized as CE /ɪ/ 62% of the time, and BP /e/ was categorized as that same vowel 45% of the time. BP /ɛ/ was categorized as CE /ɛ/ 64% of the time and 49% of the time in EP. There was a 10% difference in categorization of BP and EP /a/ to CE /æ/, with 62% categorization for the former and 52% for the latter. Finally, BP /o/ was categorized as CE /ɑ/ 54% of the time, yet, in EP, it was only categorized as /ɑ/ 26% of the time. Given the qualitative differences in classification percentages across the two dialects, we may find that there will be an interaction between the dialect that participants heard and the specific contrast that they were discriminating.

We conducted a repeated-measures analysis of variance on the categorization scores to examine the possibility of these interactions more closely. The classification percentage served as the dependent measure, and the within-participant factors were dialect (2: European, Brazilian), Portuguese vowel heard (7: i, e, ɛ, a, o, ɔ, u), and English response vowel (10: i, ɪ, e, ɛ. æ, ʌ, ɑ, o ʊ, u). We included language background (bilingual or monolingual) as a between-participants factor. The results showed a significant three-way interaction between dialect, Portuguese word heard, and English response vowel, F (54, 1296) = 4.61, $p < 0.001$. This suggests that, as predicted, participants did show differences in how they classified Portuguese vowels depending on the dialect they heard. A second three-way interaction between Portuguese vowel, English response, and language background, F (54, 1296) = 2.89, $p < 0.001$, suggests that being a bilingual or monolingual also affected how those vowels were classified into English categories. Spanish–English bilinguals may categorize some Portuguese vowels differently because their knowledge of Spanish would influence their perception of those vowels. There was no four-way interaction (F < 1).

Recall that PAM and the L2LP model both claim that discrimination difficulty is caused by perceptual similarity. Therefore, even if one observes instances where a non-native sound is perceived as two or more native categories, difficulty in discriminating vowel contrasts should only occur when each vowel in the contrast is perceived as the same multiple native categories. Following the method described in Levy (2009) and Vasiliev (2013), we calculated the amount of perceptual overlap in a given contrast in order to obtain a quantifiable prediction of the cross-linguistic influences that lead to discrimination difficulty. As mentioned above, a perceptual overlap occurs when two members of a non-native contrast are perceived (or categorized) as the same native vowel(s). The higher the perceptual overlap, the more

likely that this contrast will be poorly discriminated. If no perceptual overlap occurs (i.e., the two vowels in the target contrast are mapped onto different native categories), then there should be no challenge for listeners when distinguishing between the vowels in that contrast. The percentage of perceptual overlap in each contrast across each dialect for monolinguals and bilinguals is shown in Table 4. We obtained a perceptual overlap score by adding together the smaller percentage of the overlapping response categories. For example, for the BP contrast /i/–/e/, the CE monolinguals classified /i/ as CE /i/ 60 % of the time and as CE /ɪ/ 23% of the time. BP /e/ was also classified as CE /i/ 17% of the time and as /ɪ/ 34% of the time. Thus, by adding together the smaller percentage of the overlapping categories, the calculation of the perceptual overlap score would be as follows: 17 + 23 = 40% perceptual overlap. The perceptual overlap calculations are presented in Table 4.

Table 4. Perceptual overlap scores for monolinguals and bilinguals across the six EP and BP contrasts in percentages.

	/a/–/ɔ/	**/a/–/ɛ/**	**/i/–/e/**	**/o/–/u/**	**/e/–/ɛ/**	**/o/–/ɔ/**
Monolinguals						
BP	26	0	40	23	47	42
EP	39	0	42	39	38	23
Bilinguals						
BP	36	0	52	12	34	64
EP	22	0	56	37	39	60

The results in Table 4 clearly show no evidence of perceptual overlapping for the /a/–/ɛ/ contrast for either group in either dialect. Therefore, we predict that this contrast should be the easiest for both groups to perceive. For the bilingual participants, we predict that /o/–/ɔ/ and /i/–/e/ should be the most difficult to perceive in both dialects, whereas for our CE monolingual participants, BP /o/–/ɔ/ and /e/–/ɛ/ should be most difficult, while they should have most difficulty with /o/–/u/ and /i/–/e/ in EP.

Interestingly, from the perceptual overlap table, we can also observe that, for some contrasts, the perceptual overlap score is higher in one dialect than the other. In order to determine whether there was a difference in perceptual overlapping between dialects, we set a 5% difference as a threshold. For example, there is a higher percentage of perceptual overlap for EP /a/–/ɔ/ and /o/–/u/ than in BP for the monolingual participants, whereas the perceptual overlap scores for the /e/–/ɛ/ and /o/–/ɔ/ contrasts are much higher in BP than in EP. A similar observation can be made among the bilinguals, where BP /a/–/ɔ/ has a higher perceptual overlap score than EP and the perceptual overlap is greater in EP /o/–/u/ than BP. Duration differences between BP and EP do not seem to play a role in non-native categorization or perceptual overlap, perhaps because all vowels in American English have intermediate durations between long and short vowels. Interestingly, the perceptual overlap scores for /i/–/e/ are actually higher than /e/–/ɛ/ for the bilinguals across both dialects and for the monolinguals in EP—this does not confirm the prediction from Escudero et al. (2009a), which suggests that EP /e/–/ɛ/ may show a merger in the near future due to their acoustic proximity. Had this prediction been borne out, we would see higher overlap scores for EP /e/–/ɛ/ compared to /i/–/e/ and BP /i/–/e/ and /e/–/ɛ/. This may be explained by the fact that American English does not use duration as a cue (compared to, e.g., Australian English; see Elvin et al. 2016b), and therefore, the vowel duration and quality difference are not as salient for these listeners. Based on the above perceptual overlap scores and following previous studies (Levy 2009; Vasiliev 2013; Elvin 2016), we can qualitatively predict an ordering of discrimination accuracy from the least to most difficult (with ~referring to equal or comparable difficulty), which is shown in Example (1) below.

1. Least difficult → Most difficult
   a. Monolinguals
      BP
      /a/–/ɛ/ > /o/–/u/ ~/a/–/ɔ/ > /i/–/e/ ~/o/–/ɔ/ > /e/– /ɛ/
      EP
      /a/–/ɛ/ > /o/–/ɔ/ > /e/–/ɛ/ ~/a/–/ɔ/ ~/o/–/u/ ~/i/–/e/

   b. Bilinguals
      BP
      /a/–/ɛ/ > /o/–/u/ > /e/–/ɛ/ ~/a/–/ɔ/ > /i/–/e/ > /o/–/ɔ/
      EP
      /a/–/ɛ/ > / a/–/ɔ/ > /o/–/u/ ~/e/–/ɛ/ > /i/–/e/ ~/o/–/ɔ/

In sum, for the discrimination of BP and EP vowels, we predict an effect of contrast that may vary across listener groups, as suggested by the perceptual overlap scores. Alternatively, given that there were some instances where categorization was more consistent in one dialect than in the other, and that some perceptual overlap scores for certain contrasts were much higher in one of the two dialects, we may find an interaction between dialect and vowel contrast that may extend to both language groups. That is, some contrasts may be easier to perceive in one of the two dialects for all listeners.

3.2. Non-Native Discrimination

Participants' accuracy was measured in SPSS using a mixed-effects binary logistic regression model that included the participant, X stimulus, and trial number as random effects, with dialect, vowel contrast, and language background included as fixed effects. In addition to testing for a main effect of our fixed effects, we also tested for an interaction between vowel contrast and language background, vowel contrast and dialect, and dialect and language background, and a three-way interaction between language background, dialect, and contrast. The model revealed a main effect of vowel contrast (χ^2 (5, N = 12,480) = 341.612, $p \leq 0.001$), but no main effect of dialect (χ^2 (1, N = 12,480) = 1.555, p = 0.206), and no overall effect of language background (χ^2 (1, N = 12,480) = 2.586, p = 0.097). Additionally, the model revealed no interaction between dialect and language background (χ^2 (1, N = 12,480) = 0.151, p = 0.743). However, there was an interaction between vowel contrast and dialect (χ^2 (5, N= 12,480) = 23.768, p = 0.003), as well as a trend toward an interaction between vowel contrast and language background (χ^2 (5, N = 12,480) = 10.116, $p \leq 0.065$). We found no evidence of a three-way interaction between language background, dialect, and contrast (χ^2 (5, N = 12,480) = 4.670, p = 0.457).

To explore the main effect of vowel contrast, participants and dialects were pooled together and Fisher's Least Significant Difference (LSD)-corrected post-hoc pairwise comparisons were conducted to investigate the level of difficulty across the six Portuguese contrasts. Figure 2 shows the percentage correct across the BP and EP contrasts for all participants pooled together. The comparisons indicate that participants had higher overall accuracy for /a–ɛ/ than all the remaining contrasts ($p \leq 0.001$), and /o/–/u/ had significantly lower accuracy when compared to the other contrasts ($p \leq 0.001$). No difference was found for /a/–/ɔ/, /e/–/ɛ/, and /o/–/ɔ/ ($p \geq 0.070$), which were all significantly higher than /i/–/e/ ($p \leq 0.001$). This order of difficulty from easiest to most difficult for both groups, across both dialects, is shown in Example (2) (where ~indicates comparable accuracy). However, this main effect was qualified by an interaction with dialect, which was predicted based on previous results showing a difference between EP and BP for certain vowel contrasts.

2. Least difficult → Most difficult
   Monolinguals and Bilinguals (BP and EP)
   /a/–/ɛ/ > /a/–/ɔ/ ~/e/–/ɛ/ ~/o/–/ɔ/ > /i/–/e/ > /o/–/u/

To explore the interaction between vowel contrast and dialect and to test our predictions for level of accuracy within each dialect, participants were pooled together and Fisher's LSD-corrected post-hoc pairwise comparisons were conducted. These comparisons indicated that participants had higher

overall accuracy for BP /a/–/ɛ/ than all the remaining BP contrasts ($p \leq 0.001$) and /o/–/u/ had significantly lower accuracy than the other BP contrasts ($p \leq 0.049$). No difference was found between BP /a/–/ɔ/, /e/–/ɛ/, and /o–ɔ/ ($p \geq 0.232$), while /e/–/ɛ/ and /o/–/ɔ/ had significantly higher accuracy scores than BP /i/–/e/ ($p \leq 0.032$), and /a/–/ɔ/ had marginally higher accuracy than BP /i/–/e/ ($p = 0.075$).

For EP, the participants had higher overall accuracy for /a/–/ɛ/ than all the remaining contrasts ($p \leq 0.001$), and /o/–/u/ had significantly lower accuracy when compared to the other contrasts ($p \leq 0.005$), which mirrors the BP results. In EP, /e/–/ɛ/ and /o/–/ɔ/ did not differ in accuracy ($p = 0.522$) and had significantly higher accuracy than /i/–/e/ ($p \leq 0.001$); however, unlike in BP, EP /a–ɔ/ had significantly higher accuracy scores than all other EP contrasts ($p \leq 0.025$) except /a–ɛ/. Based on these results, our order of difficulty only partially matches the predictions based on acoustic and perceptual overlap from categorization data. The order of difficulty for BP and EP can be seen in Example (3) below.

3. Least difficult → Most difficult
   Monolinguals and Bilinguals
   a. BP
      /a/–/ɛ/ > /a/–/ɔ/ ~/i/–/e/ ~/e/–/ɛ/ ~/o/–/ɔ/ > /i/–/e/ > /o/–/u/

   b. EP
      /a/–/ɛ/ > /a/–/ɔ/ > /e/–/ɛ/ ~/o/–/ɔ/ > /i/–/e/ > /o/–/u/

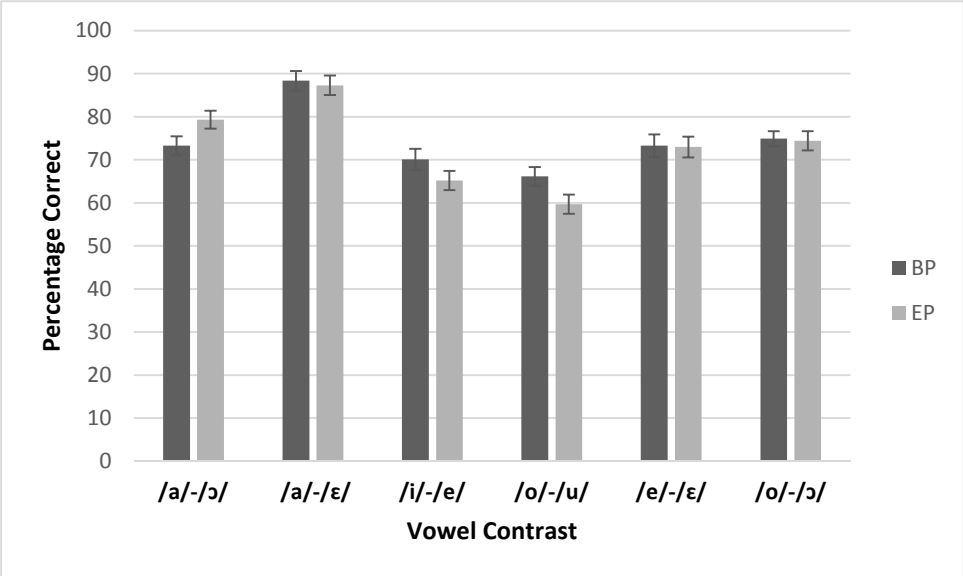


Figure 2. Percentage correct across the BP and EP contrasts for all participants pooled together.

Fisher's LSD-corrected post-hoc pairwise comparisons were also used to examine the difference in accuracy for each Portuguese contrast across dialects. Results indicated that participants had higher accuracy in the /a/–/ɔ/ contrast when it was presented in EP ($p = 0.004$), which was not predicted. Participants found both /i/–/e/ and /o/–/u/ easier to perceive when presented in BP ($p \leq 0.048$), in line with the predictions based on perceived overlap scores.

Turning to the possibility of an interaction between language background and vowel contrast, the results of the pairwise comparisons indicated that there was only one instance where the results differed significantly across language groups. Specifically, bilinguals had significantly more correct responses than monolinguals for the /o/–/u/ contrast ($p = 0.007$), which may have been driven by the fact that monolinguals were predicted to show more difficulty with EP /o/–/u/ compared to the bilinguals. Table 5 shows both groups' accuracy scores for each of the six Portuguese contrasts across the two dialects.

Table 5. Accuracy scores for BP and EP contrasts across language groups.

	Contrast	**/a/–/ɔ/**	**/a/–/ɛ/**	**/i/–/e/**	**/o/–/u/**	**/e/–/ɛ/**	**/o/–/ɔ/**
Bilingual	**BP**	73.53	89.26	71.62	68.38	71.91	73.53
	Standard Error	2.68	2.85	2.66	2.68	2.97	2.25
	EP	80.88	87.50	65.15	61.62	72.79	74.56
	Standard Error	2.28	2.90	2.97	2.92	3.28	2.52
	Contrast	**/a/–/ɔ/**	**/a/–/ɛ/**	**/i/–/e/**	**/o/–/u/**	**/e/–/ɛ/**	**/o/–/ɔ/**
Monolingual	**BP**	72.78	86.67	67.22	61.94	75.83	77.50
	Standard Error	3.90	4.04	5.08	3.40	5.24	2.80
	EP	76.39	86.94	65.28	56.11	73.33	74.17
	Standard Error	4.33	3.84	3.42	3.15	3.44	4.55

Recall that we predicted the degree of perceptual overlap in a given contrast would influence the discrimination difficulties listeners may have with certain contrasts. In other words, when there is high perceptual overlap, listeners may perceive both sounds in the contrast as the same single or multiple categories resulting in discrimination difficulties. When there is little to no overlap it should be easy. Therefore, we ran Spearman's rank correlations to confirm that discrimination difficulty increases when a greater amount of perceptual overlap is present in the contrast. For BP, we found a significant negative correlation between the amount of perceptual overlapping and the number of correct responses for the bilingual listeners ($r = -0.067$, p (two-tailed) ≤ 0.001), but not for monolinguals ($r = -0.013$, p (two-tailed) = 0.512). In the case of EP, we found significant negative correlations for both bilinguals ($r = -0.099$, p (two-tailed) ≤ 0.001) and monolinguals ($r = -0.054$, p (two-tailed) ≤ 0.001). We interpret a negative correlation as an indication that a greater amount of perceptual overlap results in lower overall accuracy. However, although significant, these correlations are fairly low (i.e., less than one in all cases), suggesting that the relationship between perceptual overlap and discrimination difficulty may also be mediated by other factors.

4. Discussion and Conclusions

We investigated CE monolingual and SpE bilingual listeners' perception of Brazilian and European Portuguese. According to the L2LP model and previous research (e.g., Escudero and Chládková 2010) and given the duration and acoustic differences in the realizations of the BP and EP vowels, there should be differences in how these vowels are filtered and categorized in the native language. We also investigated whether or not bilinguals have overall higher accuracy when perceiving the Portuguese contrasts in the two dialects. Overall, there seemed to be very little difference between the CE monolingual and bilingual participants in their categorization of the BP and EP vowels, and, for the most part, the participants systematically chose the same native CE categories for vowels in both dialects, with the exception of a few. Perhaps the most striking differences found between the categorization of the two dialects were the cases where categorization to one particular vowel was more consistent in one dialect than the other. Furthermore, the results from the perceptual overlap calculations indicated that, for both groups, /a/–/ɛ/ should be the easiest to discriminate in both dialects. The fact that the categorization was more consistent in one dialect led to the prediction that there may be an interaction for accuracy scores between a particular contrast and the dialect in which it was heard.

We found a main effect of contrast as predicted; however, we did not find a main effect of dialect or language background, indicating that, overall, there were comparable accuracy scores across the two groups and two dialects. The main effect of contrast indicated that some contrasts were easier to discriminate than others. Specifically, both groups found /a/–/ɛ/ to be the easiest to perceive and /o/–/u/ the most difficult. The fact that both groups found /o/–/u/ difficult to discriminate in both dialects is not surprising, as these two vowels are acoustically similar and were difficult to perceive in other studies of Portuguese vowel perception (Vasiliev 2013; Elvin et al. 2014; Elvin 2016). In particular, the fact that native Mexican-Spanish bilingual speakers also had trouble with the /o/–/u/

contrast here replicated Elvin et al. (2014) showing that Iberian-Spanish listeners also had trouble discriminating that contrast. As predicted, the main effect was qualified by an interaction between the dialect and vowel contrast. That is, we found that there was overall higher accuracy for specific vowel contrasts differentially in the two dialects, which was expected because the perceptual overlap scores differed across both dialects. Although BP had longer vowel duration overall (Escudero et al. 2009a), this difference did not play a significant role in categorization or discrimination. The results also indicate that /o/–/u/ was easier to discriminate when produced in BP than in EP. Although there was little difference in perceptual overlap scores of BP and EP /i/–/e/ for both groups, the results suggest that both groups did in fact find this contrast easier to perceive when produced in the BP dialect. Interestingly, our perceptual ranking from the categorization task predicted lower accuracy for /o/–/ɔ/ than /o–u/ for bilinguals in both dialects and for monolinguals in BP—Elvin et al. (2014) showed that both Iberian-Spanish and Australian-English listeners ranked /o/–/u/ as the most difficult and /o/–/ɔ/ was ranked in the middle (with /a/–/ɛ/ as the easiest). However, our participants performed worse with /o/–/u/ compared to /o/–/ɔ/ during the discrimination task. This is most likely due to the addition of EP as a dialect of interest, but may also be due to acoustic differences between the target and native dialects. In particular, our cross-linguistic acoustic comparison showed that /o/ and /ɔ/ in BP and EP were acoustically similar to two separate Mexican-Spanish vowels produced by bilingual speakers. It may be the case that bilingual speakers activated Spanish to help discriminate this contrast. This contrast may also have been easier for monolingual speakers to perceive given that the English /o/ was acoustically similar to BP and EP /o/, with no English vowels being acoustically similar to BP or EP /ɔ/. It is possible that the lack of an acoustically similar vowel to /ɔ/ made it easier for monolinguals to perceive a difference between the two vowels in this contrast. Further analyses, with more detailed acoustic analyses than those presented here should, therefore, be considered in future work examining non-native vowel perception.

Furthermore, we predicted that bilinguals would find EP /a/–/ɔ/ easier to perceive than BP, where the monolinguals would find it more difficult. Our results indicate that, in fact, both groups found this contrast easier to perceive when produced in EP. Finally, although it was a trend, we found that, overall, the bilinguals had higher accuracy than monolinguals for /o/–/u/, which is, in fact, supported by the amount of perceptual overlap in that contrast. Indeed, we found that, with the exception of BP for the CE monolinguals, there was a significant correlation between the amount of perceptual overlap and discrimination accuracy for all other participants. Thus, it seems that, despite some discrepancies, our predictions based on the amount of perceptual overlap are largely supported.

Our prediction that bilinguals would perform better than monolinguals was not supported in this study. Indeed, both groups showed similar discrimination scores and categorization patterns. While there was a trend toward an interaction between vowel contrast and language background, further investigation showed that this was due to a single contrast, namely /o/–/u/, which had higher accuracy for bilinguals. The finding of little or no advantage for bilinguals in this study patterns with previous work showing that language background does not affect how listeners perceive non-native contrasts (Mack 1989; Werker 1986). Importantly, this effect exists regardless of the languages spoken by the bilinguals: in our study, Spanish–English bilinguals did not perform better than English monolinguals even though Spanish and Portuguese share more similar vowels. We would expect that speakers of Spanish would have an advantage as they would presumably map the incoming speech signal to Spanish and show better discrimination. Importantly, this was the case in Antoniou et al. (2015), wherein the Mandarin–English bilinguals performed better than a monolingual group when learning Mandarin-like contrasts, and Korean–English bilinguals learned Korean-like contrasts better. Here, even though Spanish and Portuguese are similar languages, the addition of Spanish as a native language did not help in discrimination or categorization. The only advantage Spanish–English bilinguals had was when discrimination of the /o–u/ contrast occurred across both dialects. This may be because the production of that contrast across both dialects was the most similar to Spanish. Future acoustic analyses should investigate the distance between the Spanish and Portuguese

instances of these vowels for a more comprehensive analysis. Importantly, the categorization task was completed with English words—the native Spanish speakers may have performed differently had they been presented with Spanish words to select instead.

However, our overall results suggest that, at the level of perceiving non-native vowel contrasts, the discrimination between vowel contrasts may be led exclusively by the acoustic distance between the heard vowels. Because our stimuli are very short in duration, these tasks may not tap into the heightened linguistic and non-linguistic abilities that bilinguals seem to have in other tasks (e.g., Kaushanskaya 2012). That is, it may be the case that the task demands here were such that they did not engender differences in how monolinguals and bilinguals performed because they were very simple and did not require phonological memory or higher-order cognitive functioning. Indeed, other tasks that show a bilingual advantage tend to happen at the lexical level (e.g., word learning; Escudero et al. 2016). Therefore, while bilingualism may result in advantages in certain tasks, the present study did not show any significant differences between groups. Furthermore, bilingualism cannot be solely identified as the factor mediating the relationship between perceptual overlap and discrimination difficulty because no significant effect of language background or dialect was found in the correlations between the tasks. In general, our results do not support the idea that bilingualism heightens non-native phonetic perception across dialects, an important contribution to the literature on bilingual advantage in non-native phonetic perception.

In conclusion, our study finds that listeners do indeed filter, categorize, and discriminate the vowels of two dialectal varieties differently. Importantly, this does not vary based on language background. Furthermore, our non-native discrimination task indicates discrimination accuracy differs depending on the specific vowel contrast that they heard and that dialectal differences were only evident in a select number of vowel contrasts. Therefore, although there will be differences in the developmental trajectories of specific vowel contrasts, these differences will not carry over at a dialectal level. Language teachers may, however, choose to spend more time training vowel perception of a given vowel contrast in the specific dialect that was more difficult to perceive (e.g., /a/–/ɔ/ in BP will be more difficult than in EP and /i/–/e/ and /o/–/u/ will be more difficult in EP). We do acknowledge that our pool of data for the monolingual group was limited, and therefore, future studies should further investigate whether there are differences between bilinguals and monolinguals in their initial perception of Portuguese vowels. Furthermore, we acknowledge that our study was limited in the fact that we were unable to include Mexican-Spanish-speaking monolinguals. It would be beneficial in future studies to compare the performance of Mexican-Spanish monolinguals to the present study, as well as the Iberian-Spanish participants in Elvin et al. (2014). Finally, given that we found a difference in the listeners' difficulty levels for the perception of EP and BP, we would expect these differences to carry over into the domains of spoken word recognition and non-native productions. Future studies should be carried out to test whether or not that is indeed the case.

Author Contributions: J.E., A.T., and P.E. conceived and designed the experiments; J.E. ran the experiments; J.E. and A.T. analyzed the data; J.E., A.T., and P.E. wrote the paper.

Funding: The second (A.T) and third (P.E.) authors were supported by the ARC Centre of Excellence for the Dynamics of Language [CE140100041].

Acknowledgments: We would like to thank all of the students who voluntarily participated in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

Alispahic, Samra, Karen E. Mulak, and Paola Escudero. 2017. Acoustic properties predict perception of unfamiliar Dutch vowels by adult Australian English and Peruvian Spanish listeners. *Frontiers in Psychology* 8: 52. [CrossRef] [PubMed]

Antoniou, Mark, Eric Liang, Marc Ettlinger, and Patrick C. M. Wong. 2015. The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition* 18: 683–95. [CrossRef]

Baker, Wendy, and Laura Catharine Smith. 2010. The impact of L2 dialect on learning French vowels: Native English speakers learning Quebecois and European French. *Canadian Modern Language Review* 66: 711–38. [CrossRef]

Best, Catherine T. 1995. A direct realist perspective on cross-language speech perception. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Timonium: York Press, pp. 171–204.

Best, Catherine T., and Michael Douglas Tyler. 2007. Non-native and second-language speech perception: Commonalities and complementarities. In *Language Experience in Second-Language Speech Learning: In Honor of James Emil Flege*. Edited by Ocke-Schwen Bohn and Murry J. Munro. Amsterdam: John Benjamins, pp. 13–34.

Boersma, Paul, and David Weenink. 2018. Praat: Doing Phonetics by Computer. Computer Program, Version 6.0.37. Available online: http://www.praat.org/ (accessed on 14 March 2018).

Cenoz, Jasone. 2003. The additive effect of bilingualism on third language acquisition. *The International Journal of Bilingualism* 7: 71–88. [CrossRef]

Chládková, Kateřina, and Václav Jonáš Podlipský. 2011. Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners. *The Journal of the Acoustical Society of America* 130: EL186–EL192. [CrossRef] [PubMed]

Cohen, Stephen P., G. Richard Tucker, and Wallace E. Lambert. 1967. The comparative skills of monolinguals and bilinguals in perceiving phoneme sequences. *Language and Speech* 10: 159–68. [CrossRef] [PubMed]

Davine, M., G. Richard Tucker, and Wallace E. Lambert. 1971. The perception of phoneme sequences by monolingual and bilingual elementary school children. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement* 3: 72–76. [CrossRef]

Díaz Granado, Miriam. 2011. L2 and L3 acquisition of the Portuguese stressed vowel inventory by native speakers of English. PhD. dissertation, The University of Arizona, AZ, USA.

Elvin, Jaydene. 2016. The Role of the Native Language in Non-Native Perception and Spoken Word Recognition: English vs. Spanish Learners of Portuguese. Ph.D. dissertation, Western Sydney University, Sydney, NSW, Australia.

Elvin, Jaydene, Paola Escudero, and Polina Vasiliev. 2014. Spanish is better than english for discriminating portuguese vowels: Acoustic similarity versus vowel inventory. *Frontiers in Psychology* 5: 1–8. [CrossRef] [PubMed]

Elvin, Jaydene, Daniel Williams, and Paola Escudero. 2016a. The relationship between perception and production of Brazilian Portuguese vowels in European Spanish monolinguals. *Loquens* 3: 31. [CrossRef]

Elvin, Jaydene, Daniel Williams, and Paola Escudero. 2016b. Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English. *The Journal of the Acoustical Society of America* 140: 576–81. [CrossRef] [PubMed]

Enomoto, Kayoko. 1994. L2 perceptual acquisition: The effect of multilingual linguistic experience on the perception of a "less novel" contrast. *Edinburgh Working Papers in Applied Linguistics* 5: 15–29.

Escudero, Paola. 2005. Linguistic Perception and Second Language Acquisition. Ph.D. dissertation, Utrecht University, Utrecht, The Netherlands.

Escudero, Paola, and Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551–85. [CrossRef]

Escudero, Paola, and Kateřina Chládková. 2010. Spanish listeners' perception of American and Southern British English vowels. *The Journal of the Acoustical Society of America* 128: EL254–59. [CrossRef] [PubMed]

Escudero, Paola, and Daniel Williams. 2012. Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. *The Journal of the Acoustical Society of America* 131: EL406–12. [CrossRef] [PubMed]

Escudero, Paola, Paul Boersma, Andréia Schurt Rauber, and Ricardo A. H. Bion. 2009a. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America* 126: 1379–93. [CrossRef] [PubMed]

Escudero, Paola, Titia Benders, and Silvia C. Lipski. 2009b. Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics* 37: 452–65. [CrossRef]

Escudero, Paola, Karen E. Mulak, Charlene S. L. Fu, and Leher Singh. 2016. More limitations to monolingualism: Bilinguals outperform monolinguals in implicit word learning. *Frontiers in Psychology* 15: 1218. [CrossRef] [PubMed]

Gallardo del Puerto, Francisco. 2007. Is L3 phonological competence affected by the learner's level of bilingualism? *International Journal of Multilingualism* 4: 1–16. [CrossRef]

Grijalva, Criccely, Page E. Piccinini, and Amalia Arvaniti. 2013. The vowel spaces of Southern Californian English and Mexican Spanish as produced by monolinguals and bilinguals. *Proceedings of Meetings on Acoustics* 19: 060088. [CrossRef]

Hagiwara, Robert Elliot. 1997. Dialect variation and formant frequency: The American English vowels revisited. *The Journal of the Acoustical Society of America* 102: 655–58. [CrossRef]

Iverson, Paul, and Bronwen Evans. 2007. Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *The Journal of the Acoustical Society of America* 122: 2842–54. [CrossRef] [PubMed]

Iverson, Paul, and Bronwen Evans. 2009. Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America* 126: 866–77. [CrossRef] [PubMed]

Kaushanskaya, Margarita. 2012. Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition* 15: 470–89. [CrossRef]

Kaushanskaya, Margarita, and Viorica Marian. 2009. The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review* 16: 705–10. [CrossRef]

Kroll, Judith, and Ellen Bialystok. 2013. Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology* 25: 497–514. [CrossRef] [PubMed]

Ladefoged, Peter. 1999. American English. In *Handbook of the International Phonetics Association*. Cambridge: Cambridge University Press.

Lambert, Wallace. E., and John MacNamara. 1969. Some cognitive consequences of following a first-grade curriculum in a second language. *Journal of Educational Psychology* 60: 86–96. [CrossRef]

Levy, Erika S. 2009. On the assimilation-discrimination relationship in American English adults' French vowel learning. *The Journal of the Acoustical Society of America* 126: 2670–82. [CrossRef] [PubMed]

Mack, Molly. 1989. Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception & Psychophysics* 46: 187–200.

Norris, Dennis, and James M. McQueen. 2008. Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review* 115: 357–95. [CrossRef] [PubMed]

Rabinovitch, M. Sam, and L. M. Parver. 1966. Auditory discrimination in monolinguals and poliglots. Paper presented at the Meeting of the Canadian Psychological Association, Montreal, QC, Canada, June 1966.

van Leussen, Jan -Willem, and Paola Escudero. 2015. Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology* 6: 1–12. [CrossRef] [PubMed]

Vasiliev, Polina. 2013. The Initial State for Californian English Learners of Spanish and Portuguese Vowels. Ph.D. dissertation, University of California, Los Angeles, CA, USA.

Werker, Janet F. 1986. The effect of multilingualism on phonetic perceptual flexibility. *Applied Psycholinguistics* 7: 141–55. [CrossRef]

languages

MDPI

Article

The Lexical Development of Canadian-Born Romanian L1 Bilingual Kindergarteners

Maria Claudia Petrescu [1,*] and Rena Helms-Park [2]

1 School of Early Childhood Studies, Ryerson University, 350 Victoria Street Toronto, ON M5B 2K3, Canada
2 Centre for French and Linguistics, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON M1C 1A4, Canada; rhelms@utsc.utoronto.ca
* Correspondence: mpetrescu@ryerson.ca; Tel.: +1-416-979-5306

Received: 1 February 2018; Accepted: 7 August 2018; Published: 16 August 2018

Abstract: This study charts the lexical development of three sequential bilingual kindergarteners whose first language, Romanian, was acquired naturalistically at home, and whose second language, English, was acquired in kindergarten. The children's lexical development in English and Romanian was assessed at five different points over a two-year period via the PPVT-4 (peabody picture vocabulary test 4) and a specially adapted PPVT-4 for Romanian. The children's lexical repertoires were further analyzed to uncover home versus school and cognate versus non-cognate acquisitional differences. In addition, because there is no database of lexical items acquired by monolingual Romanian children, the PPVT-4 adapted for Romanian was administered to 22 monolingual six-year-old Romanian children in Sibiu, Romania. The findings indicate the following: (i) the three bilinguals' receptive vocabulary in English was below average when they joined kindergarten, and at or above average two years later; (ii) their lexical growth in Romanian was steady; (iii) the bilinguals' scores for words belonging to a home register reflected ceiling effects in English and Romanian (i.e., were very well known); (iv) academic words were known to an equal extent in English and Romanian, but scores were lower than for the home register; and (v) there was no definitive evidence of cognate facilitation. A comparison of the monolingual and bilingual Romanian repertoires reflects the following: (i) equally high scores for home items; (ii) differences in scores in the academic register in favour of the Romanian monolinguals; and (iii) important lifestyle and cultural differences between the groups. The Romanian children, for example, were more familiar than their Canadian counterparts with items related to home maintenance, such as *șmirghăluiește* ('sanding') and *mistrie* ('trowel'), or items probably learned in school, such as *foca* ('walrus') and *broască țestoasă* ('tortoise').

Keywords: Romanian; English; lexical development; child language acquisition; sequential bilingualism; bilingual kindergarteners

1. Introduction

The primary objective of this longitudinal study is to document the bilingual lexical development of three Romanian-speaking Canadian children between the ages of four and six. To chart their lexical development in Romanian, their first language (L1), and in English, their second language (L2), two separate measures of lexical ability were administered, namely, PPVT-4 (Peabody Picture Vocabulary Test 4) and a specially adapted PPVT-4 for Romanian. The children's language-specific development included an analysis of home versus school vocabulary differences in each language. In addition, as Romanian and English share a sizable number of Greco–Latinate cognates (Petrescu et al. 2017), possible cognate facilitation resulting from the transfer from L1 to L2 or L2 to L1 was also investigated.

For comparative purposes, especially as there are not yet any known accounts of lexical development in monolingual Romanian contexts, we also collected Romanian data via the adapted Romanian PPVT-4 from 22 monolingual kindergarteners in an urban centre in Romania. While this was not a longitudinal study, these data were collected when the children were six years old, paralleling the stage at which we collected our final data from the three Canadian bilingual children.

1.1. The Study's Rationale

Our study aims to investigate the receptive lexical knowledge of bilingual children whose L1 and L2 acquisition is successive but overlapping. In the case of simultaneous or overlapping bilingual acquisition, some argue that vocabulary growth in each of the two languages is generally delayed because of the limited input and interaction in each of the two languages; furthermore, a bilingual's lexicon in each of the two languages is smaller than that of a monolingual's for the same reasons (Bialystok et al. 2010; Werker 2012). Another interesting question is whether a bilingual's L2 lexicon is qualitatively different because of environmental and cultural factors such as the manner and purpose of L2 acquisition (Sabourin et al. 2014). Sabourin et al. (2014), for example, suggest that late L2 learners, irrespective of age, could have an integrated L1–L2 lexicon if L2 acquisition included input and interaction in a naturalistic context. As a corollary to this, Jiang (2000) hypothesizes that in a foreign-language context, where the instructor is the only source of input and interaction, learners are frequently encouraged to translate from their L2 to their L1; this gives rise to a reliance upon the L1 and to a superficial and occasionally inaccurate semantic representation of L2 words. In addition to exploring these issues, our study also aims to consider how potential disadvantages in terms of reduced lexical size or slow lexical acquisition could be mitigated by language relatedness. From a lexical point of view, for instance, it could be advantageous to have a Romance language, like Romanian, as a definitive L1 and a language like English as an L2, because numerous "everyday" Latinate words in Romanian pair up with low-frequency "academic" words (cognates) in English (Petrescu et al. 2017). Furthermore, typological similarities between an L1 and L2 could lead to more than cognate facilitation because these similarities allow for extra "processing space" to be devoted to lexical acquisition (Ard and Homburg 1983). In short, limited input and interaction could be compensated for by typological similarities.

1.1.1. Lexical Knowledge of Monolingual and Bilingual Children

Extracting words from the speech stream during the first year of L1 acquisition is a complex process ("prosodic bootstrapping") that requires segmentation at different, interacting levels: phoneme perception in syllabic units in keeping with phonemic repertoire of the ambient language; phrasal boundaries using acoustic cues such as pauses and drops to F0 (fundamental frequency); rhythm (including stresses) to acquire smaller phrases and word onsets; and so on (Cutler 1994; Guasti 2000; Jusczyk et al. 1993; Werker and Tees 1984). By nine months or so, a baby has acquired the basic phonotactics of their L1 and can begin to hold a rudimentary version of the phonological form of an L1 word (its "lexeme") in memory (Jusczyk et al. 1993), a milestone allowing them to start mapping meaning onto these forms. During their pre-school years, children begin to acquire words rapidly, mapping meaning onto phonological forms via numerous mechanisms, such as child-directed speech, semantic and syntactic bootstrapping, cognitive development, and a rapid increase in world knowledge (Clark 1995; Gleitman 1990; Pinker 1984; Snow 1977; Tomasello 2003).

In garden-variety types of successive child L2 acquisition, L2 words are acquired once the child has a well-established L1. In cases of overlapping bilingual acquisition in early childhood, the L1 lexicon might be small (and restricted to items learned in a domestic context) when L2 acquisition commences. In such a case, and most noticeably when the L1 is a minority language, while the L2 is the language of the community and school, an important number of concepts are learned through the L2 rather than the L1 (Sherkina-Lieber and Helms-Park 2015). The issue of overlapping L1 (minority) acquisition and L2 (majority) acquisition pertains to the current study, in which the participants began their L2

acquisition of English in a meaningful way at only the age of four, when they joined kindergarten. Thus, one question that the study addresses is whether the children's two lexicons differ in terms of the types of words known in each language (for example, whether words for food and personal grooming are better known in Romanian than are terms for simple arithmetic or animal houses).

Irrespective of whether L1 and L2 lexical acquisition overlap or are clearly sequential, there is inevitable interaction between the L2 and L1 lexicons (Jarvis 2009; Jiang 2000; Kroll and Stewart 1994; Kroll et al. 2010; Pavlenko 2009; Wolter 2006). In general ways, the very existence of an L1 can facilitate L2 lexical acquisition as the use of the L1 provides evidence to the user that language has structures that can convey meanings and that at least some words have referents. Furthermore, many of the concepts that accompany lexical items do not have to be re-learned. However, the L1 also complicates L2 lexical acquisition. If we consider the process of extracting words from the speech stream alone, an L1 mechanism for segmenting speech phonologically would likely be a default at the outset of L2 acquisition, and an L2 mechanism would need to be acquired alongside to segment the L2 speech stream; the degree of ease or difficulty in this regard would depend on how similar or different, respectively, the two mechanisms are (Cutler 1994; Werker 2012). Beyond phonological segmentation, there are complex ways in which related lexical entries in the two lexicons are linked. Kroll and Stewart's (1994) revised hierarchical model, for instance, posits the view that in early L2 lexical acquisition, the learner relies on the L1 translation of the L2 item to arrive at the concept ("word association"), but with growing L2 proficiency, the learner can possibly arrive at the concept directly without reliance on the L1 ("concept mediation"). However, there are other possibilities. The concept and associated lexical entry could first be acquired in an L2 context, for instance, when a Romanian-speaking child learns 'rectangle' in kindergarten without knowing the Romanian word *dreptunghi* and the shape it represents.

While L2 lexical knowledge can be both qualitatively and quantitatively different from L1 knowledge, perhaps much also depends on the mode and purpose of L2 acquisition. Based on a preliminary analysis of lexical acquisition of L1 late French bilinguals (immersion into a French environment took place between the ages of 9–19 years), Sabourin et al. (2014) suggest that naturalistic acquisition could be conducive to (i) an integrated L1/L2 lexicon (i.e., "concept mediation") versus an L2 lexicon linked to its L1 counterpart through word associations, and (ii) richer semanticity of L2 lexical representations (cf. Finkbeiner et al. 2004). Another factor that comes into play in the interaction of the two lexicons is the nature of the lexical items and their conceptual representations themselves. For example, when concepts attached to L1 and L2 words are the same, as is more often the case with items with concrete referents than abstracts ones (De Groot and Keijzer 2000), bidirectional transfer of conceptual knowledge can be positive.

Previous research suggests that bilingual children develop smaller vocabularies in each of the two languages, and more slowly in each language than monolingual children (Mahon and Crutchley 2006; Nicoladis and Genesee 1996; Oller and Eilers 2002; Oller et al. 2007; Umbel et al. 1992). As a large body of research suggests that vocabulary size is an important factor in academic success (August et al. 2005; Ouellette 2006; Rohde and Thompson 2007; Swanson et al. 2008; Verhallen and Schoonen 1993; Vermeer 2001), such findings are especially consequential for a bilingual child's school language. However, Bialystok and Feng (2009) and Bialystok et al. (2010), on the basis of tests of receptive lexical knowledge, caution that bilinguals' overall lexical scores in each of their two languages need to be broken down further to provide a clearer picture of their lexical abilities, for example, by distinguishing between every day home and community vocabulary and academic school-based vocabulary. Via a meta-analysis of existing studies on English receptive vocabulary knowledge obtained from 772 English monolingual children and 966 bilingual children between the ages of 3 and 10, Bialystok et al. (2010) uncovered significantly lower scores for bilingual children when compared with monolingual ones. Nonetheless, there was no disadvantage for academic vocabulary among the bilingual group but, instead, lower scores only in vocabulary related to home life, most likely learned by children in their L1. These findings are in line with research that found that bilingual

children are not disadvantaged in academic and literacy achievement (Bialystok et al. 2005), nor in academic uses of spoken language (Peets and Bialystok 2015). Furthermore, when taken together, the bilingual children's vocabularies in their two languages could be larger than the vocabulary of monolingual children (Bialystok et al. 2010). The potential strengths and gaps in the lexical knowledge of a child in each language is an issue that is addressed in this study.

With regard to the assertion that bilingual children develop vocabulary more slowly in each language compared with monolingual children, at least when receptive vocabulary knowledge is assessed (Mahon and Crutchley 2006; Oller and Eilers 2002), Bialystok et al. (2010) caution that this claim is too general to apply to all languages and that the specific language pairs that the children are learning will influence the rate of acquisition. For example, the existence of Greco–Latin cognates in Romanian and English could give a Romanian L1 speaker an advantage over a Vietnamese one if both are acquiring academic English vocabulary (Petrescu et al. 2017). We should note that even if cognates are excluded from consideration, the numerous typological similarities between related languages can make L2 acquisition easier for such learners, thus freeing up more resources for lexical acquisition (see, for example, Ard and Homburg (1983) study involving L1 Spanish versus L1 Arabic when it comes to acquiring L2 English vocabulary).

1.1.2. Romanian–English Cognates and Cognate Facilitation

Romanian belongs to the Romance subgroup of the Indo-European languages, and has strong lexical, grammatical, and phonological links to Italian, French, Spanish, Portuguese, and other Romance languages, being the "easternmost representative of the family of Romance languages" (Cojocaru 2003). At the morphosyntactic level, Romanian is one of the most conservative Romance languages (e.g., it is the only Romance language that has preserved all six cases and three genders inherited from Latin).

At the lexical level, Romanian is a language that is deeply rooted in Latin, developing in relative isolation as a result of geographical and historical circumstances. According to Maneca (1996), the language has a core vocabulary of Latin words (approximately 35%), which constitute around 74% of the most frequently used words in the language. Given the Latinate influence on the English lexicon through Latin as well as Old and Middle French, Romanian shares a sizeable number of cognates with English (Petrescu et al. 2017). For the purposes of this study, we define cognates psycholinguistically, that is, as words that overlap phonologically and semantically irrespective of diachronic (etymological) factors (Helms-Park and Dronjic 2016; Costa et al. 2005; Dijkstra et al. 2010; Hall 2002; Midgley et al. 2011). In this study, the degree of phonological overlap is calculated via Kohnert et al.'s COSP (crosslinguistic overlap scale for phonology) (Kohnert et al. 2004) in order to assess the probability of interlingual transfer. The theoretical basis for focusing on phonological overlap is the view that it is through automatic phonological activation of neighbours in the bilingual lexicon that cognate recognition begins; however, the learner must also ascertain immediately or subsequently that the words overlap semantically (Carroll 1992; Costa et al. 2005; Dijkstra et al. 2010; Midgley et al. 2011). For example, the Romanian word *vehicol* overlaps both phonologically and semantically with the English 'vehicle', fulfilling the condition of transfer, and thus potentially facilitating the acquisition of the English word by Romanian speaking bilingual children.

While there is a robust body of literature that demonstrates that adults are more successful in recognizing, acquiring, and retaining cognates compared with non-cognates (De Groot and Keijzer 2000; Sánchez-Casas and García-Albea 2005), research provides conflicting evidence regarding cognate facilitation when it comes to children. Umbel et al. (1992) found that children from Spanish monolingual and Spanish bilingual homes achieved similar overall scores on both the peabody picture vocabulary test (Dunn and Dunn 1981) and the Test de Vocabulario en Imágenes Peabody-Adaptación Hispanoamericana (Dunn et al. 1986), responding correctly on cognates and noncognates at about the same rate (68% vs. 67%). In a follow-up study using the same receptive tests, Umbel and Oller (1994) tested first, third, and sixth graders and obtained similar results. In contrast, Malabonga et al. (2008) report that recognition of cognates increases as the children progress academically. On the basis

of their cognate awareness test (CAT), a multiple-choice receptive test, they found no evidence of cognate facilitation effects among Spanish–English fourth graders but, a year later, the same students exhibited a cognate advantage on a multiple-choice test of low-frequency English words. The results of a variety of studies indicate that children's sensitivity to cognate recognition is associated with a range of factors; namely, amount of language exposure (Pérez et al. 2010), previous knowledge of the word concept in the L1 (Nagy et al. 1993), levels of L1 ability (Malabonga et al. 2008), and age (Kelley and Kohnert 2012).

In this study, we will examine whether cognate facilitation plays a role in lexical development in a context that has received little attention, that is, child Romanian and English bilingualism. Unlike Romance languages like French, Spanish, and Italian, Romanian has remained surprisingly under-represented in both L1 and L2 acquisition research (see Avram 2001; Buja 2008, for some notable exceptions). More generally, the current study highlights the growing importance of Romanian as a heritage (minority) language in a Canadian setting and provides an opportunity to examine the lexical knowledge of young bilinguals in this new context. Contrary to a typical French immersion scenario, which has served as the setting of most of the research on childhood bilingualism in Canada, in our study, the children's L1 and sole target language in their pre-school years is the minority language (Romanian). It is only after the age of four, when they join an English-medium kindergarten, that they acquire English in a systematic and meaningful way.

2. Method

The design and method of this two-part study aims to answer the following research questions:

- What are the patterns of lexical development in Romanian and English depicted by the three Canadian bilinguals between the ages of four and six?
- Is there any evidence of cognate facilitation among Romanian–English bilinguals during their kindergarten years?
- How do the Romanian lexicons of bilingual children compare with those of monolinguals of the same age? Apart from potential differences in vocabulary size, do the types of lexical items comprehended by these monolinguals and bilinguals differ?

In order to address the above questions, Romanian and English lexical data were collected at five different points over a period of two years from three Romanian-speaking, Canadian-born children between the ages of 4 and 6 and 22 monolingual Romanian children aged 6.

2.1. Participants

2.1.1. Bilingual Participants

The participants in this bilingual study were approximately four years old at the starting point of this project (3;11–4;2), when they started junior kindergarten (JK) in English, and approximately six years old at the end (5;11–6;2), when they finished senior kindergarten (SK) in English. The two male participants (anonymized as "Dan" and "Radu") and the female participant (anonymized as "Moni") were born in Canada to parents who spoke Romanian natively. The children's first and dominant home language was Romanian; their initial interactions and contacts were with the family members, peers, and caregivers who spoke only Romanian to them. Prior to entering school, their contact with English was limited to very sporadic media exposure and minor interaction with English-speaking children in the playground. All three were first-born children, and none of them had any siblings at the commencement of the study; however, a baby brother was born to one of the children (Radu) shortly after the study started. Parents reported that the children had no major health issues and had normal hearing and vision.

The bilingual children in this study are typical of a cohort of Canadian-born children of Romanian-speaking parents who arrived in Canada around the year 2000 and settled in the northern

end of Toronto within the municipality of Vaughan, Richmond Hill, and Aurora (Statistics Canada 2011). Their residency cohesion supports minority language maintenance as it provides varied input and plentiful opportunity for minority language use, factors that have been identified in successful language development and maintenance (De Houwer 2009). Furthermore, the three children came from university-educated families who valued both multilingualism and heritage language retention. At the age of four, the children were enrolled in English kindergarten, and at the end of the study, they were registered to begin Grade 1 in French immersion schools.

2.1.2. Monolingual Participants

As there are no standardized language tests or norms for monolingual Romanian-speaking children and no systematic study of lexical acquisition in that population, data were collected from six-year-old Romanian-speaking monolingual children in Romania (N = 22). While the scope of this part of the study was limited, we felt that it would be illuminating to examine the types of items that Romanian-speaking children of the same age would have acquired in the two different contexts, that is, the monolingual context in Romania and the bilingual one in Canada. The Romanian-born monolingual children were enrolled in the equivalent of full-time kindergarten ("Grupa Pregătitoare") in Sibiu, Romania, a historic city that is also a leading cultural centre in Europe. Like Canadian senior kindergarten, the Romanian "Grupa Pregătitoare", prepares children for the Grade 1 curriculum. In addition, as with the bilingual kindergarteners in the Canadian study, the Romanian participants belonged to middle-class families that encouraged extra-curricular activities such as sport, art, and personal development. In short, in terms of socio-economic status, the children in Romania were comparable to the three bilinguals.

2.2. Instruments

2.2.1. Peabody Picture Vocabulary Test, Fourth Edition

The measure used to assess the children's receptive vocabulary knowledge of English was the PPVT-4 (Dunn and Dunn 2007), which contains two versions, Form A and Form B. In order to avoid word recognition through repeated exposure to the same instrument, the forms were alternated between tests. Each of the 228 target items in each form consists of four coloured pictures as response options. The 228 items in each form are split into 17 sets with 12 items each. The items are listed by frequency in descending order (from the most common to the least). There are several reasons for using a receptive vocabulary test rather than a productive one in this study. First, because there is usually a lag between word comprehension and word production (Bates et al. 1994; Gershkoff-Stowe et al. 1997), we realized that there would be a problem eliciting word production in English at the beginning of this study. (Note that the children started interacting in English only at the age of four, the point at which the longitudinal study began.) Furthermore, even in an L1 (here, Romanian), children's production of words is contingent upon many factors, such as quick retrieval of the lexeme, ability to pronounce the word, or the child's level of confidence; moreover, when responding to a picture stimulus, more than one response could be correct, while others could be partially or mostly correct (e.g., if a leopard is called a cat) (Bates et al. 1994; Gershkoff-Stowe et al. 1997).

2.2.2. Romanian-Adapted Peabody Picture Vocabulary Test

As there are no standardized lexical tests for Romanian-speaking children, the peabody picture vocabulary test, fourth edition (Dunn and Dunn 2007), was adapted to Romanian and the resulting two forms, each with 228 items divided into 17 sets, were used as a measure of Romanian receptive vocabulary knowledge at the five points that coincided with the English testing.

The English stimuli were translated into Romanian with the help of two reputable dictionaries (Bantaş 1994; Leviţchi 2005), as well three native speakers of Romanian, one of whom was a linguist well versed in language acquisition, and two of whom were parents of young Romanian-speaking children.

As there is no established word frequency list for the Romanian lexicon, the Romanian-translated items closely match the English ones with respect to grammatical category, cultural importance, and relative importance in a child's life. For example, the English word *gigantic*, which is in the 6000 BNC (British National Corpus) word level was translated as *gigantic* in Romanian rather than by *uriaș* or *mare*, which are likely to be in the same frequency bands as *huge* (2000 word-level BNC) or *big* (1000 word-level BNC). In addition, culturally-biased items from the English PPVT-4 were replaced by words that are more familiar to monolingual Romanian children in a typical Romanian-speaking home. For instance, item 15 from Form A (target word *cookie*) and item 32 from Form B (target word *muffin*) were replaced by pictures and names of pastries that are encountered more frequently by Romanian children, for example *cozonac* ('sweetbread') and *biscuite* ('biscuit').

2.3. Procedure

The bilingual children were recruited through an information flyer posted on a well-known Romanian forum in Canada. Ethics approval was obtained from the University of Toronto (Protocol Reference #25110) and letters of consent were obtained from all participants (Appendix A). The monolingual children were recruited through directly contacting a kindergarten in downtown Sibiu.

The bilingual children were administered both the English PPVT and the Romanian-adapted version at each point of data collection, alternating the order of the languages tested. The tests were administered at five points over two years: at the beginning of the study in the fall (September) of junior kindergarten (Time 1—Form B), then in the spring (March) (Time 2—Form A), followed in the fall (September) of senior kindergarten (Time 3—Form B), then again in the spring (March) (Time 4—Form A), and concluded at the end of the study (August) (Time 5—Form B). The test was administered in the children's homes by the same researcher, who gave the instructions in Romanian for the Romanian test and in English for the English one.

Also, to avoid priming, the testing in the two languages took place at a minimum of two-week intervals. The task took approximately 20 to 30 min to administer individually to each child. For each item, the examiner said a word and the child pointed to the picture that best illustrated its meaning. The testing stopped when there were eight or more errors in a set of 12.

The monolingual children were administered Form B of the PPVT-4, as this was the form used with the bilingual children at the age of six, the same age as the monolingual participants. This allowed for an item analysis between the two groups.

2.4. Coding, Scoring, and Data Analysis

As the bilingual children in this study used Romanian at home and English at school, there was a possibility that different patterns of responses would emerge for the English PPVT-4 and its Romanian-adapted version. In order to explore such differences, all words from the 17 sets were first categorized as "home" or "academic" (see Appendix B). Scores were then compared at T1 and T5. Expanding on the criteria used by Bialystok et al. (2010), we included the following types of items in the home category: commonly experienced food and household items (e.g., *banana*, *lamp*), culture-specific items (e.g., *muffin*, *canoe*, etc.), frequently used clothing (e.g., *shoe*, *jacket*), household pets (e.g., *dog*, *cat*), frequent physical activities (e.g., *jumping*, *peeking*), high-frequency body parts (e.g., *mouth*, *knee*), common colours (e.g., *red*, *blue*), and words that are unlikely to appear in an academic context (e.g., *horrified*). Criteria for including the items in the academic category included the following: professions (*carpenter*, *dentist*), animals or plants (*hyena*, *cactus*), shapes (e.g., *rectangle*, *diamond*), musical instruments (*violin*, *clarinet*), low frequency body parts (e.g., *sternum*, *pelvis*), geographic locations (e.g., *peninsula*), and words used for academic tasks (e.g., *enumerating*, *composing*). Using the criteria mentioned above, two people independently classified all of the items from sets 1–17 for both forms. The inter-rater raw agreement was 97.92% and chance corrected agreement using Cohen's Kappa was 0.92, which also indicated very good inter-rater reliability. A consensus was reached on all disagreements and in the end, 54 items were classified as "home" and 138 as "academic".

Because low frequency items generally end up being assigned to the "academic" category and the high-frequency items are potentially assigned to the "home" category, the frequency level for each item was established using the British National and Lextutor corpora (www.lextutor.ca). The items in the "home" category ranged from level 1000 to 8000 frequency band, while those in the "academic" category ranged between 1000 and 17,000. Within the "academic" category, 79.71% of the items were in the 1000 to 8000 frequency band, while almost 20% (18.11%) of the items were in the 9000 to 17,000 frequency band. (Three items could not be found in the British National and Lextutor corpora, and, therefore, were not assigned to any frequency band; they were excluded from the present analysis). Although it is obvious that frequency was not an unambiguous classification factor, to eliminate confounding effects related to frequency, only those items with a frequency between 1000 and 8000 level for each of the two categories were included in the analysis of home versus academic repertoires. The final count for the home items was 52 and for the academic ones was 110.

In addition, each English word from the PPVT-4 (Form B) was rated as being a cognate or a non-cognate. Two native speakers of Romanian independently classified the items into cognates and non-cognates, basing the amount of similarity between the phonological form of the English and Romanian equivalents through the COSP (Kohnert et al. 2004). Based on Kelley and Kohnert (2012), who found COSP to be a suitable instrument in child bilingualism research, each Romanian–English pair was phonetically transcribed and assigned a value between 0 and 10, with 0 corresponding to a word pair that shared no phonological commonalities and 10 corresponding to a complete phonological overlap. To assign a COSP value, the scale in Table 1 was used for each cognate pair. For example, two consonants were considered to be similar sounds (score 1 in the category of initial sound overlap), if they shared at least one of the three features of place, manner, and voice, or at least one of the same sounds in a consonant cluster. For example, the English word *tuba* (*tuba* in Romanian) would be given a score of 1 for the initial sound overlap instead of 3 because the Romanian sound /t/ is a voiceless dental stop, unlike its English equivalent, which is a voiceless alveolar stop. The final score was determined by four features; namely, shared initial sound, shared number of syllables, shared consonants, and shared vowels (see Table 1). Note that the scale does not take into consideration fine-tuned distinctions between the acoustic and perceptual aspects of phonemes that are considered parallel in the two languages (e.g., the language-specific voice onset time of voiced and voiceless obstruent pairs in languages like Romanian and English).

Table 1. Cross-linguistic overlap scale for phonology to determine cognate status. [1]

Feature Overlap	Scoring	Example (from Romanian-and English)
Initial sound (0–3 points)	3 = Same consonant 2 = Same vowel 1 = Similar sound (e.g., same sound class or one element of a consonant cluster) 0 = Complete mismatch between initial sounds	*banana*—'banana' *injecteaza*—'injecting' *tunel*—'tunnel' *cerebral*—'cerebral'
Number of syllables (0–2 points)	2 = Equal number of syllables 1 = Different by only 1 syllable 0 = Different by more than 1 syllable	*vehicol*—'vehicle' *ferma*—'farm' *florareasa*—'florist'
Consonants (0–3 points)	3 = >70% consonant overlap 2 = 50–70% consonant overlap 1 = ≤50% consonant overlap 0 = No consonant overlap	*sedan*—'sedan' *atlet* 'athlete' *vioară*—'violin' N/A
Vowels (0–2 points)	2 = ≥80% vowel overlap 1 = 50–80% vowel overlap 0 = <50% or no vowel overlap	*harmonica*—'harmonica' *lichid*—'liquid' *cerc*—'circle'

[1] Adapted from Crossing borders: recognition of Spanish words by English-speaking children with and without language impairment, by Kohnert et al. (2004).

The PPVT-4 contained words whose COSP scores ranged from 0 to 10. For instance, banana had a score of 10; ferma, farm had a score of 7; and autobus, bus had a score of 1. The average COSP score on

the PPVT-4 Form B was 6.59 (SD = 2.43). Cognates with COSP scores between 0 and 5 were excluded from the analysis. This cut-off point was chosen based on empirical evidence from Kohnert et al. (2004), who found that the majority of the monolingual speaking adults correctly guessed the English translation for 15–50% of Spanish words with COSP scores from 6 to 9, but did not guess the English translation for Spanish words with COSP scores lower than 5. The raw agreement between raters for assigning COSP scores for Form B was 95.31% and chance corrected agreement using Cohen's Kappa was 0.90 (i.e., high inter-rater reliability). Consensus was reached on all disagreements and all test items were classified as either cognate items (total of 70 items from sets 1–16) or non-cognate items (total of 122 items from sets 1–16). After COSP scores were assigned, the semantic overlap between pairs of cognates was examined and only those items that were identical or near-identical in meaning were selected for the analysis. For instance, the pair *marsupiu*–'marsupial' with a COSP score of 9, has identical meanings and was thus selected for the analysis. In contrast, the pair *fizician*–'physician', despite having a COSP score of 9 and being etymologically related, was excluded from the analysis because *fizician* means 'physicist' in Romanian.

In order to analyze the results in more fine-tuned ways, the data were also grouped as home/non-cognate, home/cognate, academic/non-cognate, and academic/cognate. These categories were used to compare the performance of the bilinguals in Romanian and English at T5, as well as the performance of the bilingual and monolingual children in Romanian at the age of six. By breaking up the categories in this fashion, the results can be viewed without the confounding overlap between academic and cognate words because, in both English and Romanian, the academic register tends to contain more Latinate words than is the case with non-academic high frequency items (Petrescu et al. 2017).

3. Results

As indicated above, the current study aimed to address three research questions. First, the study aimed to look at patterns of lexical development in Romanian–English bilingual children. Next, the study set out to determine whether there is any evidence of cognate facilitation among bilinguals during their kindergarten years. Finally, the study aimed to identify similarities and differences in types of vocabulary acquired by monolingual Romanian-speaking children and the Canadian-born Romanian–English bilingual children.

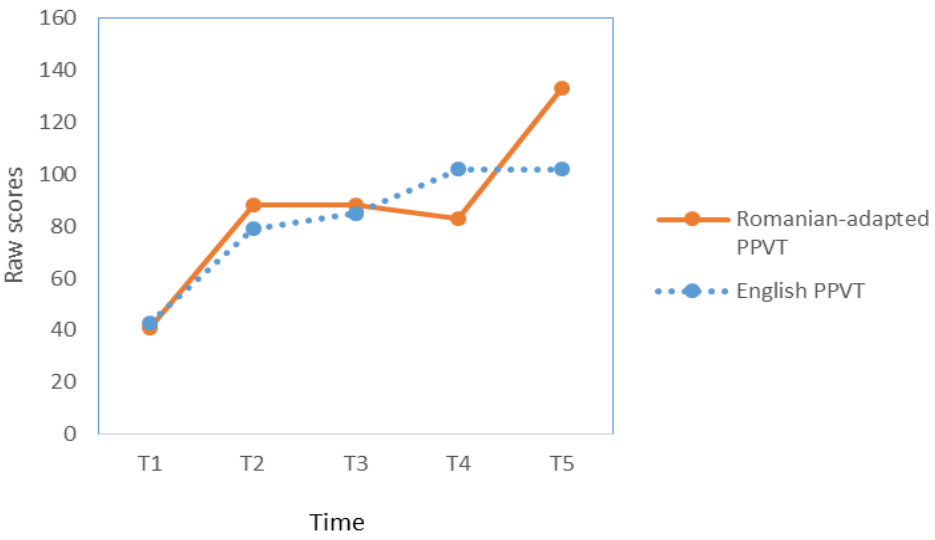
3.1. Lexical Development of the Three Romanian–English Bilinguals

3.1.1. Peabody Picture Vocabulary Test, Fourth Edition Overall Scores from T1 to T5 in English and Romanian-Adapted Versions

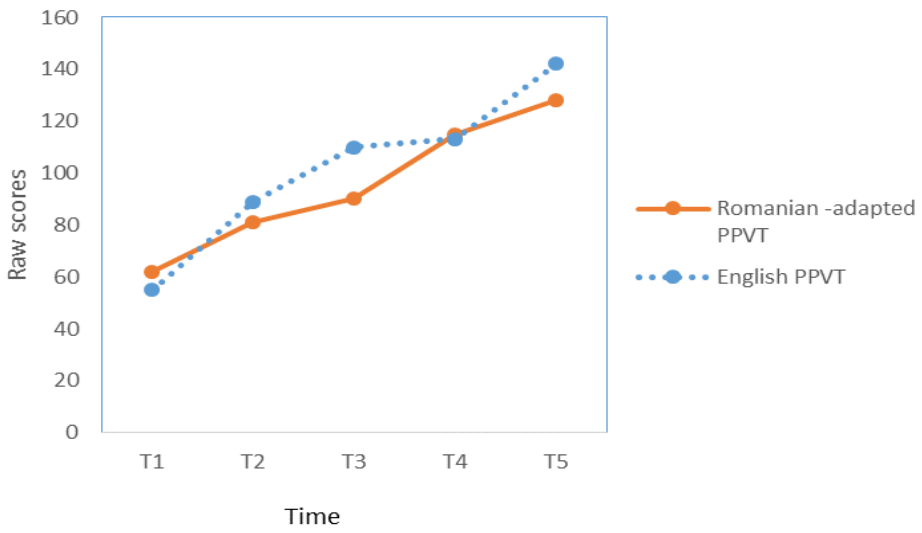
Raw scores and percentiles for the English PPVT-4 as well as the raw scores for the Romanian-adapted PPVT-4 have been obtained. In Figure 1a–c, the children's raw scores for the English PPVT-4 and Romanian-adapted PPVT-4 are plotted at each round of data collection.

Figure 1a reveals that Dan's English and Romanian raw scores were near equal at the start of the study (T1) and continued to grow over the two-year period, with Romanian surpassing English at T5. The steepest growth curve for English appeared between T1 (when Dan started kindergarten) and T2 (after six months of English instruction). For Romanian, Dan experienced a sudden surge in his receptive vocabulary between T1 to T2 and between T4 to T5 (end of kindergarten) following a four month visit to Romania.

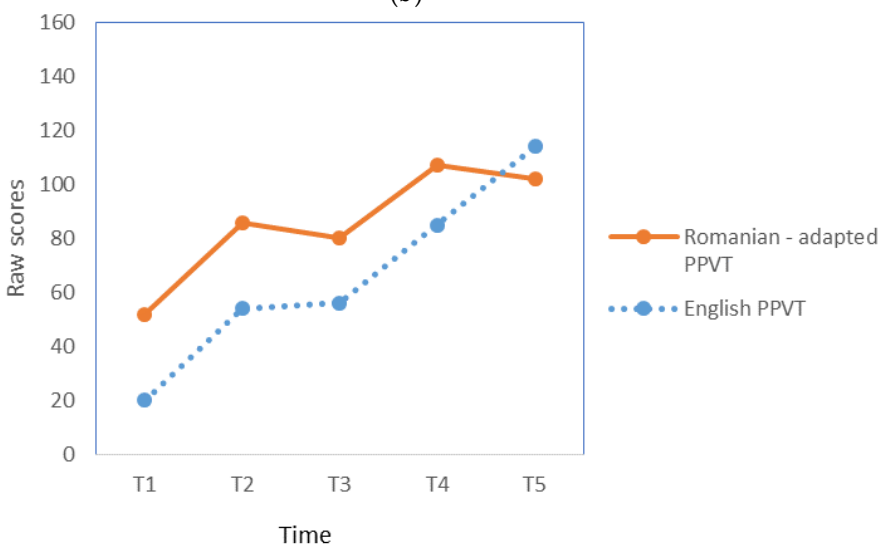
Figure 1b reveals that Radu's raw scores for both Romanian and English were near equal at T1, with Romanian showing a small advantage. Both languages continued to develop over the two-year period. Radu's English raw scores increased suddenly between T1 and T2, and then again between T4 and T5. The development of his Romanian receptive vocabulary from T1 to T5 was steady without periods of stagnation or attrition.



(**a**)



(**b**)



(**c**)

Figure 1. Peabody picture vocabulary test (PPVT)—raw scores for English and Romanian from T1 through T5 for (**a**) Dan; (**b**) Radu; and (**c**) Moni.

Figure 1c shows that at T1, Moni's raw scores for Romanian were higher than the raw scores for English and remained so until T5, when her English score surpassed the Romanian one. Both Romanian and English raw scores continued to grow over the two-year period.

Figure 2 reproduces the plot for percentile scores for the English data for all three children. Overall, all three children's percentile scores improved substantially from T1 to T5, with Radu and Moni exceeding the mean scores at T5 and Dan arriving at the mean at T5. (Note that no corresponding summary can be provided for Romanian because there are no standardized tests for receptive vocabulary).

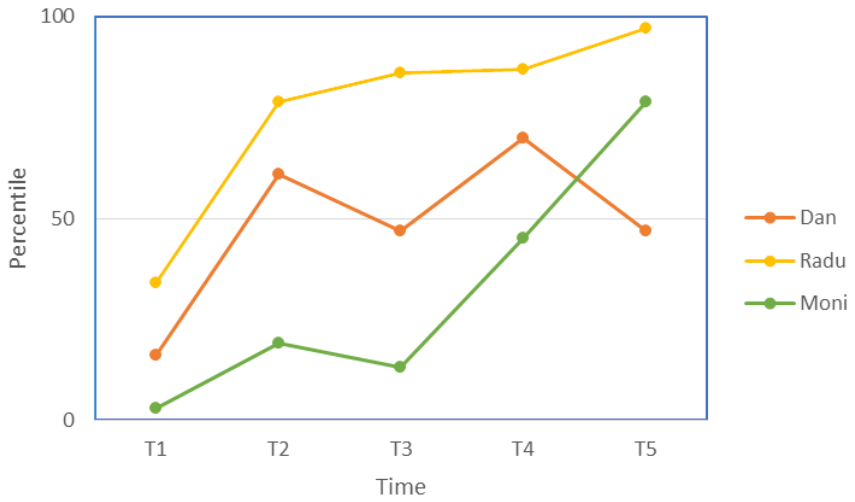


Figure 2. Peabody picture vocabulary test English percentile for each participant from T1 to T5.

3.1.2. Item Category Analysis

The PPVT-4 data were further analyzed using two conceptual frameworks: home/academic category and cognate/non-cognate category. (Note that only Form B of the PPVT-4 was used for these comparisons, making the scores at T2 and T4, based on Form A unsuitable for a comparison of congruent items.) As the means reported are raw, in each category, the maximum number of items needs to be considered as a guideline. Percentages are not being reported as they are not suitable given that the PPVT-4 works with age-appropriate percentile scores.

Scores at T1, T3, and T5 for the Home and Academic Items in Romanian and English

To explore whether certain portions of the children's vocabulary are affected by the context where they are used (home or school), the items in the PPVT-4 were classified as "home" and "academic". Of the total number of items present in Form B of the PPVT-4, 54 items were classified as "home" and 138 as "academic". However, as discussed in the method section, the analysis excluded items that were beyond the 8000-frequency level, leaving 52 home items and 110 academic items. Table 2 presents the distribution of the items in each home/academic category within the frequency levels.

Table 2. Frequency levels within the academic and home categories on the Peabody picture vocabulary test (PPVT)-4.

Form	BNC Frequency Level	Home Raw # (Max = 54)/Percentage	Academic Raw # (Max = 138)/Percentage
Form B	<8000 level	52/96.29%	110/79.71%
	>8000 level	2/3.70%	25/18.11%
	Unknown	0%	3/2.17%

BNC: British National Corpus.

As can be seen in Table 2, three items (amounting to 2.17% of the total) could not be found in the BNC and Lextutor corpora, and therefore, assigning these words to a frequency band was not possible. For this reason, they were excluded from the present analysis. The results of the analysis are presented in Figure 3a–c.

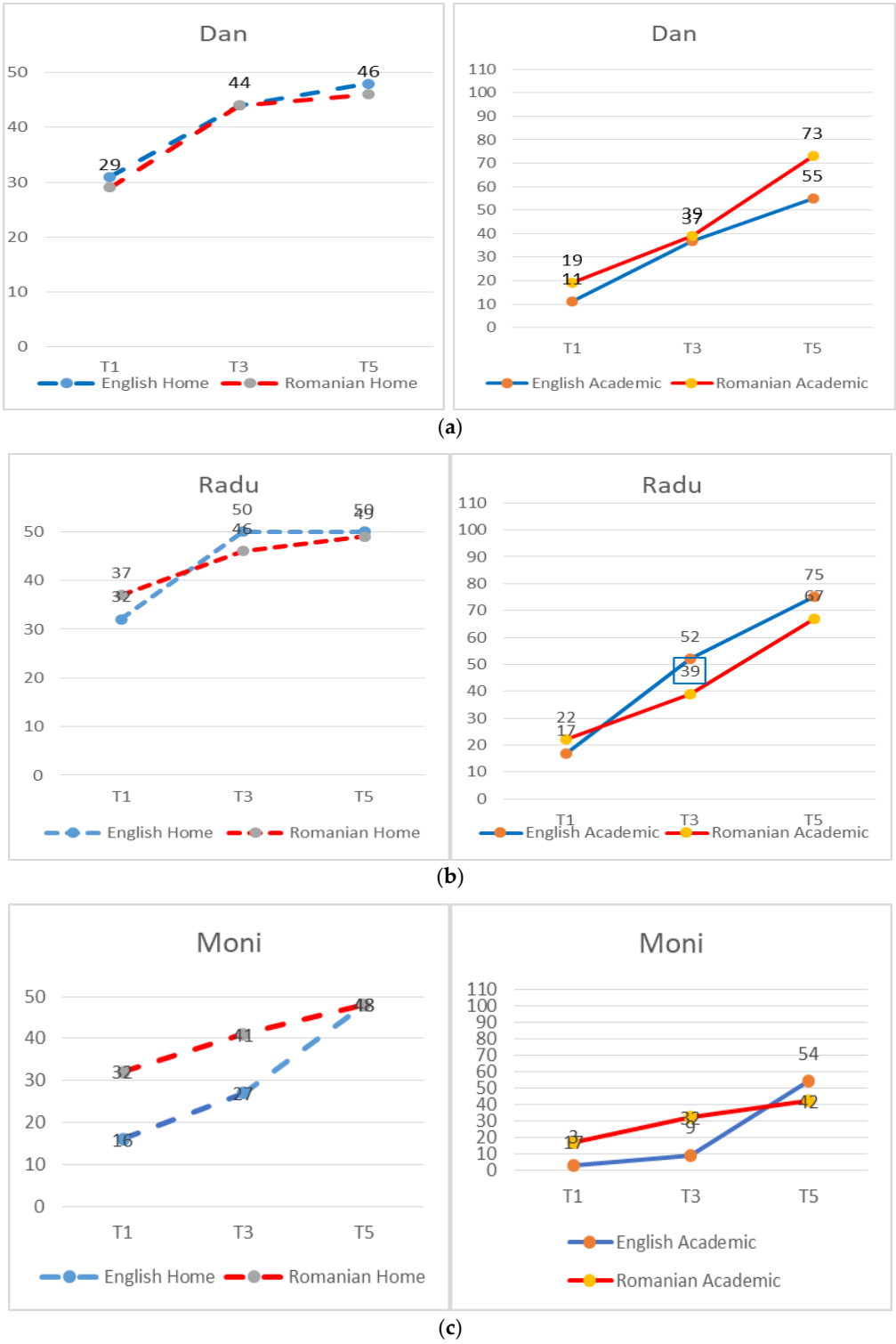


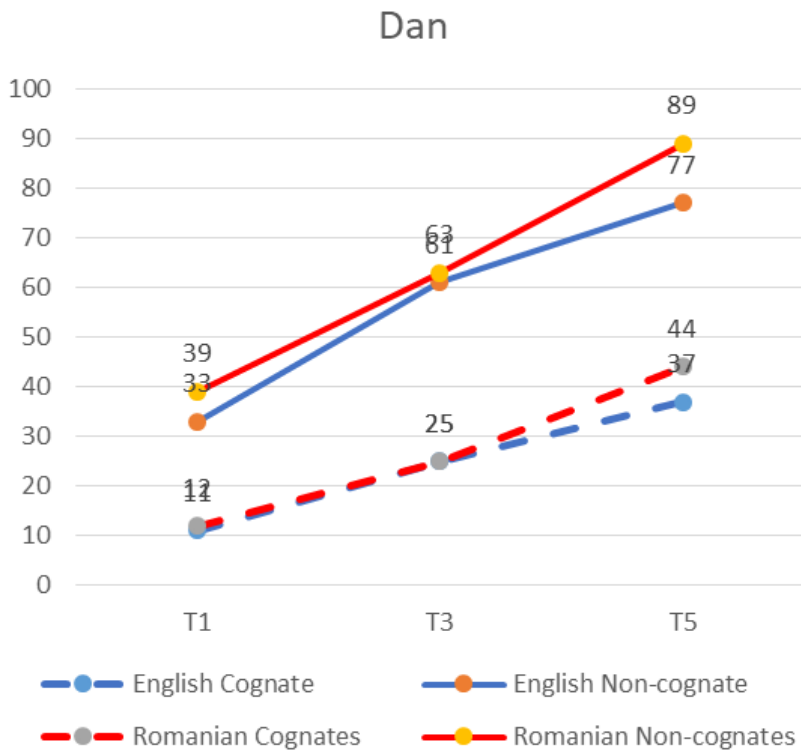
Figure 3. PPVT 4—Raw scores for Romanian and English home (max = 52), and Romanian and English academic (max = 110) for (**a**) Dan; (**b**) Radu; and (**c**) Moni.

Figure 3a–c reveals that all three children had higher scores proportionally in their home items in both English and Romanian. By T5, their scores were close to the maximum (52). All three children had lower means proportionally in the academic section in both Romanian and English. In the academic section, Dan and Radu experienced the sharpest increases between T1 and T5 in both English and Romanian. Moni experienced a sharp increase mainly between T1 and T5 mainly in English (this was true also of a sharp increase in her English home vocabulary between T1 and T5).

Scores at T1, T3, and T5 for the Cognate and Non-Cognate Items in Romanian and English

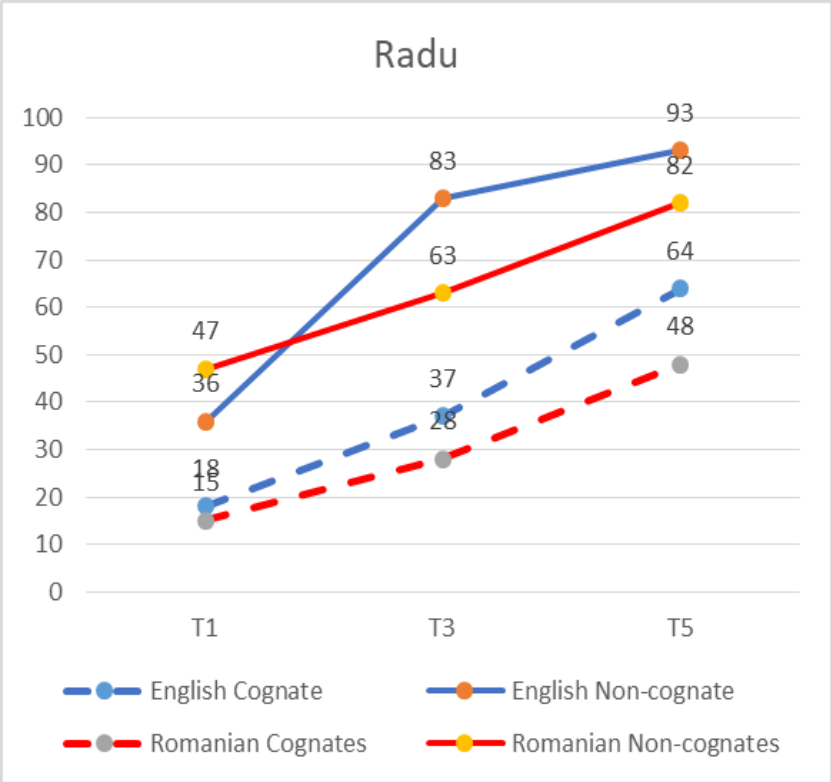
In order to investigate cross-linguistic influence in terms of possible cognate facilitation, the items from the English PPVT-4 tests were classified as either cognates or non-cognates on the basis of the cross-linguistic overlap scale for phonology (COSP) discussed in Section 2.4. As explained earlier, only the items that had a COSP score higher than 5 and those that were identical or near-identical semantically were included in the present analysis. After this screening, 122 items were classified as non-cognates and 70 items were classified as cognates.

We examined the possibility of cognate facilitation through how steep the upward trend was from T1 to T5. In the present case, the curve for non-cognates was steeper than that for cognates (see Figure 4). If anything, improvement was slightly higher for the non-cognates in some cases, for example, Moni's English scores and Radu's Romanian scores.

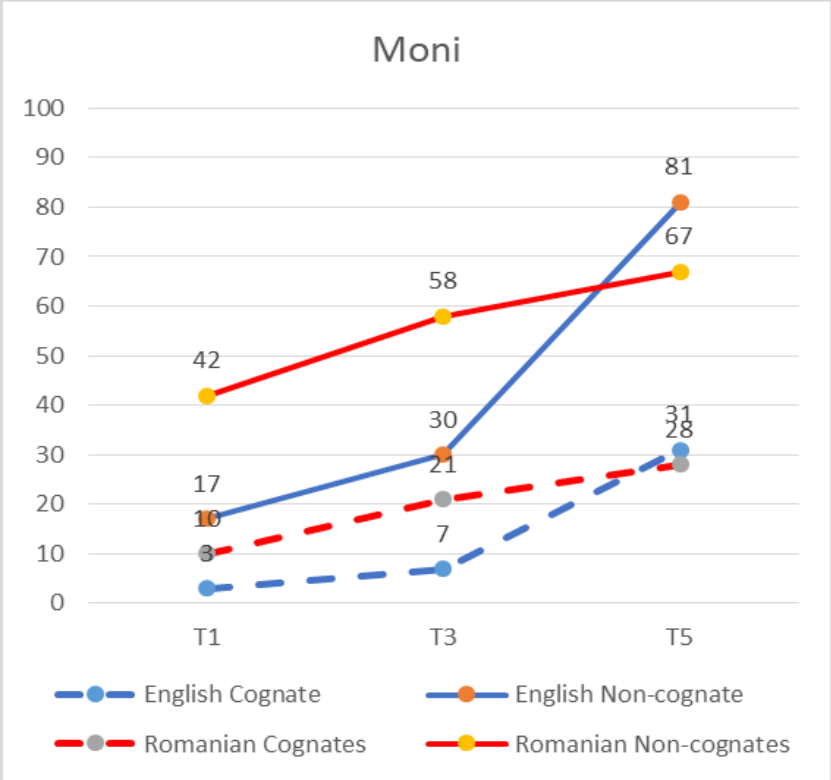


(**a**)

Figure 4. *Cont.*



(**b**)



(**c**)

Figure 4. PPVT 4—Raw scores for cognate/non-cognate categories (max 70 and 122, respectively) for Romanian and English (**a**) for Dan; (**b**) Radu; and (**c**) Moni.

3.1.3. Overview of the Bilingual Children's Scores in Romanian and English at T5

Table 3 analyzes the means and standard deviations in the four categories mentioned above: home versus academic and cognate versus non-cognate. As explained in Section 2.4, as there is overlap between academic and cognate repertoires in a language like English when examined alongside a Romance language like Romanian, the items were further broken up into four non-overlapping categories; namely, home/non-cognate, home/cognate, academic/non-cognate, and academic/cognate. A close examination of these newly created categories indicates the bilinguals' performance in the home/non-cognate and home/cognate categories were close to the maximum in English and Romanian and do not provide any evidence for cognate facilitation. The academic categories are lesser known in both English and Romanian, and once again provided no evidence for cognate facilitation. We should note that, while the raw scores for academic/non-cognates were slightly higher in English, the raw scores for academic/cognates were a little higher in Romanian. In short, the bilinguals' mean scores were very similar in their two languages.

Table 3. Mean raw scores of the bilingual children in Romanian and English at T5 (age six).

Participants	Home (Max = 54)	Academic (Max = 138)	Cognate (Max = 70)	Non-Cognate (Max = 122)	Home-Non-Cognate (Max = 39)	Home-Cognate (Max =15)	Academic-Non-Cognate (Max = 83)	Academic Cognate (Max = 55)	Total (Max = 192)
Bilinguals (*n* = 3) ENGLISH	M = 50.33 (SD = 1.53) Range: 49–52	M = 71.67 (SD = 15.89) Range: 62–90	M = 38.33 (SD = 9.45) Range: 31–49	M = 83.7 (SD = 8.33) Range: 77–93	M = 37.0 (SD = 1.0) Range: 36–38	M = 13.3 (SD = 0.58) Range: 13–14	M = 46.67 (SD = 7.37) Range: 41–55	M = 25.0 (SD = 8.89) Range:18–35	M = 122.0 (SD = 17.3) Range: 112–142
Bilinguals (*n* = 3) ROMANIAN	M = 49.33 (SD = 1.53) Range: 48–51	M = 70.0 (SD = 21.0) Range: 46–85	M = 40.0 (SD = 10.58) Range: 28–48	M = 79.33 (SD = 11.24) Range: 67–89	M = 36.33 (SD = 1.15) Range: 35–37	M = 13.0 (SD = 1.0) Range: 12–14	M = 43.00 (SD = 12.12) Range: 30–54	M = 27.0 (SD = 9.64) Range: 16–34	M = 119.33 (SD = 21.13) Range: 95–133

M: mean; SD: standard deviation.

3.2. Overview of the Lexical Knowledge of Monolingual Children Compared to Bilingual Children at Age Six

A list with all the academic items and the correct answers for each group is presented in Appendix C. All the "home" items are listed in Appendix D. Overall, the Romanian children had a bigger vocabulary size than their bilingual counterparts. Their mean score was 134.86, while the bilingual children's mean raw score was 119.33. However, a Kruksal–Wallis non-parametric test indicated that the difference was not significant ($p = 0.13$). Table 4 presents the bilingual and monolingual performance at age six on sub-classes of words on the English and Romanian PPVT-4.

Table 4. Bilingual and monolingual performance at age six on sub-classes of words on the English and Romanian PPVT-4.

Participants	Home (Max = 54)	Academic (Max = 138)	Cognate (Max = 70)	Non-Cognate (Max = 122)	Home-Non-Cognate (Max = 39)	Home-Cognate (Max = 15)	Academic-Non-Cognate (Max = 83)	Academic Cognate (Max = 55)	Total (Max = 192)
Bilinguals (*n* = 3) ROMANIAN	M = 49.33 (SD = 1.53) Range: 48–51	M = 70.0 (SD = 21.0) Range: 46–85	M = 40.0 SD = 10.58 Range: 28–48	M = 79.33 (SD = 11.24) Range: 67–89	M = 36.33 (SD = 1.15) Range: 35–37	M = 13.0 (SD = 1.0) Range: 12–14	M = 43.00 (SD = 12.12) Range: 30–54	M = 27.0 (SD = 9.64) Range: 16–34	M = 119.33 (SD = 21.13) Range: 95–133
Monolinguals (*n* = 22) ROMANIAN	M = 50.14 (SD = 2.88) Range: 42–54	M = 84.73 (SD = 19.87) Range: 41–110:	M = 45.41 (SD = 9.84) Range: 21–57	M = 89.45 (SD = 12.73) Range: 62–106	M = 36.91 (SD = 1.77) Range: 31–39	M = 13.23 (SD = 1.45) Range: 10–15	M = 52.55 (SD = 11.43) Range: 35–67	M = 32.18 (SD = 8.75) Range: 10–43	M = 134.86 (SD = 22.30) Range: 83–163
p values for bilingual/monolingual differences (Kruksal–Wallis test)	0.42	0.13	0.13	0.18	0.38	0.64	0.11	0.017	0.13

According to Table 4, the home items were known to an equal extent by both groups. The academic items were better known by the monolingual children (M = 84.73; SD = 19.87 for monolinguals vs. M = 70.0; SD = 21 for bilinguals).

For the monolingual group, a Wilcoxon test revealed a significant difference between home and academic scores in favour of home scores ($p < 0.0001$). Note that comparisons between

cognates and non-cognates within the monolingual group are not relevant. We have included the cognate/non-cognate distinction in Table 4 only for the purposes of comparing their scores with the bilinguals to determine whether the bilinguals enjoyed any cognate facilitation.

When the bilinguals were compared with the monolinguals, using the Kruksal–Wallis non-parametric test, in each of the categories labeled in Table 4, none of the differences were significant. The means and these non-significant values confirm that the bilinguals did not benefit from cognate facilitation. When we look at the means of the monolinguals and bilinguals, we see that in all of the home categories their scores are very similar. However, the monolinguals have much higher means in the academic categories.

4. Discussion

The results indicate that the receptive vocabulary of the three Canadian-born children increased with age in both Romanian and English. In English, their scores could actually be viewed in light of established norms. While all started below the 50th percentile at T1 at the beginning of the study (when they started kindergarten), they scored at or above the 50th percentile at T5 (at the end of their kindergarten years). Their sharpest increase occurred during the first six months of attending kindergarten, reflecting the rapidity of L2 lexical growth at least in the context of this study.

There were, however, some noteworthy differences in the children's patterns of lexical development. One factor that we could isolate as impacting both Romanian and English was the length and location of the children's summer vacations. Both Radu and Dan, for example, experienced their steepest growth in Romanian vocabulary after month-long visits to Romania at the beginning of kindergarten Year 1 and kindergarten Year 2, underlying the powerful impact of being fully immersed in a heritage-language milieu. Moni's visits to Romania, on the other hand, were relatively brief, which might explain the more gradual increase in her Romanian vocabulary. Conversely, Dan's English scores showed peak performance at T2 and T4 (six months into kindergarten Years 1 and 2, respectively) and lowest performance at T3 and T5 (at the beginning and end of kindergarten Year 2). A plausible explanation for this behavior is again Dan's month-long trips to Romania over the summer holidays. After full immersion in the heritage language environment, Dan's English skills seem to have regressed, attesting to the fact that the linguistic soundscape of the children is highly dynamic and plays a crucial role in the balance between the children's two languages (De Houwer 2009). The finding is also in line with previous research that suggests that summer vacations in the home country positively impact receptive language skills in the minority language of bilingual children (Hammer et al. 2008; Rojas and Iglesias 2013) and negatively impact the school language.

While two years of schooling in English appear to have erased the gap between the bilingual children and their English-speaking monolingual counterparts, there appear to be some striking differences between the bilinguals' Romanian vocabulary when compared with their 22 monolingual counterparts in Romania. As expected, the size of the monolinguals' vocabulary was greater than that of their Canadian counterparts. In short, while the English monolingual and bilingual comparisons in this study did not show a smaller English vocabulary in the bilingual children, contrary to expectations (Bialystok and Feng 2009; Werker 2012), bilingual and monolingual comparisons in the heritage language did seem to be compatible with the general belief that, as bilingual children's input and interaction is divided between two languages, their lexicons in each show gaps in some content areas (academic vocabulary). The difference in the size of the monolinguals' and bilinguals' academic vocabulary can be attributed to the bilinguals' lack of opportunity to study academic content in Romanian. The following section explores some differences between the words known by the monolinguals and bilinguals.

4.1. Comparing the Compositions of the Monolingual and Bilingual Romanian Lexical Repertoires

As the purpose of the study was more than to simply quantify the receptive lexical knowledge of the monolingual and bilingual participants, but also to look at the types of items in each groups'

repertoires, looking at the actual items in their home and academic vocabularies shed some light on the similarities and differences in the environments of Romanian lexical acquisition in Canada and Romania. A comparison of the academic items known by the bilingual children and their monolingual counterparts reveals some commonalities. Neither group was familiar with certain words related to geometry (*sferic*—'spheric', *parale*—'parallel', *concav*—'concave'); low-frequency words related to body parts (*bazin*—'pelvis'; *maxilar*—'jaw'); archaic words (*pocal*—'goblet'; *teacă*—'pod'); or words related to different trades (*scripete*—'hoisting'; *lubrifiază*—'lubricating'; *vitezometru*—'speedometer'). On the other hand, certain academic items were known by both groups, for instance, *creier* ('brain'), *astronaut* ('astronaut'), and *binoclu* ('binoculars'). However, there were many words that the monolingual children knew, probably through classroom activities, for example, *hartă* as in 'map', *marsupiu* as in 'marsupial', or *fundație* as in 'foundation'. Other words that were known by the Romanians monolinguals and not by the Romanian–English bilinguals reflected cultural differences between their communities. For example, words such as *asamblează*—'assembling', *fundație*—'foundation', *mistrie*—'trowel', *tubular*—'tubular' are often encountered in a domestic context, as many Romanian families build, repair, or modify their homes on their own. Moreover, these Romanian children have probably had the opportunity to encounter words such as *stup*—'hive', *țap*—'goat', *mânz*—'colt', and *păun*—'peacock' in their grandparents' homes (as many still live in villages and raise their own animals).

4.2. Cognate Facilitation

As discussed above, to investigate whether young bilingual children are sensitive to Romanian–English cognates with overlapping meanings in spoken form, we examined the children's scores on cognates and non-cognates in the PPVT-4. In the bilingual study conducted in Canada, the results were mixed. One of the children, Radu, seemed to show evidence of cognate facilitation at both T1 and T5, while Dan and Moni showed no cognate advantage at any stage. As suggested by Kelley and Kohnert (2012), such individual variation is frequent in the study of bilingual children "even within any well-defined relatively homogenous sample" (p. 200). In our comparison between the Canadian bilinguals and the Romanian monolinguals, there were only a small handful of cognates that the bilinguals knew, and the monolinguals had difficulty with the following: *tripleți*—"triplets", *jogging*—"jogging", *pedală*—"pedal", and *fictiv*—"fictive". However, there were numerous other potentially recognizable cognates that the monolinguals knew (despite having no second language), but that the bilinguals clearly did not recognize. The means and the non-significant statistical differences recorded in Table 4 provide further evidence of the lack of cognate facilitation among the three bilinguals.

The literature cautions that identifying cognates is a skill that develops with age; proficiency in the two related languages; and occasionally only through instruction, for example, when children begin to read for comprehension, as it is not intuitively developed at a young age (García and Nagy 1993; Kelley and Kohnert 2012; Malabonga et al. 2008). The three bilingual children in our study were younger than the age at which children are apt to recognize cognates (Grade 5, as reported in Malabonga et al. 2008). They are also very new learners of English, having only started interacting in the language at the age of four.

4.3. Liitations of This Study

The design of the two-year longitudinal bilingual study provided some much-needed insights into first language development and maintenance of Romanian in a large urban centre in Canada, where English is the mainstream language. However, it is obvious that such a study needs to be followed up by research involving larger sample groups, especially as Romanian is of growing importance as a heritage language in the Greater Toronto Area. Larger numbers would clearly lend power to any statistical tests that chart bilingual progress or make comparisons between groups. While the monolingual study featured a larger sample group (n = 22), one of its major shortcomings lays in

the fact that data collection was limited to a single point in time (i.e., when children were six years old). Longitudinal or larger cross-sectional studies on the acquisition of Romanian in Romania would certainly provide some yardsticks for viewing Canadian heritage language data in perspective.

A further limitation lies in the fact that the data analyzed in this study encompass receptive vocabulary knowledge alone. Ideally, such data can be greatly enriched by naturalistic or elicited production data because the latter provides information about the morphosyntactic, semantic and pragmatic accuracy of word use in a contextualized manner. As we know, correct responses in multiple-choice receptive tests can result from guessing through elimination, and even if not the product of guessing, correct responses do not necessarily reflect an accurate or rich mental representation of the item (Gyllstad et al. 2015).

While we made progress in creating a version of the PPVT-4 for Romanian by consulting parents of preschool children in Romanian-only households in a specific neighborhood in the Greater Toronto Area and pilot testing the items with older children, one major drawback of this newly designed test lay in the lack of reliable frequency data for the Romanian items. As the test items were, for the most part, translation equivalents of the English items, there is a distinct possibility that many were not in the same frequency range as their English equivalents. Second, the lack of any Romanian norms for tests of children's vocabulary made it impossible to consider the bilingual children's performance alongside that of standardized monolingual scores.

A related concern relates to the use of the PPVT-4 to explore cognate facilitation. While examining cognate facilitation is apropos or even necessary in a study involving two lexically proximate languages like Romanian and English, the PPVT-4 is not created to investigate cognate effects in a controlled manner. It was thus difficult to include a sufficient number of relevant items in the test, that is, those with a COSP of 5 and over, to obtain conclusive results within the bilingual study or across the monolingual and bilingual studies. Thus, what we obtained was only a gross measure of cognate facilitation (or the lack thereof).

5. Implications for Research and Testing

This study focuses on language acquisition in a unique context. There have been no previous studies involving Canadian-born children with Romanian as their first language and English as their second. Moreover, even if we consider contexts beyond a Canadian one, barring a few notable exceptions (Avram 2001; Buja 2008), Romanian has been conspicuously absent in mainstream child acquisition research. The study also makes a contribution to minority language retention in a situation where the school language, English, is not only the mainstream language of the community, but is also the language of world media, children's electronic games and TV shows, and global communication. By using a longitudinal design, the current study also captures changes in the children's minority and majority language lexicons at multiple points, allowing for new insights on the transition from home-based monolingualism to community-based bilingualism. This unique design adds to the few studies involving Romanian as a heritage language that have been conducted in different contexts and with a different focus (Buzilă 2016; Mesaros 2008; Montrul et al. 2015; Nesteruk 2010).

These insights notwithstanding, in light of the new Romanian diaspora, especially in countries like Canada, the time is ripe for larger studies investigating the acquisition of Romanian in both monolingual and bilingual contexts, and involving participants of various ages. As discussed earlier, reliable frequency data in Romanian as well as standardized Romanian tests would greatly enhance acquisition research in the language. Such studies could also provide a clearer picture of cross-linguistic influence involving Romanian L1 and English L2, including cognate facilitation, especially because Romanian, like French and other Romance languages, has the Latinate counterparts of low-frequency English academic words, but occurring in its higher frequency bands (Cummins 2005; Petrescu et al. 2017).

The development of the Romanian-adapted PPVT-4 used to test the bilingual and monolingual children was also a novel contribution of this study. However, it is clear that such a test is only

a good starting point for creating a Romanian version of the PPVT based on frequency data and validated through trial runs involving large sample groups consisting of participants of not only different ages, but also different socio-cultural milieus. Perhaps, like any test being used in both monolingual and bilingual situations, a valid and reliable Romanian PPVT needs to be based on not only general frequency data, but also frequency data from specific communities. For example, in our study, the monolingual Romanian children knew words such as *stup*—'hive', *mistrie*—'trowel', and *manz*—'colt', presumably because these words are frequently used in their cultural context. It is much less likely that pre-schoolers growing up in an urban centre in North America would be exposed to these items at home or would even find them useful in their interactions within the Romanian-speaking community. One potential solution would be to build different versions of a lexical test for bilinguals and monolingual that take ethnographic details into consideration. Another solution would be to build a certain amount of flexibility into one test package, for example, by allowing for alternatives to certain culture-specific test items.

Author Contributions: M.C.P. and R.H.-P. designed the study and selected the data collection method collaboratively; M.C.P. collected data over a period of 2 years; The first round of data collection was conducted by M.C.P., and the second by both M.C.P. and R.H.-P.; Both M.C.P. and R.H.-P. wrote the article.

Funding: This research received no external funding.

Acknowledgments: We greatly appreciate the insightful and relevant comments made by the reviewers as well as the excellent guidance and oversight provided by the academic editors. The mechanism for submitting manuscripts and responding to reviews was excellent. We would also like to thank Pablo Velázquez for his clear and quick responses to our queries throughout the review process. Finally, we would like to extend our grateful thanks to Tamara Al-Kasey for helping us edit the first draft of our paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Parent Consent Form

I have read and understood the letter describing the proposed study, titled "Minority Language Acquisition and Retention: A Study of Canadian-Born Romanian-Speaking Bilingual Children". I understand that my child will be participating in language activities so that the researcher can examine my child's language and language comprehension and production. All information collected will be kept confidential. There are no risks involved and there are financial direct benefits to me, and my child. The study will increase researchers' understanding of children's language development. I understand that my child will be asked if s/he wants to take part and will not be required to do so if s/he is shy or unwilling. Also, s/he may stop and return to the classroom at any time without penalty. I may withdraw my consent at any time. If I do not give permission or if my child does not want to participate, I understand that it will not affect me or my child in any way

Child's Name _____
(please print)

Child's Date of Birth _____
(month/day/year)

I give my permission for my child, named above, to participate in the University of Toronto study conducted by Maria Claudia Petrescu.

_____	_____	_____
Printed name of Parent/Guardian	Signature	Date

If there are any questions about this research, please contact Maria Claudia Petrescu, Ph.D. Candidate OISE/University of Toronto, 252 Bloor Street West, Toronto, Ontario M5S 1V6, Canada;

maria.petrescu@utoronto.ca, or Rena Helms-Park, Associate Professor, OISE/University of Toronto, 252 Bloor Street West, Toronto, Ontario M5S 1V6, Canada, 416-978-0277.

Ethics office:
McMurrich Building, 12 Queen's Park Cres. W, 2rd Floor Toronto, ON M5S 1S8
TEL: 416-946-3273 FAX: 416-946-5763 EMAIL: ethics.review@utoronto.ca

Appendix B.

Appendix B.1. PPVT—Home Items

- Basic colours and numbers
- Common household items
- Commonly used clothing
- Common musical instruments
- Common fruits and vegetables
- Common food items found in most homes
- Household pets
- Common domesticated animals and prevalent wild animals
- Words related to playing and other common physical activities
- Commonly discussed body parts
- Culture specific items
- Words most likely to be learned by oral dialogue in the home

Appendix B.2. PPVT—Academic Items

- Uncommon colours
- Uncommon household items (vase, safe, aquarium, luggage)
- Uncommon items of clothing
- Uncommon musical instruments
- Uncommon food type
- Uncommon domesticated and wild animals
- Plants and associated terminology
- Rarely discussed body parts (sternum)
- Uncommon means of transportation (glider, sedan)
- Geographical types (valley)
- Landscaping (fountain)
- Uncommon infrastructure (hydrant)
- Building components (chimney)
- Shapes and geometric patterns
- Professions/occupations and associated attire/occupational tools or associated terminology
- Professional activities
- Words reflecting school experience

Appendix C. Academic Items from PPVT-4 (Romanian-English)

Table A1. Academic Items from PPVT-4 (Romanian-English).

Academic Items in Romanian (with English Translations)	Correct Canadian Responses (Max = 3)	Romanian. Number (Max = 22)/Percentage
Lacăt ('Lock') *Creier* ('Brain') *Astronaut* ('Astronaut') *Magar* ('Donkey') *Cerc* ('Circle') *Steag* ('Flag') *Poarta* ('Gate') *Cangur* ('Cangaroo') *Hambar* ('Barn') *Scrie* ('Writing') *Birou* ('Desk') *Barbie* ('Chin')	3	100
Monedă ('Coin') *Statuie* ('Statue')	2	100
Rupe ('Tearing') *Dezumflat* ('Deflated') *Binoclu* ('Binoculars') *Munte* ('Mountain') *Corn* ('Horn') *Omida* ('Caterpillar')	3	95
Imens ('Huge') *Gaureste* ('Drilling') *Broasca* testoasa ('Tortoise')	1	95
Triunghi ('Triangle')	3	91
Pachet ('Package') *Radacina* ('Root') *Ingrozit* ('Terrified') *Pelican* ('Pelican') *Tornada* ('Tornado') *Lichid* ('Liquid')	2	91
Stup ('Hive')		91
Harta ('Map') *Lant* ('Chain') *Buchet* ('Bouquet')	3	86
Carlig ('Hook') *Trage* ('Tugging') *Unghi* ('Angle')	2	86
Paun ('Peacock') *Piramida* ('Pyramid')	1	86
Nara ('Nostril') *Electrician* ('Electrician')	3	82
Pereche ('Pair') *Copita* ('Hoof') *Calculeaza* ('Calculating') *Cereala* ('Grain') *Saluta* ('Salutation')	2	82
Tap ('Goat')	1	82

Table A1. *Cont.*

Academic Items in Romanian (with English Translations)	Correct Canadian Responses (Max = 3)	Romanian. Number (Max = 22)/Percentage
Raton ('Raccoon') *Insula* ('Island') *Racheta* ('Racket') *Aplaudă* ('Applauding')	3	77
Dreptunghi ('Rectangle') *Parasuta* ('Parachute')	2	77
Chitara ('Guitar') *Tubular* ('Tubular')	1	77
Coarne de cerb ('Antlers')	2	73
Asambleaza ('Assembling')	1	73
Imbratiseaza ('Embracing') *Masina de camping* (RV) *Salutare* ('Greeting')	2	68
Morsa ('Walrus') *Surprins* ('Surprised') *Dichiseste* ('Grooming')	1	68
Soseste ('Descending')		68
Dezamagit ('Disappointed') *Cursa de garduri* ('Hurdling') *Locuinta* ('Dwelling') *Eliberat* ('Liberated') *Iluminare* ('Illumination')	2	64
Medic ('Physician') *Ciocan de judecator* ('Gavel')	1	64
Manz ('Colt') *Balustrada* ('Banister') *Fundatie* ('Foundation')		64
Arici ('Porcupine') *Semnal* ('Signal') *Recolteaza* ('Harvesting') *Tropical* ('Tropical')	2	59
Cascada ('Cascade')	1	59
Pedala ('Pedal')	3	55
Stanca ('Cliff') *Neregulat* ('Irregular')	2	55
Arctic ('Arctic') *Aviatie* ('Aviation')	1	55
Marsupiu ('Marsupial')		55
Fictive ('Fictional')	3	50
Tuba ('Tuba')	2	50
Scufunda ('Immersing') *Busola* ('Compass') *Tine discurs* ('Orating') *Armonica* ('Harmonica')	1	50
Mistrie ('Trowel')		50
Balama ('Hinge')	2	45
Clopotnita ('Steeple')	1	45

Table A1. *Cont.*

Academic Items in Romanian (with English Translations)	Correct Canadian Responses (Max = 3)	Romanian. Number (Max = 22)/Percentage
Tarm ('Shore') *Compune* ('Composing')		45
Tripleti ('Triplets')	3	41
Vesmant ('Attire') *Proiectil* ('Projectile')	2	41
Tamponeaza ('Collision') *Extenuate* ('Exhausted') *Ajustabil* ('Adjustable') *Canin* ('Canine') *Fungus* ('Fungus')	1	41
Palnie ('Funnel')		41
Sugar ('Infant') *Enumereaza* ('Enumeration')	1	36
Sevalet ('Easel') *Perplex* ('Perplexed')		36
Savant ('Scholar')	1	32
Exterior ('Exterior')	1	32
Imprima ('Printing') *Smirgaluieste* ('Sanding') *Ademenitor* ('Enticing') *Monetar* ('Monetary') *Agricultura* ('Agriculture') *Recipient* ('Receptacle')		32
Bazin ('Pelvis')	1	27
Maxilar ('Jaw') *Nutritive* ('Nutritious') *Amfibiu* ('Amphibian')		27
Oaza ('Oasis')	1	23
Culinar ('Culinary') *Invapaiat* ('Blazing') *Vitezometru* ('Speedometer') *Sferic* ('Spherical')		18
Duet ('Duet')	1	14
Casete ('Capsules') *Paralel* ('Parallel')		14
Pocal ('Goblet') *Scripete* ('Hoisting')	1	9
Teaca ('Pod') *Lubrifiaza* ('Lubricating')		9
Descinde ('Descending') *Concav* ('Concave')		5
Tir ('Archery')	2	0
Escorteaza ('Escorting')	1	0

Appendix D. Home Items from PPVT-4 (Romanian-English)

Table A2. Home Items from PPVT-4 (Romanian-English).

Home Items in Romanian (with English Translations)	Correct Canadian Responses (Max = 3)			Romanian. Number (Max = 22)/Percentage
Pisică ('Cat')	1	1	1	22/100%
Măr ('Apple')	1	1	1	22/100%
Balon ('Balloon')	1	1	1	22/100%
Mână ('Hand')	1	1	1	22/100%
Pasăre ('Bird')	1	1	1	22/100%
Pom ('Tree')	1	1	1	22/100%
Masă ('Table')	1	1	1	22/100%
Bea ('Drinking')	1	1	1	22/100%
Broască ('Frog')	1	1	1	22/100%
Bani ('Money')	1	1	1	22/100%
Umbrelă ('Umbrella')	1	1	1	22/100%
Fuge ('Running')	1	1	1	22/100%
Geam ('Window')	1	1	1	22/100%
Gât ('Neck')	1	1	1	22/100%
Vorbește ('Talking')	1	1	1	22/100%
Albastru ('Blue')	1	1	1	22/100%
Deget ('Finger')	1	1	1	22/100%
Struguri ('Grapes')	1	1	1	22/100%
Inoată ('Swimming')	1	1	1	22/100%
Poștă ('Mail')	1	1	1	22/100%
Ciocan ('Hammer')	1	1	1	22/100%
Lumânare ('Candle')	1	1	1	22/100%
Tristă ('Sad')	1	1	1	22/100%
Sare ('Salt')	1	1	1	22/100%
Plantă ('Plant')	1	1	1	22/100%
Cozonac ('Sweet bread')	1	1	1	22/100%
Joc ('Game')	1	1	1	22/100%
Inel ('Ring')	1	1	1	22/100%
Fermier ('Farmer')	1	1	1	22/100%
Fermoar ('Zipper')	1	1	1	22/100%
Pară ('Pear')	1	1	1	22/100%
Cască ('Yawning')	1	1	1	22/100%
Toarnă ('Pouring')	1	1	1	22/100%
Decorat ('Decorated')				22/100%
Vapor ('Ship')	1	1	1	22/100%
Nefericit ('Unhappy')	1	1	1	22/100%
Fruct ('Fruit')	1	1	1	22/100%
Genunchi ('Knee')	1	1	1	21/95%
Plin ('Full')	1	1	1	21/95%
Masoara ('Measuring')	1	1	1	21/95%
Imens ('Immense')	1			21/95%
Bijuterie ('Jewellery')	1	1	1	20/91%
Imparte ('Dividing')	1	1	1	20/91%
Constructie ('Construction')	1	1	1	18/82%
Aspru ('Rough')	1			18/82%
Picura ('Leaking')	1	1	1	17/77%
Casier ('Chasier')	1	1		16/73%
Jogging ('Jogging')	1	1	1	15/68%
Avocado ('Avocado')	1	1		15/68%
Livada ('Orchard')				15/68%
Rama ('Frame')	1			14/64%
Timp ('Time')	1			10/45%
Termos ('Thermos')	1			10/45%
Enervant ('Annoying')	1	1		9/41%

Appendix E. Romanian-English Cognates

Table A3. Romanian-English Cognates.

Cognate Items in Romanian (with English Translations)	Correct Canadian Responses (Max = 3)	Romanian. Number (Max = 22)/Percentage
Fruct ('Fruit') *Astronaut* ('Astronaut')	3	22/100%
Statuie ('Statue')	2	22/100%
Masoara ('Measuring') *Binoclu* ('Binoculars')	3	21/95%
Imens ('Imens')	1	21/95%
Pelican ('Pelican') *Tornada* ('Tornado') *Lichid* ('Liquid')	1	20/91%
Buchet ('Bouquet')	3	19/86%
Piramida ('Pyramid')	1	19/86%
Constructie ('Construction') *Electrician* ('Electrician')	1	18/82%
Calculeaza ('Calculating') *Cereal* ('Cereal')	2	18/82%
Aplauda ('Applaud') *Racheta* ('Racket')	3	17/77%
Chitara ('Guitar') *Tubular* ('Tubular')	1	17/77%
Casier ('Cashier')	2	16/73%
Asambleaza ('Assembling')	1	16/73%
Jogging ('Jogging')	3	15/68%
Avocado ('Avocado') *Salutare* ('Greeting')	2	15/68%
Surprins ('Surprised')	2	15/68%
Eliberat ('Liberated') *Iluminare* ('Illumination')	2	14/64%
Fundatie ('Foundation')		14/64%
Tropical ('Tropical')	2	13/59%
Cascada ('Cascade')	1	13/59%
Pedala ('Pedal')	3	12/55%
Neregulat ('Irregular')	2	12/55%
Arctic ('Arctic') *Aviatie* ('Aviation')	1	12/55%
Marsupiu ('Marsupial')		12/55%
Fictive ('Fictive')	3	11/50%
Tuba ('Tuba')	2	11/50%
Armonica ('Harmonica')	1	11/50%
Termos ('Thermos')	1	10/45%
Tripleti ('Triplets')	3	9/41%

Table A3. *Cont.*

Cognate Items in Romanian (with English Translations)	Correct Canadian Responses (Max = 3)	Romanian. Number (Max = 22)/Percentage
Proiectil ('Projectile')	2	9/41%
Extenuate ('Exhausted') *Ajustabil* ('Adjustable') *Canin* ('Canine') *Fungus* ('Fungus')	1	9/41%
Enumereaza ('Enumeration')	1	8/36%
Perplex ('Perplexed')		8/36%
Exterior ('Exterior')	1	7/32%
Agricultura ('Agriculture') *Recipient* ('Recipient')		7/32%
Nutritiv ('Nutritious') *Amfibiu* ('Amphibian')		6/27%
Culinary ('Culinary') *Sferic* ('Spherical')		4/18%
Paralel ('Parallel')		3/14%
Concave ('Concave')		1/5%

References

Ard, Josh, and Taco Homburg. 1983. Verification of language transfer. In *Language Transfer in Language Learning*. Amsterdam: John Benjamins, pp. 157–76.

August, Diane, Maria Carlo, Cheryl Dressler, and Catherine Snow. 2005. The critical role of vocabulary development for English Language Learners. *Learning Disabilities Research & Practice* 20: 50–57.

Avram, Larisa. 2001. Remarks on the optional clitic stage in child Romanian. *Working Papers in Linguistics* 3: 16–28.

Bantaş, Andrei. 1994. *Dicţionar Român-Englez*. Bucureşti: Teora.

Bates, Elizabeth, Virginia Marchman, Donna Thal, Larry Fenson, Philip Dale, J. Steven Reznick, Judy Reilly, and Jeff Hartung. 1994. Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language* 21: 85–123. [CrossRef] [PubMed]

Bialystok, Ellen, and Xiaojia Feng. 2009. Language proficiency and its implications for monolingual and bilingual children. In *Language and Literacy Development in Bilingual Settings*. Edited by Aydin Yücesan Durgunoglu and Claude Goldenberg. New York: The Guilford Press, pp. 121–40.

Bialystok, Ellen, Gigi Luk, and Ernest Kwan. 2005. Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading* 9: 43–61. [CrossRef]

Bialystok, Ellen, Gigi Luk, Kathleen F. Peets, and Sujin Yang. 2010. Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition* 13: 525–31. [CrossRef] [PubMed]

Buja, Elena. 2008. *Relating Events in Narrative. A Case Study of Romanian*. Braşov: Editura Universităţii "Transilvania".

Buzilă, Paul. 2016. Particularities of Romanian as acquired by young bilingual immigrants in Spain. *Bucharest Working Papers in Linguistics* 18: 69–82.

Carroll, Susanne E. 1992. On cognates. *Interlanguage Studies Bulletin (Utrecht)* 8: 93–119. [CrossRef]

Clark, Eve V. 1995. *The Lexicon in Acquisition*. Cambridge: Cambridge University Press, vol. 65.

Cojocaru, Dana. 2003. Romanian Grammar. Available online: http://users.ox.ac.uk/~uzdh0146/compgrammar_romanian.pdf (accessed on 6 December 2013).

Costa, Albert, Mikel Santesteban, and Agnès Caño. 2005. On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language* 94: 94–103. [CrossRef] [PubMed]

Cummins, Jim. 2005. A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *The Modern Language Journal* 89: 585–92.

Cutler, A. 1994. Segmentation problems, rhythmic solutions. *Lingua* 92: 81–104. [CrossRef]

De Groot, Annette, and Rineke Keijzer. 2000. What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning* 50: 1–56. [CrossRef]

De Houwer, Annick. 2009. *Bilingual First Language Acquisition*. Buffalo: Multilingual Matters.

Dijkstra, Ton, Koji Miwa, Bianca Brummelhuis, Maya Sappelli, and Harald Baayen. 2010. How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language* 62: 284–301. [CrossRef]

Dunn, Lloyd M., and Leota M. Dunn. 1981. *The Peabody Picture Vocabulary Test-Revised*. Circle Pines: American Guidance Service.

Dunn, Lloyd M., and Leota M. Dunn. 2007. *Peabody Picture Vocabulary Test-IV*. San Antonio: Pearson.

Dunn, Lloyd M., Delia E. Lugo, Eligio R. Padilla, and Leota M. Dunn. 1986. *Test de Vocabulario en Imágenes Peabody: Adaptación Hispano-Americano [Peabody Picture Vocabulary Test-Latin American Adaptation]*. Circle Pines: Guidance Service.

Finkbeiner, Matthew, Kenneth Forster, Janet Nicol, and Kumiko Nakamura. 2004. The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language* 51: 1–22. [CrossRef]

García, Georgia Earnest, and William E. Nagy. 1993. Latino students' concept of cognates. In *Examining Central Issues in Literacy Research, Theory, and Practice*. Edited by Donald J. Leu and Charles K. Kinzer. Chicago: National Reading Conference, pp. 367–73.

Gershkoff-Stowe, Lisa, Donna J. Thal, Linda B. Smith, and Laura L. Namy. 1997. Categorization and its developmental relation to early language. *Child Development* 68: 843–59. [CrossRef] [PubMed]

Gleitman, Lila. 1990. The structural sources of verb meanings. *Language Acquisition* 1: 3–55. [CrossRef]

Guasti, Maria Teresa. 2000. An excursion into interrogatives in Early English Italian. In *The Acquisition of Syntax*. Edited by Marc-Ariel Friedemann and Luigi Rizzi. London: Routledge, pp. 115–38.

Gyllstad, Henrik, Laura Vilkaitė, and Norbert Schmitt. 2015. Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL-International Journal of Applied Linguistics* 166: 278–306. [CrossRef]

Hall, Christopher J. 2002. The automatic cognate form assumption: Evidence for the parasitic model of vocabulary development. *IRAL—International Review of Applied Linguistics in Language Teaching* 40: 69–87. [CrossRef]

Hammer, Carol S., Frank R. Lawrence, and Adele W. Miccio. 2008. Exposure to English before and after entry into Head Start: Bilingual children's receptive language growth in Spanish and English. *International Journal of Bilingual Education and Bilingualism* 11: 30–56. [CrossRef] [PubMed]

Helms-Park, Rena, and Vedran Dronjic. 2016. Crosslinguistic lexical influence: cognate facilitation. In *Crosslinguistic Influence*. Edited by Rosa Alonso. Bristol: Multilingual Matters, pp. 71–92.

Jarvis, Scott. 2009. Lexical transfer. In *The Bilingual Mental Lexicon: Interdisciplinary Approaches*. Edited by Aneta Pavlenko. Bristol: Multilingual Matters, pp. 99–124.

Jiang, Nan. 2000. Lexical representation and development in a second language. *Applied Linguistics* 21: 47–77. [CrossRef]

Jusczyk, Peter W., Angela D. Friederici, Jeanine M. Wessels, Vigdis Y. Svenkerud, and Ann Marie Jusczyk. 1993. Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language* 32: 402–20. [CrossRef]

Kelley, Alaina, and Kathryn Kohnert. 2012. Is there a cognate advantage for typically developing Spanish-speaking English-language learners? *Language, Speech, and Hearing Services in Schools* 43: 191–204. [CrossRef]

Kohnert, Katrhyn, Jennifer Windsor, and Ruth Miller. 2004. Crossing borders: Recognition of Spanish words by English-speaking children with and without language impairment. *Applied Psycholinguistics* 25: 543–64. [CrossRef]

Kroll, Judith F., and Erika Stewart. 1994. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33: 149–74. [CrossRef]

Kroll, Judith F., Janet G. van Hell, Natascha Tokowicz, and David W. Green. 2010. The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition* 13: 373–81. [CrossRef] [PubMed]

Levițchi, L. D. 2005. *Dicționar Român Englez: 60.000 de Cuvinte*. București: Teora.

Sherkina-Lieber, Marina, and Rena Helms-Park. 2015. A prototype of a receptive lexical test for a polysynthetic ancestral language: The case of Inuttitut in Labrador. *Language Testing* 32: 419–42. [CrossRef]

Mahon, Merle, and Alison Crutchley. 2006. Performance of typically-developing school-age children with English as an additional language on the British Picture Vocabulary Scales II. *Child Language Teaching and Therapy* 22: 333–52. [CrossRef]

Malabonga, Valerie, Dorry Kenyon, Maria Carlo, Diane August, and Mohammed Louguit. 2008. Development of a cognate awareness measure for Spanish-speaking English language learners. *Language Testing* 4: 495–519. [CrossRef]

Maneca, Constant. 1996. Consideratii asupra frecventei cuvintelor în limba româna literara contemporana. *Studii si Cercetari Lingvistice* XVII: 623–33.

Mesaros, Bianca Tamara. 2008. Learning English as a Third Language: The Case of the Romanian Community in Spain. Master in English Language Teaching and Acquisition in Multilingual Contexts. Master's thesis, Universitat Jaume I, Castellón, Spain.

Midgley, Katherine J., Phillip J. Holcomb, and Jonathan Grainger. 2011. Effects of cognate status on word comprehension in second language learners: An ERP investigation. *Journal of Cognitive Neuroscience* 23: 1634–47. [CrossRef] [PubMed]

Montrul, Silvina, Rakesh Bhatt, and Roxana Girju. 2015. Differential object marking in Spanish, Hindi, and Romanian as heritage languages. *Language* 91: 564–610. [CrossRef]

Nagy, William E., Georgia Earnest García, Aydin Y. Durgunoğlu, and Barbara Hancin-Bhatt. 1993. Spanish–English bilingual students' use of cognates in English reading. *Journal of Reading Behavior* 25: 241–59. [CrossRef]

Nesteruk, Olena. 2010. Heritage language maintenance and loss among the children of Eastern European immigrants in the USA. *Journal of Multilingual and Multicultural Development* 31: 271–86. [CrossRef]

Nicoladis, Elena, and Fred Genesee. 1996. Word awareness in second language learners and bilingual children. *Language Awareness* 5: 80–90. [CrossRef]

Oller, D. Kimbrough, and Rebecca E. Eilers. 2002. Balancing interpretations regarding effects of bilingualism: Empirical outcomes and theoretical possibilities. In *Language and Literacy in Bilingual Children*. Edited by D. Kimbrough Oller and Rebecca E. Eilers. Clevedon: Multilingual Matters, pp. 281–92.

Oller, D. Kimbrough, Barbara Z. Pearson, and Alan B. Cobo-Lewis. 2007. Profile effects in early bilingual language and literacy. *Applied Psycholinguistics* 28: 191–230. [CrossRef] [PubMed]

Ouellette, Gene P. 2006. What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *The Journal of Educational Psychology* 98: 554–66. [CrossRef]

Pavlenko, Aneta. 2009. Conceptual representation in the bilingual lexicon and second language vocabulary learning. In *The Bilingual Mental Lexicon: Interdisciplinary Approaches*. Edited by Aneta Pavlenko. Bristol: Multilingual Matters, pp. 125–60.

Peets, Kathleen F., and Ellen Bialystok. 2015. Academic discourse: Dissociating standardized and conversational measures of language proficiency in bilingual kindergarteners. *Applied Psycholinguistics* 36: 437–61. [CrossRef] [PubMed]

Pérez, Anita, Elizabeth D. Peña, and Lisa M. Bedore. 2010. Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Childhood Services Journal* 4: 55–67.

Petrescu, Maria Claudia, Rena Helms-Park, and Vedran Dronjic. 2017. The Impact of Frequency and Register on Cognate Facilitation: Comparing Romanian and Vietnamese Speakers on the Vocabulary Levels Test. *English for Specific Purposes* 47: 15–25. [CrossRef]

Pinker, Steven. 1984. *Language Learnability and Language Learning*. Boston: Harvard University Press.

Rohde, Treena Eileen, and Lee Anne Thompson. 2007. Predicting academic achievement with cognitive ability. *Intelligence* 35: 83–92. [CrossRef]

Rojas, Raúl, and Aquiles Iglesias. 2013. The language growth of English language learners: Change over time, individual differences, and the impact of initial status on growth. *Journal of Child Development* 84: 630–46. [CrossRef] [PubMed]

Sabourin, Laura, Jean-Christophe Leclerc, Michèle Burkholder, and Christie Brien. 2014. Bilingual lexical organization: Is there a sensitive period? In *2014 CLA Conference Proceedings, Annual Conference of the Canadian Linguistic Association, Brock University, St. Catharines, ON, Canada, 24–26 May 2014*. Edited by Laura Teddiman. Ottawa: Canadian Linguistic Association.

Sánchez-Casas, Rosa, and José E. García-Albea. 2005. The representation of cognate and noncognate words in bilingual memory: Can cognate status be characterized as a special kind of morphological relation? In *Handbook of Bilingualism: Psycholinguistic Approaches*. Edited by Judith F. Kroll and Annette M. B. de Groot. Oxford: Oxford University Press, pp. 226–50.

Snow, Catherine E. 1977. The development of conversation between mothers and babies. *Journal of Child Language* 4: 1–22. [CrossRef]

Statistics Canada. 2011. 2011 Census: Language. Available online: https://www.toronto.ca/wp-content/uploads/2017/10/8fab-2011-Census-Backgrounder-Language.pdf (accessed on 15 August 2018).

Swanson, H. Lee, Kelly Rosston, Michael Gerber, and Emily Solari. 2008. Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology* 46: 413–29. [CrossRef] [PubMed]

Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Approach to Child Language Acquisition*. Cambridge: Harvard University Press.

Umbel, Vivian M., and D. Kimbrough Oller. 1994. Developmental changes in receptive vocabulary in Hispanic bilingual school children. *Language Learning* 44: 221–42. [CrossRef]

Umbel, Vivian M., Barbara Z. Pearson, María C. Fernández, and D. Kimbrough Oller. 1992. Measuring bilingual children's receptive vocabularies. *Child Development* 63: 1012–20. [CrossRef] [PubMed]

Verhallen, Marianne, and Rob Schoonen. 1993. Vocabulary knowledge of monolingual and bilingual children. *Applied Linguistics* 14: 344–63. [CrossRef]

Vermeer, Anne. 2001. Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics* 22: 217–34. [CrossRef]

Werker, Janet F. 2012. Perceptual foundations of bilingual acquisition in Infancy. *Annals of the New York Academy of Sciences: The Year in Cognitive Neuroscience* 1251: 50–61. [CrossRef] [PubMed]

Werker, Janet F., and Richard C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7: 49–63. [CrossRef]

Wolter, Brent. 2006. Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics* 27: 741–47. [CrossRef]

languages

MDPI

Article

French Postverbal Subjects: A Comparison of Monolingual, Bilingual, Trilingual, and Multilingual French

Laia Arnaus Gil * and Natascha Müller *

Fakultät 1, Romanische Sprachwissenschaft, Bergische Universität Wuppertal, Gaußstraße 20, 42119 Wuppertal, Germany

* Correspondence: arnaus@uni-wuppertal.de (L.A.G.); nmueller@uni-wuppertal.de (N.M.); Tel.: +49-202-439-2157 (L.A.G.); +49-202-439-2160 (N.M.)

Received: 13 November 2017; Accepted: 24 July 2018; Published: 28 July 2018

Abstract: Monolingual French children have been observed to pass through a stage which is characterized by the production of target-deviant postverbal subjects of the following type (Jansen 2015, p. 272): *est tombé Philippe* 'is fallen Philippe' (Philippe, 2;2,10), *écrit bien celui-là* 'writes well this one' (Philippe, 2;2,2) (De Cat 2002; Déprez and Pierce 1993; Ferdinand 1993, 1996; Friedemann [1993] 1994; Labelle and Valois 1996; Pierce 1989). Interestingly, bilingual children who acquire French together with German produce postverbal subjects to an extremely low degree in mean length of utterance (MLU)-phases compared with monolingual children (Jansen 2015). Arguably, they skip the postverbal subject phase and are accelerated with respect to monolingual children. In our study, we tested whether multilingualism can speed up the acquisition process in French. A production test with 62 multilingual children (starting at 2;7 acquiring two, three, or four L1s) was administered in Spain and Germany to elicit finite verbs and DP (Determiner Phrase) subjects in French. The children's proficiency in French was measured on the basis of the Peabody Picture Vocabulary Test (PPVT) (Dunn et al. 1993). In comparison with monolinguals studied in the literature, the bilingual children and the children who acquired more than two languages were accelerated with respect to the placement of subjects in the postverbal position. Although the multilingual children who ranked low in the PPVT exhibited all kinds of structures as responses to the test items that are characteristic of early French, such as null-subjects, root infinites and bare nouns, they did not use postverbal subjects. The absence of postverbal subjects was observed not only in children who acquired either German (or English) at the same time as French, as did the bilingual children in Jansen's (2015) study, but also in children who acquired, in addition to French, a Romance language like Spanish, Catalan, or Italian, all null-subject languages which allow postverbal subjects in the adult language. Following (Biberauer and Richards 2006), the extended projection principle (EPP) feature of T (Tense) can be satisfied in different ways across languages: if a DP is necessary, which it is in adult French, it can be raised from Spec,*v*P (specifier of *v*P, of the small-verb Phrase) to Spec,TP (specifier of Tense Phrase) (in which case the finite verb surfaces in T). This is the default of EPP satisfaction in T. It is this option which is facilitated and boosts the acquisition of the preverbal subject position in early bilingual, trilingual, and multilingual French. The result that the children who acquired a null-subject language such as Spanish also enjoyed an advantage in French, adult Spanish being a language that allows for postverbal subjects, indicates that it is plausible that the default character of DP-raising for EPP satisfaction, instantiated in all languages involved, is the reason for its high relevance in the early French of bilingual, trilingual, and multilingual children. If our analysis of the multilingual data is plausible, monolingual French children exhibit more variation in satisfying the EPP-feature (in T) than children who acquire one or more other languages in addition to French, an observation which goes back to (Meisel 1989).

Keywords: French; postverbal subjects; EPP satisfaction; cross-linguistic influence

1. Introduction

Bilingual children who grow up with two languages from birth can separate them (cf. Genesee 1989; Lleó 2002; Lleó and Kehoe 2002; Meisel 1989), but the acquisition path is characterized by cross-linguistic influence. This influence characterizes bilingual phonology (Lleó 2016; Lleó and Rakow 2006), morphology (Nicoladis 1999, 2003), and syntax (Döpke 1992, 1997, 1998; Gawlitzek-Maiwald and Tracy 1996; Hulk and Müller 2000; Müller and Hulk 2001). Paradis and Genesee (1996) argue that it has two effects: delay and acceleration. Compared to delay effects, acceleration effects are understudied in early child bilingualism. They exist in bilingual German/English–Romance children (Kupisch 2006 for determiner realizations in German; Fernández Fuertes and Liceras 2010; Liceras et al. 2012 for copula realizations in English and for overt realization of subject pronouns in English; cf. Liceras and Fuertes 2018). Recently, Patuto et al. (2011) have considered whether these two effects, as has generally been assumed in the literature, are both rooted in cross-linguistic influence. Müller (2017) assumes that acceleration has a different source and is reminiscent of the bilingual advantage which is due to enhanced cognitive abilities.

This article will investigate whether acceleration effects are also present in children acquiring three or more languages from birth. The grammatical domain chosen, postverbal subjects in French, is accelerated in bilingual children (Jansen 2015). Since "trilingual children have to divide their exposure time across three languages and thus may have even more reduced input than many bilingual children, at least in one of their three languages" (Unsworth 2013, p. 40), the question we ask is whether, given a decreased amount of input in each language, the development of the grammatical domain is delayed as compared to bilingual as well as to monolingual French children. As a counterpoint to decreased input quantity, the trilingual child experiences increased input diversity. Thus, "infants exposed to two or more languages from birth are innately prepared to phonetically process and learn regularities from multiple native languages, and multilingualism does not delay nor compromise the development of linguistic representation" (Montanari 2010a, p. 135). This might suggest that the exposure to two or more languages from birth could lead multilingual children to experience acceleration effects.

The article is structured as follows: Section 2 gives an overview of existing studies on early child trilingualism. Section 3 describes the phenomenon observed in monolingual and bilingual French children, namely postverbal subjects. Section 4 is devoted to the present study. Elicited production was administered to 18 bilingual, 38 trilingual and 6 multilingual children who spoke French as one of their native languages. Section 5 discusses the results and concludes the article.

2. Issues in Early Child Multilingualism

The study of early child trilingualism is still in its infancy (cf. Unsworth 2013, p. 41ff), a statement which goes back to Quay (2001, p. 149), Hoffmann (2001, p. 13), Barnes (2006, p. 28) and is still valid in 2017 (but see the works of Arnaus Gil 2013 and Hager 2014 on early trilingualism). One reason might be that "for the time being, most of us are still working within the theoretical framework of bilingualism" (Hoffmann 1999, p. 16), at least in first language acquisition. Most studies on early child trilingualism underlie the view that it is "another type of bilingualism, and theories and findings from studies of bilinguals are often assumed to be applicable to trilinguals by extension" (Barron-Hauwaert 2000, p. 1; cf. Jessner 1997); however, compare the criticism in Hoffmann (2000). While similarities do exist between early child bilingualism and the acquisition of three languages from birth, "there is still a lack of data comparing bilinguals and trilinguals" (Barron-Hauwaert 2000, p. 2). Hoffmann (2001) also reports differences (cf. also Quay 2011a, 2011b) between early bi- and trilingualism. These observations should invite acquisitionists to consider the necessity of a theory of trilingual first language acquisition, growing apart from bilingual first language theories. The differences between bilinguals and trilinguals are quantitative and qualitative in nature (Hoffmann 2001, p. 1). For example, and in contrast to bilinguals, trilinguals are never assumed to be balanced; there will always be one language which is dominant, while the other two are "weaker" (cf. Montanari 2013, p. 63 for a summary of studies that defend this view). In fact, Montanari (2013, p. 63) argues that the quality of the input (in terms of the

variety of contacts in the minority language(s) in order to have not only one main language provider, as noted by Barnes (2011)) and language-internal properties play a role in how well each language is mastered by the trilingual child. This brings Quay (2011a, p. 3) to conclude that "trilingual children need to be considered as speakers in their own right". Montanari (2013, p. 63) further considers another reason for the rarity of studies on early child trilingualism, namely the time-consuming and expensive methods that must be used to document language development in trilingual children.

In contrast to young children, speakers who have acquired a third language in adulthood are the subject of several research articles dealing with successive multilingualism under the topic of third language acquisition. Yet, as one reviewer correctly notes, adult L3 acquisition studies are in fact quite limited and restricted to theoretical models that have been proposed in the last 15 years and deal with the initial stages of L3 acquisition. More importantly, third language acquisitionists have developed a sensitivity towards the speaker's chronological acquisition of foreign languages (the first one is called second language, the successive languages L3/Ln). In addition, it matters whether the language learner is bilingual from birth when s/he starts to learn a foreign language, compare the important contributions by Cenoz and Hoffmann (2003), Cenoz and Jessner (2000, 2001a, 2001b, 2001c), and Herdina and Jessner (2002). Simultaneous bilinguals can obtain higher levels of proficiency in the third language than monolinguals who acquired a second and a third language successively. The theories of third language acquisition differ with respect to the role played by the L1 and L2 in L3 learning. The Cumulative Enhancement Model (Flynn et al. 2004) assumes that all previously acquired languages will have a cumulative and positive effect on the learner's L3, making a distinction between first, second, and third language acquisition irrelevant. This assumes that the languages have an additive, positive effect on language learning.[1] Bardel and Falk (2007), on the other hand, argue in favor of the prominent role of the L2 if learned in a setting comparable to that of the L3 (tutored) within the L2 Status Factor Model. In this case, transfer from the L1 is blocked in L3 acquisition. For Rothman (2011, 2015) and his Typological Proximity Model, the language which is (psycho-)typologically similar to the L3 will have a (negative or positive) effect. To sum up, there is currently a lively debate which allows the identification of specific characteristics that distinguish third language from second language acquisition and multilingualism from early bilingualism. This debate is absent when the three first languages are simultaneously acquired (early child trilingualism). It would therefore be interesting and more than relevant to investigate the possibility of an additive effect of the three L1s in the child or, for example, whether the typology of the first and/or second L1 plays an important role in any acceleration in the third L1.

Whereas the research on third language acquisition uses a variety of empirical research methods, early child trilingualism is mainly characterized by longitudinal studies of trilingual children and questionnaires filled in by their parents or so-called parental interviews carried out by linguistic researchers. Quay (2001, p. 157) highlights methodological flaws in studies on early child trilingualism which are still present today: most studies provide minimal, if any, information about their data collection methods, and none of the studies deals systematically with early trilingual development. This is why "we should proceed with caution in evaluating the results from these studies" (Quay 2001, p. 157). The main questions which lie at the heart of early child trilingualism are whether active trilingualism, i.e., the active competence of all three first languages (cf. Chevalier 2015) is possible and what factors increase the likelihood of an active mastery of three languages. In what follows, we will summarize the main studies on trilingual first language acquisition as an illustration of the prevalence of anecdotal descriptions of trilingual language use. These descriptions function mainly as testimonies.

---

[1] As one reviewer correctly remarks, the Cumulative Enhancement Model consequently assumes that non-facilitative transfer should not be observed in the acquisitional L3 data. In order words, negative transfer in L3 acquisition should be absent or explained differently. In fact, many studies on L3 acquisition have presented counter-evidence to this model (e.g., Bardel and Falk 2007 study of the acquisition of negation in L3 Swedish and Dutch). Accordingly, positive transfer is an observable effect of the functioning of the language acquisition device.

As was the case in the first studies on early child bilingualism (cf. Leopold 1949a, 1949b; Ronjat 1913), there are parental reports (or self-reports)[2] on language acquisition in early child trilingualism. Some examples are a linguistic diary of the utterances of the researcher's daughter (Dewaele 2000) and Wang (2008) anecdotal report, which is addressed to the parents of trilingual children. These studies show that the simultaneous acquisition of three languages is possible with normal cognitive and linguistic development (Gadler and Mikeš 1986; Mikès 1991; Murrell 1966). Studies of trilingual children often analyze language mixing (Chevalier 2015; Davidiak 2010; Edwards and Dewaele 2007; Hoffmann and Stavans 2007; Ivir-Ashworth 2011; Oksaar 1977; Stavans 1992, 2001; Stavans and Swisher 2006; Widdicombe 1997 as cited by Hoffmann (1999, p. 19ff.; compare also Hoffmann and Widdicombe 1999) and most often language use according to interlocutors (Barnes 2006, 2011; Chevalier 2015, Faingold 1999, 2000). Several factors influence active trilingualism, such as the absence of the community language at home (Braun and Cline 2010, 2014; De Houwer 2004), the mastery of both minority languages by both parents (De Houwer 2004), and the variety of contacts to whom the child is exposed. Chevalier (2015, p. 137) lists eight influencing factors: (1) consistency in following the one-person–one-language strategy; (2) amount of input; (3) language constellations; (4) variety of contacts; (5) variety of media; (6) parental discourse styles; (7) parent is linguist-investigator; and (8) status of the languages involved (Helot 1988; Barron-Hauwaert 2000). In Quay (2008), we learn about the language choices of a trilingual child at 17 dinner sessions from 1;10 until 2;4, i.e., in triadic interaction. Xiaoxiao was raised in Tokyo, Japan, by a Mandarin-speaking mother and an English-speaking father. Quay (2008, p. 30) notes that "she heard English from only 20% of the input she received in her daily environment", which was enough to mainly use English with her father and not mix more in her weakest language, English. The child's dominant language was Japanese, the language of daycare eight hours each weekday. Xiaoxiao's father spoke English in 90% of his utterances to his daughter and in 85% of utterances directed to his wife, while Xiaoxiao's mother spoke Chinese in 90% of her utterances to her daughter and 80% of English to her husband. An analysis of Xiaoxiao's mixed utterances "will be reserved for a future paper" (Quay 2008, p. 24), but in general, nouns were mixed (not noun phrases) and content rather than function words figured in mixed utterances.

Rare are the studies which report on trilingual children's development of grammar and vocabulary. Hoffmann (1985) describes her two trilingual German–Spanish–English-speaking children, Cristina (until 8;5) and Pascual (until 5;6). The children heard German from their mother and Spanish from their father, and the family lived in the United Kingdom. They were regularly exposed to English when play-school started at 2;9 (Cristina) and 3;1 (Pascual). Pascual heard the three languages from birth and therefore is an infant trilingual, while Cristina is categorized as an infant bilingual (German, Spanish) and a trilingual (plus English) child (Hoffmann 1985, p. 480). Hoffmann (1985, p. 484) reports divergence from adult German by Cristina, concerning gender, case marking, the use of certain prepositions and conjunctions and word order in German subordinate clauses that exhibit finite verbs in clause-final position, but which show root word order, as in (1) and (2):

1. 3;7
   das war ich wenn ich war zwei jahre
   this was I when I was two years
   'This was me when I was two years old.'

[2] Elwert [1959] (1973) is a self-report ("Erlebnisbericht" as he writes on p. 274). He describes his own trilingual upbringing first in Italy, then in Germany, by his English-speaking mother and his father who had friends speaking German and Italian to him. It is perhaps due to the fact that the community language was present in the home that Elwert [1959] (1973, p. 284) reports that he could not understand the German-speaking doctor of the family when he addressed Elwert at the age of three. German became one of Elwert's languages when the family moved to Germany. This was the case when Elwert was eight years old.

2. 3;8
   weißt du was kann ich mal machen?
   know you what can I PART do?
   'Do you know what I can do?'

There are no quantitative analyses, however. Chevalier (2015, p. 126ff.) also finds evidence of cross-linguistic influence in the two German–French–English trilingual children she studied. Examples come from adjective placement in German (only prenominal) and French (both pre- and postnominal positions are allowed): *vert couteau* 'green knife' (Lina, 3;1,8), *vert chofette* 'green fork' (=fourchette, Lina, 3;1,8), *aff rot* 'monkey red' (=Affe, Elliot, 2;5,25), *tiger rot* 'tiger red' (Elliot, 2;5,25). The two French examples show a prenominal placement of the color adjective *vert* 'green', which predominantly prefers the postnominal position in the adult language. In adult German and English, however, only the prenominal position of attributive adjectives is possible. The next two German examples from Elliot show a postnominal position which might be due to influence from French. However, there is no quantification at all.

Maneva (2004) describes the language acquisition of her daughter Daria from birth until age 5. Daria grew up in Montreal, where she began French day care at 1;10. Her mother spoke Bulgarian to her and her father Lebanese Arabic. Daria was also in contact with English through visitors, on the radio, on television, etc. At around the age of three, the main language of communication between Daria and her father became French. From the age of 3;11 to 4;6, Daria attended a bilingual French–English preschool. Maneva (2004, p. 112) mentions that Daria used her first two-word utterances which were structured "verb+subject" in Arabic and Bulgarian and "subject+verb" in French as early as 1;7. These results show that Daria was able to produce adult-like subject–verb sequences in all three languages she was in contact with, showing syntactic differentiation.

Mikès (1991) studied lexical differentiation in multilingual children in the town of Novi Sad, where Serbo–Croatian is the dominant language. Three children in her study, Vuk, Uva (siblings) and Egon, (aged 0;10 and 1;11) grew up with Hungarian (the mother's language), Serbo-Croatian (the father's language), and German (the grandmother's language). The children had doublets and triplets in two or three languages (p. 112–13), i.e., they knew the word for a concept in at least two (doublet) or in three (triplet) of their three L1s.

Montanari (2009, 2010a, 2010b) investigated the language choice of a trilingual child named Kathryn who grew up in Los Angeles with her Tagalog-speaking mother and her Spanish-speaking father through diary records and audio recordings from 1;4 until 2;5. English daycare started at age 2;2 (eight hours a day, three days a week), but Kathryn also heard English from her parents and her sister, who was nine years older than her. Both parents were also native speakers of English.[3] Kathryn differentiated her three lexicons from early on and used 44 doublets and 10 triplets at the end of the study in the appropriate language context. Concretely, the author found that forming a triplet took half the time of learning a doublet, suggesting that adding the word for a concept in the third L1 takes less time than acquiring an equivalent word in the second L1, showing learning facilitation by the same conceptual knowledge (Montanari 2010a, p. 39). Moreover, Kathryn was able to differentiate three phonological systems (cf. Montanari 2011), and word orders (cf. Montanari 2009) in her 438 multi-word utterances: Argument–predicate sequences were differentially ordered depending on the language. Mixed utterances were mainly caused by vocabulary gaps (Montanari 2010b). She concludes that "it is not excluded that increased input diversity leads the child to an enhanced language discrimination capacity and, in turn, to a heightened attention towards

[3] Kathryn's mother is Filipino-American, her father Chilean-American. The child's mother came to the United States from the Philippines at age nine; her father moved to Los Angeles from Chile at age twelve. "From birth, Kathryn was addressed primarily in Tagalog by her mother and maternal grandparents, in Spanish by her father and paternal grandmother, and she heard English from her sister, nine years older than herself, and, more indirectly, from family conversations (English was the main medium of communication in the home)" (Montanari 2009, p. 507).

the structural properties of her languages, thereby enhancing the course and timing of differentiation" (Montanari 2010a, p. 92; cf. also Montanari 2013). In other words, early trilingualism can enhance language acquisition (i.e., acceleration).

Quay (2001) looked at the trilingual language development of a boy named Freddy from 0;11 to 1;10. Freddy grew up in Tokyo with his English-speaking mother, his German-speaking father, and Japanese via daycare. Daycare started at 11 months, but this "late contact" with Japanese played no role in the child's language choice, Japanese being the preferred language from the onset of speech. For her study, Quay (2001) made use of diaries together with the MacArthur Communicative Development Inventory (MCDI, a parental report of the child's understanding and production of the L1s) and video recordings every week. The MCDI was administered every three to four weeks. The main result of this study was that exposure to a third language at the age of 11 months (input delay) does not delay development of that language. "This calls into question distinctions between simultaneous and successive acquisition based on differences in age of exposure", specifically the time between birth and 11 months of age (Quay 2001, p. 195).

Kazzazi (2011) studied cross-linguistic influence in her two children, Anusheh (from the age of 1;0 in the form of notes, audiotapes, and videotapes until 4;9) and Irman (for the latter, only written notes are available). The children were raised in Germany by an English-speaking mother and a Farsi-speaking father. In search of cross-linguistic influence, Kazzazi (2011, p. 65) differentiates between majority and minority influence. Majority influence is observed when two languages share a feature which the third lacks. This feature will then be transferred to the third language. Minority influence takes place "where a feature in only one of the three languages is transferred to the other two languages" (ibid., p. 65). Interestingly, whereas German and English are mainly pre-modifying in attribution structures, Farsi is post-modifying. In German and English, determiners, adjectives, possessives, etc. precede the noun in the noun phrase (NP), while in Farsi, only demonstratives and numerals precede the noun (ibid., p. 67). In Farsi, modifying elements such as attributive adjectives, possessive pronouns, and modifying nouns are linked to the modified element by the so-called *ezafe* particle (which is often analyzed either as an affix or as a clitic). This particle acts as a default element that links the modifying element to the modified noun. The following examples are taken from Kazzazi (2011, p. 67). They outline that Farsi and English/German are mirror-images of each other for the relevant constructions, namely for attributive adjectives, possessive pronouns, and modifying nouns.

3. a.	English		
	red apple	my apple	apple-tree
b.	German		
	roter apfel	mein apfel	apfelbaum
c.	Farsi		
	sib-e qermez	sib-e man	derakht-e sib
	apple-ezafe red	apple-ezafe my	tree-ezafe apple
	red apple	my apple	apple tree

These examples illustrate that, if cross-linguistic influence is to be observed by the two trilingual children, it should be majority influence, i.e., if two languages share a linguistic feature not found in the third language (ibid., p. 65). Since both German and English function similarly in terms of attributive constructions as opposed to Farsi, the two Germanic languages would then influence Farsi along these constructions. In other words, prenominal attribution should be found in Farsi. This is what actually happened: at the age of 3;9,17, Anusheh is reported to use prenominal adjectives in Farsi. At the same time, however, she uses English and German compounds that actually are right-headed like the German *Lagerfeuer* 'bonfire' and the English *swimming-bath* with the head on the left (*feuerlage* 'bonfire'and *bath-swimming* 'swimming bath' at 3;3,14 and 3;3,26, respectively; cf. Kazzazi 2011, p. 70).

Thus, Anusheh seems to exhibit majority influence in the case of adjective placement, but minority influence in the case of compounds.

To summarize: the studies mentioned clearly demonstrate that it is possible for a trilingual child to reach a native competence level in all three first languages. There are several factors that influence the active production of three languages, yet 20% of parental input in one of the three languages seems to be enough for the child to become competent in a language. The studies contain evidence in favor of language separation from early on. Cross-linguistic influence is not systematically documented, so one might infer that it is rare in trilingual acquisition. When it does occur, it is not necessarily the dominant language that influences the weak language(s). The effect of trilingualism in the three acquiring languages of the child as opposed to their acquisition in a monolingual and/or bilingual context is completely unstudied. Such a comparison between monolingual, bilingual and trilingual acquisition is needed if the study of a particular grammatical phenomenon is the focus. Our study will contribute to fulfilling this need. We will investigate the phenomenon of French postverbal subjects in early monolingual, bilingual, and trilingual acquisition. The following section will first present the acquisition of French postverbal subjects by monolingual and bilingual children.

3. Postverbal Subjects in Child French

Modern French is usually considered a subject–verb–object (SVO) language. Lambrecht (1986) and many others have observed that spoken French is characterized by the dislocation of all kinds of constituents. Constituents can be dislocated to the left, as in (4a) the subject Determiner Phrase (DP), or to the right, as in (4b).

4. a. Ce livre, il est très intéressant.
   This book, he is very interesting
   'This book, it is very interesting.'

   b. Il est très intéressant, ce livre.
   He is very interesting, this book
   'It is very interesting, this book.'

Lambrecht (1986) uses the terms "topic sentence" and "antitopic sentence" for (4a) and (4b), respectively, in which a dislocated element is doubled by a resumptive pronoun in the sentence, expressing person (third person), number (singular), gender (masculine) and case (nominative, cf. 4a,b). Taking into consideration Lambrecht (1986) distinction, Ashby (1982, p. 39), in his study of French word order, found 1355 sentence tokens whose referent is given in the discourse in a corpus of 25 speakers from the age strata of 14–21 years and 51–64 years in and around Tours, recorded in the home or workplace (free conversation). In 83% of the tokens, the noun phrase that was coreferential with the subject clitic appeared to the left of the finite verb (left-dislocation), while 17% were of the type in (4b) (right dislocation). Notley et al. (2007) confirmed these results. This means that there is a preference for left-dislocations like in (4a) in spoken French. The frequencies, however, vary according to the corpora analyzed. Some authors, like De Cat (2002), observe about 30% of dislocations in spoken French. Jansen (2015, p. 151) shows that this is also true for child-directed speech, (but cf. Notley et al. 2007, who show the opposite, and van der Linden and Sleeman 2007 for early French).

Following the literature described above, we could assume that we will find left dislocations much more frequently than right dislocations in (French monolingual) child speech. In a seminal article by Déprez and Pierce (1993, p. 42), the authors investigate the productions of monolingual French children until the age of 2;2. They produce constructions like the following in (5), in which a lexical subject occurs in a VOS or a VSO construction, with a finite verb and without a resumptive pronoun that doubles the subject, or with a non-finite verb. During this period, the subject can thus occur in the postverbal position even though no resumptive pronoun that doubles the subject is observed. Notice that transitive as well as intransitive verbs are found in constructions with postverbal subjects.

5. a. Lit maman (Nathalie, 2;0,1)
reads mama
'Mama reads.'

b. Pleure clown (Daniel, 1;8,3)
cries clown
'The clown cries.'

c. Assis la poupée (Nathalie, 2;2,1)
sat-PART the doll
'The doll has sat down.'

d. Dormir là Michel (Philippe, 2;2,1)
to-sleep there Michel
'Michel sleeps there.'

e. Veut encore Adrian du pain (Grégoire, 2;1,3)
wants still Adrian of-the bread
'Adrian still wants some bread.'

f. Pousses toi sandales (Daniel, 1;8,3)
push you sandals
'You push your sandals.'

At the same time, the children have target-like utterances which contain a preverbal subject clitic in a finite sentence, like in (6a,b), a preverbal lexical subject, as in (6c), and a right-dislocated lexical object, like in (6d) (cf. Déprez and Pierce 1993, p. 45 and De Cat 2002, who show that children can use dislocations in a pragmatically proper way).[4]

6. a. Elle dort (Daniel, 1;8,1)
She sleeps
'She sleeps.'

b. Il veut un bruit (Daniel, 1;11,1)
He wants a noise
'He wants a noise.'

c. Train va tomber (Nathalie, 2;2,2)
Train falls
'The train falls.'

d. Elle la voit l'auto (Nathalie, 2;2,2)
She it-Cl wants the car
She wants it, the car.'

Déprez and Pierce (1993, p. 42) report a relatively high frequency of constructions like the ones in (5), ranging between 65% and 85%, in the four children they analyzed from the age of 1;8 onwards.

[4] One reviewer points out that subject clitics are categorically absent in the children's structures with postverbal subjects. Indeed, Déprez and Pierce (1993, p. 44) noted that pronominal subjects are absent in the children's untensed clauses and that subject clitics do not appear in the postverbal subject position. The authors explain these results by assuming that subject clitics are affixes in child grammar which are generated in Infl (Inflection) and lexically realized only when bound to a raised verb. Subjects clitics are therefore not equivalent to full NP (Noun Phrase)/DP subjects since they do not parallel the distribution of full NPs/DPs.

In her analysis of the longitudinal data of three monolingual French children from CHILDES (1;4–4;10) Jansen (2015, chapter 7.1), confining herself to finite clauses, finds evidence of postverbal subjects of the type in (5) in two additional monolingual French children and adds that the postverbal-subject stage ends at the age of 2;7.

Jansen (2015, chp. 7.1) also analyzes the data from five German–French bilingual children in French. Unlike the previous literature, she compares the amount of postverbal subjects on the basis of mean length of utterance (MLU) stages, not on the basis of age. During the whole period of investigation, which comprises 11,087 finite clauses, the monolingual children produced 213 preverbal (1.92%) and 130 postverbal (1.17%) subjects that were not doubled by a resumptive pronoun. There were 703 empty subjects (6.34%). The bilingual children produced 17,969 finite clauses and 177 (0.98%) preverbal, but only 38 (0.21%) postverbal subjects without a doubled resumptive pronoun. There were 603 empty subjects (3.35%). Postverbal subjects of the type in (5) are ungrammatical in adult French, and thus their absence in the bilingual data indicates the child's convergence to the adult norm.

Jansen (2015) concludes that the absence of postverbal subjects as in (5) in the bilingual data is characteristic of bilingual French (the amount always stays below 5%), while the monolingual children Madeleine and Philippe evidenced (different) MLU phases during which the amount of postverbal subjects was well above 5%, see Figure 1.[5]

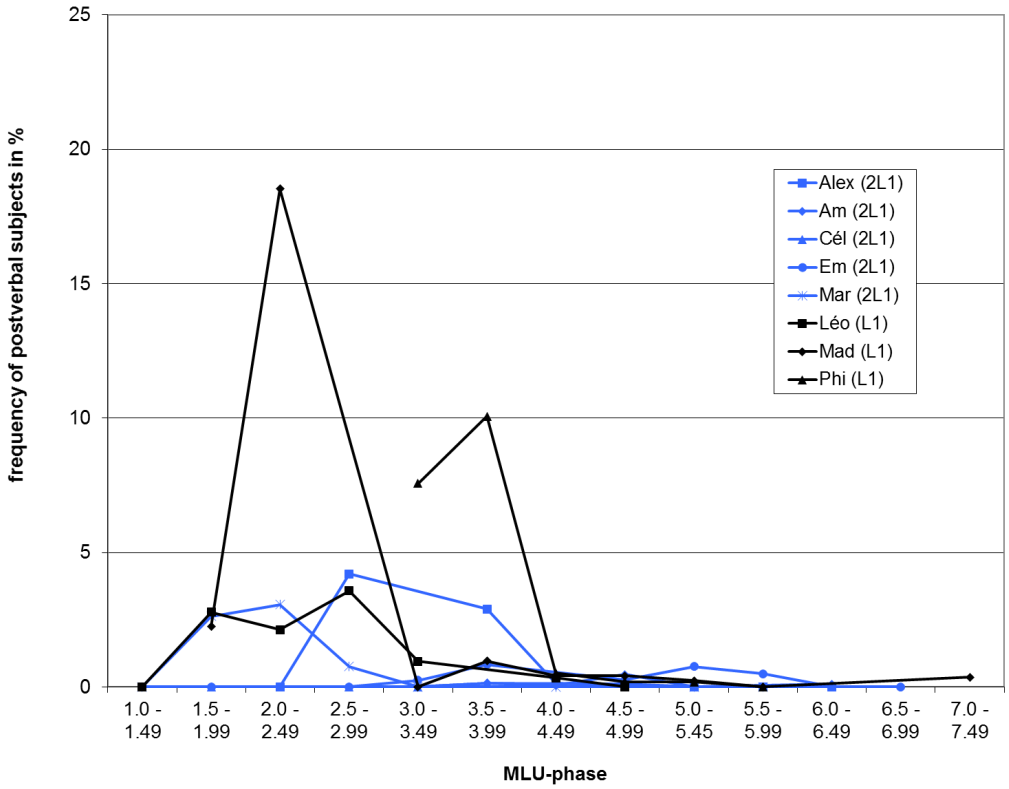


Figure 1. French postverbal subjects according to mean length of utterance (MLU) phase: Monolingual and bilingual children compared (Jansen 2015, p. 255).

[5] The monolingual child Léonard differed from his peers: He never reached the 5%-mark for postverbal subjects. Notice, however, that his corpus was small in comparison to that of the other two monolingual children: During the whole period of investigation (1;8,9–3;2,25), he used only 1251 finite clauses (Jansen 2015, p. 137). Philippe uttered approximately four times as many finite clauses during approximately the same period. The difference in the amount of production across children might be seen as part of the children's idiosyncrasies. Notice, however, that it is less likely that a particular target-deviant form appears in a small corpus.

In order to explain the results observed in monolingual French, Déprez and Pierce (1993) and Friedemann [1993] (1994) proposed the Subject-Raising Failure Approach and argued that postverbal subjects in child French are reminiscent of the child's capacity to raise finite verbs into Infl/Tense (T) in syntax. We have seen that, throughout the whole period of investigation (Jansen 2015), French bilinguals produced less than 1% of postverbal subjects, and in specific MLU stages where they appeared, they were never greater than 5%. We can thus claim that the postverbal subject stage was absent in the French of the bilingual children. It could be argued that they are accelerated with respect to the knowledge of the ungrammaticality of postverbal subjects which are not doubled by a resumptive clitic. This acceleration effect could be due to the children's second L1, German, which overwhelmingly exhibits subjects in the preverbal position. Notice, however, that the similarity of the two languages must be perused at a deep level of linguistic analysis, since the two adult systems radically differ at the surface. Jansen (2015, p. 154) shows that "dislocations" are very infrequent in spoken adult German (child-directed speech): they amount to 1%. Furthermore, she argues that "dislocations" to the right are not syntactic dislocations in German, but repairs, in which the speaker uses a lexical subject as a kind of afterthought. In other words, although both languages exhibit preverbal subjects, it is certain that they involve different syntactic machinery for handling the clear differences of their syntax.

Contrary to the Subject-Raising Failure Approach, De Cat (2002), Ferdinand (1993, 1996), and Labelle and Valois (1996) argue that postverbal subjects such as those in (5) in child French are right dislocations without a doubled overt resumptive clitic. During this stage, in which postverbal subjects (which might be analyzed as right dislocations) are produced, null subjects with transitive and intransitive verbs also appear. In a later step, the child thus has to learn that French does not allow null subjects. Under this approach, bilingual children who acquire French and German should be observed to skip the null-subject stage in French. This acceleration effect could be due to the (positive) influence of German, which is a non-null-subject language. Notice, however, that Jansen (2015) and Patuto (2012) observe (target-deviant) null subjects in the bilinguals' French, as well as in their German (Pillunat et al. 2006; Schmitz et al. 2011) until around the age of 2;6. The Right-Dislocation Approach would struggle to explain the simultaneous absence of postverbal subjects and the presence of (target-deviant) null-subjects in bilingual French.

Under both analyses, the Subject-Raising Failure Approach and the Right-Dislocation Approach to postverbal subjects, there is at least one aspect of the child data which is not accounted for, namely the optional character of postverbal subjects, as shown in (5) and (6) above. From the work by Déprez and Pierce (1993) we learn that postverbal lexical subjects co-exist with preverbal lexical subjects. Therefore, we would like to suggest the following analysis of child postverbal subjects, which has the potential to reconcile the many important observations regarding child (monolingual and bilingual) French.

In the vein of the Minimalist Program (Chomsky 1995), Biberauer and Richards (2006) seek to eliminate language-specific properties of the computational system in order to explain alternations in word order within one language which do not go hand in hand with interpretive differences. In the ideal case, different outputs with the same lexical items should have different interpretations, with the exception of the extended projection principle (EPP) feature. This feature is obligatorily present (Chomsky 2013, 2015 for a different view), but languages vary as to the size of the nominal category (the [D] category (nominal category which has the characteristic features of DPs and associates definiteness) raised with or without head/spec-pied-piping) and the source of the nominal category ([D] in a head category or a Spec) that values the EPP feature in T. All that grammar requires is that the T's EPP feature be satisfied; it does not matter how this occurs. A single grammar can therefore contain truly optional structures, i.e., semantically vacuous optionality.

Biberauer and Richards (2006) take the example of Afrikaans, where both the verb-final and the V2 structure (in which the finite verb surfaces in the position preceding the first constituent, irrespective

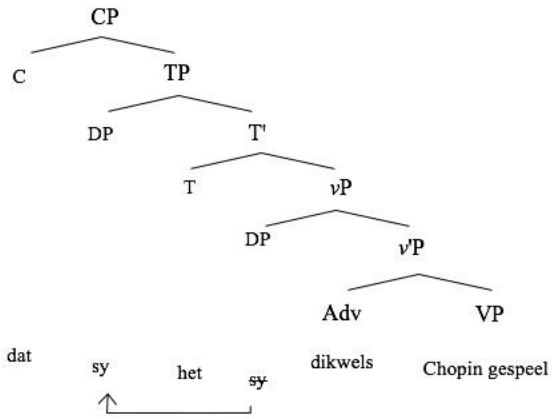
of whether this constituent is the subject or has another grammatical function in the clause) receive embedded interpretation.

7. a. Ek weet dat sy dikwels Chopin gespeel het
   I know that she often Chopin played has

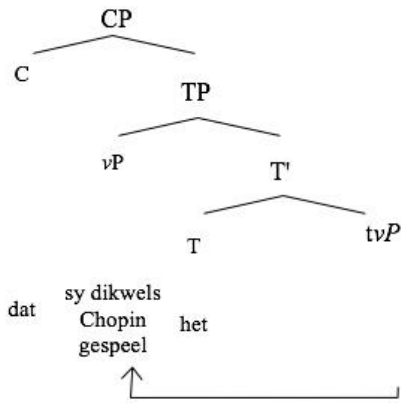
   b. Ek weet dat sy het dikwels Chopin gespeel
   I know that she has often Chopin played
   'I know that she has often played Chopin.'

The sentences in (7) share one interpretation. Biberauer and Richards (2006) assume for Afrikaans that the EPP feature in T must be satisfied by an element bearing the categorial feature [D], which is the closest accessible nominal category in T's sister domain. Afrikaans exhibits the EPP feature to be checked by a nominal category, which is either a DP raised into an outer Spec,*v*P (specifier of *v*P, of the small-verb Phrase), the specifier of TP, or by a nominal category which pied-pipes *v*P into the specifier of TP. In other words, there is either Spec-raising (7b) or Spec-pied-piping (7a). Arnaus Gil and Müller (forthcoming) analyze the alternation of German subordinate clauses (introduced by *weil* 'because', for example) which can be V2 or V-final (i.e., the finite verb appears clause-finally). In the former case, we are dealing with an instance of Spec-raising of a DP into Spec,TP (specifier of Tense Phrase, as in example (7b)), whereas a V-final structure corresponds to a process of pied-piping of *v*P into Spec,TP, as the example in (7a) has shown. The following structures in (8) below show the actual structure of the examples in (7a,b) above (cf. Arnaus Gil and Müller forthcoming).

7. a. Structure of an Afrikaans subordinate clause with V2 order



   b. Structure of an Afrikaans subordinate clause with V-final order



Let us turn to the subject domain in French in the framework of Biberauer and Richards (2006). Adult French in Table 1 is a non-pro-drop language in which an argumental (Spec) DP must obligatorily satisfy the EPP feature in T via Spec-raising. Many Romance languages, such as Italian, Spanish, Portuguese and Catalan, are pro-drop languages in which V_D has been argued to satisfy the EPP in T (Alexiadou and Anagnostopoulou 1998). In other words, movement of the finite verb to T suffices in rich agreement languages in order to satisfy the EPP feature. Of course, pro-drop languages also allow DPs in the specifier position of TP (e.g., *La ragazza mangia la mela* 'The girl eats the apple'), which neither signal a topic shift nor are contrastively stressed. Therefore, EPP satisfaction via Spec-raising is also an option for the adult system of these Romance languages, and it can be carried out by raising a DP into Spec,TP. For adult French, raising of the finite verb into T does not suffice to satisfy the EPP feature in syntax.

Table 1. Typology of extended projection principle (EPP)-in-T (Tense) satisfaction following Biberauer and Richards (2006).

Size/Type of Category Satisfying the EPP-Feature	Type: DP (Determiner Phrase)	Type: V_{finite}
Size: Raising	French, Afrikaans, German, Italian, Spanish, Catalan, Portuguese	Italian, Spanish, Catalan, Portuguese
Size: Pied-piping	Afrikaans, German	

A glance at Table 1 shows that all languages allow the EPP feature in T to be satisfied via the raising of a DP into Spec,TP. From the perspective of learnability theory, it acts as a default or a subset that all languages take into account.

However, there is another option in French for satisfying the EPP feature of T that we have not mentioned so far, namely the use of a subject clitic pronoun. Irrespective of how French subject clitics are analyzed with respect to their external categorial and their internal structure in the adult language—see Schmitz and Müller 2008 on the presence/absence of a D-layer and the head/phrase distinction, which is still an ongoing debate—the child needs time to determine the exact nature and function of subject clitics in French.[6]

Summing up: There are different kinds of EPP satisfaction in grammar. Grammar requires the EPP feature to be satisfied, but it does not matter how this occurs. EPP satisfaction in T is a case of true optionality in grammar.[7] To put it differently, the EPP feature makes the kind of optionality which is visible in child data possible to grasp, and even predictable.

Monolingual French children who are still in the stage of producing null-subjects, can thus explore raising of a subject DP into Spec,TP and raising the finite verb into T, as adult speakers of Italian, Spanish, or Catalan do (cf. Table 1). The V-raising into T would give rise to postverbal subjects (without a resumptive clitic, of course) in child grammar. We have nothing to say with respect to the possibility of analyzing these postverbal subjects as right-dislocated subjects, since a pragmatically non-neutral subject would be located outside the T-system.

---

[6] Cf. footnote 4. An investigation of the use of subject clitics in child French and in the children studied in the present cross-sectional study is in progress (Stahnke et al. 2018).

[7] One reviewer suggests that our analysis has the consequence that the difference between pro-drop and non-pro-drop languages is that pro-drop languages have more options for EPP satisfaction (namely V_D). Currently, we have nothing to say about this idea. We included a paragraph on the possibility of satisfying T's EPP feature via a subject clitic. We have seen in the data from the monolingual children that subject clitics are used from early on. The debate on the external and internal syntactic structure of French subject clitics is ongoing. Much relies on the analysis of these clitics. If the subject clitic cannot check off the EPP feature in T, then all sentences which contain a subject clitic without a doubled DP must involve a *pro* in Spec,TP. If it has the potential to check the EPP feature, it is likely that clitics are merged in T, as in Italian dialects, in which they can be obligatory. We believe that the postverbal subject stage in monolingual French is related to the unclear status of subject clitics in French. Any deeper discussion of this issue is outside the scope of the present article.

4. The Study

4.1. Participants

Table 2 illustrates the participants of a cross-sectional study of bilingual, trilingual, and multilingual children who reside in Germany or Spain.[8] The multilingual children speak more than three languages. The study included 63 children who have acquired French as one of their native languages. The 19 bilingual children range from 3;11–6;4 (mean = 5;0), the 38 trilingual children from 2;4–9;3 (mean = 5;0), and the 6 multilingual children from 3;4–6;4 (mean = 4;7). All children participating in the cross-sectional study were actually older than those in the literature. However, there is a relatively high number whose weak language is French, so a comparison can still be carried out with the children presented in the last section. Twenty-three children ranked low in French in the receptive vocabulary test, which we will present further below. The 19 bilingual children were bilingual with French and German (n = 15) or French and Spanish (n = 4). As can be seen from Table 3, the 44 trilingual and multilingual children spoke French and another Romance language, such as Catalan, Italian, Spanish, Portuguese, Galician, French and German or French and a Romance language and German. Other language(s) that the children learned were Arabic, Dutch, English, and Russian. The heterogeneity of the trilingual group is not relevant for the study of the French postverbal subjects.

Table 2. Children's language combinations in the cross-sectional study.

Number of Languages	Lang. A	Lang. B	Lang. C	Lang. D	Lang. E	Number of Children	Total
Bilingual	French	German				15	19
Bilingual	French	Spanish				4	
Trilingual	French	German	Arabic			2	38
Trilingual	French	German	English			18	
Trilingual	French	German	Russian			1	
Trilingual	French	German	Spanish			3	
Trilingual	French	Spanish	English			2	
Trilingual	French	Spanish	Italian			1	
Trilingual	French	Spanish	Catalan			10	
Trilingual	French	Spanish	Russian			1	
Multilingual	French	German	Spanish	English		1	6
Multilingual	French	German	Spanish	Catalan		2	
Multilingual	French	German	Spanish	Dutch		1	
Multilingual	French	Spanish	Catalan	Galician		1	
Multilingual	French	Spanish	German	English	Arabic	1	

Table 3. Participation of the children in the cross-sectional study.

Type of Multilingualism	Bilingual	Trilingual	Multilingual
Grammatical test participation	19	38	6
Receptive vocabulary participation	19	37[9]	6

What is important here is to be able to group the bilinguals, trilinguals and multilinguals according to the subject-type-language they acquire together with French (cf. Tables 1 and 2). Overall, 27 children spoke a null-subject language in addition to French.

[8] The study is part of a larger research project which is financed by the German Science Foundation (MU 875/12-1). The parents gave their consent for their children to take part in the study.

[9] One of the trilingual children was not tested in the receptive vocabulary task.

4.2. Methodology

We measured language proficiency with the PPVT (Peabody Picture Vocabulary Test, Dunn 1959; Dunn and Dunn 1981, 1997), from which the size of the receptive (hearing) vocabulary can be inferred. The French version of the test was designed by Dunn et al. (1993) and covers a broad range of French-language content words and syntactic categories (adjectives, adverbs, verbs, nouns). Children between the age of 2;6 and 18;0 can be tested. The administration should take 8 to 15 min. It consists of 170 items equally distributed across item sets (one item set contains 12 items). The French PPVT was designed on the basis of a representative sample of 2038 children in Canada. The children came from regions in Canada where English is spoken in addition to French (especially Québec and Ontario). A raw score is transformed into an age-dependent *t*-value which is linguistically described as:

- Extremely high "excellent"
- Moderately high "bon"
- High average "moyen"
- Low average "moyen"
- Moderately low "médiocre"
- Extremely low "faible"

The PPVT is a standardized test designed to analyze vocabulary development in typically developing French-speaking children and used in clinical assessment in order to detect language impairments and school readiness. The test offers a comparison with children who speak French as their native language (control groups) in several age groups. In the present case, it was used to measure linguistic competence in the different L1s of children who acquire more than one language from birth.[10]

The results of the bilingual and the tri- and multilingual children in the French PPVT are summarized in Figures 2 and 3, respectively.

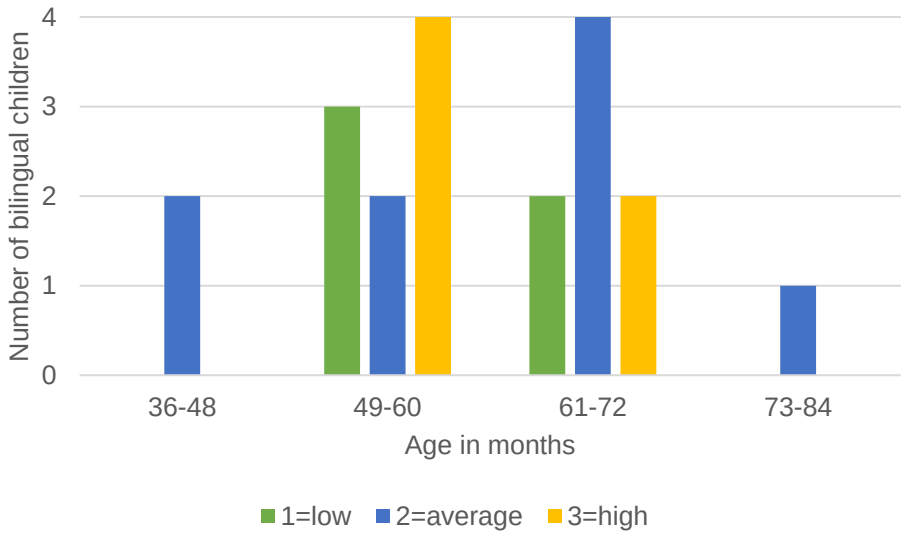


Figure 2. Results in the French Peabody Picture Vocabulary Test (PPVT): bilingual children.

[10] The PPVT was used in the languages under study for which this passive vocabulary test is available, that is, for French, German, Spanish and Catalan.

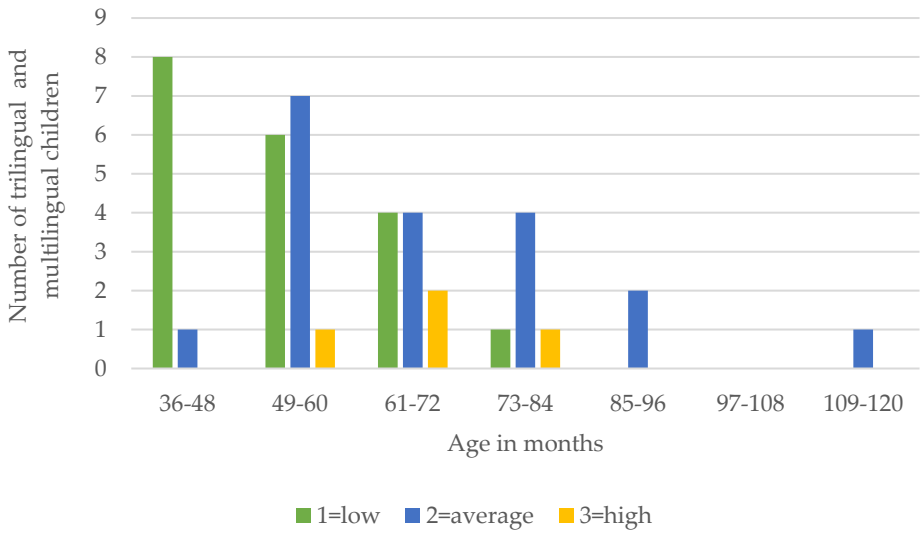


Figure 3. Results in the French PPVT: trilingual and multilingual children.

As can be seen from Figures 2 and 3 above, most bilingual children are in the average group, whereas for the trilinguals and multilinguals, there seems to be a relation between the average and the low group: the older the children get, the better their results in the French PPVT.

In addition to the PPVT, participants also had to complete a production task which aimed to elicit finite verbs and subjects in French. The grammatical test was composed of four cards which showed people performing different actions, all intransitive/unergative in order to guarantee that young children or children with low proficiency in French were able to participate in the test. The test items were the following:

9. a. clown–pleurer
   ‘clown–cry.’

   b. garçon–dormir
   ‘boy–sleep.’

   c. fille–sauter
   ‘girl–jump.’

   d. papi–rire
   ‘grandfather–laugh.’

The clown, the boy, the girl, and the grandfather were introduced to the child on separate cards before the testing began in order to be sure that the child knew the vocabulary used in the test. The child was asked by the experimenter: *Dans cette image, qu’est-ce qui se passe?* “What is happening in this picture?” to which the child was supposed to answer with a finite clause. Figure 4 illustrates the item of a boy who is sleeping. Notice that we elicited SV structures as can be seen from (9).

We deliberately used SV sentences in order to trigger broad focus and exclude the grammaticality of a right-dislocation with a clitic. Further below, it will become clear that the bilingual children never used a right dislocation as an answer to the experimenter’s question, the tri- and multilingual children used it once. De Cat (2002) claims in her work on dislocations that children know from the start of language production which construction to use in which pragmatic context, and she bases her analysis on children’s spontaneous data. This also seems to be true for the bi-, tri- and multilingual children in our experiment, since they almost never used a right dislocation.

Figure 4. French test item: *garçon–dormir* 'boy–sleep'.

4.3. Results

The French postverbal subject test successfully elicited intransitive verbs. Interestingly, bilinguals produce more intransitives than the tri/multilingual children. This result is illustrated by Figures 5 and 6.

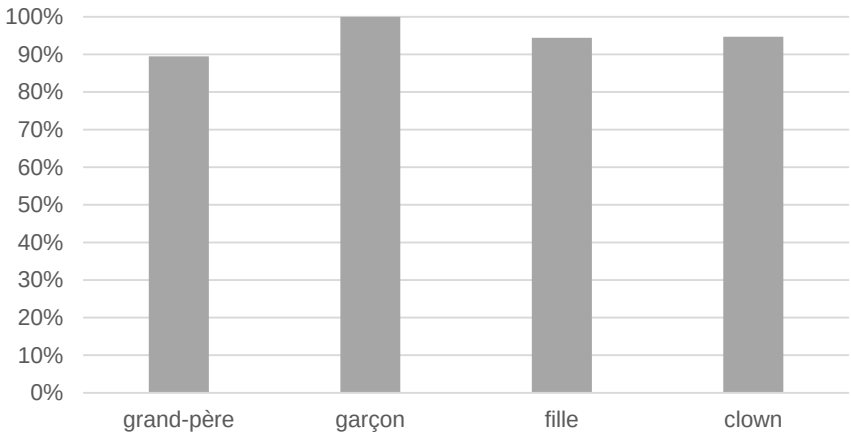


Figure 5. Usage of an intransitive verb per test item: 18 bilingual children.

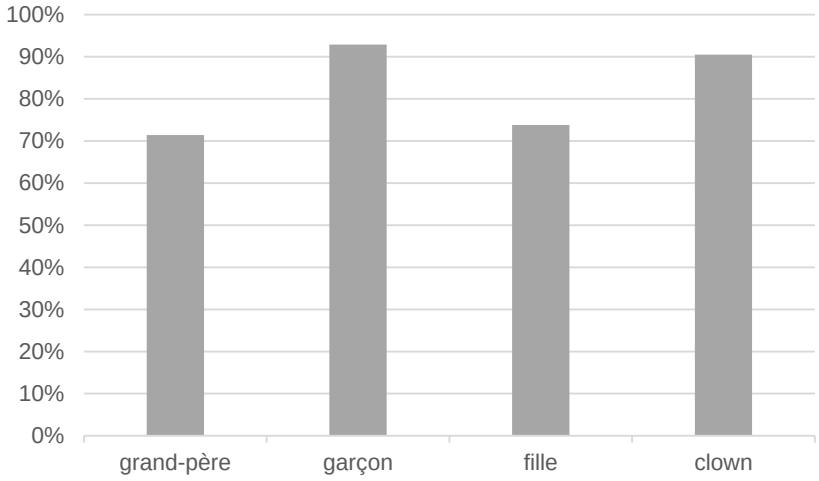


Figure 6. Usage of an intransitive verb per test item: 38 trilingual and 6 multilingual children.

The most frequent response from the bilingual children was an indefinite DP with a subject relative clause which contained the finite verb. The second most frequent response was a subject pronoun with a finite verb. Figure 7 illustrates the relative frequency of responses from the bilingual children. (10) gives an example of each response type.

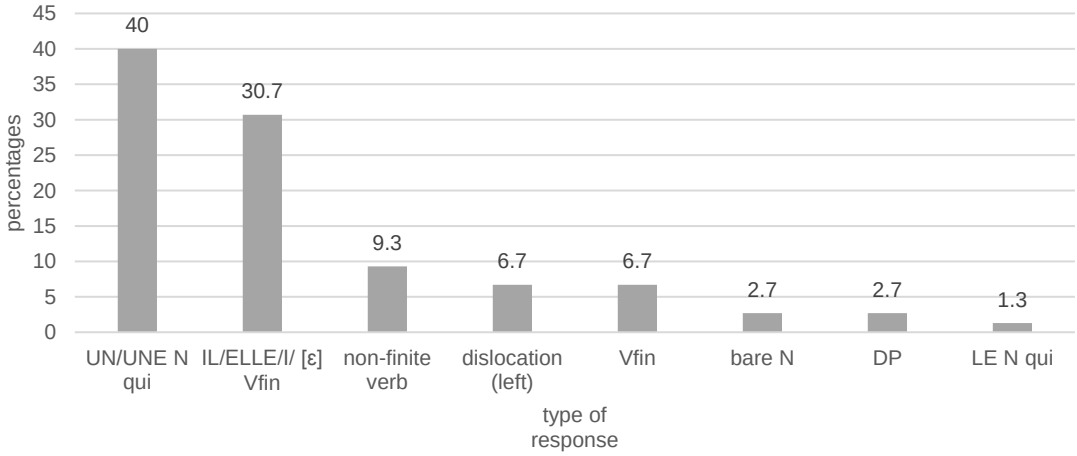


Figure 7. Responses to test items: 18 bilingual children.

10. a. LE N qui Vfin
le garçon qui dort (Simon, 5;3)
the boy who sleeps
'The boy is sleeping.'

b. UN/UNE N qui Vfin
euh un garçon qui dort (Delia, 5;7)
a boy who sleeps
'A boy who is sleeping.'

c. IL/ELLE/I/ [ɛ] Vfin
elle saute (Nadette, 3;11)
she jumps
'She is jumping.'

d. Dislocation (left)
le garçon i dortelle saute (Louisa, 5;5)
the boy he jumps
'The boy, he is jumping.'

e. Bare N
fille (Yann, 4;3)
girl

f. DP
un clown, tout triste (Lino, 3;11)
a clown, really sad

g. Vfin with null-subj
rit (Luca, 4;5)
laughs

h. Non-finite verb
rire (Sophie-Alixane, 5;9)
laugh-INF

i. UN/UNE N qui Vnon-fin
un clown qui qui pleuré (Fleur, 5;7)
a clown who who (has) cried/cry-INF

The following Figure 8 shows the percentage of responses for each type, taking into consideration whether the bilinguals acquire German or a Romance language together with French. In other words, taking, for example, the most frequent response in Figure 7 above with 40% of the total responses, Figure 8 below can be used to determine whether the bilingual group with French–German or the one with French–Spanish showed the same preference. As we can observe, only the former group exclusively answered with *UN/UNE N qui*. The latter seem to concentrate their answers mainly on responses with finite verbs only, bare N only, and *IL/ELLE Vfin*. These observations should be taken with caution, since the French–Spanish group contained only four children.

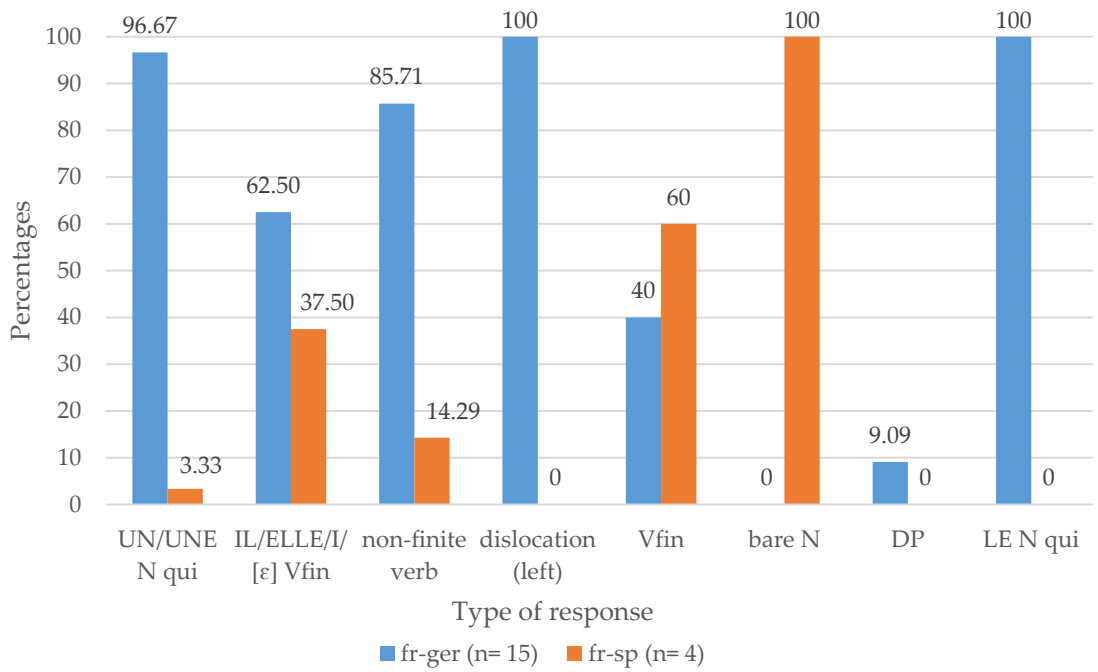


Figure 8. Responses to test items in comparison: bilingual German or Romance together with French.

In the trilingual and the multilingual children, the most frequent response was a subject pronoun with a finite verb, and the second most frequent answer was an indefinite DP with a subject relative clause that contained the finite verb. Remember that these were the same response types we observed for the bilingual group, but their order is reversed. Figure 9 illustrates the relative frequency of the responses of the trilingual and multilingual children. Figure 11 gives an example for each response type that was not used by the bilingual children.

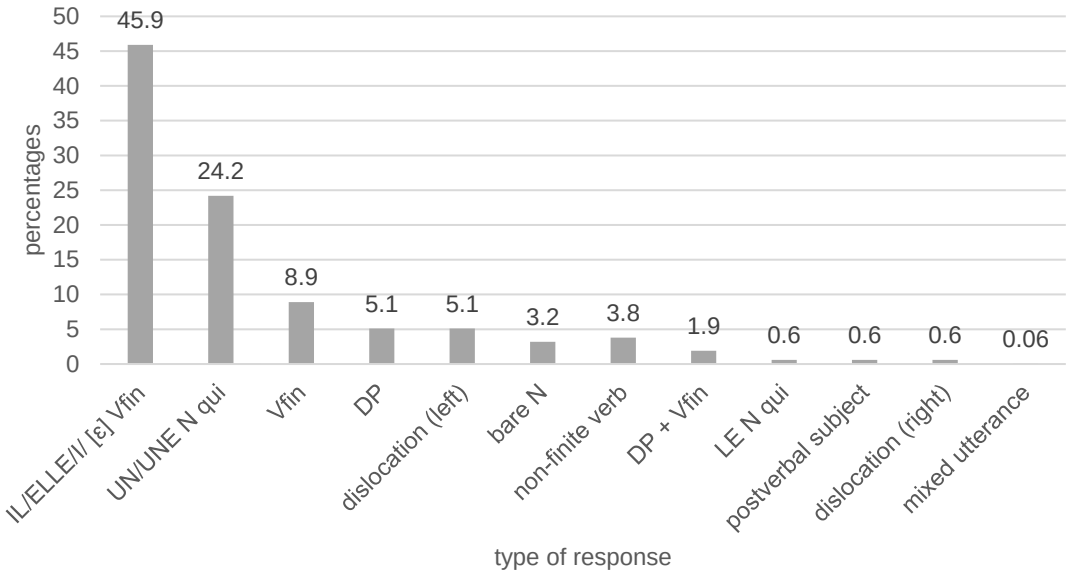


Figure 9. Responses to test items: 38 trilingual and 6 multilingual children.

11. a. Ça Vfin
ça saute (Emma, 3;9)
it jumps
'It is jumping.'

b. Dislocation (right)
il pleure le clown (Emma, 3;9)
he cries, the clown

c. Postverbal subject
fait, fille (Marla, 2;4)
makes, girl
'The girl makes (it).'

d. DP Vfin
la fille saute (Manon, 5;6)
the girl jumps
'The girl is jumping.'

e. Mixed utterance (German utterance)
schlafen (Ava, 2;7)
sleep-INF

Figure 10 shows further the percentage of responses for each type, taking into consideration whether the tri-/multilingual children are acquiring German, a Romance language, or both together with French.

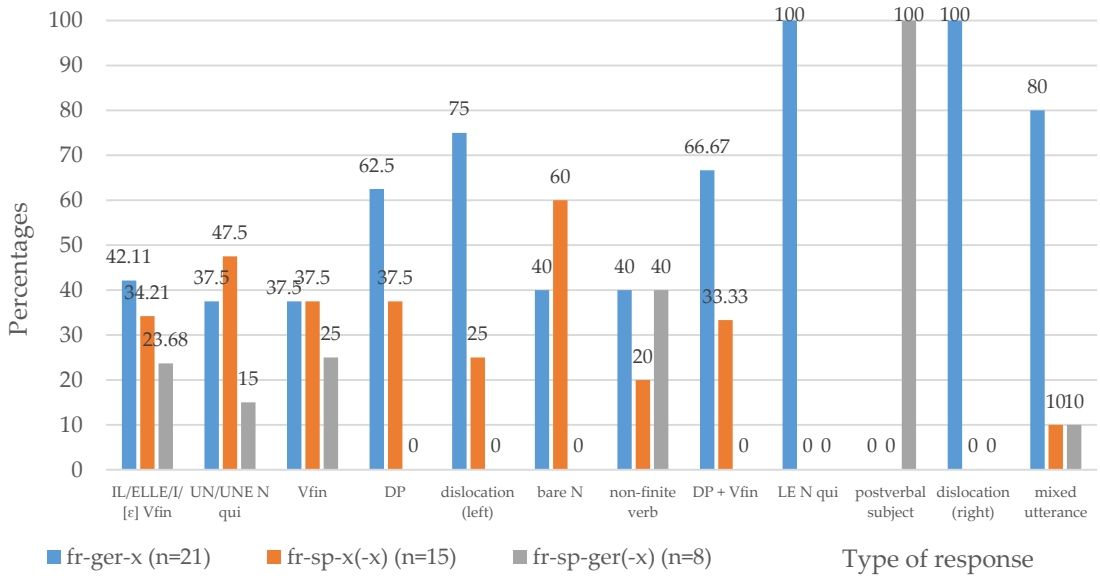


Figure 10. Responses to test items in comparison: tri- and multilingual German and/or Romance together with French.

In comparison to Figure 8 above for the bilinguals, Figure 10 shows a clear tendency for all children in the trilingual groups when it comes to the relative frequency of the most frequent responses in the test, namely subject pronoun+Vfin, indefinite DP+relative clause, and the production of a finite verb. Again, we have seen in Figure 9 that the first type of responses constituted the answer in 45.9% of the cases. Taking this into account, we see in Figure 10 that all tri- and multilinguals are represented in this response (subject pronoun+Vfin), the second (indefinite DP+relative clause), and the third most frequent answer (finite verb). This result is particularly interesting, since it implies that tri- and multilingual children show a preference for a preverbal subject, irrespective of the language combination, the number of languages (three or more), or the results achieved in the PPVT (cf. Figure 11).

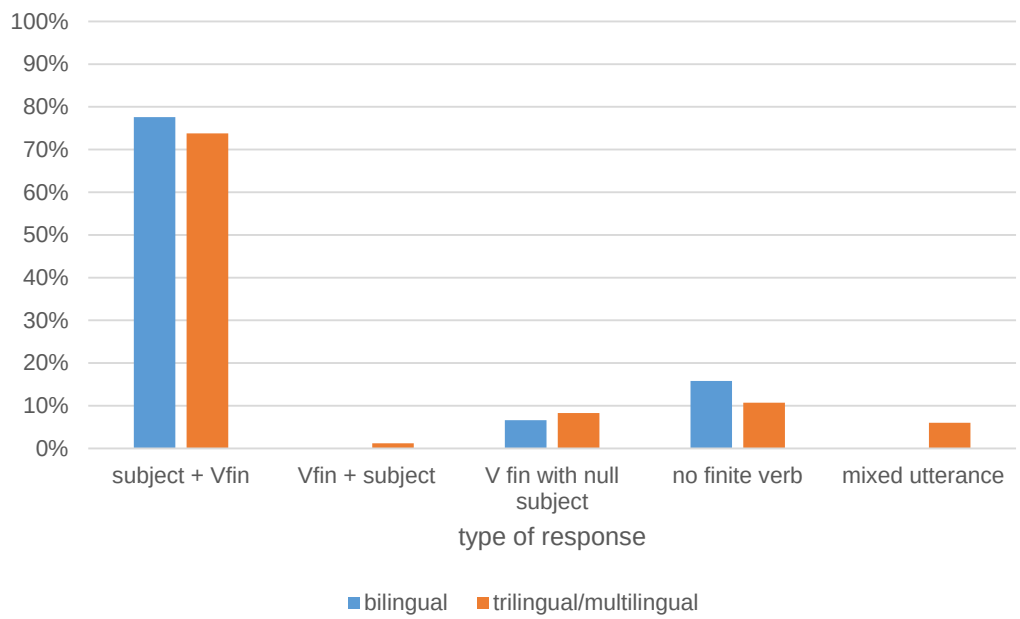


Figure 11. The position of subjects in bilingual, trilingual and multilingual children.

Figure 11 shows that the bilingual as well as the trilingual and multilingual children strongly preferred preverbal subjects. At the same time, there were bilingual and trilingual/multilingual children who exhibited structures reminiscent of early French: null subjects (10a), utterances which contained a non-finite verb form or no verb at all, and mixed utterances where the child used the non-response language (10e). These children ranked low in the PPVT. Although some structures characteristic of early French occur, there was only one utterance with a finite verb and a postverbal subject among all responses.

Interestingly, as illustrated in Figure 12, there were quite a number of children who ranked low in the French PPVT, in other words, the tested group of trilingual, multilingual and bilingual children contained a fairly high number of children who demonstrated low knowledge in the French receptive vocabulary test. Notwithstanding, the children almost never used postverbal subjects (1 of 487 total responses).

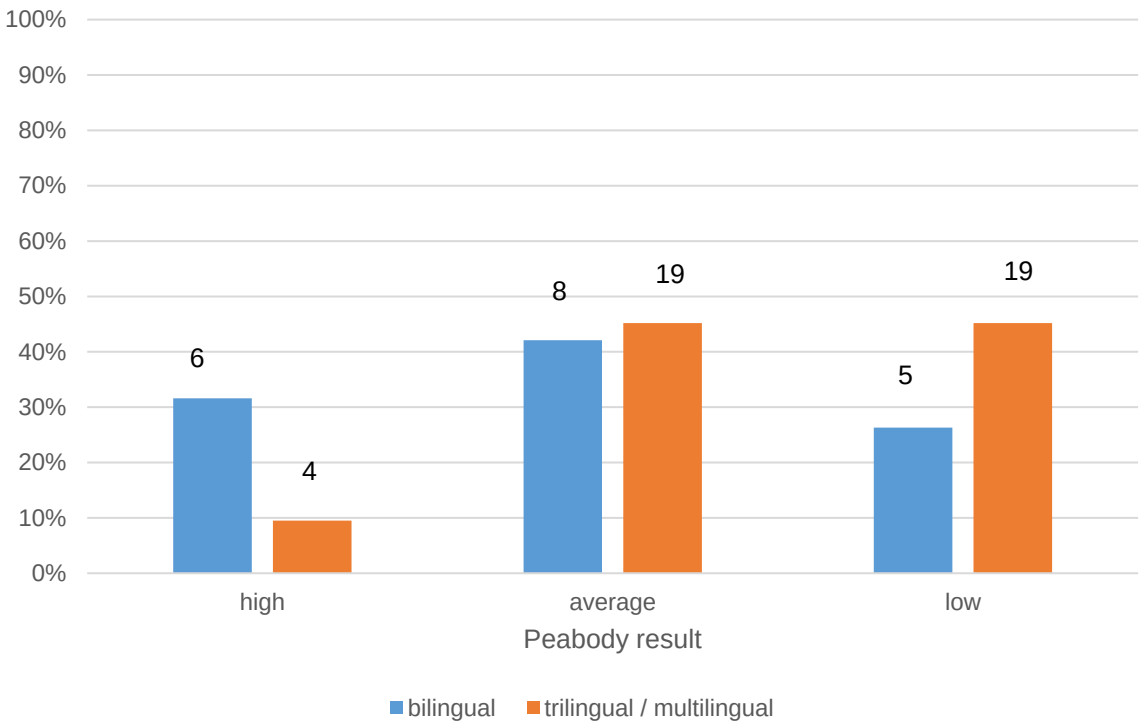


Figure 12. PPVT ranking of bilingual and trilingual/multilingual children compared.

In summary, we can conclude that, although a fairly high number of children ranked low in the PPVT and these children used structures which are reminiscent of an early stage in the development of French, they almost never used postverbal subjects, a structure characteristic of early monolingual French. The absence of postverbal subjects is observed in children who, in addition to French, acquire a spec-seeking language like German, and in children who acquire a null-subject language like Spanish.

5. Discussion

The review of the literature on early trilingualism has shown that the field is still in its infancy. Up to now, it has focused primarily on diary recollection. Code-mixing, language preference, and language balance have been issues that have captured linguists' attention in the last few years, but studies often lacked thorough analysis along an age span or among several trilingual children with a similar background. Research on specific linguistic phenomena in early trilingualism is scarce. Our article addresses both of these issues: We analyze a well-studied phenomenon in the early acquisition of monolingual French, namely the acquisition of pre- and postverbal subjects. Previous

research on this phenomenon with monolinguals and bilinguals has shown that whereas French monolinguals seem to undergo a stage (until 2;7) in which postverbal subjects seem to be one of the possible constructions (among many others, such as subjectless sentences, preverbal subjects and root infinitives), bilingual children acquiring French with German seem to skip this stage with respect to postverbal subjects. Our study contributes to this research field with a cross-sectional study carried out with 62 bilingual, trilingual, and multilingual children who had French as one of their L1s. The study was based on a production task which elicited intransitive subject–verb structures with the aid of visual material. The data show mainly preverbal responses across all children, even those who were classified as having low proficiency in French (according to the PPVT). These results replicate other studies on the acquisition of pre- and postverbal subjects in French in early bilingualism, that is to say, they show no significant amount of postverbal subjects, which seems to speak in favor of a multilingual advantage in acquiring French preverbal subjects.

From a theoretical perspective, we discussed Biberauer and Richards (2006) proposal on true optionality within the possibilities of EPP feature checking in French as well as in German and other Romance languages. Following this idea, we can observe that DP-raising is possible in all languages under study, so we could hypothesize that this is in fact the common ground, that is, the default in the early acquisition of French preverbal subjects. If we are correct, we should be able to claim that a bilingual, trilingual, or multilingual child takes advantage of the possibility of learning more than one L1 in the sense that DP-raising is facilitated and boosts the acquisition of the preverbal subject position in early bilingual, trilingual, and multilingual French. The fact that the children who acquire a null-subject language like Spanish also show the advantage in French indicates that the advantage cannot be explained with the use of the DP-raising analysis in the influencing language, since, after all, about 67% of all sentences in free conversations among adult native speakers of Spanish (and Italian) contain an empty subject, which underlies a head-seeking grammar (V_D (raised via v) can satisfy the EPP feature in T). If our syntactic analysis is correct, the children in this study have generalized preverbal subjects due to the default character of DP-raising into Spec,TP. In sum, the bilingual, trilingual, and multilingual children of this study evidence less variability in French than monolingual French children, an observation which goes back to Meisel (1989, p. 27), who observed less word order variability in children who acquire more than one native language.

Can we decide along these lines whether French postverbal subjects are indicative of the Subject-Raising Failure Approach and/or the Right-Dislocation Approach? If young monolingual children produce genuine postverbal subjects (i.e., structures in which finite verb rising into T suffices to satisfy the EPP feature), our result that bi-, tri-, and multilingual children lack structures which exhibit a postverbal subject may indicate an acceleration of knowledge concerning the subject position in French. However, if monolingual children produce right-dislocated subjects (which lack a resumptive pronoun), the fact that bi-, tri-, and multilingual children do not do so would indicate that they have a disadvantage with respect to monolingual children, who will continue to produce right-dislocations.[11] We still believe that our approach is neutral with respect to these two approaches for the following reasons: If we assume the Subject-Raising Failure Approach, we could argue within Biberauer and Richards (2006) analysis that a child who fails to raise the subject has under-appreciated the fact that the EPP feature must be satisfied, in this case via spec-seeking. If a bi, tri-, or multilingual child does not use postverbal subjects, but a monolingual child does, it might be argued that the bi-, tri-, or multilingual child is accelerated with respect to EPP satisfaction. If we assume the Right-Dislocation Approach, we could argue within Biberauer and Richards (2006) analysis that a child who fails to produce a clitic has under-appreciated the EPP feature as well, since it must be satisfied, in this case via the merging of a subject-clitic (or via the device which derives subject clitics in French, one of the most controversial topics in French linguistics). Since the focus of our analysis is the optionality of

[11] We thank one reviewer for pointing this out to us.

EPP satisfaction, a discussion of the two approaches to French postverbal subjects in the minimalist framework is outside the scope of our article. Let us not forget that there is still one problem to be solved, a problem which is difficult to solve within generative theory more generally: although the bi-, tri-, and multilingual children were accelerated with respect to postverbal subjects in that they did not produce them, they do not, like their monolingual peers, express a subject in every French finite sentence (in other words, they have subjectless finite sentences). We believe that this observation is completely unexpected under the Right-Dislocation Approach. In order to discuss the underlying syntactic analysis, we would have to focus on empty subjects as well, which, again, lies outside the scope of this study.

Concerning the elicited structures in the study, we deliberately used SV sentences, since they trigger broad focus and therefore a right-dislocation with a clitic would be ungrammatical. The bilingual children never used a right dislocation as an answer to the experimenter's question, and the tri- and multilingual children used one once. De Cat (2002) claims in her work on dislocations that children know from the start of language production which construction to use in which pragmatic context, and she bases her analysis on children's spontaneous data. This also seems to be true for the bi-, tri-, and multilingual children in our experiment, since they almost never used a right dislocation. Put differently, our results could also be argued to show merely that in broad focus contexts, the bi-, tri- and multilingual children do not produce genuine postverbal subjects, as expected from adult speakers of French. Since we did not test monolingual French children (the research project did not focus on French subjects, but tested several grammatical domains, among them copula use in Spanish and Catalan, adjective placement, and finite verb placement in several languages), and since Jansen (2015) did not analyze the longitudinal data of the monolingual and bilingual French children pragmatically, this interpretation is still possible.[12]

A particularly interesting result of our study is that, according to the literature, French monolingual children show more variability than French-speaking children who acquire more than one language. These observations add to the minimization of external factors such as the amount of exposure (input) in the acquisition process. We believe that the assumption of an EPP feature and its differing satisfaction with a default resource opens the way to a new perspective onto linguistic analyses of multilingualism. Exposure to diversity in the encoding of syntactic relations through the different L1s of the multilingual child may enhance syntactic differentiation, as Montanari (2010a, p. 4) has suggested. This differentiation in the different L1s can be perceived as a drop-out of available options in early language. In other words, children who acquire more than one language receive ideal input from a qualitative perspective. More research is needed to pursue this definitely surprising aspect of the present analysis.

A final note with respect to the syntactic approaches that exist for the description of the postverbal subject stage in monolingual French children is in order. If the present syntactic framework is plausible, monolingual children do not use postverbal subjects due to insufficient syntactic competence, but due to their grammar, which is in its initial stages and therefore allows null-subjects. Put differently, a monolingual French child can explore the raising of a subject DP into Spec,TP (producing a preverbal subject) and the raising of the finite verb into T, resulting either in a subjectless (main) clause or a clause with a subject in the postverbal position. This is a very interesting outcome of our proposal if one considers the fact that a monolingual child receives clearer and more consistent input (where the

[12] In other words, we cannot exclude the possibility that our results simply show that children who acquire more than one language know that broad-focus sentences trigger SV. Notice, however, that broad-focus sentences are extremely rare in child spontaneous speech. Therefore, we believe that more research which includes, in addition to cross-sectional data, longitudinal data of bilingual, trilingual, and multilingual children is needed. If broad focus is concerned, more monolingual French corpora should be investigated in future work. Nevertheless, a restriction to the broad focus context was necessary in this experiment to exclude dislocations and cleft-sentences as the children's responses. Future research should also exactly repeat the present production task with monolingual French children.

subject always raises to Spec,TP).[13] It remains for future research to implement the scenario depicted for the bi-, tri-, and multilingual children into a theory of first-language acquisition, in particular a theory of a parametrized syntax, namely that a bi-, tri-, or multilingual child can have a grammar which licenses null-subjects, but s/he does not explore all options to satisfy the EPP-feature (in T) as the monolingual child does (thus the lack of postverbal subjects). Put differently, it seems as if only a subset of properties of early (monolingual) French is observed in children who acquire more than one language, if we are on the right track. Since postverbal subjects arguably belong to a cluster of properties in child language, it is interesting that only one property of this cluster differs in children who acquire French together with (an)other language(s). Our research could eventually contribute to the non-parametrized view of the syntactic component in an architecture of the language faculty in which defaults are chosen in bi-, tri-, and multilingual children, all things being equal.

Our research highlights the need for more studies in the field of multilingualism beyond bilingualism with the tools used in longitudinal studies of spontaneous child speech. Furthermore, cross-sectional studies of bilingual, trilingual, and multilingual children should include comparisons with monolingual children.

Author Contributions: L.A.G. conceived and designed the experiments; a group of students all employed by the research project performed the experiments; L.A.G. and N.M. analyzed the data and wrote the paper.

Funding: The grant which made this study available comes from the German Science Foundation (MU 875/12-1).

Acknowledgments: We would like to thank the parents of the tested children who made this research possible. Also, our thanks go to the students who helped with elicitation and who transcribed and double-checked the data.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parametrizing AGR: Word order, V-movement, and EPP-checking. *Natural Language and Linguistic Theory* 16: 491–40. [CrossRef]

Arnaus Gil, Laia. 2013. *La selección Copulativa y auxiliar: Las lenguas romances (español–italiano–catalán–francés) y el alemán en contacto. Su adquisición en niños bilingües y trilingües*. Tübingen: Narr.

Arnaus Gil, Laia, and Natascha Müller. forthcoming. Acceleration and delay in bilingual, trilingual and multilingual German-Romance children: Finite verb placement in German. *Linguistic Approaches to Bilingualism*.

Ashby, William J. 1982. The drift of French syntax. *Lingua* 57: 29–46. [CrossRef]

Bardel, Camilla, and Ylva Falk. 2007. The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research* 23: 459–84. [CrossRef]

Barnes, Julia D. 2006. *Early Trilingualism. A Focus on Questions*. Clevedon: Multilingual Matters.

Barnes, Julia D. 2011. The influence of child-directed speech in early trilingualism. *International Journal of Multilingualism* 8: 42–62. [CrossRef]

Barron-Hauwaert, Suzanne. 2000. Issues surrounding trilingual families: Children with simultaneous exposure to three languages. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 5: 1–13.

Biberauer, Theresa, and Marc Richards. 2006. True optionality: When grammar doesn't mind. In *Minimalist Essays*. Edited by Cedric Boeckx. Amsterdam: Benjamins, pp. 35–67.

Braun, Andreas, and Tony Cline. 2010. Trilingual families in mainly monolingual societies: Working towards a typology. *International Journal of Multilingualism* 7: 110–27. [CrossRef]

Braun, Andreas, and Tony Cline. 2014. *Language Strategies for Trilingual Families. Parents' Perspectives*. Bristol: Multilingual Matters.

Cenoz, Jasone, and Charlotte Hoffmann. 2003. Introduction. Acquiring a third language: What role does bilingualism play? *International Journal of Bilingualism* 7: 1–5. [CrossRef]

[13] This issue was pointed out to us by one reviewer.

Cenoz, Jasone, and Ulrike Jessner, eds. 2000. *English in Europe: The Acquisition of a Third Language*. Clevedon: Multilingual Matters.

Cenoz, Jasone, Britta Hufeisen, and Ulrike Jessner, eds. 2001a. *Third Language Acquisition in the School Context*. Bristol: Multilingual Matters, vol. 4.

Cenoz, Jasone, Britta Hufeisen, and Ulrike Jessner, eds. 2001b. *Beyond Second Language Acquisition: Studies in tri- and Multilingualism*. Tübingen: Stauffenburg.

Cenoz, Jasone, Britta Hufeisen, and Ulrike Jessner, eds. 2001c. *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*. Clevedon: Multilingual Matters.

Chevalier, Sarah. 2015. *Trilingual Language Acquisition. Contextual Factors Influencing Active Trilingualism in Early Childhood*. Amsterdam: John Benjamins.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33–49. [CrossRef]

Chomsky, Noam. 2015. Problems of projection. Extensions. In *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*. Edited by Elisa Di Domenico, Cornelia Hamann and Simona Matteini. Amsterdam: John Benjamins, pp. 1–16.

Davidiak, Elena. 2010. One? ¿Dos? Drei! A Study of Code Switching in Child Trilingualism. Ph.D. dissertation, University of Iowa, Iowa City, IW, USA, May.

De Cat, Cecile. 2002. *French Dislocations*. Oxford: University of Oxford Dissertation.

De Houwer, Annick. 2004. Trilingual input and children's language use in trilingual families in Flanders. In *Trilingualism in Family, School and Community*. Edited by Charlotte Hoffmann and Johannes Ytsma. Clevedon: Multilingual Matters, pp. 118–35.

Déprez, Viviane, and Amy Pierce. 1993. Negation and functional projections in early grammar. *Linguistic Inquiry* 24: 25–67.

Dewaele, Jean-Marc. 2000. Trilingual first language acquisition: Exploration of a linguistic miracle. *La Chouette* 31: 77–86.

Döpke, Susanne. 1992. *One Parent One Language: An Interactional Approach*. Amsterdam: John Benjamins.

Döpke, Susanne. 1997. Is simultaneous acquisition of two languages in early childhood equal to acquiring each of the two languages individually? In *Proceedings of the 28th Annual Child Research Forum*. Edited by Eve V. Clark. Stanford: CSLI Publications, pp. 95–112.

Döpke, Susanne. 1998. Competing language structures: The acquisition of verb placement by bilingual German-English children. *Journal of Child Language* 25: 555–84. [CrossRef] [PubMed]

Dunn, Lloyd M. 1959. *Peabody Picture Vocabulary Test*. Circle Pines: American Guidance Service.

Dunn, Lloyd M., and Leota M. Dunn. 1981. *Peabody Picture Vocabulary Test—Revised*. Circle Pines: American Guidance Service.

Dunn, Lloyd M., and Leota M. Dunn. 1997. *Peabody Picture Vocabulary Test, Third Edition*. Circle Pines: American Guidance Service.

Dunn, Lloyd M., Claudia M. Thériault-Whalen, and Leota M. Dunn. 1993. *Échelle de Vocabulaire en Images Peabody*. Toronto: Pearson.

Edwards, Malcolm, and Jean-Marc Dewaele. 2007. Trilingual conversations: A window into trilingual competence? *The International Journal of Bilingualism* 11: 221–41. [CrossRef]

Elwert, W. Theodor. 1973. *Das Zweisprachige Individuum: Ein Selbstzeugnis*. Wiesbaden: Franz Steiner Verlag. First published 1959.

Faingold, Eduardo D. 1999. The re-emergence of Spanish and Hebrew in a multilingual adolescent. *International Journal of Bilingual Education and Bilingualism* 2: 283–95. [CrossRef]

Faingold, Eduardo D. 2000. Pro-active attitudes and educational strategies in early trilingual acquisition: Referential avoidance and parental intervention at the one-word stage. *Estudios de Lingüística Aplicada* 32: 58–73.

Ferdinand, Renata Astrid. 1993. Subject dislocations in French child language. *HIL Manuscripts* 1: 54–64.

Ferdinand, Renata Astrid. 1996. *The Development of Functional Categories—The Acquisition of the Subject in French*. Den Haag: Holland Academic Graphics.

Fernández Fuertes, Raquel, and Juana M. Liceras. 2010. Copula omission in the English developing grammar of English/Spanish bilingual children. *International Journal of Bilingual Education and Bilingualism* 13: 525–51. [CrossRef]

Flynn, Suzanne, Claire Foley, and Inne Vinnitskaya. 2004. The cumulative-enhancement model for language acquisition: Comparing adults' and children's pattern of development in first, second and third language acquisition of relative clauses. *The International Journal of Multilingualism* 1: 3–16. [CrossRef]

Friedemann, Marc-Ariel. 1994. The underlying position of external arguments in French: A study in adult and child grammar. *Language Acquisition* 3: 209–55. First published 1993. [CrossRef]

Gadler, Hanspeter, and Melaniji Mikeš. 1986. Der Erwerb der Muttersprache beim einsprachigen und beim dreisprachigen Kind. *Grazer linguistische Studien* 27: 27–43.

Gawlitzek-Maiwald, Ira, and Rosemarie Tracy. 1996. Bilingual bootstrapping. *Linguistics* 34: 901–26. [CrossRef]

Genesee, Fred. 1989. Early bilingual development: One language or two? *Journal of Child Language* 16: 161–79. [CrossRef] [PubMed]

Hager, Malin. 2014. Der Genuserwerb bei mehrsprachig aufwachsenden Kindern—Eine longitudinale Untersuchung bilingualer und trilingualer Kinder der Sprachenkombinationen deutsch-französisch/italienisch/spanisch, französisch-italienisch/spanisch und deutsch-spanisch-katalanisch. Ph.D. dissertation, Bergische Universität Wuppertal, Wuppertal, Germany, July.

Helot, Christine. 1988. Bringing up children in English, French and Irish: Two case studies. *Language Culture and Curriculum* 1: 281–87. [CrossRef]

Herdina, Philip, and Ulrike Jessner. 2002. *A Dynamic Model of Multilingualism*. Clevedon: Multilingual Matters.

Hoffmann, Charlotte. 1985. Language acquisition in two trilingual children. *Journal of Multilingual and Multicultural Development* 6: 479–95. [CrossRef]

Hoffmann, Charlotte. 1999. Trilingual competence: Linguistic and cognitive issues. *Applied Linguistic Studies in Central Europe* 3: 16–26.

Hoffmann, Charlotte. 2000. Bilingual and trilingual competence: Problems of description and differentiation. *Estudios de Sociolingüística* 1: 83–92. [CrossRef]

Hoffmann, Charlotte. 2001. Towards a description of trilingual competence. *The International Journal of Bilingualism* 5: 1–17. [CrossRef]

Hoffmann, Charlotte, and Anat Stavans. 2007. The evolution of trilingual codeswitching from infancy to school age: The shaping of trilingual competence through dynamic language dominance. *International Journal of Bilingualism* 11: 55–72. [CrossRef]

Hoffmann, Charlotte, and Steve Widdicombe. 1999. Code-Switching and language dominance in the trilingual child. In *AILE Actes du 8e Colloque EUROSLA Paris, Vol. 1: La Personne Bilingue*. Edited by Susan Foster-Cohen, Monique Lambert, Clive Perdue and Rebekah Rast. Paris: Université Paris VIII, pp. 51–62.

Hulk, Aafke, and Natascha Müller. 2000. Cross-linguistic influence at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition* 3: 227–44. [CrossRef]

Ivir-Ashworth, Ksenija Corinna. 2011. The Nature of Two Trilingual Children's Utterances: Growing Up with Croatian, English and German. Ph.D. dissertation, University of East Anglia, Norwich, UK, August 25.

Jansen, Veronika. 2015. *Dislokation im bilingualen Erstspracherwerb. Eine Untersuchung am Beispiel deutsch-französischer Kinder*. Berlin, München, Boston: de Gruyter.

Jessner, Ulrike. 1997. Towards a dynamic view of multilingualism. In *Language Choices: Conditions, Constraints and Consequences*. Edited by Martin Pütz. Amsterdam: John Benjamins, pp. 17–30.

Kazzazi, Kerstin. 2011. Ich brauche mix-cough: Cross-linguistic influence involving German, English and Farsi. *International Journal of Multilingualism* 8: 63–79. [CrossRef]

Kupisch, Tania. 2006. *The Acquisition of Determiners in Bilingual German-Italian and German-French Children*. München: Lincom Europa.

Labelle, Marie, and Daniel Valois. 1996. The status of postverbal subjects in French child language. *Probus* 8: 53–80. [CrossRef]

Lambrecht, Knud. 1986. *Topic, Focus, and the Grammar of Spoken French*. Berkeley: University of California Dissertation.

Leopold, Werner F. 1949a. *Speech Development of a Bilingual Child. A Linguist's Record III: Grammar and General Problems in the First Two Years*. New York: AMS Press.

Leopold, Werner F. 1949b. *Speech Development of a Bilingual Child. A linguist's Record IV: Diary from Age 2*. New York: AMS Press.

Liceras, Juana M., and Raquel Fernández Fuertes. 2018. Subject omission/production in child bilingual English and child bilingual Spanish: The view from linguistic theory. *International Journal of Romance Linguistics*. [CrossRef]

Liceras, Juana M., Raquel Fernández Fuertes, and Anahí Alba de la Fuente. 2012. Overt subjects and copula omission in the Spanish and the English grammar of English-Spanish bilinguals. On the locus and directionality of interlinguistic influence. *First Language* 32: 88–15. [CrossRef]

Lleó, Conxita. 2002. The role of markedness in the acquisition of complex prosodic structures by German-Spanish bilinguals. *International Journal of Bilingualism* 6: 291–313. [CrossRef]

Lleó, Conxita. 2016. Acquiring multilingual phonologies (2L1, L2 and L3): Are the difficulties in the interfaces? In *Manual of Grammatical Interfaces in Romance*. Edited by Susann Fisher and Christoph Gabriel. Berlin/München/Boston: de Gruyter, pp. 519–50.

Lleó, Conxita, and Margaret Kehoe. 2002. On the Interaction of Phonological Systems in Child Bilingual Acquisition. *International Journal of Bilingualism* 6: 233–37.

Lleó, Conxita, and Martin Rakow. 2006. The prosody of early two-word utterances by German and Spanish monolingual and bilingual children. In *Interfaces in Multilingualism*. Edited by Conxita Lleó. Amsterdam: John Benjamins, pp. 1–26.

Maneva, Blagovesta. 2004. 'Maman, je suis polyglotte!': A case study of multilingual language acquisition from 0 to 5 years. *International Journal of Multilingualism* 1: 109–22. [CrossRef]

Meisel, Jürgen M. 1989. Early differentiation of languages in bilingual children. In *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity, and Loss*. Edited by Kenneth Hyltenstam and Loraine K. Obler. Cambridge: Cambridge University Press, pp. 13–40.

Mikès, Melanie. 1991. Some issues of grammatical development in early trilinguals. *Anuario de Psicología* 48: 29–38.

Montanari, Simona. 2009. Multi-word combinations and the emergence of differentiated ordering patterns in early trilingual development. *Bilingualism: Language and Cognition* 12: 503–19. [CrossRef]

Montanari, Simona. 2010a. *Language Differentiation in Early Trilingual Development. Evidence from a Case Study*. Saarbrücken: Verlag Dr. Müller.

Montanari, Simona. 2010b. Translation equivalents and the emergence of multiple lexicons in early trilingual development. *First Language* 30: 102–25. [CrossRef]

Montanari, Simona. 2011. Phonological differentiation before age two in a Tagalog-Spanish-English trilingual child. *International Journal of Multilingualism* 8: 5–21. [CrossRef]

Montanari, Simona. 2013. Productive trilingualism in infancy: What makes it possible? *World Journal of English Language* 3: 62–77. [CrossRef]

Müller, Natascha. 2017. Different sources of delay and acceleration in early child bilingualism. *Zeitschrift für Sprachwissenschaft* 36: 7–30. [CrossRef]

Müller, Natascha, and Aafke Hulk. 2001. Cross-linguistic influence in bilingual children: Object omissions and Root Infinitives. In *Proceedings of the 24th Annual Boston University Conference on Language Development*. Edited by S. Catherine Howell, Sarah A. Fish and Thea Keith-Lucas. Somerville: Cascadilla Press, pp. 546–57.

Murrell, Martin. 1966. Language acquisition in a trilingual environment: Notes from a case-study. *Studia Linguistica* 20: 9–34. [CrossRef]

Nicoladis, Elena. 1999. Where is my brush-teeth? Acquisition of compound nouns in a French-English bilingual child. *Bilingualism: Language and Cognition* 2: 245–56. [CrossRef]

Nicoladis, Elena. 2003. Cross-linguistic transfer in deverbal compounds of preschool bilingual children. *Bilingualism: Language and Cognition* 6: 17–31. [CrossRef]

Notley, Anna, Elisabeth van der Linden, and Aafke Hulk. 2007. Cross-linguistic influence in bilingual children: the case of dislocation. In *Romance Language and Linguistic Theory 2005, Selected papers from "Going Romance"*. Utrecht, 8–10 December 2005. Edited by Sergio Baauw, Frank Drijkoningen and Manuela Pinto. Amsterdam: John Benjamins, pp. 229–59.

Oksaar, Els. 1977. On becoming trilingual. In *Deutsch im Kontakt mit anderen Sprachen*. Edited by Carol Molony. Kronberg: Scriptor Verlag, pp. 296–306.

Paradis, Johanne, and Fred Genesee. 1996. Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition* 18: 1–25. [CrossRef]

Patuto, Marisa. 2012. *Der Erwerb des Subjekts in (Nicht-)Nullsubjektsprachen: die Rolle des Spracheneinflusses und der Sprachdominanz bei bilingual deutsch-italienisch, deutsch-spanisch und französisch-italienisch aufwachsenden Kindern*. Tübingen: Narr.

Patuto, Marisa, Valentina Repetto, and Natascha Müller. 2011. Delay and acceleration in bilingual first language acquisition: The same or different? In *The Development of Grammar: Language Acquisition and Diachronic Change. Volume in Honor of Jürgen M. Meisel*. Edited by Esther Rinke and Tania Kupisch. Amsterdam: John Benjamins, pp. 231–61.

Pierce, Amy E. 1989. On the Emergence of Syntax: A Crosslinguistic Study. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, September.

Pillunat, Antje, Katrin Schmitz, and Natascha Müller. 2006. Die Schnittstelle Syntax-Pragmatik: Subjektauslassungen bei bilingual deutsch-französisch aufwachsenden Kindern. *Zeitschrift für Literaturwissenschaft und Linguistik LiLi* 143: 7–24. [CrossRef]

Quay, Suzanne. 2001. Managing linguistic boundaries in early trilingual development. In *Trends in Bilingual Acquisition*. Edited by Jasone Cenoz and Fred Genesee. Amsterdam: John Benjamins, pp. 149–99.

Quay, Suzanne. 2008. Dinner conversations with a trilingual two-year-old: Language socialization in a multilingual context. *First Language* 28: 5–33. [CrossRef]

Quay, Suzanne. 2011a. Introduction: Data-driven insights from trilingual children in the making. *International Journal of Multilingualism* 8: 1–4. [CrossRef]

Quay, Suzanne. 2011b. Trilingual toddlers at daycare centres: The role of caregivers and peers in language development. *International Journal of Multilingualism* 8: 22–41. [CrossRef]

Ronjat, Jules. 1913. *Le développement du Langage observé chez un Enfant Bilingue*. Paris: Champion.

Rothman, Jason. 2011. L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* 27: 107–27. [CrossRef]

Rothman, Jason. 2015. Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition* 18: 179–90. [CrossRef]

Schmitz, Katrin, and Natascha Müller. 2008. Strong and clitic pronouns in the monolingual and bilingual acquisition of French and Italian. *Bilingualism: Language and Cognition* 11: 19–41. [CrossRef]

Schmitz, Katrin, Marisa Patuto, and Natascha Müller. 2011. The null-subject parameter at the interface between syntax and pragmatics. Evidence from bilingual German-Italian, German-French and French-Italian children. *First Language* 31: 1–34. [CrossRef]

Stahnke, Johanna, Laia Arnaus Gil, and Natascha Müller. 2018. *Subject Clitics in the Acquisition of French: Bilingual and Trilingual Children Compared*. Unpublished Manuscript. Wuppertal, Germany: Bergische Universität Wuppertal.

Stavans, Anat. 1992. Sociolinguistic factors affecting codeswitches produced by trilingual children. *Journal of Language, Culture and Curriculum* 5: 42–53. [CrossRef]

Stavans, Anat. 2001. Trilingual narratives: Relating events in three languages. In *Narrative Development in a Multilingual Context*. Edited by Ludo Verhoeven and Sven Strömqvist. Amsterdam: John Benjamins, pp. 340–72.

Stavans, Anat, and Virginia Swisher. 2006. Language switching as a window on trilingual acquisition. *International Journal of Multilingualism* 3: 193–220. [CrossRef]

Unsworth, Sharon. 2013. Current issues in multilingual first language acquisition. *Annual Review of Applied Linguistics* 33: 21–50. [CrossRef]

van der Linden, Elisabeth H., and Petra Sleeman. 2007. Clitic dislocation: Evidence for a low topic position. In *Linguistics in the Netherlands*. Edited by Bettelou Los and Marjo van Koppen. Amsterdam: John Benjamins, pp. 173–87.

Wang, Xiao-lei. 2008. *Growing Up with Three Languages. Birth to Eleven. Parents' and Teachers' Guides*. Bristol: Multilingual Matters.

Widdicombe, Susan. 1997. Code-Switching, Coining and Interference in Trilingual First Language Acquisition: A Case Study. Master's thesis, University of Aston, Birmingham, UK.

languages

MDPI

Article

Brazilian Bimodal Bilinguals as Heritage Signers

Ronice Müller de Quadros [1,*] and Diane Lillo-Martin [2]

[1] Departamento de Libras, Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, R. Eng. Agronômico Andrei Cristian Ferreira s/n, Trindade, Florianópolis SC 88040-900, Brazil

[2] Department of Linguistics, University of Connecticut, 365 Fairfield Way, Storrs, CT 06269-1145, USA; diane.lillo-martin@uconn.edu

* Correspondence: ronice.quadros@ufsc.br; Tel.: +55-48-984544041

Received: 6 April 2018; Accepted: 6 August 2018; Published: 10 August 2018

Abstract: This paper presents an analysis of heritage signers: bimodal bilinguals, who are adult hearing children of Deaf parents who acquired sign language at home with their parents and the spoken language from the surrounding community. Analyzing heritage language with bimodal bilinguals who possess pairs of languages in different modalities provides a new kind of evidence for understanding the heritage language phenomenon as well as for theoretical issues regarding human language. Language production data were collected from four Brazilian bimodal bilinguals separately in both sign and speech, as well as from monolingual comparison Deaf signers and hearing speakers. The data were subsequently analyzed for various grammatical components. As with other types of heritage speakers, we observed a great degree of individual variation in the sign (heritage) language of balanced participants who patterned similarly to the monolingual signers, compared to those whose use of sign language differed greatly from monolinguals. One participant showed some weaknesses in the second (spoken) language. We approach the variation in language fluency in the two languages by considering the different contexts of language development and continuing use.

Keywords: bimodal; bilingual; heritage language; sign language; code-blending; Brazilian Sign Language; Brazilian Portuguese

1. Introduction

Heritage language: The Case of Hearing Children of Deaf Parents

This paper focuses on the language use of hearing adult children of Deaf[1] parents in Brazil. Children of Deaf adults—known as Codas—who are not deaf themselves may be said to acquire a sign language as a heritage language (Compton 2014; Palmer 2015; Pichler et al. 2017; Reynolds 2016). A heritage language is a minority language used in a specific socio-cultural context in which a different language is dominant in the community. Benmamoun et al. (2013, p. 132) define "heritage speakers" as, typically, second-generation immigrants who live in bilingual contexts. In this case, the heritage language is the language the children are first exposed to at home, before subsequently becoming dominant in the language of the broader society in which they live. In Brazil, for instance, where Brazilian Portuguese (BP) is the dominant societal language, any other language used in a specific community may be considered a heritage language including immigrant, ethnic and sign language communities. Similarly, in the United States, immigrant (and colonial) languages, as well as ethnic (Fishman 2001) and sign languages, are considered heritage languages in an English-dominant

[1] We use a capital D for Deaf to indicate the linguistic-cultural group rather than a medical diagnosis.

bilingual context. Because of the nature of their language acquisition, heritage speakers are often more competent speakers of their L2 (the socially dominant language) than their L1 (heritage language). The imperfect control that heritage speakers exhibit over their L1 often differs in significant and consistent ways from the imperfect control of second-language learners; frequently, heritage learners are understood to experience 'attrition' of their original language system, rather than incomplete acquisition (Polinsky 2006; Scontras et al. 2015).

Heritage speakers can shed useful light upon the current theoretical discussion about the nature of language, allowing us to adopt a novel approach to Chomsky (1986) important question: "what do we know when we know a language?". We may have intuitions about what a native speaker knows with respect to their language, but heritage language forces us to consider more deeply the question of what exactly it means for a person to be a native speaker. There is a consensus that native speakers/signers differ from non-native speakers/signers of a language because they acquired their language from a very early age within a natural input environment—this makes native speakers different from L2 speakers but identical to heritage speakers. Heritage speakers, like native speakers, acquire a home language naturally at an early age—the difference is that they also acquire a large community language at the same time, which they gradually come to rely on as their primary language in the dominant society (for example, Chinese at home and English at school and elsewhere).

Codas follow the model of heritage language acquisition when acquiring their sign and spoken language: they acquire sign language at home with their parents (and, in some cases, with the broader Deaf community) and the spoken language of their country with other people (hearing family members, colleagues at school, neighbors and other hearing people). In some cases, this acquisition model can lead to a non-native mastery of sign language in adulthood. We will follow Emmorey et al. (2008) in referring to heritage speakers of sign language as 'bimodal bilinguals', because they have two languages in two different modalities (sign and speech). Applying Benmamoun et al. (2013, p. 133) definition of heritage speakers to bimodal bilinguals, bimodal bilinguals are early bilinguals who grew up seeing (and signing) the heritage language (L1) and hearing (and speaking) the majority language either simultaneously or sequentially in early childhood (that is, roughly up to age 5) but for whom the majority language became the primary language at some point during childhood (at, around, or after the onset of schooling). As a result of language shift, by early adulthood a bimodal bilingual may be strongly dominant in the majority language (spoken language), while the heritage language (sign language) will now be the weaker language.

Although in Benmamoun et al. (2013) definition the heritage language necessarily becomes the speaker's weaker language over the course of time, in this paper, we will use the term "heritage speaker" to refer to any person whose home language differs from the dominant language in the general society. As our discussion below will show, heritage speakers of a sign language may have weak mastery of their L1, or they may pattern more like balanced bilinguals (see Pizer 2008). Lillo-Martin et al. (2014) observe that there is considerable variability among Codas in terms of the balance between the languages they are acquiring. This variation can be attributed both to the amount of spoken language used in the home (some Deaf parents may produce the spoken language vocally, while others do not—though they may follow parts of spoken language through speechreading—and most have knowledge of the written language, although this can vary in different locations) and the degree of support the children receive for their signing. Pichler et al. (2014) conclude that balanced bilingualism is most readily achieved by Deaf families who take time to encourage their children to sign with Deaf people in different contexts, as the wider society does not value sign language. Schools and the surrounding spoken language environment push these children to use English much more than their heritage language. For heritage signers, as for speakers of other heritage languages, the attitude of the children's input providers plays a role in their language choice (see (Döpke 1992) and (Lanza 1997) for unimodal bilinguals; (van den Bogaerde and Baker 2009) for Nederlandse Gebarentaal (NGT) [Sign Language of the Netherlands] and Dutch bimodal bilinguals; (Kanto et al. 2013) for Finnish Sign Language (FinSL) and Finnish bimodal bilinguals). Over time, however, in general the sign language tends to become the weaker language as

bimodal bilinguals begin to privilege the spoken language, even with Deaf interlocutors (Peyton et al. 2001; Kondo-Brown 2006).

On the other hand, since sign and spoken languages use different articulators, bimodal bilinguals differ uniquely from other heritage speakers in being able to produce both their languages simultaneously, using what is called code-blending (Emmorey et al. 2008). Code-blending differs from code-switching in that both languages are produced at the same time—whereas unimodal bilinguals have to learn to suppress one of their languages even when they code-switch, bimodal bilinguals can simultaneously use grammatical knowledge and lexical items from both languages, separately or combined, while continuing to observe language constraints (Lillo-Martin et al. 2014, p. 13). Lillo-Martin et al. (2014) observed that sociolinguistic factors influence the use of code-blending by young bimodal Codas, in the sense that they decide whether to use or avoid blending depending on who they are conversing with. On the other hand, Pyers and Emmorey (2008) found that adult Codas tend to employ grammatical sign-language facial expressions even when conversing in a speech-only modality with monolingual speaking people.

Another intriguing manifestation of bimodal bilingual production is known as bimodal "whispering". Petroj et al. (2014) describe whispering—in the case of bimodal bilinguals—as the use of the lexical items of a spoken language produced with little or no vibrations of the vocal cords during signing. The authors suggest that this practice serves to reduce the pressure of suppression of the spoken language. The authors found that the American Sign Language (ASL) and English bimodal bilinguals that they studied, whom they described as balanced bilingual language users, accommodated the grammar of their whispering to ASL structure rather than to English structure. Conversely, in contexts where heritage signers have begun to rely on the spoken language as their primary language, we might expect their "whispering" practice to take on more characteristics of the spoken language structure, rather than remaining true to the sign language structure.

All of the studies discussed above give us a good general insight into the linguistic behavior of bimodal bilinguals. The more general heritage language literature, however, provides some important models for further work. In unimodal heritage language, as in bimodal heritage language, significant variation is reported across individuals (see Benmamoun et al. 2013; and Scontras et al. 2015 for discussion). Polinsky (2008), for instance, focuses on noun categorization as a means to investigate both lexical access and sentence processing among American Russian heritage speakers. She argues for the value of noun categorization as a diagnostic tool because it bootstraps morphology, phonology, syntax, and simple semantic structures at different levels of linguistic representation. Similarly, Benmamoun et al. (2013, p. 136), following Polinsky (2005, 2006), advocate for the use of lexical proficiency as a fluency diagnostic. Polinsky (2006, p. 252) found that lexical proficiency can provide a good idea of the structural knowledge and overall competence that a heritage speaker has with each of their languages. In particular, she argues that lexical proficiency scores can be used as a basis on which to characterize the linguistic system of incomplete learners. Using this diagnostic, Polinsky (2005, 2006) found a strong correlation between a speaker's comprehension (measured in terms of oral translation of a basic word list) and grammatical phenomena (such as agreement, case marking, aspectual and temporal marking, pro-drop, co-reference, and embedding).

In a study investigating these same grammatical areas, Montrul (2011) found a tendency toward the simplification and over-regularization of complex morphological patterns in heritage language, as well as restricted word order. Likewise, Albirini et al. (2013) report that heritage speakers demonstrate significant syntactic vulnerability in their L1. The authors found that heritage language subject-verb agreement morphology is maintained more robustly than noun-adjective morphology in heritage speakers' oral production, while the unmarked singular masculine is more robust than other categories. These results suggest that asymmetries in heritage speech may be explained by a complex interaction of linguistic areas and frequency factors.

Another study from Albirini et al. (2011) analyzed the syntactic and morphological features (word order, use of null subjects, selection of prepositions, agreement, and possession) of oral narratives

produced by heritage Egyptian and Palestinian Arabic speakers in the U.S., paying particular attention to possible morphological gaps and the use of code-switching. The authors identified many non-native trends in heritage speakers' narratives, especially with respect to the use of plural and feminine nouns. In general, the speakers they analyzed were unable to apply agreement appropriately in overt or null pronominal contexts, indicating that these speakers knew the agreement morphology but lacked control over gender and number placement. The authors also found problems with number formation and agreement and the use of prepositions and possessives. All these gaps involve transfer from the dominant language to the heritage language, indicating incomplete knowledge or attrition of the L1. In terms of code-switching, Albirini et al. (2011) found that nouns were the most frequently switched category, followed by adjectives, then verbs, prepositions and adverbs. The authors concluded that the use of code-switching did not interfere with the basic sentence structure used by heritage speakers and thus did not result in ungrammatical sentences. These results are particularly interesting for our study, which includes the issue of code-blending amongst bimodal bilinguals. As we will discuss in the next section, the "Synthesis" model of language that we adopt predicts that code-mixing (switching or blending) can only felicitously apply when all feature checking in both languages is satisfied.

The goal of the present study is to expand our understanding of bimodal bilingual variation in language fluency in heritage signers' first language (sign language) and second language (spoken language).[2] Analyzing heritage language use among bimodal bilinguals who possess pairs of languages in different modalities (a sign and a spoken language) can shed new light on our understanding of heritage language phenomena more generally. The following questions are addressed by the study presented here: (i) What determines the range of variation within the heritage speaker cohort that we are considering? What is the role of language use in a given context? In considering these questions, we will take into consideration the role of sociocultural factors in the development and maintenance of the heritage sign language as well as factors of upbringing and early experience; (ii) What areas are vulnerable—that is, prone to cross-linguistic influence or errors—in both languages and what areas are robust? Why?

2. Materials and Methods

To address the research questions, we gathered data in both sign-only and speech-only contexts separately. For each language context, we assessed the following grammatical areas:

- mean length of utterance;
- word-per-minute rate;
- distribution of word types used;
- types of syntactic constructions used;
- verbal morphology production accuracy.

The diagnostic of word-per-minute output rate in spontaneous production is recommended in Benmamoun et al. (2013), following Polinsky (2008). Polinsky (2008) research has found that a heritage speaker's speech rate may be as low as 30% of the speech rate of full speakers of the same language. Polinsky (2008) also found a correlation between speech rate and grammatical knowledge (lower proficiency speakers have more difficulty accessing lexical items). In our study, we analyzed in detail the participants' word-per-minute output in each language, as well as their mean length of utterance (MLU) for each output modality. This information was compared with the participants' sign and speech proficiency scores.

Considering the correlation between word-per-minute rate and lexical proficiency reported in previous studies, we next checked for morphological asymmetries in the bimodal bilinguals' heritage

[2] We call the spoken language the second language because of chronology, even though the spoken language ends up being the dominant language (or the other language if they are balanced bilinguals).

language. Benmamoun et al. (2013) found that morphological deficits in heritage speakers occur more regularly in nominal morphology than in verbal morphology; they further showed that, within the realm of verbal morphology, agreement is the most vulnerable category while tense is the most robust (reported as well by Montrul (2011)). The authors hypothesize that heritage speakers preserve the mechanisms to generate syntactic structure but have less capacity to preserve post-syntactic operations (and therefore have difficulty mapping from syntax to Phonological Form (PF)). We predict that less-balanced heritage signers will exhibit "errors" in post-computation (phonological and morphological) materials in their heritage (sign) language, when compared to their speech production. If our findings align with those of previous researchers, these errors should occur mainly in verbal agreement. By pulling together all these elements from bimodal bilingual data, our study offers new insight into the rules and constraints that govern heritage language development and interaction.

2.1. Participants[3]

The primary participants are four heritage signers from Brazil, who are bilingual in Libras (Brazilian Sign Language) and BP. NT and JB are siblings; JB is the older child and they have a deaf mother and a hearing father. NT and JB's mother has a Deaf family with a brother, sisters, and cousins who are deaf. She is not formally educated, yet signs in a typical fashion with her deaf family and her children. NT and JB only sign with their mother and they typically do not blend sign and speech since their mother is monolingual in Libras. CL is a sign language interpreter; both of her parents are Deaf. CL's mother is sister to NT and JB's mother. Both of CL's parents had typical primary education. MR is a teacher of Deaf children; her parents are Deaf and they had a typical primary education.

Table 1 summarizes information about the participants. A Brazilian native bimodal bilingual assigned ratings of each participant's overall fluency on a scale from 1 (low)–7 (high) in Sign (Libras) and Speech (BP). The participants are listed in their order of sign proficiency. For comparison, four Deaf signers of Libras and four hearing non-signers also participated.

Table 1. Characteristics of the participants.

Group	Participant	Interpreter?	Sign Rating *	Speech Rating *
Bimodal bilinguals	CL	yes	7	7
	MR	no	6	7
	JB	no	4	7
	NT	no	3	6
Deaf signers	RM	no	7	n/a
	FR	no	7	n/a
	SD	no	7	n/a
	MS	no	7	n/a
Hearing non-signers	ZE	no	n/a	7
	SZ	no	n/a	7
	AR	no	n/a	7
	VS	no	n/a	7

* Ratings are based on the observation of a native bimodal bilingual.

2.2. Procedure

Data collection for the primary (Bimodal bilingual) participants took place within a larger study of bimodal bilingualism, the Development of Bimodal Bilingualism project (see bibibi.uconn.edu). The comparison data were collected separately, in a similar manner. For the current study, participants

[3] Research approval was granted by the Review Board at the Universidade Federal de Santa Catarina. IRB: CAAE 0014.0.242.000-09.

watched a short wordless animated movie clip (2 min) and immediately thereafter told the story they had seen to either a Deaf person (for Bimodal bilinguals in the Sign session and for Deaf participants) or a hearing person (for Bimodal bilinguals in the Speech session and for hearing non-signers). The participants were aware of the need to use Libras with the Deaf interlocutor and of the expectation that they would use BP with the hearing interlocutor. All sessions were video-recorded for later analysis.

After the videos were recorded, the processing and coding took place in the lab in several steps. The first step was to produce a complete transcription of the narrative produced by the participant, using the software program ELAN (Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands, https://tla.mpi.nl/tools/tla-tools/elan/; Crasborn and Sloetjes 2008). Signs from Libras are annotated using SignIDs, approximate BP translation equivalents written as uppercase glosses (Johnston 1991), with additional conventions for systematic representation of signed utterances (see (Pichler et al. 2010) for further information about the conventions). Each sign is represented in an individual annotation, using separate tiers for the right hand and the left hand. Spoken words are also entered individually on a tier for that purpose. Finally, a free translation is written on another tier, using both the signs and speech as well as contextual information to fill in any gaps in the utterances; the free translation also indicates the length and linguistic contents of each utterance.

Following the annotation process, each session was coded for the linguistic dimensions of interest in this study. We calculated each participant's production speed in spoken words or signs per minute using the individual annotations to indicate the number of words used. We also calculated the MLU in words (MLUw) in each language using the individual annotations to determine the number of words and the free translations to mark off utterances.

The final stage of analysis added a tier for the lexical category of each individual sign or word (verb, noun, functional category, adverb, pronoun, adjective, or WH-word) and a tier for the syntactic type (declarative, relative clause, embedded clause, coordinate clause, yes/no question, WH-question, fragment, or other). For the bimodal bilinguals, we also checked each signed verb for its morphological accuracy.

3. Results

The results of our calculation of production speed (signed or spoken words/minute), and the speech: the sign ratio of production speed is presented in Table 2. This table also presents the results of our calculation of MLUw for both sign and speech. Finally, for the bimodal bilinguals, the table provides the percentage of verbs produced with ungrammatical morphology in Libras.

Table 2. Production speed and mean length of utterance in words (MLUw) for sign and speech.

Group	Participant	Sign Words/Minute	Speech Words/Minute	Ratio Speech:Sign	Sign MLUw	Speech MLUw	%Sign VMorph Errors
Bimodal bilinguals	CL	73.84	74.29	1.01	5.50	7.46	0
	MR	86.25	155.60	1.81	3.63	6.40	3
	JB	63.42	126.00	1.99	3.70	6.47	12
	NT	48.15	41.66	0.87	2.54	5.00	59
Deaf signers	RM	129.12	n/a	n/a	7.27	n/a	n/a
	FR	79.81	n/a	n/a	6,68	n/a	n/a
	SD	77.85	n/a	n/a	6,04	n/a	n/a
	MS	100.71	n/a	n/a	5,64	n/a	n/a
Hearing non-signers	ZE	n/a	102.47	n/a	n/a	9.88	n/a
	SZ	n/a	107.14	n/a	n/a	9.00	n/a
	AR	n/a	147.64	n/a	n/a	8.96	n/a
	VS	n/a	143.63	n/a	n/a	6.37	n/a

The results of our analyses of lexical categories used in sign and speech are given in Figure 1. We found that the different participants within a group showed a similar distribution within a language, so we have collapsed the results by group/language.

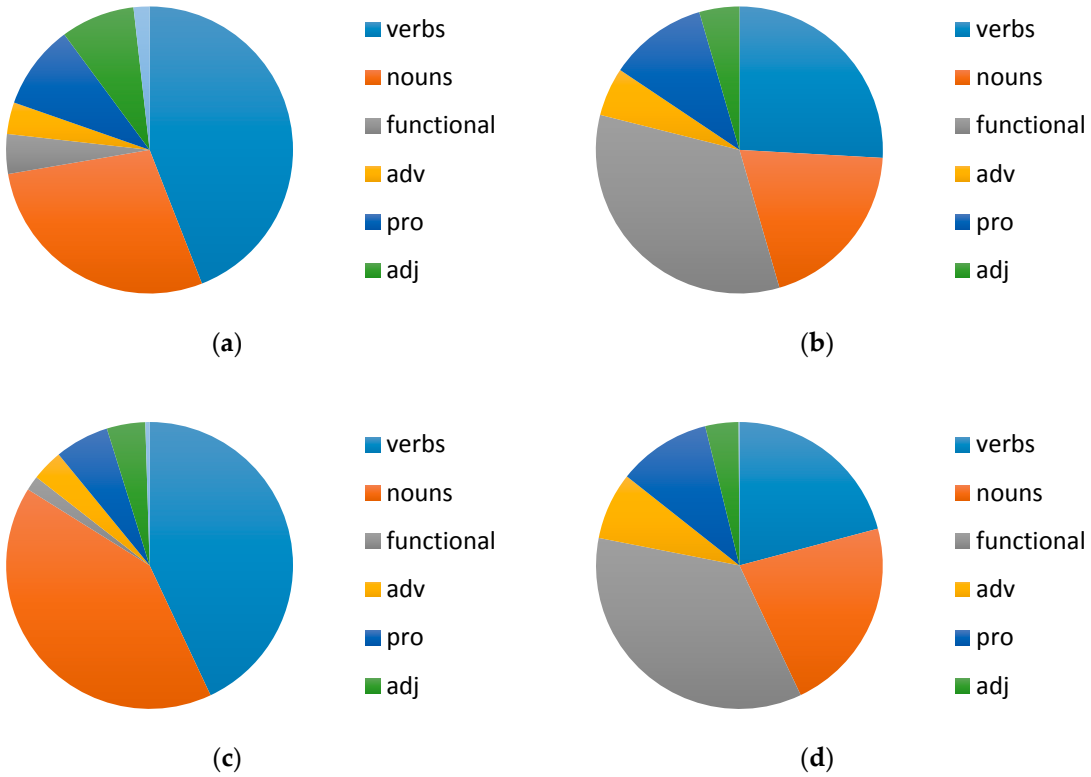


Figure 1. Lexical types used by (**a**) bimodal bilinguals in sign; (**b**) bimodal bilinguals in speech; (**c**) Deaf signers in sign; (**d**) hearing non-signers in speech.

The results of our analyses of syntactic types used in sign and speech are given in Figure 2. We found that the different participants within a group showed very different distributions within a language, so we have provided individual participant responses for each group/language.

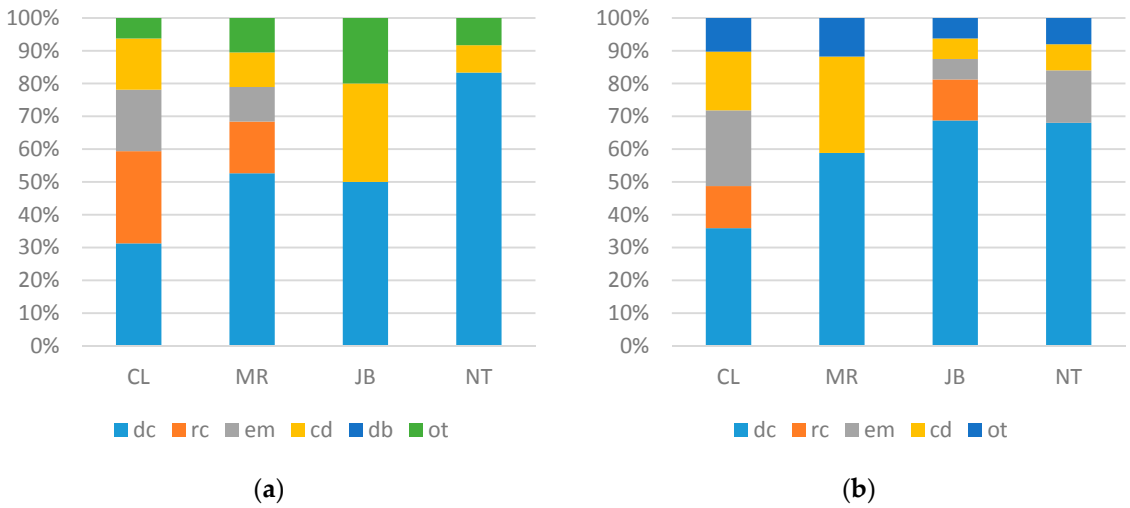


Figure 2. *Cont.*

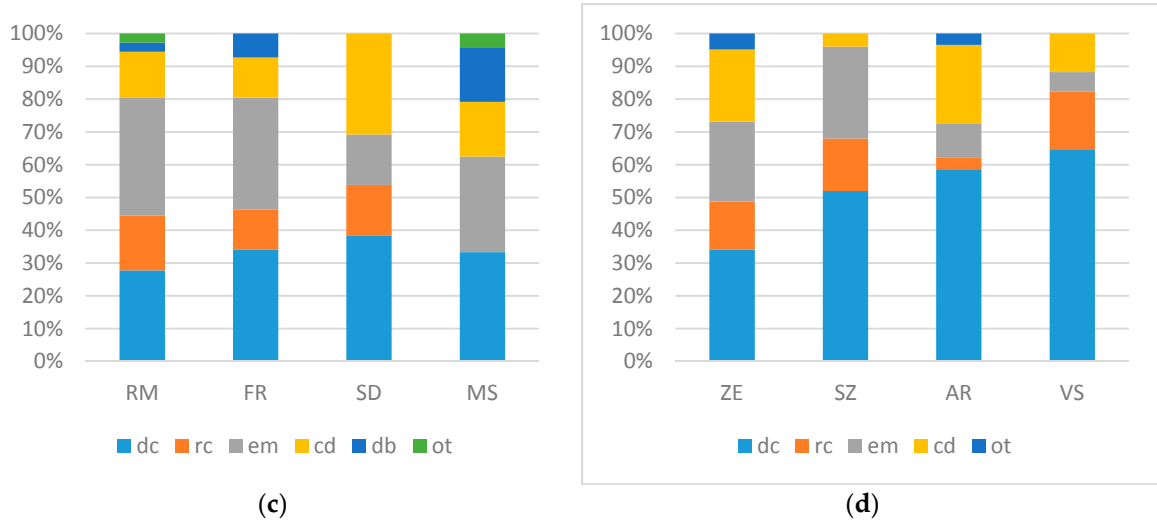


Figure 2. Syntactic types used by (**a**) bimodal bilinguals in sign; (**b**) bimodal bilinguals in speech; (**c**) Deaf signers in sign; (**d**) hearing non-signers in speech. dc: declarative; rc: relative clause; em: embedded; cd = coordination; db: double; ot: other).

4. Discussion

4.1. Empirical Discussion

Overall, the results clearly show a wide variability in proficiency with the heritage language for bimodal bilinguals. This is consistent with a number of studies discussing variable fluency in the sign language across bimodal bilinguals, including Compton (2014); Pichler et al. (2017); de Quadros (2017); de Quadros et al. (2016a, 2016b). The variability in the current data likely stems from individual differences, considering the differences in each participant's relationship with their two languages. CL, who is the highest-scoring signer and a sign language interpreter, displayed the highest MLUw in both speech and sign, and she produced sign and speech at equivalent rates. MR had a much higher word/minute rate in speech—possibly indicating that Portuguese is her strongest language, even though her signing is good—likely related to the fact that she is a teacher. JB had a lower MLUw and a slower production speed in sign (her weaker language) as compared to speech. NT exhibited the lowest MLUw and production speed in both sign and speech; however, it appeared that the test context made him somewhat uncomfortable, so this might not reflect his usual performance ability. In comparison, his sibling JB was much more comfortable when she did the test in Portuguese.

In comparison to the bimodal bilinguals, the monolingual Deaf participants showed the same or a higher rate of signs per minute and a consistently higher rate of signed MLUw. Only CL, the bimodal participant with the highest sign rating, showed an MLUw close to the minimal range of the Deaf signers. In speech, the monolinguals showed a high rate of speed comparable to MR and JB, while CL and NT were much slower. The speech of the bimodal bilinguals, except for NT, was within the wide range of MLUw in the Portuguese monolinguals. We note that the speech: sign production speed ratio for MR and JB indicates much higher rates of speech; on the other hand, the ratio is balanced for CL, who is strong in both languages, and for NT, who shows weaknesses in both languages. Given the overall lower performance by NT, it might be that, unlike the other participants, he shows the effects of having grown up in a heritage language context in both his languages, something which is also found for some hearing/speaking heritage speakers, with lower rates in both of their languages, as reported by Albirini et al. (2011, 2013); Montrul (2011); Polinsky (2006, 2008).

As for the lexical category analysis (Figure 1), we found that, for the most part, the individual participants did not differ within a language; rather, the greatest differences are observed across languages. All bimodal bilingual participants used a much greater proportion of verbs than nouns

in their sign but the monolingual signers showed this contrast less, with two individuals keeping the pattern and two going the other way. The predominance of verbs likely follows from the fact that, in Libras, as a null argument language (de Quadros 1995), it is very common for an utterance to contain only verbs, once referents have been introduced. In the experiment, the story being told centered around a particular referent that, once introduced, was clearly recoverable from the context. In speech, the participants showed a more balanced distribution of nouns and verbs, as an overt noun or pronoun must usually be used in BP.

Additionally, all participants used a much higher number of articles, prepositions and conjunctions in speech than in sign. This result is expected, since analogous structures are typically expressed in Libras by the movement of a sign, not by a separate lexical entry. Despite these overall patterns, however, we also noted a rich variability among the categories of words produced in each narrative in both languages for all participants.

The results of our analysis of sentence types (Figure 2) reveal considerable variability among the participants. NT and JB used less complex sentences in Libras than MR and CL did. This finding correlates with the word rates for each language: recall that NT and JB used fewer words per minute compared to MR and CL. Interestingly, NT also used fewer complex sentences when telling the story in Portuguese, while MR used a richer array of sentence types in Libras than in Portuguese, even though she had high word-per-minute rates in both languages. We may speculate that MR adjusted her Portuguese as if telling the story to children, which she did not do when she was signing. As MR is a teacher of children, she may have unconsciously followed this pattern of simpler spoken sentences in this context. CL demonstrated the richest distribution of sentence types in both languages; her word/minute rates were lower than MR's, but still higher than JB's and NT's.

The monolingual comparison reveals again that CL and MR pattern closely to Deaf signers, who make use of a variety of sentence types in their productions. On the other hand, JB and NT contrast with the monolinguals in showing a much reduced variety of sentence types. As for the BP speakers, the variability in their patterning is quite similar to the variability seen across the speech of the bimodal bilinguals, suggesting that the participants are generally comparable in their use of the spoken language.

Table 2 above gives the rates of ungrammatical verbal morphology in the sign of the four bimodal bilinguals. Note that the occurrences of ungrammatical verbal morphology in Libras track the overall fluency level for each participant. Such a pattern is also found in studies of hearing/speaking heritage language users—those who show weaknesses in rate and sentence complexity overall also tend to show errors in verbal morphology (Albirini et al. 2011, 2013; Montrul 2011).

Many of the observed Libras verbal morphology errors are related to the choice of handshape in depicting verbs. Verbs in this language employ particular handshapes to express semantic categorical information and it seems that NT and JB did not always know the right handshape for their categorical choices. Previous research suggests that mastering this morphological aspect is challenging for first- and second-language learners of various sign languages (e.g., (Karnopp 1994) on Libras). An example from NT is given in (1).

1. *DS(pato?-passou-na-frente-do-caminhão)
   duck?-pass-in-front-of-truck
   'The duck went in front of the truck like so.'

In this example, the notation DS refers to depicting signs, and the description within parentheses explains what the sign is conveying, but the individual words connected by hyphens should not be taken to represent separate signs or morphemes (see (Emmorey 2003) for more discussion about depicting signs, also known as classifier constructions). In the production by NT, the handshape for *PATO* ('duck') is wrong but the intention behind the sign can be recovered from the story. First, NT signed *BORBOLETA* ('butterfly') in place of *PATO* ('duck'). Then, he used a C handshape, which would be used to represent holding an object, and moved the C handshape in front of the handshape

representing the truck. The target handshape would be one indicating animal legs; there are two acceptable versions in Libras, one with only one hand (5 handshape) or one with both hands signing symmetrically (3 handshape). Since NT appeared to lack sufficient vocabulary to access while he was signing, in his storytelling he often did not use signs or relied on signs that were more gesture-like. Conversely, his sister JB knew the signs but she produced another kind of error in Libras involving the use of signing space to represent referents. Signers should use different locations in space for different referents; these spatial loci are used in the systems of pronominals and verb agreement. However, rather than using different loci for different referents, JB put multiple referents in the same location. Interestingly, this same type of error is also observed at the beginning of the sign language acquisition process among Deaf children of Deaf parents (Bellugi et al. 1990).

Another kind of error in Libras verbal morphology concerns adverbial marking, including markers for repetition and emphasis. There are no occurrences of such adverbial morphology associated with verbs in NT's data, and there are just a few occurrences of this marker in the productions of JB, but there are many occurrences produced by CL and MR. This also may indicate the poverty of the signs produced by NT and JB. In narratives, it is very typical to have adverbial markers in the signed story because the facts include a lot of information that should be marked. See the example of this marker produced by CL in (2). The sign *SUSTO* ('frighten') produced by CL is marked with an adverbial so it means 'frightened very much'.

2. HOMEM SUSTO+ PATO DS(pato-passar-frente-caminhão) DS(pato-caminhar)
   man be-afraid duck duck-cross-in-front-of-the-truck duck-walking
   'The man was frightened by the duck who was crossing in front of the truck like this.'

4.2. General Discussion

The genre of the productions analyzed in this study is narrative, reducing the opportunities for production of certain sentence types such as conditionals, yes/no questions and WH-questions. Several independent reasons may explain why there is variability in the sentence types used by each storyteller, but when we compare sentence complexity with all the other information regarding individual differences in word/minute production speed, MLU, verbal morphology and the use of adverbial modifications, we can see an indication that each participant relies more heavily on one or the other language as primary (even if this is not the dominant language), as has been reported by Benmamoun et al. (2013) for other heritage speakers.

In these heritage signers, it seems that NT and JB present a divergent Libras structure. However, it is interesting that for NT the story told in Portuguese also presents a low production speed, which is different from the others. Also, for MR, Portuguese has a very small variety of sentence types, while in Libras she presented a greater variety of different types of structures. MR has more contact with her deaf parents than with other signers but she may contact other deaf people on special occasions, and deaf children in the school where she works. CL is the most balanced signer/speaker; as she is also a sign language interpreter, it is not surprising that there are such clear differences between her and the others who sign with their families only or primarily. CL has contact with intellectual deaf signers, since she is an interpreter in the academic level, as well as in Deaf associations and in formal organizations, besides her deaf family. These different contexts may influence the signing skills of the four participants.

We now turn our discussion to addressing two overarching themes for considering bimodal bilinguals as heritage language users.

4.2.1. What Determines the Range of Variation within the Heritage Speaker Cohort that We Are Considering? What is the Role of Language Use in a Given Context?

JB and NT had contact only with their Deaf parents, and a few other Deaf people from their parents' circle of friends. Furthermore, their parents do not have formal education at all. It seems that

this lower level of education combined with access only to a more restricted community of signers may have influenced their signing skills.

Another relevant factor is the difference between participants in their professional involvement with Deaf people. This impacts directly in their bilingual status and in the way that they manage both languages. CL is a balanced bilingual and she is trained to use the two languages in different contexts. Moreover, she developed control of the uses of the languages, even in contexts using code-blending (Lillo-Martin et al. 2010, 2014, 2016). The other participants are not interpreters, but MR is a teacher of Deaf children. She has contact with other Deaf people but she usually restricts her contact to the children that she interacts with, in addition to her Deaf parents. This means that while she has access to a greater variety of signers than JB and NT do, she is still somewhat limited in use of Libras, which can have some implications for her signing skills.

4.2.2. What Areas Are Vulnerable in Both Languages and What Areas Are Robust? Why?

In this study, some bilinguals showed weaknesses in the use of Libras verbal morphology. As with spoken heritage languages, agreement marking appears to be particularly affected. In addition to agreement marking, the morphology of classifier or depicting signs—which requires specific handshape morphemes—was also vulnerable.

The weaker signers also showed a tendency to use simple sentences instead of more complex structures (declarative sentences and coordinate sentences with a few or no occurrences of embedded clauses, relative clauses, conditionals, doubling and other types). Such avoidance of embedded structures is also typical of various heritage languages (Benmamoun et al. 2013) and our results indicate that adherence to shallow structure is independent of modality. Deficiencies in syntactic domains can be investigated more systematically in comprehension tasks, and we leave this for further study.

There is some evidence for weaknesses in the spoken language for only one bimodal bilingual—NT—although it should be kept in mind that these weaknesses may be related to the fact that he was not comfortable with the tests and the cameras. In comparison to the other bilinguals and the monolinguals, NT was slower, had a lower MLUw, and used fewer sentence types in BP. The other three participants showed no weaknesses in their spoken language. This variability may be related to the kind of interactions in sign and in speech that the participants experienced while they were growing up.

5. Conclusions

This paper presented and analyzed narrative data from four bimodal bilinguals using Libras and Portuguese in Brazil. We have characterized these bimodal bilinguals as heritage speakers of their sign language. Despite this blanket characterization, there is a great degree of variance across our participants, from the more balanced ones to the ones who are more spoken-language dominant. This variation is reflected in their use of lexical material, rate of language production, and other measures.

Those speakers who can be identified as unbalanced bimodal bilinguals, with the sign language as their weaker language, show a number of properties that are typical of heritage speakers in spoken languages. In particular, their rate of production, MLU, and sentence types used are more restricted than those of the more balanced bilinguals; in addition, they have difficulty in managing material related to morphology, especially agreement and classifiers.

Turning now to the roots of variation in language maintenance, for the speakers considered here, their fluency depended on a number of social factors as well as on the amount of input they received in the sign language. In fact, we found that these factors are tightly intertwined—probably more so than they are intertwined in spoken languages—because the Deaf community is so strongly identified with Deaf culture. Relatedly, another relevant factor in language variation across our participants had to do with their professional involvement with Deaf people, since one of the participants is a professional

sign language interpreter. This research is the first step in our understanding of language variation in bimodal bilinguals; additional research considers the ways that bimodal bilinguals combine their languages in the code-blending phenomenon, which provides further means to investigate the nature of heritage bimodal bilingualism (de Quadros 2017; Lillo-Martin et al. 2016).

Author Contributions: R.M.d.Q. developed the predictions for the heritage language approach to this study, which builds on prior collaborative work between R.M.d.Q. and D.L.-M. The bimodal bilingual data were collected as part of this collaborative project. R.M.d.Q. collected the data and conducted the data analysis, with frequent input from D.L.-M. R.M.d.Q. drafted the first version of the paper, which was revised collaboratively.

Funding: This research is supported in part by Award Number R01DC009263 from the National Institutes of Health (National Institute on Deafness and Other Communication Disorders). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDCD or the NIH. Support is also provided by awards from CNPq (Brazilian National Council of Technological and Scientific Development) Grants #200031/2009-0, #470111/2007- 0, and #234255/2013-7.

Acknowledgments: We gratefully acknowledge the contribution of the bimodal bilingual participants and the Deaf and hearing monolingual comparison subjects. Maria Polinsky provided valuable discussion about some of the issues in this work while the first author was a visiting scholar at Harvard University.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

Albirini, Abdulkafi, Elabbas Benmamoun, and Eman Saadah. 2011. Grammatical features of Egyptian and Palestinian Arabic heritage speakers' oral production. *Studies in Second Language Acquisition* 33: 273–303. [CrossRef]

Albirini, Abdulkafi, Elabbas Benmamoun, and Chakrani Brahim. 2013. Gender and number agreement in the oral production of Arabic Heritage speakers. *Bilingualism: Language and Cognition* 16: 1–18. [CrossRef]

Bellugi, Ursula, Diane Lillo-Martin, Lucinda O'Grady, and Karen van Hoeck. 1990. The Development of Spatialized Syntatic Mechanisms in American Sign Language. In *The Fourth International Sympsium on Sign Language Research*. Edited by William H. Edmondson and Fred Karlson. Hamburg: Signum Verlag Press, pp. 16–25.

Benmamoun, Elabbas, Silvina Montrul, and Maria Polinsky. 2013. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39: 129–81. [CrossRef]

Chomsky, Noam. 1986. *Knowledge of Language: Its Origin, Nature and Use*. New York: Praeger.

Compton, Sarah E. 2014. American Sign Language as a Heritage Language. In *Handbook of Heritage, Community, and Native American Languages in the United States: Research, Policy, and Educational Practice*. Edited by Terrence G. Wiley, Joy Kreeft Peyton, Donna Christian, Sarah Catherine K. Moore and Na Liu. New York: Routledge, New York: Center for Applied Linguistics.

Crasborn, Onno, and Han Sloetjes. 2008. Enhanced ELAN functionality for sign language corpora. In *Construction and Exploitation of Sign Language Corpora. 3rd Workshop on the Representation and Processing of Sign Languages*. Edited by O. Crasborn, E. Efthimiou, T. Hanke, E. Thoutenhoofd and I. Zwitserlood. Parijs: ELDA, pp. 39–43.

De Quadros, Ronice Müller. 1995. As Categorias Vazias Pronominais. Master dissertation, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Brazil.

De Quadros, Ronice Müller. 2017. *Língua de herança: Língua brasileira de sinais*. Porto Alegre: Penso Editora.

De Quadros, Ronice Müller, Diane Lillo-Martin, and Karen Emmorey. 2016a. As línguas de bilíngues bimodais. *Revista de Estudos Linguísticos da Univerdade do Porto* 11: 139–60.

de Quadros, Ronice Müller, Diane Lillo-Martin, and Deborah Chen Pichler. 2016b. Bimodal Bilingualism: Sign Language and Spoken Language. In *The Oxford Handbook of Deaf Studies in Language*. Edited by Marc Marschark and Patricia Elisabeth Spencer. Oxford: Oxford University Press, pp. 181–96.

Döpke, Susanne. 1992. *One Parent—One Language: An Interactional Approach*. Amsterdam: John Benjamins Publishing Company.

Emmorey, Karen. 2003. *Perspectives on Classifier Constructions in Sign Languages*. Mahwah: Lawrence Erlbaum Associates.

Emmorey, Karen, Helsa B. Borinstein, Robin Thompson, and Tamar H. Gollan. 2008. Bimodal bilingualism. *Bilingualism: Language and Cognition* 11: 43–61. [CrossRef] [PubMed]

Fishman, Joshua. 2001. 300-plus years of heritage language education in the United States. In *Heritage Languages in America: Preserving a National Resource*. Edited by Joy Kreeft Peyton, Donald A. Ranard and Scott McGinnis. Washington, DC and McHenry: Center for Applied Linguistics and Delta Systems, pp. 81–89.

Johnston, Trevor. 1991. Transcription and glossing of sign language texts: examples from AUSLAN (Australian Sign Language). *International Journal of Sign Linguistics* 2: 3–28.

Kanto, Laura, Kerttu Huttunen, and Marja-Leena Laakso. 2013. Relationship between the linguistic environments and early bilingual language development of hearing children in deaf-parented families. *Journal of Deaf Studies and Deaf Education* 18: 242–60. [CrossRef] [PubMed]

Karnopp, Lodenir Becker. 1994. Aquisição do Parâmetro Configuração de Mão dos Sinais da LIBRAS: estudo sobre quatro crianças surdas filhas de pais surdos. Ph.D. dissertation, Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Brazil.

Kondo-Brown, Kimi, ed. 2006. *Heritage Language Development: Focus on East Asian Immigrants*. Amsterdam: John Benjamins.

Lanza, Elisabeth. 1997. *Language Mixing in Infant Bilingualism*. New York: Oxford University Press.

Lillo-Martin, Diane, Ronice Müller de Quadros, Elena Koulidobrova, and Deborah Chen Pichler. 2010. Bimodal bilingual cross-language influence in unexpected domains. In *Language Acquisition and Development: Proceedings of GALA 2009*. Edited by João Costa, Ana Castro, Maria Lobo and Fernanda Pratas. Newcastle upon Tyne: Cambridge Scholars Press, pp. 264–75.

Lillo-Martin, Diane, Ronice Müller de Quadros, Deborah Chen Pichler, and Zoe Fieldsteel. 2014. Language choice in bimodal bilingual development. *Frontiers in Psychology* 5: 1163. [CrossRef] [PubMed]

Lillo-Martin, Diane, Ronice Müller de Quadros, and Deborah Chen Pichler. 2016. The development of bimodal bilingualism: Implications for linguistic theory. *Linguistic Approaches to Bilingualism* 6: 719–55. [CrossRef] [PubMed]

Montrul, Silvina. 2011. Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition* 33: 163–92. [CrossRef]

Palmer, Jeffrey Levi. 2015. ASL Word Order Development in Bimodal Bilingual Children: Early Syntax of Hearing and Cochlear-Implanted Deaf Children From Signing Families. Ph.D. dissertation, Gallaudet University, Washington, DC, USA.

Petroj, Vanessa, Katelyn Guerrera, and Kathryn Davidson. 2014. ASL Dominant Code Blending in the Whispering of Bimodal Bilingual Children. In *BUCLD 38: Proceedings of the 38th annual Boston University Conference on Language Development*. Edited by Will Orman and Matthew James Valleau. Somerville: Cascadilla Press, vol. 2, pp. 319–30.

Peyton, Joy Kreeft, Donald A. Ranard, and Scott McGinnis, eds. 2001. *Heritage Languages in America: Preserving a National Resource*. Washington, DC and McHenry: Center for Applied Linguistics and Delta Systems.

Pichler, Deborah C., Julie Hochgesang, Diane Lillo-Martin, and Ronice Müller de Quadros. 2010. Conventions for sign and speech transcription in child bimodal bilingual corpora in ELAN. *Language, Interaction and Acquisition* 1: 11–40. [CrossRef]

Pichler, Deborah C., James Lee, and Diane Lillo-Martin. 2014. Language development in ASL-English bimodal bilinguals. In *Multilingual Aspects of Signed Language Communication and Disorder*. Edited by David Quinto-Pozos. Bristol: Multilingual Matters, pp. 235–60.

Pichler, Deborah C., Wanette Reynolds, Jeffrey Levi Palmer, Ronice Müller de Quadros, Viola Laura Kozak, and Diane Lillo-Martin. 2017. Heritage signers: Bimodal bilingual children from deaf families. In *Language Acquisition at the Interfaces: Proceedings of GALA 2015*. Edited by Jiyoung Choi, Hamida Demirdache, Oana Lungu and Laurence Voeltzel. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 247–69.

Pizer, Ginger Bianca. 2008. Sign and Speech in Family Interaction: Language Choices of Deaf Parents and their Hearing Children. Ph.D. dissertation, The University of Texas at Austin, Austin, TX, USA.

Polinsky, Maria. 2005. Word class distinctions in an incomplete grammar. In *Perspectives on Language and Language Development*. Edited by Maria Polinsky, Dorit Ravid and Hava Bat-Zeev Shyldkrot. Dordrecht: Kluwer Academic Press, pp. 419–36.

Polinsky, Maria. 2006. Incomplete acquisition: American Russian. *Journal of Slavic Linguistics* 14: 191–262.

Polinsky, Maria. 2008. Gender under Incomplete Acquisition: Heritage Speakers' Knowledge of Noun Categorization. *Heritage Language Journal* 6: 40–71.

Pyers, Jennie E., and Karen Emmorey. 2008. The face of bimodal bilingualism: Grammatical markers in American Sign Language are produced when bilinguals speak to English monolinguals. *Psychological Science* 19: 531–35. [CrossRef] [PubMed]

Reynolds, Wanette. 2016. Early Bimodal Bilingual Development of ASL Narrative Referent Cohesion: Using a Heritage Language Framework. Ph.D. dissertation, Gallaudet University, Washington, DC, USA.

Scontras, Gregory, Zuzanna Fuchs, and Maria Polinsky. 2015. Heritage language and linguistic theory. *Frontiers in Psychology* 6: 1545. [CrossRef] [PubMed]

van den Bogaerde, Beppie, and Anne E. Baker. 2009. Bimodal language acquisition in Kodas. In *Hearing, Mother-Father Deaf: Hearing People in Deaf Families*. Edited by Michele Bishop and Sherry L. Hicks. Washington, DC: Gallaudet University Press, pp. 99–131.

languages

MDPI

Article

Mood Selection in Relative Clauses by French–Spanish Bilinguals: Contrasts and Similarities between L2 and Heritage Speakers

Anahí Alba de la Fuente [1,*], Maura Cruz Enríquez [1] and Hugues Lacroix [1,2]

[1] Département de Littératures et de Langues du Monde, Université de Montréal, 3150 rue Jean-Brillant, Montréal, QC H3C 3J7, Canada; maura.cruz.enriquez@umontreal.ca (M.C.E.); hugues.lacroix@umontreal.ca (H.L.)

[2] IKER-UMR-5478 Laboratory, Université Bordeaux-Montaigne, 19 Esplanade des Antilles, 33607 Pessac, France

* Correspondence: anahi.alba.de.la.fuente@umontreal.ca; Tel.: +1-514-343-2267

Received: 20 February 2018; Accepted: 31 July 2018; Published: 6 August 2018

Abstract: In this paper, we explore three issues related to the acquisition of mood selection in Spanish relative clauses by second language (L2) and heritage (HL) speakers of Spanish: (1) whether HL speakers are more native-like than L2 learners; (2) whether the speakers' performance differs depending on task modality (written vs. oral), since HL speakers are known to perform better in oral tasks and L2 learners tend to do better in written tasks; and (3) whether knowledge of French as an L1/dominant language (DL) has an impact on the acquisition of Spanish subjunctive, since both languages include this mood in their grammars, but it is used more productively in Spanish. Results from a sentence combination felicity task (SCFT) in Spanish—in written and oral forms—and a written elicited production task (EPT) in French, administered to advanced L2 and HL speakers of Spanish whose L1/DL is French and two monolingual (Spanish and French) control groups, revealed that L2 learners pattern more closely with the control group than HL speakers in the SCFT, both written and orally. In the EPT, all bilingual speakers display higher levels of subjunctive use than the control group, showing a potential influence from the L2/weaker language on the L1/DL.

Keywords: heritage language (HL) acquisition; mood; L2 acquisition; Spanish relative clauses; task effects; oral vs. written production

1. Introduction

The contrast between the indicative and subjunctive moods in Spanish relative clauses has garnered significant attention in the fields of First Language (L1) (Blake 1983; Pérez-Leroux 1998), Second Language (L2) (Borgonovo et al. 2015), and Heritage Language (HL) (Montrul 2007; Montrul and Perpiñán 2011) Acquisition. With respect to L1, it has been found that, even though children acquire Spanish subjunctive morphology early in the language acquisition process, they only master mood selection in relative clauses around the age of 12 (Blake 1983), showing a correlation between cognitive and linguistic development (Pérez-Leroux 1998). This contrast is also lost early in attrition (Lipski 1993; Lynch 1999; Merino 1983; Silva-Corvalán 1994, 2003; Zentella 1997) and is subject to variation among native speakers (Blake 1983; Koenig 2016; Murillo 2000). Previous studies comparing L2 and HL speakers, have found that HL speakers, who acquire the language from birth, are not necessarily more native-like than L2 learners when it comes to structures that are acquired late, as is the case with mood selection in Spanish relative clauses (Montrul and Perpiñán 2011). Furthermore, previous research has also identified that, in experimental research, some of the differences between

these two groups of speakers may be task-induced. Specifically, L2 learners, who have received more instruction in the classroom, tend to obtain better results in tests that maximize explicit or metalinguistic knowledge (e.g., written tasks), whereas HL speakers, who typically acquire their heritage language at home, are more accurate in tests that measure implicit knowledge (e.g., oral tasks) (Bowles 2011; Montrul 2012; Montrul et al. 2008). This finding will be discussed in more detail in Section 2.2.

In turn, tapping into language typology, it has been discussed whether the presence of the subjunctive mood in the L1 represents an advantage in the L2 acquisition of mood selection in Spanish relative clauses. Specifically, when comparing L1 French (L1FR) with L1 English (L1EN) speakers learning Spanish as an L2, it has been found that L1FR speakers acquire the indicative-subjunctive contrast faster and pattern more closely with Spanish native speakers than L1EN speakers (Borgonovo et al. 2008; Boudreau 2007). However, as Boudreau (2007) claims, differences could be, again, task-induced, since the participants in her study performed different tasks. Namely, the L1FR group completed a written sentence combination felicity task (Boudreau 2007), while L1EN participants completed an appropriateness judgment task (Borgonovo et al. 2008). When comparing L1EN learners of Spanish with French as a third language (L3) with L1EN learners with no knowledge of French, Restorick Elordi (2012) found indications of positive cross-linguistic influence from the L3 to the L2, suggesting that L2 learners are at an advantage acquiring an interface phenomenon if they are also learning an L3 that presents the same phenomenon. However, it is important to take into account that, even though both French and Spanish include the subjunctive mood in their grammars, subjunctive is very productive in Spanish, whereas French is currently in the process of losing this mood (Menanteau 1986). Section 2.1 includes further details on the use of the subjunctive mood in French as well as in Spanish.

In order to further explore these issues, in this study we analyse mood selection in Spanish relative clauses in two groups of speakers, HL speakers and L2 learners, whose L1/dominant language (DL) is French. In addition, in order to examine possible task effects, depending on the modality of the task, we compare results from a written and an oral sentence combination felicity task (SCFT). Finally, to explore whether knowledge of French as an L1/DL has an impact on the acquisition of Spanish subjunctive, we compare participants' performances in both SCFTs with the results of a written elicited production task (EPT) in French.

The rest of the paper is organized as follows: Section 2 provides further details about mood selection in Spanish and French, and discusses previous findings regarding the acquisition of this phenomenon, with special attention to studies that explored how task modality may have yielded differences in performance; Section 3 presents the methodology of our study, including our research questions and hypotheses, participant information and materials; Section 4 includes the results and discussion, and, finally, a brief conclusion of our findings is presented in Section 5.

2. Mood Selection in Spanish Relative Clauses

2.1. Mood Selection in Spanish and French Relative Clauses

In Spanish relative clauses, the use of the subjunctive is syntactically optional, since these sentences are grammatical regardless of whether the subordinate verb is conjugated in the indicative or the subjunctive (Pérez-Leroux 1998).[1] In these sentences, mood selection is guided by the interpretation of the Determiner Phrase (DP) modified by the relative, which, in turn, depends on the [+/− existential] status of the referent. As such, as we can see in (1), existential referents trigger the use of the indicative, whereas non-existential referents require the use of the subjunctive.

[1] The examples in (1) are taken from Blake (1983). SBJV stands for subjunctive, IND for indicative.

1. a. *Busco* *una mujer que* *tenga dinero.* [−existential]
   Look.for-PRS.1SG a woman that has-SBJV money.
   'I'm looking for a woman with money.'

   b. *Busco* *a un hombre que* *es rico.* [+existential]
   Look.for-PRS.1SG a man that is-IND rich.
   'I'm looking for a man who is rich.'

As Blake (1983, p. 22) states, "the existence of the referent is frequently dependent on the individual's own subjective interpretation," which necessarily leads to variation in mood selection, both across and within individuals. However, as the author further explains, "the data show that adults and children concur substantially in their judgments of the existential status of the referent." In turn, in addition to the [+/− existential] status of the referent, we find certain contexts and operators in the language that also allow for the non-existential interpretation of the antecedent DP. In (2) and (3), we can see the contrast between a context that rejects the non-specific interpretation of the DP (2) and a context that allows for it (3).

2. * *Tengo* *un libro que* *me interese.*
   have-PRS.1SG a book that me interests-SBJV
   'I have a book I might be interested in.'

3. *Puedo* *tener* *un libro que* *me interese.*
   Can-PRS.1SG have a book that me interests-SBJV
   'I may have a book I might be interested in.'

In (2), the choice of a lexical verb such as *tener* ('to have'), which is not a modal operator, prevents the non-specific interpretation of the referent, in this case *the book*, which results in an inappropriate use of the subjunctive. In contrast, the presence of a modal verb in (3) opens up the possibility of a non-specific interpretation of the referent and, therefore, the use of the subjunctive is warranted in this sentence. Besides modal verbs, like *poder* ('can') (3), we find other operators that induce the non-specific interpretation of the indefinite DP in Spanish: negation, interrogation, future tense, strong intensional predicates and imperatives (Bosque 1999, p. 1). For the purpose of this study we focus on two of them, namely, future (4) and intensional predicates, that is, verbs with prospective lexical information (*buscar* ('to look for'), *aceptar* ('to accept'), *pedir* ('to ask for'), *esperar* ('to hope'), *querer* ('to want'), *intentar* ('to try'), *planear* ('to plan'), *perseguir* ('to pursue'), *merecer* ('to deserve'), *necesitar* ('to need'), etc.) (5). The future is the context that has the closest association with the indicative mood, while intensional verbs are more related to the subjunctive, as shown in previous studies (Boudreau 2007; Restorick Elordi 2012; among others).

4. *Contrataremos* *una persona* *que* *sepa* *idiomas.*
   hire-FUT.1PL a person that knows-SBJV languages
   'We will hire a person who can speak several languages.'

5. *Busco* *un inversor que* *tenga dinero.*
   Look.for-PRS.1SG an investor that has-SBJV money
   'I am looking for an investor that has money.'

The French verbal system, on the other hand, also presents this modal distinction in relative clauses, but is much less conservative than Spanish in its implementation. A survey conducted by Menanteau (1986) revealed that about 75% of the native speakers who participated in the study considered that the choice between subjunctive and "other mood" in sentences such as those presented in (6) did not necessarily trigger a difference in meaning.

6. a. *Je cherche* *quelqu'un qui* *soit parisien.*
I look.for-PRS.1SG someone that is-SBJV parisian
'I am looking for someone that is Parisian.'

b. *Je cherche* *quelqu'un qui* *est parisien.*
I look.for-PRS.1SG someone that is-IND parisian
'I am looking for someone that is Parisian.'

From these results, Menanteau (1986, p. 73) concluded that, in French relative clauses, the distinction between the two moods is no longer semantically relevant, and that the subjunctive is a free variant used to mark membership of, or association with, a more prestigious group. This loss of the subjunctive in French relative clauses has also been reported in other studies (Boudreau 2007; Lareau 2008).

2.2. Studying Mood Selection in L2 and Heritage Language Speakers: A Word on Task Effects

Previous studies have identified that, in experimental research, some of the differences between L2 and HL speakers may be task-induced. As mentioned in Section 1, L2 learners who have received more instruction in the classroom tend to obtain better results in tests that maximize explicit or metalinguistic knowledge (e.g., written tasks), whereas HL speakers, who typically acquire their heritage language at home, are more accurate in tests that measure implicit knowledge (e.g., oral tasks) (Bowles 2011; Montrul 2012; Montrul et al. 2008).

More specifically, Montrul et al. (2008) compared HL and L2 learners of different levels of Spanish proficiency, ranging from low to advanced. In their study, mastery of gender agreement was explored through three different tasks: written comprehension, written recognition, and oral production. Their results showed that HL speakers performed better in the oral task, while L2 learners outperformed HL speakers in both written tasks. The authors suggest that these differences are related to the type of knowledge that each task taps into—namely, the oral task is "more representative of fast, implicit and automatically processed knowledge (typically acquired early in childhood)," while the written tasks reflect ability with "metalinguistic, explicit knowledge (typically acquired later)" (Montrul et al. 2008, p. 541). According to Montrul et al. (2008, and following Paradis 2004), L2 learners use their explicit knowledge in the written tasks to compensate for their lack of implicit knowledge. However, in the oral task, they are unable to use the same strategy as they do not have enough time to do so.

In the same vein, using a battery of five tests developed by Ellis (2005) to tease apart implicit and explicit knowledge in L2 learners, Bowles (2011) explored the differences between native speakers of Spanish and HL speakers and L2 learners. Three of the five tasks relied on implicit knowledge—an oral imitation task, an oral narration task and a timed grammaticality judgment task (GJT)—while the other two, a metalinguistic awareness task and an untimed GJT, were designed to tap explicit knowledge. The native group scored close to the maximum in all tasks, with the exception of the metalinguistic awareness task, which indicates that, to a certain extent, monolingual speakers of a given language do not need metalinguistic awareness in their native language in order to master it. As expected, HL speakers scored significantly higher than L2 learners in the three tasks that relied on implicit knowledge, while L2 learners outperformed HL speakers in the metalinguistic awareness task. More specifically, the highest score for the HL group corresponded to the oral narration task, which can be considered the most implicit task in the study.

As shown in these studies, the degree of explicitness and task modality (oral vs. written) seem to have an impact on the performance of the different groups of speakers (HL vs. L2). Modality seems to be closely related to the degree of explicitness of a task, oral tasks being more implicit than written tasks. However, even within each mode, we find different degrees of explicitness. For example, an oral narration task is more implicit than an oral repetition task, which, in turn, is more explicit than an oral GJT. This gradation also applies to written tasks. Montrul and Perpiñán (2011) used two different kinds of written tasks—two elicited morphology recognition tasks and two sentence conjunction judgment tasks—to test knowledge of tense aspect and mood morphology in HL speakers and L2 learners.

Results show that, on the one hand, the L2 learners were more accurate than the HL speakers in the morphology recognition tasks, which benefit from active use of metalinguistic knowledge. On the other hand, HL speakers were more target-like in the conjunction judgment task, which is a more implicit task. Thus, we can see that the degree of explicitness of the task also varies within a given mode and that different tasks may influence results and trigger differences across populations—in this case, HL speakers and L2 learners—depending on their degree of explicitness, without altering the task mode.

In turn, van Osch and Sleeman (2016) also explored the differences between LH and L2 learners of Spanish in their use of the indicative and subjunctive mood, this time in the Netherlands, with speakers who had Dutch as their L1 or DL. Participants were administered an oral elicited production task in which they had to complete each sentence with a verb conjugated in the appropriate mood, and their results were compared to those from a GJT that participants had previously completed. In contrast with previous findings, which come mainly from studies carried out in the U.S. with L1 English speakers, Dutch HL speakers did not perform better in the oral task than in the GJT. Van Osch and Sleeman suggest that these results may be accounted for by the differences in the social contexts of the Spanish HL population in the U.S., on the one hand, and in the Netherlands on the other hand. Due to the smaller size of Hispanic communities in the Netherlands, HL speakers have fewer opportunities to speak their heritage language outside of their home. Additionally, the multilingual nature of their education system may have enhanced the metalinguistic knowledge of both HL and L2 speakers on a more general level (van Osch and Sleeman 2016, p. 15).

3. Our Study

In order to gain further insight into these issues, we present a study that analyses mood selection in relative clauses by two types of French–Spanish bilinguals, namely L2 learners and HL speakers. Specifically, we presented participants with a sentence combination felicity task (SCFT) in Spanish and a written elicited production task (EPT) in French. As discussed in the previous section, it has been argued that some of the differences between HL and L2 speakers in previous research may have been task-induced. In an attempt to further explore this issue, we administered our SCFT in both written and in oral form. In contrast with previous research using different types of tasks—with varying degrees of explicitness—and, often, different modalities as well (some tasks were presented orally and others were presented in written form), this study focuses on the issue of modality by keeping the task constant and only altering the mode in which the SCFT stimuli were presented to our participants. In the remainder of this section, we present our research questions and hypotheses, as well as detailed information about our participants and the methods and materials of our experimental tasks.

3.1. Research Questions

As discussed in Section 1, one of our research goals is to compare L2 and HL speakers whose L1/dominant language (DL) is French by examining mood selection in Spanish relative clauses. We also intend to examine possible task effects favouring either L2 or HL speakers depending on the modality of the task (written vs. oral). Finally, we aim to analyse whether knowledge of French as an L1/DL has an impact on the acquisition of Spanish subjunctive. In order to meet these objectives, we set out to answer the following research questions (RQ):

RQ 1. *Will the two experimental groups (HLSP and L2SP) differ from each other and from the Spanish native control group in terms of their mood selection patterns in Spanish?*

Based on previous studies that examine and compare these two groups, we predict that HLSP and L2SP will differ from the native control group. Specifically, we expect differences between these two groups both overall and with respect to their performance in the different types of tasks (see RQ 3, below).

RQ 2. *Will there be a positive correlation between the use of the subjunctive mood in the production task (Task 2) and the results of the sentence combination tasks (Tasks 1 and 3)? In other words, will there be an observable influence of the L1 on the L2, or vice versa?*

Based on previous studies showing positive cross-linguistic influence, we predict that there will be a positive correlation between an increased use of the subjunctive in Task 2 and higher acceptance rates of clauses using the subjunctive mood in Task 1. Furthermore, since the subjunctive is more productive in Spanish than in French (see Section 2.1), experimental groups (L2SP, HLSP) may be more inclined to use the subjunctive mood in Task 2 than the control group (L1FR with no knowledge of Spanish), which would reveal an influence from Spanish in their L1/dominant language.

RQ 3. *Will L2SP and HLSP speakers perform differently in the oral and written tasks?*

Based on previous research showing differences in the performance of HL and L2 speakers depending on the modality of the task (see Section 2.2), we predict that L2SP speakers will perform better in the written task as it favours explicit knowledge (Task 1), while HLSP will do better in the oral task (Task 3) as it favours implicit knowledge.

3.2. Participants

Four groups of participants took part in this study: two experimental groups and two control groups (see Table 1, below). The experimental groups consisted of: French–Spanish bilinguals (HLSP) and L2 learners of Spanish whose L1 and dominant language was French (L2SP). All HLSP participants had acquired Spanish from birth and all L2SP participants began learning Spanish in a formal setting after puberty. None of the heritage participants were schooled in Spanish or had received formal Spanish instruction. Both experimental groups had an advanced level of proficiency in Spanish, as evidenced by the results of a language proficiency test (see Section 3.3., below). In addition, one group of native speakers of Spanish with little or no knowledge of French (L1SP) and one group of native speakers of French with little or no knowledge of Spanish (L1FR) served as a control. All participants gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Research of the faculty of Arts and Sciences (CERFAS) of the Université de Montréal (code: CERFAS-2014-15-049).

Table 1. Participant information for each task.[2]

Group	Task 1	Task 2	Task 3
L2SP	*n* = 10 [Median age = 27.7]	*n* = 18 [Median age = 31.5]	*n* = 12 [Median age = 31.25]
HLSP	*n* = 10 [Median age = 25.9]	*n* = 13 [Median age = 24.7]	*n* = 10 [Median age = 21.9]
L1FR		*n* = 14 [Median age = 40.2]	
L1SP	*n* = 11 [Median age = 41.5]		*n* = 21 [Median age = 37]

L2SP: L2 learners of Spanish whose L1 and dominant language is French; HLSP: French–Spanish bilinguals; L1FR: native speakers of French with little or no knowledge of Spanish; L1SP: native speakers of Spanish with little or no knowledge of French.

[2] Most participants from the experimental groups completed all three tasks. However, some participants only completed two out of the three tasks (either Tasks 1 and 2, or Tasks 2 and 3), hence the difference in the composition of groups from task to task.

Table 2 presents a summary of the characteristics of each experimental group, including the age of onset of acquisition for each of the two languages involved in the study as well as the type of input and acquisition setting in each case.

Table 2. Participant information based of Age of Acquisition (AoA) of Spanish and French, acquisition setting, and input mode.

Group	AoA of Spanish	AoA of French	Spanish Acquisition Setting	Spanish Input Mode
L2SP	>15 years	since birth	instructed (classroom)	written and aural (literacy)
HLSP	since birth	since birth (simultaneous)//4–7 years (sequential)	naturalistic (home)	aural

L2SP: L2 learners of Spanish whose L1 and dominant language is French; HLSP: French–Spanish bilinguals.

As indicated in Table 2, the HLSP speakers of our study presented two different profiles in terms of their age of onset of bilingualism: simultaneous, who were exposed to both French and Spanish from birth, and sequential, who were first exposed to French between the ages of four and seven. This difference in terms of age of onset of bilingualism was not expected to be a relevant variable, since this study focuses on a late acquisition phenomenon. However, in order to present a more complete picture of our heritage speakers, these two profiles were taken into account in the codification and analysis of our data even if the type of bilingualism (simultaneous or sequential) was not part of the original experimental design. Relevant findings with respect to this variable (type of bilingualism) are discussed in detail in Section 4.

3.3. Experimental Design

All participants had to fill out a language background questionnaire in order to guarantee that the linguistic profiles of all participants matched the requirements of the study. All experimental groups also completed two Spanish proficiency tests. First, they completed a general proficiency test, based on the Diplomas de Español como Lengua Extranjera (DELE, Certificate of Spanish as a Foreign Language) exam, which has been widely used in recent second language acquisition (SLA) studies (Montrul 2007; Montrul et al. 2008; Montrul and Perpiñán 2011; among others). This test was followed by a morphology recognition task (Boudreau 2007; Slabakova and Montrul 2002; among others), which measured the participants' ability to recognize the indicative–subjunctive contrast.

In addition, three experimental tasks were administered: a sentence combination felicity task in Spanish (modified from Borgonovo et al. 2008; Boudreau 2007), both in written (Task 1) and in oral form (Task 3), and an Elicited Production task in French (Task 2). Tasks 1 and 3 were completed in Spanish by all experimental groups and the Spanish control group, and Task 2 was completed in French by all experimental groups and the French control group.

3.3.1. Task 1: Sentence Combination Felicity Task—Written Version

The goals of this task, inspired by similar ones from previous studies (Borgonovo et al. 2008; Boudreau 2007; Montrul 2007; Montrul and Perpiñán 2011), were: (1) to analyse mood selection patterns in the Spanish grammars of HL and L2 speakers; and (2) to test whether participants were able to associate each mood (indicative or subjunctive) to a given interpretation of the antecedent DP (either specific or non-specific) (Borgonovo et al. 2008). The task included 54 pairs of sentences (juxtaposed, adversative, disjunctive, and subordinate), 36 of which were experimental items and the remaining 18 were distractors. The first sentence of the experimental items always contained a main clause containing a DP that was modified by a relative clause. The second sentence in the item contextualized the information of the first sentence and determined whether the DP was specific or not. An example with an intensional verb (*buscar* 'to look for') is shown in (7). The first sentence alone (7a) could be grammatical either with

the indicative or the subjunctive mood, but when combined with the sentence in (7b), only the indicative mood would be grammatical, since it calls for a specific interpretation of the DP.

7. a. *Buscamos un restaurante que no cuesta/cueste muy caro.*
   'We are looking for a restaurant that is-IND/SBJV not too expensive.'

   b. *Me han dicho que se llama "El Marino".*
   'I was told its name was "El marino".'

The task included three different contexts (future, intensional, and modal), with 12 items per context, and four conditions per context: Appropriate Indicative (AI), Inappropriate Indicative (II), Appropriate Subjunctive (AS), and Inappropriate Subjunctive (IS). The following examples present an appropriate subjunctive (AS) item with a modal context in (8), an inappropriate indicative (II) item with an intensional context in (9), and an inappropriate subjunctive (IS) item with a future context in (10). The complete list of experimental items from Task 1 is available in Appendix A.

8. a. *Debes leer un libro que te permita comprender el subjuntivo.*
   'You must read a book that allows-SBJV you to understand the subjunctive mood.'

   b. *En esta biblioteca seguro encontrarás alguno.*
   'You will probably find one in this library.'

9. a. * *Quiero una habitación que tiene mucha luz,*
   'I want a room that is-IND very bright.'

   b. *pero solo hay habitaciones oscuras en este edificio.*
   'but there are only dark rooms in this building.'

10. a. * *Veremos una película que esté nominada a un Oscar.*
    'We are going to see a movie that is-SBJV nominated for an Oscar.'

    b. *Se titula* Boyhood.
    'It's called *Boyhood*'.

Participants had to read the sentences and judge, in each case, whether they could have been uttered by a Spanish native speaker. To do so, they had to circle on a 5-point Likert scale whether they were "Sure that it could NOT have been uttered by a native speaker" (−2), "Sure that it could have been uttered by a native speaker" (+2), or any option in between. The middle point, which was marked as 0, corresponded to "I don't know" or "I am not sure". Participants were instructed to avoid 0 as much as possible. Additionally, following Alba de la Fuente (2012), if they assigned a negative score to an item (−2 or −1), they were further instructed to correct the part of the item that they considered problematic. This additional instruction allowed us to verify whether the negative judgments were really motivated by an incorrect use of the subjunctive/indicative mood or by other factors, such as lexical limitations (Alba de la Fuente 2012, p. 128).

3.3.2. Task 2: Elicited Production Task

In the elicited production task (Task 2), participants had to fill in the blank of a series of sentences with the correct tense—and mood—of the verb provided in parenthesis. The experimental items were a French translation of the items used in tasks 1 and 3. In order to facilitate the analysis, only verbs that have different forms in the indicative and the subjunctive were selected, which required a few minor modifications to some items. For example, the original test in Spanish included the verb *criticar* ('to criticize'), which was reworded as *faire une critique* ('to do/make a criticism'). Two examples of

Task 2 items are shown in (11) and (12). The complete list of experimental items from Task 2 is available in Appendix B.

11. a. *Tu dois lire un livre qui te _____ (permettre) de comprendre le subjonctif.*
    'You must read a book that _____ (allow) you to understand the subjunctive mood.'

    b. *Dans cette bibliothèque tu en trouveras sûrement un.*
    'You will probably find one in this library.'
12. a. *Nous allons voir un film qui _____ (être) nominé pour un Oscar.*
    'We are going to see a movie that _____ (is) nominated for an Oscar.'

    b. *Il s'intitule* Boyhood.
    'It's called *Boyhood*'.

The goal of this task, which is a modified version of the one used by Boudreau (2007), was to gain further insight on the distribution of indicative and subjunctive in these structures in French and to explore whether the participants' responses were comparable to the Spanish distribution of the two moods. Thus, unlike Boudreau (2007), this task was administered to all our experimental groups (HL and L2) and not just the control group (L1FR).

3.3.3. Task 3: Sentence Combination Felicity Task—Oral Version

This task was identical to Task 1, except for the way in which stimuli were presented to the participants. While Task 1 was presented in written form, both the instructions and stimuli were presented orally in Task 3. The purpose of this task was, first, to analyse mood selection patterns in the Spanish grammar of the speakers when speakers are constrained to rely more strongly on their implicit knowledge and, in combination with the results from Task 1, to explore whether the mode in which the tasks were presented—written or oral—had an effect on participants' responses. As mentioned in Section 1.1, written tasks allow participants to make use of their explicit knowledge and metalinguistic skills, as items can be read many times and participants can take more time to think about their responses. In turn, oral tasks are considered to make participants depend more on implicit knowledge as they have to react faster and, normally, stimuli are not available to access on demand.

For the development of Task 3, the items from Task 1 were digitally recorded by a female native speaker of Spanish from Mexico. The recordings were then segmented by sentence, so that they could be presented one at a time within each item. Another speaker, a male near-native speaker of Spanish, was digitally recorded for the instructions. The introduction of two noticeably different voices was done in order to make sure that participants were able to clearly distinguish the instruction segments of the recording from the experimental items.

The task was administered through a web experiment platform developed using the jsPsych library (De Leeuw 2015) and the participants were instructed to use headphones and complete the entire experiment without interruption. In order to make sure that the participants understood the instructions and that there were no technical problems, five practice items were presented at the beginning of the task. For each item, the audio stimulus was played twice, after which a 5-point scale appeared on the screen. Participants had to indicate their judgment by clicking with their mouse on their selected answer (from −2 to 2). As in Task 1, they also had to correct what they considered to be the error in sentences that received a negative score. In these cases, the written item appeared on the screen and they had to correct the sentence using their keyboard. Again, this correction allowed us to verify that the negative judgments were really motivated by an incorrect use of the subjunctive/indicative mood. As was the case in Tasks 1 and 2, there were no time limitations on this task and participants could take as long as they needed to provide a response.

4. Results and Discussion

4.1. Sentence Combination Felicity Tasks (in Spanish)

4.1.1. Task 1—Written Version

In this task, participants had to rate the acceptability of a series of sentence pairings using a 5-point Likert scale, as indicated in Section 3.3.1. The results of this task reveal that, as expected, the native group rejected inappropriate uses of both moods and accepted their corresponding appropriate counterparts. The L2 group also presented this pattern of accepting appropriate conditions and rejecting inappropriate conditions, but at lower rates than the L1 control group. Also, the inappropriate subjunctive (IS) showed higher rejection than the inappropriate indicative (II) in this group, thus revealing more uncertainty in the II condition than in the IS condition. Finally, the HL group showed rejection of the IS condition, but not of the II condition.

Figure 1 shows the overall results of the three experimental groups.

The results of a repeated measures ANOVA reveal main effects for "mood" (($F(1, 28) = 5.541$, $p = 0.026$)), "appropriateness" ($F(1, 28) = 146.950$, $p < 0.001$) and "group" (($F(2, 28) = 4.380$, $p = 0.022$). The statistical analysis also revealed main interactions between "mood" and "group" ($F(2, 28) = 4.933$, $p = 0.015$), between "appropriateness" and "group" ($F(2, 28) = 12.928$, $p < 0.001$), between "mood" and "appropriateness" ($F(1, 28) = 29.797$, $p < 0.001$), and between "mood", "appropriateness" and "group" ($F(2, 28) = 3.955$, $p = 0.002$). Pairwise comparisons using the Bonferroni correction reveal significant differences between the L1SP and the HL groups ($p = 0.019$).

There was no main effect for "context" (intensional, future, modal) ($F(2, 56) = 2.876$, $p = 0.065$). A main interaction between "mood" and "context" was found ($F(2, 56) = 2.828$, $p < 0.001$), but none of the other interactions—either with "appropriateness" or with "group"—were significant and pairwise comparisons using the Bonferroni correction did not reveal any further statistically significant differences among the three contexts.

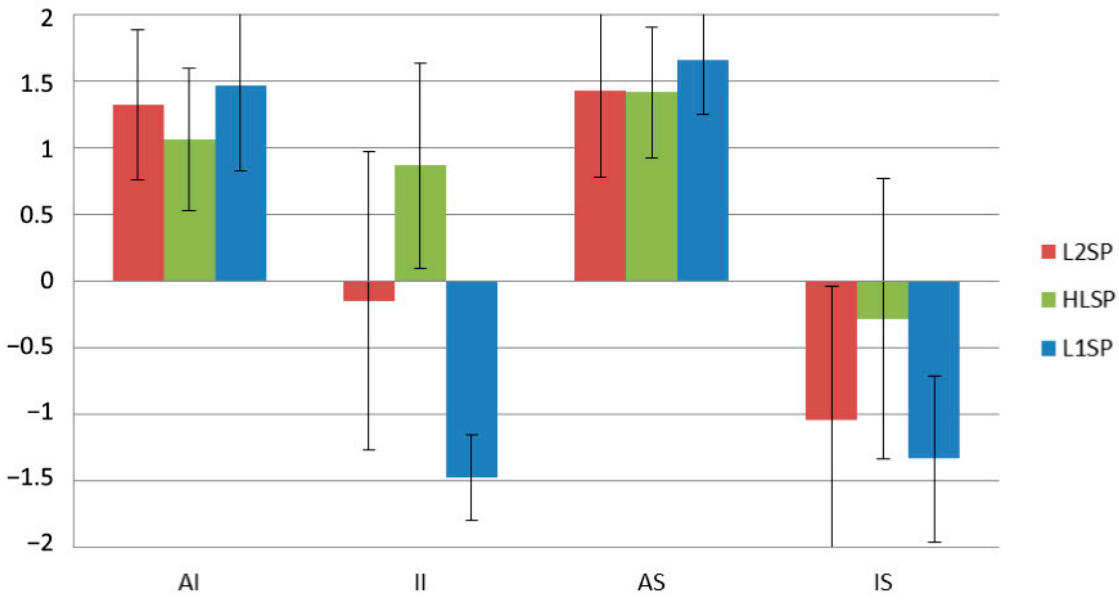


Figure 1. Overall results. Task 1—Sentence combination felicity task (written version). L2SP: L2 learners of Spanish whose L1 and dominant language is French; HLSP: French–Spanish bilinguals; L1SP: native speakers of Spanish with little or no knowledge of French. AI: appropriate indicative; II: inappropriate indicative; AS: appropriate subjunctive; IS: inappropriate subjunctive.

These results indicate that, overall, the appropriate (A) and inappropriate (I) conditions were treated differently, as were the two moods, and that the experimental groups behaved differently. Specifically, the HL group, but not the L2 group, differed significantly from the L1SP control. The different treatment of A and I conditions was expected and is apparent for all groups. However, and in contrast with the other two groups, the HL group failed to reject the II condition, thus indicating a general acceptance of the use of indicative in these relative clauses. Furthermore, the HL group tended to rate inappropriate uses of both moods more highly than both the L2SP and the L1SP groups. In terms of "context," participants did not treat the three contexts (intensional, future, mood) differently enough to reach statistical significance.

Finally, as the reader may recall, participants were asked to provide corrections for the items that were judged to be incorrect. All groups behaved very similarly in terms of their corrections and, as expected, the vast majority of the item modifications implied replacing the inappropriate mood with the appropriate one, thus validating the premise that, when participants assigned a negative score to a given item, they were indeed reacting to the inappropriateness of the verb mood in the relative clause. In a few cases, instead of replacing the inappropriate mood of the verb, participants modified another element in the sentence in an attempt to make it align with the [+/− existential] status of the referent, which also suggests that participants were reacting to the inappropriateness of the verb mood. Examples of this include the use of conditional in items with a non-existential referent, as shown in example (13) from Task 1, and clause reduction, avoiding the use of either indicative or subjunctive, as in example (14) from Task 3.

13. Original item (Participant HLSIM07)
    Compraré un traje que va bien con el tuyo, si encuentro alguno en esta tienda.
    'I will buy a suit that matches yours, if I find any at that store.'

    Correction
    <u>Compraría</u> un traje que va bien con el tuyo, si encuentro alguno en esta tienda.
    'I would buy a suit that matches yours, if I find any at that store'.
14. Original item (Participant L1SP02)
    En esta biblioteca puedes leer un libro que esté escrito en alemán. Hay muchos en el segundo piso.
    'In this library, you can read a book that is written in German. There are many on the second floor.'

    Correction
    En esta biblioteca puedes leer <u>libros escritos</u> en alemán. Hay muchos en el segundo piso.
    'In this library, you can read books written in German. There are many on the second floor.'

The few remaining cases either included modifications that were unrelated to the mood of the relative verbs (e.g., the addition of a comma or other punctuation marker, the replacement of one lexical item, etc.) or did not contain any modifications. The corrections provided by participants were essentially the same in Tasks 1 and 3, both in terms of the types of correction applied (mood changes, tense changes, clause reductions, no modification, etc.) and their distribution.

4.1.2. Task 3—Oral Version

As indicated in Section 3.3, above, Tasks 1 and 3 were identical, except for the way in which stimuli were presented to the participants. Task 1 was presented in written form and Task 3 was presented orally. The purpose of this was to explore whether the mode in which the tasks were presented—written or oral—had an effect on participants' responses. We first present the results of Task 3 and then compare them with the results of Task 1.

The results of a repeated measures ANOVA performed on these data reveal the main effects for "appropriateness" ($F(1, 40) = 63.641$, $p < 0.001$) and "group" ($F(2, 40) = 4.517$, $p = 0.017$). No effect was found for "mood" ($F(1, 40) = 2.157$, $p = 0.150$) or "context" ($F(4, 80) = 1.080$, $p = 0.344$). There were, however, significant interactions between "mood" and "group" ($F(2, 40) = 11.890$, $p < 0.001$), between

"mood" and "appropriateness" ($F(1, 40) = 10.938$, $p = 0.002$), and between "mood" and "context" ($F(2, 80) = 9.430$, $p < 0.001$). The interaction between "appropriateness" and "group" was also significant ($F(2, 40) = 14.650$, $p < 0.001$).

As in Task 1, pairwise comparisons using the Bonferroni correction reveal significant differences between the L1SP and the HL groups ($p = 0.014$).

As we can see in Figure 2, the L1SP group displays the expected acceptance and rejection pattern, which also matches the distribution found in Task 1. With respect to the other two groups, the HL group displays an overall acceptance of all conditions, albeit at varying rates, and the L2 group displays a rejection of inappropriate sentences, but only in the IS condition. Given the fact that no effect was found for "context" in either Task 1 or 3, this variable will not be further discussed.

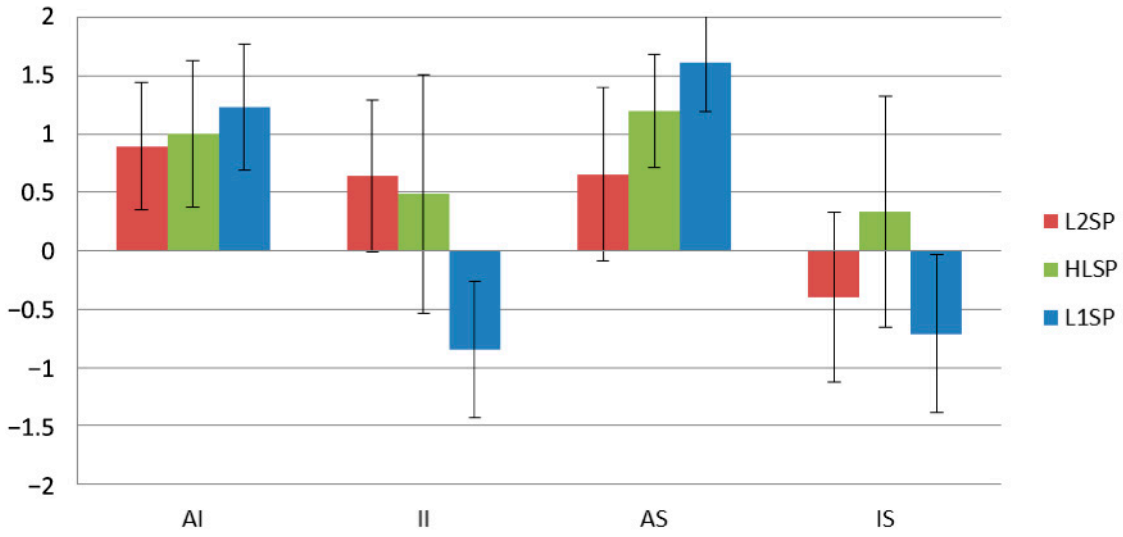
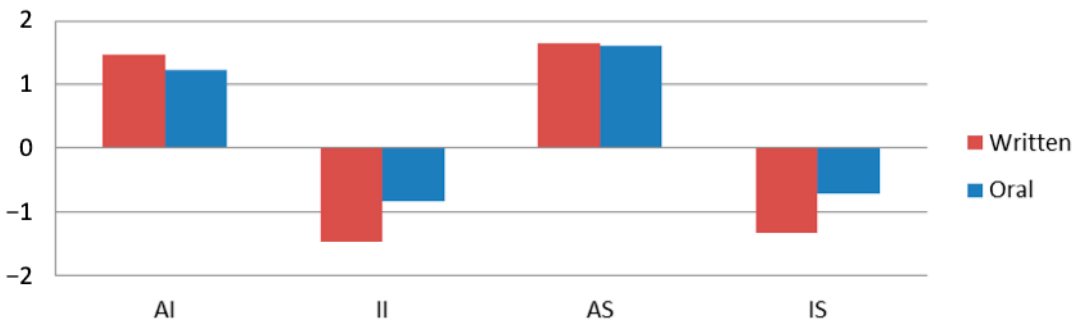


Figure 2. Overall results. Task 3—Sentence combination felicity task (oral version).

4.1.3. Task 1 (Written) vs. Task 3 (Oral)

In this section, we compare the results of Tasks 1 and 3 in order to explore whether the mode in which the task was presented (written or oral) had an effect on our participants' responses. Figure 3 presents the results of the two tasks broken down by group.



(a)

Figure 3. *Cont.*

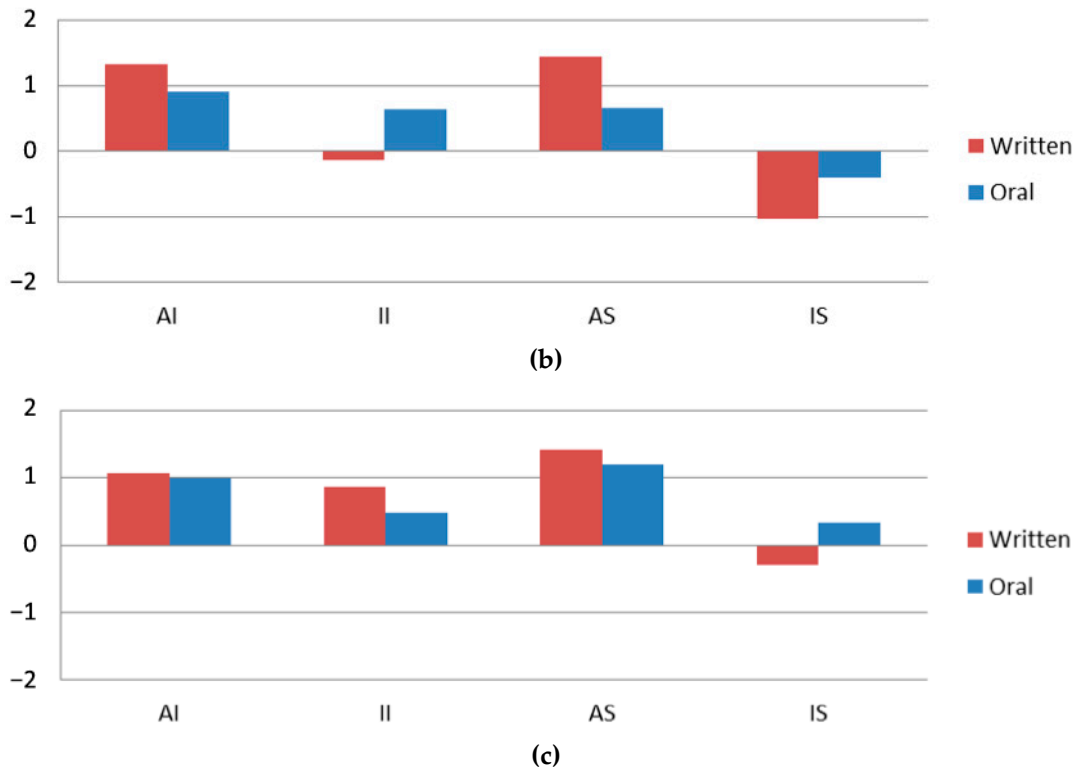


Figure 3. Task 1 vs. Task 3—Sentence combination felicity task (written vs. oral), results by group. (**a**) Task 1 vs. Task 3 (L1SP); (**b**) Task 1 vs. Task 3 (L2SP); (**c**) Task 1 vs. Task 3 (HLSP).

In general, we see that the rates were more polarized—i.e., higher acceptance and higher rejection—in the written task than in the oral task, but followed similar trends in all groups, with the exception of the II condition in the L2 group and the IS condition in the HL group. The results of a repeated measures ANOVA revealed main effects for "mood" ($F(1, 68) = 8.027$, $p = 0.006$), "appropriateness" ($F(1, 68) = 203.629$, $p < 0.001$) and "group" ($F(2, 68) = 9.375$, $p < 0.001$). The ANOVA also revealed several interactions between these factors. Namely, main interactions were found between "mood" and "group" ($F(2, 68) = 14.442$, $p < 0.001$), between "mood" and "appropriateness" ($F(1, 68) = 41.168$, $p < 0.001$), between "appropriateness" and "group" ($F(2, 68) = 24.656$, $p < 0.001$), and between "appropriateness", "mood" and "group" ($F(2, 68) = 6.947$, $p = 0.002$).

No overall effect was found for "task mode" ($F(1, 68) = 0.808$, $p = 0.372$), but significant interactions were found between "appropriateness" and "task mode" ($F(1, 68) = 10.495$, $p = 0.001$), and between "mood", "appropriateness", "group" and "task mode" ($F(2, 136) = 3.654$, $p = 0.031$). Pairwise comparisons using the Bonferroni correction reveal significant differences between the L1SP and the HL groups ($p < 0.001$) and a borderline significant difference between the HL and the L2 groups ($p = 0.049$).

Despite the absence of a main effect for task mode, in view of the presence of several significant interactions including this factor, separate repeated measures ANOVAs were performed on each group in order to further explore potential task effects in our data. The results of these analyses reveal no main effect for the L2SP group ($F(1, 20) = 0.048$, $p = 0.829$), or the HL group ($F(1, 18) = 0.000$, $p = 0.991$). In contrast, the L1SP group did show a main effect for task mode ($F(1, 30) = 7.804$, $p = 0.009$). As we can see in Figure 3, the L1SP group displays the expected acceptance/rejection patterns in both tasks. However, while the response rates for the A conditions remained similar in both tasks, there was stronger rejection for the I conditions in the written task than in the oral task, which results in an unexpected task effect for this group.

4.1.4. Some Remarks on the HL Group

Our HL group included a combination of simultaneous and sequential bilinguals. Both in the written and in the oral tasks, 40% of the HL participants were simultaneous bilinguals, and 60% were sequential bilinguals. In view of this distribution, two repeated measures ANOVAs were performed on the HL data, one for Task 1 and one for Task 3, in order to determine whether there were significant differences in the behaviour of these two sub-groups. The results of the ANOVA revealed borderline significant differences between the two subgroups in the written task ($F(1, 8) = 5.450$, $p = 0.048$), and no significant differences in the oral task ($F(1, 8) = 1.175$, $p = 0.310$). Figure 4 shows the results of the HL participants, broken down by subgroup, for both the written and the oral task.

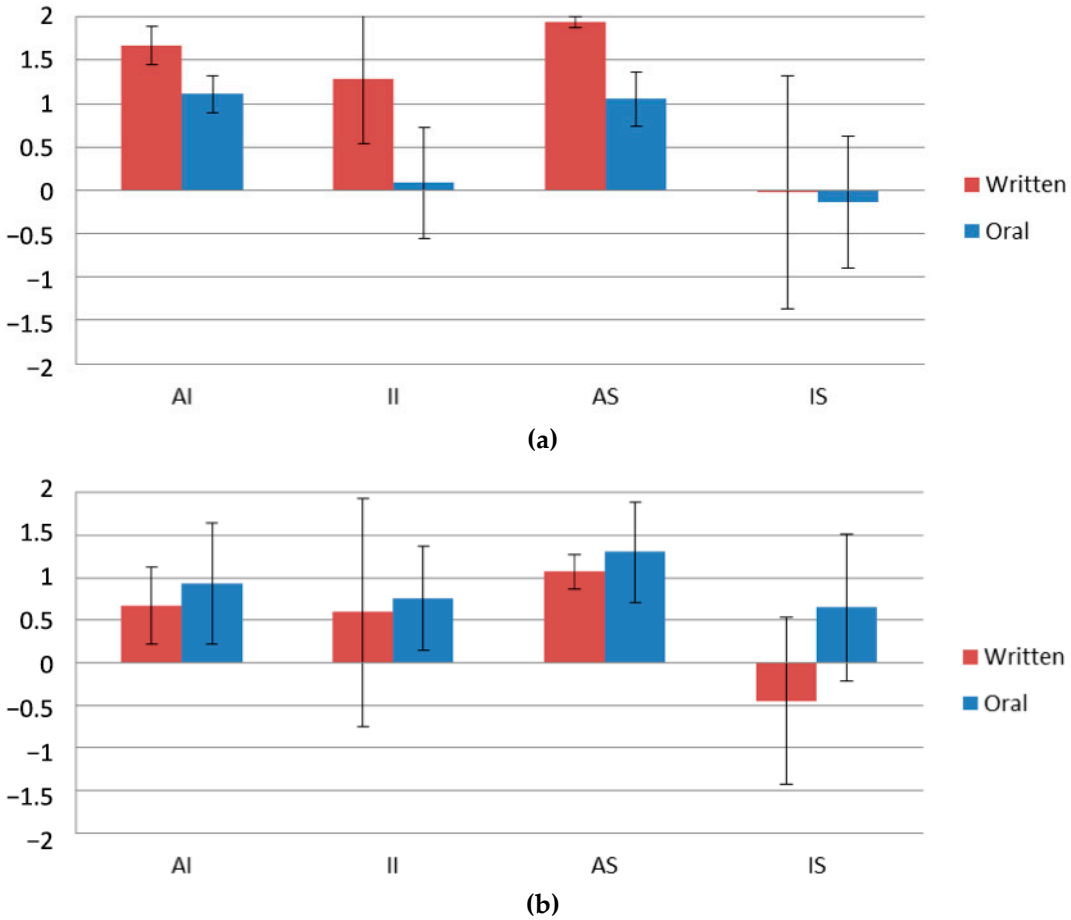


Figure 4. Sentence combination felicity task (written vs. oral), HL results, broken down by subgroup (HLSIM: simultaneous French–Spanish bilinguals, HLSEQ: sequential French–Spanish bilinguals). (**a**) Task 1 vs. Task 3 (HLSIM); (**b**) Task 1 vs. Task 3 (HLSEQ).

When comparing the two subgroups' performances in both tasks, we found no main effect for "type of bilingualism (group)" ($F(1, 16) = 0.600$, $p = 0.450$) or "task mode" ($F(1, 16) = 0.275$, $p = 0.607$). There was, however, a significant interaction between "type of bilingualism" and "task mode" ($F(1, 16) = 6.093$, $p = 0.030$), as well as between "appropriateness," "type of bilingualism," and "task mode" ($F(1, 16) = 0.153$, $p = 0.009$). These significant interactions, including "task mode" and "type of bilingualism," hint at the presence of a possible task effect that only seems to affect the simultaneous French–Spanish bilinguals (HLSIM) group. The responses from this group were most divergent from the other groups in the written task and closest in the oral task, thus showing an apparent disadvantage for HLSIM participants in the written task that disappears—or becomes less apparent—in the oral task, which is in line with previous findings. However, separate repeated

measures ANOVAs comparing the results from both tasks did not reveal significant effects for task mode for either sequential French–Spanish bilinguals (HLSEQ) ($F(1, 10) = 2.651$, $p = 0.135$) or HLSIM ($F(1, 6) = 2.737$, $p = 0.149$).

4.2. Sentence Completion Task (in French)—Task 2

As discussed in Section 3.3.2, the first goal of this task was to gain further insight on the distribution of indicative and subjunctive in native French and to obtain empirical evidence of the claim that subjunctive is being lost in these structures in French, as defended by authors such as Menanteau (1986). The second goal was to explore whether mood preferences were comparable across groups or whether knowledge of Spanish may influence participants' responses. Thus, in contrast with Boudreau (2007), this task was administered not only to the native control group but to the experimental groups (L2SP and HL) as well.

In this task, participants were presented with a sentence with a blank space, which they had to fill with the correct form of the verb provided in brackets. Participants' responses were initially coded based on both tense and mood. Then, in order to carry out the analysis, all responses were recoded as either "indicative," "subjunctive," or "other."[3] Figure 5 presents the overall results of Task 2, arranged by group and by response, including percentages for each response type.

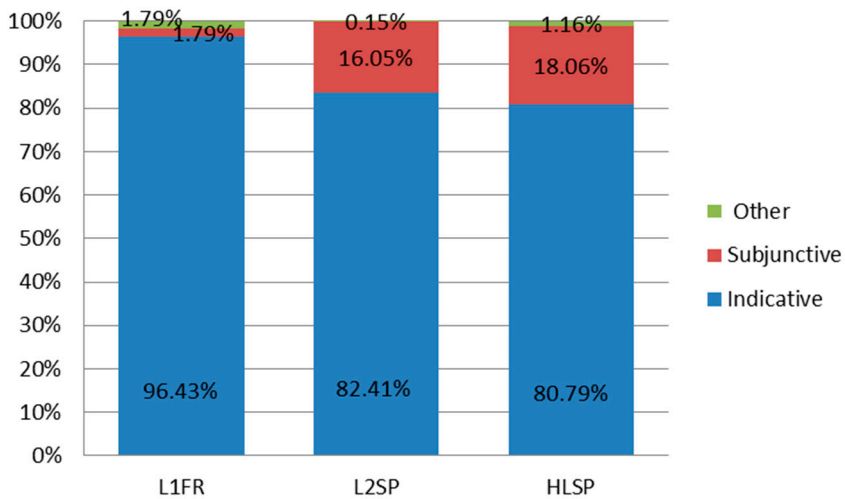


Figure 5. Sentence completion task in French (Task 2)—Overall results.

As expected, the presence of the subjunctive is practically absent in the L1FR control data, as participants provided a verb conjugated in the indicative 96.43% of the time. In contrast, the subjunctive mood was present in varying degrees in both experimental groups. Specifically, the rates of subjunctive responses were 16.05% for the L2SP group and 18.06% for the LH group. The indicative remained the most frequent mood chosen by all groups and was used at least 80% of the time. However, based on the assumption that French is in the process of losing the indicative–subjunctive contrast in these structures, and in view of the results from the L1FR control group, the presence of subjunctive

---

[3] The "other" category included the use of non-finite forms and conditional, which was the most frequent response in this category. An example of a response using conditional is presented in (i).

i. a. *Tu dois lire un livre qui te permettrait de comprendre le subjonctif.*
'You must read a book that would allow you to understand subjunctive mood.'

b. *Dans cette bibliothèque tu en trouveras sûrement un.*
'In this library, you will probably find one.' [Participant L1FR06]

responses in the other two groups was surprising, especially in the case of the L2SP group, whose L1 and dominant language was French—like the L1FR control group—and Spanish was only acquired after puberty.

The experimental design of this task included items whose Spanish equivalent would require the use of the indicative—i.e., the "indicative" condition—and items whose Spanish equivalent would require the use of the subjunctive—i.e., the "subjunctive" condition. In turn, based on this parameter, responses were also coded as "matching" if the mood chosen corresponded to the equivalent expected mood in the Spanish translation of the sentence (e.g., whether the subjunctive was used in a "subjunctive" condition), and "non-matching" if it did not (e.g., whether the indicative was used in a "subjunctive" condition). Figure 6 shows the results of the three groups broken down by condition.

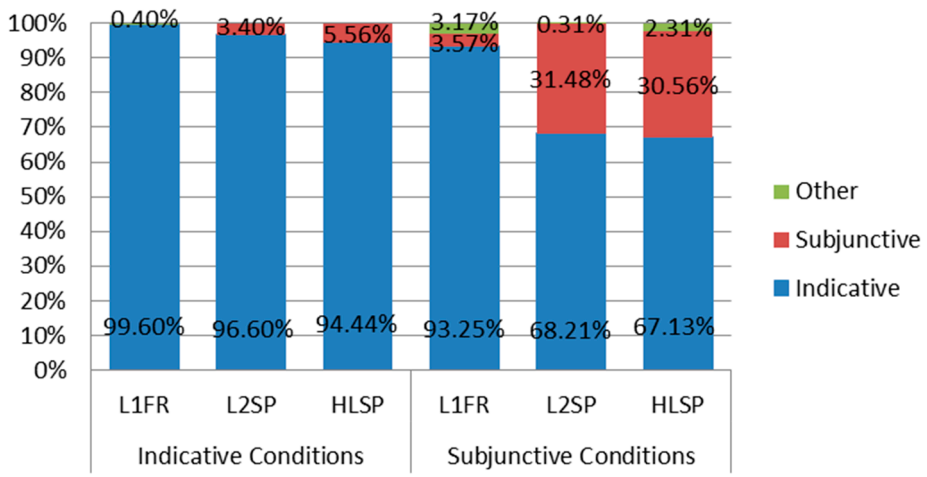


Figure 6. Overall results, broken down by condition. Task 2—Elicited production task.

As shown in Figure 6, the overwhelming majority of subjunctive responses correspond to subjunctive conditions. This suggests that, also in their dominant language, the use of the subjunctive by L2SP and HL participants matches the equivalent uses in Spanish. In other words, even though the indicative mood was overwhelmingly preferred by participants in all three groups, when L2SP and HL participants chose to use a subjunctive verb, they did so in conditions whose Spanish equivalent would also require the use of the subjunctive (i.e., "matching").

A generalized linear mixed model was performed on the data using "group", "condition" and "context" as predictors and "subjunctive use" as the outcome ("indicative" was coded as 0, "subjunctive" as (1). Both "participant" and "item" were included as random effects. The results of the statistical model revealed significant differences for all the predictors, the strongest one being "group", followed by "condition" and, then, by "context." Specifically, the differences between the L1FR control group and both the L2SP ($z = 4.007$, $p < 0.001$) and the HL ($z = 3.901$, $p < 0.001$) groups were significant, as were the differences between the indicative and subjunctive conditions ($z = 7.863$, $p < 0.001$). Finally, there were also significant differences between the "future" context and the "intensional" ($z = 4.046$, $p < 0.001$) and "modal" ($z = 3.285$, $p = 0.001$) contexts. Specifically, the results of the mixed model indicate that "future" favoured the use of the indicative, whereas "intensional" and "modal" favoured the use of the subjunctive. Interactions between the three predictors ("group," "condition," and "context") were also tested, but were removed from the model as none of them were significant.

A second mixed model was performed with the two HL sub-groups (HLSIM and HLSEQ) as separate groups, but the difference between the two models (one with HL as one group and one with HL as two groups) was not statistically significant ($p = 0.08$). In other words, as we saw in Tasks 1 and 3, the differences between the HLSEQ and HLSIM subgroups were not statistically significant.

4.3. Analysis and Discussion

In the first research question, we asked whether HL and L2 speakers' responses would differ from each other and from the native control groups. In line with our predictions, the data reported in the previous sections showed significant differences between the control and the experimental groups in all tasks. In the sentence completion task, in French (Task 2), both experimental groups differed from the L1FR control group in that they both used verbs in the subjunctive, particularly in "matching" conditions, where the Spanish equivalents would also require the use of the subjunctive. In contrast, the subjunctive mood was practically absent in the L1FR control data, which was an expected result and compatible with the idea that subjunctive is being lost in French relative clauses (Menanteau 1986). In the Sentence Combination Felicity tasks, in Spanish (Tasks 1 and 3), the difference between the L1SP control and the HL group was significant when both the stimuli were presented in written form (Task 1) and orally (Task 3). The difference between the HL and the L2 groups became more apparent and reached statistical significance only when the data from both tasks were considered. In turn, the performance of the two HL subgroups (HLSIM and HLSEQ) was comparable in the two types of task, as demonstrated by the absence of main effects regarding the "type of bilingualism" in the statistical analyses.

Research question 2 asked whether we would find a correlation between the use of subjunctive in Task 2 and the results of Tasks 1 and 3, which would indicate a possible influence of the L1 on the L2, or vice versa. As predicted, there was a positive correlation between an increased use of the subjunctive in Task 2 and higher acceptance rates of clauses using the subjunctive mood in Tasks 1 and 3. Indeed, the HL group, which presented the highest rates of subjunctive use in Task 2, was also the group that displayed the highest acceptance rates of subjunctive conditions in Tasks 1 and 3. With respect to the L2SP group, given that the use of the subjunctive is rather scarce in these constructions in French, an influence of the L1 on the L2 may have surfaced in the form of lower acceptance or higher uncertainty in the subjunctive conditions in Tasks 1 and 3. Our results indicate that this is not quite the case, since L2 participants were able to successfully accept appropriate uses of the subjunctive and reject inappropriate uses of this mood. Furthermore, we found that both HL and L2 participants use subjunctive in French as well (Task 2), and that they did so at significantly higher rates than their monolingual counterparts. Our results also indicate that the distribution of their responses in terms of mood choice displayed a certain amount of overlap with Spanish, as the conditions in which these participants provided most of their subjunctive responses were those in which the Spanish translation would also require the use of the subjunctive (i.e., "matching"). This presence of the subjunctive in French was to be expected of the HL group, but not so much of the L2SP group, given that French was their only L1, they only began learning Spanish after puberty, and they all lived in Quebec, in an environment where French is the majority language. The influence of the L1 on the L2 is a widely attested and thoroughly studied phenomenon in the field of SLA. As Silva-Corvalán (2014, p.21) notes, "[t]he weaker language is expected to evidence, among other features, more errors of production and more frequent use of structures that parallel those in the stronger language, to the detriment of alternatives attested in the input". What is less frequent is for researchers to report on the effects of the weaker/L2 language on the dominant/L1 language (a few notable exceptions are Cook 2003; Kecskes 2008; as well as studies such as Pavlenko and Jarvis 2002; Dussias and Sagarra 2007; among others), and even less so in non-immersion contexts, which is the case of our participants.

Finally, research question 3 asked whether participants would behave differently in oral and written tasks. As discussed in Section 3, having the same task—a SCFT in Spanish—administered in two different modes—orally and in written form—allowed us to explore the issue of modality in isolation, in an attempt to explore the extent to which the way in which stimuli were presented to the participants may have an influence in their responses. Based on previous studies, our prediction was that L2 speakers would perform better in the written version of our SCF task, while HL speakers would fare better in the oral version. As shown in the previous section, our results revealed the presence of a task effect, although not the way we had expected. Overall, rates tended to be higher in the oral

task than in the written task, and this difference was noticeable in all groups, including the L1SP control. However, the statistical analysis did not reveal significant differences with respect to task mode for any of the experimental groups. In contrast, it did reveal a main effect for task mode for the L1 control. This effect was caused by the aforementioned tendency to assign higher rates to inappropriate conditions in the oral task, as compared to the written task. A possible explanation for this may be found in the different degrees of formality that are associated with oral and written speech. Specifically, L1SP participants tended to "penalize" inappropriate conditions more strongly in the written task (Task 1) than in the oral task (Task 3), which is in line with the idea that written speech is normally associated with a higher degree of formality and, consequently, tolerance for errors would tend to be minimal.[4] Regardless of this unexpected task effect, the L1SP group did present the predicted response distribution pattern in both tasks, providing positive rates to all appropriate conditions and negative rates to all inappropriate conditions. Regarding the experimental groups, there was no apparent advantage for HL speakers in the oral task, which stands in contrast with previous findings by authors such as Montrul et al. (2008) and Bowles (2011).[5] Instead, our results are in line with recent research on Dutch heritage speakers of Spanish carried out by van Osch and Sleeman (2016). Like the Dutch participants, our participants' multilingual experience and the educational system may have contributed to the development of higher metalinguistic awareness in our participants. Specifically, our participants spoke English at varying levels—english as a second language (ESL) is mandatory from the beginning of primary education in Quebec, and access to the language is readily available, given its status as one of the official languages of Canada. Furthermore, it is also worth noting that, in the Quebec education system, French grammar is taught explicitly in school as part of the curriculum, which likely constitutes another relevant contributing factor in the development of our participants' presumed heightened metalinguistic awareness.

5. Conclusions

Our study explored mood selection in Spanish relative clauses by French–Spanish bilinguals by comparing two types of populations: heritage speakers and L2 learners. In addition, we addressed the question of whether any potential differences between these two groups may be task-induced. Our results show that, although the experimental groups did behave differently from each other or

---

[4] One anonymous reviewer wondered whether the saliency of the subjunctive (i.e., irregular forms, such as *tenga*, from the verb *tener*, 'to have', as opposed to regular forms, such as *venda*, from the verb *vender*, 'to sell') may have played a role in the results of the native control group. Specifically, the reviewer wondered whether the indicative–subjunctive distinction could have been more easily misperceived in the oral task given that the difference between indicative and subjunctive forms is less salient in the case of regular verbs (e.g., *vende* vs. *venda*) than irregular verbs (e.g., *tiene* vs. *tenga*). Even though the saliency of the verb form was not a variable included in the experimental design, the distribution of regular and irregular verbs in our items was rather even, with 56% of items containing a regular verb and 44% containing an irregular verb. The results of a repeated measures ANOVA reveal that there was, indeed, a main effect for saliency ($F(1, 30) = 7.587$, $p = 0.010$). Specifically, our data reveal that participants tended to reject inappropriate mood choices more strongly when they contained an irregular verb, which is in line with the reviewer's prediction. However, there was no significant interaction between saliency and task mode (oral/written) ($F(1, 30) = 2.216$, $p = 0.147$), since this pattern of higher rejection of the inappropriate condition with irregular verbs was present in both the written ($F(1, 10) = 8.461$, $p = 0.016$) and in the oral task ($F(1, 20) = 5.785$, $p = 0.026$). These results show that, while our native control group did show sensitivity to verb saliency, the extent to which this factor played a role in the modality effect found in the native control data is not so clear, as regular and irregular verbs were treated differently in both tasks. As such, the issue of verb saliency remains a topic for further investigation.

[5] One anonymous reviewer pointed out that this overall lack of difference in modality by our experimental groups may be due to the fact that, whether presented orally or in written form, the task was essentially the same. In addition, at least in the case of HL speakers, the task represented a rather uncommon linguistic activity for HL participants (judging and correcting sentences). While we agree with the reviewer that all these points merit consideration, our control group did show sensitivity to task modality, as their responses were more polarized in the written task than in the oral task. As discussed in this section, and in Section 4.1.3, above, these differences reached statistical significance. We take these results as a positive indicator that modality was a relevant intervening variable in our experimental design, all the more so if we consider that, just like HL speakers, our native control participants were not accustomed to judging and correcting sentences (whether orally or in written form), and, yet, our results indicate that their responses were, indeed, affected by modality. Thus, in our view, it is necessary to consider alternative explanations for the lack of a modality effect in our experimental groups.

from the Spanish control group, the mode in which the stimuli were presented (either in written form or orally) did not have a significant effect on the HL and L2 participants' responses. As van Osch and Sleeman (2016) noted about their Dutch participants, the absence of a task effect in our data may also be caused by an enhanced metalinguistic awareness from our participants, brought about by a combination of their multilingual experience and the specificities of the Quebec francophone educational system. In addition, Task 2 of our study revealed the apparent influence of Spanish on our participants' dominant language, signalled by the increased production of subjunctive responses in French. In contrast, our French control group (with no knowledge of Spanish) showed minimal use of the subjunctive mood, which supports the general assumption that the subjunctive is being lost in French relative clauses. Traditionally, when researchers have reported effects of the L2 on the L1, these normally occurred in contact situations where the L2 is the majority language. In contrast with this, our study constitutes an example of L2 influence on the L1 in a non-contact situation where the L1 is the majority as well as the participants' dominant language. This finding provides empirical support for the idea that bilinguals are, indeed, different from monolinguals, and that, as Kecskes (2008, p.31) explains, citing his own previous work, "(. . .) people with more than one language have different knowledge of their L1 than do monolingual people, and this difference can mainly be explained by the effect of subsequent languages on the development and use of L1 skills".

Author Contributions: All the authors contributed equally to this work.

Funding: This research received no external funding.

Acknowledgments: We would like to express our gratitude to all the participants who volunteered to take part in this study. We are also thankful to the two anonymous reviewers for their comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Item list for Task 1: Sentence combination felicity task

1. Contrataremos una persona que hable holandés, si logramos encontrar alguna.
2. Luis y Pedro quieren amueblar la casa porque les gusta más vacía.
3. Quiero una lavadora que indica la temperatura. La venden en Amazon y está barata.
4. Esteban y su hermano se parecen mucho. Siempre los confundo.
5. Compraré una casa que queda en el centro, si hay alguna a buen precio.
6. Necesitamos un director que le conviene a la empresa pero creo que ninguno de los candidatos es bueno para dirigir.
7. Tengo que alquilar una camioneta para la mudanza. Nos mudamos este sábado.
8. Quiero un motor que no haga ruido, pero no hay ninguno en esta tienda.
9. Tengo que resolver un problema difícil, por eso necesito tu ayuda.
10. Quiero un libro que contenga muchos personajes fantásticos. Se titula "La historia sin fin".
11. Viajaremos a Canadá en el verano porque queremos conocer la nieve.
12. Compraré un traje que va bien con el tuyo, si encuentro alguno en esta tienda.
13. Tenemos que encontrar una enciclopedia que muestra todos los cuadros de Picasso. No recuerdo en qué estante la puse.
14. Me gusta mucho la pera pero prefiero la manzana.
15. Podemos navegar en un barco que tenga cuatro cabinas. Es aquel que pertenece a la compañía "Caribe sol".
16. Necesitamos un autobús que tiene asientos vacíos pero están todos llenos.
17. Busco a mi perro porque no quiero encontrarlo.
18. Debo comprar una corbata que va bien con mi camisa azul. ¿Cuál me recomiendas?
19. María debe comprar un perro porque tiene ratones en la casa.

20. Buscamos un perro que tenga una sola oreja. Se llama Fido.
21. Con tu experiencia como curador artístico, puedes trabajar en una tienda que vende obras de arte. Busca en Internet, seguro hay alguna en esta ciudad.
22. Prefiero un museo que tenga dinosaurios pero si tiene dinosaurios no me gusta.
23. Debes utilizar un ordenador que esté conectado a la red. Es el que está al lado de la puerta de entrada.
24. Contrataremos un músico que toca el bajo pero por el momento solo encontramos pianistas.
25. Necesitamos un gato porque hay ratones en casa.
26. Buscamos un restaurante que no cuesta muy caro, me han dicho que se llama "El Marino".
27. Mis abuelos visitarán un lugar en el que hace buen tiempo. Se van para Cuba.
28. Iremos al cine a ver una película sobre los dinosaurios. Comienza a las 4:00.
29. Leeré un libro que habla de Beethoven. Me lo regaló mi hermano.
30. Tengo que organizar una cena para vegetarianos, por eso he comprado carne de res y de cerdo.
31. Necesito un lápiz que escriba oscuro. Es ese rojo que está sobre la mesa, ¿me lo pasas?
32. Veremos una película que esté nominada a un Oscar. Se titula *Boyhood*.
33. Busco un hotel barato pero todos son caros.
34. Tenemos que encontrar un mapa que muestre los lugares turísticos de Roma. Creo que lo puse en la maleta verde.
35. Comeremos una manzana que está barata esta semana. La cultivan aquí en Quebec.
36. Quiero una habitación que tiene mucha luz pero solo hay habitaciones oscuras en este edificio.
37. Debo vender una casa que tiene piscina y jardín. Creo que será difícil venderla porque es muy cara.
38. Quiero comprar unos zapatos para la carrera del domingo.
39. En esta biblioteca puedes leer un libro que esté escrito en alemán. Hay muchos en el segundo piso.
40. Busco un libro que tiene reproducciones de pinturas de Leonardo da Vinci. Es azul y verde, ¿no lo has visto?
41. Necesito un libro para leer pero no sé leer.
42. Te cantaré una canción que te haga feliz. ¿Tienes alguna sugerencia?
43. Tengo que cocinar un plato que le guste mucho. Dime qué le gusta comer.
44. Es suficiente con dos litros de leche pero cuatro litros no alcanzan.
45. Necesito un libro que ayude a comprender la física de manera fácil, pero creo que nadie ha escrito todavía un libro así.
46. Pediremos una comida que contenga poca grasa. ¿Me recomienda algún plato en particular?
47. Esta noche podemos ver un programa que critica el sistema de salud. Comienza a las 10:00.
48. Miguel prefiere jugar al fútbol que al tenis. Es que el fútbol es un deporte más divertido.
49. Tengo que alquilar una camioneta que carga hasta 1000 kilogramos pero no existe ninguna así.
50. Miguel busca un negocio que venda muebles antiguos. Le han dicho que en la capital podrá encontrar alguno.
51. Laura verá en el cine una película que la haga llorar. La ha visto más de diez veces.
52. Juan tiene que ir al trabajo en bicicleta porque está desempleado.
53. Debes leer un libro que te permita comprender el subjuntivo. En esta biblioteca seguro encontrarás alguno.
54. En el próximo concierto tocaremos una guitarra que tenga doce cuerdas. La acabo de comprar en la tienda.

Appendix B

Item list for Task 2: Elicited production task

1. Laura verra au cinéma un film qui la _____ (faire) pleurer. Elle l'a vu plus de dix fois.
2. Je veux un livre qui _____ (contenir) de nombreux personnages fantastiques. Il s'intitule "The Neverending Story".
3. Nous devons trouver une encyclopédie qui _____ (contenir) tous les tableaux de Picasso. J'ai oublié dans quelle étagère je l'ai mise.
4. Je vais chanter une chanson qui te _____ (rendre) heureux. J'espère être capable de la composer.
5. Nous avons besoin d' un directeur qui _____ (convenir) le mieux à l'entreprise, mais je pense qu' aucun des candidats n'est bon pour diriger.
6. Mes grands-parents iront visiter un endroit où il _____ (faire) beau. Ils iront à Cuba.
7. Avec ton expérience en tant que commissaire artistique, tu peux travailler dans un magasin qui _____ (vendre) des œuvres d'art. Recherche sur Internet, je suis sûre qu'il y en a certains dans cette ville.
8. J'ai besoin d'un crayon qui _____ (écrire) sur le verre. C'est celui qui est sur la table, peux-tu me le passer?
9. Je dois vendre une maison qui _____ (avoir) une piscine et un jardin. Je pense qu'elle sera difficile à vendre, car elle est très chère.
10. Nous allons commander un repas qui _____ (contenir) peu de graisse. Recommandez-vous un plat en particulier?
11. Je cherche un livre qui _____ (contenir) des reproductions des tableaux de Léonard de Vinci. Il est bleu et vert, l'as-tu vu?
12. Nous pouvons naviguer dans un bateau qui _____ (avoir) quatre cabines. Il appartient à l'entreprise "Les Caraïbes."
13. Je veux une laveuse qui _____ (dire) la température. On la vend sur Amazon à bon prix.
14. Miguel cherche un magasin qui _____ (vendre) des meubles anciens. On lui a dit qu'au centre-ville il pourra en trouver un.
15. Je dois louer un camion qui _____ (pouvoir) charger jusqu'à 1000 kg, mais il n'y en a aucun.
16. Nous allons embaucher un musicien qui _____ (être) bassiste, mais jusqu'à présent, nous n'avons trouvé que des pianistes.
17. Je veux un moteur qui ne _____ (faire) pas de bruit, mais il n'y en a aucun dans ce magasin.
18. Je vais lire un livre qui _____ (contenir) la biographie de Beethoven. Mon frère me l'a donné en cadeau.
19. Ce soir, nous pouvons regarder une émission qui _____ (faire) une critique du système de santé. Elle commence à 10h00.
20. J'achèterai un habit qui _____ (allez) bien avec le tien, si j'en trouve un dans ce magasin.
21. Tu dois lire un livre qui te _____ (permettre) de comprendre le subjonctif. Dans cette bibliothèque tu en trouveras sûrement un.
22. Je dois acheter une cravate qui _____ (aller) bien avec ma chemise bleue. Laquelle me recommanderiez-vous?
23. Je veux une chambre qui _____ (être) bien éclairée, mais il n'y a que des pièces sombres dans ce bâtiment.
24. Ce soir au concert, nous allons rencontrer un chanteur qui _____ (être) aussi un guitariste très doué. Il s'appelle Antonio.
25. Nous cherchons un restaurant qui _____ (être) pas très cher. On m'a dit qu'ils'appelle "El Marino".

26. Nous devons trouver une carte qui _____ (contenir) les sites touristiques de Rome. Je pense que je l'ai mise dans la valise verte.
27. Nous allons voir un film qui _____ (être) nominé pour un Oscar. Il s'intitule "Boyhood".
28. Nous avons besoin de monter à bord d'un autobus qui _____ (avoir) des sièges libres, mais tous les autobus sont pleins.
29. Je dois préparer un plat qui _____ (être) à son goût. Dites-moi ce qu'il préfère manger.
30. Je vais acheter une maison qui _____ (être) au centre-ville, si j'en trouve une pas trop cher.
31. Tu dois utiliser un ordinateur qui _____ (être) connecté au réseau. C'est celui qui est sur la table à l'entrée du laboratoire.
32. Nous cherchons un chien qui n'_____ (avoir) qu'une seule oreille; il s'appelle Fido.
33. Nous mangerons une sorte de pomme qui _____ (être) en rabais cette semaine. On la cultive ici au Québec.
34. Dans cette bibliothèque tu peux lire un livre qui _____ (être) écrit en allemand. Il y en a beaucoup au deuxième étage.
35. Nous allons embaucher quelqu'un qui _____ (comprendre) le néerlandais, si nous pouvons trouver quelqu'un.
36. J'ai besoin d'un livre qui _____ (décrire) de manière facile la physique, mais je crois qu'aucun auteur ne serait capable d'expliquer la physique de manière facile.

References

Alba de la Fuente, Anahi. 2012. Clitic Combinations in Spanish: Syntax, Processing and Acquisition. Ph.D. dissertation, University of Ottawa, Ottawa, ON, Canada. Available online: http://dx.doi.org/10.20381/ruor-5926 (accessed on 20 January 2016).

Blake, Robert. 1983. Mood Selection Among Spanish-Speaking Children, Ages 4 to 12. *The Bilingual Review/La Revista Bilingüe* 10: 21–32.

Borgonovo, Claudia, Joyce Bruhn De Garavito, and Philippe Prévost. 2008. Methodological Issues in the L2 Acquisition of a Syntax/Semantics Phenomenon: How to Assess L2 Knowledge of Mood in Spanish Relative Clauses. In *Selected Proceedings of the 10th Hispanic Linguistics Symposium*. Edited by Joyce Bruhn de Garavito and Elena Valenzuela. Somerville: Cascadilla Proceedings Project, pp. 13–24.

Borgonovo, Claudia, Joyce Bruhn De Garavito, and Philippe Prévost. 2015. Mood selection in relative clauses: Interfaces and variability. *Studies in Second Language Acquisition* 37: 33–69. [CrossRef]

Bosque, Ignacio. 1999. Sobre la gramática de los contextos modales: Entornos modales y expresiones inespecíficas en español. In *Actas del XI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina Las Palmas de Gran Canaria, del 22 al 27 de julio de 1996*. Las Palmas: Servicio de Publicaciones de la Universidad de Las Palmas, pp. 43–58.

Boudreau, Geneviève. 2007. El Modo en la Esfera Nominal: Análisis y Adquisición. Master's thesis, Université Laval, Québec, QC, Canada. Available online: http://hdl.handle.net/20.500.11794/19268 (accessed on 16 November 2016).

Bowles, Melissa A. 2011. Measuring implicit and explicit linguistic knowledge: What Can Heritage Language Learners Contribute? *Studies in Second Language Acquisition* 33: 247–71. [CrossRef]

Cook, Vivian. 2003. The changing L1 in the L2 user's mind. In *Effects of the Second Language on the First*. Edited by Vivian Cook. Clevedon: Multilingual Matters, pp. 1–18.

De Leeuw, Joshua R. 2015. jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods* 47: 1–12. [CrossRef] [PubMed]

Dussias, Paola E., and Nuria Sagarra. 2007. The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition* 10: 101–16. [CrossRef]

Ellis, Rod. 2005. Measuring implicit and explicit knowledge of a second language: A Psychometric Study. *Studies in Second Language Acquisition* 27: 141–72. [CrossRef]

Kecskes, Istvan. 2008. The effect of the second language on the first language. *Babylonia* 2: 31–34.

Koenig, Renee Elizabeth. 2016. La selección modal del subjuntivo/indicativo en el Corpus sociolingüístico Preseea-Medellín. In *Boletín de Filología*. vol. 51, pp. 89–120. [CrossRef]

Lareau, François. 2008. Vers une Grammaire D'unification Sens-Texte du Français: Le Temps Verbal Dans L'interface Sémantique-Syntaxe. Ph.D. dissertation, Université de Montréal/Université Paris 7, Montréal, QC, Canada, August. Available online: http://hdl.handle.net/1866/6421 (accessed on 5 September 2016).

Lipski, John. 1993. Creoloid phenomena in the Spanish of transitional bilinguals. In *Spanish in the United States*. Edited by Ana Roca and John Lipski. Berlin: Mouton, pp. 155–73.

Lynch, Andrew. 1999. The Subjunctive in Miami Cuban Spanish: Bilingualism, Contact, and Language Variability. Ph.D. dissertation, University of Minnesota, Minneapolis, MN, USA, August. Available online: http://works.bepress.com/andrewlynch/22/ (accessed on 10 January 2018).

Menanteau, Didier. 1986. Le mode verbal, classe grammaticale? *La Linguistique* 22: 69–80.

Merino, Barbara. 1983. Language loss in bilingual chicano children. *Journal of Applied Developmental Psychology* 4: 277–94. [CrossRef]

Montrul, Silvina. 2007. Interpreting mood distinctions in Spanish as a heritage language. In *Spanish in Contact: Policy, Social and Linguistic Inquiries*. Edited by Kim Potowski and Richard Cameron. Amsterdam: John Benjamins, pp. 23–40.

Montrul, Silvina. 2012. Is the heritage language like a second language? *Eurosla Yearbook* 12: 1–29. [CrossRef]

Montrul, Silvina, and Silvia Perpiñán. 2011. Assessing differences and similarities between instructed heritage language learners and L2 learners in their knowledge of Spanish tense-aspect and mood (TAM) morphology. *Heritage Language Journal* 8: 90–133.

Montrul, Silvina, Rebecca Foote, and Silvia Perpiñán. 2008. Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning* 58: 503–53. [CrossRef]

Murillo, Jorge. 2000. Oraciones de relativo y variación modal en el habla culta costarricense. *Revista de Filología Española* 80: 141–153. [CrossRef]

Paradis, Michel. 2004. *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.

Pavlenko, Aneta, and Scott Jarvis. 2002. Bidirectional Transfer. *Applied Linguistics* 23: 190–214. [CrossRef]

Pérez-Leroux, Ana Teresa. 1998. The acquisition of mood selection in Spanish relative clauses. *Journal of Child Language* 25: 585–604. [CrossRef] [PubMed]

Restorick Elordi, Audrey K. 2012. Interpretation of the French and Spanish Subjunctive by L1, L2, and L3 Speakers: Contexts Where Mood Can Alternate without Ungrammaticality. Ph.D. dissertation, University of Western Ontario, London, ON, Canada. Available online: https://ir.lib.uwo.ca/etd/1055/ (accessed on 29 September 2016).

Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. New York: Oxford University Press.

Silva-Corvalán, Carmen. 2003. Linguistic consequences of reduced input in bilingual first language acquisition. In *Linguistic Theory and Language Development in Hispanic Languages*. Edited by Silvina Montrul and Francisco Ordóñez. Sommerville: Cascadilla Press, pp. 375–97.

Silva-Corvalán, Carmen. 2014. *Bilingual Language Acquisition: Spanish and English in the First Six Years*. Cambrigde: Cambridge University Press.

Slabakova, Roumyana, and Silvina Montrul. 2002. On viewpoint aspect interpretation and its L2 acquisition. In *The L2 acquisition of tense-aspect morphology*. Edited by Rafael Salaberry and Yasuhiro Shirai. Amsterdam: John Benjamins, pp. 363–95.

van Osch, Brechje, and Petra Sleeman. 2016. Spanish heritage speakers in the Netherlands: Linguistic patterns in the judgment and production of mood. *International Journal of Bilingualism*. [CrossRef]

Zentella, Ana Celia. 1997. *Growing up Bilingual*. Malden: Blackwell.

languages

MDPI

Article

Continuity in the Adult and Children's Comprehension of Subject and Object Relative Clauses in French and Italian

Maria Teresa Guasti [1,2,*], Mirta Vernice [1] and Julie Franck [3]

1 Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo, 20126 Milano, Italy; mirta.vernice@unimib.it
2 Australian Research Council (ARC) Centre of Excellence in Cognition and its Disorders (CCD), Macquarie University, 16 University Avenue, NSW 2109 Sydney, Australia
3 Faculty of Psychology and Education Sciences, University of Geneva, Boulevard du Pont-d'Arve 40, 1211 Geneva 4, Switzerland; Julie.Franck@unige.ch
* Correspondence: mariateresa.guasti@unimib.it; Tel.: +39-02-6448-3812

Received: 5 January 2018; Accepted: 29 June 2018; Published: 5 July 2018

Abstract: Subject and object relative clauses have been studied from the point of view of language acquisition and adult sentence processing. In the adult sentence processing literature, subject relative clauses (RCs) are read faster than object RCs (e.g., Frauenfelder et al. 1980 for French; King and Kutas 1995 for English; Schriefers et al. 1995 for Dutch). Similarly, children understand and produce subject RCs earlier and with greater accuracy than object RCs in a variety of languages with head-initial relative clauses, as English, Hebrew and Italian. These findings cannot be a coincidence but reflect the fact that what children acquire first is also easier to process by adults. In this article, we support this observation by investigating subject and object RCs in children and adults speaking French and Italian. These languages display subject and object relatives as in (1), but they also have a type of object relative in which the subject is postverbal. We replicate the observation that subject relatives are easier than object and show that object relatives as in (1b), with the embedded subject in preverbal position are easier than those with the embedded subject in postverbal position, both for children and adults. We offer an account of these findings in terms of Fodor and Inoue's (2000) diagnosis model in light of the fact that acquisition involves processing.

Keywords: relative clauses; children; adults; processing; French; Italian

1. Introduction

Language acquisition and language processing research have generally proceeded separately, although in the last 15 years, phenomena studied in the adult psycholinguistic literature have been explored in children to document how children solve ambiguities, whether/how they differ from adults and whether aspects of grammatical development may be due to children's limited sentence processing abilities (see Phillips and Ehrenhofer 2015 for a review). One phenomenon that illustrates this trend is subject and object relative clauses.

In the adult sentence processing literature, subject relative clauses (RCs), as in (1a), are read faster than object RCs, as in (1b) (e.g., Frauenfelder et al. 1980 for French; King and Kutas 1995 for English; Schriefers et al. 1995 for Dutch).

1. a. The boy [that *t* kissed the girl] went home

   b. The boy [that the girl kissed *t*] went home.

This subject RC advantage in English is demonstrated by a large number of studies performed with various methods. When processing sentences such as (1b) in comparison with (1a), subjects display poor comprehension in lexical decision tasks (Ford 1983) with up to 35% errors (Villata et al. 2018; see also Carminati et al. 2006), longer reading times in self-paced reading experiments (King and Just 1991), longer fixations in eye-tracking experiments (Traxler et al. 2002), both quantitative and qualitative differences in event-related potentials (King and Kutas 1995), dissimilar brain activations in functional magnetic resonance imaging (Caplan et al. 2002; Cooke et al. 2002; Just et al. 1996) and in positron emission tomography (Stromswold et al. 1996). The preference for subject RCs is uniformly observed in several languages with head-initial relative clauses: Dutch (Frazier 1987; Mak et al. 2002), German (Schriefers et al. 1995; Mecklinger et al. 1995), French (Frauenfelder et al. 1980; Cohen and Mehler 1996; Schelstraete and Degand 1998; Casalis et al. 2013), Spanish (Betancort et al. 2009).

Similarly, children understand and produce subject RCs earlier and with greater accuracy than object RCs in a variety of languages with head-initial relative clauses, such as English, Hebrew and Italian, among others (Adams 1990; Friedmann et al. 2009; Guasti and Cardinaletti 2003; Utzeri 2007; Belletti and Contemori 2010; Contemori and Garraffa 2010; Guasti 2017). Several experiments carried out with an act-out task (De Villiers et al. 1979; Tavakolian 1981), a picture-selection task (Adani et al. 2010; Adani et al. 2014; Arnon 2005; Arosio et al. 2009; Friedmann and Novogrodsky 2004; Volpato and Vernice 2014), a referent/agent selection task (Adani 2011), and self-paced reading and listening tasks (Booth et al. 2000; Arosio et al. 2011) have documented a marked subject/object asymmetry at all ages. While children understand subject RCs quite well at age 3–4, they only understand object RCs above chance around age 5 and are still improving around age 10, with differences among studies depending on the procedure and material used and on the specific properties of RCs in the language. Given the above-mentioned studies indicating that object RCs seem to be critical not only in child language acquisition, but in adult sentence processing, one could propose the following generalization:

2. Child–adult generalization (CHAG): if structure B is more difficult to process (e.g., elicits longer processing time) than structure A by adults, then structure B is mastered later than structure A by children.
   The boy [that *t* kissed the girl] went home

Relative clauses are an emblematic example of this generalization, but they are not the only one; the same holds true for wh-questions (e.g., adults: De Vincenzi 1991; Kaan et al. 2000 and children: Durrleman et al. 2016) and for some cases of anaphora resolution (see Phillips and Ehrenhofer 2015 for a review).[1] Based on the generalization in (2), if adults and children struggle on the same structures, the same common factors may be responsible for their difficulties.

In this article, we investigate the processing and the comprehension of subject and object RCs by French- and Italian-adults and children. This allows us to bring another type of object RC present

[1] One reviewer wonders whether our generalization in (2) holds in the other direction, i.e., if all structures that are mastered later are also more difficult to process. This is an empirical question. Certainly, this holds in a banal way for relative clauses and wh-questions. We can reframe the question by asking whether what matters is the sequence of acquisition in production and comprehension or both. The available evidence allows us to say that CHAG holds in both directions when comprehension and production are at stake (as it is the case of wh-questions and relative clauses). The production of clitic pronouns in some Romance languages reaches adult level at 3 or 4 years of age and thus qualifies as a late acquisition. Little is known about comprehension and even less about adults' processing. This would be an area to further investigate the bidirectionality of CHAG. A second reviewer asks whether the noun phrase accessability hierarchy (NPHA, Keenan and Comrie 1977) could fall under CHAG. We think that it does. According to the NPAH, the ease of a given relative clause depends on the grammatical function of the relative head. Thus, subjects are easier than objects, these than oblique and these last than genitive relative clauses. Several studies on second language acquisition have supported the hierarchy (see Shirai and Ozeki 2007 for review). Although there is no systematic investigation of relative clauses spanning the whole hierarchy, we know that, in head-initial languages, subject relatives are acquired before all the other relatives and genitive relative clauses are the last (Guasti and Cardinaletti 2003 for evidence from French and Italian; see also Diessel and Tomasello 2000 for evidence from English and German).

in these languages to the stage, one with the subject in a postverbal position, and investigate it with respect to the object RC with a preverbal subject and with respect to generalization (2), which we will show holds for this structure as well. Capitalizing on generalization (2), we offer an account of the underlying difficulties experienced by adults and children. The paper is organized as follows. First, we discuss some typological properties of French relative clauses in the light of previous literature and make some predictions about French- and Italian-speaking children's acquisition of RCs and French- and Italian-speaking adults processing of RCs, which we test in three experiments. On the basis of our results, we propose a processing account, whereby children and adults are more willing to revise a preferred and simpler analysis when they have positive information about how to perform the correct analysis (Fodor and Inoue 2000), rather than when they lack this piece of information.

2. Setting the Stage: French Adults' Processing of Relative Clauses

We start by recalling some typological properties of French RCs. First, in French, subject and object RCs are obligatorily introduced by distinct relative markers (or complementizers), unlike in the English examples in (1). Consider the examples in (3a) and (3b). In (3a), the subject RC is introduced by the relative marker *qui* 'that', which is used in all cases of subject extraction whereas the object RC in (3b) is introduced by the relative marker *que* 'that', which is used in cases of object extraction. In spite of this early marker of the syntactic nature of the extracted head, object RCs remain more difficult to process than subject RCs (Frauenfelder et al. 1980; Cohen and Mehler 1996). Second, like in English, French object RCs are usually constructed with the subject preceding the verb (OSV RCs), as in (3b). However, and unlike in English, object RCs can also have the subject follow the verb, as in (3c) (OVS RCs).[2]

3. a. Le gardien de but qui critiquait les joueurs de football est tombé
   'The goalkeeper that criticized the soccer players fell.'

   b. Le gardien de but que les joueurs de football critiquaient est tombé
   'The goalkeeper that the soccer players criticized fell.'

   c. Le gardien de but que critiquaient les joueurs de football est tombé
   'The goalkeeper that (whom) the soccer players criticized fell.'

Schelstraete and Degand (1998) performed an experiment on adults' comprehension of these three types of French RCs (3a–c) using a self-paced reading task. They found that both kinds of object RCs were harder to process than subject RCs. In addition, they observed that participants showed shorter reading times in OVS RCs such as (3c) than in OSV RCs (3b). They interpreted this outcome as evidence supporting a functionalist approach: the competition model (Bates 1999; Bates and MacWhinney 1987). According to this model, the form–function mapping in sentence comprehension emerges from a competition based on the strength of language specific cues and on the memory costs of processing these cues. In the subject RC (3a), there is no competition, as the first noun (the head of the relative) comes before the verb and takes the subject role, which canonically occupies

[2] Subject inversion in French occurs in a limited set of syntactic environments, among which in contexts of extraction, as in OVS relative clauses. OVS RCs are an instance of stylistic inversion in French (Kayne and Pollock 1978). In the 1980s, the agreement was that the subject was in a low position of the clause, possibly right-adjoined to the Verb Phrase (VP), regardless of the type of structure. In more recent times, it has been assumed that some instances of postverbal subjects in stylistic inversion contexts are in a high topic position in the Complementizer Phrase (CP) area and the whole clause has moved past to it. It is also possible that not all postverbal subjects in stylistic inversion environments are in the same position (Kayne and Pollock 2001). OVS RCs are rarely used (Noizet et al. 1972) and may be misinterpreted as subject RCs (Holmes and O'Regan 1981). In Italian, postverbal subjects are more common; however, not all postverbal subjects are in the same position. In particular, the postverbal subject occurring in extraction contexts, in contrast to declarative contexts, is in a sort of dislocated position (Guasti 1996), possibly right adjoined to the VP or Inflectional Phrase (IP). We assume that the same holds for French relatives. According to Belletti and Chesi's (2011) corpus analysis, OVS RCs are rare (9%), while subject RCs are much more frequent (66%) and OSV RCs are in between (25%).

a sentence-initial position in French. The OSV RC in (3b) should be more difficult to solve than (3a) for two reasons. First, the word order cue is not valid, as the sentence does not have the canonical word order: the first noun, i.e., the head of the relative, must receive the object role, rather than the subject role. Second, it requires the parser to keep the two nouns (the object and the subject) in working memory before the verb is reached. The OVS RC in (3c) is expected to be harder than (3a), but easier than (3b). As in (3b), word order does not provide a valid cue since the head of the relative, which comes before the verb, has to take the object role, while the post-verbal NP has to take the subject role. However, in terms of memory load, OVS RCs are like subject RCs in that they only require one noun to be stored (the head of the relative) before encountering the verb. In line with these predictions, Schelstraete and Degand (1998) found that participants took more time to read object RCs than subject RCs, and that OSV were slower than OVS RCs. The authors argued that the memory load component, whose burden is heavier in the OSV relative, was the factor responsible for the processing difficulties observed in OSV RCs and that the memory load counted more than the word order cue.

The competition model makes predictions concerning the course of language acquisition: the order of acquisition of a given structure is determined by the validity of its form-function mapping cues. Hence, subject RCs are expected to be acquired earlier than object RCs since word order cues are not valid in the latter. This is in line with the various findings from the cross-linguistic literature. Since OSV sentences involve higher memory demands than OVS sentences and children's memory resources are known to develop over time, the former are expected to be acquired later, in line with the finding that adults are faster to read OVS compared to OSV. Thus, according to the competition model, the hierarchy of difficulties in French, based on adult data, is: subject > OVS > OSV relatives. Data form Italian-speaking children discussed by Arosio et al. (2009), Adani (2011) yield a different hierarchy of difficulties: subject > OSV > OVS relatives. Given that postverbal subjects are less stylistic marked in Italian than in French and likely they are more common in the former than in the latter language, the hierarchy of difficulties ensuing from Schelstraete and Degand's (1998) study is surprising. However, since Italian data are from children and French data are from adults, our first step is to establish whether the Italian hierarchy holds for French-speaking children or not. In the positive case, there would be a discrepancy between French-speaking children and adults and our generalization would be disconfirmed; in the second, there would be a discrepancy between French and Italian (at least as far as children are concerned).

3. Experiment 1: French Children

In this experiment, we tested 5 to 7 year-old French-speaking children's comprehension of RCs using the same method used by Adani (2011) with Italian children; i.e., a referent selection task. The research was approved by the school commission of the Département de l'Instruction Publique de Genève and parents had to provide written consent for their child.

3.1. Participants

Fourteen monolingual French-speaking children attending various schools in Geneva, Switzerland participated in the experiment, ranging in age from 5;5 to 7;4 with a mean age 6;5 (standard deviation SD = 6;3). Two participants were discarded because they answered incorrectly to filler items most of the time (at least 83% incorrect answers).

3.2. Materials and Design

The experiment consisted of 24 spoken right-branching RCs, 8 subject RCs (4a), 8 OSV RCs (4b) and 8 OVS RCs (4c).

4. a. SVO relative clause (Subject RC)
   Montre-moi [le lion qui mord les chameaux]
   'Show me the lion that bites the camels.'

   b. OSV relative clause (Object RC)
   Montre-moi [l'oie que les lapins poursuivent]
   'Show me the goose that the rabbits chase.'

   c. OVS relative clause (Object RC)
   Montre-moi [les lions que bat le cheval]
   Show me the lions that hits the horse.
   'Show me the lions that the horse hits.'

All sentences included transitive reversible verbs. The relative head and the other argument in the RCs were always mismatched in number; one in the singular and the other in the plural, with singular and plural being counterbalanced. All RCs included a verb inflected for either third person singular or third plural agreement (counterbalanced across items) and plural agreement was always phonetically audible. Sentence length was similar across items. We used verbs with audible agreement to make this experiment parallel to the Italian one, where agreement is always audible. We acknowledge that in Schelstraete and Degand's (1998) experiment agreement was not audible and this may be a limitation, although in unpublished work we found that there is no difference between relatives with or without audible agreement. Experimental items were interspersed with 12 fillers, which were either subject RCs with intransitive verbs (3 items) or simple descriptions of individuals (e.g., *Les hommes sur les arbres* 'The men on the trees'). All the sentences were recorded by a female native speaker of French. All the sentences were introduced by *show me* . . . and were spoken along with a picture displaying three sets of characters, as in Figure 1.

Figure 1. Sample of stimuli used in Experiment 1.

For instance, (4c) appeared together with the event represented in Figure 1, involving a horse in the center that beats two lions on its left and that is bitten by two lions on its right. Therefore, for each picture, the target response was either the character(s) on the left or that/those on the right, and never the one in the center. Left and right were counterbalanced across pictures. Correct responses for fillers involved, instead, the character in the center. The relevance of these pictures, as compared to classical pictures used in experiments testing comprehension of RCs and involving only two sets of characters, is that they make relativization pragmatically felicitous: two sets of identical characters are depicted that can be singled out by the relative clause (Crain and Thorton 1998).

3.3. Procedure

Before the experiment started, children's lexical knowledge of the nouns and verbs used in the experiment was assessed. Then, children went through a familiarization session during which they heard intransitive RCs and received feedback on their responses. The experiment was run on a computer and stimuli were presented with E-Prime 1.0 (Psychology Software Tools, Sharpsburg, PA, USA). Responses were recorded via a touch screen. The experiment took place in a quiet room in the school. While pictures were displayed on the screen, children heard sentences presented binaurally via headphones. Children were instructed to carefully listen to sentences, at the end of which they were invited to touch with their finger to the character that matched the sentence they heard. Responses were registered through the E-prime software.

3.4. Results and Discussion

The means and SD of correct responses for each type of sentence are presented in Table 1. Correct responses are more frequent for subject (SRC) than for both object RCs. We observed also a difference between the two types of object RCs with a better performance for OSV than for OVS.

Table 1. Mean correct responses and standard deviations for each type of sentence in Experiment 1.

Type of RC	Mean (SD)
SRC	0.87 (0.33)
OSV RC	0.58 (0.49)
OVS RC	0.23 (0.42)

SRC: subject relative clause (RC); OSV RC: object RC with preverbal subject; OVS RC: object RC with post-verbal subject.

Performance was considered above chance when the child answered correctly to at least 6 out of the 8 items in each type of sentence, as predicted by the binomial distribution ($p = 0.01$). Looking at children's individual performance, all children are above chance in the comprehension of SRCs, less than half (5/12) were above chance in the comprehension of OSV RCs, and none were above chance in the comprehension of OVS RCs. As response accuracy is a categorical variable, we submitted our data to a series of mixed effects logit models, run in R (R Development Core Team 2011). We established whether one predictor significantly contributed to the model's fit by comparing a model including that predictor against another that did not involve it using an χ-square test (Jaeger 2008). Then, we computed the z value, based on the Wald statistic, allowing for an estimation of the statistical significance of each predictor in the model. Our analyses included sentence type as a within-participant factor and participants and items as random factors. In the analyses performed, the reference category was subject RC. Sentence type ($\chi(2) = 35.4$, $p < 0.001$ added significant information to the model. Subject RCs were easier to understand than OSV RCs (estimate = -1.92, st.err = 0.49, $Z = -3.92$, $p < 0.001$) and than OVS RCs (Estimate = -3.93, St.Err = 0.55, $Z = -7.1$, $p < 0.001$) (log-likelihood = -165.2, $N = 336$, SD for subject = 0.58, SD for item = 1.03). By changing reference category, we have been able to establish that OVS was more difficult than OSV RCs (estimate = 2.00, st.err = 0.46, $Z = 4.3$, $p < 0.001$).

When children did not provide the correct answer, they typically provided two types of errors: middle errors, which consist in choosing the middle character (either the subject or the object), and reversal errors, which consist in reversing the thematic roles. The distribution of these two types of errors in the various structures tested is illustrated in Table 2. Reversal errors were the most frequent error in object RCs, especially in OVS structures, suggesting that they were interpreted as SRCs. The level of individual analysis reveals that 6 children (out of 12) were above chance in making these errors (6 out of 8 times) when presented with an object OVS RC, showing that they consistently went for the SRC interpretation.

Table 2. Percentages and raw middle errors and reversal errors for each type of sentence in Experiment 1.

SVO SRC		OSV ORC		OVS ORC	
Middle	Reversal	Middle	Reversal	Middle	Reversal
4% (5)	8% (9)	13% (15)	26% (31)	8% (9)	68% (77)

To sum up, the results show a subject/object asymmetry, with subject RCs being comprehended more accurately than object RCs, in line with the prediction of the Ccmpetition model. However, OSV object RCs are comprehended more accurately than OVS object RCs, contrary to the predictions of the competition model. Most of the errors, especially with object OVS RCs, were reversal errors consisting in interpreting the first noun phrase as the subject and the last noun phrase as the object, which amounts to re-establishing the canonical order. This latter finding fails to replicate the results of French-speaking adults (Schelstraete and Degand 1998) but is in line with those reported for Italian-speaking children (Adani 2011; Arosio et al. 2009, Arosio et al. 2011; Volpato and Vernice 2014).

In order to preserve the competition approach, one may hypothesize that children's unexpected performance in object RCs is due to a developmental change in the ranking of cues. Schelstraete and Degand (1998) argued that OSV RCs are harder than OVS RCs for adults because the former require keeping two NPs in memory before the verb is encountered. One may entertain the possibility that the memory load involved in OSV structures is not an issue for children; however, this is not plausible given the wide evidence showing lower memory resources for children than adults. Another more plausible hypothesis is that children and adults make different use of verbal cues. Indeed, OSV and OVS object RCs differ in that the critical information that the sentence is an object relative (besides the complementizer 'que', which is common to both OSV and OVS) is the presence of the second NP (the subject) following the object in OSV sentences, while in OVS it is the agreement cue for the verb (provided that the object and the subject mismatch in number, which was the case in our materials). Therefore, one could stipulate that, whereas children grant a stronger role to the word order cue, available in OSV (NP (comp) NP V) than to the agreement cue, available in OVS (NP (comp) V NP), adults would grant a stronger role to the verbal agreement cue. This would explain why French-speaking children acquire OSV RCs earlier than OVS RCs, while French-speaking adults process OVS faster than OSV. Under that hypothesis, the developmental change consists in a change in the ranking of cues. The hypothesis, grounded in the competition model, that language development involves changes in the ranking of cues predicts that Italian-speaking children should behave like French-speaking children, which we already know is the case. It also predicts that Italian-speaking adults should behave like French-speaking adults and demonstrate easier processing of OVS than OSV RCs. If this prediction would hold, it would falsify our generalization in (2), according to which children and adults should show the same hierarchy of difficulty, such that a structure that is more difficult to understand by children, should also be slower to process by adults. Since generalization (2) holds for other cases, as discussed earlier, it is important to establish whether there are exceptions and why. Therefore, in order to examine the fate of our generalization, we carried out two experiments. Experiment 2 was conducted on French-speaking adults: we wanted to replicate the faster processing for OVS compared to OSV observed by Schelstraete and Degand (1998) and determine whether this finding may be due to the particular data analyses conducted by the authors. Experiment 3 was conducted on Italian-speaking adults, since no data were available on adults in that language.

4. Experiment 2: French Adults

4.1. Participants

We tested a group of 32 French university students of the University of Geneva. All subjects were native speakers of French, had normal vision, and received credits for their participation. Written consent was collected from them.

4.2. Materials and Design

We constructed 36 sets of three center-embedded relative clauses each, a subject RC (5a), an OSV RC (with a preverbal subject) (5b) and an OVS RC (with a postverbal subject) (5c). Each sentence started with the head of the relative clause phrase followed by the relative marker. Two prepositional phrases (PPs) were introduced before the main verb in each of the three conditions, and all sentences ended with a PP. The first PP was introduced in order to control for spillover effects. It can be expected that the most difficult structure is read slower not only at the disambiguating segment but at the following segment as well, i.e., the first PP. At the second PP, there is no reason to expect an additional effect and all structures should elicit similar reading times. The embedded verb was conjugated in the past perfect (an auxiliary plus the past participle). The main verb was conjugated in the passive present perfect. The subject and the object of the relatives always mismatched in number, and their number was counter-balanced across items. All verbs in the relative clauses were reversible and only animate nouns were used in the test items. Verbs and nouns were chosen so that no pragmatic disambiguation could favor one interpretation over the other. Examples with the segmentations used during the self-paced reading are given below.

5. a. Subject RC

[$_{NP}$ Les mannequins $_1$]	[$_{RC}$ qui avaient séduit le couturier $_2$]	[$_{PP1}$ à chaque défilé $_3$]
The models	who had seduced the designer	at each show
[$_{PP2}$ de mode $_4$]	[ont été interviewés $_5$]	[$_{PP3}$ à la TV $_6$]
of fashion	have been interviewed	on the tv

'The models who had seduced the designer at each show of fashion have been interviewed on the tv.'

b. OSV RC (Object RC)

[$_{NP}$ Les mannequins $_1$]	[$_{RC}$ que le couturier avait séduit $_2$]	[$_{PP1}$ à chaque défilé $_3$]
The models	who the designer had seduced	at each show
[$_{PP2}$ de mode $_4$]	[ont été interviewés $_5$]	[$_{PP3}$ à la TV $_6$]
of fashion	have been interviewed	on the TV

'The models who the designer seduced at each fashion show have been interviewed on TV.'

c. OVS RC (Object RC)

[$_{NP}$ Les mannequins $_1$]	[$_{RC}$ qu'avait séduit le couturier $_2$]	[$_{PP1}$ à chaque défilé $_3$]
The models	who had seduced the designer	at each show
[$_{PP2}$ de mode $_4$]	[ont été interviewés $_5$]	[$_{PP3}$ à la TV $_6$]
of fashion	have been interviewed	on the TV

'The models who the designer seduced after the fashion show have been interviewed on TV.'

We constructed three experimental lists such that each condition was represented by 12 items in each list, but participants only saw one version of each item. Each list also contained 45 fillers, 15 of which were declarative sentences, 15 right-branching RCs and 15 center-embedded RCs, containing either one inanimate NP or a thematically biased verb. Experimental sentences were divided into six segments as indicated to the right of the square brackets in (5). The critical segment containing the relative marker, the verb and the NP in the relative clause was segment 2 for all three types of relatives. We adopted this particular segmentation such that segment 2 always includes the same material across the three conditions, allowing for a direct comparison among them. An alternative, which is adopted in the literature, consists in segmenting the verb and the NP in the relative clause, such that reading times can be measured more specifically at the verb. However, this way of segmenting has the disadvantage

that whereas the verb comes before the object in the subject relative, it comes after the object in the object relatives OSV and OVS, which makes comparison across structures non-minimal and therefore problematic. Segment 3 contains the spillover region, which may or may not extend until segment 4. To make sure that participants paid attention during the experiment, they were asked to respond to a yes/no comprehension question following 75% of the experimental sentences. Participants chose the "yes" or "no" response by pressing the left and right arrow buttons on the keyboard. Half of the questions were true, half were false. These questions never concerned the thematic relations between noun phrases so no attention was drawn to the goal of the experiment. Line breaks were always placed after PP1 and PP2 to make sure that no critical region would occur at the beginning of a new a line.

4.3. Procedure

Subjects were tested individually in a quiet room. Sentences were presented in the middle of the screen on a portable computer using E-Prime. Subjects were informed about the procedure and asked to pace their way through the sentences reading as normally as possible. Subjects were requested to press the space bar only after they had read and understood each segment. The experiment started with 6 practice trials. Each sentence was preceded by an asterisk after which the first segment appeared. The space bar was used to move to the next segment; after pressing the space bar the first segment was replaced by lines (one line for each word). Reading times and response accuracy to the comprehension questions were recorded.

4.4. Results and Discussion

Reading times beyond or above 3 SD from the mean were discarded, representing 2.5% of the data. Data were log-transformed to correct for normality. In general, subjects read the sentences carefully, showing an average response accuracy of 96%. Mean reading times by segment and by Sentence Type are reported in Figure 2.

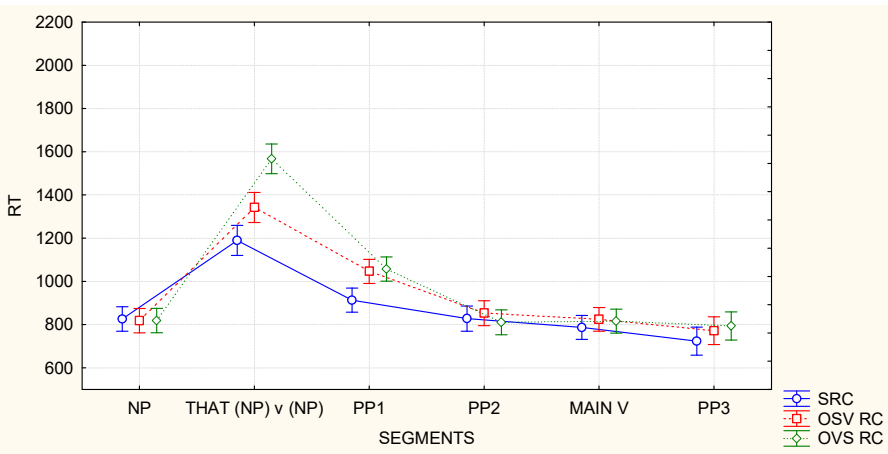


Figure 2. Mean reading times (RT) per segment in each experimental condition of Experiment 2 (French). SRC: subject relatives; OSV RC: object relatives with a pre-verbal subject; OVS RC: object relatives with a post-verbal subject. NP: noun phrase; PP: prepositional phrase. Vertical bars indicate confidence intervals at 0.95.

The data of each segment were submitted to a series of mixed effects logit models with subjects and items as random factors and sentence type and number of characters in the segment as fixed factors. The reference category for sentence type was SVO RCs. By changing it, we were also able to compare the two object RCs. At the critical segment 2, sentence type contributed to the model's fit, ($\chi^2(2) = 33.99$, $p < 0.001$), indicating that the three structures significantly differed: subject RCs were read faster than the two object RCs and OSV RCs were read faster than OVS RCs. Number of

characters also contributed significantly to the model ($\chi^2(1) = 10.52$, $p < 0.01$). That is, the higher number of characters the segment contained, the longer it took to read it. At the following segment 3, the PP1, sentence type again contributed to the model's fit ($\chi^2(2) = 16.65$, $p < 0.001$) as well as number of characters ($\chi^2(1) = 10.65$, $p < 0.01$). Subject RCs elicited shorter RTs than the two object RCs, which did not differ. No sentence type effect was found at the PP2 (segment 4) or at the main verb (segment 5). At these two segments, only number of characters contributed to the fit of the model: $\chi^2(1) = 8.65$, $p < 0.01$, $\chi^2(1) = 10.79$, $p < 0.01$, respectively. At the last segment 5, number of characters ($\chi^2(1) = 14.08$, $p < 0.01$) and sentence type contributed to the model's fit ($\chi^2(2) = 8.41$, $p < 0.05$) with subject RCs eliciting shorter RTs. In Appendix A, we report a summary of the fixed effects for those segments where sentence type contributed a significant fit and had to be included as a predictor in the model. To sum up, in line with the children's data reported in Experiment 1, French-speaking adults appear to read subject RCs faster than object RCs and OSV RCs faster than OVS RCs. The finding that OSV relatives are read faster than OVS relatives is at odds with the data reported by Schelstraete and Degand (1998). Where does the discrepancy come from? We suggest that it lies in the way reading times were measured. In our experiment, the critical region (segment 2) involved the same content across the three conditions (6b), although the linear order and the syntactic status of the NPs varied. Schelstraete and Degand (1998) used a different segmentation: the region corresponding to our segment 2 was segmented in 3 regions, as illustrated in (6a):

6. a. Schelstraete and Degand's (1998) segmentation
   OSV [RC . . . that $_2$] [NP$_{\text{-subject 3}}$] [**V**$_4$]
   OVS [RC . . . that $_2$] [V$_3$] [**NP**$_{\textbf{-subject 4}}$]

   b. Segmentation in this study
   OSV [RC . . . that NP$_{\text{-subject}}$ **V**$_2$]
   OVS [RC . . . that V **NP**$_{\textbf{-subject 2}}$]

Schelstraete and Degand (1998) found that the verb in OSV RCs (V_4) elicited slower reading times than the verb in OVS (V_3) and capitalized on that finding to conclude that OSV RCs were harder to process than OVS RCs. However, the authors also compared reading times at the subject noun, which is in segment 3 in OSV RCs (NP$_{\text{-subject 3}}$), i.e., before the verb, and in segment 4 in OVS, i.e., after the verb (NP$_{\text{-subject 4}}$). In that segment, reading times were actually longer in OVS than in OSV RCs, although the difference is not as high as that found at the verb.[3] That is, participants took longer to read the post-verbal subject in OVS than to read the pre-verbal subject in OSV. Since comprehension is incremental, it is possible that the precise locus of the difficulty differs across the two object relatives. The authors capitalized on the reading times at the verb and concluded that OVS relatives are easier than OSV relatives. However, their finding that the effect is reversed at the subject noun, which is read faster in OSV than in OVS, suggests that the reason why reading times were so fast at the verb in OVS may be that the parser did not yet have to integrate the subject as an argument of the verb or did not do anything. This integration would come at a cost, which would manifest in terms of longer reading times once the subject is reached. In contrast, in OSV the difficulty lies at the verb, since both its arguments have been encountered and therefore have to be integrated once the verb is reached.

In summary, Schelstraete and Degand (1998) found opposite effects at the verb and at the subject when comparing OSV and OVS, but capitalized only on the effect observed at the verb, to conclude that OVS are easier. By measuring reading times on a bigger region including both the verb and the subject, we were able to estimate the overall cost of having to process the OVS structure compared to the OSV structure. Before discussing the theoretical consequences of this finding, we first wanted

[3] By looking at Schelstraete and Degand's (1998) Figure 1, we can notice that the verb takes 1000 ms in OSV RC and about 650 ms in OVS RCs. The subject takes about 650 ms in OVS RCs and 580 ms in OSV RCs.

to determine whether Italian adult speakers would also process relative clauses as French-speaking adults do in order to have the full picture. This was the aim of Experiment 3.

5. Experiment 3: Italian Adults

5.1. Participants

A group of 28 Italian university students of the Università degli studi di Milano-Bicocca took part in the study. All subjects were native speakers of Italian, had normal vision, and received credits for their participation. Written consent was collected from them.

5.2. Materials and Design

The materials consisted of the Italian translation of the French materials from Experiment 2. Since the subject and the object of the relatives were always mismatched in number, the number morphology on the embedded auxiliary, which agrees with the subject, was the only cue to disambiguate SVO subject relatives from OVS object relatives since the relativizer is always *che* 'that' in Italian. This contrasts with French, in which the relativizer provides an additional cue (*qui* 'that' in SVO vs. *que* 'that' in OVS). Examples of items in the three conditions are presented in (7). Forty students of the University of Milano-Bicocca (none of which participated in the self-paced reading experiment) ranked the plausibility of the Italian items on a 1 to 5 scale. One item turned out to be significantly different from the others ($p < 0.05$), so this sentence was adjusted before running the experiment. This plausibility experiment was done before running the French experiment and the change was included in the French material.

7. a. Subject RC

[$_{NP}$ Le modelle $_1$]	[$_{RC}$ che avevano sedotto il designer $_2$]	[$_{PP2}$ dopo la sfilata $_3$]
The models	who had seduced the designer	after the show
[$_{PP2}$ di moda $_4$]	[sono state intervistate $_5$]	[$_{PP3}$ alla TV $_6$]
of fashion	have been interviewed	on the tv

'The models who seduced the designer after the fashion show have been interviewed on TV.'

b. OSV RC (Object RC)

[$_{NP}$ Le modelle $_1$]	[$_{RC}$ che il designer aveva sedotto $_2$]	[$_{PP2}$ dopo la sfilata $_3$]
The models	who the designer had seduced	after the show
[$_{PP2}$ di moda $_4$]	[sono state intervistate $_5$]	[$_{PP3}$ alla TV $_6$].
of fashion	have been interviewed	on the TV

'The models who the designer seduced after the fashion show, have been interviewed on TV.'

c. OVS RC (Object RC)

[$_{NP}$ Le modelle $_1$]	[$_{RC}$ che aveva sedotto il designer $_2$]	[$_{PP2}$ dopo la sfilata $_3$]
The models	who had seduced the designer	after the show
[$_{PP2}$ di moda $_4$]	[sono state intervistate $_5$]	[$_{PP3}$ alla TV $_6$]
of fashion	have been interviewed	on the TV

'The models who the designer seduced after the fashion show, have been interviewed on TV.'

Lists, segments definition and comprehension questions were identical to the French materials in Experiment 2.

5.3. Procedure

The same procedure as Experiment 2 was adopted.

5.4. Results and Discussion

Reading times beyond or above 3 DS from the means were discarded, representing 2% of the data. Raw data were log-transformed to correct for normality. In general, subjects were accurate in answering to the comprehension questions, displaying an overall response accuracy of 97%. Mean reading times (RTs) for each segment by sentence type are reported in Figure 3.

The data were submitted to mixed logit models with subject and item as random factors and sentence type and number of characters as fixed factors. The reference category for sentence type was subject RCs. By changing it, we have been able to compare the two object RCs.

At segment 1, only numbers of characters contributed to the fit of the model: $\chi^2(1) = 22.37$, $p < 0.001$. At the critical segment 2, sentence type contributed to the model's fit ($\chi^2(2) = 34.6$, $p < 0.001$): subject RCs were read faster than the two object RCs, and OSV object RCs were read faster than OVS object RCs. Number of characters ($\chi^2(1) = 7.34$, $p < 0.01$) also contributed to the model's fit.

In all the following segments, only number of characters contributed to the fit of the models (Segment 3: $\chi^2(1) = 33.45$, $p < 0.01$; segment 4: $\chi^2(1) = 21.09$, $p < 0.01$; segment 5: $\chi^2(1) = 36.90$, $p < 0.01$; segment 6: $\chi^2(1) = 28.89$, $p < 0.01$). Sentence type was never significant. In Appendix B, we report a summary of the fixed effects only for segment 2, as only in that region did the sentence type factor contribute a significant fit. The findings of Experiment 3 show that Italian-speaking adults process subject relatives faster than object RCs, and that they process OSV RCs faster than OVS RCs. The data are in line with the results from Experiments 1 and 2 for French-speaking children and adults, as well as with the data reported in the literature for Italian-speaking children. Experiments 1, 2 and 3 all invalidate the prediction of the competition model that OVS sentences should be easier given the reduced memory demands (only one argument has to be kept in memory until the verb is reached). Moreover, results from Experiments 2 and 3 also invalidate the hypothesis, developed within the framework of the competition model, that children and adults grant different ranks to cues. Rather, the data are compatible with our generalization in (2), according to which processing time in adults reflects order of acquisition in children. In the next section, we develop an alternative account of our results in line with that generalization.

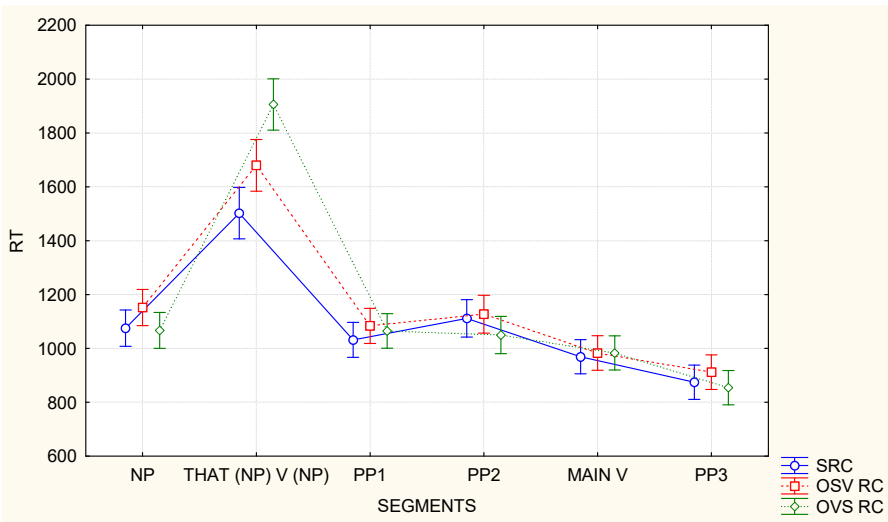


Figure 3. Mean RTs per segment in each experimental condition of Experiment 3 (Italian). Vertical bars indicate confidence intervals at 0.95.

6. Discussion

We carried out three experiments aiming at examining how French-speaking children as well as French- and Italian-speaking adults understand and process relative clauses. In particular, we were interested in the relative difficulty in the processing and acquisition of OSV and OVS relatives in

these two languages. Our starting points were (i) the generalization that structures that are more difficult to process for adults are mastered later by children (Phillips and Ehrenhofer 2015), and (ii) the findings by Schelstraete and Degand (1998) that French-speaking adults process OVS object relative clauses more easily than OSV ones. The authors took their finding as evidence for an approach of relative clause processing in terms of the competition model, according to which the processing of OSV should be harder due to higher memory costs. Specifically, they claimed that OSV RCs are harder than OVS because two arguments have to be kept in memory before the verb is reached (Bates and MacWhinney 1987). In contrast, our results from three experiments all point to a greater difficulty in the processing (adults) or comprehension (children) of OVS compared to OSV relatives. This finding was consistently found in 6;5 year-old French-speaking children and in French- and Italian-speaking adults and replicates what was previously established for Italian-speaking children in several studies using different tasks (Adani 2011; Arosio et al. 2009; Arosio et al. 2012). We argued that the conclusions by Schelstraete and Degand (1998) were based on the fact that (1) they compared smaller segments, involving only the verb, that critically differed in their position in OSV and OVS sentences, and (2) they focused on only one aspect of their results, namely reading times at the verb (longer for OSV than OVS), neglecting another aspect, reading times at the subject, which pointed to the opposite direction (longer reading times for OVS than for OSV). By measuring reading times on a wider segment including both the verb and the subject (preverbal in OSV, postverbal in OVS), as we did here, we were able to overcome the problem of comparing different segments and capitalize on an overall measure of difficulty for the two object RCs. This led us to conclude that OVS sentences are globally harder to process than OSV sentences.

Our finding that OSV ORCs are easier to process than OVS ORCs does not align with the prediction of the competition model. The fact that it is observed for both children and adults, of both French and Italian, together with the well-established finding that both children and adults struggle with object relatives more than with subject relatives, suggests that common explanatory factors underlie both OSV and OVS relative clause processing in the two age groups. What are these factors? Our data support a parsing account whereby the parser prefers to assign a subject relative clause analysis (Minimal Chain Principle, De Vincenzi 1991). That is, at the relative head, the subject role is assigned to the relative head (as this involves the shortest dependency). This analysis is abandoned and revision must start, when some material incongruent with the preferred analysis is encountered. This is the embedded subject in OSV relatives, and the embedded verb in OVS relatives.[4] The revision process is

---

[4] An alternative proposal by Friedmann et al. (2009) holds that object RCs are more difficult than subject RCs because of the intervention of the DP subject between the relative head and its copy or the position it has moved from (and is interpreted) indicated in brackets (the gap) in (i).

i.	Show me	[the goose	that	**the rabbits**	chase	<the goose>]
		+R+NP		+NP		+R +NP

In particular, they argue that the difficulty arises when the intervening subject shares a subset of the features of the relative head. In (i), the relative head has the feature +R +NP (R for relative feature and NP for descriptive content) and the intervening subject has a subset of these features (+NP). To interpret (i), a chain has to be established between the relative head and the position where it comes from. Following relativized minimality (Rizzi 1990), this chain can be established provided there is no closer antecedent to the gap that qualifies for this dependency (or has the same features). In (i), a closer antecedent could be the embedded subject, which, however, cannot qualify because it does not have the exact same features as the relative head: the subject (+NP) only has a subset of the head's features (+R, +NP). Friedmann et al. (2009) hypothesize that adults have the resources to compute such a subset relation and conclude that the subject is not the antecedent, although this comes at a cost (in terms of reaction times). That is, the greater processing time to read an OSV object relative compared to a subject relative is argued to be due to the computation of the subset relation. Young children's inability to comprehend OSV object relatives is argued to be due to the fact that children sometimes fail to compute the subset relation. The authors also show that if the relation between the head and the intervening subject is one of disjunction, in which the two elements are maximally distinct, computation succeeds, as for example in Hebrew free relatives. This analysis is more difficult to apply to relatives with postverbal subjects, as shown in Guasti (2017; see Guasti et al. 2012). In addition, data from Greek that we discuss later in the text suggest that morphological case matters but its ameliorating role cannot easily be integrated into Friedmann et al. (2009) proposal. In fact, we would have to postulate that the relevant NPs have morphological case features and that these enter into the computation carried out to decide whether an antecedent qualifies or not for a given relation. However, morphological case has no status in linguistic theory and all NPs have case, whether abstract or lexically realized.

modulated by the mode of disambiguation or the type of information encountered. Following Fodor and Inoue's (2000) diagnosis model (see also Bader and Meng 1999; Arosio et al. 2009), we propose that OSV relatives are easier to process than OVS relatives because, whereas the former ones involve positive information as to how the initial parse has to be revised, the latter does not. In OSV relatives, the subject, situated in the preverbal canonical subject position, informs the parser that a subject RC analysis is not correct and it also indicates what expression must be the subject: the preverbal NP itself. The additional time needed to process this structure, as compared to subject relatives, is the time needed to reassign the thematic role of the subject (see also Frauenfelder et al. 1980). The processing cost observed by Schelstraete and Degand (1998) at the verb in OSV may reflect a spillover effect of the reanalysis initiated at the subject of the relative clause. In OVS, instead, the information that comes after the first noun phrase (despite the complementizer) is the verb. Here, verbal agreement provides information for reanalysis, if the verb fails to agree with the initial NP. In that case, the information comes in a negative form: it tells the parser that the relative head is not the subject, but it does not tell the parser what expression the subject is. The parser must look for a subject in this case. In other words, the parser knows that the analysis is wrong, but no information is provided about how to repair the initial parse. It is only at the postverbal NP that the parser can engage in reanalysis: if that NP agrees with the verb, it can be analyzed as its subject, and the relative head can thus be reanalyzed as the object. In other words, parsing is doubly difficult in OVS: on the one hand, the parser is left without a solution at the verb (while it has a solution at the preverbal subject, i.e., at the point, which triggers reanalysis); on the other, when a solution is found, an additional operation of agreement checking has to be performed. The fact that Schelstraete and Degand (1998) found a processing cost at the postverbal subject in OVS may reflect the fact that only when the subject is found the parser engages in a new analysis. Under this view, the course of acquisition could be explained by assuming that children also start with the simpler analysis, that of subject RCs. It is known that at young ages, children are reluctant to abandon a preferred response, i.e., they have weak inhibitory skills (Mazuka et al. 2009; Choi and Trueswell 2010). Hence, their overall difficulty with object RCs may be due to a difficulty to engage in a revision process. As they grow older, they may be better equipped to abandon the subject RC analysis, because inhibitory skills have improved. The higher difficulty they show with OVS structures may be due to the negative nature of the information that leads to the revision. This view is congruent with results from Guasti et al. (2012) on Greek and by Arosio et al. (2012) on German for children and by Meng and Bader (2000) for adult German. We illustrate this point by relaying on Greek. Guasti et al. (2012) investigated children's comprehension of subject and OVS RCs in Greek, a language that can mark nominal elements with morphological case (as German). Two types of OVS RCs were tested, illustrated in (8):

8. a.	Dhikse	mou	To	alogo	pou	kinigoun	ta	liontaria
	Show	me	the.NEUT.SG	horse	that	pull.3PL	the.NEUT.PL	lion
	'Show me the horse that the lions are pulling.'							
b.	Dhikse	mou	ti	maimou	pou	pleni	I	arkouda
	Show	me	the.ACC.FEM.SG	monkey	that	wash.3SG	the.NOM.FEM.SG	bear
	'Show me the monkey that the bear is washing.'							

(8a) is like the OVS RC that we tested in French and Italian: the two nouns are marked with neuter case (specifically on the article), which is compatible with a subject or an object role. The information that indicates that the sentence is an object RC is verbal agreement: the verb does not agree with the relative head. This is a piece of negative information. (8b), instead, has the relative head morphologically marked with accusative case and the postverbal NP marked with nominative case. This information provides positive evidence that the preverbal NP is the object (it has accusative case) and the postverbal NP is the subject (it has nominative case). Guasti et al. (2012) found that 5-year-old Greek-speaking children were better at comprehending subject than OVS RCs, as in Italian. In addition, they found that the same children comprehended OVS RCs as in (8b) better than OVS RCs as in (8a). The fact that subject RCs

are easier indicates that children engage in a subject RC analysis (even if the relative head is marked accusative case (by the main verb). The difference between the two types of OVS RCs can be explained if we assume that morphological case is a positive piece of information that tells the parser who the subject and object are, while verbal agreement is a negative piece of information, that signal that the preferred subject relative clause analysis is wrong, but still requires the parser to find the subject and check agreement. On a similar vein, Meng and Bader (2000), based on German, found that adults are quicker at revising wh-questions disambiguated by morphological case than those disambiguated by agreement. Notice that this fact further supports our generalization (2), in that it also shows that children and adults are similarly affected by the type of information that disambiguate a given sentence.

So far, we have capitalized on the similarities between Italian and French RCs. However, there is one relevant difference. The complementizer in French RCs tells one whether a subject or an object RCs is to be expected. If this piece of information would immediately have been taken into account, we should not have observed a subject/object asymmetry. Whether or not they did at some point, we cannot tell on the basis of our data.[5] In any event, the finding that even adults still struggle with object relatives in French suggests that the information provided by the complementizer is actually not, or only weakly, taken into account. One reviewer notices that the reading times are generally higher all over the segments in Italian than in French and wonders whether this depends on the fact that, in French, the complementizer signals what type of relative clause is to be expected. We cannot exclude this, but we find it unlikely that the different reading times depend on the different nature of the complementizers in the two languages, because reading times are already shorter in French, at the first NP before the complementizer. In order to find out if and how the information concerning the complementizer is used by the parser, we would need to carry out an experiment with German relatives and manipulate the information on the complementizer. In (9a), the complementizer *den* ('that'+ACC) indicates that the relative head is an object. In (9b), the complementizer *das* ('that'+NEUTER) can be nominative or accusative. It is only at the embedded subject *den Mann* that the sentence is disambiguated. If the information on the complementizer is immediately used, (9a) should be easier or elicit shorter reaction time than (9b).

9.	a.	Zeig	mir	den	Mann,	den	am Montag	das	Kind
		Show	me	the.ACC	man	that. ACC	on Monday	the.NEUTER	child
		getroffen hat							
		found has							
		'Show me the man that the child found on Monday.'							
	b.	Zeig	mir	das	Kind,	das	am Montag	der	Mann
		Show	me	the.NEUTER	Child	that.NEUTER	on Monday	the.NOM	man
		getroffen hat							
		found has							
		'Show me the child that the man found on Monday.'							

In conclusion, we have shown that French- and Italian-speaking children and adults behave similarly when they have to comprehend subject and object RCs. In particular, beyond the well-known subject advantage, we have provided evidence that OSV RCs are easier than OVS RCs for adults to process and for children to comprehend. We have also contributed to reinterpreting the previous report on French-speaking adults by Schelstraete and Degand (1998) that suggested that OVS structures were easier to process: a fine analysis of the data suggests that their conclusion was invalid. Although based on different type of measures (reading time and accuracy), our data converge toward the conclusion

[5] We observed that, in adults, the revision process spanned several segments in French, but not in Italian. This is likely due to the fact that there are differences in the verbal system between Italian and French. For example, agreement is always audible in Italian, but not in French.

that children and adults are more likely to abandon a preferred analysis (subject RC analysis) when they encounter positive evidence as to what they have to do, i.e., the information that leads to abandon the preferred analysis also indicates how to revise it, rather than when they encounter a negative piece of information that only indicates that the preferred analysis is wrong. The difference between adults and children can be traced back to the ability to inhibit a preferred response. Children are known to have weak inhibitory skills. However, when they are able to abandon a preferred analysis, for them, as for adults, it is less difficult if they can rely on information that indicates what to do.

Author Contributions: M.T.G. conceived and designed the experiments with the help of J.F.; students performed the experiments; M.V. and M.T.G. analyzed the data; M.T.G. author wrote the paper with the help of J.F.

Funding: This research received no external funding.

Acknowledgments: We would like to thank the child and adult participants in the study. We would also like to thank the students who collected the data.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Summary of the Fixed Effects of the Mixed Logit Models in Segments 2, 3 and 6 of Experiment 2 (French)

Table A1. Summary of the analyses.

Predictor	Coefficient	SE	df	*t* Value	*p*
		Segment 2 (that (NP) V (NP))			
		Log-likelihood = −4939.8 (*N* = 1114)			
Intercept	57.35	8.1	80.3	8.25	<0.01
		Sentence (ref. cat. = subject RC)			
OSV	3.50	1.38	1046.3	2.54	<0.01
OVS	8.05	1.39	1060.8	5.76	<0.001
		Sentence (ref. cat. = OSV)			
OVS	4.52	1.38	1058	3.26	<0.001
		Subjects and items had SD of 15.14 and 6.34 respectively.			
		Segment 3 PP1			
		Log-likelihood = −4802.7 (*N* = 1105)			
Intercept	71.10	3.72	55.7	19.08	<0.001
		Sentence (ref. cat. = subject RC)			
OSV	4.47	1.3	1038.3	3.44	<0.01
OVS	4.76	1.3	1038.7	3.64	<0.001
		Sentence (ref. cat. = OSV)			
OVS	0.28	1.3	1038	0.22	0.82
		Subjects and items had SD of 10.16 and 4.03 respectively.			
		Segment 6 PP3			
		Log-likelihood = −3333.8 (*N* = 819)			
Intercept	53.91	4.55	48.7	11.83	<0.001
		Sentence (ref. cat. = subject RC)			
OSV	2.5	1.12	771.2	2.25	<0.05
OVS	3.06	1.13	763.9	2.71	<0.01
		Sentence (ref. cat. = OSV)			
OVS	0.51	1.12	769.7	0.46	0.64
		Subjects and items had SD of 9.64 and 5.58 respectively.			

Appendix B. Summary of the Fixed Effects of the Mixed Logit Models in Segment 2 of Experiment 3 (Italian)

Table A2. Summary of the analyses.

Predictor	Coefficient	SE	Df	*t* Value	*p*
		Segment 2 (that (NP) V (NP)			
		Log-likelihood = −7569.8 (*N* = 1023)			
Intercept	712.8	203.20	71.4	3.5	<0.001
		Sentence (ref. cat. = Subject RC)			
OSV	70.99	28.5	965.4	2.48	<0.001
OVS	169.9	28.32	963.7	5.99	<0.001
		Sentence (ref. cat. = OSV)			
OVS	98.92	28.44	963.3	3.47	<0.001
		Subjects and items had SD of 234.93 and 91.16 respectively.			

References

Adams, Catherine. 1990. Syntactic comprehension in children with expressive language impairment. *British Journal of Disorders of Communication* 25: 149–71. [CrossRef] [PubMed]

Adani, Flavia. 2011. Re-thinking the acquisition of Relative Clauses in Italian: Towards a grammatically-based account. *Journal of Child Language* 22: 141–65. [CrossRef] [PubMed]

Adani, Flavia, Heather K.J. Van der Lely, Matteo Forgiarini, and Maria Teresa Guasti. 2010. Grammatical Feature Dissimilarities Make Relative Clauses Easier: A Comprehension Study with Italian Children. *Lingua* 120: 2148–66. [CrossRef] [PubMed]

Adani, Flavia, Matteo Forgiarini, Maria Teresa Guasti, and Heather K.J. Van Der Lely. 2014. Number dissimilarities facilitate the comprehension of relative clauses in children with (Grammatical) Specific Language Impairment. *Journal of Child Language* 41: 811–41. [CrossRef] [PubMed]

Arnon, Inbal. 2005. Relative clause acquisition in Hebrew: Toward a processing-oriented account. In *BUCLD 29: Proceedings of the 29th annual Boston University Conference on Language Development*. Edited by Alejna Brugos, Manuella R. Clark-Cotton and Seungwan Ha. Somerville: Cascadilla Press, vol. 1, pp. 37–48.

Arosio, Fabrizio, Flavia Adani, and Maria Teresa Guasti. 2009. Grammatical Features in the comprehension of Italian relative clauses by Children. In *Merging Features: Computation, Interpretation and Acquisition*. Edited by José M. Brucart, Anna Gavarró and Jaume Solà. Oxford: Oxford University Press, pp. 138–58.

Arosio, Fabrizio, Maria Teresa Guasti, and Natale Stucchi. 2011. Disambiguating Information and Memory resources in Children's Processing of Italian Relative Clauses. *Journal of Psycholinguistic Research* 40: 137–54. [CrossRef] [PubMed]

Arosio, Fabrizio, Kazuko Yatsushiro, Matteo Forgiarini, and Maria Teresa Guasti. 2012. Morphological information and memory resources in the acquisition of German Relative Clauses. *Language, Learning and Development* 8: 340–64. [CrossRef]

Bader, Markus, and Michael Meng. 1999. Subject-object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research* 28: 121–43. [CrossRef]

Bates, Elizabeth. 1999. Processing complex sentences: A cross-linguistic study. *Journal of Language and Cognitive Processing* 14: 69–123. [CrossRef]

Bates, Elizabeth, and Brian MacWhinney. 1987. Functionalism and the competition model. *The Crosslinguistic Study of Sentence Processing* 3: 73–112.

Belletti, Adriana, and Cristiano Chesi. 2011. Relative clauses from the input: Syntactic considerations on a corpus-based analysis of Italian. In *Studies in Linguistics*. Siena: University of Siena.

Belletti, Adriana, and Carla Contemori. 2010. Intervention and attraction. On the production of subject and object relatives by Italian (young) children and adults. In *Language acquisition and development. Proceedings of GALA 2009*. Edited by João Costa, Ana Castro, Maria Lobo and Fernanda Pratas. Cambridge: Cambridge Scholars Press.

Betancort, Moises, Manuel Carreiras, and Patrick Sturt. 2009. Short article: The processing of subject and object relative clauses in Spanish: An eye-tracking study. *Quarterly Journal of Experimental Psychology* 62: 1915–29. [CrossRef] [PubMed]

Booth, James R., Brian Mac Whinney, and Yasuaki Harasaki. 2000. Developmental Differences in Visual and Auditory Processing of Complex Sentences. *Child Development* 71: 981–1003. [CrossRef] [PubMed]

Caplan, David, Sujith Vijayan, Gina Kuperberg, Caroline West, Gloria Waters, Doug Greve, and Anders Dale. 2002. Vascular responses to syntactic processing: Event-related fMRI study of relative clauses. *Human Brain Mapping* 15: 26–38. [CrossRef] [PubMed]

Carminati, Sara, Maria Teresa Guasti, Hans Schadee, and Claudio Luzzatti. 2006. Subject and object relative clauses in Italian: Normal subjects and an agrammatic patient. *Brain and Language* 99: 164–65. [CrossRef]

Casalis, Séverine, Christel Leuwers, and Heather Hilton. 2013. Syntactic comprehension in reading and listening: A study in French dyslexic children. *Journal of Learning Disabilities* 46: 201–19. [CrossRef] [PubMed]

Choi, Youngon, and John C. Trueswell. 2010. Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology* 106: 41–61. [CrossRef] [PubMed]

Cohen, Laurent, and Jacques Mehler. 1996. Click monitoring revisited: An on-line study of sentence comprehension. *Memory & Cognition* 24: 94–102.

Contemori, Carla, and Maria Garraffa. 2010. Comparison of modalities in SLI syntax: A study on the comprehension and production of non-canonical sentences. *Lingua* 120: 1940–55. [CrossRef]

Cooke, Ayanna, Edgar Zurif, Christian DeVita, David Alsop, Phyllis Koenig, John Detre, James Gee, Maria Piñango, and Murray Grossman. 2002. Neural basis for sentence comprehension: Grammatical and short-term memory components. *Human Brain Mapping* 15: 80–94. [CrossRef] [PubMed]

Crain, Stephen, and Rosalind Thorton. 1998. *A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge: MIT Press.

De Villiers, Jill G., Helen B. Tager Flusberg, Kenji Hakuta, and Michael Cohen. 1979. Children's comprehension of relative clauses. *Journal of Psycholinguistics Research* 8: 499–518. [CrossRef]

De Vincenzi, Marica. 1991. *Syntactic Parsing Strategies in Italian*. Dordrecht: Kluwer Academic.

Diessel, Holger, and Michael Tomasello. 2000. The development of relative clauses in spontaneous child speech. *Cognitive Linguistics* 11: 131–52. [CrossRef]

Durrleman, Stephanie, Theodoros Marinis, and Julie Franck. 2016. Syntactic complexity in the comprehension of wh-questions and relative clauses in typical language development and autism. *Applied Psycholinguistics* 37: 1501–27. [CrossRef]

Fodor, Janet Dean, and Atsu Inoue. 2000. Syntactic features in reanalysis: Positive and negative symptoms. *Journal of Psycholinguistic Research* 29: 25–36. [CrossRef] [PubMed]

Ford, Marily. 1983. A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior* 22: 203–18. [CrossRef]

Frauenfelder, Ulrich, Juan Segui, and Jacques Mehler. 1980. Monitoring around the Relative Clause. *Journal of Verbal Learning and Verbal Behaviour* 19: 328–37. [CrossRef]

Frazier, Lyn. 1987. Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory* 5: 519–59.

Friedmann, Naama, and Rama Novogrodsky. 2004. The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language* 31: 661–81. [CrossRef] [PubMed]

Friedmann, Naama, Adriana Belletti, and Luigi Rizzi. 2009. Relativized Relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119: 67–88. [CrossRef]

Guasti, Maria Teresa. 1996. On the controversial status of Romance interrogatives. *Probus* 8: 161–80. [CrossRef]

Guasti, Maria Teresa. 2017. Language Acquisition. The Growth of Grammar, 2nd ed.MIT Press: Cambridge.

Guasti, Maria Teresa, and Anna Cardinaletti. 2003. Relative clause formation in Romance child's production. *Probus* 15: 47–88. [CrossRef]

Guasti, Maria Teresa, Chiara Branchini, and Fabrizio Arosio. 2012. Interference in the production of Italian subject and object wh-questions. *Applied Psycholinguistics* 33: 185–223. [CrossRef]

Guasti, Maria Teresa, Stavroula Stavrakaki, and Fabrizio Arosio. 2012. Crosslinguistic differences and similarities in the acquisition of relative clauses. Evidence from Greek and Italian. *Lingua* 122: 700–13. [CrossRef]

Holmes, Virginia M., and J. Kevin O'Regan. 1981. Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior* 20: 417–30. [CrossRef]

Jaeger, T. Florian. 2008. Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language* 59: 434–46. [CrossRef] [PubMed]

Just, Marcel Adam, Patricia A. Carpenter, Timothy A. Keller, William F. Eddy, and Keith R. Thulborn. 1996. Brain activation modulated by sentence comprehension. *Science* 274: 114–16. [CrossRef] [PubMed]

Kaan, Edith, Anthony Harris, Edward Gibson, and Phillip Holcomb. 2000. The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15: 159–201. [CrossRef]

Kayne, Richard, and Jean-Yves Pollock. 1978. Stylistic inversion, successive cyclicity, and Move NP in French. *Linguistic Inquiry* 9: 595–621.

Kayne, Richard, and Jean-Yves Pollock. 2001. New thoughts on stylistic inversion. In *Subject Inversion in Romance and the Theory of Universal Grammar*. Edited by Aafke Hulk and Jean-Yves Pollock. Oxford: Oxford University Press, pp. 107–62.

Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63–99.

King, Jonathan, and Marcel Adam Just. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30: 580–602. [CrossRef]

King, Jonathan W., and Marta Kutas. 1995. Who did what and when? Using word- and causal-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience* 7: 376–95. [CrossRef] [PubMed]

Mak, Willem M., Wietske Vonk, and Herbert Schriefers. 2002. The influence of animacy on relative clause processing. *Journal of Memory and Language* 47: 50–68. [CrossRef]

Mazuka, Reiko, Nobuyuki Jincho, and Hiroaki Oishi. 2009. Development of executive control and language processing. *Language and Linguistics Compass* 3: 59–89. [CrossRef]

Mecklinger, Axel, Herbert Schriefers, Karsten Steinhauer, and Angela D. Friederici. 1995. Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition 23*, 477–94.

Meng, Michael, and Markus Bader. 2000. Mode of disambiguation and garden-path strength: An investigation of subject-object ambiguities in German. *Language and Speech* 43: 43–74. [CrossRef]

Noizet, Georges, Francoise Deyts, and Jean-Pierre Deyts. 1972. Producing complex sentences by applying relative transformations: A comparative study. *Linguistics* 10: 49–67. [CrossRef]

Phillips, Colin, and Lara Ehrenhofer. 2015. The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism* 5: 409–53. [CrossRef]

R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing*. Vienna: The R Foundation for Statistical Computing, ISBN 3-900051-07-0.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge: The MIT Press.

Schelstraete, Marie-Anne, and Liesbeth Degand. 1998. Assignment of grammatical functions in French relative clauses. *Language Sciences* 20: 163–88. [CrossRef]

Schriefers, Herbert, Angela D. Friederici, and Katja Kuhn. 1995. The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language* 8: 499–520. [CrossRef]

Shirai, Yasuhiro, and Hiromi Ozeki. 2007. Introduction. *Studies in Second Language Acquisition* 29: 155–67. [CrossRef]

Stromswold, Karin, David Caplan, Nathaniel Alpert, and Scott Rauch. 1996. Localization of syntactic comprehension by positron emission tomography. *Brain and Language* 52: 452–73. [CrossRef] [PubMed]

Tavakolian, Susan. 1981. The conjoined-clause analysis of relative clauses. In *Language acquisition and linguistic theory*. Cambridge: The MIT Press, pp. 167–187.

Traxler, Matthew, Robin Morris, and Rachel Seely. 2002. Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47: 69–90. [CrossRef]

Utzeri, Irene. 2007. The production and acquisition of subject and object relative clauses in Italian. *Nazan Linguistics Special Issue* 3: 283–314.

Villata, Sandra, Whitney Tabor, and Julie Franck. 2018. Encoding and retrieval interference in sentence comprehension: Evidence from agreement. *Frontiers in Psychology* 9: 2. [CrossRef] [PubMed]

Volpato, Francesca, and Mirta Vernice. 2014. The production of relative clauses by Italian cochlear-implanted and hearing children. *Lingua* 139: 39–67. [CrossRef]

languages

MDPI

Article

Acquisition of French Causatives: Parallels to English Passives

Jason Borga [1,*] and William Snyder [2,*]

[1] Department of Linguistics, McGill University, 1085 Dr. Penfield, Montreal, QC H3A0G4, Canada
[2] Department of Linguistics, University of Connecticut, 365 Fairfield Way, Storrs, CT 06269-1145, USA
* Correspondence: jason.borga@mail.mcgill.ca (J.B.); william.snyder@uconn.edu (W.S.); Tel.: +1-514-398-4222 (J.B.); +1-860-486-0157 (W.S.)

Received: 3 December 2017; Accepted: 22 May 2018; Published: 31 May 2018

Abstract: Guasti (2016) notes similarities between English *get*- and *be*-passives, and Romance causatives of the *faire-par* and *faire-infinitif* types, respectively. On this basis she conjectures that *faire-infinitif* will show an acquisitional delay similar to that found for English *be*-passives, which are not mastered until sometime after the age of four. Here, this prediction is tested and supported for French *faire-infinitif* causatives of transitive verbs. To explain the delay, the Universal Freezing Hypothesis (UFH) of Snyder and Hyams (2015) is extended to this type of causative: a restriction on movement is recast as a restriction on AGREE. A novel prediction, that *faire* causatives involving unergative or unaccusative verbs will be acquired much earlier, is also tested and supported. Finally, English *get*-passives and French "reflexive causative passives" are examined in light of the fact that both are acquired substantially earlier than age four.

Keywords: first language acquisition; relativized minimality; causatives

1. Introduction

Children acquiring English often start using *get*-passives before they turn three, but most studies find that children have difficulties producing and understanding *be*-passives until much later ages. The details vary by study but, in general, children do not reliably succeed on the *be*-passives of eventive verbs (especially if a *by*-phrase is present) until around the fourth birthday, and they continue to have substantial difficulties with the *be*-passives of non-eventive verbs (e.g., *see*, *like*) until age five or six. Guasti (2016) conjectures that a similar acquisitional time course might exist for certain causative constructions found in Romance languages like French. Her idea is that the delay seen for English *be*-passives might also exist for French causatives of the *faire-infinitif* type.

In this paper we will first test, and strongly support, Guasti's (2016) conjecture as it applies to French causatives of transitive verbs. Next, we will propose that the Universal Freezing Hypothesis (UFH) of Snyder and Hyams (2015), which seeks to explain the delay in English *be*-passives, can be extended in a natural way to explain the observed delay in this type of causative. The proposed extension leads to novel predictions for the acquisition of causatives involving intransitive verbs, and we will provide evidence supporting these predictions. Finally, we will examine two early-acquired structures, English *get*-passives and French reflexive causative passives. The key point will be that, in both cases, the external argument associated with the lower verb does not have to be present in the syntactic structure and, as a consequence, children's early successes are fully consistent with the UFH.

2. French Causatives and Guasti's (2016) Conjecture

2.1. Causatives and Passives

Within the Generative tradition there exists an extensive literature on French *faire*-causatives, dating to Kayne's (1975) seminal analysis dividing them into two main types, *faire-infinitif* and *faire-par.* In *faire-infinitif* causatives, if the causativized verb is transitive, the causee is obligatorily realized as either a dative-marked Determiner Phrase (DP) following the embedded object, as in (1a), or as a dative clitic preceding *faire*, as in (1b):

1. a.

Jean	a	fait	laver	la voiture	à	Paul
John	has	made	to-wash	the car	DAT	Paul

'John made Paul wash the car.'

b.

Jean	lui	a	fait	laver	la voiture
John	him.DAT	has	made	to-wash	the car

'John made him wash the car.'

c.

Jean	a	fait	laver	la voiture	(par Paul)
John	has	made	to-wash	the car	(by Paul)

'John had the car washed (by Paul).'

The causee in the *faire-par* causative, however, appears in an optional *par*-phrase, which is located after the object of the lower verb, as shown in (1c).

In the present paper we will employ a somewhat larger, five-way classificatory scheme, as indicated in (2):

2. a. *Faire*-Dative (FD) = *faire* + $V_{transitive}$ with a <u>dative</u> causee
Jean a fait laver la voiture à Paul. Jean lui a fait laver la voiture.

b. *Faire-Par* (FP) = *faire* + $V_{transitive}$ with a <u>*par*</u>-phrase
Jean a fait laver la voiture par Paul.

c. *Faire*-Null (FN) = *faire* + $V_{transitive}$ with a <u>null</u> (i.e., omitted) causee
Jean a fait laver la voiture.

d. *Faire*-Unaccusative (FuA) = *faire* + $V_{\underline{unaccusative}}$ (and an accusative causee)
Jean a fait partir Paul. Jean l'a fait partir.
John has made leave Paul John him.ACC has made leave

e. *Faire*-Unergative (FuE) = *faire* + $V_{\underline{unergative}}$ (and an accusative causee)
Jean a fait danser Paul. Jean l'a fait danser.
John has made dance Paul John him.ACC has made dance

Thus, the *faire-infinitif* sentences in (1a,b) will both be referred to as FDs. The *faire-par* sentence in (1c) will be referred to as an FP if the *par*-phrase is pronounced, or an FN if the *par*-phrase is omitted.

Alongside the differences in the causee phrases, there exist a number of semantic differences across the various subtypes of causatives. Crucially, Guasti (2016) notes that certain semantic restrictions found in FP/FNs are also found in English *get*-passives: First, stative verbs (*loved*, *aimer* in 3a,b) are rejected in both *get*-passives (e.g., Hirsch and Wexler 2004) and FP/FNs:

3. a. ?* John got loved

   b. ?* Ils ont fait aimer Jean (par Marie)
   they have made love John (by Mary)
   (?*) 'They had John loved (by Mary).'

Second, Guasti (2016) notes that both *get*-passives and FP/FNs require the logical object to be [+affected]. For example, in (4a,b) *the answer/la solution* is [-affected]:

4. a. ?* The answer got found

   b. ?* Ils ont fait trouver la solution (par le chercheur)
   they have made find the answer (by the researcher)
   (?*) 'They had the answer found (by the researcher).'

These restrictions are absent in *be*-passives (5a,b), and in FDs (6a,b):

5. a. John was loved

   b. The answer was found

6. a. Ils ont fait aimer Jean à Marie
   they have made to-love John DAT Mary
   'They made Mary love John.'

   b. Ils ont fait trouver la solution au chercheur
   they have made find the answer DAT+the researcher
   'They made the researcher find the answer.'

On the basis of these parallels, Guasti (2016, p. 185) conjectures that the acquisitional time course of French causatives will mirror that of English *get*- and *be*-passives: just as *be*-passives are delayed relative to *get*-passives, FDs will be delayed relative to FP/FNs.[1]

2.2. Testing Guasti's (2016) Conjecture

In order to assess Guasti's (2016) conjecture we examined 11 longitudinal corpora from CHILDES (MacWhinney 2000) for children acquiring French in France: Anaïs, Marie, Marilyn, Nathan, Théotime (Demuth and Tremblay 2008); Anaé, Antoine, Léonard, Madeleine, Théophile (Morgenstern and Parisse 2007); and Anne (Plunkett 2002). The ages covered for the children in question are listed in Table 1, in Y;MM,DD format.

Table 1. Age Ranges of Corpora.

Child	Ages
Anaé	1;04,20–5;01,21
Antoine	1;00,24–6;03,08
Leonard	0;11,19–5;03,19
Madeleine	1;00,05–6;11,27
Théophile	1;00,09–4;11,11
Anais	1;00,23–3;00,15
Marie	1;00,02–4;00,05
Marilyn	1;06,13–2;11,14
Nathan	1;00,12–3;00,03
Théotime	0;11,17–3;00,03
Anne	1;10,12–3;05,04

[1] Strictly speaking, Guasti (2016) states her conjecture in the broader terms of *faire-infinitif* versus *faire-par* causatives, but we will initially focus on the subtypes of these causatives in which the causativized verb is transitive (hence, FDs and FP/FNs). We will turn to the causatives of intransitive verbs in Section 4.

For our analysis we first used a computer search to identify all child utterances containing any form of the verb *faire*, and then we manually removed all non-causative uses. Next we classified each of the causatives as either "Clear FD", or "Other", according to the criteria in (7) and (8) (these criteria are based directly on the discussion in Guasti 2016).

7. **Clear FD:** A causative containing an overt dative argument, and/or a transitive verb that is semantically incompatible with FP/FN (i.e., a non-eventive verb, or a verb with a direct object that is [-affected]).

8. **Other:** A causative containing an overt *par*-phrase, or a lower verb that is either intransitive or semantically compatible with being an FN.

Note that children often omit material that is obligatory for adult speakers. Hence, our criteria do not rely exclusively on the form of the causee. An utterance like (9), where the causee is omitted, would nonetheless be classified as a clear FD, since the verb *trouver* 'to find' does not satisfy the semantic constraints on FP/FNs:[2]

9. Ils ont fait trouver la solution (. . .)

Our prediction is that no instances of a clear FD will appear in the children's utterances prior to the late age-range associated with the appearance of *be*-passives: any child using clear-cut FDs prior to age four will be a counterexample.[3]

2.3. Results

Our examination of the 11 children's data found no use of a clear FD prior to age four. As expected, shortly after the age of four some of the children did begin using clear-cut FDs, as seen in (10). Additionally, as expected, other types of causatives appeared much earlier. In fact, the data included two examples of FPs (11a,b), and both were produced well before the age of three:[4]

10. Je faisais faire quoi à mes trois enfants?
I made do what DAT my three children?
'What did I make my three children do?' (Madeleine 4;01,27)

11. a. Il va [se] faire gronder par sa maman et papa
he goes [self] make scold by his mom and dad
'He's going to get (himself) scolded by his mom and dad.' (Antoine 2;09,16)

b. Elle se fait tirer par la boulle comme ça fait du bruit
she self makes pull by the ball as that makes of-the noise
'She is getting (herself) pulled (=becoming irritated) by the ball since it makes noise.' (Madeleine 2;05,12)

---

[2] Note that in principle a child might fail to respect the semantic constraints on FNs, and produce an early use of an FN that would be coded, under (7), as "clearly FD". For example, in principle a child might have a non-adult-like grammar that allowed her to say the FP/FN, *Je l'ai fait aimer (par Paul)* 'I had her loved (by Paul)'. If the *par*-phrase were omitted, this utterance would be coded as "clearly FD", since the verb *aimer* 'love' should be incompatible with an FP/FN, according to the judgments reported in the literature. Moreover, as brought to our attention by an anonymous reviewer, it seems that certain adult speakers of French may sometimes violate the claimed semantic constraints on FP/FNs. Yet, as we will see below, the children in our study did not produce anything that was coded as "clearly FD" until after the age of four. Hence, if any child in our study ever violated the semantic constraints on FP/FNs (and simultaneously omitted the *par*-phrase), the utterance in question must have occurred at too late an age to affect our results (i.e., it must have been produced after the child had turned four).

[3] We chose the age of four years (48 months) as a cut-off point because (i) studies reporting a delay in *be*-passives report that difficulties persist at least until this age; and (ii) in absolute terms, by the age of four children are remarkably successful at most aspects of their target grammar; hence, any delay beyond 48 months is notable in its own right.

[4] Interestingly, both instances of FPs were reflexive. This point will be discussed in Section 5.

Table 2 provides the results by child. For example, the child Madeleine produced her first example of a (non-FD) *faire*-causative at the age of two years, one month (2;01), and produced an additional six such utterances before her fourth birthday. Her first recorded use of an FD was at the age of 4;02, and was followed by two additional FDs before the end of her corpus. In Figure 1 we have graphed Madeleine's use of causatives as a function of her age.

Table 2. Results by Child.

	Onset of Non-FDs	#Non-FDs before Age 4	Onset of FD	#FDs (to End of Corpus)
Anaé	2;00	12	>5;01	0
Anaïs	2;08	5	>3;00	0
Anne	2;04	14	>3;05	0
Antoine	2;06	11	4;02	1
Léonard	2;01	2	>3;02	0
Madeleine	2;01	7	4;02	3 (see Figure 1)
Marie	2;05	18	>4;00	0
Marilyn	2;09	4	>2;11	0
Nathan[5]	>3;00	NA	>3;00	NA
Théophile	2;10	11	>4;11	0
Théotime	2;03	13	>3;00	0

FD = *Faire*-Dative: *faire*-causative + transitive V with dative causee.

To check for statistical significance of the observed delay in FDs, we employed an absolute-frequency binomial test, as described in Snyder (2007, chp. 5): as indicated in (12–14), we first calculated the total number (U) of child utterances in our sample that were produced after the given child was already producing both datives and (non-FD) causatives, but before the child had reached the age of four years:

12. U = 33,244 utterances

To estimate the per-utterance frequency of FDs in speakers whose grammar allows them, we analyzed all parental utterances in our 11 corpora, calculated each parent's per-utterance frequency of FDs, and then took the median, F:

13. F = (7 FDs)/(22,778 utterances by the given parent)

To calculate the probability that U would be as high as observed (or even higher), under the null hypothesis that FDs were always available to children as soon as they had both non-FD causatives and dative arguments, we calculated the binomial probability, $p = (1-F)^U$:

14. $p = 0.00003651 < 0.0001$

As seen in (14), the age gap between the onset of non-FD causatives and dative-marked arguments, on the one hand, versus the first clear FD, was robustly significant. Therefore, Guasti's (2016) conjecture (at least as it applies to FDs) is strongly supported by our longitudinal data. In the next section we look at the connection between these results and the UFH, before moving on to more fine-grained predictions regarding the acquisition of causativized intransitives.

[5] Nathan's corpus ends (at age 3;00,03) before he uses either dative arguments or causatives of any kind.

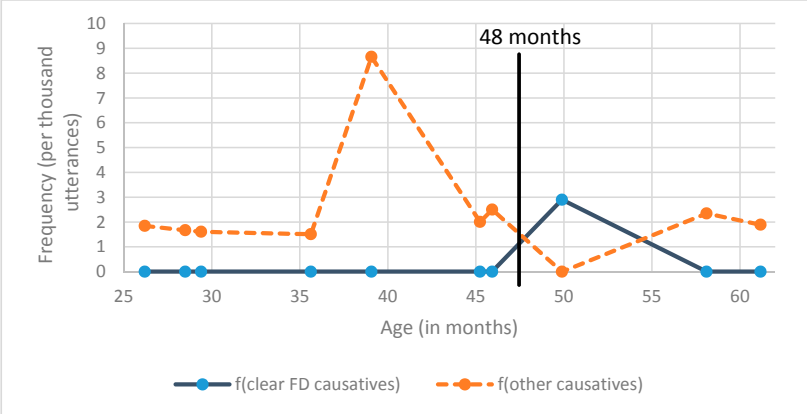


Figure 1. Frequency of Causatives by Age (Madeleine).

3. Explaining the Acquisitional Delay for *Faire*-Datives

In this section we will first review Snyder and Hyams' (2015) account of the acquisitional delay found in English *be*-passives. Then we will show that their account, the UFH, can be extended in a simple way to account for the observed delay in FDs.

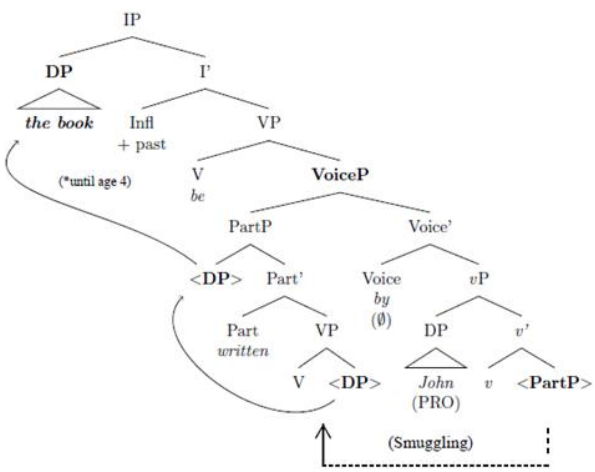
3.1. English Be-Passives and the Universal Freezing Hypothesis

Snyder and Hyams (2015) propose that the lateness of English *be*-passives is due to a freezing effect. For the syntax of the passive, they adopt the analysis of Collins (2005a), which is based on a strict interpretation of UTAH (i.e., the Uniformity of Theta-Assignment Hypothesis; Baker 1988, p. 46; 1997, p. 74). Under Collins (2005a) interpretation, a verb's external θ-role must be assigned in exactly the same way in active and passive sentences, namely to the specifier of *v*P. In a passive the external argument is realized either as a PRO, in the case of a short passive; or as an overt DP (preceded by the overt voice-head *by*), in the case of a long passive.

Consider the *be*-passive in (15). Given the presence of the external argument *John* in Spec-*v*P, simple argument-movement of the object DP *the book* into Spec-IP would violate relativized minimality (RM; Rizzi 2001, 2004). The solution proposed by Collins (2005a) is that the object is "smuggled" past the external argument, via movement of a larger phrase, after which it raises to subject position without a minimality violation, as shown in (16).

15. The book was written by John.

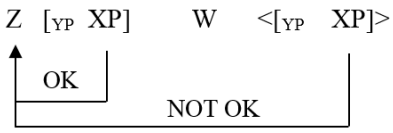
16. Cf. Collins 2005a, pp. 90, 95



In (16), the V *write* raises to the head of PartP and forms the participle. Smuggling occurs when the entire PartP, including the object, raises past the external argument in Spec-*v*P and lands in the specifier of the passive voice phrase. From here the object DP can undergo cyclic movement into the subject position in a fashion compatible with minimality.

Collins (2005b, p. 292) defines "smuggling" in terms of the illustration in (17): The constituent XP cannot be directly related to Z, due to an intervening element W. Yet, after movement of the larger constituent YP to a position from which it c-commands W, XP can be related to Z. In this case, YP "smuggles" XP past W.[6]

17. Smuggling (Collins 2005b, p. 292, ex. 6)



In the cases of smuggling discussed by Collins (2005a, 2005b), the relation between XP and Z is always one of movement: the relation created by moving XP to position Z.

The entire concept of smuggling hinges on the idea that the Freezing Principle (which was proposed independently in Ross (1967, 1986) and Wexler and Culicover (1980)) allows for some exceptions. One formulation of the principle, based on Müller (1998), is given in (18).

18. Freezing Principle: In the following configuration, no operation (such as Move) may relate X and Z: * Z . . . [$_Y$ . . . X . . .] . . . <Y>

According to this principle, the direct object inside the PartP in (16), which corresponds to the X inside Y in (18), should be frozen to further syntactic operations, and therefore unable to undergo raising into subject position. Hence, Collins (2005a, 2005b) concludes that there must exist situations in which the Freezing Principle fails to apply, although he does not attempt to spell these out in a general theory.

Finally, the role of smuggling in the late acquisition of the *be*-passive, according to Snyder and Hyams (2015), is connected to a difference in the capacity of adults and children under four to make the required exception to the Freezing Principle. In children that young, the Freezing Principle blocks smuggling, and thereby blocks the *be*-passive: without smuggling, the underlying object cannot get past the intervener that separates it from subject position. Snyder and Hyams (2015) formulate this hypothesis as the UFH, as stated in (19).

19. The Universal Freezing Hypothesis: For the immature child (until about age four), the Freezing Principle always applies. No subpart of a moved phrase can ever be extracted.

The UFH thus accounts for the observed delay in children's mastery of *be*-passives, which depend on smuggling to circumvent a Minimality violation.

3.2. The Universal Freezing Hypothesis and the Locus of Maturational Change

As formulated above, the UFH describes a change in the availability of exceptions to the Freezing Principle, but it does not address why such a change would occur. The answer presumably depends on the precise nature of the maturational change that the UFH describes. For example, one possibility is that there are essentially two Freezing Principles provided by Universal Grammar (UG). Under this

---

[6] Collins (2005a, 2005b) does not require the movement of YP in (17) to be feature-driven, except insofar as XP may need to enter into a feature-checking relation with Gehrke and Grillo (2009) have worked out a possible feature-driven version of smuggling for the case of English *be*-passives, in which a VP-shell is attracted to Spec-VoiceP by a feature of passive voice. Snyder and Hyams (2015) assume that some version of this feature-based approach can be maintained, and we will make the same assumption here.

view, a "maximally restrictive" version of the Freezing Principle, which is active in the grammar of very young children, is replaced by a "selective" version sometime around the age of four.

Yet, this is not the approach that we advocate. For one thing, there are reasons to doubt that the Freezing Principle itself (whether selective or restrictive) is actually a primitive element of UG. For example, Uriagereka (1999) argues that freezing effects can be made to follow as a deductive consequence from mechanisms that are independently needed for cyclic spell-out. Alternatively, Keine (2016) argues that they can be derived from limitations on the search space of probes.

Furthermore, in Borga and Snyder (forthcoming) we review recent experimental evidence suggesting that an absolutely restrictive version of the Freezing Principle would be too strong, even for children younger than four. As an alternative, we propose that the syntactic structures resulting from a smuggling-based derivation impose excessive demands on the child's computational resources for language processing. This proposal accommodates recent evidence suggesting that children's ability to produce and comprehend *be*-passives improves in experiments that either (i) eliminate the need for the use of a smuggling derivation in the first place, for example by adding a feature like [+*wh*] or [+Topic] to the derived subject (cf. Rizzi 2004); or (ii) take steps to reduce the processing load that a smuggling-based structure creates, for example by providing a structural prime for the *be*-passive (e.g., Messenger 2010). In the case of (i), information structure is manipulated via context provided in the experimental task, resulting in a derived subject which is either discourse old ([+Topic]) or which undergoes *wh*-movement, avoiding Minimality violations and the need for smuggling in either case (cf. footnote 8). As regards (ii), one analysis of syntactic priming studies is that the comprehension of a sentence can aid in activating its underlying syntactic representation, which can then be re-used in further production and comprehension, at a reduced processing cost (Guasti 2016, p. 203). The point which follows from both observations is that the change taking place around the age of four is not a change in the grammar itself, but rather in the level of computational complexity (and, hence, grammatical complexity) that the child's language processing system can handle.[7]

3.3. Extending the Universal Freezing Hypothesis to Faire Datives

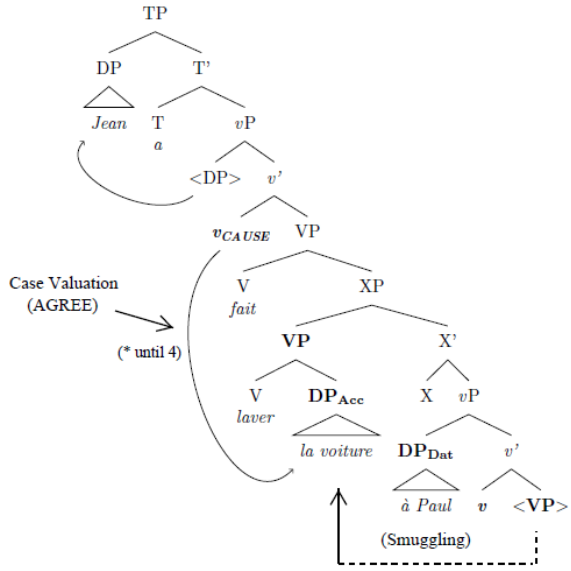
Despite the existence of parallels between English passives and Romance causatives, it is not immediately clear that Snyder and Hyams (2015) freezing-based approach to passives can account for the acquisitional timing of causatives. The maturational delay described in the UFH appears to be tied to DP movement, which does not play any role in standard analyses of *faire*-causatives (e.g., Kayne 1975; Rouveret and Vergnaud 1980; Burzio 1986). Yet, there is reason to think a form of smuggling might also be needed in certain causatives.

A number of authors, going back to Kayne (1975), have proposed that a derivation along the general lines of (20) (which is based on a discussion in (Belletti and Rizzi 2012)) is found in French and Italian FDs.

---

[7] As noted in Borga and Snyder (forthcoming), it is not entirely straightforward to harmonize the assumptions of (Collins (2005a, 2005b) for smuggling, with those of (Rizzi (2004)) for Relativized Minimality (RM). According to this version of RM, in a configuration [. . . X . . . Z . . . Y], a local relation cannot hold between X and Y if Z belongs to the same structural type as X, where structural type is determined in terms of feature classes. Under Rizzi's (2004) assumptions it should be possible to derive an English *be*-passive without risk of an RM violation and, hence, without smuggling, when the logical object bears a feature such as [+Topic] or [+WH]. Yet, on the assumptions of Collins (2005a, 2005b), without smuggling the surface word order comes out wrong (e.g., *The book was by John written*). Borga and Snyder (forthcoming) very briefly sketch an approach that might resolve this conflict. In particular, it would allow a child to avoid the need for a smuggling derivation altogether, when the logical object bears one of the aforementioned features.

20. FD Cf. Belletti and Rizzi 2012



In (20) we have labeled as 'smuggling' the step of the derivation in which the VP headed by *laver* 'wash' is moved past the external argument, *à Paul* 'to Paul'. (In the version of this analysis sketched by Belletti and Rizzi (2012, p. 135), the VP undergoes feature-driven movement to the specifier of a functional head, X, which the authors, suggest is the head of a "(small) clausal complement" to the causative verb).

As emphasized by Belletti and Rizzi (2012), this is an instance of smuggling, even though nothing moves out of the moved constituent. The VP containing the lower verb's direct object moves past the intervening dative causee, into a position where the direct object can be case-valued by v_{cause}.[8] To repeat the definition from Section 3.1, stated in terms of the diagram in (17) (repeated as (21), with the elements from the FD inserted), smuggling is the movement of a constituent YP (here the VP containing the lower object) into a position where XP (the lower object) can be "related to" Z (the higher functional v_{CAUSE} head).

21. Smuggling in the *faire dative:*

v_{CAUSE} $[_{VP}$ $DP_{Acc}]_i$ $[DP_{Dat}]$ $<[_{VP}$ $DP_{Acc}]_i>]$

OK NOT OK

In the cases of smuggling discussed in Collins (2005a, 2005b), the relation between XP and Z was always one of movement, but if we take movement to require the establishment of an AGREE relation, then the crucial relation between XP and Z can be seen as one of AGREE. Correspondingly, in (20) the crucial relation is an instance of AGREE that simply values a case feature of XP, with no resulting movement.

---

8 The nature of the dative case on the transitive causee is the subject of a number of competing analyses, including those where the causee argument is located in an additional Applicative projection (Pylkkänen 2008). Belletti and Rizzi's (2012) smuggling approach is consistent with any analysis in which the causee argument is generated in a position intervening between the case-assigning causative head and the object in need of case valuation. Bellucci (2015) proposes an analysis based on Manzini and Savoia's (2011) analysis of oblique case as interpretable; in this case the dative-causee remains in the specifier of *v*P. In Kayne's (2004) analysis, the causee is still merged in Spec-*v*P initially, and *à* itself is responsible for checking dative case; the causee Goal must raise to establish an Agree relation with the prepositional Probe, and the VP chunk containing the object still must raise to a position local to the accusative case-assigning causative head. Both analyses make the same acquisitional prediction regarding the need for smuggling of the object-containing VP.

Hence, to extend the UFH to FDs we really only need to make one change: freezing effects are not fundamentally about movement, but rather about AGREE. As a first approximation we can state this view in the form of a "Modified Freezing Principle", as in (22):

22. Modified Freezing Principle: In the following configuration, an AGREE relation (usually) cannot be established between X and Z: * Z . . . [$_{Y}$ . . . X . . .] . . . <Y>

The formulation in (22) includes the qualification "(usually)" as an explicit indication that certain exceptions will be possible for adults, but—precisely as before—the UFH will mean that these exceptions are unavailable to children younger than four. Hence, the UFH now works well for FDs: children will have substantial difficulties until the age of about four, because the lower verb's direct object cannot be case-valued without making an exception to the Modified Freezing Principle.

Indeed, potential support for an interaction between freezing and AGREE can be found in recent work by Keine (2016, 2018), who notes a type of freezing effect on long distance agreement in Hindi.[9] Agreement between a matrix verb and the object of a nonfinite embedded clause is usually optional in Hindi, as in (23), but is substantially degraded in instances where the lower clause has undergone extraposition, as in (24):

23. shiksakõ-ne [raam-ko kitaab parhne] d-ii
teachers-$_{ERG}$ Ram-$_{DAT}$ book.$_{F}$ read.$_{INF}$ let.$_{PFV.F.SG}$
'The teachers let Ram read a book.'

24. ?? shiksakõ-ne t_i d-ii [raam-ko kitaab parhne]$_i$
teachers-$_{ERG}$ let.$_{PFV.F.SG}$ Ram-$_{DAT}$ book.$_{F}$ read.$_{INF}$
'The teachers let Ram read a book.'

Thus, it appears that Hindi may provide support for the idea that freezing effects can be found in the domain of AGREE, per se, even when extraction from the moved constituent is not at issue.

In order to interpret the UFH in terms of the computational demands of language processing, as described in Section 3.2, and in terms of a maturationally-timed change in children's computational abilities, we will need one more adjustment to our assumptions. In terms of movement operations, a smuggling-based derivation requires the language processing systems to represent a chain whose "tail" is properly contained within the "head" of another chain (for example: *The book was* [*written* <*the book*>] by John <*written the book*>.) This configuration is plausibly responsible for a substantial increase in computational complexity. Yet, in order to subsume FDs under this account, we will need to assume that this configuration is a special case of a more general source of computational complexity: structures involving AGREE into a moved constituent. In other words, regardless of whether any phrase moves out of the moved constituent, simple AGREE into a moved constituent is something that is both necessary to represent, and difficult to represent, during the computations that subserve language production and comprehension.

Finally, note that our account crucially relies on the idea (shared with Snyder and Hyams (2015)) that *get*-passives can, at least in some cases, be derived without recourse to smuggling. We will discuss this point in detail in Section 5. First, however, we will derive and test a novel prediction of the proposals that we have made in the current section.

---

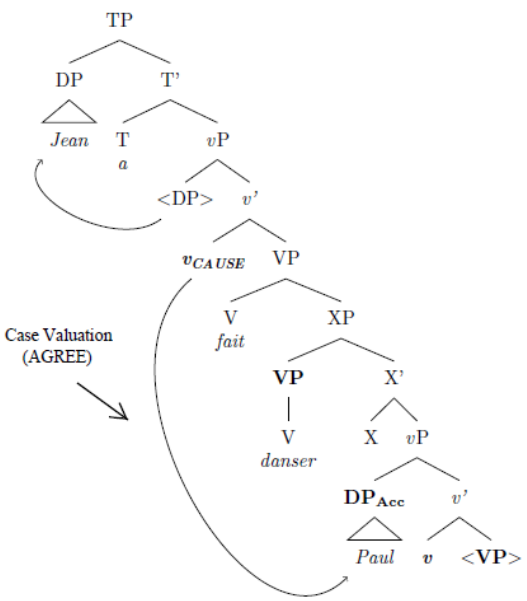
[9] We only recently learned of Keine's (2016, 2018) work, and were pleased to discover that he had independently arrived at—and extensively developed—the same general idea discussed above: namely that there should exist closely related restrictions on AGREEment into, and movement out of, a moved constituent.

4. The Universal Freezing Hypothesis and Causatives of Intransitive Verbs

4.1. New Predictions for Causativized Intransitives

Further examination of the syntax literature on Romance causatives has led us to a novel prediction: the causative of an intransitive verb, whether unaccusative (FuA) or unergative (FuE), will be within the grammatical capabilities of a two- to three-year-old child. Specifically, we note that Guasti (1996) and Folli and Harley (2007) have argued that, in terms of Kayne's (1975) distinction between *faire-infinitif* and *faire-par*, FuEs necessarily belong to the *faire-infinitif* type, with a derivation along the lines of (25):[10]

25. *Faire* + unergative = '*faire-infinitif*'



In (25), the assumption is that FuEs, like FDs, require movement of the lower VP, but there is nothing to intervene between v_{cause} and the accusative-marked causee, *Paul*. Note, too, that in terms of the Modified Freezing Principle, the UFH does not create any problems for the child under four. This is because there is no AGREE operation into the VP that underwent movement.

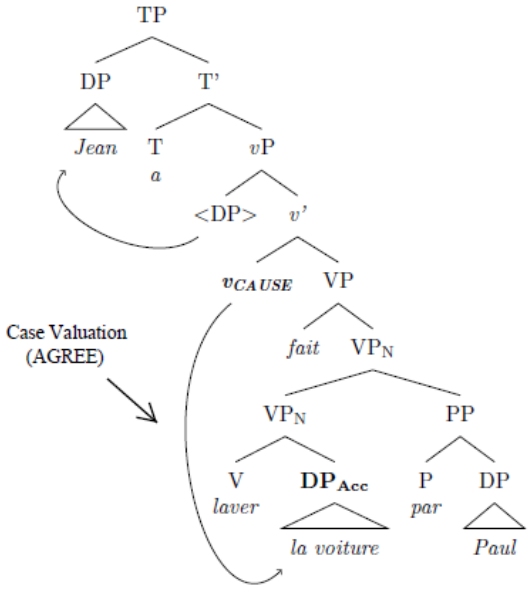
In the case of FuAs, Folli and Harley (2007) and Guasti (1996) argue that *faire*-causatives without an external argument are at least potentially instances of Kayne's (1975) *faire-par* causative (recall that, in present terms, these correspond to FP or FN causatives, depending on whether the optional *par*-phrase is present). For a typical FP, (26) shows a version of the analyses proposed by Folli and Harley (2007) and Guasti (1996): the complement of *faire* is not a *v*P but rather a nominalized VP.[11] The *par*-phrase is a PP adjunct similar to the *by*-phrase in English derived nominals (Folli and Harley 2007) and does not count as an intervener in either movement or case valuation. Accordingly, neither smuggling nor the Modified Freezing Principle plays any role in FP/FN causatives. Similarly, when the causativized verb

[10] Folli and Harley (2007) and Guasti (1996) base their argument on passivization facts in Italian. There the *faire-infinitif* does not passivize, and passivized causatives containing verbs with alternating unergative/unaccusative interpretations are only acceptable on an unaccusative reading. For a discussion of the corresponding passivization facts in French, see Belletti (2016a, pp. 15–18).

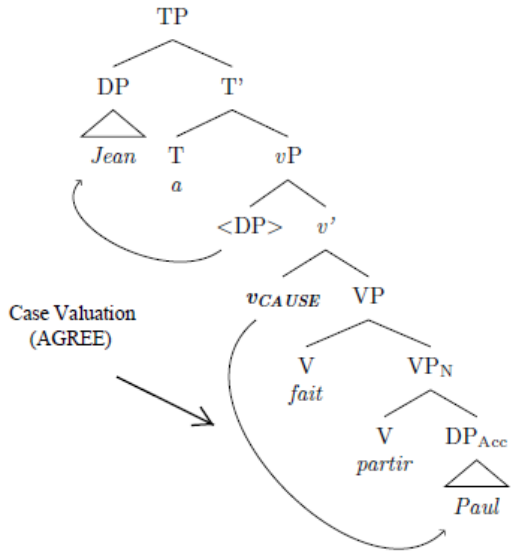
[11] A reviewer points out issues with Folli and Harley's (2007) analysis of the embedded VP in FPs as nominalized; in particular, the allegedly nominalized constituent does not behave like a DP as regards restrictions on extraction. Labelle (2013) presents an analysis based on Folli and Harley (2007) in which *faire* simply embeds a non-external argument assigning *v*P, with an agent optionally assigned in a PP. As regards the acquisitional prediction, both analyses critically posit the absence of an intervening external argument in the FP.

is an unaccusative, as in (27), there is nothing to intervene between v_{cause} and the accusative-marked causee, *Paul*. No smuggling occurs, and the Modified Freezing Principle is respected. Hence, the prediction of the UFH is that children under four will succeed.

26. *Faire-par (FP):* Cf. Guasti 1996; Folli and Harley 2007



27. *Faire* + unaccusative = '*faire-par*'



4.2. Testing the Predictions

Note that Guasti's (1996) acquisitional conjecture, which she expressed in terms of Kayne's (1975) *faire-infinitif* / *faire-par* distinction, linked causatives of the *faire-infinitif* type with English *be*-passives. In Section 2, when we tested her conjecture, we focused on FDs, which are *faire-infnitif* causatives in which the causativized verb is transitive. Yet, as discussed in the preceding section, FuEs also seem to belong to the *faire-infinitif* type of causative. If so, one might expect them to show the same acquisitional delay seen for FDs in Section 2.3. Yet, as indicated above, the UFH actually leads to the opposite prediction: FuEs (as well as FuAs) are expected not to show any particular delay.

For the corpus results reported in Section 2.3, we simply distinguished between FDs and other causatives, but to evaluate the new predictions for causativized intransitives we will now examine utterances in the "Other" category more closely. In brief, both types of causative are attested at early ages, although at all ages FuAs are produced much more frequently than FuEs. Ten of the 11 children produced at least one example of an FuA in their recorded speech, and in all ten cases the first use was prior to the age of three. For FuEs, 6 of the 11 children produced at least one example, and in four cases the first use was prior to the age of three. Some representative examples of FuEs are provided in (28), and details for each child are provided in Table 3.

28. a. Il me fait pleurer
he me makes cry
'He is making me cry.' (Marilyn 2;11,14)

b. Ça me fait rire
that me makes laugh
'That makes me laugh.' (Théotime 2;7,28)

Thus, our evidence from the acquisition of French *faire* causatives with intransitive verbs, both FuAs and FuEs, is fully in line with the predictions of the UFH.

Table 3. Causativized intransitives, by Child.

	First Use of FuA (Age)	#FuA (Total)	First Use of FuE (Age)	#FuE (Total)
Anaé	2;00	7	NA	0
Anaïs	2;08	3	NA	0
Anne	2;04	10	2;05	2
Antoine	2;06	11	4;02	1
Léonard	2;01	2	NA	0
Madeleine	2;01	10	NA	0
Marie	2;06	12	2;05	1
Marilyn	2;09	3	2;11	1
Nathan	NA	0	NA	0
Théophile	2;10	8	3;07	3
Théotime	2;03	9	2;05	4

5. *Get*-Passives and Reflexive Causative Passives

In this section we return to two structures that are acquired substantially earlier than the age of four, *get*-passives and reflexive causative passives, to clarify how they differ from *be*-passives and FDs, respectively.

5.1. Get-Passives

In contrast to *be*-passives, English *get*-passives are produced and understood by the age of two or three years (Turner and Rommetveit 1967a, 1967b; Harris and Flora 1982; Crain and Fodor 1993; Slobin 1994). While the underlying structure of *get*-passives is controversial, one influential view is that there exist multiple types of *get*-passives, each with its own structure (Reed 2011, p. 42). Snyder and Hyams (2015) contend that the types of *get*-passive that are produced and understood by children younger than four do not contain an external argument.

Tests for the syntactic presence of a phonetically null, external agent argument include compatibility with a purpose clause, as in (29), and compatibility with the adverbial modifier *on purpose*, as in (30) (these examples are based on Fox and Grodzinsky 1998, p. 327).

29. a. The ship was sunk [PRO to collect the insurance money].

    b. *The ship sank [PRO to collect the insurance money].

    c. *??The ship got sunk [PRO to collect the insurance money].

30. a. The book was torn on purpose.

    b. *The book tore on purpose.

    c. *??The book got torn on purpose.

The *be*-passives in (29a) and (30a) clearly have an understood agent in their syntactic structures, because the agent is available to control the PRO subject in a purpose clause (29a), or to license the use of the agent-oriented adverbial *on purpose* in (30a). The anticausative counterparts in (29b) and (30b), which uncontroversially lack any external argument, are sharply ungrammatical (at least on the interpretations available in 29a and 30a). Crucially, the *get*-passives in (29c) and (30c) appear to pattern with the anticausatives, which indicates that these particular examples of the *get*-passive are actually incompatible with the syntactic presence of an understood Agent.

Yet, while the native-speaker judgments indicated in (29) and (30) are the ones that are usually reported, there does exist some inter-speaker variation for the (c) examples.[12] Indeed, Reed (2011) and Alexiadou (2012) report that *get*-passives are sometimes judged to be fully compatible with a *by*-phrase, or with other diagnostics for the syntactic presence of an Agent, especially if contextual support is provided. According to Alexiadou's (2012) account, English actually has two different non-actional voice heads, corresponding to passive voice and middle voice, that are possible in a *get*-passive (in contrast, the *be*-passive always contains the 'passive' voice head). Crucially, passive voice assigns an external argument, but middle voice does not.

For Alexiadou (2012), two factors contribute to the structural ambiguity of *get*-passives. First, English middle voice is not associated with any overt morphology.[13] Second, *get* is a semantically bleached aspectual verb introducing a change-of-state event, but lacking its own argument structure. Hence, the verbs that typically form *get*-passives are semantically unspecified for a distinction between external and internal causation. In contexts that strongly favor an agentive interpretation of the event associated with the participle, an underlying passive-voice construal of the *get*-passive is, at least, marginally possible for some speakers, and this yields the less-than-straightforward judgments for examples like (29c) and (30c).

For present purposes, the essential part of Alexiadou's (2012) account is simply that the "passivized" verb in an English *get*-passive routinely lacks the external argument that would be present in a *be*-passive. If there is no external argument to intervene, the promotion of the logical object into subject position will not require smuggling. Hence the Freezing Principle will not be violated, and the UFH is fully compatible with children's early success on *get*-passives.

5.2. Reflexive Causative Passives

As noted above, the earliest examples of FPs that we found were actually reflexive causatives; these were already present in the speech of two-year-olds. As discussed by Belletti (2016a) for Italian, these "reflexive causative passives" are, intuitively, an extremely complex grammatical construction, yet

---

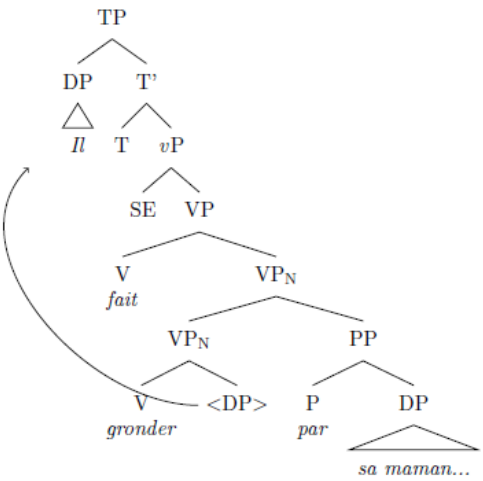
[12] We thank an anonymous reviewer for bringing this to our attention.

[13] See Alexiadou (2012) and Alexiadou and Doron (2012) for discussion of cross-linguistic variation in the morphological realization of the functional head for middle voice. There the *get*-passive is specifically compared to the medio-passive forms in Greek and Hebrew, which do not require the presence of an external argument and share morphology with the dispositional middle.

people's intuitive ideas about grammatical complexity are a poor predictor of acquisitional timing, and in Italian (much as in French), children produce Reflexive Causative Passives (RCPs) remarkably early.

For RCPs (see 31), we again follow Guasti (1996) and Folli and Harley (2007) in their analyses of FPs. For the object-to-subject analysis of "formally reflexive" clitic constructions in French, we follow Sportiche (2010). Snyder and Hyams (2015) have provided extensive evidence for early mastery of reflexive and non-reflexive clitics in children acquiring French and Italian. Moreover, despite the passive-like interpretation, there is no passive morpheme in reflexive FPs, and thus—once again—no need for smuggling. Hence, children's early success on RCPs is entirely as expected.

31. Reflexive Causative Passive (RCP):



Belletti (2016a, 2016b), however, offers an alternative analysis of RCPs, in which the reflexive morpheme represents the external argument of the causative verb, and the lower argument does have to be smuggled past it into surface subject position. Belletti (2016b) further notes that the *da*-phrase in the Italian reflexive *si-fare* is compatible with non-agent arguments (32a), which are illicit in the regular, non-reflexive *fare-da* (32b).

32. a. Maria si fa capire da tutti
Maria SI makes understand by everybody
'Maria makes herself understood by everybody.' (Belletti 2016b, p. 5)

b. *Questa spiegazione ha fatto capire il problema da tutti
this explanation has made understand the problem by everybody
'This explanation has made the problem understood by everybody.'

Belletti (2016b) takes this latter observation to indicate that the reflexive FP causative, at least in Italian, is not always as closely related to the non-reflexive FP as the presence of the *da*-phrase would suggest, and indeed she leaves open the possibility that the construction is potentially ambiguous between a more complex, semantically reflexive construction requiring smuggling, and a simpler derivation along the lines of (31). Hence, one possibility is that Belletti's (2016a) analysis (with smuggling) is the correct one when a true, semantic reflexive is present (as it often is, in adult speech), while the structure is simpler (and smuggling is not involved) when the construction is only "formally" (not semantically) reflexive.

There are several reasons to assume that a simple structure of this kind is indeed available to children who have not yet mastered smuggling. As regards the reflexive morpheme, Snyder and Hyams (2015) provide detailed evidence for two-year-old French and Italian children's mastery of formally (but not semantically) reflexive constructions with *se*. In addition, Labelle (2013) observes that the reflexive morpheme in the RCP (*se-faire*) construction in some cases must not be associated

with the external argument of the causative verb, as the surface subject cannot be construed as playing a role in bringing about the caused event (33).

33.	Les habitants	se sont	fait	sorprendre	pendant	leur sommeil	par
	the inhabitants	SE AUX	made	surprise	during	their sleep	by
	l'eruption	du volcan					
	the eruption	of-the volcano					
	'The inhabitants were taken by surprise during their sleep by the eruption of the volcano.'						(Labelle 2013, p. 238)

Hence, at least in some cases, *se* clearly is not the external argument of the causative verb. In our view it is quite reasonable to assume that the RCPs in early child speech are not semantically reflexive, that *se* is not an intervener for purposes of RM, and that no smuggling is required.

Finally, we would like to note that Labelle (2013) posits that *se* in RCPs is actually the overt realization of a non-active voice head which prevents the merge of an external argument, resulting in the promotion of the internal argument from the embedded predicate and the absence of a semantically reflexive interpretation. The causative verb itself is described as a semantically underspecified lexical realization of a verb which brings about an event. In light of the acquisitional evidence presented in this paper, we believe the parallels between these ideas and the structure Alexiadou (2012) formulates for the *get*-passive in English definitely warrant further exploration.

6. Conclusions

In this acquisitional study of French causatives, we first tested and found support for Guasti's (2016) conjecture regarding the acquisition of *faire-par* versus *faire-infinitif* causatives, at least as it applies to causativized transitives: there was strong evidence for a delay in FDs. Next, drawing on the syntax literature for Italian, and updating the Freezing Principle to take AGREE relations into account, we showed that the UFH can explain the observed delay for FDs. We then derived a novel prediction: even though *faire*-causatives with unergative verbs require an analysis as *faire-infinitif* causatives, they will not be delayed like the FDs, since the unergatives have no internal argument requiring case valuation. This prediction was tested and supported. Finally, we examined certain parallels between RCPs and *get*-passives with respect to non-assignment of an external theta-role, which help to explain why these structures are not affected by the UFH.

The above findings raise several questions for future research. As described by Hirsch and Wexler (2006), children at least as old as age 5 struggle with accurate comprehension of the *by*-phrase in English nominals. One potential take-away from this observation is that children have difficulties acquiring the default Agent interpretation conveyed by the by-phrase, outside of those actional passives where young children do succeed, given adequate context or a structural prime. The *par*-phrase in both the FP and RCP constructions has been analyzed as similarly conveying a default Agent role (Guasti 1996; Folli and Harley 2007; Labelle 2013). As Guasti (2016) notes, this would seem to predict that children acquiring these constructions might experience difficulties in the accurate comprehension of the *par*-phrase. As several instances of *par*-phrases did appear in the spontaneous speech data prior to age 3, this prediction deserves further review. In light of Belletti's (2016b) observation regarding the differences between *da*-phrases in the Italian FP and RCP constructions (the former involving a default Agent role assignment, the latter transmission of a thematic role from the verb), future research might test whether children acquiring French display asymmetries in the comprehension/production of these two constructions.

Another area in which the UFH makes predictions concerns the inverse/specificational copular construction in English. Specificational copular sentences (34a) appear to invert the typical order of arguments in a predicational copular construction (34b).

34. a. The doctor is Maria

    b. Maria is the doctor

Under several analyses of specificational copular sentences including that developed in Mikkelsen (2005), the predicational argument (*the doctor* in (34)) is Merged in a structurally lower position than the referential argument (*Maria)* prior to being promoted to subject position, resulting in what appears to be a Minimality violation. If this violation must be circumvented via a more complex smuggling derivation, then a delay in the acquisition of specificational copular sentences is predicted under the UFH.

In conclusion, evidence from the acquisition of Romance causatives provides further evidence for the UFH, and contributes to a developing understanding of the role of freezing effects in acquisition. The similarities between *get*-passives and RCPs merits further acquisitional and theoretical research, particularly given the very early appearance of the latter in spontaneous speech data.

Author Contributions: J.B. and W.S. shared responsibility for the following aspects of this project: Conceptualization, Methodology, Resources, Funding Acquisition, Investigation, Supervision, Project Administration, Data Curation, Software, Formal Analysis, Validation, Visualization, Writing-Original Draft Preparation, and Writing-Review & Editing.

Funding: This research was funded by NSF grant IGERT DGE-1144399 to the University of Connecticut, and by a Seed Grant to W.S. from the Connecticut Institute for the Brain and Cognitive Sciences.

Acknowledgments: The authors are grateful to the Editors for their work on this Special Issue, and to the reviewers for their many helpful suggestions. The authors would also like to thank Adriana Belletti, Giulia Bellucci, Kamil Deen, Stéphanie Durrleman, Nina Hyams, Luigi Rizzi, and Virginia Valian for helpful discussion.

Conflicts of Interest: The authors declare no conflicts of interest.

References

Alexiadou, Artemis. 2012. Noncanonical passives revisited: Parameters of nonactive Voice. *Linguistics* 50: 1079–110. [CrossRef]

Alexiadou, Artemis, and Edit Doron. 2012. The syntactic construction of two non-active Voices: passive and middle. *Journal of Linguistics* 48: 1–34. [CrossRef]

Baker, Mark. 1988. *Incorporation*. Chicago: Chicago University Press.

Baker, Mark. 1997. Thematic roles and syntactic structure. In *Elements of Grammar*. Edited by Liliane Haegeman. Dordrecht: Kluwer, pp. 73–137.

Belletti, Adriana. 2016a. On the Acquisition of Complex Derivations, with Related Considerations on Poverty of the Stimulus and Frequency. Unpublished manuscript, Université de Genève, Geneva.

Belletti, Adriana. 2016b. Reflexive Si as a Route to Passive in Italian. Unpublished manuscript, Université de Genève, Geneva.

Belletti, Adriana, and Luigi Rizzi. 2012. Moving verbal chunks in the low functional field. In *Cartography of Syntactic Structures*. Edited by Laura Brugè, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro and Cecilia Poletto. New York: OUP, vol. 7, pp. 129–37.

Bellucci, Giulia. 2015. Acquiring Romance causatives: Evidence from child Italian. *Quaderni di Linguistica e Studi Orientali* 1: 89–116.

Borga, Jason, and William Snyder. Forthcoming. *French Experiencer Verbs and the Universal Freezing Hypothesis*. Edited by Anna Gavarró. Selected Proceedings of The Romance Turn 8. Amsterdam: John Benjamins.

Burzio, Luigi. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.

Collins, Chris. 2005a. A smuggling approach to the passive in English. *Syntax* 8: 81–120. [CrossRef]

Collins, Chris. 2005b. A smuggling approach to raising in English. *Linguistic Inquiry* 36: 289–98. [CrossRef]

Crain, Stephen, and Janet D. Fodor. 1993. Competence and performance in child language. In *Language and Cognition: A Developmental Perspective*. Edited by Esther Dromi. Norwood: Ablex, pp. 141–71.

Demuth, Katherine, and Annie Tremblay. 2008. Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language* 35: 99–127. [CrossRef] [PubMed]

Folli, Raffaella, and Heidi Harley. 2007. Causation, obligation, and argument structure: On the nature of little *v*. *Linguistic Inquiry* 38: 197–238. [CrossRef]

Fox, Danny, and Yosef Grodzinsky. 1998. Children's Passive: A view from the by-phrase. *Linguistic Inquiry* 29: 311–332. [CrossRef]

Gehrke, Berit, and Nino Grillo. 2009. How to become passive. In *Explorations of Phase Theory: Features, Arguments, and Interpretation at the Interfaces*. Edited by Kleanthes K. Grohmann. Berlin: De Gruyter Mouton, pp. 231–68.

Guasti, Maria Teresas. 1996. Semantic restrictions in Romance causatives and the incorporation approach. *Linguistic Inquiry* 27: 294–313.

Guasti, Maria Teresa. 2016. Voice alternations (active, passive, middle). In *Oxford Handbook of Developmental Linguistics*. Edited by Jeffrey Lidz, William Snyder and Joe Pater. Oxford: Oxford University Press, pp. 179–205.

Harris, Frances N., and June A. Flora. 1982. Children's use of get passives. *Journal of Psycholinguistic Research* 11: 297–311. [CrossRef]

Hirsch, Christopher, and Kenneth Wexler. 2004. Who seems to rescue raising. Paper presented at Generative Approaches to Language Acquisition North America (GALANA), University of Hawai'i at Manoa, Honolulu, HI, USA, December 17.

Hirsch, Christopher, and Kenneth Wexler. 2006. By the way, children don't know by. In *BUCLD 30: Proceedings of the 30th Annual Boston University Conference on Language Development*. Edited by David Bamman, Tatiana Magnitskaia and Colleen Zaller. Somerville: Cascadilla, pp. 249–61.

Kayne, Richard S. 1975. *French Syntax: The Transformational Cycle*. Cambridge: The MIT Press.

Kayne, Richard S. 2004. Preposition as probes. In *Structures and Beyond. The Cartography of Syntactic Structure*. Edited by Adriana Belletti. Oxford: Oxford University Press, vol. 3, pp. 192–97.

Keine, Stefan. 2016. Probes and their Horizons. Ph.D. dissertation, University of Massachusetts Amherst, Amherst, MA, USA.

Keine, Stefan. 2018. The Ups and Downs of Agreement. Paper presented at Colloquium Speech, McGill University, Montreal, QC, Canada, January 31.

Labelle, Marie. 2013. The passive *se faire* construction in French: The anticausative of a causative. In *Non Canonical Passives*. Edited by Artemis Alexiadou and Florian Schäfer. Amsterdam: John Benjamins, pp. 235–60.

MacWhinney, Brian. 2000. *The CHILDES project: Tools for Analyzing Talk*, 3rd ed. Mahwah: LEA.

Manzini, Maria Rita, and Leonardo M. Savoia. 2011. *Grammatical Categories. Variation in the Romance Languages*. Cambridge Studies in Linguistics 128. Cambridge: Cambridge University Press.

Messenger, Katherine. 2010. Syntactic priming and children's production and representation of the passive. *Language Acquisition* 17: 121–23. [CrossRef]

Mikkelsen, Line. 2005. *Copular Clauses. Specification, Predication, and Equation*. Amsterdam: John Benjamins.

Morgenstern, Aliyah, and Christophe Parisse. 2007. Codage et interprétation du langage spontané d'enfants de 1 à 3 ans. *Corpus* 6: 55–78.

Müller, G. 1998. *Incomplete Category Fronting*. Dordrecht: Kluwer.

Plunkett, Bernadette. 2002. Null Subjects in child French interrogatives: A view from the York Corpus. In *Romanistische Korpuslinguistik: Korpora und Gesprochene Sprache (=ScriptOralia; 126)*. Edited by Claus D. Pusch and Wolfgang Raible. Tübingen: Narr, pp. 441–52.

Pylkkänen, Liina. 2008. *Introducing Arguments*. Cambridge: MIT Press.

Reed, Lisa A. 2011. *Get*-passives. *The Linguistic Review* 28: 41–78. [CrossRef]

Rizzi, Luigi. 2001. Relativized Minimality effects. In *The Handbook of Contemporary Syntactic Theory*. Edited by Mark Baltin and Chris Collins. Oxford: Blackwell, pp. 89–110.

Rizzi, Luigi. 2004. Locality and Left Periphery. In *Structures and Beyond. The Cartography of Syntactic Structure*. Edited by Adriana Belletti. Oxford: Oxford University Press, vol. 3, pp. 223–51.

Ross, John R. 1967. Constraints on Variables in Syntax. Ph.D. dissertation, MIT, Cambridge, MA, USA.

Ross, John R. 1986. *Infinite Syntax!* Norwood: Ablex.

Rouveret, Alain, and Jean-Roger Vergnaud. 1980. Specifying reference to the subject: French causatives and conditions on representations. *Linguistic Inquiry* 11: 97–202.

Slobin, Dan I. 1994. Passives and alternatives in children's narratives in English, Spanish, German, and Turkish. In *Voice: Form and Function*. Edited by Barbara Fox and Paul J. Hopper. Amsterdam and Philadelphia: John Benjamins, pp. 341–64.

Snyder, William. 2007. *Child Language: The Parametric Approach*. Oxford: Oxford University Press.

Snyder, William, and Nina Hyams. 2015. Minimality effects in children's passives. In *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*. Edited by Elisa Di Domenico, Cornelia Hamann and Simona Matteini. Amsterdam: John Benjamins, pp. 343–68.

Sportiche, D. 2010. French Reflexive *se*: Binding and Merge Locality. Available online: http://ling.auf.net/lingBuzz/001273 (accessed on 13 August 2015).

Turner, Elizabeth A., and Ragnar Rommetveit. 1967a. The acquisition of sentence voice and reversibility. *Child Development* 38: 649–60. [CrossRef]

Turner, Elizabeth A., and Ragnar Rommetveit. 1967b. Experimental manipulation of the production of active and passive voice in children. *Language and Speech* 10: 169–80. [CrossRef] [PubMed]

Uriagereka, Juan. 1999. Multiple spell-out. In *Working Minimalism*. Edited by Samuel David Epstein and Norbert Hornstein. Cambridge: MIT Press, pp. 251–82.

Wexler, Kenneth, and Peter Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge: MIT Press.

languages

MDPI

Article

Null Subject Occurrence in Monolingual Spanish SLI: A Discriminant Function Analysis

John Grinstead [1,*], Paij Lintz [1], Amy Pratt [2], Mariana Vega-Mendoza [3], Juliana De la Mora [4], Myriam Cantú-Sánchez [5] and Blanca Flores-Avalos [6]

1 Department of Spanish and Portuguese, The Ohio State University, 298 Hagerty Hall, 1775 College Road, Columbus, OH 43210, USA; Paijlintz3@gmail.com
2 Department of Speech and Hearing Sciences, The Ohio State University, 1070 Carmack Road, 110 Pressey Hall, Columbus, OH 43210, USA; Pratt.110@osu.edu
3 School of Philosophy, Psychology and Language Sciences, The University of Edinburgh, Edinburgh EH8 9YL, UK; m.vega-mendoza@ed.ac.uk
4 Facultad de Lenguas y Letras, Universidad Autónoma de Querétaro, Cerro de las Campanas, Centro Universitario, 76010 Santiago de Querétaro, Mexico; juliana.delamora@gmail.com
5 Departament de Filologia Catalana, Universitat Autonoma de Barcelona, Campus UAB, Plaça Cívica, Bellaterra, 08193 Barcelona, Spain; myriam.cantu.s@gmail.com
6 Instituto Nacional de Rehabilitación, Calz México-Xochimilco 289, Arenal Tepepan, 14389 Ciudad de México, Mexico; bflores@inr.gob.mx
* Correspondence: grinstead.11@osu.edu

Received: 14 November 2017; Accepted: 17 May 2018; Published: 23 May 2018

Abstract: Background: Child Spanish-speakers appear to use more null subjects than do adults. Null subject use, like the use of tense marking, is sensitive to discourse-pragmatics. Because tense marking has been used to identify child Spanish-speakers with specific language impairment (SLI) with near good sensitivity and specificity (89%), null subject use may as well, following the predictions of the Interface Deficit Hypothesis. We investigate the possibility that null subject occurrence may form part of a useful discriminant function for the identification of monolingual child Spanish-speakers diagnosed with specific language impairment. **Methods**: We evaluate the rate of null subject expression from spontaneous production data, together with results from independent measures of another discourse-sensitive construction, verb finiteness, in child Spanish. We perform a discriminant function analysis, using null subject expression as a target variable, among others, to classify monolingual child Spanish-speakers (N = 40) as SLI or as typically-developing (TD). **Results**: The SLI group is shown to have significantly higher scores than the TD group on null subject expression. Multiple discriminant functions, including the null subject variable with tense measures, and in combination with mean length of utterance in words (MLUw), are shown to provide good sensitivity and specificity (<90%) in the classification of children as SLI vs. TD. **Conclusion**: Our findings support the contention that null subject occurrence is a plausible reflection of the Interface Deficit of SLI for Spanish-speaking children.

Keywords: Spanish; null subjects; SLI; Interface Deficit; Interface Delay; discriminant function analysis

1. Introduction

Children acquiring overt subject languages such as English produce fewer overt subjects than adults do, with rates of overt subject use that vary from 30–100%, depending on age and mean length of utterance (MLU, see Hyams 2011 for review). Developmental linguistic studies of this phenomenon in overt subject languages have explored syntactic (e.g., Kramer 1993; Rizzi 2005), phonological (e.g.,

Gerken 1991, 1994) and language processing (Bloom 1991, 1993; Valian 1990) accounts of null subject utterances in these languages, including child English, French, Dutch and German (Hamann 1996; Hamann and Plunkett 1998). Researchers concerned with specific language impairment (SLI) have explored the possibility that this non-adult-like predisposition of English-speaking children to use overt subjects less than adults might distinguish children with SLI from those without it (Grela 2003a, 2003b; Grela and Leonard 1997).

In null subject languages, such as Spanish, Catalan and Italian, there is also evidence (though a debate persists) that typically-developing children use more null subjects than adults. Those arguing that null subject use in these languages is non-adult-like (Grinstead 1998, 2000, 2004; Grinstead and Spinner 2009) maintain that neither language processing nor developing syntax explain non-adult-like subject use in Spanish, but rather that the cross-linguistic predisposition of children to overuse definites, including phonetically null subject pronouns, is at work. For example, child English speakers say things like *Have you seen her, Mommy?* when Mommy does not know who *her* is, or *Have you seen the dinosaur, Daddy?* when the child is surrounded by 75 plastic dinosaurs and Daddy cannot possibly know which dinosaur is prominent in the child's discourse representation, and should thus be referred to with definite *the*. Piaget (1929) famously referred to this phenomenon as characteristic of the "Egocentric" developmental stage. Grinstead (1998, 2004) and Grinstead et al. (2013, 2014) contend that this overuse of definites manifests itself in Spanish and Catalan as the overuse of null subjects, among other things. The idea is that null subjects are used in a language such as Spanish when the speaker assumes that the listener requires no clarification as to what antecedent to this pronoun the speaker believes to be prominent in discourse. Overt subjects may be used in adult Spanish when such clarification is required (e.g., topicalization—foregrounding of old information; focus—making prominent new information, etc.) or for other discourse-pragmatic reasons having nothing to do with grammar, as in variationist studies (see Otheguy and Zentella 2012 for review). From this perspective, children's use of fewer overt subjects than adults in English would seem to be a different kind of phenomenon than in Spanish. Child English-speakers may lack the adult-like syntactic, phonological or language processing resources required to produce overt subjects. In contrast, child Spanish-speakers would appear to possess these resources, but fail to deploy them as a result of the general, cross-linguistic developmental tendency of children to assume that their interlocutors share their presuppositions about the salience of referents in the Conversational Common Ground, to use Stalnaker (1974) term. In the same way that this predisposition plays out as, for example, overuse of definite articles in English (see e.g., Maratsos 1974, 1976), it plays out as the overuse of null subjects in Spanish.

To date, no one has investigated the possibility that child Spanish-speakers with SLI use more null subjects than do typically-developing age-matched children. In what follows, we investigate this question, with a comparative eye towards overt subject child languages and with the overall objective of determining whether null subject occurrence, by itself, or in combination with other variables, can help us identify child Spanish-speakers with SLI, using a discriminant function analysis. Though the very interesting question of which discourse-pragmatics variables (e.g., joint attention, physical presence of antecedents and others) determine the use of overt subjects should also be investigated, we will limit the scope of this study to measuring proportions of use of null subjects, which is the elsewhere condition in most varieties of non-Caribbean Spanish. Doing this narrows our focus to grammatical contexts in which children allow the Spanish null pronominal subject to be used. This coheres with our previous work in that it is fundamentally about anaphora, which typically-developing children across languages struggle with (Interface Delay), and which children with SLI have severe problems with (Interface Deficit). In a more practical vein, it is also true that such easy-to-calculate rates of null subjects could be useful to speech-language pathologists in identifying children with SLI in Spanish. First, we turn to children's subject use in an overt subject language, English, and to studies of subject use by child English speakers with SLI.

2. Subject Occurrence in an Overt Subject Language

Overt subject use in child English has been extensively studied since the 1960s (e.g., Antinucci and Parisi 1973; Bates 1976; Bloom et al. 1975a, 1975b; Brown and Fraser 1964; Guilfoyle 1984; Hyams 1986; Orfitelli and Hyams 2008; Valian 1991). As an illustration of the phenomenon, Valian (1991) shows that in a cross-sectional sample of US English-speaking children, the least grammatically developed, between MLU 1.5 and 2.0 (n = 6, age range 1;10–2;2), used overt subjects in roughly 70% of non-imperative sentences, while a sample of Italian-speaking children used overt subjects in roughly 30% of such utterances (n = 5, age range 2;0–2;5). Examples of such null subject sentences in child English include the following from Bloom et al. (1975a, 1975b) and Brown (1973).

1. Want more apple.
2. Tickles me.
3. No play matches.
4. Show mommy that.

Among accounts of the null subject phenomenon for child English that have been proposed is the prosodic account of Gerken (1991, 1994), which claims that sentences consisting of iambic metrical structures (two syllables, the first of which is weak), as in [she KISSED] + [the DOG], are most likely to have the weak syllables (critically, the *she* subject) deleted. While this account seems promising for child English, it seems less so for child Spanish and Catalan, in which subject pronouns are stressed (e.g., *Ella besó al perro* 'She kissed the dog') and would not be predicted to delete for prosodic reasons.

Language processing accounts of child English null subjects claim that overt subjects are dropped because language processing resources are less available at the beginning of sentences (Bloom 1991) and that the phonetically heavier the subject, the more likely it is to be deleted. Problems for this account of child English include the fact that children seem much better able to produce full lexical noun phrase (NP) subjects than they are pronominal subjects in elicited imitations studies (Gerken 1991; Valian et al. 1996). To our knowledge, such an explanation has not been pursued for child null subject languages.

A wide range of developmental syntactic accounts of this non-adult-like behavior have been proposed, many of which have in common an assumption of underdeveloped syntactic structure housing subjects in child grammars (Guilfoyle 1984; Guilfoyle and Noonan 1992; Hyams 1986; O'Grady et al. 1989; Radford 1990; Rizzi 1993). Radford (1990), for example, proposes that functional structure associated with inflection, assumed to house subjects on influential accounts (e.g., Baker 1985, 1996; Pollock 1989), is initially missing, accounting for documented contingencies between verbal nonfiniteness and null subjects (see Hyams and Wexler 1993; Sano and Hyams 1994). On this account, the null subject used in child English is PRO (Chomsky 1981), the null subject used in the subject position of nonfinite verbs, as in (5) and (6).

5. [PRO to leave now] would be ill-advised.
6. Juan quiere [PRO salir del auto ya].

Similar proposals have been made for the gramatical nature of the null subjects used in other child overt subject languages, including Dutch (Kramer 1993); as well as for child Spanish and Catalan (Grinstead 1998), when they occur with nonfinite verbs, which are a fixture in child language cross-linguistically. On such proposals, nonfinite PRO is discourse-identified by the same principles that govern the discourse identification of the adult-like *pro* of null subject languages. Thus, whether they occur with nonfinite verbs and are PRO, or with finite verbs and are *pro*, null subjects in child Spanish would seem to require that interlocutors be able to recover the identity of the antecedent, which children appear to generally overassume. While there is a diverse array of child null subject accounts, to date, no comprehensive study has been done to determine empirically the degree to which each variable contributes to an explanatory account of the phenomenon. In what follows, we will

assume that child Spanish-speakers are using some combination of *pro* and PRO null subjects, as a function of the finiteness of their verbs.

3. Subject Occurrence in English-Speaking Children with Specific Language Impairment

While there does not appear to be a consensus account of this phenomenon in child overt subject languages, it seems to be uncontroversial that child English speakers are doing something with subjects that is not adult-like. In the domain of language disorders, Loeb and Leonard (1988) observe that child English speakers with SLI appear to use more null subjects than do typically-developing children. Exploring this observation further, Grela and Leonard (1997) report that in spontaneous production, child English speakers with SLI use more null subjects than do MLU controls. In a sentence completion task, Grela (2003a) shows that a sample of SLI children produces more null subjects than do age controls, but not more than do MLU controls. We see then in child English that null subject production may serve to differentiate children with SLI from typically developing age and MLU controls, depending on methodology. If, as some authors have argued, child Spanish speakers produce more null subjects than do adults, could we find that null subject production is a similarly useful grammatical characteristic of children with SLI, such that it could distinguish them from typically developing children? Before answering this question, we turn to the question of subject use in typically-developing child Spanish.

4. Subject Occurrence in Child Spanish and Catalan

There is evidence that typically-developing child Spanish-speakers also use more null subjects than do adults, though many generative child language researchers (e.g., Hyams 2011; Camacho 2013) appear committed to the position that null subject child languages are adult-like, in spite of this evidence. Grinstead (2000) shows that four longitudinally-studied child speakers of Catalan from the Serra and Solé (1986) Corpus, at the beginning of two-word speech, pass through an early period during which they use no overt subjects at all with verbs in non-imperative sentences. Grinstead (2004) shows that the same phenomenon holds of three longitudinally-studied Spanish-speaking children.

Attempts to dismiss the empirical claim that children learning null subject languages pass through a No Overt Subject Stage have presented putative counterevidence in the form of rates of overt subject occurrence averaged over developmentally large spans of time or have depended on children who have higher MLUs and higher levels of grammatical sophistication even in their first recording sessions than the children reported in Grinstead and Spinner (2009). For example, Bel (2003) presents data from six children aggregated over 8–12 months and also gives month-by-month breakdowns for two of them (Júlia and María). In the aggregate reports (Table 3, p. 7) she shows, for example, that a Catalan-speaking child, Pep, between 1;06 and 2;06 (12 months of language development) uses overt subjects 32% of the time; arguably adult-like proportions. Presenting data in this way, however, disguises the fact that before Pep begins using overt subjects at adult-like rates, or in fact any overt subjects at all, he produces 48 non-imperative verbal utterances without overt subjects (Grinstead 2000, Table 1, p. 125). The other six children reported in Grinstead (2000, 2004) have similar early No Overt Subject Stages.

The month-by-month data presented by Bel (2003), and Aguado-Orea and Pine (2002) to argue against the no-overt subject claim critically depends on the data of children such as María of the López-Ornat corpus. María's MLU in words (MLUw) in her earliest file is already 1.93, substantially greater than the MLUw of the Catalan-speaking children (mean MLUw = 1.55) and the MLUw of the Spanish-speaking children (mean MLUw = 1.55) in Grinstead and Spinner (2009). More importantly, as pointed out in Grinstead and Spinner (2009), María appears to be using more grammatically sophisticated constructions than the children studied in Grinstead (2004) in that she uses verbs with past imperfect verb tense (e.g., *Tenía pupa* 'I had an owie', María, 1;7), wh- questions (e.g., *¿Dónde está el miau?* 'Where's the kitty?', María 1;7) and fronted objects (e.g., *Este a pi* 'This, let's paint', María 1;7),

none of which are present in the spontaneous speech of the seven children studied by Grinstead and colleagues.[1]

Other studies regularly cited as counterevidence to the No Overt Subject Stage claim, for example in Hyams (2011) and Camacho (2013), suffer from these same two shortcomings. Obviously, if one's goal is to determine whether or not there is a no-overt subject stage in a child's longitudinal data, one has to begin looking at data collected at the end of the 1-word stage. Serratrice (2005), for example, cited in both Hyams (2011) and Camacho (2013), is a study of the discourse-pragmatic influences on overt subject expression. Its child data begins when children are already at MLUw 1.5—the exact point at which the children studied in Grinstead (2004) begin to use overt subjects—making it irrelevant to the no overt subject claim, though it is a very interesting study of developmental discourse-pragmatics. Similarly uninformative for the question of whether there is a No Overt Subject Stage, but interesting in their own right, are the data in Lorusso et al. (2004), which was about unergative vs. unaccusative intransitive subjects, and Valian (1991), which is about the plausibility of any syntactic account of null subjects in child English, both cited in Hyams (2011) work.

In sum, studies that purport to contradict the claim of a No Overt Subject Stage either present or depend upon data that has been aggregated in a way that obscures the no overt subject observation or ignores the possibility that a child's first recording session—even though the child may appear to be of a very young chronological age—may have occurred after overt subjects have begun being used. The stage is evident if researchers pay attention to aspects of children's grammatical development, reflected in MLU values, which are the relevant measures of grammatical, and not chronological, development. See Liceras et al. (2006) and Grinstead and Spinner (2009) for a fuller discussion.[2]

Besides the obvious null vs. overt subject distinction between Spanish and English, there also appears to be a difference in the syntactic nature of their overt subjects. The account of the No Overt Subject Stage presented in Grinstead (2004) and Grinstead and Spinner (2009) claims, following Alexiadou and Anagnostopoulou (1998) and Ordóñez and Treviño (1999), that overt subjects in Spanish may be housed in the part of the clause that is closely associated with verb finiteness, as in English, but that they primarily occur with other syntactically optional constituents, higher up in the structure, in the "left periphery" in Rizzi (1997) terms. This region of the clause is argued to house other discourse-sensitive constituents including wh- elements in wh- questions and fronted objects in focus and topicalization structures. To illustrate the relevant differences, in Figure 1 the English subject John occurs only in the specifier of the Inflection Phrase (IP), while the Spanish subject Juan occurs in the either the specifier of IP or in specifier of the Complementizer Phrase (CP) (e.g., Ordóñez 1997; Zubizarreta 1998). This syntactic distinction between the two language types is relevant to our claim that overt subjects in Spanish are initially missing for the same reason that children fail to use other syntactically optional, left-peripheral constituents early on.

To show that subjects, wh- questions and fronted objects[3] begin to be used at the same time, Grinstead and Spinner (2009), following Snyder and Stromswold (1997) and Snyder (2001), used the binomial test to demonstrate that there was no significant difference between the moment at which subjects began to be used and the moment when the other constructions began to be used. Table 1 reports the ages of onset of overt subjects, wh- question and fronted objects in the longitudinal study of four Catalan-speaking children and three Spanish-speaking children (compiled

---

1 Bel (2003) month-by-month presentation of the Catalan data of Júlia, in which she uses 7.2% overt subjects in her earliest recording, at MLUw 1.5, is much more consistent with the no-overt subject stage claim than with the claim that she is using overt subjects in adult-like proportions (38% according to Casanovas 1999).

2 A similar early No Overt Subject Stage is reported for Italian in Grinstead (1998) and for Romanian in Avram and Coene (2009).

3 The generic term "fronted objects" is used instead of the more specific focused objects (*ESTO quiero yo* 'This I want') or topicalized object (*Esto, lo quiero* 'This I want') because the two constructions are distinguished in adult Spanish by the presence of an accusative clitic in topicalization constructions, and children in the 2 year-old range are not consistent with their production of such clitics. In the available spontaneous production transcripts, in the absence of a clitic, the two constructions are indistinguishable.

from (Grinstead and Spinner 2009, Tables 3 and 4, pp. 66–67)). The third column gives the ages of the children in the recording session in which the first overt subject was used. The fourth and fifth columns give the children's ages when wh- questions and fronted objects were first used. Critically, the fourth and fifth columns also either gives a *p* value, if there was a significant difference between the onset of overt subjects and the onset of the non-subject construction, or it gives nothing, if there was no significant difference between the onset of overt subjects and the non-subject construction.

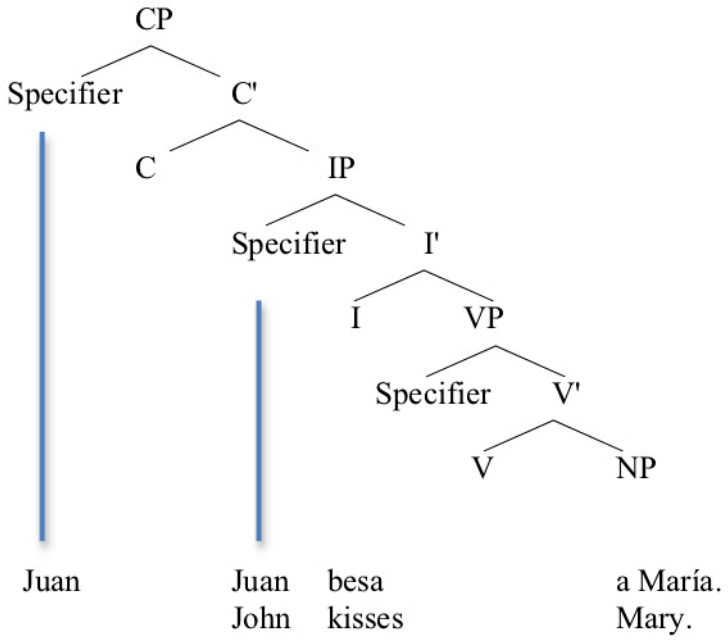


Figure 1. Left-Peripheral subjects in Spanish; IP-internal subjects in English.

Table 1. Ages of onset and binomial test result of overt subjects, wh- questions and fronted objects in 4 Catalan-speakers and 3 Spanish-speakers, compiled from Grinstead and Spinner (2009, Tables 3 and 4, pp. 66–67).

	Language	Overt Subjects	Wh- Questions	Fronted Objects
Gisela	Catalan	2;1.23	2;8.0 ($p < 0.001$)	2;1.23
Guillem	Catalan	1;11.13	2;3.28	1:11.13
Laura	Catalan	2;4.11	2;4.11	2;8;30
Pep	Catalan	1;10.06	1;11.06	1;10.06
Carlos	Spanish	1;10.13	2;9.15 ($p < 0.001$)	2;2.07
Eduardo	Spanish	2;9.10	2;9.10	2;10.28
Graciela	Spanish	1;11.29	2;0.25	2;3.25 ($p < 0.001$)

The evidence in Table 1 shows that each of the seven children had either a correlated emergence of overt subjects and fronted objects or between overt subjects and wh- questions and that 4 of the 7 children had both. Confirmation from another research group is presented in Villa-García and Snyder (2009), summarized in Table 2 (compiled from Villa-García and Snyder 2009, Tables 2 and 3, pp. 6–8). Table 2 presents evidence from four more Spanish-speaking children that there is a contingency between the onset of overt subjects and other left-peripheral constituents. In this case, the evidence is strongest for fronted objects, for which the binomial test shows no difference between their onset and the onset of overt subjects. For wh-questions, this same contingency held only for one of the four children. Summarizing over all 11 children from both research groups, six out of 11 showed a significant contingency between the onset of overt subjects and wh-questions and 10 out of 11 showed a contingency between the onset of overt subjects and the onset of fronted objects. We take this to be

strong confirmation of the hypothesis that left peripheral constructions in Spanish, including overt subjects, begin to be used at the same time. It is worth noting that, as pointed out above, the earliest transcripts of María, presented in Bel (2003) and Aguado-Orea and Pine (2002), include all three of these constructions, as Grinstead and Spinner (2009) claim would predict.

Table 2. Ages of onset and binomial test result of overt subjects, wh- questions and fronted objects in 4 Spanish-speakers, compiled From Villa-García and Snyder (2009, Tables 2 and 3, pp. 6–8).

	Language	Overt Subjects	Wh- Questions	Fronted Objects
Emilio	Spanish	1;09.19	2;03.01 ($p = 0.018$)	2;03.01
Inés	Spanish	1;06.05	1;09.03	1;09.03
Irene	Spanish	1;07.05	1;08.26 ($p = 0.001$)	1;11.13
Juan	Spanish	1;09.02	2;03.0 ($p = 0.045$)	2;03.0

In contrast to Spanish and Catalan, Grinstead and Spinner (2009) show that in child German, an overt subject language in which subject position, but not subject overtness, is discourse-sensitive, children begin using overt subjects significantly earlier, again by the binomial test, than they do wh- questions or fronted objects, as illustrated in Table 3. This comparison makes it clearer that the contingencies found between the occurrence of overt subjects and other left peripheral syntactic constituents in child Spanish and Catalan reflect stark differences in the syntactic nature of the subjects in the two language types and are not simply a function of all syntactic constructions emerging at the same time in all child languages.

Table 3. Ages of onset and binomial test result of overt Subjects, wh- questions and fronted objects in 2 German-speakers, compiled from Grinstead and Spinner (2009, Table 6, p. 73).

	Language	Overt Subjects	Wh- Questions	Fronted Objects
Caroline	German	1;10.17	2;02.10 ($p < 0.001$)	2;03.0 ($p < 0.001$)
Simone	German	1;10.20	2;10.04 ($p < 0.001$)	2;01.16 ($p < 0.001$)

In sum, child English-speakers use more null subjects than do adult English speakers and child Spanish speakers also use more null subjects than do adult Spanish-speakers. Studies that are offered as evidence that child Spanish-speakers use overt subjects in proportions similar to adult proportions either aggregate data in a way that obscures the No-Overt Subject Stage or ignore the possibility that chronologically early recording sessions may not be grammatically early with respect to overt subject occurrence. Presumably earlier data collection would have shown subject occurrence patterns in these studies similar to the patterns attested in the children referred to here. More importantly, null subject use in Spanish is discourse-sensitive and 11 longitudinally studied child Spanish-speakers and child Catalan-speakers, in contrast to child German speakers, begin to use overt subjects at the statistically equivalent developmental moment at which other discourse-sensitive, left-edge constructions (wh-questions and fronted objects) begin to be used. A theory of why null subjects are overused, Interface Delay (Grinstead 1998, 2004), and of why this overuse is particularly severe in children with SLI, Interface Deficit (Grinstead et al. 2009, 2013), may shed some light on the patterns observed so far, and make predictions for what we expect to find with respect to null subject use in child Spanish-speakers with SLI. Before moving on the Interface Deficit, let us review what is known about SLI in Spanish.

5. Interface Delay and Interface Deficit

Jackendoff (1987, 1997) proposes a model of cognitive architecture he refers to as Representational Modularity. On this view, mental faculties have representations that are specific to their domains and the computations that are performed upon them are similarly faculty-specific. In particular, he takes

evidence from primate auditory and visual perception (Bradley et al. 1996; Rauschecker et al. 1995; Weinberger 1995) to argue that informational encapsulation (originally suggested by Fodor (1983)) is a natural product of the fact that the mental representations of particular types of sense data are specific to the kind of data that they represent. Following this view, Grinstead (1998) proposes that the domain of discourse-pragmatics and the domain of syntax are fundamentally constituted of different kinds of representations and that the delay in children's use of discourse-sensitive constructions follows from the faculty-specific nature of each domain's representations. Interface Delay occurs because the intermediate links that must exist between faculty-specific representations appear to take longer to develop than faculty-internal representations do. A specific consequence of this phenomenon is that any cognitive process that requires an interface among multiple domains will develop more slowly than processes that do not require this type of interaction.[4]

In the domain of language disorders, Grinstead et al. (2013) claim that this failure to communicate between syntax and discourse-pragmatics is responsible not only for the prolonged delay in tense marking in children with SLI, reported by Rice and Wexler (1996) among others, but also for the more general pattern of all discourse-sensitive constructions being slow to develop in the language of children with SLI, while discourse-insensitive constructions are acquired with relatively little difficulty. Specifically, they review evidence that discourse-sensitive constructions including definite articles (e.g., Anderson and Souto 2005), direct object clitics (e.g., De la Mora et al. 2003) and verb tense (Grinstead et al. 2013) are problematic for child Spanish-speakers, diagnosed with SLI, while discourse-insensitive constructions, such as nominal plural marking and noun-adjective agreement (Grinstead et al. 2008), are not problematic. This special difficulty in the development of discourse-sensitive constructions for children with SLI is referred to as Interface Deficit.

What all of the discourse-sensitive constructions referred to here have in common is a reliance on speakers accessing a representation of the Conversational Common Ground (Stalnaker 1974). Noun phrases are referred to with a definite article if the speaker presupposes that interlocutors are familiar with the referent, i.e., that the referent is salient in the Conversational Common Ground. Direct object (clitic) pronouns, similarly, can only be produced under the presupposition that interlocutors are familiar with the antecedent. Verb tense, in an analogous way, expresses a kind of anaphora from the perspective of the speaker between the speech time-event time relationship prominent in the discourse representation shared with interlocutors. To illustrate the discourse-dependent property of tense interpretation, and its similarity to nominal anaphora, consider the following examples. In these felicitous adult English root infinitives, the grammatical aspect of the verbs in (13b), (14b) and (15b) is morphologically expressed, but the absolute temporal interpretation of the expressions in b. are only interpretable by virtue of temporal anaphora with the expressions in (13a), (14a) and (15a).

13. a. What is/was Wallace doing, Gromit?

    b. Eating cheese.

14. a. What does/will Wallace want to do, Gromit?

    b. Eat cheese.

15. a. What has/had Wallace done, Gromit?

    b. Eaten cheese.

Discourse-sensitive constructions, such as tense marking, contrast with discourse-insensitive constructions, including nominal plural marking and noun-adjective agreement. In order for nouns

[4] See Grinstead et al. (1998) for an argument that Interface Delay occurs between language and number in the development of the counting process.

to be marked plural, no access to Conversational Common Ground is required. Similarly, in order for nouns and adjectives in Spanish and other languages to agree in their markings for number and gender, grammar appears to use strictly "local" morphosyntactic relationships that do not depend on the larger discourse. Oetting (1993) and Oetting and Rice (1993) show that child English speakers with SLI do not have problems with plural marking on nouns and Grinstead et al. (2008) show that child Spanish-speakers with SLI do not have problems with either plural marking on nouns or noun-adjective agreement.

If the Interface Deficit account of SLI is correct, discourse-sensitive constructions including tense marking and null subject use should be particularly problematic for children with SLI, and might be helpful in identifying them.

6. Discriminant Function Analysis

The objective of analyzing the language systems of children with language disorders with a discriminant function analysis is to determine whether particular linguistic properties or collections of properties can be useful for identifying children with the disorders. Bedore and Leonard (1998) show that discriminant functions derived from a verbal morpheme composite, a nominal morpheme composite and MLU in morphemes (MLUm), generated from spontaneous production data, can identify 4 year-old English-speaking children with SLI with fair (<80% accuracy) to good (<90% accuracy) levels of sensitivity (identification of the SLI children) and specificity (identification of the typically-developing age control children). In particular, their Verb Morpheme + MLUm function yielded good and relatively balanced sensitivity (94.7%) and specificity (94.7%). In contrast, Moyle et al. (2011) showed that the same measures, when applied to a sample of older children (mean age = 7;9), produced less effective classification in the fair accuracy range, e.g., their Verb Morpheme + MLUm function produced 74% sensitivity and 84% specificity (79% overall). This is consistent with findings suggesting that tense marking in child English SLI reaches near-typically-developing levels at 8;0 (Rice et al. 1998, 1999).

Discriminant function analysis is particularly important for research on SLI, which is unfortunately a condition diagnosed using primarily exclusionary, as opposed to inclusionary, criteria (Tager-Flusberg and Cooper 1999). Following our Interface Deficit theory of SLI, using the discourse-sensitive construction of tense marking in child Spanish, Grinstead et al. (2013) demonstrate 89% sensitivity and 89% specificity in the identification of children with SLI, approaching the "good" level of precision specified by Plante and Vance (1994), using an elicited production measure of tense in a sample of monolingual Spanish-speaking children in Mexico City. Given that this discourse-sensitive grammatical construction was useful for the identification of Spanish-speaking children with SLI, we ask whether another discourse-sensitive construction, overt subject use, might also prove useful as a clinical marker of SLI in Spanish. Further, we ask whether the addition of a "mixed" discourse-sensitive/discourse-insensitive measure, such as MLUw, might improve accuracy, as it appeared to in both Bedore and Leonard (1998) and Moyle et al. (2011).[5]

7. Research Questions

1. Do monolingual child Spanish-speakers diagnosed with SLI use more null subjects than do typically-developing age controls, as Interface Deficit predicts?
2. If so, can null subject rates play a role in the formation of a discriminant function for the selective identification of children with SLI?
3. Will the addition of MLUw to more strictly discourse-sensitive measures improve discriminant accuracy, as in previous research?

---

[5] Note that MLU measured in words and MLU measured in morphemes correlate at 0.9 in previous studies of Spanish language development (e.g., Aguado Alonso 2000) and are taken to be roughly equivalent measures here.

8. Methods

8.1. Participants

Forty monolingual Spanish-speaking children participated in this study. Each child produced a spontaneous speech sample, interacting with adult native speakers of the Spanish of Mexico City, trained as speech-language pathologists, neuropsychologists or as pediatric neuropsychologists. The children in the study came from a daycare center/preschool and a speech and hearing clinic that service a broad socioeconomic spectrum of children in Mexico City. Twenty of the children were diagnosed with SLI (age range = 58–76 months, mean age = 66 months, SD = 5.8 months), using conventional criteria (following Leonard 2014) and 20 of them were typically-developing (age range = 58–79 months, mean age = 67 months, SD = 5.3 months). The ages of the two groups were not significantly different (t(38) = 0.695, $p = 0.491$).

Following Leonard (2014), our inclusive criterion for the SLI group was a score on a standardized language test of 1.25 SD below the mean. To that end, children took four subtests of the Batería de Evaluación de Lengua Española or BELE (Rangel et al. 1988), which was normed in Mexico City. Children were given the BELE receptive grammar and vocabulary tests ("Comprensión Gramatical" and "Adivinanzas") as well as its expressive grammar and vocabulary tests ("Producción Dirigida" and "Definiciones"). Though there are no published validity studies of the BELE, earlier work (Grinstead et al. 2013) reports significant correlations between children's ($n = 29$) BELE scores and spontaneous speech measures, including the Subordination Index ($r = 0.570$, $p = 0.001$), MLUm ($r = 0.717$, $p < 0.001$) and Number of Different Words ($r = 0.702$, $p < 0.001$). Children had to score 1.25 standard deviations below the mean on at least one receptive and one expressive subtest to be included in our SLI sample. All TD children were within 1 standard deviation of the mean for their ages. Mean BELE scores for children in each group with p values resulting from independent samples t-tests are summarized in Table 4.

Table 4. Mean BELE scores, with standard deviations, comparison results and partial Eta squared values.

	SLI (SD)	TD (SD)	*p* Value	Partial Eta Squared
BELE 4 Subtest Composite	18.89 (4.05)	40.05 (10.03)	<0.001	0.654
Elicited Production	1 (1.57)	9.09 (2.57)	<0.001	0.785
Grammaticality Judgment	6.17 (2.71)	11.90 (2.86)	<0.001	0.525
Expressive Vocabulary	6.5 (2.28)	11 (4.59)	0.001	0.278
Receptive Vocabulary	5.22 (1.87)	8.05 (5.90)	0.059	0.093

SLI: specific language impairment; TD: typically developing.

For exclusive criteria, children were given a Spanish translation of the WPPSI (Weschler Preschool and Primary Scale of Intelligence) to measure non-verbal intelligence and had to receive a score of 85 or above to be included in the study. Children were also given a phonological screen in which they were asked to repeat 24 two-syllable nonce words that included the segments used in Spanish to represent tense in word-final position, with appropriate stress. In order to be included in the study, children had to produce at least four out of five correctly from each category. Children were given thorough hearing tests and had to pass them at conventional levels. In addition, parental report and medical history had to suggest no recent episodes of otitis media with effusion in order for a child to be included. Similarly, neurological tests determined that the children had no frank neurological damage. With respect to oral structure and oral motor function, initial examination ruled out structural anomalies and assured normal function. Parental report and family history interviews ruled out concerns pertaining to social and physical interactions.

8.2. Procedures

A parent of each child signed U.S. and Mexican institutional review board-approved informed consent documents in order to participate. Children were video-recorded interacting with researchers for 20–30 min, until roughly 100 utterances had been produced. The interaction was unstructured and consisted of conversations about friends in preschool, siblings, favorite movies, etc. and answers were largely narrative in character.

8.3. Subject Occurrence Coding

As alluded to above, our dependent variable is the proportion of non-imperative verbal utterances that occurred with a null subject. Our assumption is that where a child does not use an overt subject, they have chosen to instead use a null subject pronoun, which is a definite. As we have claimed above, this is of theoretical interest inasmuch as children with SLI have been shown to struggle with definites, including definite articles (Restrepo and Gutiérrez-Clellen 2001; Anderson and Souto 2005), direct object clitics (Merino 1983; De la Mora 2004) and verb tense (Grinstead et al. 2013). While the ways in which discourse-pragmatic considerations predict the expression of an overt subject in particular dialects is very interesting,[6] they were not the focus of our inquiry. Our goal was instead to determine whether a more rapidly calculable rate of overt subject occurrence might be clinically useful, in contrast with the complexity of identifying discourse-pragmatic determinants of subject occurrence, which is substantially more laborious and consequently less likely to be of clinical utility.[7] A noun phrase was counted as an overt subject of a verbal predicate if it occurred in a discourse context that was semantically compatible with its association with that predicate. Subject–verb agreement was taken to be a sufficient, but not a necessary, condition for this association, given children's potential to produce utterances that do not follow subject–verb agreement, such as the following from Grinstead (1998).

16. Eduardo—3;0.28

Yo	quiere	hacerlo
I	want (root + e theme vowel)	do-INF
		CL-ACC-SG-MASC

'I wants to do it.'

17. Carlos—3;3.28

Yo	va	a	buscar
I-NOM	go	STEM	to look for-INF

'I goes to look for.'

18. Graciela—2;6.5

Hace	esto	yo
do (root + e theme vowel)	this	I-NOM

'I does this.'

Utterances including verbal predicates that appeared nonfinite, such as 16–18, were included in our counts.

---

[6] It appears that there are no absolute predictors of subject occurrence (see Otheguy et al. 2007 on predictors of subject pronoun expression) and children may be substantially older than our participants before they attain adult-like command of these variables (see Shin and Cairns 2012, for example).

[7] See Clancy (1997); Allen (2000) and Serratrice (2005) for discourse-pragmatic analyses of some factors influencing overt subject occurrence in null subject child languages.

8.4. Reliability

Spontaneous speech samples were transcribed by native speakers of the Spanish of Mexico City, the same dialect spoken by the children in the sample. Transcribers were initially normed on a series of common transcriptions. Then, each recording session was transcribed by a single transcriber and checked by a second transcriber. Finally, half of all transcripts were randomly selected to have 10% of their utterances re-transcribed by a second transcriber to ensure accuracy. Agreement between transcribers as to each word transcribed ranged from 90–99%, with a mean agreement percentage of 95.4%. A Krippendorf's Alpha Interrater Reliability Coefficient (Hayes and Krippendorf 2007) was then calculated for the interval data represented by each transcriber's number of words per utterance, for each transcript. Alpha values ranged between 0.904 and 0.998, with a mean Krippendorf's alpha value of 0.974.

Each non-imperative verb was also coded for person, number, and tense, as well as for whether or not it was accompanied by an overt subject. Imperative verbs were not examined because of their cross-linguistic tendency to lack overt subjects (see Rivero and Terzi 1995). Fleiss' (1971) Kappa was used to determine the inter-rater reliability among the three Spanish-speaking coders. The calculated Kappa of 0.68 was in the category of substantial agreement per the standard set forth by Landis and Koch (1977). Transcripts were further analyzed for MLUw, using the CLAN programs from the CHILDES Project (MacWhinney 2000).

8.5. Previous Experimental Measures

To construct our discriminant functions, we included two independent tests of tense marking, a discourse-sensitive construction, reported in earlier work. Each child had a score from the Grammaticality Choice Task of tense (n = 35), reported in Grinstead et al. (2009), and a score from an elicited production task measuring tense (n = 29), reported in Grinstead et al. (2009). The Grammaticality Choice Task presents children with two sentences, each uttered by a different puppet. One sentence is the adult-like version (e.g., *Yo abro la boca* 'I open my mouth'), while the other is the putative child-particular version (e.g., *Yo abre la boca* 'I opens my mouth'). The elicited production task has children correct a sentence such as *El gato está rompiendo los platos* ('The cat is breaking the dishes') with a sentence that matches the image presented, such as *El gato está lavando los platos* ('The cat is washing the dishes').

9. Results

9.1. Comparisons

An independent samples t-test was performed to compare the rates of null subject use of the SLI and typically-developing age control groups, which are given in Table 5.[8] The SLI rate was significantly higher than was the TD rate ($t(38) = 3.141$, $p = 0.003$). Notice in the accompanying histogram that the distributions are normal for both groups and shifted towards one (100% null subjects) in the SLI group. Measures of tense by Elicited Production and Grammaticality Choice are also given in Table 5, together with MLUw values. We report p values in the last column, which corresponds to a t-test for MLUw and subject use, which was normally distributed, but to Mann-Whitney tests for the tense measures, which were not and for which the Z scores are reported. Note that the Quadratic Discriminant Function analysis reported below is robust to non-normality as long as it results from skewness and not outliers, which is the case in our data.

[8] These rates are roughly consistent with the 81% null subject expression rate reported for subject pronouns in a small sample of Mexican immigrants to New York in Otheguy et al. (2007), though note that our rate corresponds to all subject noun phrases and not just pronouns, which is plausibly a consequential difference.

Table 5. Mean rates of null subject use, mean length of utterance in words (MLUw), grammaticality choice measures of tense and elicited oroduction measures of tense across SLI and TD age control groups.

	SLI (SD)	TD (SD)	Test Statistic	*p*	Eta Squared
Overt Subject Expression *n*	0.81 (0.09) 20	0.72 (0.09) 20	$t = -3.141$	0.003	0.206
MLUw *n*	3.005 (0.724) 20	5.499 (1.831) 20	$t = 5.895$	<0.001	0.458
Tense—Grammaticality Choice *n*	10.647 (3.481) 17	14.389 (2.704) 18	$Z = -3.096$	0.002	0.117
Tense—Elicited Production *n*	10.643 (3.499) 14	14.400 (0.737) 15	$Z = -3.565$	<0.001	0.206

9.2. Discriminant Functions

A discriminant function analysis was performed to classify participants into groups on the basis of the variables tested by producing a discriminant criterion that maximizes the differences between the groups. Figure 2 illustrates the normal distribution of null subjects in the two populations, which is shifted rightwards in the SLI population, towards 100%. In addition to our subject expression measure, we also have MLUw measures for the same children and verb finiteness measures (grammaticality choice and elicited production) for a subset of these children (grammaticality choice $n = 35$; elicited production $n = 29$, all four measures $n = 26$). Linear discriminant function analysis assumes equal within group variance-covariance matrices, which our data do not have. For this reason, we instead used a quadratic discriminant function analysis, which does not depend upon this assumption (Klecka 1980). Further, discriminant function analysis assumes normally distributed data. Two of our variables are normally distributed (subjects and MLUw) and two are not, resulting from skewness. Because quadratic discriminant function analysis is robust to non-normality caused by skewness and not the presence of outliers (which are absent from our data), we proceeded with the analysis. Various combinations of these four measures (null subject use, MLUw, elicited production tense and grammaticality choice tense) produced the five quadratic discriminant functions given in Table 6.

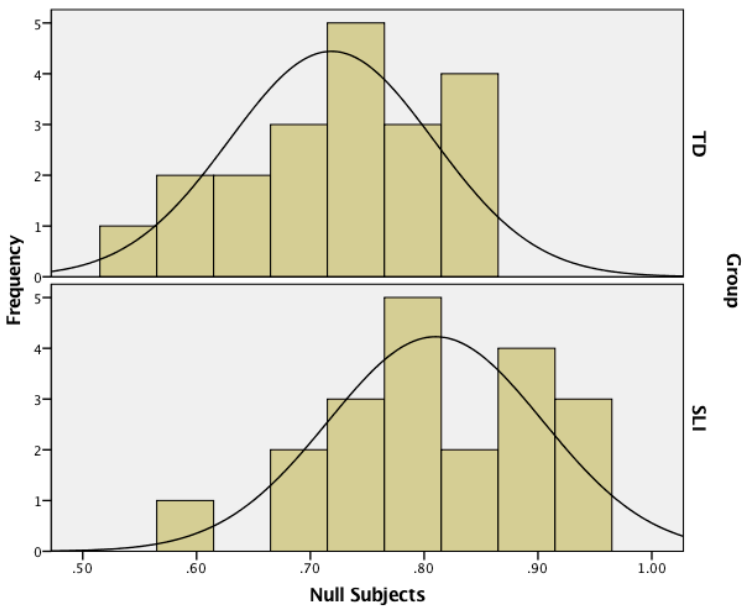


Figure 2. Histogram of rates of null subject use in the specific language impairment (SLI) and age-matched typically-developing control group, in normal distribution (Kolmogorov-Smirnov with Lilliefors significance correction and Shapiro–Wilk *p* values are >0.05 for both groups). TD: typically developing.

Table 6. Five discriminant functions including subject expression, MLUw, grammaticality choice and elicited production tense results.

Variables	*n*	Sensitivity (% Correct SLI)	Specificity (% Correct TD)	Mean % Correct
Subjects	40	70 (14/20)	65 (13/20)	67.5
MLUw	40	95 (19/20)	80 (16/20)	87.5
Subjects, MLUw	40	95 (19/20)	75 (15/20)	85.0
Grammaticality Choice, Elicited Production	26	66.7 (8/12)	92.9 (13/14)	80.8
Subjects, Grammaticality Choice, Elicited Production	26	83.3 (10/12)	85.7 (12/14)	84.6
Subjects, MLUw, Grammaticality Choice, Elicited Production	26	91.7 (11/12)	92.9 (13/14)	92.3

In Table 6, the classification statistics of five distinct discriminant functions computed from our variables are reported. In the second row, we see that null subject use by itself has an accuracy of 67.5%. Thus, this discourse-sensitive measure, by itself, does not identify children with high levels of accuracy. In the third row, we see that MLUw by itself, which is a hybrid measure of both discourse-sensitive and discourse-insensitive morphemes, produces 87.5% correct identification of children, with the SLI identification (sensitivity) being very high (95%, 19/20) and with correct identification of TD children (specificity) substantially lower (80%, 16/20). In the fourth row, we see that two tense measures produce fair accuracy (80.8%), which, in contrast to MLUw, is more accurate in identifying the TD children (specificity—92.9%, 13/14) than the SLI children (sensitivity—66.7%, 8/12). In the fifth row, we see discourse-sensitive constructions only (subjects and the two tense measures) produce 84.6% accuracy, which is relatively balanced between sensitivity (83.3%, 10/12) and specificity (85.7%, 12/14). In the last row, we see that a function consisting of our three discourse-sensitive measures as well as our hybrid discourse-sensitive/discourse-insensitive measure, MLUw, produces a highly accurate 92.3% classification that is balanced between sensitivity (91.7%, 11/12) and specificity (92.9%, 13/14).

This fifth discriminant function yields discrimination above the 90% level, which Plante and Vance (1994) characterize as a "good" level of sensitivity and specificity. For this function, the overall Chi-square test was significant (Wilks $\lambda = 0.296$, Chi-square = 26.790, df = 4, Canonical correlation = 0.839, $p < 0.001$) and the function accounted for 70% of the variance in diagnosis. The standardized canonical discriminant function coefficients, illustrated in Table 7, were, in descending order, subjects (−0.603), elicited production of tense (0.597), MLUw (0.529) and grammaticality choice test of tense (0.159). The structure matrix values, given in the third column of Table 7 are roughly consistent with the standardized coefficients in the sense that subjects, MLUw and the elicited production measure of tense were the most effective variables at distinguishing children in the two diagnostic categories.

Table 7. Standardized canonical discriminant function coefficient and structure matrix for the discourse sensitive function.

	Standardized Canonical Discriminant Function Coefficient	Structure Matrix
Subjects	−0.603	−0.459
Tense-Elicited Production	0.597	0.507
MLUw	0.529	0.688
Tense-Grammaticality Choice	0.159	0.357

10. Discussion

In this study, we have reviewed evidence for the claim that typically-developing child Spanish-speakers pass through a no-overt subject stage. Two different research groups looking at the question from before children get to MLUw levels of 1.5, or so, indeed find in 11 children's spontaneous production data that such a stage exists. Perhaps of more interest to understanding the differences between the syntax of overt subjects in the two language types, the same two groups show that these 11 children begin using overt subjects at the same time that they begin using other plausibly left-peripheral constructions, including wh- questions and fronted objects. To make it clear that not all

syntactic constructions arise at the same time in all languages, Grinstead and Spinner (2009) show that overt subjects in child German, an overt subject language, begin to be used significantly earlier than do wh- questions or fronted objects. We take this to be strong confirmation that the Interface Delay Hypothesis is correct in its predictions that definite NPs, including null subjects, should be overused in child language, as has been documented at least since Piaget's observation of "egocentric" linguistic behavior (Piaget 1929).

Note that this situation is starkly different from child English. Child Spanish-speakers are presented with input that overwhelmingly uses null subjects–pronominal definites. Child English-speakers are overwhelmingly presented with overt subject utterances. What unifies the two phenomena is that child English speakers probably use a nonfinite PRO, driven by non-adultlike tense marking. They cannot license overt subjects because they lack the means to assign structural nominative Case to subject position. The tense deficit, on our account, stems from their overuse of a kind of definite tense form, to wit, the *Eat cheese* type we referred to in examples (13b), (14b), and (15b). This kind of production from children has highly visible, noticeable syntactic consequences in child English because adult listeners expect an overt subject every time. Something similar is highly likely to occur in Spanish, but it is much less noticeable because it is impossible to phonetically distinguish between PRO and *pro*. Further, the syntax of null subjects in spanish is different and far from settled science, though it may involve incorporation of tense and a pronominal subject into the verb, if Ordóñez (1997) is correct, which could have dramatic consequences for acquisition. We do not ultimately know if occurring in a specifier position is harder to learn or easier to learn than is pronominal incorporation. Our ability to make this determination empirically seems limited. As to why child English speakers do not choose overuse of null subjects as an instance of overuse of definites, it would be because they are not exposed to them. They are not in the input, at least not of the same quality or in the same quantity that exists in Spanish.

Turning to our research questions, we can say that (1) yes, monolingual child Spanish-speakers with SLI use significantly more null subjects than do typically-developing age controls; (2) yes, null subject use appears able to play a role in forming a discriminant function capable of identifying Spanish-speaking children with SLI and (3) yes, adding MLUw to a discriminant function that consists of discourse-sensitive grammatical morphemes increases its accuracy, as in Bedore and Leonard (1998) and Moyle et al. (2011).

The fact that monolingual child Spanish-speakers with SLI use more null subjects than do age-matched controls is very reminiscent of the results presented in Grinstead et al. (2013) showing that child Spanish-speakers with SLI use more root nonfinite verbs than do typically-developing age-matched controls. In both cases, there is a linguistic phenomenon of delayed development relative to other dimensions of language (e.g., nominal plural marking and noun-adjective agreement) that is characteristic of typically developing child Spanish-speakers, which is more severe and prolonged in child Spanish-speakers with SLI. As we have attempted to make clear, we believe that both phenomena stem from the same root cause, namely, Interface Deficit. On that note, it is interesting to observe that the discriminant function in Table 6 consisting of Subjects and the two tense measures achieved relatively balanced, "fair" sensitivity (83.3%) and specificity (85.7%) of 84.6%. Crucially, the function created by the sum of these three measures is superior to the two tense measures by themselves (80.8%) or subjects by themselves (67.5%). This is theoretically interesting because though null subject use is a function of the nominal domain and tense is a function of the verbal domain, they share anaphoric reference as a common trait for determining use of a morpheme. As alluded to above, null subjects are used in Spanish, among other reasons, when the speaker's model of the interlocutor's representation of the Conversational Common Ground presupposes shared knowledge of the identity of the antecedent. Similarly, verb tense seeks to express the relationship between speech time and event time, to maintain discourse coherence with conversational participants. Once a speech time-event time relationship is fixed for a stretch of discourse, verb tense refers back to that event time anaphorically (e.g., Reichenbach 1947; Bittner 2011). Perhaps the difference in nominal vs. verbal anaphora is such

that the variance in children's use of each construction is somewhat unique, which allows for greater precision of identification in a combined discriminant function. Though SLI likely has more than one locus and more than one subtype, we nonetheless aspire to add additional discourse-sensitive constructions to future discriminant functions in hopes of improving accuracy to see how far this dimension of SLI can take us on the road to improved diagnosis.

On another note, the fact that child Spanish-speakers with SLI showed differences from typically-developing children even in unstructured discourse may mean that this is a particularly sensitive construction for them. That is, structured spontaneous production techniques, including story retell, tend to produce high rates of errors per t-unit in Spanish, as in Restrepo (1998), which were high enough to produce significant differences between TD and SLI groups. In Grinstead et al. (2013), however, we showed that errors per t-unit in transcripts resulting from our unstructured spontaneous production technique did not produce significant differences between TD and SLI groups. Thus, children with SLI appear to produce fewer errors in unstructured discourse than they do in structured discourse. That being the case, if they produce more null subjects than do TD controls, even in unstructured discourse, it suggests that the use of the construction may not be something that can be easily compensated for.

In future studies of subject use, we hope to look more in detail at the discourse-pragmatic functions of overt subjects. While we have shown that raw proportions do vary across SLI and TD groups, it remains to be seen whether the distributions of these subjects are similar or different across the same groups.

Author Contributions: B.F.-Á, M.C.-S., M.V.-M. and J.D.l.M. collected the data. P.L., A.P., M.C.-S., M.V.-M. and J.D.l.M. analyzed certain parts of the data. J.G. conceived of the project, did the statistical analysis and wrote the manuscript, in conjunction with the rest of the team.

Funding: This study was supported by an Arts and Sciences Innovation Grant from The Ohio State University to the first author, as well as by an Undergraduate Research Fellowship to the third author.

Acknowledgments: We are grateful to Silvia Romero for her discussion of the BELE and to Laida Restrepo for discussions of our errors per t-unit results.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results'.

References

Aguado Alonso, Gerardo. 2000. *El desarrollo de la morfosintaxis en el niño*. Madrid: Ciencias de la Educación Preescolar y Especial.

Aguado-Orea, Javier, and Julian M. Pine. 2002. There is no evidence for a 'no overt subject' stage in early child Spanish: A note on Grinstead (2000). *Journal of Child Language* 29: 865–74. [CrossRef] [PubMed]

Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parameterizing AGR: Word order, v-movement and EPP checking. *Natural Language & Linguistic Theory* 16: 491–539.

Allen, Shanley E. M. 2000. A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics* 38: 483–521. [CrossRef]

Anderson, Raquel T., and Sofía M. Souto. 2005. The use of articles by monolingual Puerto Rican Spanish-speaking children with specific language impairment. *Applied Psycholinguistics* 26: 621–47. [CrossRef]

Antinucci, Francesco, and Domenico Parisi. 1973. Early Language Acquisition: A Model and Some Data. In *Studies in Child Language Development*. Edited by Charles A. Ferguson and Dan I. Slobin. New York: Holt, Rinehart & Winston, pp. 607–19.

Avram, Larisa, and Martine Coene. 2009. Early subjects in a null subject language. In *New Directions in Language Acquisition: Romance Languages in the Generative*. Edited by Pedro Guijarro-Fuentes and Laura Domínguez. Cambridge: Cambridge Scholars Publishing, pp. 189–217.

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373–415.

Baker, Mark. 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.

Bates, Elizabeth. 1976. *Language and Context: The Acquisition of Pragmatics*. New York: Academic Press.

Bedore, Lisa M., and Laurence B. Leonard. 1998. Specific Language Impairment and Grammatical Morphology: A Discriminant Function Analysis. *Journal of Speech, Language, and Hearing Research* 41: 1185–92. [CrossRef] [PubMed]

Bel, Aurora. 2003. The syntax of subjects in the acquisition of Spanish and Catalan. *Probus* 15: 1–26. [CrossRef]

Bittner, Maria. 2011. Time and Modality Without Tense or Modals. In *Tense Across Languages*. Edited by Renate Musan and Monika Rathert. Tübingen: Niemeyer, pp. 147–88.

Bloom, Lois, Peggy J. Miller, and Lois Hood. 1975a. Variation and reduction as aspects of competence in language development. In *Minnesota Symposia on Child Psychology*. Edited by Anne D. Pick. Minneapolis: University of Minnesota Press, vol. 9, pp. 3–55.

Bloom, Lois, Patsy M. Lightbown, and Lois Hood. 1975b. Structure and variation in child language. *Monographs of the Society for Research in Child Development* 40: 1–97. [CrossRef]

Bloom, Paul. 1991. Subjectless sentences in child language. *Linguistic Inquiry* 21: 491–504.

Bloom, Paul. 1993. Grammatical continuity in language development: The case of subjectless sentences. *Linguistic Inquiry* 24: 721–34.

Bradley, David C., Marsha Maxwell, Richard A. Andersen, Martin S. Banks, and Krishna V. Shenoy. 1996. Mechanisms of heading perception in primate visual cortex. *Science* 273: 1544–47. [CrossRef] [PubMed]

Brown, Roger W. 1973. *A First Language: The Early Stages*. Cambridge: Harvard University Press.

Brown, Roger W., and Colin Fraser. 1964. The acquisition of syntax. In *The Acquisition of Language*. Edited by Ursula Bellugi and Roger W. Brown. Chicago: University of Chicago Press, pp. 43–78.

Camacho, José A. 2013. *Null Subjects*. Cambridge: Cambridge University Press.

Casanovas, Lourdes. 1999. El sujeto en el Catalán coloquial. *Revista Española de Lingüística* 29: 105–31.

Chomsky, Noam. 1981. *Lectures On Government And Binding: The Pisa Lectures*. Berlin: Mouton de Gruyter.

Clancy, Patricia. 1997. Discourse motivations of referential choice in Korean acquisition. In *Japanese/Korean Linguistics*. Edited by Ho-min Sohn and John Haig. Stanford: CSLI, vol. 6, pp. 639–59.

De la Mora, Juliana, Johanne Paradis, John Grinstead, Blanca Flores, and Myriam Cantú-Sánchez. 2003. Object Clitics in Spanish Speaking Children with and without SLI. Presented at the Symposium for Research on Child Language Disorders, Madison, WI, USA.

De la Mora, Juliana. 2004. Direct Object Clitics and Determiners in Spanish Specific Language Impairment. Unpublished Master's Thesis, University of Alberta, Edmonton, AB, Canada.

Fleiss, Joseph L. 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin* 76: 378–82. [CrossRef]

Fodor, Jerry A. 1983. *Modularity of Mind*. Cambridge: MIT Press.

Gerken, LouAnn. 1991. The metrical basis for children's subjectless sentences. *Journal of Memory and Language* 30: 431–51. [CrossRef]

Gerken, LouAnn. 1994. Young children's representations of prosodic phonology: Evidence from English speakers' weak syllable productions. *Journal of Memory and Language* 33: 19–38. [CrossRef]

Grela, Bernard. 2003a. Production-Based Theories May Account for Subject Omission in Both Normal Children and Children with SLI: A Case Study. *Journal of Speech-Language Pathology and Audiology* 27: 221–28.

Grela, Bernard. 2003b. The omission of subject arguments in children with specific language impairment. *Clinical Linguistics & Phonetics* 17: 153–69.

Grela, Bernard, and Laurence B. Leonard. 1997. The Use of Subject Arguments by Children with Specific Language Impairment. *Clinical Linguistics & Phonetics* 11: 443–53.

Grinstead, John. 1998. Subjects, Sentential Negation and Imperatives in Child Spanish and Catalan. Ph.D. dissertation, UCLA, Los Angeles, CA, USA.

Grinstead, John. 2000. Case, Inflection and Subject Licensing in Child Catalan and Spanish. *Journal of Child Language* 27: 119–55. [CrossRef] [PubMed]

Grinstead, John. 2004. Subjects and Interface Delay in Child Spanish and Catalan. *Language* 80: 40–72. [CrossRef]

Grinstead, John, Myriam Cantú-Sánchez, and Blanca Flores-Avalos. 2008. Canonical and Epenthetic Plural Marking in Spanish-Speaking Children with Specific Language Impairment. *Language Acquisition* 15: 329–49. [CrossRef]

Grinstead, John, Juliana De la Mora, Mariana Vega-Mendoza, and Blanca Flores-Avalos. 2009. An Elicited Production Test of the Optional Infinitive Stage in Child Spanish. In *Proceedings of the 3rd Conference of*

Generative Approaches to Language Acquisition—North America. Edited by Jean Crawford, Koichi Otaki and Masahiko Takahashi. Storrs: Cascadilla Press, pp. 36–45.

Grinstead, John, Alisa Baron, Mariana Vega-Mendoza, Juliana De la Mora, Myriam Cantú-Sánchez, and Blanca Flores. 2013. Tense marking and spontaneous speech measures in Spanish specific language impairment: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research* 56: 352–63. [CrossRef]

Grinstead, John, Paij Lintz, Mariana Vega-Mendoza, Juliana De la Mora, Myriam Cantú-Sánchez, and Blanca Flores-Avalos. 2014. Evidence of optional infinitive verbs in the spontaneous speech of Spanish-speaking children with SLI. *Lingua* 140: 52–66. [CrossRef]

Grinstead, J., Jeff MacSwan, Susan Curtiss, and Rochel Gelman. 1998. The Independence of Language and Number. In *BUCLD 22: Proceedings of the 22nd annual Boston University Conference on Language Development*. Edited by Annabel Greenhill, Mary Hughes, Heather Littlefield and Hugh Walsh. Somerville: Cascadilla Press, pp. 303–13.

Grinstead, John, and Patti Spinner. 2009. The Clausal Left Periphery in Child Spanish and German. *Probus* 21: 51–82. [CrossRef]

Guilfoyle, Eithne. 1984. *The Acquisition of Tense and the Emergence of Lexical Subjects*. Montreal: McGill University.

Guilfoyle, Eithne, and Máire Noonan. 1992. Functional categories and language acquisition. *Canadian Journal of Linguistics/Revue canadienne de Linguistique* 37: 241–72. [CrossRef]

Hamann, Cornelia. 1996. Null arguments in German child language. *Language Acquisition* 5: 155–208. [CrossRef]

Hamann, Cornelia, and Kim Plunkett. 1998. Subjectless sentences in child Danish. *Cognition* 69: 35–72. [CrossRef]

Hayes, Andrew F., and Klaus Krippendorf. 2007. Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures* 1: 77–89. [CrossRef]

Hyams, Nina. 1986. *Language Acquisition and the Theory of Parameters*. Dordrecht: D. Reidel Publishing Company.

Hyams, Nina. 2011. Missing subjects in early child language. In *Handbook of Generative Approaches to Language Acquisition*. Edited by Jill G. De Villiers and Thomas Roeper. New York: Springer, vol. 41, pp. 13–52.

Hyams, Nina, and Kenneth Wexler. 1993. On the Grammatical Basis of Null Subjects in Child Language. *Linguistic Inquiry* 24: 421–59.

Jackendoff, Ray S. 1987. *Consciouness and the Computational Mind*. Cambridge: MIT Press.

Jackendoff, Ray S. 1997. *The Architecture of the Language Faculty*. Cambridge: Massachusetts Instit Technology Press.

Klecka, William R. 1980. *Discriminant Analysis*. Quantitative Applications in the Social Sciences Series; Thousand Oaks: Sage Publications.

Kramer, Irene. 1993. The licensing of subjects in early child languages. In *Papers on Case and Agreement II*. Edited by Colin Phillips. Cambridge: MIT Working Papers in Linguistics, vol. 19, pp. 197–212.

Landis, J. Richard, and Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33: 159–74. [CrossRef] [PubMed]

Leonard, Laurence. 2014. *Children With Specific Language Impairment*. Cambridge: MIT Press.

Liceras, Juana M., Aurora Bel, and Susana Perales. 2006. 'Living with Optionality': Root Infinitives, Bare Forms and Inflected Forms in Child Null Subject Languages. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*. Edited by Nuria Sagarra and Almeida Jacqueline Toribio. Somerville: Cascadilla Press, pp. 203–16.

Loeb, Diane F., and Laurence B. Leonard. 1988. Specific Language Impairment and Parameter Theory. *Clinical Linguistics & Phonetics* 2: 317–27.

Lorusso, Paolo, Claudia Caprin, and Maria Teresa Guasti. 2004. Overt Subject Distribution in Early Italian Children. In *A Supplement to the Proceedings of Boston University Conference on Language Development*. Edited by Alejna Brugos, Manuella R. Clark-Cotton and Seungwan Ha. Somerville: Cascadilla Press.

MacWhinney, Brian. 2000. *The CHILDES Project*. Mahwah: Lawrence Erlbaum.

Maratsos, Michael P. 1974. Preschool children's use of definite and indefinite articles. *Child Development* 45: 446–55. [CrossRef]

Maratsos, M.P. 1976. *The Use of Definite and Indefinite Reference in Young Children: An Experimental Study of Semantic Acquisition*. New York: Cambridge University Press.

Merino, Barbara J. 1983. Language development in normal and language handicapped Spanish-speaking children. *Hispanic Journal of Behavioral Sciences* 5: 379–400. [CrossRef]

Moyle, Maura J., Courtney Karasinski, Susan Ellis Weismer, and Brenda K. Gorman. 2011. Grammatical morphology in school-age children with and without language impairment: A discriminant function analysis. *Language, Speech, and Hearing Services in Schools* 42: 550–60. [CrossRef]

O'Grady, William, Ann M. Peters, and Deborah Materson. 1989. The transition from optional to required subjects. *Journal of Child Language* 16: 513–29. [CrossRef] [PubMed]

Oetting, Janna B. 1993. Language-Impaired and Normally Developing Children's Acquisition of English Plural. Ph.D. dissertation, University of Kansas, Lawrence, KS, USA.

Oetting, Janna B., and Mabel L. Rice. 1993. Plural Acquisition in Children with Specific Language Impairment. *Journal of Speech and Hearing Research* 36: 1236–48. [CrossRef] [PubMed]

Ordóñez, Francisco. 1997. Word Order and Clause Structure in Spanish and Other Romance Languages. Ph.D. dissertation, City University of New York, New York, NY, USA.

Ordóñez, Francisco, and Esthela Treviño. 1999. Left Dislocated Subjects and the Pro-Drop Parameter: A Case Study of Spanish. *Lingua* 107: 39–68. [CrossRef]

Orfitelli, Robyn, and Nina Hyams. 2008. An Experimental Study of children's comprehension of null subjects: Implications for grammatical/performance accounts. In *BUCLD 32: Proceedings of the 32nd annual Boston University Conference on Language Development*. Edited by Harvey Chan, Heather Jacob and Enkeleida Kapia. Somerville: Cascadilla Press, pp. 335–46.

Otheguy, Ricardo, Ana C. Zentella, and David Livert. 2007. Language and Dialect Contact in Spanish in New York: Toward the Formation of a Speech Community. *Language* 83: 770–802. [CrossRef]

Otheguy, Ricardo, and Ana C. Zentella. 2012. *Spanish in New York: Language Contact, Dialectal Leveling and Structural Continuity*. Oxford: Oxford University Press.

Piaget, Jean. 1929. *The Child's Conception of the World*. London: Kegan Paul Trench Trubner.

Plante, Elena, and Rebecca Vance. 1994. Selection of preschool language tests: A data-based approach. *American Journal of Speech-Language Pathology* 4: 70–76. [CrossRef]

Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar and the Structure of IP. *Linguistic Inquiry* 20: 365–424.

Radford, Andrew. 1990. *Syntactic Theory and the Acquisition of English Syntax*. Cambridge: Basil Blackwell.

Rangel, Elena, Silvia Romero, and Margarita Gómez. 1988. *Batería de evaluación de la lengua española para niños de 3 a 11 años: Manual de aplicación, calificación e interpretación*. Mexico City: Secretaría de Educación Pública, Dirección General de Educación Especial.

Rauschecker, Josef P., Biao Tian, and Marc Hauser. 1995. Processing complex sounds in the macaque nonprimary auditory cortex. *Science* 268: 111–14. [CrossRef] [PubMed]

Reichenbach, Hans. 1947. *Elements of Symbolic Logic*. New York: MacMillan.

Restrepo, María A. 1998. Identifiers of Predominantly Spanish-Speaking Children with Language Impairment. *Journal of Speech, Language, and Hearing Research* 41: 1398–411. [CrossRef] [PubMed]

Restrepo, María A., and Vera F. Gutiérrez-Clellen. 2001. Article Use in Spanish-Speaking Children with Specific Language Impairment. *Journal of Child Language* 28: 433–52. [CrossRef] [PubMed]

Rice, Mabel L., and Kenneth Wexler. 1996. Toward Tense as a Clinical Marker of Specific Language Impairment in English-Speaking Children. *Journal of Speech and Hearing Research* 39: 1239–57. [CrossRef] [PubMed]

Rice, Mabel L., Kenneth Wexler, and Scott Hershberger. 1998. Tense over Time: The Longitudinal Course of Tense Acquisition in Children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research* 41: 1412–31. [CrossRef] [PubMed]

Rice, Mabel L., Kenneth Wexler, and Sean M. Redmond. 1999. Grammaticality Judgments of an Extended Optional Infinitive Grammar: Evidence from English-Speaking Children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research* 42: 943–61. [CrossRef] [PubMed]

Rivero, Maria Luisa, and Arhonto Terzi. 1995. Imperatives, V-movement and Logical Mood. *Journal of Linguistics* 31: 301–32. [CrossRef]

Rizzi, Luigi. 1993. Some Notes on Linguistic Theory and Language Development: The Case of Root Infinitives. *Language Acquisition* 3: 371–93. [CrossRef]

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar: Handbook in Generative Syntax*. Edited by Liliane Haegeman. Dordrecht: Kluwer, pp. 281–337.

Rizzi, Luigi. 2005. On the grammatical basis of language development: A case study. In *The Oxford Handbook of Comparative Syntax*. Edited by Guglielmo Cinque and Richard S. Kayne. New York: Oxford University Press, pp. 70–109.

Sano, Tetsuya, and Nina Hyams. 1994. Agreement, Finiteness and the Development of Null Arguments. In *NELS 24: Proceedings of the North East Linguistic Society*. Edited by Mercè González. Amherst: University of Massachusetts, pp. 544–58.

Serra, Miquel, and Rosa Solé. 1986. Language Acquisition in Catalan and Spanish Children. Unpublished Research Project Manuscript. Universitat de Barcelona; Universitat Autònoma de Barcelona, Barcelona.

Serratrice, Ludovica. 2005. The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied Psycholinguistics* 26: 437–62. [CrossRef]

Shin, Naomi L., and Helen S. Cairns. 2012. The development of NP selection in school-age children: Reference and Spanish subject pronouns. *Language Acquisition: A Journal of Developmental Linguistics* 19: 3–38. [CrossRef]

Snyder, William. 2001. On the Nature of Syntactic Variation: Evidence from Complex Predicates and Complex Word-Formation. *Language* 77: 324–42. [CrossRef]

Snyder, William, and Karin Stromswold. 1997. The Structure and Acquisition of English Dative Constructions. *Linguistic Inquiry* 28: 281–317.

Stalnaker, Robert. 1974. Pragmatic Presuppositions. In *Semantics and Philosophy*. Edited by Milton K. Munitz and Peter K. Unger. New York: New York University Press, pp. 197–219.

Tager-Flusberg, Helen, and Judith Cooper. 1999. Present and Future Possibilities for Defining a Phenotype for Specific Language Impairment. *Journal of Speech, Language, and Hearing Research* 42: 1275–78. [CrossRef] [PubMed]

Valian, Virginia. 1990. A problem for parameter-setting models of language acquisition. *Cognition* 35: 105–22. [CrossRef]

Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40: 21–81. [CrossRef]

Valian, Virginia, James Hoeffner, and Stephanie Aubry. 1996. Young children's imitation of sentence subjects: Evidence of processing limitations. *Developmental Psychology* 32: 153–64. [CrossRef]

Villa-García, Julio, and William Snyder. 2009. On the Acquisition of Overt Subjects, Topics and Wh-Questions in Spanish. In *Language Acquisition and Development: Proceedings of GALA 2009*. Edited by Ana Castro, João Costa, Maria Lobo and Fernanda Pratas. Cambridge: Cambridge Scholars Publishing, pp. 493–504.

Weinberger, Norman M. 1995. Dynamic regulation of receptive fields and maps in the adult sensory cortex. *Annual Review of Neuroscience* 18: 129–58. [CrossRef] [PubMed]

Zubizarreta, María Luisa. 1998. *Prosody, Focus and Word Order*. Cambridge: MIT Press.

languages

MDPI

Article

Acquisition of L2 French Object Pronouns by Advanced Anglophone Learners

Julia Herschensohn [1,*] and Randall Gess [2,3]

1 Department of Linguistics, University of Washington, Box 352425, Seattle, WA 98195-2425, USA
2 Department of French, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada; randall.gess@carleton.ca
3 School of Linguistics and Language Studies, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada
* Correspondence: herschen@uw.edu; Tel.: +1-206-543-2046

Received: 22 December 2017; Accepted: 19 April 2018; Published: 3 May 2018

Abstract: The role native language transfer plays in L2 acquisition raises the question of whether L1 constitutes a permanent representational deficit to mastery of the L2 morphosyntax and prosody or if it can eventually be overcome. Earlier research has shown that beginning and low intermediate Anglophone L2 French learners are insensitive to French morphosyntactic and prosodic constraints in using in situ pronouns transferred from the L1. The prosodic transfer hypothesis (PTH) proposes that native prosodic structures may be adapted to facilitate acquisition of L2 prosodic structure. Our study presents new evidence from three Anglophone advanced learners of L2 French that indicates ceiling performance for pronoun production (99% accuracy in 300 tokens over nine interviews) and grammaticality judgment (98% accuracy). This native-like performance demonstrates target French morphosyntax and prosody, built—as predicted by the PTH—by licensing pronominal free clitics in a new pre-verbal L2 position distinct from post-verbal L1. Furthermore, the learners' data confirms accurate prosody by way of appropriate prominence patterns in clitic + host sequences, correct use of clitics with prefixed verbs, use of stacked pronouns, as well as correct prosodic alternations involving liaison and elision. These results counter impaired representation approaches and suggest early missing inflection may be overcome.

Keywords: second language acquisition; native language transfer; French object pronouns; advanced learners; prosodic transfer hypothesis; French object clitics; strong pronouns

1. Introduction

The role of the native language in L2 acquisition has been a topic of inquiry for decades (e.g., Lado 1957; Schwartz and Sprouse 1996) and is of particular interest in cases where the two languages differ in several dimensions. This notion of transfer raises the question of whether the native language constitutes a permanent representational deficit to mastery of the new morphosyntax (Hawkins and Chan 1997; Tsimpli and Dimitrakopoulou 2007) or if it can eventually be overcome (Lardiere 2007; Rothman 2011). In the arena of prosody, one proposal concerning native language influence is the prosodic transfer hypothesis (PTH, Goad and White 2006, 2008, 2009), which argues that learners' native language prosodic structures may be used to accommodate L2 functional material.

The acquisition of French pronouns is a challenge for Anglophone second language (L2) learners, since the two languages differ morphosyntactically and prosodically for object pronouns.

1. a. I see John

   b. I see him

   c. *I him see

   d. I see'm

2. a. Je vois Jean

   b. Je le vois

   c. *Je vois le

In terms of morphosyntax, English pronouns are located in the canonical object position of the full noun phrase (1a, b) and may optionally be right-cliticized to the verb (1d). In contrast, French object pronoun clitics are left-adjoined to the inflected verb (2b), and are not in the canonical direct object position (2a, c). The acquisition of L2 Romance pronouns has been a subject of investigation to test morphosyntactic transfer hypotheses (White 1996; Hawkins and Franceschina 2004; Granfeldt and Schlyter 2004; Arche and Domínguez 2011) in studies that indicate that mastery of clitic pronouns varies substantially according to proficiency level.[1] Beginning learners of L2 French clearly transfer English structures in producing in situ pronouns as in (2c) (Herschensohn 2000; Hawkins 2001), while intermediate proficiency learners show transitional responses, both behavioral and cortical, to ungrammatical in situ pronouns (Sneed German et al. 2015). Little research has been done on advanced learners of L2 French with respect to object clitic pronouns.[2]

Nor have object clitics been examined with respect to the prosodic transfer hypothesis. In terms of prosodic structure, English nouns and pronouns and French nouns constitute independent prosodic words, whereas French object pronouns constitute free clitics attached to the left of the verb, so that the prosodic structure of (2b) is as shown in (3c), where the clitic is attached as a sister to the prosodic word (PWd) and a daughter to the prosodic phrase (PPh) (Goad and Buckley 2006). English pronouns, normally independent prosodic words, may also cliticize, but as right-branching affixal clitics, so that the prosodic structures of (1a/b) and (1d) are as shown in (3a) and (3b) (Selkirk 1996).

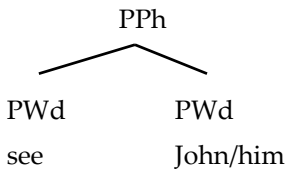
---

[1] Even early bilinguals may show variability. Perpiñán (2017) describes non target-like responses by Spanish-dominant bilinguals for Catalan clitics non-existent in Spanish, indicating a divergent Catalan grammar despite very early exposure (before age four) to Catalan.

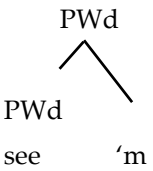
[2] Hoover and Dwivedi (1998) and Duffield et al. (2002) study the use of clitics (in reading and sentence matching tasks, respectively) in French students that they label as advanced.

3. English

a.

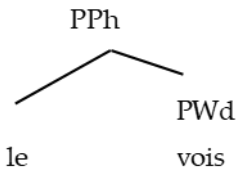


b.



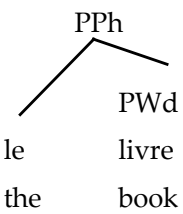
French

c.



The French clitic is both directionally (left, not right) and prosodically (free, not affixed) distinct from the English pronominal clitic. In their nominal morphosyntax, however, both languages share a similar prosodic free clitic structure, that of the leftward article plus noun as in (4).

4. French/English



The availability of this structure is interesting, given the PTH's proposal that target L2 prosody can be built when L1 structures are licensed in new positions ("minimal adaptation", Goad and White 2004).

In this article, we explore implications of the PTH for Anglophone acquisition of L2 French clitics by examining new data from advanced learners with extensive experience in the target environment. The first background section describes French and English pronouns (Section 2.1), prosodic structure and the prosodic transfer hypothesis (Section 2.2), and previous research on the L2 acquisition of French clitics (Section 2.3), which leads to our research questions for advanced learners (Section 2.4). The next sections present new evidence from advanced learners of L2 French (participants and methodology described in Sections 3.1 and 3.2, respectively), coming from three case studies that include data on production and grammaticality judgments (results presented in Section 3.3). We discuss the new evidence in light of the PTH (Section 4), arguing that the data support the hypothesis in demonstrating the learners' shift to L2 French prosody. Our analysis adopts the concept of minimal adaptation of native prosodic structures to accommodate L2 functional material. We close with some brief concluding remarks (Section 5).

2. Background

2.1. French and English Object Pronouns

Cross-linguistically, pronouns fall into strong, weak and clitic classes (Cardinaletti and Starke 1999). Strong pronouns project the most complex syntactic, semantic and phonological structure, followed by weak pronouns and finally by clitics.[3] French pronouns include parallel sets of strong and clitic pronouns in six persons, while English pronouns are strong unless cliticized to the verb as in (1d) (Table 1).[4]

Table 1. Pronouns of English and French.

	1st s.	2nd s.		3rd s.	1st p.	2nd p.	3rd p.
				French			
Strong	*moi*	*toi*	*vous*	*lui, elle*	*nous*	*vous*	*eux, elles*
Cl nom.	*je*	*tu*	*vous*	*il, elle, on*	*nous*	*vous*	*ils, elles*
Cl acc.	*me*	*te*	*vous*	*le, la*	*nous*	*vous*	*les*
Cl dat.	*me*	*te*	*vous*	*lui*	*nous*	*vous*	*leur*
				English			
Stg nom.	I	you	you	he, she, it	we	you	they
Stg obj.	me	you	you	him, her, it	us	you	them

Strong pronouns in French and English are phonetically free standing and fully referential, whereas clitics in French are morphosemantically minimal (person, gender and number features) and structurally deficient, requiring phonological attachment to a host. Strong pronouns act like full determiner phrases (DPs) in French in that they are syntactically mobile and phonologically stressed. They, but not clitic pronouns, can appear in coordination (5), in isolation (6) and in prepositional phrases (7). They are quite frequent in left and right dislocation (8) and are usually animate/ human.

5. *Elle* et le prochain, je les ai écoutés.
   **La* et le prochain, je les ai écoutés.
   'Her and the next one, I listened to them.'

6. Qui a-t-il aidé?
   'Who did he help?'
   Lui/Toi/Elles
   'him/you/them.'
   **Le/*Te/*Les*
   'him/you/them.'

7. *Toi* seul va au cínema avec *lui*.
   **Tu* seul va au cínema avec *le*.
   'You alone go to the movies with him.'

8. *Moi,* je trouve qu'il est stupide
   **Me,* je trouve qu'il est stupide
   'Me, I think he's stupid.'

---

[3] Romance clitic syntax has been discussed for decades within generative theory, beginning with Kayne (1975). See recent summaries by Roberts (2010) and Grüter (2008). Generally, there are two approaches, one group treating clitics as agreement markers on the verb (Grüter 2008; Culbertson 2010), and the other arguing that clitics must move from within the Verb Phrase (Kayne 1975; Roberts 2010; Cardinaletti and Starke 1999; De Cat 2005).

[4] Our focus here is on strong and clitic pronouns that are relevant for French and English. In Italian in *Telefono a loro* 'we phone them' the pronoun is strong, in *Telefono loro* it is weak, and in *Gli telefono* it is a clitic (Ordoñez and Repetti 2014).

In contrast, clitic pronouns are unstressed, may not remain in situ, and refer to both animate and inanimate referents. Roberts (2010, p. 41) analyzes French clitics as "simultaneoulsy maximal and minimal elements", bundles of features (number, gender, person, case) in the head D of the Determiner Phrase in which they occur, whose defective nature requires cliticization to an inflected verb or infinitive. English strong pronouns resemble French ones in syntactic distribution and stress, as the translations indicate. The clitic forms differ between English and French, however, in terms of placement, morphological variation, feature specification, as well as prosody, as outlined in the next section.

To be able to use pronouns, Anglophone learners need to master the implicit knowledge that French object pronouns must be placed in non-canonical position to the left of the verb; that strong and weak pronouns are in syntactic complementary distribution (nominative and objective obligatory clitic); that French pronouns are marked for person, number and gender (as English), but that gender is grammatical, not semantic (and thus applies to every inanimate noun as well as animate). Learners also need to prosodically integrate object pronouns in a way compatible with L2 target structures.

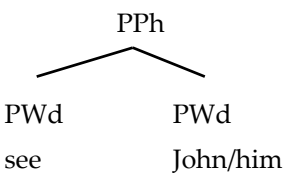
2.2. Prosodic Structure

Which prosodic representations accommodate object pronouns in English and French? English object pronouns are typically realized as full prosodic words (9a), but may cliticize as right-branching affixal clitics (9b) (Selkirk 1996).

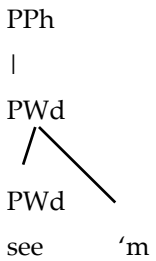
When an object pronoun in English is realized as a full prosodic word (9a), it attaches as a PWd sister to the verb, itself a PWd, with both of these linking to a phonological phrase node. A different structure is projected under phonological reduction of the pronoun. In this case (9b), failing to project its own PWd node, the pronoun is instead incorporated into an embedded prosodic word structure as a structurally deficient sister to the PWd verb.

9. Prosodic structure of English object pronouns

a.

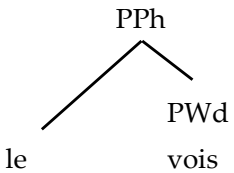


b.



In French, on the other hand, (as for all "morphologically-free functional items [in French] which precede their host" (Goad and Buckley 2006, p. 114)), object pronouns are prosodically left-branching free clitics (10).

10. Prosodic structure of French object pronouns



The free clitic—which can be separated from its host by other elements—differs from its affixal counterpart by its direct relation to the PPh: "the function word is sister to PWd and daughter to phonological phrase" (Selkirk 1996). The English affixal clitic cannot be separated from its host (11), whereas the French free clitic can, as we see in examples involving clitic stacking in Tense Phrases (12).

11. a. I saw'm

    b. *I saw'm'n'er [I saw him and her]

12. a. J'ai offert le cadeau
    I have offered the-M gift
    'I offered the

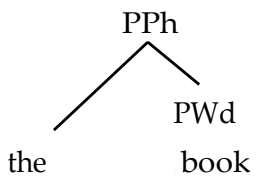
    b. Je ne l'ai pas offert
    I NEG have NEG offered
    'I didn't offer it.'

    c. Je ne le lui ai pas offert
    I NEG it to him have NEG offered
    'I didn't offer it to him.'

In (12a) the subject clitic *je* 'I' (which elides before the vowel-initial auxiliary) is attached to the inflected verb; in (12b, c) it is separated from the verb by additional pronominal and negative clitics that can intervene. French prosodic alternations include elision of vowels before following vowels and liaison of final consonants in function words when they precede vowel-initial content words/PWds. The nominal and verbal structures in which liaison and elision occur are the PPh, which may be iterated (Buckley 2005). Acquisition of French prosody is partially diagnosed by the learner's use of these prosodic alternations.

With respect to the free clitic structure in French, it is noteworthy that it is one which occurs in English, albeit in a different context—in article + noun sequences. In other words, English articles are, like French object pronouns, prosodically left-branching free clitics (13). Note that in both English and French, other elements may be inserted (e.g., adjectives) between the free clitic article and the head noun in the DP, just as clitics may be stacked before the verb in the TP. It is well known that articles are acquired before object clitics in L1A and L2A (Prévost 2009; Paradis et al. 2014).

13. Prosodic structure of English articles

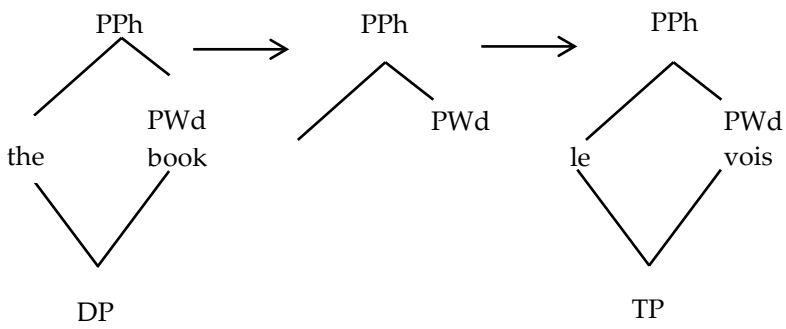


According to the prosodic transfer hypothesis (Goad et al. 2003; Goad and White 2006, 2008, 2009), learners have difficulty with L2 prosody distinct from L1. Goad and White (2004, 2006, 2008, 2009) and Goad et al. (2003) study L1 Mandarin and L1 Turkish learners of L2 English to test the PTH with respect to distinctive L2 prosodic structures in the past tense (Mandarin does not mark tense and does not allow affixal prosody) and articles (neither Mandarin nor Turkish have articles and do not allow free affixes-left). The series of articles suggests that learners accommodate the L2 English prosodic structures either by omission (of tense or article) or by adapting an L1 prosodic realization. For example, some of the L1 Turkish participants (Goad and White 2009) used an L1-licit prosodic structure of a stressed article in L2 English (even when L2 stressing was not appropriate). This tendency is confirmed by Snape and Kupisch (2010) in their phonetic analysis (using Praat software, www.praat.org) of an L1 Turkish L2 English learner's article production in which she overused stressed articles.

We have seen that English pronouns vary in prosodic realization between prosodic words and affixal clitics-right, whereas French consistently requires object pronouns to be free clitics-left. The two languages differ, therefore, not only in the directionality of object clitic placement (driven by morphosyntax) but also in their prosodic structure (affixal vs. free clitics). The PTH offers a means by

which a distinct new L2 prosodic structure might be built by licensing L1 structures in new positions ("minimal adaptation", Goad and White 2004). Goad and White (2009, p. 9) point out two conditions for minimal adaptation: "(a) when they [structures] can be built through combining pre-existing (L1) licensing relations; or (b) when they involve L1 structures being licensed in new positions." Given the availability to our learners of the left-branching free clitic structure in (13), we assume that minimal adaptation in this case would involve the extension of the L1 free clitic-left structure for nominal DP morphosyntax, to verbal TP configurations. This process is illustrated in (14).

14. Minimal adaptation in the L2 acquisition of object clitics



The possible prosodic transfer illustrated in (14) is similar to the one discussed in Goad and White (2009), which involved some of the L1 Turkish participants' inappropriate use of a stressed article prosodic structure in their L2 English, except that in the case at hand the prosodic transfer would be facilitative.[5]

2.3. L2A of Clitics

This section reviews previous research on the L2 acquisition of French clitic pronouns, and then looks at investigations of prosodic structure, mainly of English determiners. Over the past 40 years, there have been several studies of French pronouns, indicating that beginning learners of L2 French manifest four types of pronoun realizations: in situ placement of strong and clitic pronouns (15), use of null pronouns (16), past participle cliticization (17) and correct left clitic placement (Herschensohn 2000; Hawkins 2001; Granfeldt and Schlyter 2004; Herschensohn 2004; Prévost 2009).[6]

---

[5] We thank a reviewer for pointing out that the syntactic literature on clitics has used the distinction phonological versus syntactic clitics to refer to the English clitic-right type versus the French clitic-left type (cf. Kayne 1975; Auger 1994).

[6] Herschensohn (2004) reports on Chloe, one of the subjects of the current study, at an earlier stage of her L2 French acquisition, at which she makes numerous production and comprehension errors.

15. In situ

a.	J'ai	vu	elle	(= je	l'	ai vue)	(Herschensohn 2004)
	I	seen	her-STG	(I	her-CL	have seen)	
	'I saw her.'						
b.	Il	veut	les	encore	(= les	veut)	(Selinker et al. 1975)
	He	wants	them-CL	still	(= them-CL	wants	
	'He wants them still.'						

16. Null pronoun

a.	T'	as placé [e]	sur le lit	(= tu l'as placé)	(Herschensohn 2004)
	You-SG	have placed	on the-M	(= you-SG it have placed)	
	'You put it on the bed.'				
b.	J'ai	mangé [e]		(= je l'ai mangé)	(Adiv 1984)
	I have	eaten		(= I it have eaten)	
	'I ate it.'				

17. Past participle

a.	Vous	avez	la	pris	(= vous l'avez pris)	(Herschensohn 2004)
	You-PL	have	it-F	taken	(= you-PL it have taken)	
	'You have taken it.'					
b.	J'ai	le	vu		(= je l'ai vu)	(Granfeldt and Schlyter 2004)
	I have	it-M	seen		(I it have seen)	
	'I saw it.'					

The most prevalent errors in early interlanguage are the use of null pronouns (16)—similar to that of children acquiring French as an L1 (Hamann 2004; Hulk 2004; Paradis et al. 2014)—and in situ pronouns (15). Paradis et al. (2014, p. 170) note for L1A "regarding incorrect structures, over 90% were omitted objects". In situ French pronouns reflect native language transfer of English morphsyntactic structure, while null objects could represent a well-documented tendency to missing inflection and missing functional items (Prévost and White 2000; Lardiere 2007; Arche and Domínguez 2011). As learners advance in proficiency they show more native-like command of pronouns (Arche and Domínguez 2011; Cuza et al. 2013; Grüter and Crago 2012; Sneed German et al. 2015).

Acquisition of L2 Romance pronouns has been a subject of investigation to test morphosyntactic transfer hypotheses (Granfeldt 2005; Granfeldt and Schlyter 2004; Arche and Domínguez 2011). Impaired representation approaches (Snape et al. 2009) assume that adult learners will have persistent difficulties mastering grammatical (uninterpretable) features of an L2, and that the syntactic deficit will be represented by persistent morphosyntactic errors. In contrast, missing inflection approaches (Prévost and White 2000; Lardiere 2007) attribute morphological errors to other problems such as processing overload or prosodic difficulties. In any case, evidence from earlier studies indicates that mastery varies according to proficiency level: beginning learners of L2 French transfer from English in producing in situ pronouns, whereas intermediate learners show transitional responses, both behavioral and cortical, to ungrammatical in situ pronouns (Sneed German et al. 2015).

Sneed German et al. (2015) use both behavioral measures and event related potentials (ERPs) to investigate comprehension of object clitics by two groups of intermediate Anglophone learners of L2 French and native speakers. Event related potentials, which record electrophysiological cortical activity using scalp electrodes, have been used to determine reactions to language phenomena in real time, with temporal resolution to the millisecond (Friederici et al. 1993; Hagoort and Brown 1999; Osterhout et al. 2004, 2006; Steinhauer 2014). Lexico-semantic and morpho-syntactic anomalies reliably

elicit two distinct native responses, the N400 and the P600 respectively.[7] For example, sentence (18b), would elicit a negative trending electrophysiological wave 400 ms after the event of interest (cream vs. batteries), compared to (18a).

18. a. I drink my tea with cream

    b. I drink my tea with batteries

In contrast, (19b) would elicit a positive going wave 600 ms after the area of interest compared to (19a).

19. a. I drink my tea with cream

    b. I drinks my tea with cream

Event related potentials research with L2 learners has documented mixed results, with beginners usually insensitive to violations, but showing increasingly native-like cortical responses with greater proficiency (Foucart 2008; McLaughlin et al. 2010; Tanner et al. 2013).

Given extensive production data documenting Anglophone in situ L2 French pronouns, Sneed German et al. (2015) tested reactions to ungrammatical in situ strong and clitic pronouns (20a, b), as well as grammatical sentences (20c) that serve as a baseline.[8]

20.	a.	Après	le	match	elle	a rencontré	***le./ *les**	**dans**	la	rue
		After	the-M	match	she	has met	him-CL/them-CL	in	the-F	street
		'After the match she met him/them in the street.'								
	b.	Après	le	match	elle	a rencontré	***lui/ *eux**	**dans**	la	rue
		After	the-M	match	she	has met	him-STG/ them-STG	in	the-F	street
		'After the match she met him/them in the street.'								
	c.	Après	le	match	elle	l'a/les a	rencontré(s)	**dans**	la	rue
		After	the-M	match	she	him-CL/them-CL	has met	in	the	street
		'After the match she met him/ them in the street.'								

From a prosodic perspective, strong pronouns are licit in a stressed PWd position (20b), but they are morphosyntactically ungrammatical. In contrast, the in situ clitics (20a) are both morphosyntacitcally and prosodically illicit. Participants did two behavioral tasks (grammaticality judgments and general proficiency) and their cortical reactions were measured. The native French speakers showed predictable behavioral results (near ceiling accuracy) and a robust P600 to ungrammatical in situ pronouns, both strong and clitic. The high-intermediates' results were qualitatively similar in behavioral and ERP responses to natives and significantly better than low-intermediates. In contrast, the low-intermediates' results on the behavioral grammaticality judgments (GJ) task showed chance judgement of in situ clitics, (rejected 49% of the time), while ungrammatical strong pronouns were dismissed only 23% of the time, suggesting that low-intermediates are not sensitive to the ungrammaticality of these constructions.[9]

An important inference for further investigation concerns the response distinction between strong and clitic pronouns for the two groups of intermediates. Recall that the English pronoun system uses strong pronouns as the default and differs in morphosyntactic form and word order from the French

---

7 For discussion, see Gouvea et al. (2010); Steinhauer et al. (2010).

8 Given the incremental appearance of the clitic *le*, the reader's first reaction to (20a) would be to process *le* as an article of an anticipated noun phrase.

9 Buckley (2005), in a phonological analysis of a clitic production task given to intermediate French learners, found a small sub-group of her participants who stress the clitics; most of the participants did not assign prominence to the clitics.

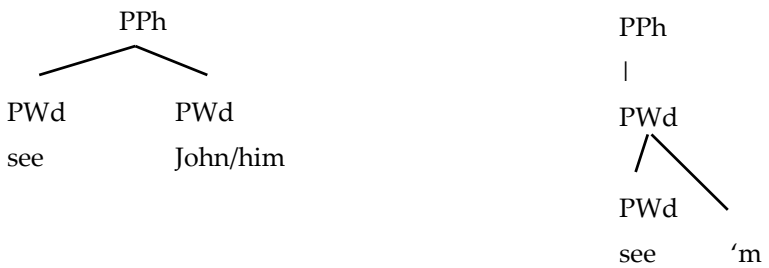
system, in which object pronouns may only be clitic. Both groups of intermediates showed higher accuracy and more French-like cortical responses to the clitic violations than the strong pronouns. The less proficient group persists in their acceptance of prosodically licit strong pronouns, while showing an emerging implicit understanding of object clitics; one might conclude that while still apparently using their L1 English settings for pronoun placement, their cortical response demonstrates emerging sensitivity to French prosody. In particular, learners may be demonstrating sensitivity to the apparent prosodic deficiency of the clitic pronouns in L2 input. This in turn may be part of a general emerging sensitivity to French schwa (the vowel of highly frequent object clitics *me*, *te*, and *le*). If this is the case, part of leaners' higher intolerance for the clitic pronoun in situ results from an emergent awareness that the schwa vowel is incompatible with a prominent position. This may facilitate the acquisition of morphosyntax, while not necessarily driving it.

The more proficient group, while showing some residual behavioral acceptance of strong pronouns, already were approaching the French natives in their responses to the ungrammatical in situ clitics. The research cited earlier on beginning and intermediate Anglophone learners of L2 French provides evidence that beginners and low intermediates transfer native pronoun word order, but with increasing proficiency, learners are able to gain mastery of L2 French morphosyntax, as indicated by comprehension, production and cortical measures. We infer that L2 learners can acquire features and word order different from the L1 system, but need to further investigate prosody.

2.4. Research Questions for Advanced Learners

Turning now to a proficiency level that has been less explored, that of advanced L2 French, we are led to the following research questions. Recall the prosodic structures involved in French and English clitic constructions ((9) is repeated here):

9. Prosodic structure of English object pronouns



English full pronouns are PWds under PPh, whereas their reduced clitic forms are affixal clitics right-attached to the PWd verb. In the early stages of L2A, the in situ pronouns that have been observed in production can be analyzed prosodically as PWd, as in (9a). There is no phonological evidence (phonological reduction or prosodic weakness) to suggest an affixal clitic (9b) analysis. In contrast, French object clitics are left free clitics under PPh, just as French and English articles are.

Beginning L2 French learners persist in Anglophone transfer with in situ placement; intermediate learners appear to gain prosodic sensitivity, distinguishing licit strong pronouns in situ from illicit clitics. We investigate advanced learners of L2 French pronouns to determine if they have adapted their object clitics to a French-like prosodic representation of left free clitics (left sister to PWd, daughter of PPh), or if they persist in using an English-like PWd structure (right sister to PWd). We pose the following research questions.

Research questions (RQs)

In order to determine that advanced learners have adapted to French prosody in their comprehension and production of pronominal clitics, we will need to prove that they are not using the

English-like representations in (9).[10] A lack of prosodic prominence on object clitics (RQ2) will rule out a PWd analysis, while clitic "stacking" (RQ3) will constitute evidence against an affixal clitic analysis. The production of clitics before verbs with prefixes will constitute further evidence against an affixal analysis (RQ4) (Buckley 2005, p. 127, citing Buckley (2005)). We pose the following research questions.

RQ1. Do advanced Anglophone learners of L2 French show any evidence of prosodic transfer in their comprehension (grammaticality judgments) or production (morphosyntax) of object pronouns?
RQ2. Do advanced Anglophone learners of L2 French produce object pronouns with any prosodic prominence?
RQ3. Do advanced Anglophone learners of L2 French produce object pronouns in "stacked" sequences?
RQ4. Do advanced Anglophone learners of L2 French produce object pronouns before verbs with prefixes?

3. New Evidence from Advanced Learners

In this section we present L2 French data on production and GJ drawn from oral interviews with Chloe (age of acquisition onset 13), Eleanor (age of onset 17) and Max (age of onset 48), all post-puberty L2 learners.[11] As noted earlier, French object clitic usage can be problematic for L2 learners, as indicated by several studies cited in the previous section. We examine the use of French object clitics and strong pronouns in three interviews for each individual conducted before, during, and after a period abroad of seven to nine months (a total of nine interviews over a nine-month period). The following section provides a description of the subjects and the data collection.[12]

3.1. Participants

Max (age 59–60 at interview) began studying French at age 48, first on his own (completing the "French in Action" video program of first-year French) and later (with his wife Eleanor) with the help of a native French tutor who met with them one hour weekly for conversational exchanges over a period of 11 years prior. Eleanor (age 53 at interview) studied for six years in high school and college and at age 28 spent two months with a family in France. She and Max vacation annually in France and do extensive reading, independent vocabulary/grammar study, audio listening, and television viewing in French for 16-18 h per week at home. At the time of the interviews, Max and Eleanor were spending four months in Paris and three months in Lyon, where they had daily contact with French in a variety of contexts. Chloe, (age 22–23 at interview), studied French for nine years in high school and college; spent six months in a family in France at age 16 (Herschensohn 2001, 2003) and four months in study abroad at age 20. Subsequently, she became an "assistante d'anglais" (English teaching assistant) in the French overseas department of Réunion for nine months. Table 2 summarizes these points.

Table 2. Characteristics of subjects.

Subject	Age of Onset	Age, Interview	Immersion
Max	48	59–60	13 months
Eleanor	17	53	15 months
Chloe	13	22–23	19 months

[10] An anonymous reviewer notes "in order to show that clitics are free clitics in interlanguage French, it must be shown that no other analysis is possible." We thank this reviewer for suggesting the following diagnostics.
[11] This research protocol was approved by the University of Washington IRB committee, in compliance with human subjects guidelines. All subjects gave their informed consent to participate in interviews and grammaticality tasks. Pseudonyms are used.
[12] Chloe's L2 French mastery at an earlier stage is reported in Herschensohn (2001, 2003, 2004); the three advanced learners are documented (for TP and DP phenomena) in Herschensohn and Arteaga (2007, 2009, 2016).

3.2. Materials and Methods

Although the project was ostensibly longitudinal (over seven to nine months), the interviews showed that the learners were sufficiently advanced to have reached a fairly steady state, and so the differences among the interviews were rather minimal. No notable change was observed over time in the error rate (but Chloe's production increases over the nine month period), and for this reason, we have collapsed the results from the three interviews together. Interviews were conducted by one of the authors (who has been certified by the American Council on the Teaching of Foreign Languages), who informally evaluated the three as being at an advanced level according to the American Council on the Teaching of Foreign Languages (ACTFL) Guidelines (cf. http://www.actfl.org/i4a/pages/index.cfm?pageid=4236). All three were able to use a range of tenses and moods, to discuss hypothetical situations and to carry out conversations on abstract and theoretical topics. They nevertheless made errors, particularly of gender assignment and aspect (Herschensohn and Arteaga 2009, 2016). Conducted in a university office or in the learner's residence in France, each interview included elicitations of the following areas: present tense, descriptions of everyday routines and environments; past and future tenses; hypothetical situations; role play including a problem (e.g., confronting a dry cleaner for damage or introducing a speaker at a lecture). The interviewer gave leading questions to carry forward the conversation, but it was the interviewees who provided most of the dialogue.

After the interview, subjects did written grammaticality tasks related to pronouns and other topics. The GJ task was adapted from Hawkins et al. (1993) and modified to have three versions for the three interviews. Comprising 50 sentences, half of which were ungrammatical, it had been used in the earlier studies reported (Herschensohn 2001, 2003, 2004; Herschensohn and Arteaga 2007, 2009, 2016). It included sentences that targeted both pronoun use (20% of the sentences) and verb placement (80% of the sentences); representative sentences for the pronouns are included in Appendix A. Participants were asked to mark the sentences as G (grammatical), NS (not sure) or * (ungrammatical) and to correct the latter.

The data from the interviews were transcribed and checked for accuracy by three linguists fluent in French. The pronoun production data were coded and compiled by a graduate research assistant. The evidence was in the present case evaluated for all obligatory contexts of non-nominative pronouns, and for all uses of non-nominative pronouns, which were coded as strong or clitic. The production data from the nine interviews included 17,000 words, of which 217 were object clitics and 91 strong pronouns.

3.3. Results from Production and Grammaticality Judgments

The evidence from the interview production and the grammaticality judgments indicates near ceiling performance on object pronoun production and the grammaticality judgment task. Table 3 documents the suppliance in obligatory contexts (the denominator in Table 3) of the 217 accurate object clitics, Table 4 portrays the correct use of the 91 strong pronouns (plus one error), while Table 5 gives the results of the GJ task. The three interviewees produced no prosodic errors of clitics in situ or strong pronouns as preverbal; more importantly, they produced consistent and accurate elision and liaison of function words, indicating ease with the L2 prosodic alternations (Tables 3 and 4).

Table 3. Object clitic suppliance in obligatory contexts.

	Interview I	Interview II	Interview III
Chloe	5/5	21/21	25/25
Max	24/24	24/24	44/44
Eleanor	20/20	28/29	26/27

The only errors in object clitic suppliance or word order were Eleanor's in the second and third interviews (21).

21. Eleanor (II, III)

a.	vous	nous	n'	aimez	pas	(= vous ne nous aimez pas)
	you-PL	us	NEG	like-2PL	NEG	(= you-PL NEG us like-2PL NEG)
	'You do not like us.'					
b.	dis	à	lui			(= dis-lui)
	tell-2SG	to	him-ST			(tell him-DAT-CL)
	'Tell him.'					

Chloe made lexical errors of case assignment (22), which were not counted as lack of supply or wrong morphological form.

22. Chloe (II, II, III)

a.	me	visitent	(= me rendent visite)
	me-CL	visit-3PL	(= me-ACC pay visit))
	'They visit me.'		
b.	les	répondre	(= leur répondre)
	them-ACC	to answer	(them-DAT to answer)
	'To answer them.'		
c.	leur	aider	(= les aider)
	them-DAT	to help	(them-ACC to help)
	'To help them.'		

As for the strong pronouns, most were used in PPs (obligatory suppliance indicated in Table 4 as the denominator); in addition, the three participants used freestanding strong pronouns (indicated by "+" in the table; these were supplemental, not required). These correct tokens numbered 91. Eleanor mistakenly used *tu* for *toi* in interview II (23), that is the additional "+*1" indicated in her Interview II column.

Table 4. Strong pronoun use.

	Interview I	**Interview II**	**Interview III**
Chloe	3/3 PP	3/3 PP + 4	4/4 PP + 1
Max	3/3 PP + 6	3/3 PP + 2	14/14 PP + 4
Eleanor	6/6 PP + 4	15/15 PP + 5 + *1	11/11 PP + 3

23. Strong pronoun errors (Eleanor II)

vous,	tu	et Michel	(= vous, toi et Michel)
you-PL	you-SG-CL	and Michel	(= you-PL you-SG-STG and Michel)
'Y'all, you and Michel.'			

The errors include mainly lexical errors (21b, the verb *dire* 'to say' requires an indirect object clitic, not a PP complement), and two case reversals on the object clitics (22b, c). Eleanor misaligns her negative clitic *ne* (21a) and misuses strong and clitic pronouns (21b, 24). The word order error in 21a is noteworthy in that the object clitic is placed (and "stacked") to the left of the negative clitic particle *ne*—a placement incompatible with an affixal clitic analysis of object pronouns on the part of this speaker.

As the first section has shown, beginning and intermediate Anglophone learners of L2 French transfer native pronoun word order, but the evidence here indicates that these advanced learners appropriately use pronominal morphosyntax and word order. The pronominal errors produced by the

interviewees do not resemble those of beginners; there are no in situ clitics or missing objects. Rather, the few mistakes that do appear are almost all lexical errors or wrong case.

As for the grammaticality judgments (Table 5), the advanced learners' all indicated rejection of in situ clitics and across the board acceptance of properly placed object clitics. There were no errors for Eleanor, and one not sure response in Interview I for Max for ungrammatical sentence (24).[13]

24. *Ils deux sont partis à midi (= tous les deux, ils sont partis)
    They two are gone at noon
    'They both left at noon.'

Table 5. Accuracy of pronouns in GJT.

	Interview I	Interview II	Interview III
Chloe	NA	7/10	9/10
Max	9/10	10/10	10/10
Eleanor	10/10	10/10	10/10

In contrast, Chloe did not complete the GJT for Interview I, and made wrong judgements for the sentences in (25) in the other sessions.

25. a. Eve-Anne s'est brossé les cheveux (NS, Int II)
    'EA brushed her hair.' (G)

    b. Est-ce que vous leur avez parlé ? (*, Int II)
    'Have you spoken to them?' (G)

    c. Ils deux sont partis à midi. (G, Int II, III)
    'They both left at noon.' (*)

The three participants correctly rejected in situ clitics and accepted correctly placed clitics (cf. Appendix A). In contrast, Chloe and her peer Emma, taking the same GJT at an earlier stage of development (Herschensohn 2004), averaged less than 70% accuracy on the task, with mistaken judgements about in situ clitics and those ungrammatically attached to past participles.

4. Discussion

Recall the research questions posed earlier.

RQ1. Do advanced Anglophone learners of L2 French show any evidence of prosodic transfer in their comprehension (grammaticality judgments) or production (morphosyntax) of object pronouns?
RQ2. Do advanced Anglophone learners of L2 French produce object pronouns with any prosodic prominence?
RQ3. Do advanced Anglophone learners of L2 French produce object pronouns in "stacked" sequences?
RQ4. Do advanced Anglophone learners of L2 French produce object pronouns before verbs with prefixes?

Do Max, Eleanor and Chloe indicate in their use of L2 French pronouns that they have adapted their object clitics to a French-like prosodic representation of left free clitics (left sister to PWd, daughter of PPh), or do they persist in using the English-like PWd structure (right sister to PWd)? As discussed

[13] Sentence (25) was not varied for the three versions; it is ungrammatical since the subject clitic is stressed, but can only be corrected through paraphrase.

earlier, if the clitics are never stressed, the PWd analysis is eliminated, and if clitics are "stacked", or used before prefixed verbs, the affixal analysis is invalid.

Our three learners show no evidence of prosodic prominence (stress) on object clitics that are preverbal; the only pronouns used in the stressed position (freestanding and in PPs) are strong pronouns. We therefore eliminated the PWd analysis (which may be appropriate for beginning learners who transfer in situ pronouns from English). As we saw earlier, the advanced learners' GJs all indicated rejection of in situ clitics and across the board acceptance of properly placed object clitics. Furthermore, Max, Eleanor and Chloe show evidence of stacked pronouns (see Appendix B) as in (26).

26. a. Je ne le connais pas (Chloe II)
    'I don't know him.'

    b. il s'en fichait (Max III)
    'he doesn't care a whit.'

    c. je ne l'ai jamais visité (Eleanor III)
    'I have never visited it.'

Their ability to stack clitics shows that they are not using a flipped affixal clitic analysis (which permits no stacking), but rather are treating these leftward clitics attached to inflected and infinitival verbs as free clitics. Finally, Max at least uses preverbal object clitics with verbs containing prefixes such as *ré-*, and *dé-* (Appendix B).

The accuracy of both production and comprehension (GJs) indicates mastery of the L2 French morphosyntax of clitic pronouns. As Arche and Domínguez (2011, p. 305) note, "if learners have acquired the morphosyntactic properties of Spanish clitics, then high rates in both the production and comprehension tasks will be observed." We answer the first question in the negative, as the learners have overcome the transfer of English word order and morphosyntax; their only errors are lexical. We consequently see no impaired representation of their L2 syntax with respect to the use of strong and clitic pronouns in French. Although beginning and low intermediate French L2 learners make persistent errors with in situ clitics in behavioral responses and non-native neural responses, advanced Anglophone learners of L2 French master French pronoun word order and morphosyntax. These advanced learners demonstrate target-like mastery of French pronoun word order and prosody in production, and sensitivity to ungrammatical uses in grammaticality judgments. Similarly to Gabriele et al. (2013, p. 225) observations, we believe "our results show that by an advanced level of proficiency, learners can process features not instantiated in the native language similarly to native speakers".

As for prosodic stucture, recall that English pronouns vary in prosodic realization between prosodic words and affixal clitics-right, while French consistently requires object pronouns to be free clitics-left. However, we noted the availability of leftward free clitics in both L1 and L2 DPs. It is our point that learners are incapable of gaining the free clitic construction for French clitic pronouns for quite a long time. Anglophone learners are able to gain free clitic articles at the beginning and intermediate stages, so the free clitic prosodic structure is presumably available to beginning–intermediate learners in L1 English and L2 French articles (which are morphologically identical to third person clitic pronouns). The difficulty that the Anglophone learners have with French object pronouns is clearly not a difficulty with the prosody at the beginning and early intermediate stages, since they master the article plus noun left-clitic structure early on (indeed, they can simply transfer the L1 to L2 DP prosody). Our point is that a combination of factors that differentiate English and French pronouns—the contrasts noted by Cardinaletti and Starke (1999) in terms of semantic, morphological, syntactic and phonological properties—seem to converge to inhibit command of French object pronouns by early Anglophone learners. We propose that it is not until high intermediate (cf. Sneed German et al. 2015) and advanced levels that learners finally gain the implicit mastery that enables them to transfer this article free clitic prosody from DP to TP. Despite its availability, the

beginning and intermediate learners were unable to adapt that structure to TP in L2 French. Eventually, though, at the advanced level, learners attain the ability to adapt the DP structure of prosodic free clitics to the TP domain, supporting the PTH (transfer of the L1 free clitic article) and the related concept of minimal adaptation (adapting the DP clitic to a TP domain) of native prosodic structures to construct the new L2 structures. The prosodic transfer hypothesis maintains that target prosody can be built when L1 structures can be licensed in new positions, and it is clear that the advanced learners have mastered the leftward clitics, licensing the pronominal objects in L2 preverbal instead of L1 postverbal positions. A lack of stressed preverbal clitics, together with the presence of stacked clitics and clitics before prefixed verbs, demonstrates that the correct surface word order corresponds to the correct prosodic structure as well; the clitics are prosodified as free clitics, and not as PWds or affixal clitics.

Together with previous research, our results show that target prosody achieved via minimal adaptation, that is the transfer of the free clitic construction (article + noun) to a new context (pronoun + verb), cannot take place in an immediate transfer, but rather, a certain threshold of proficiency must be reached before the prosodic structures are adapted. Morphosyntax and prosody must apparently develop in parallel.

5. Conclusions

Beginning and low intermediate Anglophone L2 French learners are insensitive to French prosodic constraints in using in situ pronouns transferred from the L1, as proposed by the prosodic transfer hypothesis.[14] The PTH proposes that native prosodic structures may be adapted to facilitate the acquisition of L2 prosodic structure. This study has presented new evidence from three Anglophone advanced learners of L2 French that indicates ceiling performance for pronoun production (99% accuracy in 300 tokens over nine interviews) and grammaticality judgments (98% accuracy). This native-like performance demonstrates that target French morphosyntax and prosody, are built—as predicted by the PTH—by licensing pronominal free clitics in a new pre-verbal L2 position distinct from post-verbal L1. Furthermore, accurate prosody is confirmed by correct pronoun placement, pronoun use in stacked sequences and with prefixed verbs, as well as target-like production of prosodic alternations (liaison, elision) in the learners' data. These results counter impaired representation approaches and suggest that early missing inflection may be overcome.

Author Contributions: The authors contributed equally to this work.

Acknowledgments: We thank Deborah Arteaga, Elisa Sneed German, Cheryl Frenck Mestre, Neal Snape and audiences at EuroSLA 2016 and GALA 2017 for comments and questions. We thank Kristen Piepgrass for transcribing and coding the data.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix Examples of Grammaticality Judgment Sentences with Pronouns/Clitics*

1. *La television? Nous la regardons tous les jours.* ('The television? We watch it every day.')
2. **Marie? Est-ce que vous avez vu la? (=vous l'avez vue?)* ('Marie? Have you seen her?')
3. *Eve-Anne s'est brossé les dents.* ('Eve-Anne brushed her teeth.')
4. **Marc a lave se avec du savon de Marseille (=Marc s'est lavé).* (Marc washed with soap from Marseille.)
5. *Est-ce que tu leur as parlé?* ('Did you speak to them?')
6. **Marie et Paul, je la et le connais.* (=je les connais) ('Marie and Paul, I know them.')

* indicates ungrammaticality; the correct sentence follows.

[14] We thank a reviewer for pointing out that in order to show that the difficulty is for prosodic reasons, further research should include the three levels of proficiency using the same experimental tasks.

Appendix Examples of Stacked Clitics and Clitics before Prefixed Verbs

Stacked clitics

1. *Je ne le connais pas.* (Chloe II) ('I don't know him.')
2. *Je vois aucune façon de s'en sortir*. (Chloe III) ('I see no way of getting out of here.')
3. *Tout le monde s'y intéresse.* (Max II) ('everyone is interested in that.')
4. *Il s'en fichait.* (Max III) ('he didn't care a whit about it.')
5. *Le soleil ne s'est levé jusqu'à 8h30.* (Eleanor II) ('the sun didn't come up until 8:30.')
6. *Je ne l'ai jamais visité.* (Eleanor III) ('je have never visited it')

Clitics before prefixed verbs

1. *Mais tout le monde est en train de se déplacer.* (Max I) ('. . . everyone is moving')
2. *Pour assurer qu'aucun étudiant ne se démarque pas par rapport aux autres.* (Max II) ('to make sure that no student stands out with respect to the others.')
3. *Son désir est de se réintégrer.* (Max III) ('her desire is to reintegrate herself.')

References

Adiv, Ellen. 1984. Language learning strategies: The relation between L1 operating principles and language transfer in L2 development. In *Second Languages: A Cross-Linguistic Perspective*. Edited by Roger Andersen. Rowley: Newbury House, pp. 125–42.

Arche, María J., and Laura Domínguez. 2011. Morphology and syntax dissociation in SLA: Evidence from L2 clitic acquisition in Spanish. In *Morphology and Its Interfaces*. Edited by Alexandra Galani, George Tsoulas and Glyn Hicks. Amsterdam and Philadelphia: John Benjamins, pp. 291–320.

Auger, Julie. 1994. Pronominal Clitics in Québec Colloquial French: A Morphological Analysis. Ph.D. dissertation, University of Pennsylvania, Philadelphia, PA, USA.

Buckley, Meaghen. 2005. Prosodic constraints and the syntax-phonology interface: The phonology of object clitics in L2 French. In *BUCLD 29 Proceedings*. Edited by Alejna Brugos, Manuella R. Clark-Cotton and Seungwan Ha. Somerville: Cascadilla Press, pp. 122–33.

Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In *Clitics in the Languages of Europe*. Edited by Henk van Riemsdijk. Berlin: Mouton de Gruyter, pp. 145–233.

Culbertson, Jennifer. 2010. Convergent evidence for categorical change in French: From subject clitic to agreement marker. *Language* 86: 85–132. [CrossRef]

Cuza, Alejandro, Ana Teresa Pérez-Leroux, and Liliana Sánchez. 2013. The role of semantic transfer in clitic drop among simultaneous and sequential Chinese-Spanish bilinguals. *Studies in Second Language Acquisition* 35: 93–125. [CrossRef]

De Cat, Cécile. 2005. French subject clitics are not agreement markers. *Lingua* 115: 1195–219. [CrossRef]

Duffield, Nigel, Lydia White, Joyce Bruhn de Garavito, Silvina Montrul, and Philippe Prévost. 2002. Clitic placement in L2 French: Evidence from sentence matching. *Journal of Linguistics* 38: 487–525. [CrossRef]

Foucart, Alice. 2008. Grammatical Gender Processing in French as a First and Second Language. Unpublished Doctoral dissertation, 2008, Université de Provence & University of Edinburgh, Provence, France, Edinburgh, Scotland.

Friederici, Angela D., Erdmut Pfeifer, and Anja Hahne. 1993. Event related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research* 1: 183–92. [CrossRef]

Gabriele, Alison, Robert Fiorentino, and José Alemán Bañón. 2013. Examining second language development using event-related potentials: A cross-sectional study on the processing of gender and number agreement. *Linguistic Approaches to Bilingualism* 3: 213–32. [CrossRef]

Goad, Heather, and Meaghen Buckley. 2006. Prosodic structure in child French: Evidence for the foot. *Catalan Journal of Linguistics* 5: 109–42. [CrossRef]

Goad, Heather, and Lydia White. 2004. Ultimate attainment of L2 inflection: effects of L1 prosodic structure. In *EuroSLA Yearbook 4*. Edited by Susan H. Foster-Cohen, Michael Sharwood Smith, Antonella Sorace and Mitsuhiko Ota. Amsterdam: J. Benjamins, pp. 173–200.

Goad, Heather, and Lydia White. 2006. Ultimate attainment in interlanguage grammars: A prosodic approach. *Second Language Research* 22: 243–68. [CrossRef]

Goad, Heather, and Lydia White. 2008. Prosodic structure and the representation of L2 functional morphology: A nativist approach. *Lingua* 118: 577–94. [CrossRef]

Goad, Heather, and Lydia White. 2009. Prosodic transfer and the representation of determiners in Turkish-English interlanguage. In *Representational Deficits in SLA: Studies in Honor of Roger Hawkins*. Edited by Neal Snape, Yan-kit Ingrid Leung and Michael Sharwood Smith. Amsterdam and Philadellphia: J. Benjamins, pp. 1–26.

Goad, Heather, Lydia White, and Jeffrey Steele. 2003. Missing inflection in L2 acquisition: defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics* 48: 243–63. [CrossRef]

Gouvea, Ana. C., Colin Phillips, Nina Kazanina, and David Poeppel. 2010. The linguistic processes underlying the P600. *Language and Cognitive Processes* 25: 149–88. [CrossRef]

Granfeldt, Jonas. 2005. The development of gender attribution and gender concord in French: A comparison of bilingual first and second language learners. In *Focus on French as a Foreign Language*. Edited by Jean-Marc Dewaele. Clevedon: Multilingual Matters.

Granfeldt, Jonas, and Suzanne Schlyter. 2004. Cliticisation in the acquisition of French as L1 and L2. In *The Acquisition of French in Different Contexts: Focus on Functional Categories*. Edited by Philippe Prévost and Johanne Paradis. Amsterdam and Philadelphia: J. Benjamins, pp. 334–70.

Grüter, Theres. 2008. When learners know more than linguists: (French) direct object clitics are not objects. *Probus* 20: 211–34. [CrossRef]

Grüter, Theres, and Martha Crago. 2012. Object clitics and their omission in child L2 French: The contribution of processing limitations and L1 transfer. *Bilingualism: Language and Cognition* 15: 531–49. [CrossRef]

Hagoort, Peter, and Colin M. Brown. 1999. Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research* 28: 715–28. [CrossRef] [PubMed]

Hamann, Cornelia. 2004. Comparing the development of the nominal and the verbal functional domain in French Language Impairment. In *The Acquisition of French in Different Contexts: Focus on Functional Categories*. Edited by Philippe Prévost and Johanne Paradis. Amsterdam and Philadelphia: J. Benjamins, pp. 109–46.

Hannahs, S.J. 1995. The phonological word in French. *Linguistics* 33: 1125–44. [CrossRef]

Hawkins, Roger. 2001. *Second Language Syntax*. Oxford: Blackwell.

Hawkins, Roger, and Cecilia Yuet-hung Chan. 1997. The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research* 13: 187–226. [CrossRef]

Hawkins, Roger, and Florencia Franceschina. 2004. Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In *The Acquisition of French in Different Contexts. Focus on Functional Categories*. Edited by Philippe Prévost and Johanne Paradis. Philadelphia and Amsterdam: John Benjamins, pp. 175–207.

Hawkins, Roger, Richard Towell, and Nives Bazergui. 1993. Universal Grammar and the acquisition of French verb movement by native speakers of English. *Second Language Research* 9: 189–233. [CrossRef]

Herschensohn, Julia. 2000. *The Second Time Around: Minimalism and L2 Acquisition*. Philadelphia and Amsterdam: John Benjamins.

Herschensohn, Julia. 2001. Missing inflection in L2 French: Accidental infinitives and other verbal deficits. *Second Language Research* 17: 273–305. [CrossRef]

Herschensohn, Julia. 2003. Verbs and rules: Two profiles of French morphology acquisition. *Journal of French Language Studies* 13: 23–45. [CrossRef]

Herschensohn, Julia. 2004. Functional categories and the acquisition of object clitics in L2 French. In *The Acquisition of French in Different Contexts: Focus on Functional Categories*. Edited by P. Prévost and J. Paradis. Philadelphia and Amsterdam: John Benjamins, pp. 207–42.

Herschensohn, Julia, and Deborah Arteaga. 2007. Marquage grammatical des syntagmes verbaux et nominaux d'un apprenant avancé. *AILE* 25: 159–78.

Herschensohn, Julia, and Deborah Arteaga. 2016. Parameters, features and re-assembly in L2 French DP. In *The Acquisition of French in Multilingual Contexts*. Edited by Pedro Guijarro-Fuentes, Katrin Schmitz and Natascha Müller. Bristol, Buffalo and Toronto: Multilingual Matters, pp. 215–33.

Herschensohn, Julia, and Deborah Arteaga. 2009. Tense and Verb Raising in advanced L2 French. *Journal of French Language Studies* 19: 291–318. [CrossRef]

Hoover, Michael L., and Veena D. Dwivedi. 1998. Syntactic processing by skilled bilinguals. *Language Learning* 48: 1–29. [CrossRef]

Hulk, Aafke. 2004. The acquisition of the French DP in a bilingual context. In *The Acquisition of French in Different Contexts: Focus on Functional Categories*. Edited by Philippe Prévost and Johanne Paradis. Amsterdam and Philadelphia: J. Benjamins, pp. 243–74.

Kayne, Richard. 1975. *French Syntax*. Cambridge: MIT Press.

Lado, Robert. 1957. *Linguistics across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor: University Michigan Press.

Lardiere, Donna. 2007. *Ultimate Attainment in Second Language Acquisition: A Case Study*. Mahwah: L. Erlbaum.

McLaughlin, Judith, Darren Tanner, Ilona Pitkänen, Cheryl Frenck-Mestre, Kayo Inoue, Geoffrey Valentine, and Lee Osterhout. 2010. Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning* 60 S2: 123–50. [CrossRef]

Ordoñez, Francisco, and Lori Repetti. 2014. On the morphological restrictions of hosting clitics in Italian and Sardinian dialects. *L'Italia Dialettale: Rivista di Dialettologia Italiana* 75: 173–200.

Osterhout, Lee, Judith McLaughlin, Albert Kim, Ralf Greenwald, and Kayo Inoue. 2004. Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In *The On-Line Study of Sentence Comprehension: Eyetracking, ERP, and Beyond*. Edited by Manuel Carreiras and Charles E. Clifton Jr. New York: Psychology Press.

Osterhout, Lee, Judith McLaughlin, Ilona Pitkänen, Cheryl Frenck-Mestre, and Nicola Molinaro. 2006. Novice learners, longitudinal designs, and event-related potentials: A paradigm for exploring the neurocognition of second-language processing. *Language Learning*. 56: 199–230. [CrossRef]

Paradis, Johanne, Antoine Tremblay, and Martha Crago. 2014. French-English bilingual children's sensitivity to child-level and language-level input factors in morphosyntactic acquisition. In *Input and Experience in Bilingual Development*. Edited by Theres Grüter and Johanne Paradis. Amsterdam and Philadelphia: J. Benjamins, pp. 161–80.

Perpiñán, Silvia. 2017. Catalan-Spanish bilingualism continuum: The expression of non-personal Catalan clitics in the adult grammar of early bilinguals. *Linguistic Approaches to Bilingualism* 7: 477–513. [CrossRef]

Prévost, Philippe. 2009. *The Acquisition of French: The Development of Inflectional Morphology and Syntax in L1 Acquisition, Bilingualism, and L2 Acquisition*. Amsterdam and Philadelphia: J. Benjamins.

Prévost, Philippe, and Lydia White. 2000. Accounting for morphological variation in second language acquisition: Truncation or missing inflection? In *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics*. Edited by Marc-Ariel Friedemann and Luigi Rizzi. Harlow: Longman, pp. 202–35.

Roberts, Ian. 2010. *Agreement and Head Movement: Clitics, Incorporation and Defective Goals*. Cambridge: MIT Press.

Rothman, Jason. 2011. L3 syntactic transfer selectivity and typological determinacy: The Typological Primacy Model. *Second Language Research* 27: 107–27. [CrossRef]

Schwartz, Bonnie D., and Rex A. Sprouse. 1996. L2 cognitive states and the full transfer/full access model. *Second Language Research* 12: 40–72. [CrossRef]

Selinker, Larry, Merrill Swain, and Guy Dumas. 1975. The interlanguage hypothesis extended to children. *Language Learning* 25: 139–52. [CrossRef]

Selkirk, Elisabeth. 1996. The prosodic structure of function words. In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Edited by James L. Morgan and Katherine Demuth. Mahwah: Lawrence Erlbaum, pp. 187–213.

Snape, Neal, Yan-kit Ingrid Leung, and Michael Sharwood Smith, eds. 2009. *Representational Deficits in SLA: Studies in Honor of Roger Hawkins*. Amsterdam and Philadelphia: J. Benjamins.

Snape, Neal, and Tanja Kupisch. 2010. Ultimate attainment of second language articles: A case study of an endstate second language Turkish-English speaker. *Second Language Research* 26: 527–48. [CrossRef]

Sneed German, Elisa, Julia Herschensohn, and Cheryl Frenck-Mestre. 2015. Pronoun processing in Anglophone late L2 learners of French: Behavioral and ERP evidence. *Journal of Neurolinguistics* 34: 15–40. [CrossRef]

Steinhauer, Karsten. 2014. Event related potentials (ERPs) in second language research: A brief introduction to the technique, a selected review and an invitation to reconsider critical periods in L2. *Applied Linguistics* 35: 393–417. [CrossRef]

Steinhauer, Karsten, John E. Drury, Paul Portner, Matthew Walenski, and Michael T. Ullman. 2010. Syntax, concepts and logic in the temporal dynamics of language comprehension: Evidence from event-related potentials. *Neuropsychologia* 48: 1525–42. [CrossRef] [PubMed]

Tanner, Darren, Judith McLaughlin, Julia Herschensohn, and Lee Osterhout. 2013. Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition* 16: 367–82. [CrossRef]

Tsimpli, Ianthi Maria, and Maria Dimitrakopoulou. 2007. The Interpretability Hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research* 23: 215–42. [CrossRef]

White, Lydia. 1996. Clitics in L2 French. In *Generative Perspectives on Language Acquisition*. Edited by Harald Clahsen. Amsterdam and Philadelphia: J. Benjamins, pp. 335–68.

languages

MDPI

Article

The Mixed Effects of Phonetic Input Variability on Relative Ease of L2 Learning: Evidence from English Learners' Production of French and Spanish Stop-Rhotic Clusters

Laura Colantoni [1] and Jeffrey Steele [2,*]

[1] Department of Spanish and Portuguese, University of Toronto, 91 Charles St W, Toronto, ON M5S 2C7, Canada; laura.colantoni@utoronto.ca

[2] Department of French Studies, University of Toronto, 50 St Joseph Street, Toronto, ON M5S 1J4, Canada

* Correspondence: jeffrey.steele@utoronto.ca; Tel.: +1-416-926-1300 (ext. 3352)

Received: 9 January 2018; Accepted: 13 April 2018; Published: 24 April 2018

Abstract: We examined the consequences of within-category phonetic variability in the input on non-native learners' production accuracy. Following previous empirical research on the L2 acquisition of phonetics and the lexicon, we tested the hypothesis that phonetic variability facilitates learning by analyzing English-speaking learners' production of French and Spanish word-medial stop-rhotic clusters, which differ from their English counterparts in terms of stop and rhotic voicing and manner. Crucially, for both the stops and rhotics, there are differences in within-language variability. Twenty native speakers per language and 39 L1 English-learners of French (N = 20) and Spanish (N = 19) of intermediate and advanced proficiency performed a carrier-sentence reading task. A given parameter was deemed to have been acquired when the learners' production fell within the range of attested native speaker values. An acoustic analysis of the data partially supports the facilitative effect of phonetic variability. To account for the unsupported hypotheses, we discuss a number of issues, including the difficulty of measuring variability, the need to determine the extent to which learners' perception shapes intake, and the challenge of teasing apart the effects of input variability from those of transferred L1 articulatory patterns.

Keywords: input; variability; relative difficulty; second language production; consonant clusters; stops; rhotics; phonetic parameters; voicing; manner

1. Introduction

Research on the role of input in L2 speech learning has first and foremost examined the effects of overall input quantity and quality (e.g., Flege 2009; Freed et al. 2004; Moyer 2009, 2011; Saito 2015, for phonetics and phonology). A smaller body of work has studied the effects of input variability, namely, the extent to which the latter may facilitate speech learning. As we will see in greater detail in Section 2, findings have overwhelmingly found a positive role for this variable. Laboratory studies investigating native talker variability—specifically, learners' exposure to one versus many speakers, including speakers of different varieties—have shown that such variability not only improves the discrimination of some native contrasts (the English /l/-ɹ/ contrast in particular) but also the retention of categories in long-term memory (e.g., Logan et al. 1991; Pisoni and Lively 1995).

We seek to contribute further to the study of the role of input variability in L2 speech learning by expanding on previous research in several ways via the analysis of the acquisition of French and Spanish word-medial stop-rhotic (SR) clusters (e.g., French *sucré* /sykʁe/ 'sweet', *degré* /dəgʁe/ 'degree';

Spanish: *sobra* /sobra/ 'excess', *sidra* /sidra/ 'cider') by native speakers of English. First, whereas the majority of previous research has looked at variability's effects on the perception of phonemic contrasts, here, we seek to determine whether such effects are also observed in the production of subphonemic properties. Second, in contrast to some researchers (e.g., Logan et al. 1991; Pisoni and Lively 1995), we focus on within-category variability within a given language as well as between speakers as opposed to between-speaker variation alone. Third, we analyze the interaction of variability and proficiency by comparing intermediate versus advanced speakers. The present research thus differs from much previous research, which has looked at L2 learners at the initial stages of L2 speech learning. Finally, we make two new empirical contributions by comparing the relative ease of acquisition of consonant sequences (in contrast to individual liquids or vowels) and by investigating the acquisition of two Romance languages (as opposed to English).

The remainder of the paper is structured as follows. In the following section, we review previous empirical research on the role of acoustic variability in L2 phonetic and lexical learning. We will see that, overwhelmingly, such studies reveal the positive effect of this variable. We then present a detailed overview of the phonetic properties of French and Spanish SR clusters mentioned above, focusing particularly on patterns of variability in stop and rhotic voicing and manner. This is followed by the presentation of a set of specific hypotheses for the acquisition of these clusters by native English speakers for each of the phonetic parameters in question. These hypotheses are then investigated using data from an experimental study that tested L2 learners of intermediate and advanced proficiency on their production of these clusters via a carrier-sentence reading task. We conclude with a discussion of the importance of variability as well as other factors, including first language influence and articulatory complexity, as predictors of relative difficulty in L2 speech learning.

2. Empirical Evidence for the Role of Input Variability in L2 Acquisition

Laboratory studies have generally demonstrated that exposure to multiple talkers or larger (i.e., more variable) stimuli sets seems not only to facilitate category formation in perception (e.g., Brousseau-Lapré et al. 2013; Nishi and Kewley-Port 2007, for vowels; Logan et al. 1991; Pisoni and Lively 1995; Pruitt et al. 2006; Sadakata and McQueen 2013; Zhang et al. 2009, for consonants; Hardison 2003, for prosody) but also the retention of these categories in memory (Lively et al. 1994; Pisoni and Lively 1995) and the ability to extend training to novel contrasts/speakers (e.g., Clopper and Pisoni 2004; Nishi and Kewley-Port 2007; Pruitt et al. 2006; Sadakata and McQueen 2013). Indeed, (Brousseau-Lapré et al. 2013, p. 420) claim that "there is now a consensus that highly variable natural speech input provides the best foundation for learning in second language speech perception interventions". Some caution in accepting such consensus may be warranted given that other research has either shown no advantage or even disadvantage for more heterogeneous speech input. Iverson et al.'s (2005) training study showed no greater improvement using input with signal manipulation than with high-variability phonetic training; this suggests that it may be training in general—as opposed to stimuli speaker variability in particular—that shapes learning. Giannakopoulou et al. (2017), in a study of native Greek speakers' perception of the English /i-ɪ/ contrast, found greater advantage for training with a single, as opposed to multiple, talkers and that this benefit increased over the course of the 10 training sessions. Finally, Bohn and Bundgaard-Nielsen (2009, p. 218) demonstrated that the least intelligible vowels produced by their Danish-speaking L2 learners were the same vowels that vary the most across English (American, Southern British, and Australian) dialects. These authors propose explicitly that "an additional source of learning problems for *non-native* speakers is inherent to a learning target that is highly variable".

Mixed findings regarding the effects of input variability on L2 learning can also be observed in a series of studies that tested the effects of acoustic variability on speakers' vocabulary learning (Barcroft and Sommers 2005, 2014; Sommers and Barcroft 2007). Barcroft and Sommers (2005) found improved performance in terms of both accuracy and reaction time for learners trained with either greater between-talker variability (1 versus 3 (moderate) versus 6 speakers (high variability)) and

within-talker variability related to voice type (neutral, excited, whispered, and nasal as well as digitally edited high-pitched and elongated variants). This study also demonstrated that the beneficial effects of greater input variability may be blocked under certain conditions. Specifically, the positive effects of voice-type variability on L2 vocabulary learning were only observed, once potential between-speaker differences in intelligibility were controlled for. Sommers and Barcroft (2007) present a similar study, differing principally in the acoustic parameters targeted. A positive effect for rate of speech—but not for overall amplitude and fundamental frequency—was found. Following (Sommers et al. 1994, p. 232), the authors argue that the fact that the positive effect on vocabulary learning was limited to variability in speech rate is consistent with the phonetic-relevance hypothesis that "acoustic variability will impair spoken word identification only if the source of variability in question alters phonetically relevant properties of the speech signal". This hypothesis echoes the general explanation for the effectiveness of high-variability training on phonetic contrasts, namely, that the "experience of variation allows the formation of generalized representations that include only phonetically relevant cues and exclude irrelevant talker identity cues" (Giannakopoulou et al. 2017, p. 6). In other words, greater variability directs learners' attention to those aspects of the input that are relevant and most consistent across exemplars. Barcroft and Sommers (2014) tested the phonetic-relevance hypothesis directly, examining the effects of fundamental frequency (f0) variability. Consistent with the hypothesis, speakers of a tonal language (Zapotec) benefitted from such training, whereas speakers of a nontonal language (Spanish) did not. In summary, taken together, these three studies demonstrate the potential effects of input phonetic variability on L2 lexical learning, including that such effects may only manifest themselves when the phonetic parameter manipulated is relevant to an L1 phonemic contrast.

Thus, experimental studies on the L2 acquisition of phonetic and lexical competence demonstrate the mainly positive effects for input variability. However, such effects may be mitigated by the type of phonetic variability. As demonstrated in Barcroft and Sommers (2005, 2014) and Sommers and Barcroft (2007), in keeping with the phonetic-relevance hypothesis, input variability's effects may be restricted to L1 contrastive features.

We now turn to the French and Spanish SR clusters targeted in the present experiment, with the goal of highlighting differences in the degree of variability in stop and rhotic voicing and manner in order to be able to propose a set of variability-based hypotheses.

3. The Phonetics of French and Spanish Stop-Rhotic Clusters

Stop-rhotic clusters differ in English, French, and Spanish in at least three respects.[1] First, English speakers must master Romance stop voicing; this involves eliminating aspiration in voiceless stops and realizing phonemically voiced stops as fully voiced.[2] Beyond these general differences, voiceless and voiced stops are realized variably in both languages. In French, both voiced and particularly voiceless stops may be variably voiced throughout their realization. In Spanish, voiced stops vary in manner: in intervocalic position, they are realized most often as approximants, but they are realized as stops after nasals and in utterance-initial position.[3] Second, the English alveolar approximant [ɹ] must be replaced by the velar/uvular fricative [ɣ]/[ʁ] in French versus the alveolar tap [ɾ] in Spanish. Once again, these are not the only possible realizations in each language. In French, [ɣ]/[ʁ] may be variably realized as an approximant (e.g., Colantoni and Steele 2007a; O'Shaughnessy 1982) and may be devoiced following a voiceless stop (e.g., Léon 1992; Tranel 1987; Walker 1984, 2001). In Spanish, the rhotic may be realized as a tap or an approximant (e.g., Blecua 2001, 2008).

---

[1] Along with the differences to be discussed immediately, these languages also contrast in the phonetic realization of the sequences. In French and Spanish, but not in English, the clusters may be broken up by an epenthetic vowel (Colantoni and Steele 2005, 2007a). This vowel is quasi-categorically present in voiceless and voiced stop-rhotic clusters in Spanish, whereas in French it only appears in voiced stop-rhotic clusters.

[2] Learners should also acquire different durational parameters. Although in previous studies (e.g., Colantoni and Steele 2007b, 2008) we have discussed the role of duration, we will not deal with this parameter here, given the complexity of the comparison.

[3] The voiced dental is also realized as a stop following a lateral.

In the present study, we focused on four phonetic parameters of word-medial SR clusters, namely, stop voicing and manner, and rhotic voicing and manner. The choice to restrict the focus to SR clusters in this position was doubly motivated. First, testing word-medial as opposed to word-initial clusters ensures comparable phonetic context across speakers. When reading carrier sentences in which only the target word changes, as was done in the present study, speakers—particularly those of lower proficiency—may sometimes pause before the target word for emphasis. In contrast, word-medial clusters are realized intervocalically without exception. As pauses affect the realization of voicing, one of the parameters measured here, restricting the focus to word-medial contexts was necessary. Second, as will become obvious when discussing the results, investigating four phonetic parameters in two languages creates a large, complex data set. Focusing on two prosodic contexts (word-initial and word-medial) would only increase this complexity.

Many researchers have investigated the phonetics of stops and rhotics in singletons and clusters, but there are no previous studies comparing the relative degree of variability of stops and rhotics in such clusters along these four phonetic parameters. Thus, in order to develop the specific hypotheses for the present study, we re-analyzed data from the studies reported in Colantoni and Steele (2005, 2007a, 2011). In these studies, 40 native speakers (10 each of Quebec and European French; Argentine and Chilean Spanish) were tested on their production of obstruent-liquid clusters via the same sentence-reading task used with the L2 learners in the study reported here. Stimuli were controlled for obstruent place and manner (stops and fricatives) and liquid type (laterals and rhotics).[4] For our present needs, we focused on the subset of word-medial SR clusters (French n = 12; Spanish n = 10) in Table 1.

Table 1. French and Spanish word-medial stop-rhotic (SR) clusters.

Language	Stop Voicing					
	Voiceless			Voiced		
French (n = 12)	*mépris*	/meˈpʁi/	'disdain'	*cobra*	/kɔˈbʁɑ/	'cobra'
	vitré	/viˈtʁe/	'glass, ADJ'	*poudrer*	/puˈdʁe/	'powder, INF'
	sucré	/suˈkʁe/	'sweet'	*degré*	/dəˈgʁe/	'degree'
	soprano	/sɔpʁaˈno/	'soprano'	*librairie*	/libʁeˈʁi/	'bookstore'
	détraquer	/detʁaˈke/	'upset, INF'	*redresser*	/ʁədʁeˈse/	'straighten up, INF'
	décrasser	/dekʁaˈse/	'clean up, INF'	*dégrader*	/degʁaˈde/	'degrade, INF'
Spanish (n = 10)	*lepra*	/ˈlepɾa/	'leprosy'	*cobrá*	/koˈbɾa/	'charge, IMP, 2PS'
	letra	/ˈletɾa/	'letter'	*sobra*	/ˈsobɾa/	'left over, PRES, 3PS'
	sacra	/ˈsakɾa/	'sacred'	*podré*	/poˈdɾe/	'be able, FUT, 1PS'
	lucrá	/luˈkɾa/	'speculate, IMP, 2PS'	*sidra*	/ˈsidɾa/	'cider'
				negra	/ˈnegɾa/	'black, FEM'
				lograá	/loˈgɾa/	'achieve, IMP, 2PS'

Note: INF: infinitive; 2PS: 2nd person singular; IMP: imperative; 3PS: 3rd person singular; FUT: future; 1PS: 1st person singular; FEM: feminine.

All of the target words were read in a carrier sentence (French: *Je dis____encore une fois*; Spanish: *Digo____otra vez*; 'I say____ again') three times, each time in random order. The values reported in the following sections are based on the 1236 elicited tokens (French voiceless: 333, voiced: 374; Spanish voiceless: 237, voiced: 292) involving word-medial SR clusters.

For each member of the cluster, we measured the % voicing (proportion of visible f0 over the duration of the segment) using the same procedure as in Snoeren et al. (2006). Pulses were also displayed to verify the analysis. Manner was transcribed and transcriptions were verified based on the overall acoustic characteristics of the sounds, as determined by the visual inspection of the waveform

[4] These variables were controlled for in order to test a series of hypotheses that concern liquid duration and the effects of voicing, place, and stress on the duration of both the liquid and the epenthetic vowel breaking up the clusters. No effect of stress on either manner realization or percentage voicing is expected.

and spectrogram. Results were measured for each of the 40 native speakers. In what follows, we report means as well as measures of dispersion (standard deviation, range).

3.1. Stop Voicing

As a measure of phonemic stop voicing, we will focus on one particular phonetic parameter, namely, laryngeal voicing as measured by the percentage of the stop's articulation during which the fundamental frequency (f0) is present in the spectrogram. We have chosen this parameter, as opposed to the widely used voice onset time (VOT), because it facilitates both within- and between-category comparisons in two respects. First, underlying voiceless stops may be fully voiced in Romance languages, and % voicing, as opposed to VOT, better captures this reality (Möbius 2004; Snoeren et al. 2006). Second, voiced stops may be variably realized as approximants: using VOT would not allow us to compare voicing across these different manner realizations.

Figure 1 presents the density plots of the distribution of % voicing for the stop of word-medial SR clusters in both varieties of French and Spanish, controlling for the phonemic voicing of the stop (voiceless versus voiced); Table 2 displays the summary statistics for this parameter.[5] Both the density plots and the summary statistics reveal a larger dispersion in the observed values for voiceless stops. Indeed, as indicated by the results of a two-sample test of variance, there is significantly greater variability in French voiceless (SD = 32) than voiced clusters (SD = 21; F = 2.25, $p < 0.0001$).[6] A similar pattern is found with the Spanish native speakers (voiceless: SD = 29; voiced: SD = 19; F = 2.43, $p < 0.0001$). In both languages, variability is greater in voiceless clusters (French 69%; Spanish 50%) than voiced ones (French 26%; Spanish 27%). These two measures taken together show that, when determining the prototypical percentage of laryngeal voicing necessary to realize the stops of SR clusters, there is greater input variability in terms of how individual speakers produce this parameter on average as well as how such mean values vary across a range of native speakers.

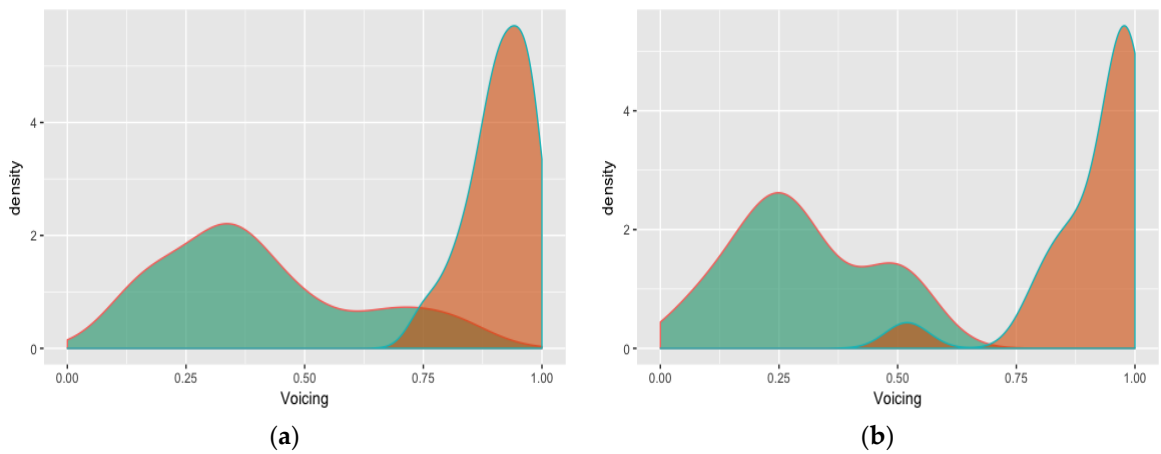


Figure 1. Density plots of stop voicing (%) in French (**a**) and Spanish (**b**) voiceless (green, left) and voiced (orange, right) word-medial stop-rhotic (SR) clusters.

---

5 See Section 4.3 "Data Preparation and Analysis" for discussion of the segmentation and measurement of the phonetic parameters of the SR clusters.

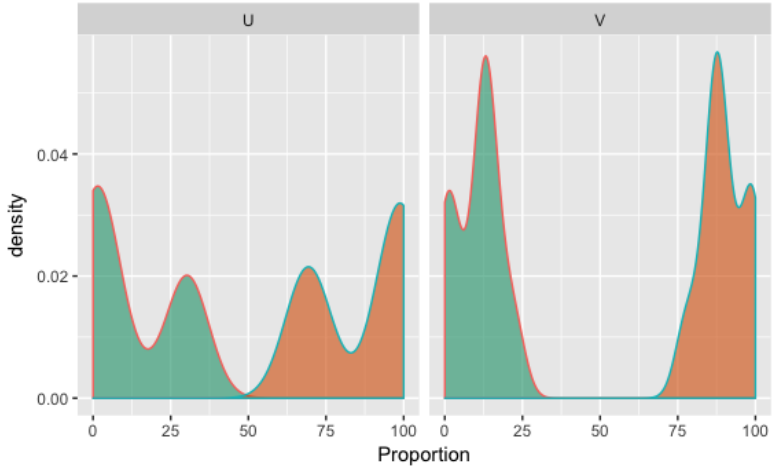
6 In order to compare the relative degree of variability for a given parameter, we used paired-sample *t*-tests of variance for quantitative variables and paired-sample *z*-tests of proportions for categorical variables following Wade et al. (2007).

Table 2. Stop % voicing: summary statistics.

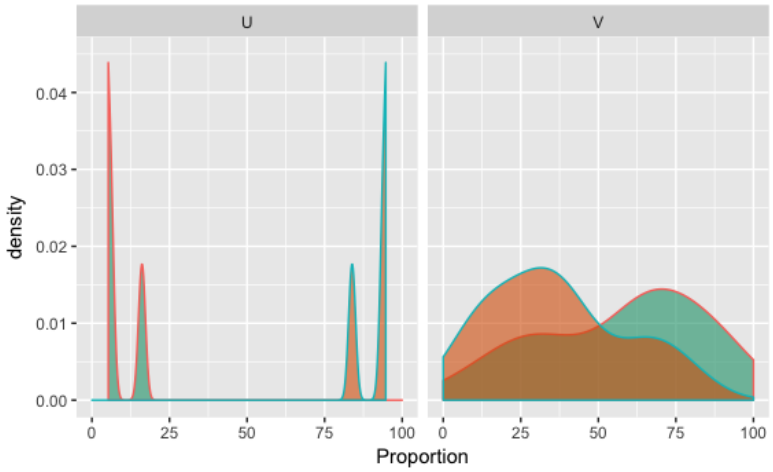
	Voiceless			Voiced		
Language	**M**	**SD**	**Min–Max (Range)**	**M**	**SD**	**Min–Max (Range)**
French	51	32	23–92 (69)	91	21	74–100 (26)
Spanish	36	29	11–61 (50)	95	19	73–100 (27)

3.2. Stop Manner

The density plots in Figure 2 below present the variation in manner for phonemically voiceless and voiced stops in native French and Spanish; Table 3 provides the summary statistics.



(**a**)



(**b**)

Figure 2. Density plots of stop manner (%) in French (**a**) and Spanish (**b**) voiceless (U) and voiced (V) word-medial SR clusters. Approximant (green); stop (orange).

Table 3. Stop manner: summary statistics.

Language	Voiceless				Voiced			
	M		SD	Min–Max (Range)	M		SD	Min–Max (Range)
	Stop	App			Stop	App		
French	81	19	23	50–100 (50)	86	14	12	62–100 (38)
Spanish	98	2	4	90–100 (10)	17	83	18	0–67 (67)

In French, voiceless stops are characterized by a higher degree of within-category variation (SD = 23) than voiced stops (SD = 12).[7] The relative degree of variability is also evidenced by the range of values observed; voiceless stops display more interspeaker variability (range = 50) than voiced stops (range = 38). When compared to French, Spanish seems to maintain a clear between-category distinction: voiceless stops display almost no variation (SD = 4; range = 10), whereas Spanish voiced stops display both a higher standard deviation and, particularly so, range (SD = 18; range = 67). Results of a two-sample test for proportions indicate that the proportion of Spanish stop versus approximant realizations is significantly higher than the proportion obtained for Spanish voiceless stops (z = 18.62, $p < 0.0001$).

Based on these measures of manner, within French, voiceless stops should present the least difficulty for learners. In Spanish, in contrast, it is the voiced stops that should be easier for learners to master—manner in Spanish voiceless stops should present relatively greater challenge due to the low interspeaker variability. While this latter hypothesis is strictly in keeping with the general hypothesis that L2 speech learning is facilitated by greater input variability, the degree of variation in manner in Spanish voiceless stops is so small that one could classify this parameter as invariant and thus propose that a prediction based on input variability alone is unwarranted here. We nonetheless tested the hypothesis that manner should be acquired more readily in voiced stops in order to push our general hypothesis to its limits. In the Discussion (Section 6.2), we will return to the question of what degree of variability is relevant for determining the relative difficulty of L2 speech learning.

3.3. Rhotic Voicing

Both the French voiced dorsal fricative /ʁ/ and the Spanish voiced alveolar tap /ɾ/ differ from the English voiced alveolar approximant /ɹ/. Accordingly, unlike stops where it is a matter of adjusting English phonetic parameters to target French and Spanish ones, in the case of rhotics, learners must acquire completely new articulatory patterns for French—while English has voiced fricatives (e.g., *themselves* [ð̠əmsɛlv̠z]), none are dorsal—and relatively new patterns for Spanish; the Spanish tap resembles the North American English flap allophone of intervocalic unstressed /t/ (e.g., *bottom* /bɑtəm/ → [ˈbɑɾəm]).

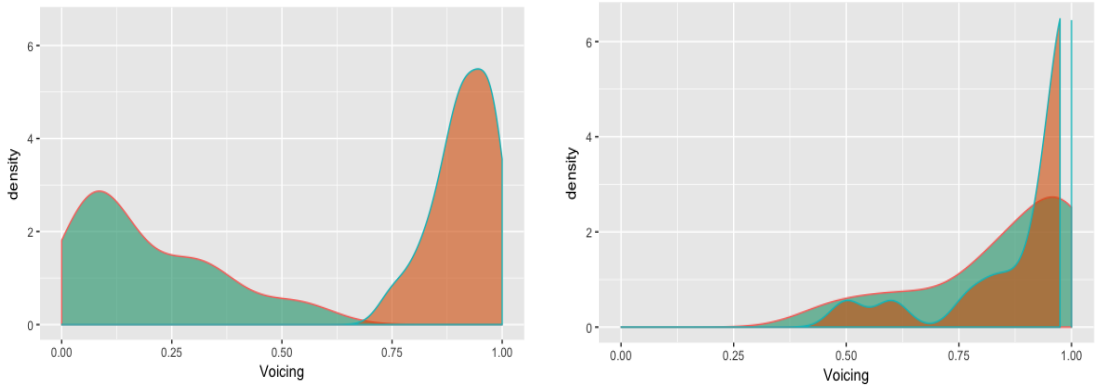
Figure 3 and Table 4 illustrate the degree of variability in the data for the first of the two rhotic parameters, rhotic % voicing.

In French (Figure 3a), significantly more variability is attested in voiceless clusters (F = 1.65, $p < 0.0001$): while this difference is minimal in terms of the standard deviation in the group mean (voiceless 34; voiced 28), the difference in the range of individual mean values is large (voiceless 81; voiced 27). In Spanish (Figure 3b), the variability in voiceless and voiced clusters is more similar. While the standard deviation is significantly greater in voiceless clusters (34 versus 26 for voiced; F = 1.73, $p < 0.0001$), the individual mean value is greater in voiced clusters (64 versus 51).

Table 4. Rhotic % voicing: summary statistics.

Language	Voiceless			Voiced		
	M	SD	Min–Max (Range)	M	SD	Min–Max (Range)
French	25	34	4–85 (81)	89	28	73–100 (27)
Spanish	85	34	49–100 (51)	91	26	36–100 (64)

[7] Results of a two-sample test for proportions indicate that the proportion of stops versus approximant realizations with French voiceless stops is not significantly higher than for French voiced ones ($z = -0.55$; $p = 0.58$). It is important to keep in mind that this test measures differences in proportions as opposed to differences in variability, which is reflected in the standard deviation and the range.

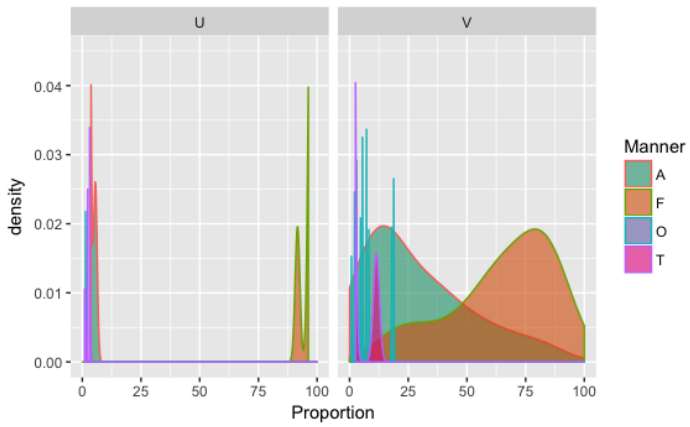


(**a**) (**b**)

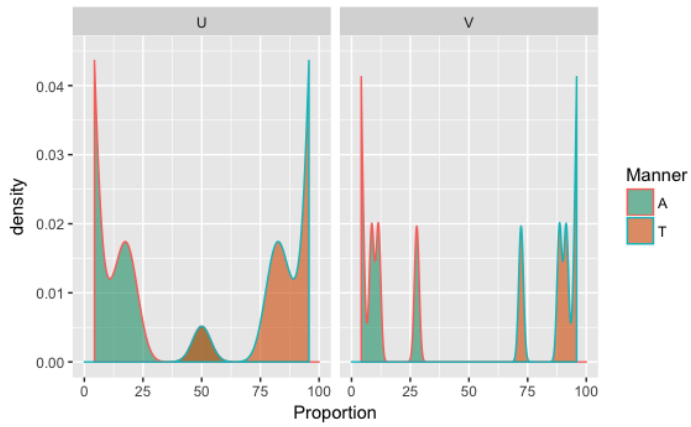
Figure 3. Density plots of rhotic voicing (%) in French (**a**) and Spanish (**b**) voiceless (green) and voiced (orange) word-medial SR clusters.

3.4. Rhotic Manner

Spanish is once again characterized by a lesser degree of variability with rhotic manner (Figure 4 and Table 5), especially with voiced clusters. Spanish /ɾ/ is realized as a tap in the majority of cases (voiceless 93%; voiced 98%) although a two-sample test of proportion reveals that the proportion of taps is significantly smaller in voiceless versus voiced clusters ($z = -2.65$, $p = 0.008$).



(**a**)



(**b**)

Figure 4. Density plots of rhotic manner (%) in French (**a**) and Spanish (**b**) voiceless (U; left) and voiced (V; right) word-medial SR clusters.

Table 5. Rhotic manner: summary statistics.

Language	Voiceless					Voiced				
	M			SD	Min–Max (Range)	M			SD	Min–Max (Range)
	Tap	Fric	App			Tap	Fric	App		
French		99	1	2	94–100 (6)		56	44	25	11–88 (77)
Spanish	93	6	1	10	73–100 (27)	98	2		6	73–100 (27)

Note: "SD" and "Min–Max (Range)" values are based on the percentage of fricatives in French and taps in Spanish.

In stark contrast, while the French rhotic is generally a fricative in voiceless clusters (99%), in voiced contexts, it is realized as an approximant in almost half (44%) of realizations. The difference in proportion of fricatives between voiceless and voiced contexts is significant ($z = 12.71$, $p < 0.0001$).

3.5. Summary of Variability in Stop and Rhotic % Voicing and Manner

Table 6 provides a summary of the relative within-language variability for the four phonetic parameters discussed. While the measurements presented in the preceding sections also allow for between-language comparisons, in order to test hypotheses based on differences between target languages, the characteristics of the learner groups (target language (TL) proficiency at time of testing, and quantity and quality of input encountered over the course of acquisition, among other variables) would have to be extremely similar. As we acknowledge differences in the profiles of the groups of English-speaking French and Spanish learners who participated in the present study (see Section 4.1 "Participants" below for further details), no between-language hypotheses will be proposed or tested here. However, we underline the interest and importance of doing so in future research: examining the acquisition of two different languages by learners sharing the same L1 allows for the teasing apart of the respective contribution of transfer and universal developmental effects and, as we will discuss in Section 6, both types of effects may mitigate the role of input variability on L2 speech learning.

For each combination of parameter (stop % voicing and manner; rhotic % voicing and manner) and language (within French, within Spanish), the cluster type indicated (voiceless/voiced) is the one for which there is greater variability in native speakers' production. Based on input variability alone, these are consequently the clusters for which the parameters in question are predicted to be relatively easier to acquire. Note that in the case of Spanish rhotic manner, the difference in the proportion of taps (voiceless 93%; voiced 98%), while statistically significant, is so small in absolute terms that we consider this difference not to be real.

Table 6. Summary of relatively greater within-language (French, Spanish) phonetic variability in word-medial SR clusters by stop and rhotic phonetic parameters (% voicing, manner).

Parameter	Relative Greater Variability	
	Within French	Within Spanish
Stop		
% Voicing	Voiceless	Voiceless
Manner	Voiceless	Voiced
Rhotic		
% Voicing	Voiceless	N/A
Manner	Voiced	N/A

On the general assumption that relatively greater within-category variability in the input has a positive effect on the acquisition of target structures, we made the following specific predictions concerning the relative difficulty of acquiring the above four phonetic parameters in voiceless versus voiced SR clusters for French and Spanish respectively:

Hypothesis 1. *In French, learners will acquire stop voicing and manner as well as rhotic % voicing more readily in voiceless clusters; rhotic manner will be easier to acquire in voiced clusters.*

Hypothesis 2. *In Spanish, acquiring stop voicing should be easier in voiceless clusters, whereas acquiring stop manner should be easier in voiced clusters. No variability-related differences are predicted for rhotic % voicing and manner.*

With these hypotheses in mind, we now turn to the experimental study designed to test them.

4. Materials and Methods

In order to test the variability-based hypotheses outlined in the previous section, intermediate and advanced English-speaking learners of French and Spanish were tested on their production of word-medial SR clusters via the same sentence-reading task used to elicit the native French and Spanish speaker data just examined. The study outlined here received approval from the University of Toronto Research Ethics Board and all participants provided written consent. In the following sections, we outline the methodological aspects of the study not already discussed with reference to the native speakers, namely, the learners' profiles, the way in which proficiency was determined via an additional production task, and further details of the data analysis.

4.1. Participants

Thirty-nine English-speaking learners participated (10 each of intermediate and advanced proficiency for French; 9 intermediate and 10 advanced Spanish learners). Given that the learners were to be tested on low-level phonetic parameters, it was important to ensure that they had been exposed to sufficient native speaker input. Indeed, in contrast to the majority of previous research on the role of input phonetic variability that has looked at the L2 acquisition of phonemic contrasts or lexical items by inexperienced learners, in order to become aware of the type of variability studied here, learners arguably require considerable experience with the target language, including with a variety of native speakers. Thus, during recruitment, learners were asked to have spent a minimum immersion period in a French- or Spanish-speaking milieu (intermediate: 3 months; advanced: 6 months). The lower range of the French intermediate speakers is due to the fact that two learners had not spent time in a French-speaking context. However, both individuals had undertaken French immersion schooling in Canada during which they would have been exposed to a wide variety of speakers over many years. In Spanish, we find quite the opposite distribution; namely, the intermediate speakers have spent, on average, more time in Spanish-speaking contexts than the advanced speakers. This difference is due to the presence of one outlier: one of the participants in the intermediate group had lived in Chile for approximately 30 years. Table 7 summarizes other relevant aspects of the learners' profiles.

Table 7. Learner profiles.

Group	Age (Years)				Immersion Experience (Months)	
	Onset of Acquisition		At Testing			
	M	Range	M	Range	M	Range
French						
Intermediate	11	5–21	33	21–52	8	0–36
Advanced	9	4–13	40	26–61	49	1–180
Spanish						
Intermediate	21	14–34	39	18–66	66	1–408
Advanced	18	11–29	30	21–35	14	3–36

Proficiency levels, which were first established prior to testing based on learners' self-evaluation, were verified via the information gathered through a conversation with each speaker and the use of a background questionnaire, as well as through a reading passage administered as part of the experiment. As is typical of many L2 phonetic production studies (e.g., Bongaerts et al. 2000; Birdsong 2008; Saito et al. 2016), accentedness scores were used as an objective measure of target language oral proficiency. Each learner was asked to read aloud the French or Spanish version of the text "The North Wind and the Sun" (Appendix A). Following the experiment, for each language, the 20/19 learner readings were interspersed with those of five native speakers of each language. The two groups of recordings were then randomized and presented to two panels (one per language) of three native speaker judges who had no training in linguistics and were unaware of the goals of the study. Following the methodology outlined in Bongaerts et al. (2000) and Birdsong (2008), the judges were asked to rate each of the readings on a scale from 1 ("heavy accent; clearly non-native") to 5 ("no foreign accent; definitely native"). The judges' ratings (Appendix B) were consistent with the proficiency level determined during subject recruitment. The average rating for the intermediate and advanced groups was noticeably different for both target languages. (French—intermediate: mean 2.1, range 1.5–2.8; advanced: mean 3.8, range 2.8–4.5; Spanish—intermediate: mean 2.3, range 1.5–3.2; advanced: mean 3.9, range 2.8–4.7).

4.2. Tasks

The learner-participants were tested individually in a quiet room. They performed three tasks: (i) the carrier sentence-reading task involving the word-medial SR stimuli in Table 1; (ii) a mirror English task included to test a set of unrelated hypotheses and thus discussed no further here; and (iii) the reading passage discussed above to determine proficiency. Stimuli for (i) were read three times, each time in random order. Out of a possible maximum of 1290 clusters (French: 12 clusters × 20 speakers × 3 rounds = 720; Spanish: 10 clusters × 19 speakers × 3 rounds = 570), once learner realizations that involved obvious misreadings or mispronunciations (e.g., metathesis or deletion; n = 84) were removed, a total of 1206 clusters were available for analysis (French: voiceless 347 (intermediate 175, advanced 172), voiced 351 (intermediate 175, advanced 176); Spanish voiceless 226 (intermediate 108, advanced 118), voiced 282 (intermediate 134, advanced 148)).

Between each round, the subjects were given approximately 5 minutes of break. During the first break, they completed a background questionnaire, which provided, among other things, the information summarized in Table 7. Following the three rounds, the participants read the language-appropriate version of 'The Northwind and the Sun'. Finally, the English stimuli were read once. All four rounds were recorded (44,100 Hz; 32-bit; stereo) using a Marantz CDR300 CD recorder (Marantz, Kawasaki, Japan) and unidirectional Audiotechnica AT803B lavaliere microphone (Audio-Technica U.S., Inc., Stow, OH, USA). Participants were remunerated $10 CDN.

4.3. Data Preparation and Analysis

Sound files were downsampled (22,050 Hz; 16-bit; mono) and tokens involving word-medial SR clusters were extracted and labeled. All such tokens were analyzed acoustically using Praat 4.0.41 (www.praat.org). As stated previously, voicing was measured in terms of the percentage of the stop or rhotic's articulation involving the presence of f0[8] in the spectrogram. In order to measure the duration of the stop, we took into account several parameters (see Figure 5). The onset of the stop was determined by a drop in intensity and a lowered first formant, whereas the stop offset was signaled by the presence of a burst, a rise in F1, and a rise in intensity (voiceless stops) or by the two latter parameters alone in the case of voiced stops, and, particularly, the approximant realizations. Values for percentage voicing were then averaged over the three repetitions of each

[8] The autocorrelation method—the default method in PRAAT for the analysis of intonation—was used.

token. Manner—stop, fricative, tap, approximant, trill, vocalization, or realizations of mixed manner (e.g., initial approximantization followed by frication or vice versa)—was transcribed and evaluated based on examination of periodicity and noisiness in the waveform and spectrogram (i.e., acoustic information was used to verify the transcription; no specific acoustic parameters were measured). Given the low proportion of vocalizations and mixed manner realizations, they were recoded as "Other" in the results presented in Tables 8–11. Results were exported to a spreadsheet and statistics were calculated with the Statistical Analysis Software (SAS).

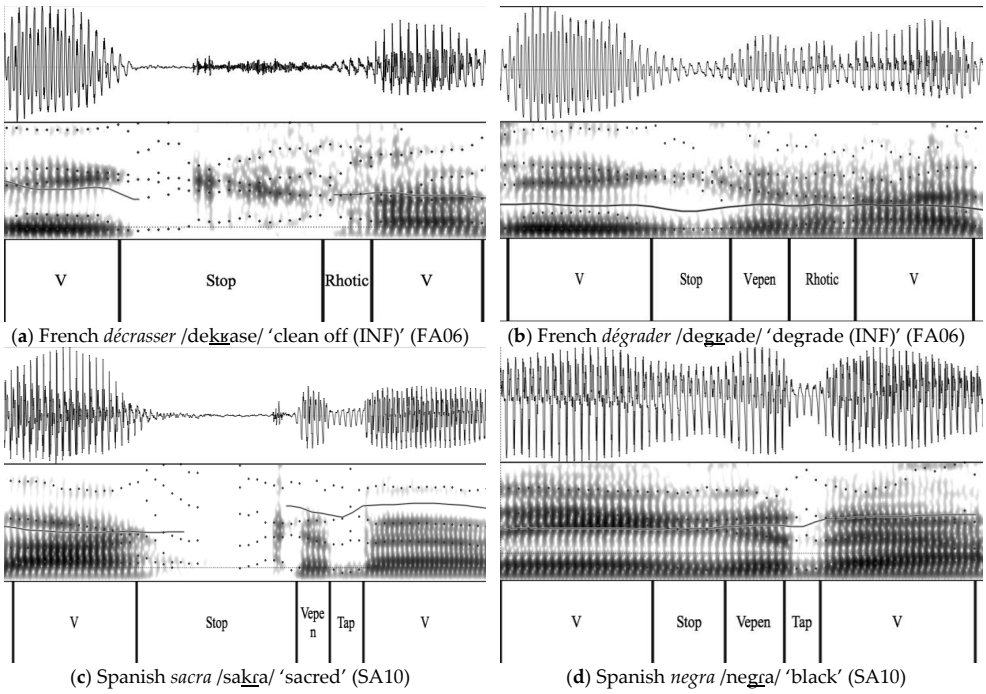


Figure 5. Sample segmentations of French (**a**) voiceless and (**b**) voiced; and Spanish (**c**) voiceless and (**d**) voiced SR clusters as realized by two advanced learners.

In order to determine whether learners had acquired a given parameter, we compared their production with that of the native speakers examined in Section 3. While cognizant of the importance of keeping in mind speaker characteristics—including the type of bilingualism—when constituting the control group in L2 acquisition studies (e.g., Grosjean 1988; Hulstijn 2012), our choice to use more French-/Spanish-dominant native speakers was triply motivated. First, the nature of the within-category voicing and manner variability in SR clusters has previously not been documented for French or Spanish with any population. It thus seemed logical to start with more monolingual, French-/Spanish-dominant speakers of the target languages. Second, all of the native speakers in both control groups had some knowledge of English and, thus, their speech was typical of the type of input to which our learners would have been exposed during their classroom learning experience. Finally, we are unaware of any literature having demonstrated L2-to-L1 influence on stop and rhotic % voicing and manner; as such, there was no empirical ground for believing that balanced bilinguals would constitute a better comparison group.

The quantitative criterion for determining whether a structure has been acquired has been the object of considerable debate in L2 acquisition research (see e.g., White 2003, pp. 77–78, for morphosyntax). In the present study, we started by using the criterion of ±2 standard deviations from the native speaker group mean. Adopting this criterion respected a central assumption concerning category formation and variability in L2 acquisition: interlanguage categories are based on archetypical values reflecting the totality of the heterogeneous input to which learners are exposed (e.g., Ellis 2003, in general; Flege 2009, p. 175, for L2 phonetics). The archetypical values for a parameter are reflected by the mean,

while the heterogeneity of the input can be measured by the standard deviation. However, for all but two of the comparisons (stop and rhotic manner in Spanish voiced SR clusters), this resulted in range minima that fell below the lowest of the attested individual native speaker means. Had the criterion of ±2 standard deviations from the native speaker group mean been kept, this would have resulted in overestimating the learners' ability, as some learners whose mean values for parameters resembled no native speaker control would nonetheless have been evaluated as having acquired those parameters. Accordingly, it was decided to use the attested native speaker group minima and maxima as the lower and upper values for a parameter to be deemed as acquired. On this criterion, learners were deemed to have acquired a phonetic parameter when their mean value for the parameter fell within the range of mean values of the native speaker controls.

5. Results

In the following sections, we will review the French then Spanish learners' production of the four phonetic parameters of the SR clusters analysed—stop % voicing and manner followed by rhotic % voicing and manner.

5.1. Stop Voicing and Manner

Tables 8–11 present the individual learner results for French and Spanish, respectively. Here and elsewhere, within the tables, learners are presented in increasing order of proficiency based on the average of the three judges' accentedness scores on 'The North Wind and the Sun' reading passage (see Appendix B for the accentedness scores). The shaded cells highlight the learner means that fall within the range of the control group's means for the parameter in question, that is, those parameters deemed to have been acquired.

Table 8. English-speaking learners' realization of stop % voicing and manner, and rhotic % voicing manner in French voiceless SR clusters. Here and elsewhere, shaded cells highlight learner means falling within the range of the individual means of the control group (i.e., that learners have acquired the parameter).

Learner	Stop			Rhotic				
	% Voicing	Manner (%)		% Voicing	Manner (%)			
		Stop	App		Tap	Fric	App	Other
FI02	16	100		93	24	76		
FI07	26	100		4		100		
FI01	20	100		88	17	11	56	16
FI10	15	100		5		100		
FI11	18	100		5		100		
FI04	75	100		28		100		
FI06	11	100		34		94	6	
FI08	13	100		7		100		
FI05	08	100		1		100		
FA06	24	100		34	19	75	6	
FI03	03	100		27		81	13	6
FA05	59	100		9		100		
FA09	40	100		4		100		
FA01	20	100		22		100		
FA10	62	100		19		100		
FA02	62	100		17		100		
FA07	28	100		15		94	6	
FA08	43	100		22		100		
FA04	49	100		3		100		
FA03	49	100		4		100		
NS Range	23–92	50–100		4–85		94–100		

App = approximant, Fric = fricative, NS = Native speaker.

Table 9. English-speaking learners' realization of stop % voicing and manner, and rhotic % voicing and manner in French voiced SR clusters.

Learner	Stop			Rhotic				
	% Voicing	Manner (%)		% Voicing	Manner (%)			
		Stop	App		Tap	Fric	App	Other
FI02	95	100		100	41	59		
FI07	30	94	6	1		100		
FI01	89	100		86	33	11	39	17
FI10	24	100		43		83	17	
FI11	57	100		62		94	6	
FI04	73	100		73		89	11	
FI06	31	94	6	65		83	17	
FI08	65	100		68		100		
FI05	67	100		56		94	6	
FA06	75	100		100	50	6	39	5
FI03	13	100		65		81	13	6
FA05	79	100		68		94	6	
FA09	56	100		97		87	13	
FA01	74	100		99		78	22	
FA10	77	100		72		89	11	
FA02	100	100		100		78	22	
FA07	100	100		95		94	6	
FA08	94	100		58		100		
FA04	82	100		72		100		
FA03	83	100		64		94	6	
NS Range	74–100	62–100		73–100		11–88		

Table 10. English-speaking learners' realization of stop % voicing and manner, and rhotic % voicing and manner in Spanish voiceless SR clusters.

Learner	Stop			Rhotic				
	% Voicing	Manner (%)		% Voicing	Manner (%)			
		Stop	App		Tap	Fric	App	Other
SI08	21	100		68	17		17	66
SI06	13	100		41	92			8
SI02	20	100		88	37	26	37	
SI10	3	100		86	87	13		
SI07	13	100		46	42			58
SI09	13	100		44	58			42
SA04	10	100		77	84	8	8	
SI04	39	100		29			17	83
SI01	2	100		59	100			
SI03	15	100		45	50	13	25	12
SA03	7	100		27	9	82	9	
SA02	52	100		67	100			
SA09	45	100		76	100			
SA10	40	100		92	58	33	9	
SA06	84	100		100	73		27	
SA07	30	100		97	100			
SA05	47	100		79	100			
SA08	37	100		51	100			
SA01	9	100		90	100			
NS Range	11–61	90–100		49–100	73–100			

Table 11. English-speaking learners' realization of stop % voicing and manner, and rhotic % voicing and manner in Spanish voiced SR clusters.

Learner	Stop			Rhotic				
	% Voicing	Manner (%)		% Voicing	Manner (%)			
		Stop	App		Tap	Fric	App	Other
SI08	33	100		70	17		17	66
SI06	96	47	53	73	100			
SI02	97	80	20	100	60		40	
SI10	55	100		100	60	20	20	
SI07	92	73	27	90	73			27
SI09	97	86	14	100	93		7	
SA04	89	86	14	96	93		7	
SI04	85	73	27	53	80	13	7	
SI01	88	60	40	72	100			
SI03	35	93	7	60	80	10	10	
SA03	84	47	53	89	80	7	7	6
SA02	100	7	93	80	100			
SA09	95	67	33	82	93			7
SA10	100	91	9	100	80		20	
SA06	99	87	13	100	60	7	33	
SA07	98	43	57	100	100			
SA05	93	88	12	100	74		13	13
SA08	100	13	87	80	100			
SA01	89	33	67	93	100			
NS Range	73–100	0–67		36–100	73–100			

Results for French % stop voicing in voiceless clusters (Table 8) indicate that three intermediate as well as nine advanced learners met the criterion for acquisition. In the case of voiced stops (Table 9), 11 learners of French (2 intermediate; 9 advanced) had acquired this parameter. As shown in Tables 8 and 9, stop manner did not present a problem for French learners, all of whom had acquired this parameter in both voiceless and voiced SR clusters. It is interesting to observe that, although manner varied in the realization of voiceless and voiced stops in many of the native speakers' production (see the density plots in Figure 2), the English-speaking learners almost categorically used stops. This may indicate that learners were using L1-based categories, which could also explain why some of the learners failed to produce target-like % voicing. It is also plausible that learners are realizing the most frequent, prototypical manner realization they encountered across native speakers; this would reinforce the use of L1 categories.

Overall, the Spanish learners performed better than the French learners. In voiceless clusters (Table 10), 13 of the learners' mean values (7 intermediate, 6 advanced) fell within the range of the control means.

With phonemically voiced stops (Table 11), only three intermediate learners produced insufficient phonetic voicing. Similar to their French counterparts, all Spanish learners were successful at acquiring stop manner in voiceless clusters (Table 10). However, the acquisition of manner in voiced clusters (Table 11), where all learners produced some percentage of approximants, proved to be extremely problematic. Recall from Figure 2 that stops are realized as approximants in 87% of the native speaker realizations of voiced SR clusters. Only 8 of the 19 learners (2 intermediate, 6 advanced) had acquired this parameter. All others produced a majority of stops.

5.2. Rhotic Voicing and Manner

Results obtained for % rhotic voicing by the L2 French speakers resemble those reported for stop voicing in terms of the learners performing relatively better as a whole with voiceless stop-rhotic clusters (Table 8) than with voiced ones (Table 9). Indeed, with voiceless clusters, 16 of the 20 learners

(7 intermediate, 9 advanced) had mean values within the range of those attested for the controls. Results for voiced clusters were much worse; only 8 of the 20 learners (3 intermediate, 5 advanced) reached criterion.

An opposite asymmetry is observed for rhotic % voicing among the L2 Spanish speakers. Here, learners were more successful at mastering rhotic % voicing in voiced (all learners) than in voiceless clusters (13 of 19; 4 intermediate, 9 advanced). The relatively higher success in voiced clusters may be attributed to voicing assimilation, given that the rhotic is preceded and followed by voiced segments.

French rhotic manner in voiced stop-rhotic clusters proved to be the most difficult of the French parameters—indeed the most difficult of any of the parameters in either language—to acquire; only 6 of 20 of the L2 learners (3 intermediate, 3 advanced) matched the controls. In voiceless clusters, 16 of 20 L2 speakers (7 intermediate, 9 advanced) had mastered this parameter.

Spanish learners were less accurate with rhotic manner in voiceless than voiced clusters, where 11 (3 intermediate, 8 advanced) versus 15 of the 19 L2 speakers (6 intermediate, 9 advanced) had acquired this parameter, respectively. In addition, Spanish learners showed a relatively clear proficiency-based pattern, with advanced speakers outperforming the intermediates with both voiceless and voiced SR clusters. This was not the case in French voiced clusters, where 3 intermediate and 3 advanced proficiency learners had target-like production values.

5.3. Summary of Results: Overall Accuracy

In presenting our general hypothesis that within-category variation facilitates the acquisition of new phonetic patterns, we proposed that ease of acquisition could be measured in terms of learners' accuracy with the different parameters in question. Consequently, for each of the target languages, in this section, we will summarize the results in terms of the total number of learners of intermediate and advanced proficiency who mastered a given parameter (overall accuracy).

Table 12 provides a summary of the total number of learners out of a maximum of 20 (French) or 19 (Spanish) who mastered each of the four parameters for both voiceless and voiced SR clusters for the two learner proficiency groups.

Table 12. Number (percentage) of English speaking-learners acquiring stop and rhotic parameters (% voicing, manner) in French (Maximum = 20) and Spanish (Maximum = 19) voiceless and voiced medial SR clusters.

Language	Voiceless				Voiced			
	Stop		Rhotic		Stop		Rhotic	
	% V	Manner	% V	Manner	% V	Manner	% V	Manner
French								
Intermediate	3	10	7	7	2	10	3	3
Advanced	9	10	9	9	9	10	5	3
All	12 (60%)	20 (100%)	16 (80%)	16 (80%)	11 (55%)	20 (100%)	8 (40%)	6 (30%)
Spanish								
Intermediate	7	9	4	3	6	2	9	6
Advanced	6	10	9	8	10	6	10	9
All	13 (68%)	19 (100%)	13 (68%)	11 (58%)	16 (84%)	8 (42%)	19 (100%)	15 (79%)

Note: % V = % Voicing.

Having now established the intermediate and advanced English-speaking learners' accuracy with each the four parameters for both target languages, we are ready to evaluate the specific hypotheses underlying the study.

5.4. Hypotheses Evaluation

The experimental study set out to test the two specific hypotheses concerning the contribution of within-category variability to learning difficulty, which we repeat below:

Hypothesis 1. *In French, learners will acquire stop voicing and manner as well as rhotic % voicing more readily in voiceless clusters; rhotic manner will be easier to acquire in voiced clusters.*

Hypothesis 2. *In Spanish, acquiring stop voicing should be easier in voiceless clusters, whereas acquiring stop manner should be easier in voiced clusters. No variability-related differences are predicted for rhotic % voicing and manner.*

Table 13 provides a summary of the evaluation of both of the hypotheses organized by target language (Hypothesis 1: French; Hypothesis 2: Spanish). Based on the summary of the results for the L2 learners presented in Table 12, each hypothesis was evaluated as follows: a given parameter was deemed to be more easily acquired for the cluster (voiceless or voiced) for which there were more learners whose mean production value fell within the range of those of the native speakers. For example, based on the fact that, for French, there was greater variability in stop % voicing in voiceless as opposed to voiced SR clusters in the native speaker controls' production, it was predicted that voicing in the former should be easier for the L2 learners to acquire. As shown in Table 13, this hypothesis was supported.

Table 13. Evaluation of the variability-related hypotheses with parameters listed predicted to be easier to acquire due to higher rates of within-category variability based on overall accuracy (intermediate and advanced learners).

Comparison	Parameter			
	Stop		Rhotic	
	% Voicing	Manner	% Voicing	Manner
Hypothesis 1: French				
Predicted greater ease	Voiceless	Voiceless	Voiceless	Voiced
Intermediate	√		√	X
Advanced			√	X
Hypothesis 2: Spanish				
Predicted greater ease	Voiceless	Voiced	N/A	N/A
Intermediate	√	X		
Advanced	X	X		

Note: A checkmark "√" indicates that the prediction was supported; "X" indicates that the prediction was rejected; an empty cell indicates that the data were inconclusive (i.e., equal number of learners mastering the parameter in both voiceless and voiced clusters). The Spanish rhotic parameters for which no significant difference in variability existed, and thus for which no difference in acquisitional difficulty was predicted, are indicated with "N/A".

Of the predictions in Hypothesis 1 made for French, one was supported by both the intermediate and advanced groups' performance (rhotic % voicing), another was refuted by both groups' results (rhotic % manner), a third was supported only by the intermediate learners' data (stop % voicing), and, for the final prediction concerning the greater ease of acquiring stop manner in voiceless SR clusters, the results were inconclusive: as highlighted earlier, all learners mastered this parameter in both voiceless and voiced clusters. In summary, for French, when the results for each of the proficiency-based groups are considered separately, the data relevant to evaluating the general hypothesis that within-category input variability leads to greater acquisitional ease more often support or are inconclusive (three cases each) than refute the specific predictions (two cases). In contrast, of the four predictions made in Hypothesis 2 for Spanish, three out of four comparisons refute the greater ease of acquiring stop % voicing in voiceless clusters (advanced learners) and stop manner

in voiced clusters (both proficiency levels). When the predictions made for both languages are considered together, there is slightly less support for the hypothesized positive correlation between within-category variability and ease of acquisition (four of nine conclusive comparisons).

In summary, based on both the intermediate and advanced learners' accuracy, the majority of the predictions concerning Spanish were refuted, as were the predictions regarding the acquisition of rhotic manner in French. In contrast, the data support or are consistent with the variability-based predictions that French stop and rhotic voicing should be more readily acquired in voiceless clusters. In the final section, we review those cases where our general input-variability-based hypothesis was not supported, discussing factors such as L1-based influence, universal articulatory constraints, and target language proficiency as well as other challenges inherent in integrating the role of variability when modeling L2 speech learning.

6. Discussion

Overall, the findings of the present study differ from the majority of previous research that has investigated the role of input phonetic variability in ease of L2 learning. Whereas this body of research has overwhelmingly found a positive effect for variability, the evidence presented here from the study of intermediate and advanced English-speaking learners' production of French and Spanish SR clusters is very much mixed. In an attempt to explain such differences, we begin by discussing the ways in which the current data set differs from those of previous studies as concerns target language proficiency and learners' awareness of the variability under investigation. Then, we propose that L1-based transfer and universal articulatory constraints must be taken into account, at least when making predictions for production. We also discuss the need to consider further some of the assumptions made here regarding input variability, including the types of variability learners are likely to encounter in real learning situations, their ability to parse the full range of variability in order that it become intake, as well as the potential need to distinguish between variability involving discrete variables with few variants versus continuous variables with a much higher degree of variability characteristic of the phonetic parameters that were investigated here.

6.1. Differences between the Present and Previous Research: Target Language Proficiency and Learners' Awareness of Variability

In the Introduction, we highlighted that the design of the present study differs from that of much, if not most, previous research on input phonetic variability's effect on ease of L2 learning in several ways. We discuss here those differences that might help to explain the lesser support for the general hypothesis found in the present study.

As discussed in Section 2, the effect of input variability has been studied most often via laboratory perceptual training studies involving participants with little to no experience with the target language. Accordingly, such studies provide insights first and foremost into the effect of this variable on phonemic categorization at the very earliest stages of learning. In contrast, the present study tested the same general hypothesis using production data from learners with considerable learning experience both in terms of the number of years of study and, with two exceptions, having a minimum of 3 (intermediates) or 6 months (advanced learners) of French-/Spanish-language immersion. The lack of support for many of the variability-based hypotheses found in the present study parallels the findings of Bohn and Bundgaard-Nielsen (2009). Recall that these researchers found that the least intelligible vowels produced by their Danish-speaking L2 learners were the same vowels that vary the most across English (American, Southern British, and Australian) dialects. The learners in this latter study also had formal target language learning experience (5–8 years in Denmark) although, unlike our learners, had little to no immersion experience. It may be that the effects of input variability on ease of learning are strongest at the beginning of L2 acquisition when one of the main learning objectives is target-like category formation. This hypothesis is supported by the current study's data in that three of the four

parameters that provided support for a positive role for acoustic input variability involved the less experienced intermediate learners' production.

A second important way in which the present study differs from most previous research is that, in laboratory perception studies employing high variability phonetic training, learners' attention is directed towards the contrast of interest via the task. For example, the Japanese-speaking learners of English in Logan et al. (1991) were trained using a minimal pair /l-ɹ/ identification task involving explicit feedback. In contrast, it is unclear how the reading task in the present study could have heightened learners' awareness of the type of phonetic variability involved, which is arguably below the level of conscious awareness. Given the important role attributed to awareness in L2 learning (e.g., Robinson et al. 2011), this task-based difference may be relevant. We will discuss further limitations on L2 learners' analysis of the input, including the role of awareness in Section 6.3.

6.2. Determining the Relative Importance of Input Variability versus Transfer and Articulatory Constraints when Predicting Relative Ease of Acquisition

When formulating the specific hypotheses tested in the present study, the sole factor considered was the relative degree of acoustic input variability. In the case of L2 acquisition, such variability is parsed using existing L1-influenced categories. Production too is influenced by a learner's L1 as well as by universal articulatory constraints. As such, hypotheses must consider both of these factors. Teasing the role of variability apart from transfer and articulatory complexity is not a straightforward task. In some cases, they make opposite predictions. For example, while stop voicing is more variable in French voiceless than voiced SR clusters, the learners' L1 and the target language categories are more similar in voiceless contexts. The English-speaking learners' relatively greater success with stop and rhotic % voicing in voiceless contexts—a result that is in keeping with our input-variability-based hypothesis—may indeed be related to transfer. In the case of stops in both languages, learners seem to be using their L1 voiced stop categories, which are insufficiently voiced. In the case of French rhotics, the English-speaking participants may be simply transferring their native articulatory patterns, which also involve devoicing of the following liquid.

Their accuracy with rhotic % voicing may also be related to the fact that aerodynamic constraints favor voicing assimilation of the rhotic, leading naturally to target-like low levels of voicing in /ʁ/.[9] The use of L1 articulatory patterns allows for rapid accuracy, in spite of the challenge present in analyzing the target language variability. The same can be said about the challenge of considering articulatory complexity. French L2 speakers were less successful at acquiring the rhotic parameters in voiced than in voiceless clusters; this was particularly the case of rhotic manner, which was only acquired by six speakers in the former context. This could be attributed to the fact that realizing a voiced fricative is problematic from an aerodynamic point of view: whereas fricatives require an open glottis for sufficient airflow, this glottal configuration disfavors voicing. Finally, variability in L2 production does not always mirror input variability. This is illustrated with the variation in rhotic manner observed with the intermediate speakers of Spanish. As indicated in Tables 10 and 11, these learners produced a high proportion of rhotics that fell within the category 'Other', which included retroflexed rhotics not attested in Spanish that clearly resulted from transfer from English. As such, even if input variability increases the range of possible targets, complicating the analysis of categories and thus the process of category formation, input cannot be the sole source of these differences between L2 and target variability.

---

[9] The parallel failure to produce fully voiced stops may be due to durational differences. Normally, L2 learners are less fluent than native speakers (e.g., Munro and Derwing 1998; Towell et al. 1996). As such, L2 learners tend to produce longer segments and greater duration impedes the maintenance of voicing.

6.3. Considering Learners' Perception and Analysis of the Input

Numerous studies have shown that L2 learners fail to perceive some consonantal and vocalic contrasts (e.g., Cebrian 2006; Flege et al. 1996; Flege and MacKay 2004; Guion et al. 2000) as well as consonant sequences (e.g., Kabak 2003; Matthews and Brown 2004) in the same way as native speakers. Some of these difficulties may be overcome with experience and training, whereas others present greater challenge. Given the existing evidence that L2 learners sometimes fail to perceive contrasts in the L2 that are absent in their L1, it is possible that they may fail to notice variation in the target language. This is supported by Strange's (2009) proposal that, when task complexity increases, L2 learners use a primarily phonological level of processing. This would preclude learners—at least those of lower levels of proficiency and having less experience with the target language—from being able to parse the range of phonetic variation attested in real-world communicative situations. Failure to notice the type of subphonemic phonetic variation studied here is also in keeping with the phonetic-relevance hypothesis discussed in Barcroft and Sommers' (2005, 2014; Sommers and Barcroft 2007) studies that proposes that variability is relevant only when it targets contrastive parameters. If this is the case, the types of analyses of variation that were undertaken here in order to formulate the specific hypotheses tested would simply not be possible for many learners, even those having considerable target language experience. We propose that selective, non-native-like perception offers a partial explanation for the failure of learners of Spanish to realize underlying stops as approximants in voiced clusters: not only is approximantization of underlying stops not a feature of their L1, but transferred English grapheme-phoneme mappings would only enhance the probability of learners adopting a stop analysis. It is also possible that, even if such variation is noticed, non-native speakers may interpret it differently from native speakers. Indeed, it may be the case that, particularly at initial stages, learners interpret allophonic variation in the target language as phonemic and that, only with large quantities of input over longer periods of time, are target-like analyses possible. This interpretation may be reinforced by the existence of some overlap in the phonetic realization of the phonemic categories. For example, in French, the upper range of stop % voicing in phonemically voiceless stops overlaps with that of voiced stops (23–92% versus 74–100%; Table 2).

In a related vein, in the present study, variability in the target language was said to exist whenever statistically significant differences existed in the mean value and/or range of values for a given phonetic parameter between voiceless and voiced clusters. It may be important to nuance this operationalization. It may be the case that, even when statistically significant, differences are too small to be perceptible. Recall that, when formulating the specific hypothesis concerning Spanish stop manner, we highlighted that while greater variability was attested in voiced stops, the degree of variability in their voiceless counterparts was so limited as to be able to characterize voiceless stop manner as invariant. In such cases, it is arguably the case that using between-category variability (here, voiceless versus voiced) is not motivated.

It may also be the case that L1-shaped cue weighting may result in learners being more sensitive to certain types of variability and, consequently, relatively insensitive to others. For example, Escudero (2000) demonstrated that L2 learners may focus on the primary L1 phonetic cues to a phonological contrast even when such a cue is not used by native speakers of the target language in production. In a parallel fashion, it may be the case that learners fail to perceive phonetic variability, at least at earlier stages of acquisition, when L1 realizations of the same phonological category do not vary along the same phonetic parameter(s). In such cases, input variability could not be used to predict acquisitional difficulty, as it would not constitute part of learners' intake. In future work, it will be important to define variability not only in terms of production measures—including whether there are thresholds below or above which variability is too limited to be relevant to predicting relative difficulty—but also as determined by learners' sensitivity to existing differences via perception tests. Finally, it could also be assumed that, even when learners are sensitive to these differences and are able to establish native-like long-term phonetic representations, they lack the articulatory control necessary to realize the patterns attested.

6.4. What Is Learners' Input?

Perhaps the greatest challenge to predicting the effects of variability is determining the actual input to which learners are exposed over the course of acquisition. For both French and Spanish, the specific predictions formulated here were based on the production of 20 native speakers of two different varieties. One might question whether it is reasonable to assume that all of the L2 learners in the present study, even with the requirement of a minimum of 3 or 6 months of immersion, would have encountered such a degree of variability. For some learners, particularly the most proficient, this is likely. For other learners, the range of variability encountered might be much less than that measured in the native speaker controls in the present study. Furthermore, the extent to which variability predicts ease of acquisition might depend upon the point at which it is encountered. As proposed in Section 6.1, if learners' sensitivity to variation is indeed related to phonetic properties of the L1, they might be less sensitive at earlier stages of learning. However, it is also possible that the instantiation of phonetic categories in long-term memory over the course of acquisition might desensitize L2 learners to variability if it is not encountered early enough. This would be true of those classroom learners having little exposure to speakers other than their (non-native) instructors during the first years of learning. In summary, it may be necessary not only to know what degree of variability existed in a learner's input, but also at which point along the acquisitional path it was encountered. Determining the degree of input variability at various stages of learning may be, unfortunately, impossible (see (Flege 2009) for further discussion).

6.5. Types of Variability

In future research, it will be necessary to distinguish between different types of variability and the relatively different challenges they may pose for learners. Determining the role that input variability plays is complicated by the fact that the label 'within-category' variation includes inter- and intraspeaker variation in the realization of a given category. For example, a given phoneme such as a French voiceless stop may be realized categorically as a stop by some speakers or variably as a stop or approximant by others. Another phoneme, like the French rhotic, may show a higher degree of within-category variation, varying in manner among all speakers. These two types of within-category variation should pose different problems for L2 learners. The first type should be easier for learners to manage, as they have a wider range of options for realizing a target structure, all of which constitute target-like production. The second type may be more difficult to perceive and produce. In addition to the range of within-category variation, languages vary in how sharp the contrasts between categories are, and native speakers vary on how the perceive and produce these contrasts (Perkell et al. 2006). Although not analyzed here, it is expected that sharper phonemic contrasts will pose fewer difficulties to L2 learners than cases where there is some degree of overlap between members of the categories (see Wade et al. 2007).

6.6. Measuring Acquisition and Relative Ease of Learning

We wish to conclude our discussion by exploring further the criterion used to determine whether a given structure has been acquired. In the present study, acquisition was equated with learner values falling in the range of the control means for a given phonetic parameter. Other criteria have been proposed and used in the past, as was highlighted in Section 4.3. First, we could have simply used the control group's mean as the point of reference. However, while using means would have been relatively meaningful for some of the parameters, this would not have been the case for others. For example, stop or rhotic manner in French and voiceless stop and rhotic manner in Spanish are relatively less variable than the other parameters that were explored here. While we could argue that learners may be able to calculate and use means in such cases, means are not representatives of the native speakers' performance with the other highly variable parameters. Indeed, very few native speakers have values close to the group mean; the median may indeed be more representative of typical possible values in

such cases. Moreover, in the case of highly variable parameters, it is not clear that acquiring the mean equals having acquired the target category, which is intrinsically variable. Second, as we have done in past research (Colantoni and Steele 2006), we could have undertaken a speaker-by-speaker analysis, comparing each of the L2 learners against each of the native speakers for all of the parameters. In this case, acquisition would be deemed to have occurred when a given learner matches any native speaker control for all parameters. While a useful method for evaluating native-like ultimate attainment, in the study of variability, using a speaker-to-speaker comparison would go against the idea that learners use the aggregate input in forming their categories. If this is true, and if this input is more varied than for L1 learners in terms of the range of native and non-native speakers who serve as models, it is arguably unwise to compare learner production with individual native grammars.

7. Conclusions

Explaining the great degree of variation in L2 production has been, and will continue to be, a central goal in developing theories of non-native acquisition. In the present work, we have sought to contribute to this line of research by investigating the degree to which within-category target language variability predicts the difficulty of acquisition of four phonetic parameters in French and Spanish stop-rhotic clusters. The general hypothesis that a greater degree of variability in the input results in more accurate acquisition was partially supported, in that it predicted a subset of the specific hypotheses formulated. In those cases where hypotheses were not supported, such as the acquisition of French rhotic manner or Spanish stop manner, we highlighted areas that require further consideration including (i) the difficulty in measuring the variability in the actual input to which a given learner is exposed; (ii) the need to determine the extent to which learners' perception may lead to intake which makes them less sensitive to variability including how this might be affected by target language experience and proficiency; and (iii) teasing the effects of input variability from those of transferred L1 articulatory patterns and universal production constraints.

In future research, it will be necessary to build on the findings here. This will include expanding the target structures to take into account prosodic phenomena such as stress and intonation. In particular, it will be of interest to test learners on sets of structures for which a clearly defined hierarchy of variability can be established, so as to be able to further test the general hypothesis proposed here. It would also be of interest to conduct studies involving learners with a range of target language experience in order to test directly the hypothesis that the influence of input variability may wane with time as learners move from primarily categorizing to producing the target language. On a methodological level, we have highlighted the need to couple production studies on the acquisition of variability with perception studies that will allow us to determine to what extent L2 hearers are sensitive to low-level phonetic variation and, thus, refine the general proposal that variability increases the learning challenge. In conclusion, pushing forward our understanding of the role of variability in L2 learning will require insights from both perception and production formalized in a multifactor model that takes into account all of the aspects discussed here.

Author Contributions: Both authors have contributed equally to all aspects of this paper.

Acknowledgments: Funds for data collection were received from the University of Toronto Connaught Fund (both authors) and from Victoria College, University of Toronto (Colantoni).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. French and Spanish Texts (Proficiency Measure)

French: "La bise et le soleil"

La bise et le soleil se disputaient, chacun assurant qu'il était le plus fort. Quand ils ont vu un voyageur qui s'avançait, enveloppé dans son manteau, ils sont tombés d'accord que celui qui arriverait le premier à le lui faire ôter serait reconnu comme le plus fort. Alors, la bise s'est mise à souffler de toutes ses forces mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui. Finalement, elle renonça à le lui faire ôter.

Alors, le soleil commença à briller et au bout d'un moment le voyageur, réchauffé, ôta son manteau. Ainsi, la bise dut reconnaître que le soleil était le plus fort.

Spanish: "El viento del norte y el sol"

El viento del norte y el sol discutían acerca de cuál de los dos sería el más fuerte, cuando, de repente, pasó un viajero envuelto en una amplia capa. Al verlo, convinieron en que el primero que consiguiera quitarle la capa sería el más fuerte. El viento del norte comenzó a soplar con mucha furia, pero, cuanto más soplaba, más se aferraba el viajante a su capa, hasta que el viento norte desistió. El sol brilló entonces con todo su esplendor, e inmediatamente, el viajante arrojó su capa. Así, el viento norte tuvo que reconocer la superioridad del sol.

Appendix B. Native Speaker Judges' Evaluation of French and Spanish Learners' Proficiency Level Based on Reading Passage

Table A1. French learners.

Intermediate					Advanced				
	Judge			Average		Judge			Average
	1	2	3			1	2	3	
FI01	2	2	1	1.7	FA01	4	4	2.5	3.5
FI02	1.5	2	1	1.5	FA02	4.5	3.5	4	4
FI03	3.5	3	2	2.8	FA03	5	4.5	4	4.5
FI04	2	2.5	2	2.2	FA04	4	4.5	4.5	4.3
FI05	3	3	1.5	2.5	FA05	3	3.5	3.5	3.3
FI06	2	2.5	2	2.2	FA06	3	3.5	2	2.8
FI07	1	2.5	1	1.5	FA07	4.5	4	4	4.2
FI08	2	2.5	2.5	2.3	FA08	4	4	4.5	4.2
FI10	2	2.5	1.5	2	FA09	4	4	2	3.3
FI11	2	2.5	1.5	2	FA10	4.5	3.5	3.5	3.8
Average				2.1	Average				3.8

Table A2. Spanish learners.

Intermediate					Advanced				
	Judge			Average		Judge			Average
	1	2	3			1	2	3	
SI01	3	3.5	3	3.2	SA01	4.5	5	4.5	4.7
SI02	2	2	1.5	1.8	SA02	3.5	4	3.5	3.7
SI03	3.5	3	3	3.2	SA03	3	4	3.5	3.5
SI04	2.5	4	2.5	3	SA04	2.5	3	3	2.8
SI06	1.5	1.5	2	1.7	SA05	4	4	4.5	4.2
SI07	2.5	2	2	2.2	SA06	4	4	4	4
SI08	1	2	1.5	1.5	SA07	4	4	4	4
SI09	1	3	2.5	2.2	SA08	4.5	4	4.5	4.3
SI10	1.5	2	2	1.8	SA09	3.5	4	4	3.8
					SA10	3.5	4	4	3.8
Average				2.3	Average				3.9

References

Barcroft, Joe, and Mitchell Sommers. 2005. Effects of acoustic variability on second language vocabulary learning. *Studies in Second Language Acquisition* 27: 387–414. [CrossRef]

Barcroft, Joe, and Mitchell Sommers. 2014. Effects of variability in fundamental frequency on L2 vocabulary learning. *Studies in Second Language Acquisition* 36: 423–49. [CrossRef]

Birdsong, David. 2008. Nativelike pronunciation among later learners of French as a second language. In *Language Experience in Second Language Speech Learning. In Honor of James Emil Flege*. Edited by Ocke-Schwen Bohn and Murray J. Munro. Amsterdam and Philadelphia: John Benjamins, pp. 99–116.

Blecua, Beatriz. 2001. Las vibrantes del español: Manifestaciones acústicas y procesos fonéticos. Ph.D. thesis, Universidad Autónoma de Barcelona, Barcelona, Spain.

Blecua, Beatriz. 2008. Los sonidos vibrantes: Aspectos comunes y variación. *Language Design* 1: 23–30.

Bohn, Ocke-Schwen, and Rikke Bundgaard-Nielsen. 2009. Second language speech learning and diverse inputs. In *Input Matters in SLA*. Edited by Thorsten Piske and Martha Young-Scholten. Bristol, Buffalo and Toronto: Multilingual Matters, pp. 207–18.

Bongaerts, Theo, Susan Mennen, and Frans van der Slik. 2000. Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia Linguistica* 54: 298–308. [CrossRef]

Brousseau-Lapré, Françoise, Susan Rvachew, Meghan Claynards, and Daniel Dickson. 2013. Stimulus variability and perceptual learning of nonnative vowel categories. *Applied Psycholinguistics* 34: 419–41. [CrossRef]

Cebrian, Juli. 2006. Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics* 34: 372–87. [CrossRef]

Clopper, Cynthia, and David Pisoni. 2004. Effects of talker variability on perceptual learning of dialects. *Language and Speech* 47: 207–39. [CrossRef] [PubMed]

Colantoni, Laura, and Jeffrey Steele. 2005. Phonetically-driven epenthesis asymmetries in French and Spanish obstruent-liquid clusters. In *Experimental and Theoretical Approaches to Romance Linguistics*. Edited by Randall Gess and Edward J. Rubin. Amsterdam and Philadelphia: John Benjamins, pp. 77–96.

Colantoni, Laura, and Jeffrey Steele. 2006. Native-like attainment in the L2 acquisition of Spanish stop-liquid clusters. In *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages, Minneapolis, October 2004*. Edited by Carol A. Klee and Timothy L. Face. Somerville: Cascadilla Press, pp. 59–73.

Colantoni, Laura, and Jeffrey Steele. 2007a. Voicing-dependent cluster simplification asymmetries in Spanish and French. In *Segmental and Prosodic Issues in Romance Phonology*. Edited by Pilar Prieto, Joan Mascaró and Maria-Josep Solé. Amsterdam and Philadelphia: John Benjamins, pp. 109–29.

Colantoni, Laura, and Jeffrey Steele. 2007b. Acquiring /R/ in context. *Studies in Second Language Acquisition* 29: 381–406. [CrossRef]

Colantoni, Laura, and Jeffrey Steele. 2008. Integrating articulatory constraints into models of second language phonological acquisition. *Applied Psycholinguistics* 29: 489–534. [CrossRef]

Colantoni, Laura, and Jeffrey Steele. 2011. Synchronic evidence of a diachronic change: Voicing and duration in French and Spanish stop-liquid clusters. *Canadian Journal of Linguistics* 56: 147–77.

Ellis, Nick. 2003. Constructions, chunking, and connectionism: The emergence of second language structure. In *The Handbook of Second Language Acquisition*. Edited by Catherine J. Doughty and Michael H. Long. Malden: Blackwell Publishing, pp. 63–103.

Escudero, Paola. 2000. Developmental Patterns in the Adult L2 Acquisition of New Contrasts: The Acoustic Cue Weighting in the Perception of Scottish Tense/Lax Vowels in Spanish Speakers. Master's thesis, University of Edinburgh, Edinburgh, UK.

Flege, James. 2009. Give input a chance! In *Input Matters in SLA*. Edited by Thorsten Piske and Martha Young-Scholten. Bristol, Buffalo and Toronto: Multilingual Matters, pp. 175–90.

Flege, James, and Ian MacKay. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition* 26: 1–34. [CrossRef]

Flege, James, Anne Marie Schmidt, and Glenn Wharton. 1996. Age of learning affects rate-dependent processing of stops in a second language. *Phonetica* 53: 143–61. [CrossRef] [PubMed]

Freed, Barbara F., Norman Segalowitz, and Dan Dewey. 2004. Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition* 26: 275–301. [CrossRef]

Giannakopoulou, Anastasia, Helen Brown, Meghan Claynards, and Elizabeth Wonnacott. 2017. High or low? Comparing high and low-variability phonetic training in adult and child second language learners. *PeerJ* 5: e3209. [CrossRef] [PubMed]

Grosjean, François. 1988. Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1: 131–49. [CrossRef]

Guion, Susan, James Flege, Reiko Akahane-Yamada, and Jesica C. Pruit. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* 107: 2711–24. [CrossRef] [PubMed]

Hardison, Debra. 2003. Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics* 24: 495–522. [CrossRef]

Hulstijn, Jahn. 2012. The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition* 15: 422–33. [CrossRef]

Iverson, Paul, Valerie Hazan, and Kerry Bannister. 2005. Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *Journal of the Acoustical Society of America* 118: 3267–78. [CrossRef] [PubMed]

Kabak, Baris. 2003. The Perceptual Processing of Second Language Consonant Clusters. Ph.D. thesis, University of Delaware, Delaware, MD, USA.

Léon, Pierre. 1992. *Phonétisme et Prononciations du Français*. Paris: Nathan.

Lively, Scott, David Pisoni, Reiko Yamada, Yoh'ichi Tohkura, and Takashi Yamada. 1994. Training Japanese listeners to identify English /r/-/l/. III. Long term retention of new phonetic categories. *Journal of the Acoustical Society of America* 96: 2076–87. [CrossRef] [PubMed]

Logan, John, Scott Lively, and David Pisoni. 1991. Training Japanese listeners to identify English /r/-/l/. A first report. *Journal of the Acoustic Society of America* 89: 874–86. [CrossRef]

Matthews, John, and Cynthia Brown. 2004. When intake exceeds input: Language specific perceptual illusions induced by L1 prosodic constraints. *International Journal of Bilingualism* 8: 5–27. [CrossRef]

Möbius, Bernd. 2004. Corpus-based investigations on the phonetics of consonant voicing. *Folia Linguistica* 38: 5–26.

Moyer, Alene. 2009. Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment. In *Input Matters in SLA*. Edited by Thorsten Piske and Martha Young-Scholten. Bristol, Buffalo and Toronto: Multilingual Matters, pp. 159–74.

Moyer, Alene. 2011. An investigation of experience in L2 phonology: Does quantity matter more than quality? *The Canadian Modern Language Review* 67: 191–216. [CrossRef]

Munro, Murray, and Tracey Derwing. 1998. The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning* 48: 159–82. [CrossRef]

Nishi, Kanae, and Diane Kewley-Port. 2007. Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research* 50: 1496–509. [CrossRef]

O'Shaughnessy, Douglas. 1982. A study of French spectral patterns for synthesis. *Journal of Phonetics* 10: 377–99.

Perkell, Joseph S., Frank H. Guenther, Harlan Lane, Nicole Marrone, Melanie L. Matthies, Ellen Stockmann, Mark Tiede, and Majid Zandipour. 2006. Production and perception of phoneme contrasts covary across speakers. In *Speech Production: Models, Phonetic Processes and Techniques*. Edited by Jonathan Harrington and Marija Tabain. New York: Psychology Press, pp. 69–84.

Pisoni, David, and Scott Lively. 1995. Variability and invariance in speech perception: A new look at some old problems in perceptual learning. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Timonium: York Press, pp. 433–59.

Pruitt, John, James Jenkins, and Winifred Strange. 2006. Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese. *Journal of the Acoustical Society of America* 119: 1684–96. [CrossRef] [PubMed]

Robinson, Peter, Alison Mackey, Susan M. Gass, and Richard Schmidt. 2011. Attention and awareness in second language acquisition. In *The Routledge Handbook of Second Language Acquisition*. Edited by Susan M. Gass and Alison Mackey. London and New York: Routledge, pp. 247–67.

Sadakata, Makiko, and James McQueen. 2013. High stimulus variability in non-native speech learning supports formation of abstract categories: Evidence from Japanese geminates. *Journal of Acoustical Society of America* 134: 1324–35. [CrossRef] [PubMed]

Saito, Kazuya. 2015. Experience effects on the development of late second language learners' oral proficiency. *Language Learning* 65: 563–95. [CrossRef]

Saito, Kazuya, Pavel Trofimovich, and Talia Isaacs. 2016. Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics* 37: 217–40. [CrossRef]

Snoeren, Natalie, Pierre Hallé, and Juan Segui. 2006. A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics* 34: 241–68. [CrossRef]

Sommers, Mitchell, and Joe Barcroft. 2007. An integrated account of the effects of acoustic variability in first and second language acquisition: Evidence from amplitude, fundamental frequency, and speaking rate variability. *Applied Psycholinguistics* 28: 231–49. [CrossRef]

Sommers, Mitchell, Lynne Nygaard, and David Pisoni. 1994. Stimulus variability and spoken word recognition. I. Effects of variability in speaking rate and overall amplitude. *Journal of the Acoustical Society of America* 96: 1314–24. [CrossRef] [PubMed]

Strange, Winifred. 2009. Automatic Selective Perception (ASP) of L1 and L2 speech: A working model. Paper presented at 2nd Acoustical Society of America (ASA) Special Workshop on Speech Cross-Language Speech Perception and Variations in Linguistic Experience, Portland, OR, USA, May 22–23.

Towell, Richard, Roger Hawkins, and Nives Bazergui. 1996. The development of fluency in advanced learners of French. *Applied Linguistics* 17: 84–119. [CrossRef]

Tranel, Bernard. 1987. *The Sounds of French*. Cambridge: Cambridge University Press.

Wade, Travis, Allard Jongman, and Joan Sereno. 2007. Effects of acoustic variability in the perceptual learning of non-native accented speech sounds. *Phonetica* 64: 122–44. [CrossRef] [PubMed]

Walker, Douglas. 1984. *The Sounds of Canadian French*. Ottawa: University of Ottawa Press.

Walker, Douglas. 2001. *French Sound Structure*. Calgary: University of Calgary Press.

White, Lydia. 2003. *Universal Grammar and Second Language Acquisition*. Cambridge: Cambridge University Press.

Zhang, Yang, Patricia Kuhl, Toshiaki Imada, Paul Iverson, John Pruitt, Erica Stevens, Masaki Kawakatsu, Yoh'ichi Tohkura, and Iku Nemotoc. 2009. Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *NeuroImage* 46: 226–40. [CrossRef] [PubMed]

languages

MDPI

Article

Does Typological Proximity Really Matter? Evidence from Mandarin and Brazilian Portuguese-Speaking Learners of Spanish

Alejandro Cuza [1,*], Jian Jiao [1] and Julio César López-Otero [2]

[1] School of Languages and Cultures, Purdue University, West Lafayette, IN 47907, USA; jjiao@purdue.edu
[2] Department of Spanish and Portuguese, Rutgers University, Academic Building, 15 Seminary Place, 5th Flr. College Ave., New Brunswick, NJ 08901, USA; julio.lopez@rutgers.edu
* Correspondence: acuza@purdue.edu

Received: 30 December 2017; Accepted: 12 April 2018; Published: 25 April 2018

Abstract: The present study examines the role of typological proximity in the acquisition of Differential Object Marking (DOM) in Spanish among eighteen (n = 18) Mandarin-speaking second language (L2) learners and sixteen (n = 16) Spanish heritage speakers (HSs) with Brazilian Portuguese (BP) as their dominant language. Specifically, we investigate the extent to which language proximity (languages are members of the same family) plays a role in the complete specification of the relevant features constraining DOM marking in Spanish. Results from an elicited production task and an acceptability judgment task (AJT) showed no support for the typological proximity model (Rothman 2010). There were also no age of onset of acquisition effects, in contrast to what was expected. The post-puberty Mandarin L2 learners outperformed the BP HSs in most of the conditions examined, suggesting a role for language instruction. Results are discussed along the lines of Liceras and Alba de la Fuente's (2015) proposal whereby the locus of transfer is more related to the typological similarity between the languages at the microparametric level than to language proximity itself.

Keywords: Spanish Differential Object Marking; Chinese learners of Spanish; typological proximity model; Spanish heritage speakers

1. Introduction

The present study examines the role of typological proximity in the acquisition of Differential Object Marking (DOM) in Spanish among Mandarin-speaking second language (L2) learners and Spanish heritage speakers (HSs) with Brazilian Portuguese (BP) as their dominant language. Specifically, we investigate the extent to which language proximity/closeness (languages are members of the same family) plays a role in the complete specification of the relevant features constraining DOM marking in Spanish. We extend previous work by investigating (1) the role of typological proximity in language development, as postulated by Rothman's 2010 Typological Primacy Model (TPM), and (2) by analyzing the acquisition process among Mandarin/Spanish bilinguals, a language pair so far underexplored in the L2 acquisition literature (Cuza et al. 2013; Jiao 2017). Furthermore, we aim to inform current debate on the role of age and maturational constraints in the acquisition process by testing Spanish heritage speakers born and raised in Brazil and with BP as their L1 (Giancaspro et al. 2015; Montrul 2008; Montrul et al. 2011).

Animate and specific objects in Spanish are obligatorily marked (i.e., *Dora visitó **a** su padre* ('Dora visited her father')), while inanimate and non-specific objects are not (i.e., *Dora visitó el museo* 'Dora visited the museum')) (Aissen 2003; Bossong 1991; Leonetti 2004). However, BP does not typically mark direct objects, leading to potential cross-linguistic influence effects and patterns of

optionality in L2 learners and HSs of Spanish. Mandarin, in contrast, is argued to mark direct objects in a much more limited/restricted way than Spanish (only in preverbal contexts) (Thompson 1973; Yang and Bergen 2007).

Previous research among Spanish-English bilinguals learning BP as a third language (L3) shows illicit overextension of DOM to BP, leading to non-facilitative transfer (Giancaspro et al. 2015; Montrul et al. 2011). Rothman (2010, 2011) argued that typological proximity determines the source and directionality of transfer in the initial state of L3 acquisition. However, Liceras and Fuente (2015) argued quite convincingly that L2 grammar is not affected by typological proximity per se but rather, by typological similarity. The authors state that languages like Spanish and French are typologically-close but instantiate significant microparametric distinctions that result in cross-linguistic interference and optionality. This proposal is along the lines of Kayne (2005) definition of microparameters (microparametric syntax) that differentiate closely-related languages or dialects. Intralanguage differences or similarities can lead to either negative or facilitative transfer, respectively, regardless of the typological proximity of the languages in contact.

We contribute to previous work by investigating the Spanish grammars of Mandarin L1/Spanish L2 speakers and BP L1/Spanish HSs. These two groups diverge not only in belonging to two typological distant languages but also in regard to (1) the age of onset of acquisition (adulthood vs. birth respectively) and (2) the type of input received. By type of input, we mean formal vs. naturalistic input. The Mandarin speakers were exposed to classroom instruction in Spanish in China as part of their major; the HSs were exposed to Spanish at home from their parents. However, the two groups are similar in that both of them were exposed to reduced Spanish input in either China or Brazil. If language typology and age of onset of acquisition play major roles in bilingual development, we would predict BP speakers to outperform the Mandarin-speaking L2 learners, as BP is typologically closer to Spanish, and the HSs were exposed to Spanish from birth at home. However, it is also possible that there will be no role for typological proximity as Mandarin is typologically similar to Spanish, microparametrically speaking, allowing some sort of DOM.

In what follows (Section 2), we discuss previous work on the syntax and semantics of DOM in Spanish, Mandarin and BP. This is followed by a summary of previous research on the acquisition of DOM in L2 and heritage Spanish, the issue of typological proximity vs. similarity and the research questions and hypotheses (Section 3). Section 4 presents the study, including participant characteristis and tasks. The results are presented in Section 5, followed by the discussion and the conclusions in Section 6.

2. Direct Object Marking in Spanish, Mandarin and Brazilian Portuguese

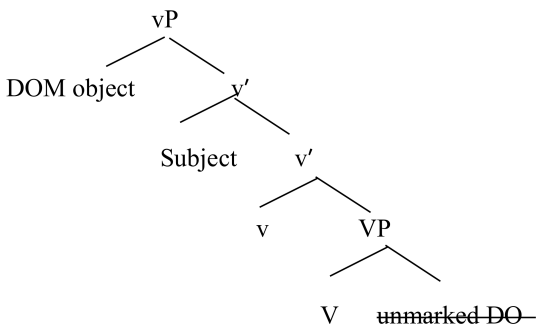
2.1. Object Case Marking in Spanish

Differential Object Marking refers to the overt morphological marking of direct objects to contrast the direct object with the subject (Aissen 2003; Bossong 1991). This cross-linguistic phenomenon is argued to be regulated by the semantic features of animacy and specificity (Aissen 2003). However, there is no consensus on which of the two semantic features is the driving force of DOM in Spanish. Some authors argue for animacy (e.g., García 2007; Leonetti 2004; Rodríguez-Mondoñedo 2007), whereas others argue for specificity (e.g., Fábregas 2013; Laca 2006; Torrego 1998). In Spanish, animate and specific direct objects are overtly marked with the preposition *a* (functioning as the case marker), whereas other direct objects are not usually marked. Additionally, the telicity of the predicate, the agentivity of the subject and the ability to select different complements of the verb are also argued to constrain Spanish DOM (Fábregas 2013; García 2007; Von Heusinger 2008). Currently, there is no consensus on one specific factor nor on the exact constraints that regulate this grammatical phenomenon (Torrego 1998; Zagona 2002).

Syntactically, marked direct objects in Spanish are argued to be at a higher position than unmarked objects (Torrego 1998). Torrego (1998) suggested that in a VP configuration, there are two

specifiers that are occupied, respectively, by the subject and the marked object. The marker *a* has a D-feature, and the marked accusative is raised to the specifier position by the D-feature in v, as shown in (1). Torrego argued that only specific direct objects (DOs) can rise to the higher specifier position occupied by *a*-marked DOs. She also indicated that this prominent position licenses a secondary agentive role to the DOs and denotes the eventive reading of the verb. Thus, the DOs in this position are usually animate and the predicates are frequently eventive. This is represented in (1).

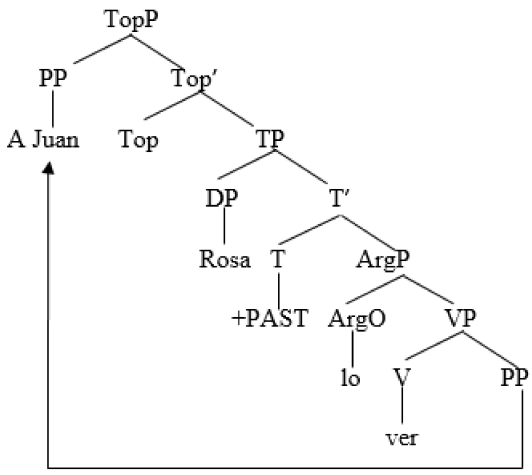
1.



Spanish DOM is closely related with clitic left-dislocated (CLLD) structures.[1] In CLLD sentences, the dislocated object is doubled by a clitic that checks the accusative case and shares the *phi*-features with a dislocated noun phrase (NP), as shown in (2) (Sportiche 1996; Zapata et al. 2004).

2.

A Juan,	Rosa	lo	vio
to John,	Rosa	him	saw
'Rosa saw John.'			



Diesing (1992) argued that a position above v is associated with a specific reading. Leonetti (2004) acknowledged that it is the topicality what forces the specific reading of the proposed NP and that the marker *a* in CLLD structures also serves as a topic marker. Evidence of this is that there are cases

[1] The accusative case marking in Spanish occurs in sentences with topicalized personal pronouns (Pensado 1995; Von Heusinger and Kaiser 2005). Von Heusinger and Kaiser (2005) noticed that Spanish DOM extends from personal pronouns in topicalized sentences to mark all definite DOs, followed by specific indefinite DOs. In a recent study, Iemmolo (2010) examined the relationship between DOM and topicality in Romance languages like Catalan, Italian dialects and two French varieties. The author found that for these languages, DOM is motivated by the markedness of a topicalized noun phrase.

where the marking of the NP in simple sentences is optional, but marking is obligatory in CLLD sentences, as shown in (3).

3. a. A muchos estudiantes, ya los conocía
   to-DOM many students, already them I knew
   'I already knew many students.'

   b. Ya conocía (a) muchos estudiantes
   already I knew (to)-DOM many students
   'I already knew many students.' (Leonetti 2004, p. 86)

To summarize, in Spanish, among other constraints related to the verb and the predicate, a direct object is overtly marked if it is [+animate, +specific]. This phenomenon exists both in simple sentences and syntactically more complex CLLD sentences where the *a*-marking also functions as a topic marker.

2.2. Object Case Marking in Mandarin

Mandarin does not exhibit DOM of the Spanish type in simple sentences, as shown in (4a and 4b). In contrast to Spanish, Mandarin only overtly marks direct objects with the particle *ba* when it is moved to the preverbal position, as shown in (5a–5c) (Thompson 1973).

4. a. Xiaoli kanjian-le Xiaozhang (simple sentence) [+animate, +specific] -DOM*
   Xiaoli see-PERF Xiaozhang
   'Xiao Li saw Xiao Zhang.'

   b. Xiaozhang, Xiaoli kanjian-le (topicalized sentence) [+animate, +specific] -DOM*
   Xiaozhang, Xiaoli see-PERF
   'Xiao Zhang, Xiao Li saw.'

5. a. Ta ba yi zhi laohu chi-le [+animate, −definite] -DOM√
   he ACC a CLASSIFIER tiger eat-PERF
   'He ate a tiger.'

   b. Ta ba yi ge li chi-le [-animate, −definite] -DOM√
   he ACC a CLASSIFIER pear eat-PERF
   'He ate a pear.'

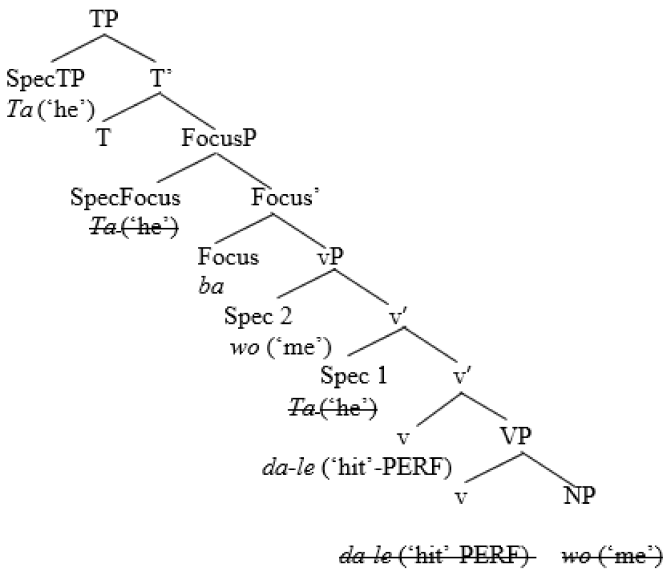
   c. Ta ba wo da-le [+animate, +definite] -DOM√
   he ACC me hit-PERF
   'He hit me.'

The status of the marker *ba* and the derivation of *ba* sentences have attracted a great deal of attention in Chinese linguistics. The *ba* constructions correspond to agentive or causative sentences where the subject has the thematic role of agent or cause (Shi 2010; Weng 2012). The accusative case marker *ba* (Huang 1982; Yang and Bergen 2007) has also been analyzed as a verb (Bender 2000; Hashimoto 1971), or as a preposition (Li 2001; Travis 1984). Yang and Bergen (2007) provided a detailed analysis on the marking of the direct object in *ba* sentences. The authors argued that the constraints diverge from Aissen (2003) proposal in that the marking is also constrained by syntactic position, and that animacy and definiteness features function differently: animate indefinite and definite objects (as in (5a) and (5c) above) and inanimate indefinite objects (as in (5b)) are usually marked. However, the marking of inanimate definite objects is optional. De Swart (2007) argued that the animacy constraint of the object marking serves to distinguish the object from the subject, and that the definiteness constraint tends to license the marking of indefinite DOs.

Furthermore, DOs are marked only in a preverbal position (i.e., in Subject-Object-Verb (SOV) sentences) and there is no marking when the NP is moved to the clause's initial position. Syntactically,

the dislocation of the DO marked by *ba* to the preverbal position is argued to be a focalization structure and the marker *ba* is the head of FocusP (Wang 2014). According to Wang (2014) analysis, the marked DO is moved to a higher position than the complement of the VP. Thus, the example (5c) is generated as illustrated in (6): the marked DO originates as the complement of the verb. The core functional category v is chosen and the VP merges to it form v′. The VP thus moves to v′, and v assigns the accusative case to the DO and the external argument thus becomes the specifier 1, which merges with v′, forming another v′ layer. Subsequently, the EPP (Extended Projection Principle) feature of v requires the marked DO to be specifier 2 of the higher v′, thus it merges into the higher v′ forming a vP. The functional category FocusP extracts the specifier 1 of v′ to the SpecFocus position, while the Focus head *ba*, checks the [Focus] feature with the nearest constituent specifier 2 of vP (i.e., *wo* 'me'). Finally, the EPP feature of TP extracts the SpecFocus to the SpecTP position as the subject of the sentence.

6. Ta ba wo da-le
   he ACC me hit-PERF
   'He hit me.'



This analysis shows that in contrast with Spanish, in Mandarin, direct objects are only marked when focalized. Another peculiarity of Mandarin is that, unlike subject-prominent languages, such as English, Mandarin is a topic-prominent language (Li and Thompson 1976). It exhibits concurrence of multiple topics in one sentence, both clause external and clause internal (Huang 1982), as shown in (7).

7. [Topic$_1$ Zhangsan [Topic$_2$ neixie ren [Topic$_3$ lian yige [ta dou bu renshi]]]]
   Zhangsan those people every one he does not know
   'As for Zhangsan, of those people, he does not know a single one.' (Huang 1982, p. 88)

The different case-marking behavior between Mandarin and Spanish languages raises some learnability issues for Mandarin L2 learners of Spanish. On one hand, the null marking in SVO and OSV sentences in Mandarin and the different constraining conditions associated with the overt marking option suggests difficulty for Mandarin speakers in properly mapping Spanish markers to DOs in simple and CLLD sentences. On the other hand, however, the productivity of topicalization in Mandarin and the overt case marking in preponed DOs for the pragmatic-driven operation of focalization may facilitate the learning of Spanish DOM in CLLD sentences. The topic-prominent nature endows Mandarin speakers with more exposure to topicalized sentences; hence, they are more sensitive to the overt marker *a* as a topic marker (Leonetti 2004).

2.3. Object Case Marking in Brazilian Portuguese

Despite the typological proximity existing between BP and Spanish, BP does not typically instantiate DOM. However, it can differentially mark direct objects in some rare cases (Perini 2002; Ramos 1989; Thomas 1969). First, DOM can be found before *Deus* 'God' and other religious nouns (Perini 2002) and in verbs that express feelings (Ramos 1989), as shown in (8).

8. E então se ajoelhou e adorou a Jesus
   and then REFL knelt down and praised DOM Jesus
   'And then he knelt down and praised Jesus.'

Furthermore, DOM in BP can appear before pronouns. For example, as seen in (9) and (10), DOM can be found "in the written language, before object pronouns when modified or when emphasized" (Perini 2002, p. 444).

9. Ela ama a mim, não a ele
   she loves DOM me, not DOM him
   'She loves me, not him.'

10. Eles me escolheram a mim
    they me chose DOM me
    'They chose me.'

Overall, the uses of Brazilian DOM are very restricted, which renders it very infrequent. Ramos (1989) showed that, in a corpus of Brazilian letters, the accusative *a* decreased across the centuries. While it was used 14.8% of the time in the 16th century, in the 20th century, it was only used 0.7% of the time.

3. Previous Research on the Acquisition of Differential Object Marking

The acquisition of DOM in Spanish has been found challenging among HSs and adult L2 learners with English as their L1 or dominant language (Guijarro-Fuentes and Marinis 2009; Montrul 2004, 2010a). Research has shown significant omission and commission errors, suggesting lack of attainment of the animacy and specificity features constraining DOM use and interpretation due to several factors, including cross-linguistic influence from English (Montrul and Bowles 2009; Montrul and Sánchez-Walker 2013), L2 proficiency effects (Guijarro-Fuentes 2012; Guijarro-Fuentes and Marinis 2007; Montrul et al. 2015; Nediger et al. 2016) and complexity issues (Cuza et al. forthcoming) among other factors.

In regard to L2 learners, Guijarro-Fuentes and Marinis (2007) investigated the grammar of 33 L2 learners of Spanish at different levels of proficiency and found that, despite advanced levels of proficiency, all speakers performed significantly differently to the native speaker controls. The participants showed low omission rates of personal *a* use in animate specific contexts. In a subsequent study, Guijarro-Fuentes (2012) found that advanced L2 learners converged with native speakers regarding the acquisition of animacy features as a constraining force in DOM use and intuition, but not in terms of specificity or verbal semantics. This suggests that some features are easier to acquire than others in interlanguage grammar, and that near native acquisition of DOM is possible at higher levels of proficiency. The author argued that the L2 learners start by learning the lexical features of [-animate] and gradually expand their knowledge to more complex clustering of features. Guijarro-Fuentes account for his findings on the basis of the Feature Reassembly Hypothesis (Lardiere 2008, 2009). According to this hypothesis, grammatical features are assembled differently across languages. If features in the L1 and the L2 are distributed differently at the morphosyntactic level, L2 learners have to reassemble the features of their L1 into their L2. The learnability issue for the L2 learners is to reorganize the distinctly distributed features and map them to the *a*-marking in the appropriate contexts.

Montrul (2010a) examined the acquisition of DOM among 72 L2 learners and 67 heritage speakers of Spanish via an oral narration task and an acceptability judgment task. Results from the narration task

found more omission errors among the L2 learners compared to the heritage speakers (about twice as many), and significant differences in terms of personal *a* expression compared to the controls. However, the advanced heritage speakers showed much fewer omission errors, and their speaking capability was not significantly different from the controls. These results suggest that native-like attainment among heritage speakers is plausible, confirming previous work (Montrul and Bowles 2009), and that there are clear age of onset of acquisition effects that provide an advantage for heritage speakers. However, the L2 learners outperformed the HSs in the acceptability judgment task (AJT), which is expected given their higher level of metalinguistic awareness.

In more recent work, Montrul and Sánchez-Walker (2013) investigated the acquisition of personal *a* among English/Spanish bilingual children and compared their results with adult heritage speakers, a group of long-term immigrants akin to the children's parents and monolingual speakers from Mexico. Results from an oral narration task and a picture description task showed significant omission errors among the child and adult heritage speakers compared to the controls. However, there were no significant differences between the adult heritage speakers and the long-term immigrants. The adult heritage speakers were able to learn the Spanish system despite some optionality. Similar results were found by Montrul et al. (2015), who attributed the difficulties with target DOM use to the complexity of the structure and the multi-functionality of the Spanish marker *a*.

In relation to child heritage speakers, Cuza et al. (forthcoming) examined the production of DOM in simple and CLLD contexts among Spanish/English child bilinguals and their parents via an elicited production task. The results showed significant omission errors in animate specific contexts among the bilingual children and long-term immigrants suggesting cross-linguistic influence effects and structural complexity issues. The authors argued that the animacy features constraining DOM among the bilingual children remain underspecified, crucially in CLLD contexts, due to syntactic complexity issues.

3.1. Typological Proximity vs. Similarity

The acquisition of DOM in BP as L3 among Spanish/English bilinguals was investigated recently by Giancaspro et al. (2015), following Rothman (2010) TPM. The TPM model proposes that the extent to which two or three languages in contact are typologically closer/more similar leads to either facilitative transfer or non-facilitative transfer due to general economy principles at the earlier stages of L3 acquisition. The more structurally similar the L3 is to the L1 or the L2, the more facilitative transfer there will be, as the language learners will draw upon their knowledge of known language(s) when learning a new one. Therefore, in the case of English L1/Spanish L2 bilinguals learning BP as L3, there will be facilitative transfer from Spanish given the structural similarity between Spanish and BP—both Romance languages—at the lexical and grammatical levels (Rothman 2011, 2013). Within this view, Giancaspro et al. (2015) examined the role of transfer from Spanish DOM into L3 BP among three groups of Spanish/English bilinguals (English L1/Spanish L2, Spanish L1/English L2, and English-speaking heritage speakers of Spanish) using a grammaticality judgment task. The participants rejected the use of DOM in the English tokens, showed consistent acceptance of unmarked inanimate objects and rejected unmarked animate objects. However, all three groups accepted BP tokens containing DOM, unlike the BP native speakers. The authors claimed that these findings support Rothman (2010) TPM, arguing that the similar results found across the three groups indicate transfer from Spanish, regardless of its status as L1 or L2, due to its similar typologically to BP. These findings are consistent with Montrul et al. (2011), who tested the acquisition of clitics in L3 BP among two groups of successive Spanish/English bilinguals using an oral production task and a written acceptability judgment task. The results showed transfer effects from Spanish (as L1 and L2), particularly with respect to DOM and clitics. The authors claimed that these results suggest that structural similarity matters in L3 acquisition.

In contrast with the TPM proposal, Liceras and Fuente (2015) proposed that the main contribution to bilingual speakers' optionality and transfer is not necessarily on typological proximity but rather typological similarity. The authors argue that the concept of typological proximity is elusive in the sense

that languages belonging to the same family (i.e., Spanish and French) can have striking differences micro-parametrically (i.e., different instantiations of the null-subject parameter or the presence of clitic doubling in Spanish but not in French), causing significant acquisition difficulties and consistent cross-linguistic influence effects due to divergent options in the L1 and the L2 (Liceras 1986; Han 2013; Selinker 1972). For the authors, there must be a difference, linguistically speaking, between typological proximity and typological similarity. Typological proximity is when the two languages share the same options of macro-parameters. Thus, in this sense, typological proximity can be applied to Spanish and French more so than to Spanish and English, but Spanish and French are not necessarily typologically similar, according to the authors. Typological similarity refers to cases when "a typological or formal universal is equally realized in these two typologically-close languages" (Liceras and Fuente 2015, p. 8). Thus, Spanish and French are typologically close, but not typologically similar in the way that they realize formal universals; typological proximity does not equal necessarily typological similarity. For this reason, it is imperative that when testing the TPM, we test a linguistic domain where both languages behave similarly, despite the existing typological proximity or distance between the two languages. For example, Korean-speaking L2 learners of Spanish might have an easier time learning the distribution of subject pronouns in Spanish than English-speaking learners, since both Korean and Spanish are pro-drop languages. In the case of Spanish and BP, this is evident by the fact that, for example, although both languages have a rich morphological system [+strong], the BP present tense selects only habitual aspectual meanings (as in English), whereas the Spanish present tense can select both habitual and ongoing meanings (Schmitt 1996), resulting in transfer and optionality. Similarly, although Spanish and BP are typologically close, they are not similar in the instantiation of DOM, leading to non-facilitative transfer as far as personal *a* is concerned.

More recent research within a Linguistic Proximity Model (LPM) (Westergaard et al. 2017) also provided evidence against a facilitative role from typologically closer languages in L3 acquisition. The LPM proposes that cross-linguistic influence in L3 acquisition comes from both previously-learned languages, rather than only from the typologically closer one, regardless of the order of acquisition. Thus, it is possible in a trilingual individual to have both facilitative and non-facilitative transfer from one or the two languages already learned (also see Slabakova 2017). Westergaard et al. (2017) examined the directionality of transfer in the acquisition of subject-auxiliary inversion in English L3 among Norwegian-Russian simultaneous bilingual children, and compared the results with age-matched Russian and Norwegian L2 learners of English. Particularly, they investigated whether the sole source of transfer always comes from the typologically more similar language. Results did not confirm a role of typological proximity in the directionality of transfer. The L1 Norwegian children over-accepted ungrammatical sentences in English with a word order that reflected verb movement (V2). In contrast, the bilingual children and the L1 Russian learners showed higher sensitivity to these types of errors. Moreover, the L1 Russian learners outperformed the bilingual children in the target acceptance of grammatical items, which suggests that Norwegian does not play a facilitative role in the acquisition of English despite the typological similarity between the two languages. They concluded, as in the case of Liceras and Fuente (2015), that it is not only an issue of typological proximity but also of structural similarity at the abstract level. The acquisition patterns will be different depending on the linguistic phenomena and not on the typological proximity between the L3 and the previously-acquired languages (Westergaard et al. 2017).

3.2. Research Questions and Hypotheses

If the tenets of the TPM are generalizable to other populations and stages of language acquisition, and are therefore open to further scrutiny, heritage speakers of Spanish with BP as L1 would be expected to outperform Mandarin-speaking learners as far as the acquisition of DOM is concerned, due to the typological proximity between the two Romance languages. However, as discussed by Liceras and Fuente (2015), this might not be necessarily the case, as BP and Spanish do not behave

similarly as far as this particular structure is concerned, leading to optionality and cross-language interference. We postulate the following research questions (RQs):

- RQ1: To what extent do heritage speakers of Spanish with BP as L1 and Mandarin-speaking L2 learners have knowledge of the semantic constraints regulating DOM in Spanish? If difficulties arise, will there be a difference between matrix and CLLD sentences?
- RQ2: Does typological proximity play a role in the acquisition of this domain for the advantage of BP learners, as predicted by the Typological Primacy Model?
- RQ3: In addition to typologically similarity effects, would the HSs benefit from being exposed to Spanish from an early age?

We predict that the HSs speakers will perform more target-like than the L2 learners given that (1) BP is typologically closer to Spanish than Mandarin; and (2) due to the fact that the HSs were exposed to Spanish from an early age at home. However, it is also possible that there will be no facilitative effects from BP, or that the Chinese learners will actually outperform the HSs given the fact that Mandarin is typologically similar to Spanish in allowing DOM in certain contexts. We also expect the Mandarin speakers to perform better with CLLD structures than with simple sentences as Mandarin is a topic-prominent language. Specifically, we hypothesize the following:

Hypothesis 1. *There will be significant differences between both experimental groups and the control participants regarding: (1) the target proportion of personal a use in animate specific contexts with simple and CLLD structures; (2) their acceptance of ungrammatical use of personal a in inanimate contexts in the AJT (errors of commission).*

Hypothesis 2. *The heritage speakers will outperform the L2 learners across both tasks given the typological proximity between Spanish and Brazilian Portuguese.*

Hypothesis 3. *The heritage speakers will outperform the L2 leaners due to age of onset of acquisition effects. The heritage speakers were exposed to Spanish from birth and will therefore outperform post-puberty L2 learners in both production and acceptability judgments.*

Hypothesis 4. *The Chinese learners will perform better with CLLD structures than with simple sentences as Mandarin is a topic-prominent language.*

4. The Experiment

4.1. Participants

Sixteen (n = 16) heritage speakers of Spanish with BP as their L1 (henceforward, HSs) (age range, 21–55; M = 33; SD = 8.6), and eighteen (n = 18) Mandarin-speaking L2 learners of the Spanish language (henceforward, L2 learners) (age range, 19–21; M = 20.1; SD = 0.75) participated in the study as experimental groups. A group of fifteen (n = 15) Spanish monolinguals from Spain served as the control group (age range = 19–48, M = 25.9; SD = 6.5). All of the participants completed a language background questionnaire (Cuza 2013) and an adapted version for Latin American Spanish of the DELE (Diploma de Español como Lengua Extranjera) proficiency test (Cuza et al. 2013) (HSs: M = 45/50; L2: M = 38/50). The DELE exam, the language background questionnaire, and the AJT were completed online via Qualtrics (Qualtrics, LLC, Provo, UT, USA). All of the participants signed a participant consent form and were compensated for their participation.

The HSs (n = 16), except for two participants, were born and raised in São Paulo, Brazil, and had been exposed to both Spanish and BP from birth. Their parents were born in Chile, Argentina, Spain, Paraguay, Bolivia or El Salvador. Their mean score in the DELE test ranged from 39 to 48 points out of 50 (M = 45.21; SD = 2.29). Regarding their pattern of language use, 73% (11/15) reported speaking

"mostly Portuguese" or "only Portuguese", while 20% (3/15) reported speaking "equal Portuguese and Spanish" or "slightly more Portuguese". Only 7% (1/15) of the participants reported speaking "Spanish" or "mostly Spanish" at home. Most of the participants reported using more Portuguese at school, work, and social situations, and 87% (13/15) indicated feeling more comfortable in Portuguese; the other 13% (2/15) indicated feeling comfortable in both Portuguese and Spanish.

The L2 learners (n = 18) were all born and raised in China. They were second year university students (Spanish majors) at a major university in Northeastern China. At the time of the experiment, the L2 learners were finishing their fourth semester. Mandarin was their L1, except for two participants who reported having a local dialect as their L1. In regard to patterns of bilingual language use, Mandarin was reported to be the language used at home and in social situations. Only one participant (6%) reported using "more Spanish" or "mostly Spanish" in social situations. Four participants (22%) reported speaking "slightly more Spanish" or "only Spanish" at work. The rest of the participants reported speaking Mandarin mostly in most contexts. Their mean score in the DELE test ranged from 33 to 44 points out of 50 (M = 38; SD = 3.29). The participants also reported knowledge of English which they had learned at high school. The control participants (n = 15) were all native speakers of Spanish, recruited and tested in Seville, Spain. Six of the speakers had a university degree, while nine had completed technical/professional education.

4.2. Tasks

4.2.1. Elicited Production Task

The elicited production task (EPT) consisted of a Question and Answer task (which was intended to elicit DOM use in simple sentences), and a Sentence Completion task (which was intended to elicit DOM production in CLLD sentences). In the Question and Answer task, the participants were presented with a preamble together with an image followed by a question, as in (11). In the Sentence Completion task, the participant was also presented with preamble and then asked to complete a sentence using the verb provided between parenthesis (12).

11. Question and Answer Task

Preamble:	*Juan está muy feliz hoy.* 'Juan is very happy today.'	
	(here appeared a photo of Santa Claus greeting a child)	
Prompt:	*¿Por qué está tan feliz?* 'Why is he so happy?	(*conocer* 'meet')
Expected response:	*Porque conoció **a** Papá Noel* 'Because he met Santa Claus.'	

12. Sentence Completion Task

Preamble:	*Hoy los periodistas no hicieron entrevistas . . .* 'Today the journalists didn't do any interviews . . . '	
	(here appeared a photo of Pablo being interviewed by some reporters)	
Prompt:	*pero . . .* 'but'	(Pablo)
Expected response:	***a** Pablo sí lo entrevistaron.* 'they did interview Pablo.'	

There were 20 experimental tokens and 22 distractors. The experimental tokens were distributed equally between two conditions: [+animate, +specific] and [−animate, +specific], in which the use of DOM was expected and unexpected, respectively. Answers were coded as 1 if they were considered expected (i.e., the use of DOM preceding animate DPs and DOM omission before inanimate DPs) or as 0 if they were unexpected (i.e., DOM omission before animate DPs and use of DOM before inanimate DPs). Responses containing other structures unrelated to DOM were coded as "other". Participants were interviewed and recorded individually.

4.2.2. Acceptability Judgment Task

The acceptability judgement task had 32 experimental tokens and 44 distractors. All the tokens were introduced by a preamble followed by a question and a response (test token). The participants had to judge the response using a Likert scale ranging from 1 to 5, as represented in (13) below.

13. **Preamble:** *Juan está muy feliz hoy. ¿Por qué está Juan tan feliz?*
'Juan is very happy today. Why is Juan so happy?'

Prompt: *Porque Juan conoció a Papá Noel*
'Because Juan met Santa Claus'

1	2	3	4	5
completely odd	odd	neither good nor bad	good	completely good

If the participants found the token to be odd or slightly odd, they had to say why. The experimental tokens were distributed between eight conditions (*4 tokens each*) in which three variables were considered: animacy (animate or inanimate), sentence type (matrix or CLLD), and grammaticality (grammatical, ungrammatical). Grammaticality was controlled by the use or omission of DOM in relation to the animacy of the object. Table 1 below lists the various conditions tested:

Table 1. Acceptability judgment task (AJT): list of conditions tested.

Sentence Type	Animacy	Grammaticality	Token Sample
Matrix	Animate	Grammatical	*Juan conoció a Papá Noel.* 'Juan met Santa Claus.'
		Ungrammatical	**Ayer Rosa saludó Manuel.* 'Yesterday Rosa greeted Manuel.'
	Inanimate	Grammatical	*Cujo está mordiendo el zapato.* 'Cujo is biting the shoe.'
		Ungrammatical	**Ayer Laura visitó al Museo de Chicago.* 'Yesterday Laura visited the Chicago Museum.'
CLLD	Animate	Grammatical	*Hoy los periodistas no hicieron muchas entrevistas, pero a Pablo sí lo entrevistaron.* 'The reporters did not do many interviews today, but Pablo, they did interview him.'
		Ungrammatical	**El perro no apareció, pero el niño lo encontraron en el parque.* 'They could not find the dog, but the child, they did find him in the park.'
	Inanimate	Grammatical	*Garfield no hizo nada con la cerca, pero la casa sí la pintó.* 'Garfield didn't do anything to the fence, but the house, he did paint it.'
		Ungrammatical	**Pilar siempre lleva sus libros en la mano, pero al diccionario lo lleva en la maleta.* 'Pilar always carries her books in her hand, but the dictionary, she takes it in the suitcase.'

Note: The asterisk symbol means ungrammatical utterance.

In the following section, we present the group and individual results for each task.

5. Results

5.1. Elicited Production Task: Simple Sentences

Results from the production task showed low levels of DOM use in simple sentences among the BP learners (63%), compared to the Mandarin-speaking learners (82%) and the control group (100%). The BP learners also showed higher instances of omission (28%) in animate contexts. With inanimate objects, both experimental groups behaved similarly (81% vs. 83%) (Figure 1).

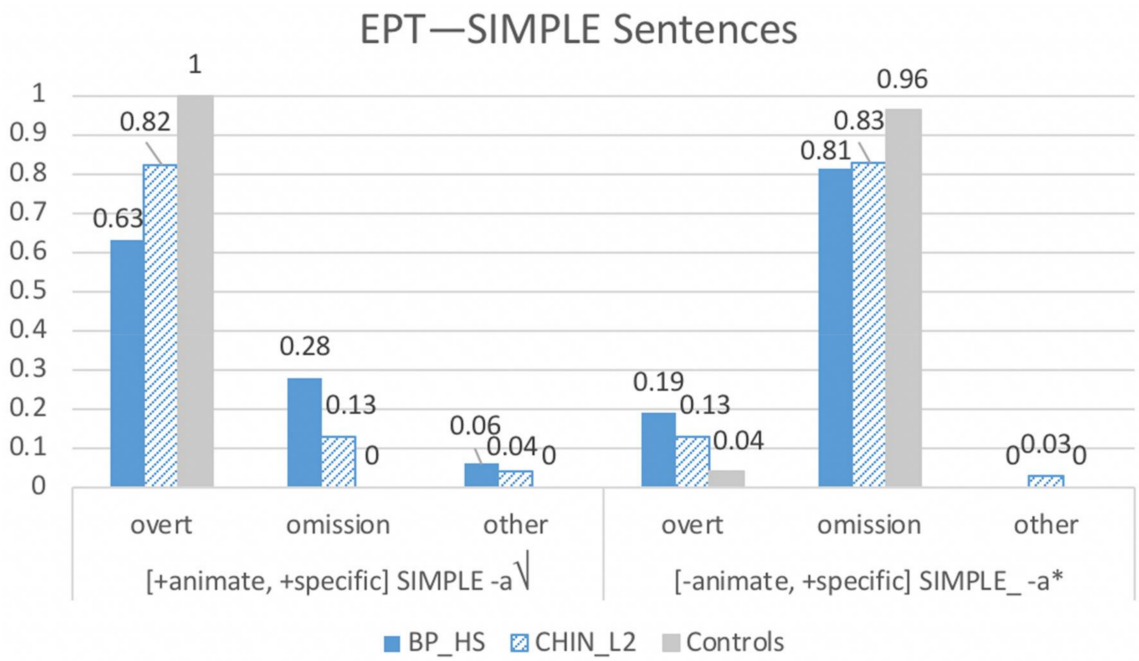


Figure 1. Elicited production task (EPT): proportion of Differential Object Marking (DOM) use, omission and other structures realized per group in simple sentences. BP_HS: Brazilian Portuguese Heritage Speakers; CHIN_L2: Chinese L2 learners.

A one-way ANOVA analysis comparing the means of personal *a* use in animate contexts was conducted to examine if there were any significant differences between groups (between-subject effects) in contexts where DOM was required ([+animate, +specific]). The results showed highly significant differences between groups ($F(2, 49) = 7.87$, $p < 0.001$). To examine the locations of the differences, independent sample *t*-tests were conducted. The results showed significant differences between the HSs and the L2 learners ($t(32) = -2.18$, $p = 0.037$). Both experimental groups behaved significantly different from the controls, as predicted in H1 (HS-control, $t(29) = -4.25$, $p = 0.001$; L2-control, $t(31) = -4.13$, $p = 0.001$). Despite the typological differences existing between BP and Mandarin, and the fact that the HSs were exposed to Spanish from birth, the Chinese L2 learners significantly outperformed the HSs on this condition, disconfirming H2. In regard to [−animate, +specific] contexts (where personal *a* is not required), a one-way ANOVA analysis showed no significant differences between groups ($F(2, 46) = 2.96$, $p = 0.062$).

To examine whether the differences between groups on each condition regarding simple sentences were also observable at the individual level, we conducted an individual analysis. For this analysis, we classified the participants according to their number of personal *a* uses as follows: upper range (4–5/5 instances); mid-range (3/5 instances); low range (1–2/5 instances) and zero production (0/5 instances). With animate objects, almost half of the HSs were in the mid-range (44%), while most of the L2 learners were in the upper range (83%). This confirms the group results, showing an advantage for the L2 learners. In regard to inanimate contexts, most of the HSs and the L2 learners had low range or

zero production. The control participants were all in the upper range for animate contexts and had low range or zero production for inanimate contexts. This is represented in Table 2.

Table 2. EPT: proportion of overt DOM use in simple sentences.

			[+animate]	[−animate]
Group		**#Items**	**#Participants**	**#Participants**
Heritage speakers (HSs) (n = 16)	*Upper range*	4–5	25% (4/16)	6% (1/16)
	Mid-range	3	44% (7/16)	0% (0/16)
	Low range	1–2	13% (2/16)	50% (8/16)
	Zero production	0	19% (3/16)	44% (7/16)
L2 learners (n = 18)	*Upper range*	4–5	83% (15/18)	0% (0/18)
	Mid-range	3	11% (2/18)	5% (1/18)
	Low range	1–2	5% (1/18)	44% (8/18)
	Zero production	0	0% (0/18)	50% (9/18)
Controls (n = 15)	*Upper range*	4–5	100% (15/15)	0% (0/15)
	Mid-range	3	0% (0/15)	0% (0/15)
	Low range	1–2	0% (0/15)	20% (3/15)
	Zero production	0	0% (0/15)	80% (12/15)

EPT: Elicited production task; DOM: Differential Object Marking.

5.2. Elicited Production Task: Clitic Left-Dislocated Sentences

Regarding CLLD structures, the results showed low use of personal *a* marking among the Mandarin-speaking learners (66%), compared to the BP learners (83%) and the controls (80%), with animate specific contexts. With inanimate contexts, where omission was required, again, the Mandarin-speaking learners showed lower levels of omission (49%) compared to the BP learners (68%) and the controls (92%). These results are shown in Figure 2.

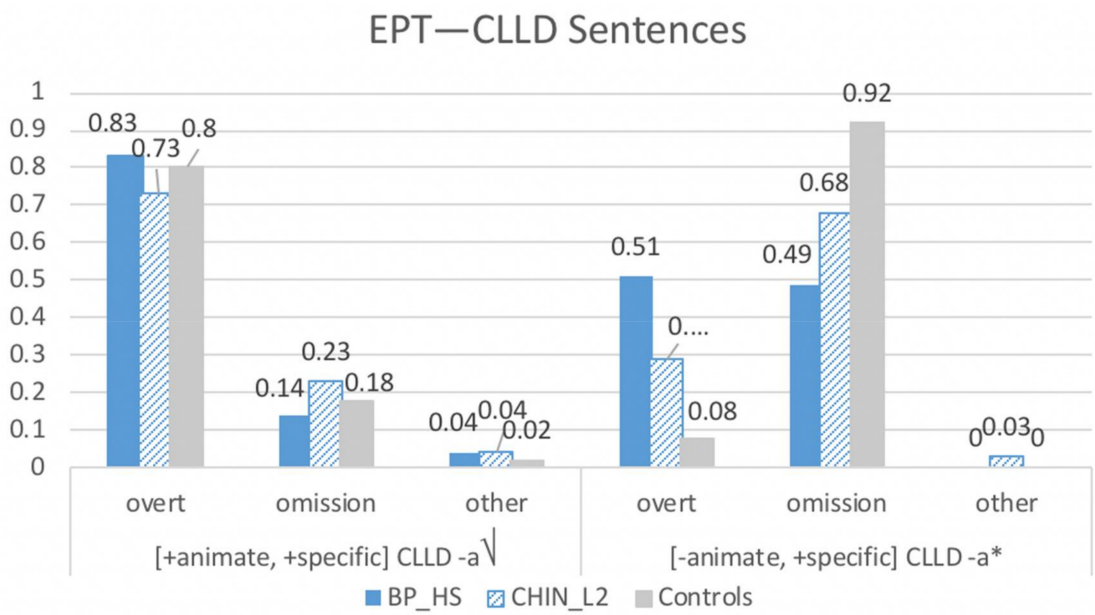


Figure 2. Elicited production task: proportion of DOM use, omission and other structures realized per group in clitic left-dislocated (CLLD) sentences.

A one-way ANOVA analysis comparing the means of personal *a* use in animate CLLD contexts per group showed no significant differences ($F(2, 46) = 0.664$, $p = 0.52$). All of the participants behaved

statistically similarly on this condition. In regard to inanimate CLLD contexts (where personal *a* is not required), a one-way ANOVA analysis revealed significant differences between groups (F(2, 46) = 8.24, $p = 0.001$). An independent sample *t*-test showed no significant differences between the HSs and the L2 learners (t(32) = 1.79, $p = 0.082$). However, there were significant differences between HSs compared to controls (t(29) = 4.45, $p = 0.001$) and L2 learners compared to controls (t(31) = 2.39, $p = 0.023$).

As in the case of the simple sentences, we conducted an individual analysis to examine if the group results were also observable at the individual level. With animate objects, more than half of the HSs and the L2 learners were in the upper range (63% and 67%, respectively). In this specific condition, the control participants did not behave target-like either, with only 47% of them in the upper range. In regard to inanimate contexts, both experimental groups were had low range or zero production. This was particularly so with the L2 learners. The control participants were target-like with this condition. These results are represented in Table 3.

Table 3. EPT: proportion of overt DOM use in CLLD sentences.

			[+animate]	**[−animate]**
Group		**#Items**	**#Participants**	**#Participants**
HS	*Upper range*	4–5	63% (10/16)	38% (6/16)
(*n* = 16)	*Mid-range*	3	25% (4/16)	6% (1/16)
	Low range	1–2	13% (2/16)	38% (6/16)
	Zero production	0	0% (0/16)	19% (3/16)
L2 learners	*Upper range*	4–5	67% (12/18)	22% (4/18)
(*n* = 18)	*Mid-range*	3	11% (2/18)	25% (2/18)
	Low range	1–2	17% (3/18)	22% (4/18)
	Zero production	0	5% (1/18)	44% (8/18)
Controls	*Upper range*	4–5	47% (7/15)	0% (0/15)
(*n* = 15)	*Mid-range*	3	27% (4/15)	0% (0/15)
	Low range	1–2	27% (4/15)	6% (1/15)
	Zero production	0	0% (0/15)	94% (14/15)

CLLD: clitic left-dislocated.

In contrast with simple sentences, there was more variability in regard to CLLD structures at the individual level. There was also an item effect with the item *pero al cajero lo robaron y le quitaron su dinero* 'but they stole money from the cashier'. Seven L2 learners and eight controls omitted the personal *a* in this item. Another item that caused difficulty for most of the participants was *pero el tren lo detuvo superman* 'but the train was stopped by superman', with eight HSs and five L2 learners overextending the personal *a* in this context, together with five control participants.

5.3. Acceptability Judgment Task: Simple Sentences

Results from the AJT with simple sentences showed no difficulties with grammatical sentences. The two experimental groups accepted grammatical items in animate contexts (DOM required) and in inanimate contexts (DOM not allowed). The control participants behaved almost at a ceiling level. In regard to ungrammatical sentences, both experimental groups were undecided in regard to animate and inanimate contexts (commission errors). With inanimate contexts, the HSs tended to accept DOM omission more than the L2 learners. The control participants behaved target-like in all conditions (Figure 3).

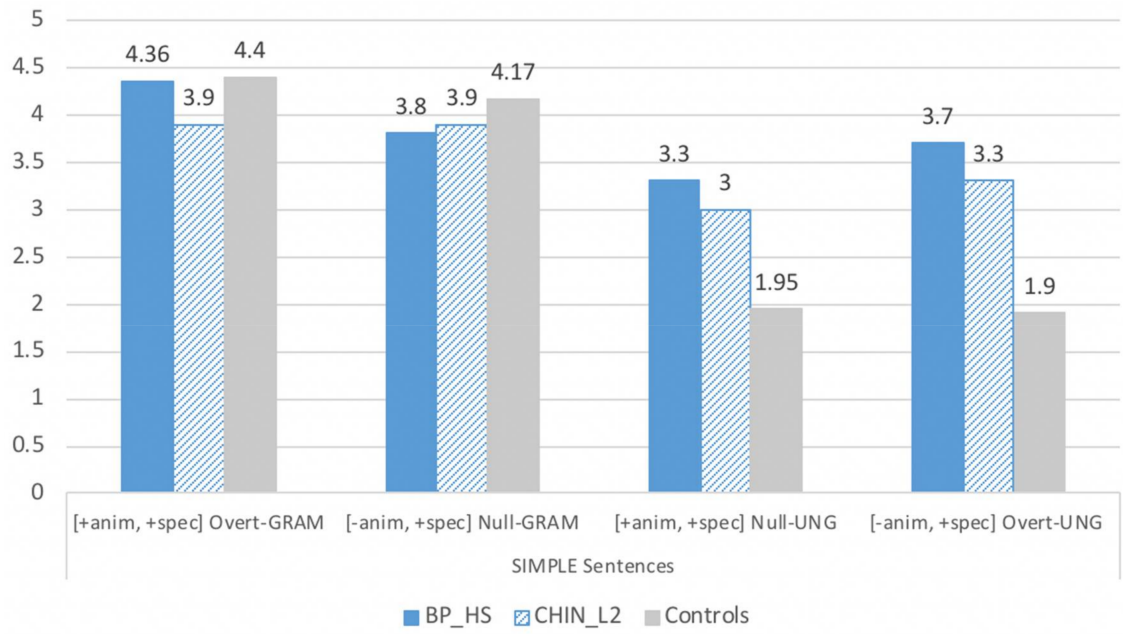


Figure 3. Acceptability judgment task (AJT): proportion of acceptance of grammatical and ungrammatical items in simple sentences per group. Overt: overt NP; Null: null object; GRAM: grammatical sentences; UNG: ungrammatical sentences.

Result from a one-way ANOVA comparing the means of personal *a* use in animate grammatical contexts showed significant differences per group ($F(2, 43) = 4.96$, $p = 0.012$). Independent samples *t*-tests showed significant differences between the HSs and the L2 learners ($t(29) = 2.30$, $p = 0.029$) and between the L2 learners and the controls ($t(30) = -3.197$, $p = 0.003$). In this condition, the L2 learners were outperformed by the HSs and the control group. There was no interaction between the HSs and the controls ($p = 0.777$). The results also showed no significant differences between groups with inanimate grammatical contexts, where the personal *a* was not required ($p = 0.392$). In regard to ungrammatical sentences, an ANOVA analysis showed significant differences between groups with both animate ($F(2, 43) = 8.12$, $p = 0.001$) and inanimate contexts ($F(2, 43) = 8.12$, $p = 0.001$). Although the experimental groups behaved statistically similarly to each other, the HSs and the L2 learners behaved significantly differently from the controls in the two ungrammatical conditions ($p = 0.001$). It is interesting to note that both experimental groups behaved similarly to each other despite their typological differences, and the fact that the HSs were exposed to Spanish from birth.

A closer look at the individual results with ungrammatical sentences disconfirmed the group results with the L2 learners outperforming the HSs. In regard to animate contexts, 4/17 L2 learners accepted the ungrammatical items, and another 3/17 were undecided. However, the majority of the participants (8/17) rejected the sentences in contrast with HSs. The L2 learners also outperformed the HSs in terms of inanimate contexts (71% of the HSs accepted the ungrammatical items). This is represented in Table 4.

Table 4. AJT: proportion of acceptance of ungrammatical items in SIMPLE sentences.

			[+animate]	[−animate]
Group		**#Items**	**#Participants**	**#Participants**
HS	*High Acceptance*	3–4	21% (3/14)	71% (10/14)
(*n* = 14)	*Undecided*	2	36% (5/14)	14% (2/14)
	Low Acceptance	1	36% (5/14)	0% (0/14)
	Zero Acceptance	0	7% (1/14)	14% (2/14)
L2 Learners	*High Acceptance*	3–4	24% (4/17)	29% (5/17)
(*n* = 17)	*Medium Acceptance*	2	18% (3/17)	29% (5/17)
	Low Acceptance	1	12% (2/17)	29% (5/17)
	Zero Acceptance	0	47% (8/17)	12% (2/17)
Controls	*High Acceptance*	3–4	0% (0/15)	0% (0/15)
(*n* = 15)	*Medium Acceptance*	2	20% (3/15)	0% (0/15)
	Low Acceptance	1	20% (3/15)	13% (2/15)
	Zero Acceptance	0	60% (9/15)	87% (13/15)

5.4. Acceptability Judgment Task: Clitic Left-Dislocated Sentences

Regarding CLLD sentences, both the experimental groups and the controls behaved target-like in their acceptance of grammatical items. The Chinese learners were not as target-like as the controls or the BP learners, but they did not differ much from each other. In regard to ungrammatical sentences, both experimental groups were undecided, with a mean acceptance rate of around 3/5. This level of indeterminacy with ungrammatical sentences was more salient with inanimate contexts (*a* overextension/commission errors). The control group behaved as expected (Figure 4).

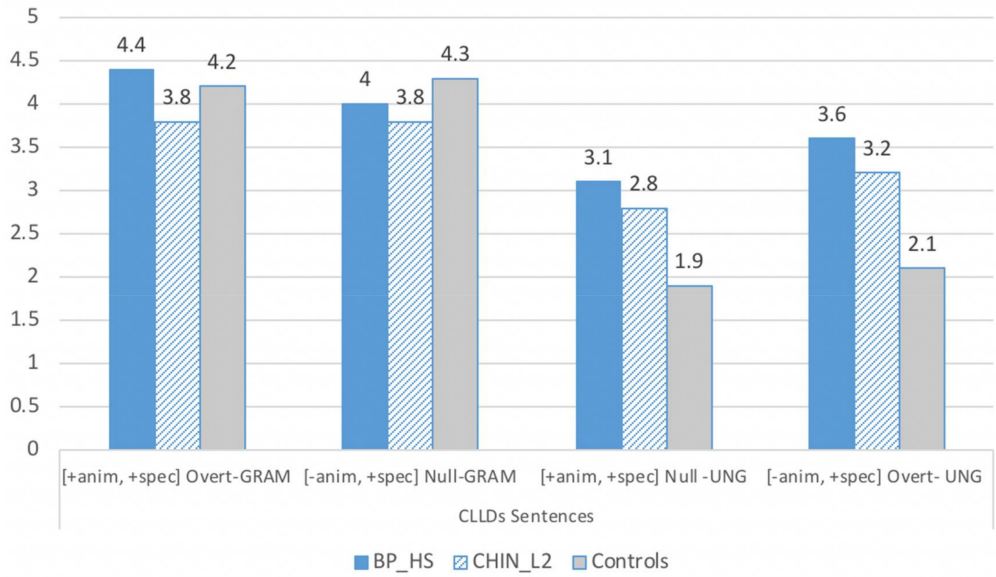


Figure 4. AJT task: proportion of acceptance of grammatical and ungrammatical items in CLLD contexts per group.

A one-way ANOVA comparing the means of personal *a* use in an animate grammatical CLLD context showed significant differences per group (F(2, 43) = 4.39, p = 0.018). However, significant differences occurred only between the L2 learners and the HSs (t(29) = 2.88, p = 0.007). The HSs outperformed the L2 learners There were no significant differences between the L2 learners and the controls (p = 0.099) or between the HSs and the controls (p = 0.244). In regard to the inanimate grammatical context, results also showed significant differences between groups (F(2, 43) = 4.02,

$p = 0.025$); however, significant differences occurred only between the L2 learners and the controls ($t(30) = -3.42$, $p = 0.002$). There were no significant differences between HS-L2 ($p = 0.259$) or between the HSs compared to the controls ($p = 0.173$). In regard to ungrammatical CLLD sentences, the results showed significant differences between groups with animate ($F(2, 42) = 5.61$, $p = 0.007$) and inanimate contexts ($F(2, 43) = 12.25$, $p = 0.001$). In both cases, the L2 learners and the HSs behaved significantly differently from the controls, but there were no significant differences between the two experimental groups. This is represented in Table 5.

Table 5. AJT: independent samples *t*-tests of ungrammatical (UNG) sentences (CLLDs).

Group	Animate_UNG	Inanimate_UNG
HS–L2	$p = 0.617$	$p = 0.272$
L2–control	$p = 0.008$	$p = 0.001$
HS–control	$p = 0.003$	$p = 0.001$

An individual analysis on the ungrammatical sentences confirmed the group results. Both experimental groups accepted a high number of ungrammatical items, crucially in the inanimate condition, compared to the controls. These results are shown in Table 6.

Table 6. AJT: proportion of acceptance of ungrammatical items in CLLD sentences.

			[+animate]	[−animate]
Group		**#Items**	**#Participants**	**#Participants**
HS ($n = 14$)	*High Acceptance*	3–4	29% (4/14)	50% (7/14)
	Undecided	2	21% (3/14)	29% (4/14)
	Low Acceptance	1	21% (3/14)	14% (2/14)
	Zero Acceptance	0	29% (4/14)	7% (1/14)
L2 learners ($n = 17$)	*High Acceptance*	3–4	41% (7/17)	18% (3/17)
	Medium Acceptance	2	6% (1/17)	35% (6/17)
	Low Acceptance	1	24% (4/17)	24% (4/17)
	Zero Acceptance	0	29% (5/17)	24% (4/17)
Controls ($n = 15$)	*High Acceptance*	3–4	0% (0/15)	0% (0/15)
	Medium Acceptance	2	0% (0/15)	13% (2/15)
	Low Acceptance	1	40% (6/15)	33% (5/15)
	Zero Acceptance	0	60% (9/15)	53% (8/15)

6. Discussion and Conclusions

The goal of the present study was to examine two main issues relevant to current research in second language acquisition and heritage language bilingualism. First, we wanted to investigate the extent to which typological proximity plays a significant role in the specification of the relevant features constraining DOM marking in Spanish, a syntax-semantics interface structure so far unexplored among Mandarin-speaking L2 learners of Spanish. Rothman (2010, 2011) has argued that typological proximity determines the directionality of transfer in earlier stages of L3 acquisition. Thus, we wanted to examine whether this is also the case in intermediate stages of L2 acquisition and heritage language development; and if not, whether it is in fact typological similarity rather than proximity that makes a relevant contribution in the degree of optionality vs. final attainment (Liceras and Fuente 2015). To achieve this goal, we tested a group of Mandarin-speaking L2 learners of Spanish and compared them with heritage speakers of Spanish with BP as their dominant language. Mandarin and Spanish are typologically distant languages, compared to BP and Spanish. However, Mandarin is similar to Spanish in that it allows DOM in certain contexts, in contrast with BP where DOM is generally not allowed. If Rothman's hypothesis on typological similarity is applicable to other instances of language development, we would expect HSs of Spanish with BP as their L1 to outperform Chinese-speaking L2

learners. Second, we discussed the issue of age effects as an advantage of the heritage speakers, as they were exposed to Spanish from an early age, in contrast with the Chinese learners. As pointed out by a reviewer, Mandarin-speaking L2 learners and BP-speaking heritage speakers of Spanish are probably not the most ideal group to examine issues relative to age effects as the dominant language of these two groups of participants is a confounding factor. However, the fact that the BP speakers were exposed to Spanish from birth at home and that the Mandarin speakers were exposed to Spanish via formal input is interesting to investigate in light of formal approaches to heritage language theorizing that place a significant weight on age effects as one of the main differences between second and heritage language development in relation to both maturational constraints but also to the quality and quantity of input (Cuza and Frank 2015; Cuza and López-Otero 2016; Montrul 2002, 2010b, 2011; Montrul et al. 2008). If the acquisition of Spanish morphosyntax in heritage language development is constrained by the age of onset of bilingualism, as has been argued quite profusely, we would expect the HSs to do better than the L2 learners, regardless of their L1. However, it is possible that the L2 learners might outperform the HSs in tasks that maximize explicit metalinguistic knowledge, like the acceptability judgment task, rather than in tasks that maximize implicit knowledge, like the elicited production task (Bowles 2011). It is also possible that since Brazilian Portuguese and Spanish are so similar, this might make it harder for the heritage speakers to tease the two languages apart in cases where they do not overlap. So, language proximity might turn out to be more of a curse than a blessing depending on the type of structure under consideration.

The results from the elicited production task showed significant differences between the experimental groups and the control participants in the proportion of personal *a* use in animate contexts with simple sentences, confirming Hypothesis 1. However, the L2 learners significantly outperformed the HSs (82% vs. 63%), disconfirming Hypothesis 2 and Hypothesis 3. Despite Mandarin being typologically distant from Spanish compared to BP, the Chinese learners behaved significantly better than the HSs. The results were also confirmed at the individual level, where 83% of the Mandarin-speaking L2 learners were in the upper range of DOM use compared to only 25% of the HSs. These results suggest that (1) language proximity is not a determinant factor in L2 or heritage language acquisition, as it has been argued to be for initial stages of L3 acquisition (Rothman 2011); and that (2) exposure to Spanish as a heritage language during early childhood is not a categorical contributor to native-like development, at least not as far as DOM is concerned (Montrul and Sánchez-Walker 2013). Although the HSs were exposed to Spanish at home during childhood, the input might not have been sufficient enough to activate this morphosyntactic knowledge. Furthermore, the HSs had not received formal instruction in the classroom, in contrast with the Mandarin speakers.

In regard to CLLD structures, the results showed no differences between groups in relation to their proportion of target DOM use in animate contexts, disconfirming Hypothesis 1. However, the individual results showed a small advantage for the L2 learners compared to the HSs in terms of target use (67% vs. 63%). In addition, as in the case of simple sentences, the L2 learners outperformed the HSs as a group with their target rate of personal *a* omission in inanimate contexts (68% vs. 49%), disconfirming Hypothesis 2 and Hypothesis 3. In this condition—where personal *a* was not required—51% of the HSs overextended the marker (errors of commission) compared to only 29% of the L2 learners. At the individual level, however, both experimental groups behaved similarly in regard to inanimate objects in CLLD contexts (Table 3). In contrast with Rothman (2013) proposal, these results suggest no effect for typological proximity, confirming more recent research (Liceras and Fuente 2015; Westergaard et al. 2017). The results also disconfirm Hypothesis 3, which expected age of onset of acquisition effects to be an advantage for the HSs. Furthermore, the results show more difficulty with CLLD structures in inanimate contexts than in simple sentences among the L2 learners, disconfirming Hypothesis 4. This might be related to the fact that CLLD structures are syntactically more complex than simple sentences (larger number of syntactic derivations) (Cuza et al. forthcoming).

In sum, the results of the EPT go against the typological proximity model. The advantage of the Mandarin-speaking learners over the HSs can be accounted for along the lines of

Liceras and Fuente (2015) proposal, whereby crosslinguistic interference and optionality are more affected by typological similarity than by proximity. This is also along the lines of the *Linguistic Proximity Model* (Westergaard et al. 2017). The fact that Mandarin is structurally similar to Spanish in that it instantiates some sort of DOM might have played a role in the participants' performance, compared to the BP speakers. Furthermore, we found no role for the age of onset of acquisition, in contrast with previous work (Johnson and Newport 1989; Montrul 2008). Although the HSs were exposed to Spanish from birth in Brazil, they seem to have undergone underspecification of the animacy features regulating DOM in Spanish, crucially in CLLDs contexts.

It is possible that the level of Spanish input these HSs received at home was not sufficient to activate certain morphosyntactic features, including DOM. This is not unusual among HSs of Spanish born and raised in contact with other dominant languages and with little access to the heritage language outside the home environment. As discussed earlier, 73% of them indicated speaking "mostly" Portuguese or "only" Portuguese, while only 20% (3/15) reported speaking "equal Spanish and Portuguese". Just one participant reported speaking mostly Spanish at home. These patterns of language use, in addition to the fact that 87% of them indicated feeling more comfortable in Portuguese, account for their low levels of performance. It is interesting though that the mean score in the DELE test was 45 points out 50 (advanced proficiency).

Results from the AJT showed significant differences between the experimental groups and the controls in regard to ungrammatical simple sentences, confirming Hypothesis 1. Both experimental groups behaved similarly at the group level, disconfirming Hypothesis 2. As in the case of the EPT, there were no advantages for the HSs over the L2 learners despite being exposed to Spanish from birth and despite BP being typologically closer to Spanish than Mandarin. This goes against Hypothesis 3 as well, which predicted age of onset of acquisition effects to be an advantage for the HSs.

A closer look at the individual data revealed that the L2 learners, in fact, outperformed the HSs in regard to animate and inanimate ungrammatical contexts (Table 4). In regard to CLLD ungrammatical sentences, the results show significant differences between groups for animate and inanimate sentences, confirming Hypothesis 1. However, the two experimental groups behaved similarly, disconfirming Hypothesis 2 and Hypothesis 3 (Table 5). The individual results, however, showed an advantage for the L2 learners with their rejection of ungrammatical sentences in inanimate contexts. Overall though, both experimental groups behaved relatively similarly for these two conditions. In contrast to what was predicted in Hypothesis 4, the L2 learners treated ungrammatical simple and CLLD sentences in the same way. The fact that Mandarin is a topic-prominent language did not play any role in their judgments of CLLDs sentences.

To conclude, in contrast to what was expected, we found no role for typological proximity in the acquisition of DOM by heritage speakers with BP as their L1, compared to Mandarin-speaking L2 learners. Overall, our results show either an advantage for the Chinese L2 learners or no differences between the two groups. This behavior is more along the lines of Liceras and Fuente (2015) account in that the crucial factor is not necessarily typological proximity in terms of language families, but rather, typological similarity in terms of the types of features that the languages in contact instantiate. The results also confirm the postulates of the LPM which argue that cross-linguistic influence might occur regardless of typological proximity. This prediction is more advantageous in that it is not limited to one specific population of learners or to one specific stage of language development. Although the predictions of the TPM model were made specifically for initial stages of L3 acquisition, one would expect that if in fact there is an advantage in language development due to typological proximity, this would be evident across different stages of acquisition and bilingual populations (L1, L2 or heritage). A theory that is limited to one particular population of language learners or to one particular stage in the acquisition process is difficult to falsify. Future research would benefit from comparing Mandarin-speaking L2 learners of Spanish with BP as L3, and examine other properties related to DOM in CLLD structures, such as clitic doubling, a property available in Spanish but not in Mandarin or BP.

Author Contributions: Alejandro Cuza contributed with the overall vision of the project, writing the introduction, the study, running the statistical analyses and writing the discussion and conclusions. Jian Jiao contributed with the theoretical section and literature review regarding Spanish and Mandarin, and conducted data collection in China. Julio César López-Otero contributed with the theoretical analysis and the literature review regarding Portuguese and data collection in Brazil and Spain.

Acknowledgments: We would like to thank the audience of the 2017 Hispanic Linguistics Symposium at Texas Tech University for their comments and suggestions. We are also very grateful to Peng Zhang from Tianjin Foreign Studies University for her assistance with data collection in China. Finally, we would like to thank the two anonymous reviewers for their thorough comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21: 435–83.

Bender, Emily. 2000. The syntax of mandarin. *Journal of East Asian Linguistics* 9: 105–45. [CrossRef]

Bossong, Georg. 1991. Differential object marking in Romance and beyond. In *New Analyses in Romance Linguistics*. Edited by Dieter Wanner and Douglas A. Kibbee. Philadelphia: John Benjamins, pp. 143–70.

Bowles, Melissa. 2011. Exploring the role of modality: L2-heritage learner interactions in the Spanish language classroom. *The Heritage Language Journal* 8: 30–65.

Cuza, Alejandro. 2013. Cross-linguistic influence at the syntax proper: Interrogative subject-verb inversion in heritage Spanish. *The International Journal of Bilingualism* 17: 71–96. [CrossRef]

Cuza, Alejandro, and Joshua Frank. 2015. On the role of experience and age-related effects: Evidence from the Spanish CP. *Second Language Research* 31: 3–28. [CrossRef]

Cuza, Alejandro, and Julio César López-Otero. 2016. The acquisition of the semantic values of the Spanish present and progressive forms in heritage speakers and L2 learners. *Spanish Journal of Applied Linguistics* 29: 462–86.

Cuza, Alejandro, Lauren Miller, Rocío Pérez-Tattam, and Mariluz Ortiz-Vergara. Forthcoming. Structure complexity effects and vulnerable domains in child heritage Spanish: The case of Spanish personal *a*.

Cuza, Alejandro, Ana Teresa Pérez-Leroux, and Liliana Sánchez. 2013. The role of semantic transfer in clitic-drop among simultaneous and sequential Chinese-Spanish bilinguals. *Studies in Second Language Acquisition* 35: 93–125. [CrossRef]

De Swart, Petrus Jacobus Franciscus. 2007. *Cross-Linguistic Variation in Object Marking*. Utrecht: LOT.

Diesing, Molly. 1992. Bare plural subjects and the derivation of logical representations. *Linguistic Inquiry* 23: 353–80.

Fábregas, Antonio. 2013. Differential object marking in Spanish: State of the art. *Borealis: An International Journal of Hispanic Linguistics* 2: 1–80. [CrossRef]

García, Marco. 2007. Differential Object Marking with inanimate objects. In *Proceedings of the Workshop "Definiteness, Specificity and Animacy in Ibero-Romance Languages"*. Edited by Georg A. Kaiser and Manuel Leonetti. Fachbereich Sprachwissenschaft (Arbeitspapier 122); Konstanz: Universitat Konstanz, pp. 63–84.

Giancaspro, David, Becky Halloran, and Michael Iverson. 2015. Transfer at the initial stages of L3 Brazilian Portuguese: A look at three groups of English/Spanish bilinguals. *Bilingualism: Language and Cognition* 18: 191–207. [CrossRef]

Guijarro-Fuentes, Pedro. 2012. The acquisition of interpretable features in L2 Spanish: Personal *a*. *Bilingualism: Language & Cognition* 15: 701.

Guijarro-Fuentes, Pedro, and Theodoros Marinis. 2007. Acquiring phenomena at the syntax/semantics interface in L2 Spanish: The personal preposition *a*. *Eurosla Yearbook* 7: 67–87. [CrossRef]

Guijarro-Fuentes, Pedro, and Theodoros Marinis. 2009. The Acquisition of the Personal Preposition *a* by Catalan-Spanish and English-Spanish Bilinguals. In *Selected Proceedings of the 11th Hispanic Linguistics Symposium*. Edited by Joseph Collentine, Maryellen Garcia, Barbara Lafford and Francisco Marcos Marin. Somerville: Cascadilla Proceedings Project, pp. 81–92.

Han, Zhao Hong. 2013. Forty years later: Updating the Fossilization Hypothesis. *Language Teaching* 46: 133–71. [CrossRef]

Hashimoto, Anne Y. 1971. Mandarin syntactic structures. *Unicorn* 8: 1–149.

Huang, C. T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Cambridge: Cambridge University.

Iemmolo, Giorgio. 2010. Topicality and differential object marking: Evidence from Romance and beyond. *Studies in Language. International Journal Sponsored by the Foundation "Foundations of Language"* 34: 239–72. [CrossRef]

Jiao, Jian. 2017. The Acquisition of Spanish Differential Marking in Chinese-Speaking L2 Learners. Unpublished Master's thesis, Purdue University, West Lafayette, Indiana.

Johnson, Jacqueline S., and Elissa L. Newport. 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60–99. [CrossRef]

Kayne, Richard. 2005. Some notes on comparative syntax, with special reference to English and French. In *The Oxford Handbook of Comparative Syntax*. Edited by Giglielmo Cinque and Richard Kayne. New York: Oxford University Press, pp. 3–69.

Laca, Brenda. 2006. El objeto directo. In *Sintaxis histórica del español, Parte 1. La frase verbal*. Edited by Concepción Company. México D. F.: Universidad Autónoma de México, pp. 423–79.

Lardiere, Donna. 2008. Feature-assembly in second language acquisition. In *The Role of Formal Features in Second Language Acquisition*. Edited by Juana Liceras, Helmut Zobl and Helen Goodluck. Mahwah: Lawrence Erlbaum, pp. 106–40.

Lardiere, Donna. 2009. Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research* 25: 173–227. [CrossRef]

Leonetti, Manuel. 2004. Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics* 3: 75–114. [CrossRef]

Li, Yen-Hui Audrey. 2001. *The ba-Construction*. Los Angeles: University of Southern California.

Li, Charles N., and Sandra A. Thompson. 1976. Subject and Topic: A New Typology of Language. In *Subject and Topic*. Edited by Charles Li. New York: Academic Press, pp. 459–89.

Liceras, Juana. 1986. *Linguistic Theory and Second Language Acquisition: The Spanish Non-Native Grammar of English Speakers*. Tübingen: Gunter Narr.

Liceras, Juana, and Anahí Alba de la Fuente. 2015. Typological proximity in L2 acquisition. The Spanish non-native grammar of French speakers. In *The Acquisition of Spanish in Understudied Language Pairings*. Edited by Tiffany Tiffany and Silvia Perpiñán. Amsterdam and Philadelphia: John Benjamins, pp. 329–58.

Montrul, Silvina. 2002. Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition* 5: 39–68. [CrossRef]

Montrul, Silvina. 2004. Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition* 7: 125–42. [CrossRef]

Montrul, Silvina. 2008. *Incomplete Acquisition in Bilingualism: Re-Examining the Age Factor*. Amsterdam: John Benjamins.

Montrul, Silvina. 2010a. Dominant language transfer in adult second language learners and heritage speakers. *Second Language Research* 26: 293–327. [CrossRef]

Montrul, Silvina. 2010b. How similar are L2 learners and heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics* 31: 167–207. [CrossRef]

Montrul, Silvina. 2011. Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition* 33: 155–61. [CrossRef]

Montrul, Silvina, and Melissa Bowles. 2009. Back to basics: Incomplete knowledge of Differential Object Marking in Spanish heritage speakers. *Bilingualism: Language and Cognition* 12: 363–83. [CrossRef]

Montrul, Silvina, and Noelia Sánchez-Walker. 2013. Differential object marking in child and adult Spanish heritage speakers. *Language Acquisition* 20: 109–32. [CrossRef]

Montrul, Silvina, Rebecca Foote, and Silvia Perpiñán. 2008. Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning* 58: 503–53. [CrossRef]

Montrul, Silvina, Rejanes Dias, and Hélade Santos. 2011. Clitics and object expression in the L3 Acquisition of Brazilian Portuguese: Structural similarity matters for transfer. *Second Language Research* 27: 21–58. [CrossRef]

Montrul, Silvina, Rakesh Bhatt, and Roxana Girju. 2015. Differential object marking in Spanish, Hindi, and Romanian as heritage languages. *Language* 91: 564–610. [CrossRef]

Nediger, Will, Acrisio Pires, and Pedro Guijarro-Fuentes. 2016. An Experimental Study of the L2 Acquisition of Spanish Differential Object Marking. In *Proceedings of the 13th Generative Approaches to Second Language Acquisition*. Cascadilla Proceedings Project; Somerville: Cascadilla Press, pp. 151–60.

Pensado, Carmen. 1995. El complemento directo preposicional: estado de la cuestión y bibliografía comentada. In *El Complemento Directo Preposicional*. Edited by Carmen Pensado. Madrid: Visor, pp. 11–59.

Perini, Mário. 2002. *Modern Portuguese: A Reference Grammar*. New Haven: Yale University Press.

Ramos, Jânia Martins. 1989. O emprego de preposições no português do Brasil. In *Fotografias Sociolingüísticas*. Edited by Fernando Tarallo. Campinas: Pontes Editores, pp. 83–93.

Rodríguez-Mondoñedo, Miguel. 2007. The Syntax of Objects: Agree and Differential Object Marking. Unpublished Ph.D. dissertation, University of Connecticut, Storrs, CT, USA.

Rothman, Jason. 2010. On the typological economy of syntactic transfer: Word order and relative clause attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics* 48: 245–73. [CrossRef]

Rothman, Jason. 2011. L3 syntactic transfer selectivity and typological determinacy: The Typological Primacy Model. *Second Language Research* 27: 107–28. [CrossRef]

Rothman, Jason. 2013. Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance. In *Romance Languages and Linguistic Theory*. Edited by Sergio Baauw, Frank Dirjkoningen, Luisa Meroni and Manuela Pinto. Amsterdam: John Benjamins, pp. 217–48.

Schmitt, Cristina. 1996. Aspect and the Syntax of Noun Phrases. Unpublished Ph.D. dissertation, University of Maryland, College Park, MD, USA.

Selinker, Larry. 1972. Interlanguage. *IRAL* 10: 209–31. [CrossRef]

Shi, Chun Hong. 2010. Cong ju shi qin zhong kan "bǎ" zi ju ji xiang guan ju shi de yu fa yi yi [A study of the grammatical meanings of the bǎ -construction and other relevant constructions from the perspective of construction group]. *Chinese Teaching in the World* 3: 291–309.

Slabakova, Roumyana. 2017. The scalpel model of third language acquisition. *International Journal of Bilingualism* 21: 651–65. [CrossRef]

Sportiche, Dominique. 1996. Clitic constructions. In *Phrase Structure and the Lexicon*. Edited by Johan Rooryck and Laurie Zaring. Dordrecht: Kluwer, pp. 213–76.

Thomas, Earl W. 1969. *The Syntax of Spoken Brazilian Portuguese*. Nashville: Vanderbilt University Press.

Thompson, Sandra A. 1973. Transitivity and the ba construction in mandarin Chinese. *Journal of Chinese Linguistics* 1: 208–21.

Torrego, Esther. 1998. *The Dependencies of Objects*. Cambridge: MIT Press.

Travis, Lisa DeMena. 1984. *Parameters and Effects of Word Order Variation*. Cambridge: Cambridge University.

Von Heusinger, Klaus. 2008. Verbal semantics and the diachronic development of DOM in Spanish. *Probus* 20: 1–31. [CrossRef]

Von Heusinger, Klaus, and Georg. A. Kaiser. 2005. The evolution of differential object marking in Spanish. Paper presented at the Workshop: Specificity and the Evolution/Emergence of Nominal Determination Systems in Romance, Berlin, Germany, October; pp. 33–69.

Wang, Shi Long. 2014. Yu duan li lun xia "bǎ" zi ju sheng cheng yan jiu [A study of *bǎ–construction under Phase theory]. Modern Chinese: Language Research* 5: 60–4.

Weng, Shan Shan. 2012. Xian dai han yu fei dian xing "bǎ" zi ju yan jiu [A Study of Non-Typical *bǎ*–Construction in Modern Mandarin]. Unpublished Ph.D. dissertation, Peking University, Beijing, China.

Westergaard, Marit, Natalia Mitrofanova, Roksolana Mykhaylyk, and Yulia Rodina. 2017. Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism* 21: 666–82. [CrossRef]

Yang, Ning, and Geertje van Bergen. 2007. Scrambled objects and case marking in Mandarin Chinese. *Lingua* 117: 1617–35. [CrossRef]

Zagona, Karen. 2002. *The Syntax of Spanish*. Cambridge: Cambridge University Press.

Zapata, Gabriela C., Almeida Jacqueline Toribio, and Liliana Sánchez. 2004. The (In) Stability of Linguistic Systems: Evidence from the Performance of Spanish Monolinguals, Heritage Speakers, and Second Language Learners. Working Paper, University of Alberta Edmonton, Edmonton, AB T6G 2R3, Canada. Available online: http://www.rci.rutgers.edu/~lsanchez/publications/LL.pdf (accessed on 30 December 2017).

languages

MDPI

Article

The Interpretation of Pronouns across Spanish-Speaking Populations

Sergio Baauw

Department of Languages, Literature and Communication, Utrecht Institute of Linguistics OTS, Universiteit Utrecht, Heidelberglaan 8, 3584 CS Utrecht, The Netherlands; s.baauw@uu.nl; Tel.: +31-(0)30-253-9689 or +31-(0)61-540-5374

Received: 24 December 2017; Accepted: 11 April 2018; Published: 20 April 2018

Abstract: In this paper, we will present data from both Spanish acquisition and aphasia on the Pronoun Interpretation Problem (PIP), according to which children allow pronouns to be identified with local c-commanding antecedents. Although it has recently been claimed that the PIP is, to a great extent, an experimental artifact, there are good reasons to believe that there is something "real" in the effect. As with many phenomena from acquisition, researchers have tried to explain this development in terms of "learning", or more concretely, in terms of "parameter setting". Children either must set the right local domain for the application of Principle B or they must set a +/− Principle B parameter. However, considering the PIP as an acquisition problem is problematic since it is difficult to see how children can converge on the target grammar without negative evidence. In this paper, we will defend an alternative approach, according to which the PIP is portrayed as the result of interplay between properties of predicates and different kinds of pronouns on the one hand, and language processing factors on the other.

Keywords: binding; co-reference; Principle B; A-Chains; agrammatism; processing; economy; clitic pronouns; Exceptional Case Marking constructions; picture selection; picture verification

1. Introduction

It is a well-known finding that English five-year-old children have problems with the interpretation of personal pronouns, as in (1a), allowing them to be identified with local subjects. Typically, at the same age they do not seem to have the same problems with reflexives (1b) (Chien and Wexler 1990, a.o.):

1. a. Mama bear touched her [her = Mama bear, 50% acceptance]

   b. Mama bear touched herself [herself = Goldilocks, 10% acceptance]

This phenomenon, which we will call the Pronoun Interpretation Problem (PIP), normally disappears when the child reaches the age of seven.

Although it has recently been claimed that the PIP is, to a great extent, an experimental artifact (Conroy et al. 2009), in this paper we will discuss several published studies that show that the PIP is a real effect and that it is the result of the interplay between properties of different kinds of predicates and pronouns on the one hand, and language processing factors on the other. We will argue that this approach fares better in explaining the available data than alternative approaches that considered the PIP to be the result of incorrect or incomplete parameter setting.

We will make our case using data from Spanish and other languages, and from both acquisition and aphasia.

2. The Pronoun Interpretation Problem as a Real Phenomenon

Conroy et al. (2009) argued that the PIP found in child language development is an experimental artifact. They argue that previous experiments using a Truth Value Judgment Task (TVJT), such as Thornton and Wexler (1999), did not really respect the condition of plausible dissent (Crain and Thornton 1998). In the context of a TVJT, this condition states that a "no" response is only felicitous if the situation that would have led to a "yes" response has been considered by the child. This only happens if this situation has been a live option in the storyline of the trial. According to Conroy et al. (2009), this was not the case in many previous studies in which the stories of the trials revolved completely around one individual (e.g., *Bert*) who was the sentence internal antecedent in the test sentences with referential subjects (2a). Interestingly, in the sentences with a quantified subject (2b), this individual was the sentence external antecedent. According to Conroy et al. (2009), the prominence of the sole antecedent (*Bert*) led five-year-old children to accept co-reference in 50% of the cases in (2a) but in only 15% of the cases in (2b). This difference in performance, found in some studies, was argued to be the result of children's observance of Principle B (violated if *every reindeer* binds the pronoun *him*), and their frequent acceptance of *local co-reference* between a pronoun and a local c-commanding referential Determiner Phrase (DP), due to children's difficulties with the pragmatic constraints on local co-reference (Chien and Wexler 1990).

2. a. Bert brushed him

   b. Every reindeer brushed him

Conroy et al. (2009) designed an experiment in which they "repaired" the flaws of earlier studies, carefully respecting plausible dissent, giving all potential antecedents equal prominence. The results showed a radical drop of the percentage of non-adult-like "yes" responses in sentences such as (2a) to 11%, and the disappearance of a significant contrast between (2a) and (2b), which led to 14% non-adult-like "yes" responses. The PIP reappeared in (2a) when the flaws of earlier studies were reintroduced in the design.

Although Conroy et al. (2009) convincingly showed that five-year-old children master Principle B and the principles regulating local (pragmatic) co-reference, this does not necessarily mean that the PIP is exclusively an artifact. As they admit, earlier studies showed cross-linguistic differences with respect to the presence of a PIP in acquisition. Spanish-speaking children, for instance, did not show a PIP in sentences such as (3), despite an experimental setup that did not respect plausible dissent (Baauw 2000):

3. La niña la señala
   'The girl pointed at her.'

Similar results were found for languages such as French (Hamann et al. 1997), Catalan (Escobar and Gavarró 1999) and Italian (McKee 1992; Berger 1999), all languages in which the weak pronoun is a syntactic clitic, an element reported to resist local co-reference (Delfitto 2002).

Also, if the PIP were an experimental artifact—the result of the unbalanced prominence that is given to the sentence internal antecedent—it is not clear why it showed up in older picture verification studies (Chien and Wexler 1990; Koster 1993; Philip and Coopmans 1996; a.o.), in which the potential antecedents were not introduced by any story, and for that reason had similar prominence (or lack of it).

3. The Pronoun Interpretation Problem as Incomplete Learning

Like many phenomena from acquisition, researchers have tried to explain child language development in terms of "learning", or more concretely, in terms of "parameter setting". Binding and co-reference were no exception to it.

One specific proposal along these lines is McKee (1992). Her proposal is based on Manzini and Wexler's (1987) parameter setting proposal. They argue that the local domain in which Principle B is applied, the Governing Category (GC) is parameterized. In some languages, such as English, Dutch or Spanish, it is the minimal XP (i.e., a maximal projection) that contains the pronoun, an accessible subject, and its governor (Chomsky 1981). In other languages the GC may be different. The learning task of children is to set this parameter at the language-specific value. McKee (1992) argues that English-speaking children initially assume the GC to be the Verb phrase (VP) (4b). This means that for them the pronoun must be free inside the VP but can be freely bound by the subject in the specifier of the Inflection Phrase ([Spec, IP]), leading to the acceptance of the reflexive interpretation of (1). McKee (1992) also accounts for the absence of a PIP in the acquisition of Italian (and by extension, in Spanish). Since in Italian the weak pronoun is a syntactic clitic, it is not inside the VP, but in a position adjoined to I (4a). This extends automatically the GC to the IP, making binding by the subject in [Spec, IP] impossible without violating Principle B, which is considered to be innate and inviolable.

4. a. [IP Lo gnomo [I' lo [VP lava]]]
   the gnome him washes

   b. [IP The gnome [VP washed him]]

A different approach is taken by Elbourne (2003; cited by Elbourne 2005). He proposes that Principle B itself may be subject to parameterization. In some languages it is active, such as in Modern English, and in others it is inactive, such as in Middle English. Elbourne (2003) assumes that English-speaking children start out with a grammar in which Principle B is inactive, as in Middle English. A slightly different approach is taken by Fodor (1992; cited by Elbourne 2005). She argues that children may analyze pronouns such as *him* initially as ambiguous between a pronoun and a weak reflexive, such as Dutch *zich* in (5). When they analyze *him* as *zich*, the pronoun can be bound without violating Principle B.

5. a. Jan waste zich
   John washed REFL
   'John washed (himself).'

Chien and Wexler (1990) argued that children obey (innate) Principle B, rejecting local binding of the pronoun, but that they often accept local co-reference in contexts in which adults would reject such an interpretation. This explained why English-speaking five-year-olds accepted the reflexive interpretation much more often in (6a) than in (6b).

6. a. Mama Bear is touching her [50% target-like performance]

   b. Every bear is touching her [85% target-like performance]

In (6a) the pronoun can be identified with the local subject through variable binding, which would lead to a violation of Principle B, but also through local (pragmatic) co-reference. In the adult language, local co-reference is limited to special contexts, such as in (7) (Heim and Kratzer 1998).

7. a. Everybody hates John. Mary hates him, Peter hates him, even John hates him.

   b. Do you know what Mary and John have in common? Mary admires him and John admires him too.

Children, on the other hand, are argued to often accept local co-reference also outside these special contexts. According to Chien and Wexler (1990), this is because they still must learn the pragmatic constraints on local co-reference, constraints that Chien and Wexler (1990) call Principle P.

In (6b), on the other hand, the pronoun can only be identified with the local subject through variable binding, since quantified DPs cannot co-refer. However, since binding violates Principle B, and Principle B is assumed to be respected by children, a reflexive reading of (6b) cannot be construed.

Crucially, the assumption is that pragmatic principles, such as Principle P, must be acquired by children, unlike syntactic principles, such as Principle B, which are assumed to be universal and innate.

Most accounts that assume some sort of incomplete acquisition, either as incorrect parameter setting (McKee 1992; Elbourne 2003) or as incomplete pragmatic acquisition (Chien and Wexler 1990), face the problem of having to explain the optionality of children's acceptance of local co-reference. If a pragmatic principle is not in place or a parameter is incorrectly set, the child is expected to perform consistently in accordance with this stage of language-specific competence, accepting the reflexive interpretation of (1a) across the board, and not roughly 50% of the time, as is found in many studies.

Another problematic aspect of this approach is the fact that the PIP is only found in comprehension, not in production. Bloom et al. (1994) have shown that the PIP is absent in spontaneous production of English-speaking children. Experimental language production studies show no evidence of a PIP either. An elicited production study by De Villiers et al. (2005) showed that English-speaking children with a mean age of 6;2 produced pronouns, in conditions eliciting reflexives, only between 0% and 14% of the time (depending on the kind of sentence).

Finally, incomplete acquisition accounts face a learnability problem. A child that allows both local and non-local reference of the pronoun at some point must retreat to a system in which only non-local reference is possible. The question is how the child can do this in the absence of negative evidence. Some authors have argued that there may be triggers in the input that signal the target parameter setting (McKee 1992). However, the question remains why children are not sensitive to these triggers from the beginning.

4. The Pronoun Interpretation Problem as a Processing Problem

In this section, we will defend an alternative approach to the PIP, according to which it is the result of the interplay between properties of predicates and different kinds of pronouns, on the one hand, and language processing factors, on the other (Avrutin 1994, 2004; Avrutin and Baauw 2013; Baauw et al. 2011).

4.1. The Pronoun Interpretation Problem Cross-Linguistically

We will adopt Grodzinsky and Reinhart's (1993) account of the PIP in sentences such as (6). They propose that children obey Principle B, which explains why they perform highly target-like on (6b). The reason they often accept the reflexive interpretation of (6a) is because children often accept local co-reference in contexts in which adults with no language pathologies reject this reading. This is due to children's difficulties with the application of a discourse-syntax interface constraint on local co-reference, called Rule I. Rule I is invoked when a local co-reference interpretation is an option, such as in (6a). Children have difficulties with the application of this rule because it requires the comparison of two possible construals of the same sentence—the binding construal and the co-reference construal—to rule out the co-reference construal if it turns out to be indistinguishable from the binding construal. Grodzinsky and Reinhart (1993) argue that this reference set computation requires too many resources from the child's immature brain, causing Rule I to "break down". As a result, children start to guess to establish the reference of the pronoun, which leads them to accept co-reference in (6a) roughly 50% of the time.

However, there is considerable cross-linguistic variation with respect to the presence of a PIP in acquisition. The phenomenon has been found for English (Jakubowicz 1984; Chien and Wexler 1990), Dutch (Koster 1993; Philip and Coopmans 1996), Icelandic (Sigurjónsdóttir 1992) and Russian (Avrutin and Wexler 1992; Avrutin 1994), but not in many other languages, such as French (Hamann et al. 1997), Italian (McKee 1992), Catalan (Escobar and Gavarró 1999), Spanish (Baauw et al. 1997; Baauw 2000), Greek (Varlokosta 2000), Norwegian (Hestvik and Philip 2000) and Hungarian (Margócsy 2000). For Spanish, several picture verification task studies showed that five-year-old children accept the reflexive interpretation of (8) about 10% of the time (Baauw 2000; Baauw and Cuetos 2003; Alija and Baauw 2005).

8. La niña la señala [la = la niña, 10% acceptance]
   the girl her points-at
   'The girl pointed at her.'

In Baauw (2000) it was argued that the factor that distinguishes languages with a PIP in their acquisition and those that do not show a PIP is the status of their weak pronouns. Concretely, no PIP is expected in the acquisition of languages in which the weak pronoun is a syntactic clitic (Baauw 2000). Baauw (2000) adopts Chien and Wexler's (1990) and Grodzinsky and Reinhart's (1993) proposal that the PIP is the result of children's non-adult-like acceptance of local co-reference. In languages with syntactic clitics, such as Spanish, the PIP does not show up because syntactic clitics resist local (pragmatic) co-reference, which prevents Rule I from being invoked. Baauw (2000) adopts Sportiche's (1992) view that clitics are VP external functional projections that host in their specifier a moved null object. The trace that is left by this movement is argued to convert the structure c-commanded by the specifier of the clitic phrase into a predicate (Heim and Kratzer 1998; Neeleman and Weerman 1999; Delfitto 2002). Subsequent application of Quantifier Raising of the subject DP and identification of the subject trace with the object trace under agreement (i.e., feature sharing) will lead to a clear violation of Principle B (9).[1]

9. La niña la señala
   the girl her points-at
   [λx (x señala x)] (la niña) → violation of Principle B

Assuming that children respect Principle B, Spanish-speaking children will reject an interpretation in which the object is identified with the subject in (9). In sum, Spanish-speaking children's target-like performance on the interpretation of sentences such as (9) is similar to English-speaking children's target-like interpretation on (6b); the unavailability of a co-reference construal prevents children from showing a PIP.

The absence of a PIP in Spanish-speaking children's interpretation of sentences such as (8) shows that they respect Principle B, but also that they are aware of the syntactic clitic properties of Spanish weak pronouns. This is no surprise, since there is ample evidence that clitics are acquired early in acquisition. Research shows that Spanish- and Italian-speaking children learn very early that weak pronouns are syntactic clitics, and they hardly make errors with respect to the position of these elements (Guasti 1994; Ezeizabarrena 1996; Schaeffer 1997; Lyczkowsky 1999). Since the referential properties of syntactic clitics depend on their syntactic position (Baauw 2000; Di Sciullo and Agüero-Bautista 2008), rejection of local co-reference is expected as soon as they acquire the syntactic properties of clitics.

4.2. What about Strong Pronouns in Romance?

If the absence of a PIP in Romance (and some other) languages is due to the clitic status of their weak pronouns, then it is predicted to show up in structures with strong pronouns. Baauw (2000) carried out two experiments to discover a PIP in Spanish sentences containing strong pronouns.

In the first picture verification experiment, 32 Spanish-speaking children between 4;4 and 7;2 years of age (mean age 5;11) were tested on the reflexive interpretation of sentences such as (10a), with a

---

[1] Note that the fact that clitic pronouns resist local (pragmatic) co-reference does not mean that they cannot be used for reference to non-c-commanding or sentence external DPs. In fact, sentences such as (i) clearly show that they can:

i. Juan saludó a *María*. Pedro también *la* saludó. [la = María]
   Juan greeted María. Pedro also her greeted

Baauw (2000) argues that in contexts in which disjoint reference of the object and the subject is intended, an empty topic is generated, which binds the object trace created by movement of the null object. This null topic can be identified in discourse with other DPs. For further elaboration of this proposal, see Delfitto (2002).

clitic direct object pronoun, (10b), with a strong direct object pronoun, and (10c), in which the strong pronoun is the complement of a preposition.[2]

10.	a.	El niño	le	dibujó	
		the boy	$\text{him}_{\text{clitic}}$	drew	
	b.	El niño	le	dibujó	a él
		the boy		drew	ACC-$\text{him}_{\text{strong}}$
	c.	El niño	miró	hacia	él
		the boy	looked	at	$\text{him}_{\text{strong}}$

The results showed no evidence of a PIP in sentences containing strong direct object pronouns. The reflexive interpretation was rejected 91% of the time in (10a) and 83% of the time in (10b). This difference was not significant. Baauw (2000) argues that the absence of a PIP in (10b) may be due to the obligatory presence of clitic doubling in Spanish structures involving strong direct object pronouns. Baauw (2000) proposes, following Aoun's (1999) analysis of doubled elements in Lebanese Arabic, that strong pronouns may be right-dislocated elements, which means that the object position in these constructions is occupied by the trace of a moved null-object, just as in the case of other pronominal clitic constructions (see also Fernández-Soriano (1989) for Spanish).

Although for Spanish a PIP could not be found in constructions containing strong direct object pronouns, in Italian Berger (1999) did find such an effect. She tested Italian-speaking children on the interpretation of sentences with direct object clitics, such as (11a), and sentences containing strong direct object pronouns, such as (11b).

11.	a.	Il ragazzo	lo	sta indicando
		the boy	$\text{him}_{\text{clitic}}$	is pointing-at
	b.	Il ragazzo	sta indicando	lui
		the boy	is pointing-at	$\text{him}_{\text{strong}}$

The results showed that children rejected a reflexive interpretation of (11a) but accepted it 61% of the time in (11b). This suggests that in constructions without clitic doubling, a PIP shows up with strong pronouns.[3]

However, Baauw (2000) also tested Spanish constructions in which strong pronouns are not doubled by clitics, namely constructions in which the strong pronoun is the complement of a preposition, as in (10c). The results show that, in this construction, the reflexive interpretation was rejected only 43% of the time by Spanish-speaking children. However, adult control participants showed a highly similar 51% rejection of (10c). This shows that the child performance cannot been interpreted as a PIP. In fact, in Spanish, strong pronouns can be identified with the subject of the verb in a wide range of prepositional constructions (13), whereas in languages such as English this possibility is limited to locative prepositions (12).

---

[2] Since the testing took place in Madrid, and Madrid is a *leísta* area, *le* was used instead of *lo* as the masculine direct object form (both are translated as *him* in English).

[3] Some caution should be taken in interpreting these results. Berger (1999) also found that Italian children rejected the extra-sentential interpretation of the strong pronoun in (11b) 19% of the time. This may be due to the fact that strong pronouns are "marked" and usually used in contrastive contexts in which the pronoun tends to be stressed. As McDaniel and Maxfield (1992) notice, pronominal stress facilitates local, or "pragmatic", coreference. Researchers may try to pronounce the strong pronoun without stress, but this may sound unnatural, and could affect the results in unpredictable ways.

12. a. Mary$_i$ put a chair behind her$_i$

    b. *Mary$_i$ only thought about her$_i$

13. a.

María$_i$	puso	una silla	detrás de ella$_i$
Mary	put	a chair	behind of her

    b.

María$_i$	solo	pensaba	en ella$_i$
Mary	only	thought	about her

Apparently, locative Prepositional Phrases (PPs) can be analyzed as separate binding domains in both Spanish and English, but in Spanish some non-locative PPs should be considered separate binding domains too, allowing pronouns to be bound by the subject of the sentence. However, the possibility of non-locative PPs to be separate binding domains is subject to particular linguistic constraints; with some prepositions, a reflexive interpretation of the verb is more marginal than with others. Sentences containing these prepositions might allow us to see whether Spanish-speaking children show a PIP in sentences containing strong pronouns. However, this is only the case if we can somehow exclude the possibility that (more frequent) acceptance of a reflexive interpretation is due to children's possibly more "relaxed" constraints on the ability of PPs to be separate binding domains.

In order to control for this possibility, Baauw (2000) carried out a second picture verification experiment in which 30 Spanish-speaking children were tested (age range 4;8–7;3, mean age 6;0) on the interpretation of sentences containing strong pronouns as the complement of non-locative prepositions that, according to native speakers, do not easily allow for a reflexive interpretation. These were presented either with a referential subject DP (14a), or with a quantified subject DP (14b).[4]

14. a.

la niña	sueña	con ella
the girl	dreams	about her

    b.

cada niña	sueña	con ella
every girl	dreams	about her

The reason sentences with quantified subjects were included is that quantified DPs cannot be identified with pronouns through co-reference; they can only bind them. This allows us to determine whether the acceptance of a reflexive reading is due to an over-acceptance of "pragmatic" co-reference by children, just like what happens in Dutch- and English-speaking children interpreting sentences such as (1a), or whether this acceptance is due to their across-the-board acceptance of PPs as separate binding domains. In that case, a reflexive interpretation could be a case of (non-local) binding instead of co-reference. The results showed that Spanish-speaking children rejected the reflexive interpretation of (14a) 42% of the time, and of (14b) 53% of the time. Adults rejected the reflexive interpretation of (14a) 72% of the time and of (14b) 89% of the time. An ANOVA showed that there was a main effect of sentence type (quantified DP/referential DP) and of group (children/adults). However, there was no interaction effect. This means that children rejected a reflexive interpretation less often than adults did, but that the extent to which they rejected this reading in sentences with quantified subject more often was not greater than in adults.[5] This indicates that Spanish-speaking children's acceptance of the reflexive interpretation of (14a) (and of 14b) is due to children's less strict constraints on the

---

[4] The experiment also tested strong pronouns in locative PPs. These data will not be discussed in this article. See Baauw (2000) for details.

[5] Some native speakers resist binding of a strong pronoun by a quantified DP, also when the PP is allowed to be a separate binding domain, as in (ii) and (iii) (see Menuzzi 1996):

ii.

%nadie$_i$	vio	una culebra	cerca de él$_i$
nobody	saw	a snake	next to him

establishment of PPs as separate binding domains, and not to problems with the application of the pragmatic constraints on local (pragmatic) co-reference (Rule I), as was argued to be the case in Dutch- and English-speaking children. Probably, the availability of a binding construal that does not violate Principle B prevents Spanish-speaking children from establishing a co-reference relation between the strong pronoun and the subject.

4.3. Structural Variation

Philip and Coopmans (1996) noted that the intensity of the PIP depended on structural factors. They noticed that Dutch-speaking children show a much stronger PIP in Exceptional Case Marking (ECM) sentences, such as (15a) than in simple transitive sentences, such as (15b). Even 7-year-old rejected the reflexive interpretation of (15a) only 16% of the time, while they rejected (15b) 55% of the time.

15.	a.	het jongetje	ziet	hem	dansen	[hem = het jongetje, ±80% acceptance]
		the boy	sees	him	dance	
	b.	het jongetje	aait	hem		[hem = het jongetje, ±50% acceptance]
		the boy	pets	him		

A comparable difference in performance between ECM sentences and simple transitive sentences was also found in Spanish 5-year-olds (16) (Baauw et al. 1997; Baauw 2000; Baauw and Cuetos 2003), and in the acquisition of several other languages, such as French (Hamann et al. 1997), Italian (McKee 1992), Catalan (Escobar and Gavarró 1999), Greek (Varlokosta 2000), Norwegian (Hestvik and Philip 2000; Philip, pc) and Hungarian (Margócsy 2000). Below we provide the results of Baauw and Cuetos (2003), which used a picture verification task (disguised as a guessing game), with 38 five-year-old European Spanish-speaking children. The results are from the trials eliciting "no" responses in adults.

16.	a.	la niña	la	ve	bailar	[la = la niña, 40% acceptance]
		the girl	her	sees	dance	
		'The girl sees her dance.'				
	b.	la niña	la	señala		[la = la niña, 10% acceptance]
		the girl	her	points at		
		'The girl is pointing at her.'				
	c.	cada niña	la	señala		[la = cada niña, 11% acceptance]
		every girl	her	pointed at		
		'Every girl is pointing at her.'				

This difference in performance is not easy to explain in a framework that adopts standard Binding Principle B (Chomsky 1981), since this principle treats both sentence types alike; binding of the embedded pronominal ECM subject by the main clause subject would lead to a Principle B violation. In addition, the application and breakdown of Rule I in case the child tries to establish a co-reference relation should lead to a 50% adult-like performance rate in both kinds of sentences in children acquiring languages such as Dutch, which has no Romance-like clitic pronouns, and around 90% adult-like performance in children acquiring languages with syntactic clitics, such as Spanish.

---

iii.	%todo el mundo$_i$una culebra	hablaba	de él$_i$
	everybody	talked	about him

This explains why even adults rejected the reflexive interpretation more often in (14b) than in (14a).

However, the results fit easily in a framework that assumes an alternative formulation of the binding constraints. Reinhart and Reuland (1993) propose that the constraints on local binding are captured by two independent principles: a constraint on the creation of reflexive predicates, which they call Principle B, and a constraint on A-Chains, which they call the A-Chain Condition:

17. Principle B: A reflexively interpreted predicate must be reflexive-marked.

18. A-Chain Condition: A maximal A-Chain ($a_1 \ldots a_n$) contains exactly one link –a_1– that is both [+R] and case-marked.

Reformulated Principle B, as in (17), can be seen as an interface filter on valency reduction (Reuland 2001). It limits the possibility of verbs to allow a reflexive interpretation ($\lambda x\ \lambda y\ (xRy) \rightarrow \lambda x\ (xRx)$) to those verbs in which reflexivity is somehow marked, either by means of a reflexive-marking morpheme or as a result of particular semantic properties. This principle can account for the ungrammaticality of the interpretation in which the object pronoun is bound by the subject of the same verb, as in (15b) or (16b). However, it cannot account for the ungrammaticality of the reflexive interpretation of (15a) and (16a). The reason is that in ECM constructions the subject of the main clause (*het jongetje/la niña* ('the boy'/'the girl')) and the pronominal subject of the embedded clause (*hem/la* ('him'/'her')) are not part of the same predicate; while *het jongetje/la niña* receives its thematic role from the main clause verb *ziet/ve* ('sees'), the pronoun *hem/la* receives its role from the embedded verb *dansen/bailar* ('to dance'). This means that when the main clause subject binds the subject of the embedded sentence, no reflexive predicate is created, rendering Principle B irrelevant. However, adult speakers reject such an interpretation. According to Reinhart and Reuland (1993), this is the result of the A-Chain Condition, as in (18). They argue that local binding relations lead to the creation of an A-Chain between the binder and the bindee. The A-Chain Condition states that the tail of this chain must be [−R], i.e., a non-referential element. Simple reflexive pronouns such as Dutch *zich* are [−R], but DPs and third person pronouns, such as *hem* 'him' or *la* 'her' are [+R]. As a result, binding of *hem* 'him' by *het jongetje* 'the boy' will lead to a violation of the A-Chain Condition, since both the head and the tail of the A-Chain are [+R]:

19. <u>het jongetje</u> ziet [<u>hem</u> dansen]
    [+R] [+R]
    'the boy sees [him dance]'

Therefore, it turns out that Dutch-speaking children show a stronger PIP in constructions where only the A-Chain Condition prevents an acceptable reflexive interpretation, and this turns out to be the only context in which Spanish-speaking children show a PIP. Philip and Coopmans (1996) and Baauw (2000) propose that this is due to children's ability to, optionally, interpret third-person pronouns as [−R]. When children interpret third-person pronouns as [−R], the establishment of an A-Chain between the main clause subject and the embedded clause subject will not lead to a violation of the A-Chain Condition. Philip and Coopmans (1996) proposed that children's ability to treat third-person pronouns as [−R] is due to incomplete feature acquisition. They follow Reinhart and Reuland's (1993) proposal that [+R] is related to full specification for features such as number and structural case. Ambiguities in the Dutch pronominal system would cause a delay in case feature specification. Baauw (2000) proposed that incomplete acquisition of the feature [number] might be the cause of both Dutch- and Spanish-speaking children's ability to treat third-person pronouns as [−R]. However, as we argued before, incomplete acquisition proposals face the problem of having to explain the optionality that the child displays with respect to the interpretation of pronouns.

Baauw et al. (2011) and Avrutin (2006) therefore propose an alternative account, which is based on computational economy. Following Reuland (2001) and Avrutin (2004), they propose that the ungrammaticality of the reflexive reading of (15a) and (16a) is due to the violation of an economy condition on referential dependencies. Reuland (2001) argues that referential dependencies can be established in discourse (co-reference), in semantics (pronominal bound-variable construal) or in

syntax (A-Chain). A-Chain formation is only possible between simple reflexives, such as Dutch *zich*, and a local c-commanding DP. This is because pronominal elements underspecified for [number], such as Spanish *se* and Dutch *zich*, forcibly establish a checking relation with local DPs. He further argues that syntactic dependencies are cheaper than dependencies built in semantics or discourse. This means that (20a) is ungrammatical because it is blocked by the more economical (20b).

20. a. *Jan$_i$ ziet [hem$_i$ dansen] (bound variable)
'John sees him dance.'

b. Jan ziet [zich dansen] (A-Chain)
A-Chain
'John sees REFL dance.'

Baauw et al. (2011) and Avrutin (2006) propose that in populations with a reduced syntactic processing capacity, such as young children, bound variable construals and A-Chains may be equally economical, which will lead children to often accept (20a) as an alternative of (20b).

Note, finally, that the approach proposed here, in which the PIP is explained as the result of the interplay between structural and processing factors, accounts for the differences in performance between Dutch- and Spanish-speaking children in different syntactic contexts. In simple transitive sentences, Dutch-speaking children show a PIP, as a result of their limited processing ability, which affects the application of Rule I, an interface constraint on local co-reference. Spanish-speaking children do not show a PIP in these contexts, because weak pronouns are syntactic clitics in Spanish, and syntactic clitics do not allow local co-reference. This prevents Rule I from being invoked in Spanish, so that it cannot break down either. In ECM sentences Dutch-speaking children show an extra-strong PIP, since two independent causes of the PIP coincide; on the one hand, children may accept a bound-variable construal (20a) as an equally economical alternative for A-Chain formation (20b), due to their more limited syntactic processing capacity. On the other hand, children may accept local co-reference as an alternative for a bound-variable construal, due to a breakdown of Rule I. Spanish-speaking children, on the other hand, are only affected by the former cause; their more limited processing capacity will lead them to accept (21a) as an equally acceptable alternative of (21b), but they will not be affected by a breakdown of Rule I, because this constraint is never invoked in structures containing syntactic clitics, such as in Spanish.[6]

21. a. *la niña$_i$ la$_i$ ve bailar (bound variable)
'The girl sees her dance.'

b. la niña se ve bailar (A-Chain)
A-Chain
'The girl sees REFL dance.'

---

[6] Note that the PIP in Spanish ECM sentences cannot be argued to be the result of the fact that the ECM sentence is longer. In Baauw (2000) it is shown that Spanish-speaking children reject the reflexive interpretation of (iv) 85% of the time.

iv. el niño trata de [PRO lavarle]
the boy tries to wash-him

Similar results were found for Catalan (Escobar and Gavarró 1999).

4.4. The Pronoun Interpretation Problem in Agrammatic Patients

A strong argument in favor of a processing account of the PIP are populations that already have acquired language but lost their ability to use it in a target-like way. Patients with agrammatic Broca aphasia are such a population. They were in full command of a language but lost the ability to speak fluently because of brain damage. Typically, agrammatic patients have problems with the use of functional elements such as articles and tense, and with the interpretation of syntactic operations such as passive formation and wh-questions (Avrutin 2006). Grodzinsky et al. (1993) found a PIP in English-speaking agrammatic patients in sentences such as (1a), identical to children acquiring language. Vasic (2006) found a similar result in Dutch patients (see Ruigendijk et al. 2006 for an overview of several studies). Several researchers have claimed that agrammatic patients do not lose linguistic knowledge, but have problems with the applications of this knowledge, which leads them to show variability in their performance on syntactic operations (Avrutin et al. 1999; Kolk 1987).

Concretely, Avrutin (1994, 2004, 2006) proposes that the similarities between child language acquisition and adult agrammatism are the result of a common problem with the processing of syntactic configurations. He argues that this processing limitation affects specifically those syntactic operations that are used to structure extra-syntactic (discourse, pragmatic) information. Although information of that kind can in principle also be conveyed by making use of other, extra-syntactic operations, the use of syntax is considered to be more economical in adult speakers with no language pathologies. This explains why adult speakers with no language pathologies normally use articles to express specificity and tense morphemes to express time and use Chain formation (21b) instead of variable binding (21a) to establish referential dependencies. They further argue that in populations with weakened syntactic processing abilities, such as young children and patients with agrammatic Broca aphasia, syntax is equally economical as other, extra-syntactic, modules. This explains their omission of articles, use of root infinitives, the PIP and other non-target-like performance at the syntax-discourse interface. On the other hand, syntactic operations or constructions that do not interface with discourse, such as the syntactic position of functional heads, are predicted to be relatively robust.

If the PIP is not the result of incomplete acquisition, nor of a loss of linguistic structures, it is predicted that agrammatic patients that speak Spanish should show a similar response pattern to Spanish-speaking children. They should show a PIP in those constructions in which cross-modular economy considerations play a role, but not in other constructions. Baauw and Cuetos (2003) performed a picture verification task in which four patients, diagnosed with agrammatic Broca's aphasia, participated. They were aged between 48 and 64 years. In the experiment the interpretation of reflexives and pronouns was tested. The pronouns were tested in simple transitive sentences (22a) and in ECM sentences (22b). The results showed that the patients rejected the reflexive interpretation of (22a) 79% of the time, but only 21% of the time in (22b).

22.	a.	La niña	la	tocó		[la = la niña, 21% acceptance]
		the girl	her	touched		
	b.	La niña	la	vio	bailar	[la = la niña, 79% acceptance]
		the girl	her	saw	dance	

This shows that agrammatic patients, like children, accept a bound variable interpretation of the clitic pronoun in contexts in which Reinhart and Reuland's (1993) Principle B plays no role, as in (22b), but that they reject a local co-reference interpretation in (22a), where binding of the clitic pronoun would lead to a violation of Principle B. This shows that agrammatic patients, like children, do not violate Principle B, but also that both populations analyze Spanish weak pronouns as syntactic clitics, making a local co-reference construal impossible. As has been argued before, children acquire the syntactic properties of clitic pronouns early. Since the syntactic position of the clitic functional head is not subject to any cross-modular economy consideration, this property of syntax is relatively robust.

Therefore, children hardly make mistakes with this aspect of grammar and agrammatic patients do not lose these properties.[7]

4.5. Task Effects

Another argument that supports a processing account of the PIP is its sensitivity to task effects. Baauw et al. (2011) argued that children perform more target-like on experimental methods that require fewer processing resources. Concretely, they show that the intensity of the PIP, both in children and agrammatic patients, differs according to the used research method, picture verification tasks, leading to higher rates of non-target-like performance than less demanding picture selection tasks.

Following Baauw et al. (2011), we assume that children, when they hear a sentence with a pronoun, will try to resolve the pronoun in a way that avoids violating constraints on local binding (Principle B, but also economy considerations involving bound variable construals) and, depending on the language, constraints on local co-reference (Rule I). This makes the extra-sentential DP the preferred antecedent but, depending on contextual or task-related factors, the sentence internal antecedent may also be considered.

Given this strategy, the picture selection task and the picture verification task impose different processing demands on the child. In the picture selection task that we will consider, the child is presented a test sentence (for example *John is washing him*) and tree pictures. One picture represents the "reflexive" action (for instance, *John washing himself*), another one represents a "transitive" action (*John washing Peter*), and a third one a non-related action (*John filming somebody*). The child must point at the picture that best matches the meaning of the test sentence. In the picture verification task, on the other hand, the child is presented a test sentence and one picture. This picture represents either a "reflexive" action or an action performed by one individual to another individual, and the child must decide whether the content of the picture matches the test sentence by saying "yes" or "no". When the test sentence is a pronoun, adult speakers with no language pathologies will say "yes" when the picture represents a non-reflexive action, and "no" when the picture represents a reflexive action.

It is clear that in the case of the picture selection task the child can avoid considering an interpretation that involves cross-modular economy considerations or reference set calculations, by choosing the picture representing the "transitive" interpretation. In the case of a picture verification task, this freedom is not offered to the child. The child is "forced" to consider the possibility of a local co-reference or binding interpretation, which will often lead to non-target-like responses because of the processing demands that this interpretation imposes on the child.

In Baauw et al. (2011) this prediction was tested by comparing results on picture selection experiments with Spanish-speaking children and agrammatic patients, with earlier results of studies that used a picture verification task.

In a picture selection study, 38 Spanish-speaking children ranging between 5;3 and 6;2 years old (mean age 5;9) were tested on sentences containing reflexives, pronouns in simple transitive sentences (such as 16b), and pronouns in ECM sentences (such as 16a). The results show that they scored 94% target-like on the interpretation of pronouns in simple transitive sentences, and 75% target-like on the interpretation of pronouns in ECM sentences. These results were compared with the percentage correct "no" responses on trials testing pronouns that elicited "no" responses in adults without language pathology, provided by several picture verification (PV) studies: PV-1 (Baauw et al. 1997; Baauw and Cuetos 2003), PV-2 (Baauw 2000) and PV-3 (Alija and Baauw 2005). In Figure 1 the results from both the Picture Selection (PS) and Picture Verification (PV) studies are presented.

---

[7] It has been reported that Spanish agrammatic patients, unlike Spanish children (Wexler et al. 2004), often omit object pronouns. Importantly though, when they use them they are in the correct position (Reznik et al. 1995). This indicates that the underlying syntactic structure of clitics is preserved in agrammatism.

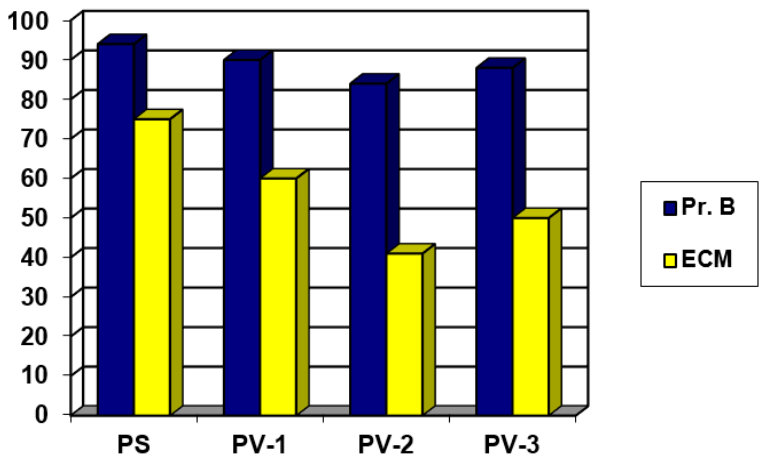


Figure 1. Percentage of target-like performance. Pr. B: simple transitive sentences; ECM: Exceptional Case Marking sentences; PS: Picture Selection; PV: Picture Verification.

As can be seen, children score at ceiling in simple transitive sentences, in all studies, independently of the used method. In the ECM sentences, on the other hand, we see that children scored systematically less target-like in the PV studies (around 50% target-like) than in the PS study (75% target-like).

A similar comparison was carried out with Spanish-speaking agrammatic patients. In a picture selection study, 7 agrammatic patients ranging from 42 to 68 years of age were tested with the same materials as in the child study reported above (only the number of trials per sentence type was expanded). The patients scored 89% target-like in simple transitive sentences and 61% target-like in the ECM sentences. In Figure 2 these results are compared with the PV-task results on the pronominal "no" trials from Baauw and Cuetos (2003).

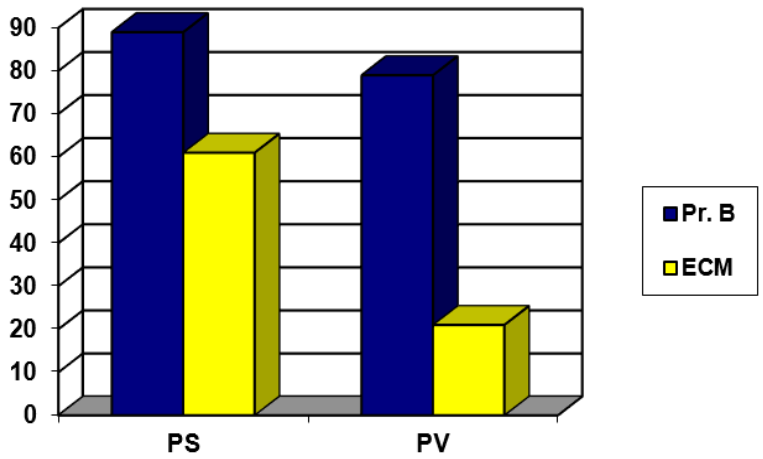


Figure 2. Percentage of target-like performance.

As we can see, the patients scored highly target-like on the simple transitive sentences, in both studies, but scored considerably less target-like on ECM sentences in the PV study than in the PS study.

These results are confirmed by a combined PS/PV experiment presented in Baauw et al. (2011). In this experiment 20 Dutch-speaking children (ranging from 4;11 to 6;0, mean age 4;10) were presented a test in which PS and PV items alternated. In the test, three sentence types were tested: reflexives, pronouns in simple transitive sentences, and pronouns in ECM sentences. The results on the pronominal trials are presented in Figure 3.

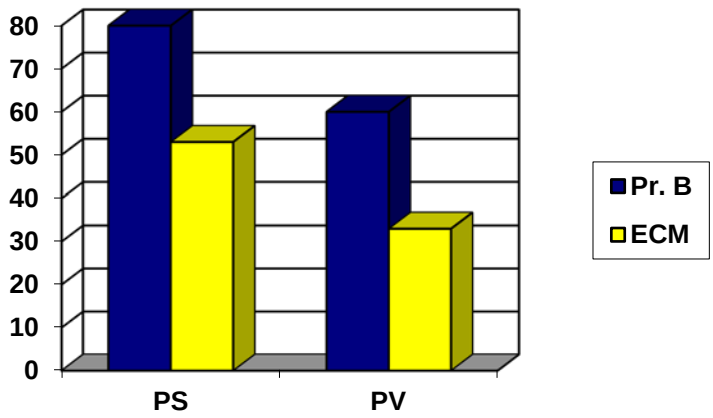


Figure 3. Percentage of target-like performance.

The results clearly show a statistically significant effect of both methodology (PS vs. PV) and construction (Pr. B vs. ECM).

Finally, the effect of task is also confirmed by van Rij et al. (2010). They tested 75 Dutch-speaking children in a picture verification task, in which the test sentences were presented at different speech rates. The assumption is that at high speed rates the analysis and interpretation of pronouns require more processing resources. The results showed that a reduction in speech rate affected the interpretation of pronouns positively.

5. Conclusions

In this paper, we have argued that the PIP can best be viewed as the result of an interaction between language specific properties of predicates and pronouns on the one hand, and a processing limitation that affects children's ability to carry out reference set calculations and to apply cross-modular economy principles in a target-like way (Grodzinsky and Reinhart 1993; Avrutin 2004, 2006).

The language-specific properties of pronouns account for the observed cross-linguistic differences between languages without syntactic clitic pronouns, such as Dutch and English, on the one hand, and languages that have syntactic clitics, such as Spanish, on the other hand. This property of Spanish syntax, which is acquired early, accounts for the virtual absence of a PIP in simple transitive sentences, such as the Spanish counterparts of *the girl pointed at her*, and the weaker PIP in ECM sentences such as *the girl saw her dance*. The impossibility of clitic pronouns (or the null-object that they are associated with) to enter a local co-reference relation with the subject prevents Rule I from being invoked, to the effect that it cannot "break down" either. Romance strong pronouns, on the other hand, are predicted to give rise to a PIP in young children. Evidence from Italian acquisition suggests that strong direct object pronouns do indeed give rise to such an effect. Spanish strong direct object pronouns do not give rise to a PIP, probably because they are obligatorily doubled by a clitic pronoun. When strong pronouns are the complement of a preposition, the increased acceptance of the reflexive interpretation by Spanish-speaking children is most likely due to their more frequent analysis of PPs as separate binding domains.

The syntactic properties of ECM sentences explain why Dutch-speaking children show a stronger PIP in this construction. In addition to their problems with the application of Rule I, their weaker syntactic processing capacity changes the economy relation between referential dependencies formed in syntax (A-Chains between simple reflexives and local DPs) and pronominal bound variable relations. This leads them to accept the embedded subject pronoun of the ECM sentence to be bound by the main clause subject DP. Spanish-speaking children are not affected by a breakdown of Rule I, due to the clitic status of Spanish weak pronouns, but they are affected by the changed economy relation between syntactic and extra-syntactic referential dependencies. This explains the presence of a PIP in their ECM sentences, but a weaker one than Dutch-speaking children show in these constructions.

A processing account of the PIP explains many of the experimental results presented in this paper. It accounts for the variability of the PIP, for its presence in adults affected by agrammatic Broca's aphasia, for its absence in production, and for its sensitivity to task effects. It also has the advantage that it does not face the kind of learnability problems that alternative approaches face. This does not mean that the child does not need to learn anything to perform in an adult-like manner on the interpretation of pronouns; the child needs to learn the syntactic properties of pronouns, such as whether the pronoun to be learned is a syntactic clitic or not. However, to acquire these properties, the child arguably receives sufficient relevant examples from the input. In fact, research shows that children acquire these properties at an early stage of acquisition (Ezeizabarrena 1996; Guasti 1994).

Conflicts of Interest: The author declares no conflict of interest.

References

Alija, Maira, and Sergio Baauw. 2005. La Interpretación de los Pronombres Masculinos y Femeninos por Niños Neerlandeses y Españoles [The Interpretation of Masculine and Feminine Pronouns by Dutch and Spanish Children]. *Cognitiva* 17: 3–28.

Aoun, Joseph. 1999. Clitic doubled arguments. In *Beyond Principles and Parameters*. Edited by Kyle Johnson and Ian Roberts. Dordrecht: Kluwer Academic Publishers.

Avrutin, Sergey. 1994. Psycholinguistic Investigations in the Theory of Reference. Unpublished Ph.D. dissertation, MIT, Cambridge, MA, USA.

Avrutin, Sergey. 2004. Optionality in child and aphasic speech. *Lingue e Linguaggio* 1: 67–69.

Avrutin, Sergey. 2006. Weak Syntax. In *Broca's Region*. Edited by Yosef Grodzinsky and Katrin Amunts. Oxford: Oxford University Press, pp. 49–62.

Avrutin, Sergey, and Sergio Baauw. 2013. A processing view on agrammatism. In *Diagnosing Syntax*. Edited by Lisa Lai-Shen Cheng and Norbert Corver. Oxford: Oxford University Press, pp. 397–411.

Avrutin, Sergey, and Kenneth Wexler. 1992. Development of Principle B in Russian: Coindexation at LF and Coreference. *Language Acquisition* 4: 259–306. [CrossRef]

Avrutin, Sergey, Stuart Lubarsky, and Jennifer Greene. 1999. Comprehension of Contrastive Stress by Agrammatic Broca's Aphasics. *Brain and Language* 70: 163–86. [CrossRef] [PubMed]

Baauw, Sergio. 2000. Grammatical Features and the Acquisition of Reference. A Comparative Study of Dutch and Spanish. Ph.D. dissertation, Utrecht University, Utrecht, The Netherlands. Utrecht: LOT.

Baauw, Sergio, and Fernando Cuetos. 2003. The interpretation of pronouns in Spanish language acquisition and breakdown: Evidence for the "Principle B Delay" as a Non- unitary phenomenon. *Language Acquisition* 11: 219–75. [CrossRef]

Baauw, Sergio, María A. Escobar, and William Philip. 1997. A Delay of Principle B Effect in Spanish Speaking Children: The Role of Lexical Feature Acquisition. In *Proceedings of the GALA '97 Conference on Language Acquisition*. Edited by Conxita Lleó, Antonella Sorace, Caroline Heycock and Richard Shillcock. Edinburgh: Human Communication Research Centre.

Baauw, Sergio, Shalom Zuckerman, Esther Ruigendijk, and Sergey Avrutin. 2011. Principle B Delays as a processing problem: Evidence from task effects. In *Production- comprehension Asymmetries in Child Language*. Edited by Angela Grimm, Anja Müller, Cornelia Hamann and Esther Ruigendijk. Berlin: De Gruyter, pp. 247–72.

Berger, Carola V. 1999. De Verwerving van Anaforische Relaties in het Italiaans: Een Onderzoek naar Lexical Feature Learning. [The Acquisition of Anaphoric Relations in Italian: An Investigation in Lexical Feature Learning]. Master's Thesis, Utrecht University, Utrecht, The Netherlands.

Bloom, Paul, Andrew Barss, Janet Nicol, and Laura Conway. 1994. Children's knowledge of binding and coreference: Evidence from spontaneous speech. *Language* 70: 53–71. [CrossRef]

Chien, Yu-Chin, and Kenneth Wexler. 1990. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1: 225–95. [CrossRef]

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Conroy, Anastasia, Eri Takahashi, Jeffrey Lidz, and Colin Phillips. 2009. Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry* 40: 446–86. [CrossRef]

Crain, Stephen, and Rosalind Thornton. 1998. *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge: MIT Press.

Delfitto, Denis. 2002. On the Semantics of Pronominal Clitics and Some of Its Consequences. *Catalan Journal of Linguistics* 1: 41–69. [CrossRef]

Di Sciullo, Anna Maria, and Calixto Agüero-Bautista. 2008. The Delay of Principle B Effect (DPBE) and its absence in some languages. *Language and Speech* 51: 77–100. [CrossRef] [PubMed]

Elbourne, Paul. 2003. Are our children speaking Middle English? Paper presented at CUNY Syntax Supper.

Elbourne, Paul. 2005. On the acquisition of Principle B. *Linguistic Inquiry* 36: 333–65. [CrossRef]

Escobar, Linda, and Anna Gavarró. 1999. *The Acquisition of Catalan Clitics and Its Implications for Complex Verb Structure*. Report de Recerca, Rep. No. GGT-99-3; Barcelona: Universitat Autònoma de Barcelona.

Ezeizabarrena, María José. 1996. Adquisición de la Morfología Verbal en Euskera y Castellano por Niños Bilingües. Unpublished Ph.D. thesis, University of Hamburg, Hamburg, Germany.

Fernández-Soriano, Olga. 1989. Strong pronouns in null-subject languages and the avoid pronoun principle. In *Proceedings of the Student Conference on Linguistics 1989. MIT Working Papers in Linguistics 11*. Edited by P. Branigan, J. Gaulgin, M. Kubo and K. Murasugi. Cambridge: MIT Press.

Fodor, Janet. 1992. *Designated Triggers versus the Subset Principle*. Unpublished manuscript; New York: CUNY Graduate Center.

Grodzinsky, Yosef, and Tanya Reinhart. 1993. The Innateness of Binding and Coreference. *Linguistic Inquiry* 24: 69–102.

Grodzinsky, Yosef, Kenneth Wexler, Yu-Chin Chien, Susan Marakovitz, and Julie Solomon. 1993. The breakdown of binding relations. *Brain and Language* 45: 396–422. [CrossRef] [PubMed]

Guasti, Maria Teresa. 1994. Verb syntax in Italian child grammar: finite and non-finite verbs. *Language Acquisition* 3: 1–40. [CrossRef]

Hamann, Cornelia, Odette Kowalski, and William Philip. 1997. The French 'Delay of Principle B Effect'. In *Proceedings of the Annual Boston University Conference on Language Development 21*. Edited by Elizabeth Hughes, Mary Hughes and Annabel Greenhill. Somerville: Cascadilla Press.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Hestvik, Arild, and William Philip. 2000. Binding and Coreference in Norwegian Child Language. *Language Acquisition* 8: 171–235. [CrossRef]

Jakubowicz, Celia. 1984. On Markedness and Binding Principles. In *Proceedings of the Northeastern Linguistic Society*. Edited by Charles Jones and Peter Sells. Amherst: GLSA, University of Massachusetts, vol. 14.

Kolk, Herman. 1987. A theory of grammatical impairment in aphasia. In *Natural Language Generation*. Edited by Gerard Kempen. Dordrecht: Kluwer.

Koster, Charlotte. 1993. Errors in Anaphora Acquisition. Ph.D. dissertation, Utrecht University, Utrecht, The Netherlands.

Lyczkowsky, David Alan. 1999. Adquiéretelo: On the Acquisition of Pronominal Object Clitics in Spanish. Bachelor's thesis, Harvard University, Cambridge, MA, USA.

Manzini, M. Rita, and Kenneth Wexler. 1987. Parameters, binding theory, and learnability. *Linguistic Inquiry* 18: 413–44.

Margócsy, D. 2000. Pronouns in Hungarian. Master's thesis, Utrecht University College, Utrecht, The Netherlands.

McDaniel, Dana, and Thomas L. Maxfield. 1992. Principle B and contrastive stress. *Language Acquisition* 2: 337–58. [CrossRef]

McKee, Cecile. 1992. A comparison of pronouns and anaphors in Italian and English acquisition. *Language Acquisition* 2: 21–54. [CrossRef]

Menuzzi, Sergio. 1996. Constraint interaction in binding. In *Linguistics in the Netherlands*. Edited by Crit Cremers and Marcel den Dikken. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Neeleman, Ad, and Fred Weerman. 1999. Flexible Syntax: A Theory of Case and Arguments. In *Studies in Natural Language and Linguistic Theory*. Dordrecht: Kluwer Academic Publishers, vol. 47.

Philip, William, and Peter Coopmans. 1996. The role of lexical feature acquisition in the development of pronominal anaphora. In *Amsterdam Series on Child Language Development*. Edited by William Philip and Frank Wijnen. Amsterdam: Instituut Algemene Taalwetenschap, vol. 5.

Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24: 657–720.

Reuland, Eric. 2001. Primitives of Binding. *Linguistic Inquiry* 32: 439–92. [CrossRef]

Reznik, Magdalena, Silvia Dubrovsky, and Silvia Maldonado. 1995. Agrammatism in Spanish: A Case Study. *Brain and Language* 51: 355–68. [CrossRef] [PubMed]

Ruigendijk, Esther, Nada Vasić, and Sergey Avrutin. 2006. Reference Assignment: Using Language Breakdown to Choose Between Theoretical Approaches. *Brain and Language* 96: 302–17. [CrossRef] [PubMed]

Schaeffer, Jeannette. 1997. On the acquisition of object placement in Dutch and Italian. In *Proceedings of the GALA '97 Conference on Language Acquisition*. Edited by Antonella Sorace, Caroline Heycock and Richard Shillcock. Edinburgh: Human Communication Research Centre.

Sigurjónsdóttir, Sigríður. 1992. Binding in Icelandic: Evidence from Language Acquisition. Ph.D. dissertation, University of California, Los Angeles, CA, USA.

Sportiche, Dominique. 1992. *Clitic Constructions*. Unpublished manuscript; Los Angeles: University of California.

Thornton, Rosalind, and Kenneth Wexler. 1999. *Principle B, VP Ellipsis and Interpretation in Child Grammar*. Cambridge: MIT Press.

van Rij, Jacolien, Hedderik Van Rijn, and Petra Hendriks. 2010. Cognitive architectures and language acquisition: A case study in pronoun comprehension. *Journal of Child Language* 37: 731–66. [CrossRef] [PubMed]

Varlokosta, Spyridoula. 2000. Lack of Clitic-Pronoun Distinctions in the Acquisition of Principle B in Child Greek. In *Proceedings of the 24th Annual Boston University Conference on Language Development 24*. Edited by S. Catherine Howell, Sarah A. Fish and Thea Keith-Lucas. Somerville: Cascadilla Press.

Vasic, Nada. 2006. Pronoun Comprehension in Agrammatic Aphasia. Ph.D. dissertation, Utrecht University, Utrecht, The Netherlands. Utrecht: LOT.

De Villiers, Jill, Jacqueline Cahillane, and Emily Altreuter. 2005. Touchy subject: Optimality and coreference. Unpublished manuscript. Smith College.

Wexler, Kenneth, Anna Gavarró, and Vicenç Torrens. 2004. Feature checking and object clitic omission in child catalan and Spanish. In *Romance Languages and Linguistic Theory 2002*. Edited by Reineke Bok-Bennema, Bart Hollebrandse, Brigitte Kampers Manhe and Petra Sleeman. Amsterdam: John Benjamins.

languages

MDPI

Article

Language Mixing in the Nominal Phrase: Implications of a Distributed Morphology Perspective †

Michèle Burkholder

Department of Linguistics, University of Ottawa, 70 Laurier Ave. E., Ottawa, ON K1N 6N5, Canada; mburk048@uottawa.ca

† Some of the ideas and data that are discussed here were presented at the 14th Generative Approaches to Second Language Acquisition (GASLA) Conference, Southampton, UK, 7–9 April 2017.

Received: 14 January 2018; Accepted: 9 April 2018; Published: 12 April 2018

Abstract: This paper investigates a pattern found in Spanish–English mixed language corpora whereby it is common to switch from a Spanish determiner to an English noun (e.g., *la house*, 'the house'), but rare to switch from an English determiner to a Spanish noun (e.g., *the casa*, 'the house'). Unlike previous theoretical accounts of this asymmetry, that which is proposed here follows assumptions of the Distributed Morphology (DM) framework, specifically those regarding the relationship between grammatical gender and nominal declension class in Spanish. Crucially, and again in contrast to previous accounts, it is demonstrated that this approach predicts no such asymmetry for French–English. This hypothesis is tested experimentally using an acceptability judgment task with self-paced reading, and as expected, no evidence is found for an asymmetry. This experiment is also used to test predictions regarding how English nominal roots in mixed nominal phrases are assigned grammatical gender, and the impact of language background factors such as age of acquisition. Evidence is found that bilinguals attempt to assign analogical gender if possible, but that late sequential bilinguals have a stronger preference for this option than do simultaneous bilinguals.

Keywords: Distributed Morphology; language mixing; nominal phrase; Spanish–English; French–English; analogical criterion; age of acquisition; language dominance; bilingualism

1. Introduction

Current mainstream linguistic theory has been built upon language data obtained from native speakers of a large sample of the world's diverse languages. What is often overlooked, however, is that most people in the world speak more than one language, and that such speakers often combine elements from both of their languages within a single utterance, and sometimes even within a single word. This type of language mixing is spontaneous in that it is not possible to predict when a switch or insertion will occur; however, it is also highly constrained (Lipski 1978; Pfaff 1979; Poplack 1980; Woolford 1983).[1] Considerable research has consequently been dedicated to understanding the formal linguistic principles that underlie such language mixing (e.g., Poplack 1980, 1981; Sankoff and Poplack 1981; Joshi 1985; Di Sciullo et al. 1986; Mahootian 1993; Belazi et al. 1994; MacSwan 2000; Liceras et al. 2008).

[1] The term "language mixing" will be used to refer to a phenomenon termed "intra-sentential code-switching" by some (e.g., MacSwan 2005), and "nonce borrowing" by others (e.g., Poplack and Dion 2012).

Within the framework of the minimalist program (Chomsky 1993, 1995, 2005), a productive line of research has assumed that the principles that govern bilingual language mixing are the same as those that govern monolingual unmixed language; that there should be no specific constraints that apply only to language mixing (MacSwan 2000). For example, if linguistic competence involves choosing and combining lexical items via operations such as select and merge to form convergent syntactic derivations, then the simple difference between unilingual unmixed language and bilingual mixed language is that the latter operates over lexical items from two languages rather than one. Consequently, mixed-language data can provide a unique testing ground for linguistic theory and, in turn, it can offer important insights into the nature of the underlying linguistic system(s) a bilingual has acquired.

While investigations within this line of research have often taken a lexicalist view of morphosyntax (e.g., Cantone and Müller 2008; Van Gelderen and MacSwan 2008; Van Dulm 2009; Moro Quintanilla 2014), other researchers (e.g., González-Vilbazo and López 2011; Grimstad et al. 2014; Alexiadou et al. 2015; Riksem 2017) have also been exploring the predictions and explanatory power of non-lexicalist frameworks such as Distributed Morphology (DM; Halle and Marantz 1993; Marantz 1997; Embick and Noyer 2007). As such, the first goal of this paper is to focus on a particular language mixing pattern that has been reported in the literature, and to re-examine it within the theoretical framework of DM. Specifically, the object of investigation is a fairly robust asymmetry whereby switches between a Spanish determiner (D) and an English noun (N), such as (1a), are considered to be grammatical, but those between an English D and a Spanish N, as in (1b), are not (Jake et al. 2005; MacSwan 2005; Liceras et al. 2008, 2016).[2]

1.	a.	*la*	house
		the.DF.SG.F	house
	b.	* the	*casa*
		the.DF	house

In previous minimalist accounts of this asymmetry, some have argued that it is a consequence of an underlying asymmetry in the morphosyntactic features of the two languages; specifically, that Spanish has grammatical gender features whereas English does not (e.g., MacSwan 2005; Liceras et al. 2008; Moro Quintanilla 2014). While the account that will be proposed here also attributes a central role to the gender feature asymmetry, it will be argued that it is the relationship between gender and nominal declension classes in Spanish that is the true source. Interestingly, the proposed theoretical account makes the prediction that not all language pairs with a gender feature asymmetry will exhibit a mixed determiner phrase (DP) asymmetry. It will be argued that this is the case for French–English.

It is important to note that this account will assume that the bilinguals in question have the same underlying linguistic system in their two languages as do monolingual native speakers of each of their respective languages. Indeed, this may be the case for individuals who acquired both of their languages from birth, and who are perfectly balanced with respect to proficiency and frequency of use.[3] An important issue with this assumption, however, is that most bilinguals do not fit these "ideal" characteristics. As such, the theoretical proposal will also be discussed with respect to the different predictions it makes for bilinguals who have acquired their two languages under different learning conditions (e.g., simultaneous versus late sequential bilinguals). The second goal of this paper is then to present an experiment investigating some of these predictions for French–English mixed

[2] The following abbreviations for features will be used throughout the text: plural = [PL]; singular = [SG]; definite = [DF]; feminine = [F] or [+FEM]; masculine = [M] or [-FEM]; neuter = [N].

[3] Although it is often argued that even such bilinguals are not necessarily "two monolinguals in one" (e.g., González-Vilbazo and López 2011), this can be an adequate initial hypothesis.

DPs. Experimental data, however, are not directly reflective of a bilingual's underlying linguistic competence, but may also reflect performance factors related to language processing. For bilinguals, the relative ease with which they process and access words from their two languages can depend on external factors such as language proficiency and dominance (i.e., whether they use one language preferentially). As such, the impact of these factors will also be taken into consideration when forming hypotheses.

The remainder of this paper will be organized as follows. Section 2 will provide the necessary background, including a more detailed look at the mixed DP asymmetry in Spanish–English (2.1), a brief overview of the DM framework (2.2), and the particular view on gender that will adopted (2.3). This view will be demonstrated in the context of unmixed Spanish DPs. In Section 3, this framework will be extended to account for Spanish–English mixed DPs (3.1), and to make predictions for French–English mixed DPs (3.2). This proposal will then be discussed with respect to the predictions that such a model makes for bilinguals with a late age of acquisition (3.3), and with respect to how language proficiency and language dominance might influence the patterns observed in actual mixed-language data (3.4). Section 4 will then present the details of an experimental study designed to test the predictions made for different types of French–English bilinguals. Results will be presented in Section 5 and discussed in Section 6. Section 7 will conclude.

2. Background

2.1. The Mixed Determiner Phrase Asymmetry in Spanish–English

The claim that it is ungrammatical to switch between an English D and a Spanish N is based on the observation that such constituents are quite rare in the speech of Spanish–English bilinguals. For example, in Liceras et al.'s (2006) analysis of the Moyer's (1993) Spanish–English corpus, they found that only two of the 215 mixed DPs consisted of an English D and a Spanish N; and in the corpus reported on in Jake et al. (2002), there are no reported instances of such mixed DPs (Jake et al. 2005).[4] Further, Liceras et al. (2016) report a similar pattern found in the spontaneous productions of simultaneous bilingual children, based on corpus data from from the FerFuLice corpus (Fernández Fuertes and Liceras 2009) available in CHILDES MacWhinney (2000).

MacSwan (2005) accounts for this asymmetry based on Moro's (2001) analysis, which takes a minimalist point of view with lexicalist assumptions. They propose that mixed DPs such as *the casa* crash because a Spanish N cannot successfully value and delete the uninterpretable features of an English D "in one fell swoop". This is because the Spanish N has PERSON, NUMBER and GENDER features, but there is no corresponding GENDER feature on an English D. In contrast, mixed DPs such as *la house* converge because the English N's features, PERSON and NUMBER, are a subset of those of the Spanish D. These can therefore be valued and deleted, leading to convergence (MacSwan 2005).

Liceras et al. (2008, 2016), however, argue that this proposal is too categorical, and as such, it is not able to account for the fact that this asymmetry has not been replicated in acceptability judgment tasks. For example, they present data showing that adult L1-Spanish L2-English sequential bilinguals (Liceras et al. 2008; Liceras 2014), as well as both sequential and simultaneous bilingual children (Fernández Fuertes et al. 2011), evaluate mixed DPs with an English D to be as acceptable as those with a Spanish D, as long as the Spanish D matches in gender with the English N's translation equivalent.[5] In order to account for this data as well as the patterns found in the corpus data, they propose the grammatical features spell-out hypothesis. This hypothesis is based on the theoretical assumptions of Pesetsky and Torrego (2001) regarding feature checking, and those of Kihm (2005)

[4] Note that these researchers expect that mixed DPs with English Ds are possible and would be more frequent in a corpus where the speakers' dominant language is English, as there would be a greater tendency towards English providing the morphosyntactic frame (i.e., the "matrix language"), which they argue is the source of functional items such as determiners.

[5] Note that "L1" refers to an individual's native language, and "L2" to their second language.

regarding gender as a feature on the nominalizing head, *n*. Importantly, it provides a solution to the conflicting experimental and naturalistic data by suggesting that in processing, but not in production, the gender valuation mechanism between an N and a D can be cancelled in order to be more economical with processing resources.

What these two accounts have in common, however, is that they predict a similar mixed DP asymmetry for any language pair that has an underlying gender feature asymmetry. Indeed, this asymmetry has been reported for other language pairs, such as Welsh–English (Deuchar 2006) and German–English (Jorschick et al. 2011). On the other hand, there is evidence from a Norwegian–English mixed language corpus that switches in both directions are attested (Alexiadou et al. 2015; Riksem 2017). Further, there is little in the literature reporting on French–English mixed DPs, and it does not appear that the mixed DP asymmetry has been investigated here. As such, it is unclear if it exists for this language pair. The DM-based theoretical account that will be proposed aims to reconcile these conflicting observations. Crucially and in contrast with previous accounts, it does not predict such an asymmetry for French–English.

2.2. The Distributed Morphology Framework

From a traditional lexicalist perspective of minimalist syntax, the lexicon includes items that are specified for phonological form, word category information, and morphosyntactic features such as PERSON, NUMBER and GENDER (e.g., Lieber 1992; Chomsky 1995; Müller and Wechsler 2014). Morphologically complex words are derived in the lexicon and enter the derivation with their features. Syntactic structure is then projected from those features. This "lexicalist" approach has been extended to account for bilingual language mixing, where it has been proposed to operate in exactly the same way, the only difference being that items from either of a bilingual's two lexicons can be selected into the numeration (MacSwan 1999, 2000). MacSwan (1999, p. 235) explicitly makes reference to such lexicalist assumptions when he defines borrowing from one language into another as "an operation whereby a new stem is introduced into a specific lexicon where morphologically complex items are formed before entering the numeration".

In contrast, the DM framework (Halle and Marantz 1993; Marantz 1997; Embick and Noyer 2007), as a non-lexicalist approach, assumes not a single lexicon, but one that is divided up into three separate "lists". Each of these lists is accessed at different points in the derivation. List 1 consists of bare lexical roots and abstract morphemes, which are the elements that can be selected to form a numeration. Roots (e.g., $\sqrt{\text{HOUSE}}$, $\sqrt{\text{BOOK}}$) have no morphosyntactic features and they are not specified for word categories. It has also been argued that they are not specified for phonological forms (e.g., Siddiqi 2009), nor even for semantics (e.g., Harley 2014). Abstract morphemes, on the other hand, consist of morphosyntactic features (e.g., [SG], [DF], $n_{[+\text{FEM}]}$). List 1 elements that are selected into the numeration combine with each other in the computational syntactic component to form morphologically complex words via the same operations as those which combine words to form syntactic structures (i.e., merge, agree). For example, bare roots can become nouns by combining with the nominalizing head, *n*.

Syntactic terminal nodes at spell-out therefore consist of roots fused together with the abstract morphemes that have merged above them (Siddiqi 2009), and still lack phonological and semantic information. This brings us to the vocabulary items (VIs) of List 2, which are the mappings of these root and feature bundles to their corresponding phonological forms. List 3 items are the mapping of terminal nodes to their semantic exponents.[6] Crucially, VIs are morphologically underspecified and compete for insertion into the fully specified syntactic terminal nodes, such that the exponent with the most compatible features (but no incompatible features) is selected. This selection process is known as

[6] This is somewhat of a simplification; however, because List 3 is not relevant to the topic of this paper, it will not be discussed further.

the subset principle (Halle 1997). For example, of the Spanish VIs that compete at a D terminal node with [M], [SG] and [DF] features, (2d) provides the best match and so is inserted.

2. List 2 Spanish VIs competing for D[M.SG.DF]
   a. F.PL.D → las
   b. PL.DF → los
   c. F.DF → la
   d. **M.DF** → **el**
   e. DF → lo

2.3. A Distributed Morphology Account of Unmixed Spanish Determiner Phrases

Inanimate nouns in Spanish are arbitrarily assigned one of two grammatical genders, feminine and or masculine, and typically fall into one of three declension classes (I, II, III), which are each associated with a particular declension class suffix (*−o*, *−a*, or *−e*, respectively). However, while declension classes are strongly correlated with gender, such as in (3a) and (3b), they are indeed distinct from gender as there are exceptions, for example (3c). Also, while some nouns take no overt suffix, such as established loanwords, for example (3d), nouns that do take a declension class suffix never surface bare; they are bound roots.

3. a. la cas-a
   the.F house.F-II

   b. el libr-o
   the.M book.M-I

   c. la man-o
   the.F hand.F-I

   d. el béicon
   the.M bacon.M

Turning now to how unmixed Spanish DPs are derived from a DM perspective (see Figure 1a), it will assumed that bare lexical roots must merge with a nominalizing head, *n*, in order to surface as a noun. In languages without grammatical gender, such as English, *n* is plain; however, in Spanish, *n* is bundled together with either an uninterpretable *u*[+FEM] or *u*[-FEM] feature (Kihm 2005; Kramer 2015).[7] Spanish roots are licensed under one or the other, and this is what ensures that nouns consistently surface with the correct gender, feminine or masculine, respectively.[8] Importantly, Kramer (2015) proposes that, for Spanish, there are also "theme nodes" which are adjoined to *n* after spell-out. This is done in accordance with theme node insertion rules, such as those in (4). The rules in (4a) capture all the listed exceptions, such as (3c) above, and those in (4b) and (4c) ensure that, aside from those exceptions, declension class is indeed dependent on the underlying gender feature.

---

[7] See Kramer (2015) and also Kučerová (2017) for a more complete view on the different types of gender that are involved, for example, in the derivation of animate nouns with biological gender.

[8] Nouns that surface with masculine gender could alternatively be licensed under a plain *n* as default, as assumed by Kramer (2015); however, it will be assumed here that there is an uninterpretable [-FEM] for inanimate nouns just as there is an interpretable [-FEM] feature for animate nouns with male biological gender.

4. Theme node insertion rules for Spanish
   a. (i) Insert [THEME, III] in the context of √ESTANT-, √CARN- ...
      (ii) Insert [THEME, II] in the context of √DI- ...
      (iii) Insert [THEME, I] in the context of √MAN- ...

   b. Insert [THEME, II] in the context of *n*[+FEM]

   c. Insert [THEME, I] in the context of *n*[-FEM]

   d. Insert Ø elsewhere.

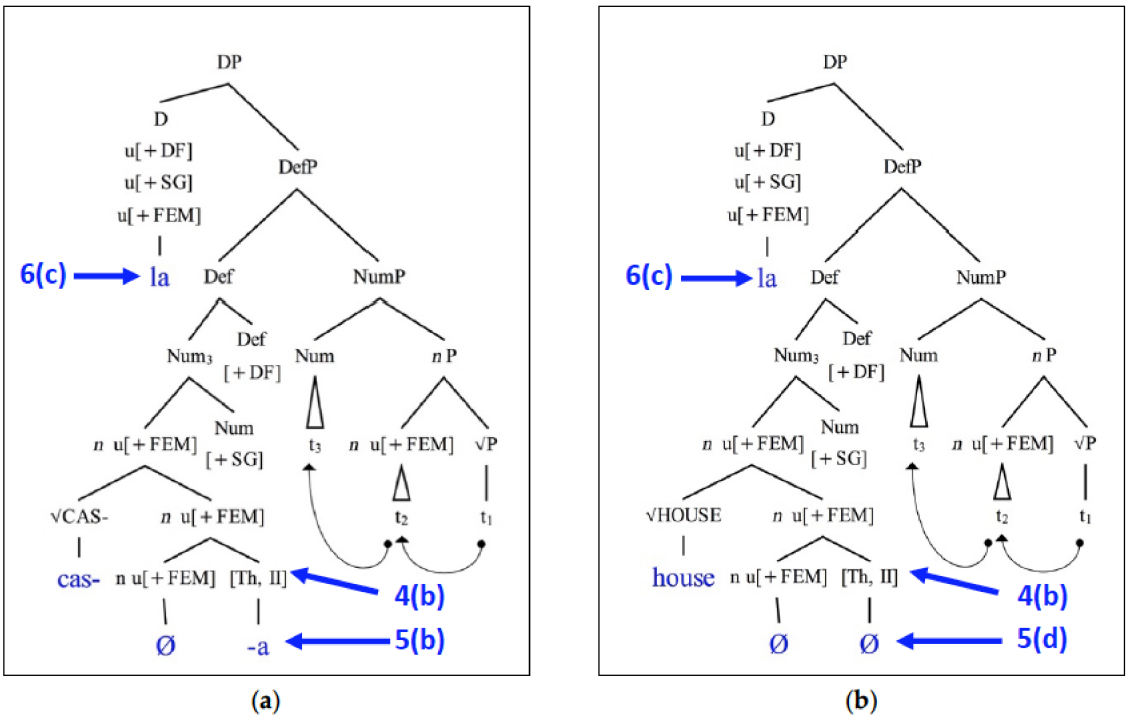


Figure 1. (**a**) Syntactic structure for the unmixed Spanish determiner phrase (DP) *la casa* 'the house'; (**b**) Syntactic structure for the mixed Spanish–English DP *la house* 'the house'. Numbers indicated by arrows show which insertion rules apply in the given syntactic context.

Then, when vocabulary insertion applies, List 2 theme vowels are inserted into the theme nodes. This is done in accordance with the rules in (5). Crucially, it will be proposed here that theme node insertion only applies in contexts where there is a bound root.[9] This ensures that athematic nouns, such as established loanwords (e.g., 3d), do not get a declension class suffix. It also captures the idea that the purpose of declension class suffixes is to render bound roots pronounceable by completing their phonological form.

[9] Kramer (2015) also proposes that theme vowels are only inserted when the theme node is left-adjacent to the number (Num) head in order to account for the fact that roots do not have theme vowels inserted when there is another affix (such as the diminutive) directly following the root.

5. List 2 vocabulary insertion rules for theme nodes
   a. [THEME, III] → -e/√X- __ Num
   b. [THEME, II] → -a/√X- __ Num
   c. [THEME, I] → -o/√X- __ Num
   d. [THEME] → Ø

3. Proposed Distributed Morphology Account of Mixed Determiner Phrases

Extending the DM model to language mixing in general, we can make the following two assumptions. First, it will be assumed that bilinguals have functionally separate distributed lexicons, such that in unilingual language mode they can access the set of lists associated to one language with limited interference from the other.[10] Second, during language mixing, neither language is inhibited. As such, List 1 elements from either language can be selected into the numeration, and List 2 VIs from both languages can compete for insertion after spell-out.

By applying these assumptions to the above account for unmixed Spanish DPs, an analysis for Spanish–English mixed DPs can now be proposed. This will then be extended to make predictions for French–English. As mentioned in the introduction, this proposal will assume that the bilinguals in question have "perfect" competence in each of their two languages. It will also assume that their innate knowledge is reflected in mixed language data. As such, the discussion of the theoretical proposal will focus on the impact that age of acquisition (AoA) and factors related to language processing may have on the expected patterns.

3.1. Analysis for Spanish–English

With respect to the specific structure involved in DPs, it will be assumed that English differs from Spanish in two relevant ways. First, because English has no grammatical gender system, it has only a plain *n*, and the definite determiner, *the*, has only a [DF] feature. Second, like established loanwords in Spanish, it will be assumed that English nominal roots are free, not bound.

3.1.1. The Mixed Determiner Phrase Asymmetry

Looking first at how mixed DPs such as *la house* are derived, Figure 1b shows that when an English root merges with a Spanish *n*, this operates in exactly the same way as with unmixed Spanish DPs. The crucial difference occurs during vocabulary insertion. Because the theme node is not right adjacent to a bound root in this case, rule (5d) applies, and no theme vowel is inserted. Because D values its uninterpretable φ-features (PERSON, NUMBER and GENDER) against a feature bundle containing a feminine gender feature, the Spanish determiner *la* (6c) wins the competition for insertion, resulting in *la house*. Note that an important question in this case is whether an English root will merge an *n* with the *u*[+FEM] feature or that with the *u*[-FEM] feature. This will be addressed below in Section 3.1.2.

We turn now to the case of *the casa*, where a Spanish root merges with an English *n*. Because *n* has no gender feature, theme node insertion rule (4d) applies and so nothing is inserted. As such, the bound root cannot get a theme vowel to complete its phonological form. In order to get a theme node inserted and thus a theme vowel, the Spanish root has to merge with a Spanish *n* with a gender feature; however, if that is the case, then the D terminal node will also have a feminine feature by virtue of agree, and so the feminine Spanish determiner *la* (6c) is the VI that

[10] It is assumed that language membership is a psychologically emergent property that allows for the functional separation of a bilingual's two languages, such that activation (of one, the other, or both) is managed via cognitive control mechanisms. Importantly, the computational system does not see language membership as a formal feature.

wins the competition for insertion, not *the* (6f). As such, *the casa* is blocked via the subset principle.

6. List 2 VIs for Spanish (a–e) and English (f) competing for D[DF.SG.F]
   a. F.PL.DF → las
   b. PL.DF → los
   c. **F.DF → la**
   d. M.DF → el
   e. DF → lo
   f. DF → the

It is therefore not so much the case that mixed DPs consisting of an English D and a Spanish N violate any morphosyntactic requirements of the computational system. Rather, they are blocked by the availability of a more highly specified Spanish VI. Interestingly, such an account is not incompatible with the grammatical features spell-out hypothesis of Liceras et al. (2008), which also assigns a preference to the more grammaticalized functional element, but leaves room for the acceptability of the less grammaticalized option (the English D) in situations where it is more efficient with respect to processing resources.

3.1.2. The Analogical Criterion

As it is assumed that Spanish has two possible *n*s, ($n_{u[\text{-FEM}]}/n_{u[\text{+FEM}]}$), there is an option with respect to which Spanish *n* an English root will be licensed under in mixed DPs. Indeed, there are three possible outcomes for root licensing. First, it could be entirely arbitrary; because such English roots are combined spontaneously with Spanish elements and not repeatedly, it is unlikely that licensing conditions exist for these roots in the same way that they do for Spanish roots, whose licensing conditions would have been established during lexical acquisition via cues in the input. This would predict that all the options presented in (7) would be equally acceptable. Second, there could be a preference for licensing all English roots under the $n_{u[\text{-FEM}]}$ by default, leading to preferences for (7b) and (7d). Finally, there could be a preference for licensing English roots under the *n* of their translation equivalent, leading to preferences for (7a) and (7d).

7. a. *la* house
   the.DF.F house (c.f. house = casa.F)

   b. *el* house
   the.DF.M house

   c. *la* book
   the.DF.F book (c.f. book = libro.M)

   d. *el* book
   the.DF.M book

This third option is known as the "analogical criterion", and Liceras et al. (2008) found strong evidence for this with the L1-Spanish L2-English sequential bilinguals that they tested using an acceptability judgment task. Data from the corpora, however, suggest that the analogical criterion may be more of a tendency than a rule. For example, in the Spanish–English corpus-based studies presented in Table 1, it is clear that an English N with a masculine translation equivalent almost categorically appeared with a masculine Spanish D; whereas an English N with a feminine translation equivalent had considerably more optionality in terms of the gender of the Spanish D with which it appeared.

This suggests that, in production, analogical gender is applied to all English Ns if it can be easily retrieved from memory. If it is not retrieved, then masculine gender is applied as default.[11]

Table 1. The analogical criterion in Spanish–English corpora.

Study	Gender of Translation = M		Gender of Translation = F	
	Matching D	Mismatching D	Matching D	Mismatching D
	e.g., *el book*	e.g., *la book*	e.g., *la house*	e.g., *el house*
Poplack et al. (1982)	97%	3%	78%	22%
Clegg and Waltermire (2009)	96%	4%	21%	79%
Jake et al. (2002)	97%	3%	19%	81%

D: determiner.

Experimental results from a study employing eye-tracking with the visual world paradigm support this pattern. Kroff et al. (2016) found that lexical access to English Ns with a feminine Spanish translation was facilitated by the presence of a feminine D, indicating that this was a useful cue during comprehension. This could be because feminine Ds rarely co-occur in the input with English Ns that have a masculine translation (i.e., only about 3–4%). On the other hand, they found that lexical access to English Ns that have a masculine translation was not facilitated by the presence of a masculine D, suggesting that the masculine D is not a useful cue. Again, this is perhaps because masculine Ds co-occur in the input with English Ns that have either a masculine translation (96–97%) or a feminine translation (22–81%), which appears to be due to the default masculine status.

3.2. Predictions for French–English

French and Spanish have very similar gender systems, with two grammatical genders, masculine and feminine, that are assigned arbitrarily to inanimate nouns. It will therefore be assumed that inanimate roots in French are also licensed under either $n_{u[\text{+FEM}]}$ or $n_{u[\text{-FEM}]}$. Unlike Spanish, however, French has no declension class suffixes. As such, roots in French are typically free as opposed to bound. While French has both feminine and masculine singular definite determiners, as does Spanish, there are some differences with respect to the VIs that French–English bilinguals have available for insertion into D, as shown in (8). Importantly, the feature specification of *le* (8c) is the same as that of *the* (8d), because unlike in Spanish, there is no other gender-unspecified determiner (i.e., Spanish *lo*). These properties of French sometimes lead to different predictions with respect to the mixed DP asymmetry and the analogical criterion.

8. List 2 VIs for French (a–c) and English (d) competing for D[DF.SG.F]
   a. PL.DF → les
   b. **F.DF → la**
   c. DF → le
   d. DF → the

3.2.1. The Mixed Determiner Phrase Asymmetry

Since French nouns do not have obligatory affixes that are dependent on a gender feature (c.f. declension class suffixes in Spanish), this means that there is nothing prohibiting French roots from combining with a genderless English *n*. As such, one of the main predicted differences between French–English and Spanish–English mixed DPs is that combinations of an English D and a French N,

[11] Note that, of the three groups, the bilinguals in Poplack et al.'s (1982) corpus appear to be much better at retrieving the analogical gender of English Ns, as they do so for 78% of those with a feminine translation. It can be assumed that they do so at a similar rate for those with a masculine translation, though this is obscured by the correspondence between analogical gender and the default gender.

as in (9a), are expected to be grammatical, as well as those involving a French D and an English N, as in (9b). In other words, no mixed DP asymmetry is expected.

In support of this prediction is data from a Norwegian–English mixed language corpus (Alexiadou et al. 2015; Riksem 2017), which suggest that the mixed DP asymmetry is not present in this language pair despite their having a similar underlying gender feature asymmetry; both DPs in (10) are attested (see (Burkholder et al. 2017) for more details). Importantly, Norwegian roots, like French roots and unlike Spanish roots, are free.[12]

9.	a.	the	*maison*
		the.DF	house
	b.	*la*	house
		the.DF.F	house

10.	a.	*det*	*andre*	crew-*et*
		the.DF.N	other	crew-DF.SG.N
	b.	the	*by*	
		the.DF	city	

3.2.2. The Analogical Criterion

Because it is assumed that French and Spanish have *n*s with the same gender features ($n_{u[-FEM]}/n_{u[+FEM]}$), this leads to the prediction that French–English, like Spanish–English, has an option with respect to which French *n* an English root will be licensed under in mixed DPs. There is indeed evidence of this from French–English corpus data (Poplack et al. 1982). After controlling for biological gender, it was determined that 78% of English Ns with masculine translation equivalents took analogical gender (e.g., 11d), compared to only 42% of those with feminine translation equivalents (e.g., 11a). While this pattern is very similar to what they found in the Spanish–English corpora (see Table 1), it is notable that the application of analogical gender for masculine Ns was not as categorical in this French–English corpus. The authors note that this increased variability could be due to a higher number of cases where it was unclear to them which analogical gender should be assigned, which may have been partly related to their own lesser familiarity with French compared to Spanish.

11.	a.	*la*	house
		the.DF.F	house (c.f. house = maison.F)
	b.	*le*	house
		the.DF.M	house
	c.	*la*	sun
		the.DF.F	sun (c.f. sun = soleil.M)
	d.	*le*	sun
		the.DF.M	sun

[12] Note that even though Norwegian nouns can take definite suffixes that have been argued to mark declension class rather than gender (Lohndal and Westergaard 2016), these nouns can surface without any suffixes, such as in indefinite constructions. This is evidence that Norwegian nouns involve free roots. It should also be noted that overt definite determiners are only present under certain conditions, such as when there is an adjective (e.g., see Julien 2005).

There is also evidence that L1-French speakers may apply different strategies than L1-Spanish speakers with respect to the analogical criterion. While Liceras et al. (2008) found strong evidence for the analogical criterion with their L1-Spanish L2-English sequential bilinguals, they found that L1-French L2-Spanish bilinguals (with English as a second L1) tended to prefer applying masculine as default, at least in Spanish–English mixed DPs.

3.3. Impact of Age of Acquisition

The DM-based analysis suggests that there may be different mixed DP patterns for simultaneous versus late sequential bilingual acquisition. For bilinguals acquiring two L1s simultaneously, it could be expected that new Spanish/French roots encountered in nominal contexts would become licensed under the appropriate gender-bearing *n* (either $n_{\text{u[+FEM]}}$ or $n_{\text{u[-FEM]}}$), whereas new English roots would become licensed under the plain *n*. As such, in language mixing contexts, both gender-bearing *n*s and the plain *n* would be available to these bilinguals when selecting the numeration. This is what has been assumed in the above theoretical analyses. In contrast, if it is assumed that after a particular age, functional features not present in the L1 are no longer available to L2 learners (e.g., Hawkins and Chan 1997), then it is possible that late L2 learners may not have all three *n*s in their repertoire.

For late L1-Spanish/L1-French L2-English bilinguals, this may mean that they do not have a plain *n*. If so, this predicts that both English and French/Spanish roots can only merge with a gender-bearing *n*. This has several possible consequences.

First, in Spanish–English mixed DPs, when Spanish roots merge with a plain English *n*, this already creates issues due to declension classes. As such, not having an English *n* in the repertoire of abstract morphemes does not change anything for Spanish roots. On the other hand, this possibility would mean that English roots (e.g., √HOUSE) appearing in Spanish–English mixed language contexts would always have to merge with either $n_{\text{u[+FEM]}}$ or $n_{\text{u[-FEM]}}$. This predicts that such English Ns should only appear with Spanish Ds. This is because *el* and *la* are both more specified than *the*, meaning that they would be preferentially chosen at vocabulary insertion if available, as per the subset principle.[13] There is actually good evidence for this prediction. In the Spanish–English corpora presented in Table 1, full English DPs (e.g., *the book*) are extremely rare compared to those consisting of a Spanish D and an English N (e.g., *el book*).

Second, for DPs with French roots appearing in mixed language contexts, not having a plain English *n* could mean that D will always have a gender feature obtained from agreeing with a gender-bearing French *n*. As such, mixed DPs with a feminine noun, such as *the maison,* should be blocked in favour of unmixed DPs, such as *la maison,* because the more specified French VI *la* (8b) will always win over the English *the* (8d). However, because *le* (8c) has the same specification as *the* (8d), optionality is expected for mixed DPs such as *the soleil* and unmixed DPs such as *le soleil* (11a). In other words, a partial mixed DP asymmetry would actually be expected for a group of late L1-French L2-English bilinguals, whereby only feminine French Ns cannot appear with English Ds.

Finally, for English roots appearing in French–English mixed language contexts, no preference would be expected for a French D to be inserted. Consequently, full English DPs would be considered acceptable, unlike in Spanish–English. This is because *the* is assumed to have the same feature specification as *le*, and because English roots can be merged with $n_{\text{u[-FEM]}}$ by default regardless of the gender of its translation.

With respect to late L1-English L2-French/L2-Spanish bilinguals, they may not have gendered *n*s in their repertoire of functional features. These bilinguals would also not likely have any gender features specified on the VIs for determiners, and therefore Spanish *la* and *el*, as well as French *la* and

---

[13] Note that *el/la* would not be available in unilingual English mode, allowing for full English DPs in that case.

le, would all have the same feature specifications as English *the*.[14] This would predict that, in mixed language contexts, English roots would be acceptable with all three of the available determiners, and as such, there would be considerable optionality. This prediction is consistent with the acceptability judgment data for L1-English L2-Spanish bilinguals presented by Liceras et al. (2008), which indicated that English nouns were found to be equally acceptable with either Spanish determiner, regardless of the gender of the noun's Spanish translation equivalent.

For such late L1-English bilinguals, it would also be expected that they would not represent Spanish declension classes in the same way as L1-Spanish speakers, because they would not have the gendered *n*s on which declension class suffixes are dependent. As such, this also predicts that such bilinguals would not display mixed DP asymmetry. Indeed, the L1-English bilinguals in the Liceras et al. (2008) study show a strong preference for mixed DPs consisting of an English D and a Spanish N (e.g., *the casa*) over those consisting of a Spanish D and an English N (e.g., *la house*), and rated them considerably higher than the L1-Spanish group.

3.4. Effects of Language Dominance and Proficiency on Processing

Experimental data collected using methods other than acceptability judgment tasks have actually found that there may be greater language switch costs associated with mixed DPs such as *la house*, for example using a picture naming task (Fairchild and Van Hell 2017), self-paced reading (Litcofsky and Hell 2017), as well as event-related potentials combined with rapid serial visual presentation (idem.). In Fairchild and Van Hell's picture naming task, it was proposed that this cost was associated with difficulty retrieving the appropriate gender-marked Spanish D rather from difficulty retrieving the English N, as even though participants were L1-Spanish, they were currently dominant in English. They hypothesize that balanced or Spanish-dominant bilinguals might not exhibit the same processing asymmetry. However, in the experiments conducted by Litcofsky and Hell (2017), also with L1-Spanish L2-English bilinguals, they found that that switch costs were related to language dominance such that they were greater in the dominant-to-weak direction, regardless of the L1. In other words, it was Spanish-dominant bilinguals who had greater processing costs when reading mixed DPs such as *la house*. This is what was found in the Fairchild and Van Hell study for English-dominant bilinguals. These different results could reflect the fact that the former study tested production, while the latter tested reading.

Relative language proficiency may also impact the extent to which the analogical criterion is applied in language use. In order for an English noun to surface with the gender of its translation equivalent, the mental representation of that translation must be accessed from memory. The ease with which this is done may depend on the way in which translations are stored, which may in turn depend on the bilingual's language proficiency. For example, one of the most prominent models of the bilingual mental lexicon, Kroll and Stewart (1994) revised hierarchical model (RHM), suggests that L1 words are learned in direct association with their concepts, whereas L2 words (i.e., in sequential L2 acquisition) tend to be learned by directly associating them to their L1 translation. As such, L2 links to concepts are only indirect, mediated through the L1. As learners become more proficient in their L2, the direct links between L2 words and concepts become stronger. Assuming this type of psycholinguistic model, it is possible that unbalanced sequential bilinguals who have these strong lexical links between L2 and L1 words might have a stronger tendency to transfer the licensing conditions of the L1 (French or Spanish) root to that of the L2 (English) root than would balanced simultaneous bilinguals who do not have such strong associations between L2 and L1 lexical representations.

---

[14] That which would ensure that French/Spanish nouns typically appear with the appropriate gender-marked determiner is not licensing conditions under the appropriate *n*, but proceduralized lexical associations between nouns and the particular determiners with which they frequently co-occur in the input.

The following section presents a study designed to test some of the predictions and issues presented in the previous sections. Two primary research questions are addressed: first, whether there is evidence of a French–English mixed DP asymmetry; and second, how the analogical criterion is applied in mixed DPs with an English lexical root. In particular, the impact of language background factors such as AoA and language dominance will be investigated by comparing the results of simultaneous bilinguals to those of late sequential L1-French L2-English bilinguals. The specific hypotheses will be presented in Section 4.4.

4. Methods and Materials

The experiment combines two methodologies: an acceptability judgment task (AJT) and self-paced reading (SPR). The SPR paradigm was chosen as it was expected to elicit more implicit grammaticality assessments of targeted structures. This paradigm is typically used to measure the processing load of words during reading comprehension, where reading times (RTs) are correlated with lexical properties such as frequency and word length (Haberlandt 1994). Importantly, the processing load can also be increased due to the presence of a morphosyntactic violation (e.g., subject–verb agreement error). Indeed, previous studies using this methodology have demonstrated that morphosyntactic violations lead to increased RTs at the point of the anomaly (e.g., De Vincenzi et al. 2003). This effect has also been correlated with the P600 event-related potential (ERP) component, which indexes morphosyntactic violations (Ditman et al. 2007). In the context of language mixing, it is expected that RTs in an SPR will be longer at points of illicit or less preferred switches between languages. Increased RTs, however, may also reflect increased processing effort due to factors other than the implicit detection of grammatical anomalies. For example, it could be more costly to switch to a less activated language (Litcofsky and Hell 2017). Having participants also judge the acceptability of each sentence allows for the results of two dependent measures to be directly compared, and for the results of this study to be more easily related to those of previous studies that employed an AJT.

4.1. Participants

A total of 49 French–English bilinguals participated in this study. Of these, 10 were excluded for the following reasons: (1) not being a native speaker of just French, or French and English; (2) having early exposure to a language other than French or English; (3) having a high level of proficiency in a language other than French or English; and (4) not being a self-described language mixer.[15] This was determined by their responses to a detailed language background questionnaire (Sabourin et al. 2016) and a language-mixing questionnaire (adapted from (Byers-Heinlein 2013)). One additional participant was later excluded for rating sentences significantly lower than all other participants (see Section 5). The remaining 38 participants (30 female) were all between the ages of 18 and 35 (mean = 20.5, SD = 4.1). They were divided into groups based on their age of immersion (AoI) in English, which was operationalized at the age at which their daily exposure to English reached 20%. Table 2 shows the linguistic profiles of each of the three resulting groups: simultaneous bilinguals (AoI = 0), early L2 learners (AoI < 7), and late L2 learners (AoI > 7).

[15] Being a native speaker was defined as having acquired that language from a primary caregiver during infancy (i.e., between birth and the age of two) in an immersive environment (daily exposure > 20%).

Table 2. Language background of participants. Means are shown with standard deviations (SD).

	Simultaneous	Early L2 English	Late L2 English
	(*n* = 18)	(*n* = 7)	(*n* = 13)
1. Age of Acquisition (AoA)			
Age of first exposure to French	0 (0)	0 (0)	0 (0)
Age of Immersion (AoI) in French	0 (0)	0 (0)	0 (0)
Age of first exposure to English	0 (0)	1.3 (1.3)	6.9 (3.3)
Age of Immersion (AoI) in English	0 (0)	3.0 (1.4)	14.6 (6.2)
2. Language exposure < age 2			
French	55% (22%)	87% (6%)	91% (11%)
English	47% (24%)	13% (5%)	7% (11%)
3. Current self-reported language proficiency[16]			
French Total Proficiency/29	27.8 (1.7)	28.4 (1.1)	27.5 (4.0)
English Total Proficiency/29	28.0 (2.5)	27.7 (1.1)	19.3 (6.7)
French Reading Proficiency/6	5.9 (0.2)	5.9 (0.4)	5.7 (0.6)
English Reading Proficiency/6	5.8 (0.3)	6.0 (0.0)	4.5 (1.4)
4. Frequency of Language Mixing[17]			
General mixing	2.5 (1.4)	1.9 (1.1)	2.6 (1.6)
English to French switches	4.1 (1.6)	3.1 (2.2)	3.8 (1.5)
French to English switches	3.3 (1.3)	2.1 (1.1)	3.1 (1.6)
English borrowing	2.4 (1.2)	2.1 (1.1)	2.0 (1.3)
French borrowing	2.6 (1.5)	1.6 (0.5)	3.3 (1.9)

4.2. Stimuli

4.2.1. Critical Stimuli

Forty-eight nouns were selected to form part of the critical DPs. All were mono-morphemic and commonly known non-cognates, and approximately half had feminine gender in French (23/48). Nouns beginning with vowels were generally excluded in order to avoid élision between the definite determiner and the noun, which neutralizes the overtly marked gender of the determiner. No established loan words (such as the English words *job, fun*, which are commonly used in French) were included. The properties of these critical nouns in both languages are compared in Table 3.

A two-way repeated measures analysis of variance (ANOVA) was conducted for both word frequency and word length, using Language (English vs. French) as the within-item independent variable, and Gender (masculine vs. feminine) as the between-item independent variable. For word frequency, there were no significant main effects or interaction, indicating that word frequency was matched across language and gender. In terms of word length, neither Gender nor the interaction between Language and Gender were determined to be significant; however, the main effect of Language was determined to be significant ($F(1,46) = 10.497$, $MSE = 10.109$, $p = 0.002$), indicating that French nouns were significantly longer on average than English nouns overall. Little impact, however, is expected, as the magnitude of the difference is quite small (0.7 graphemes), and the words in French and English are all relatively short (both have a maximum of seven graphemes).

[17] On a scale of 1–7, where 1 means 'very true' and 7 means 'not true at all' with respect to statements such as "In general, I often mix English and French" and "I often start a sentence in English then switch to French".

[16] Measured with respect to five skill areas—reading, writing, listening, speaking, and pronunciation—with the first four given on a scale of 1–6 (very low to native) and the last on a scale of 1–5 (very accented to native).

Table 3. Lexical properties of critical nouns.

		Mean	SD	Min.	Max.
Word Frequency[18] (words per million)	**French**	72.0	123.0	0.7	570.3
	English	65.7	87.8	2.2	514.0
Word Length (graphemes)	**French**	5.4	1.1	4	7
	English	4.7	1.0	3	7

To create the critical DPs, nearly all of these nouns were paired with a singular definite determiner (43/48). The remaining five appeared with a singular indefinite determiner, either due to a better fit with the sentence context, or to avoid élision, as mentioned above. No differences were expected between the results for definite and indefinite determiners, as their features are comparable in both languages. One sentence was created for each of the 48 critical DPs, and three dependent variables were manipulated in order to produce six conditions for each sentence (see Table 4 for an example and the Supplementary Materials for a complete list). First, the language of the determiner in the critical DP was either French or English. Second, the language of the noun was either the same as the determiner (unmixed DP), or the opposite (mixed DP). Finally, for conditions with French determiners, the gender of the determiner either matched or mismatched the gender of the noun or its translation equivalent. Sentences were also coded for the gender of the noun or its translation equivalent.

The critical DP was always the third DP of the sentence, where the first was a language neutral proper name, and the second was a fully switched DP. This was done to establish the language mixing of the sentence before the critical DP was encountered, potentially reducing the processing cost of the critical language switch by priming it with a previous language switch. Also, the words preceding the critical determiner were always in the same language as that determiner, and the words following the critical noun were always in the same language as that noun. This was done in order to ensure that the language switch between the determiner and the noun was more along the lines of a multi-word "code-switch" as opposed to a single word "nonce borrowing" or insertion. Finally, the critical noun was never the last constituent in the sentence in order to avoid any sentence wrap-up effects.

Table 4. Conditions for critical stimuli (critical Determiner Phrase (DP) bolded). *

D	DP Type	Congruence	Example
FR	Mixed	Matched	Paul bought trois livres pendant **le month** of May.
		Mismatched	Paul bought trois livres pendant **la month** of May.
	Unmixed	Matched	Paul bought trois livres pendant **le mois** de mai.
		Mismatched	Paul bought trois livres pendant **la mois** de mai.
EN	Mixed	N/A	Paul a acheté three books during **the mois** de mai.
	Unmixed	N/A	Paul a acheté three books during **the month** of May.

* All sentences mean 'Paul bought three books during the month of May'.

4.2.2. Filler Items

Three types of filler sentences were included as stimuli (see Table S1). First, 15 sentences were created such that they contained a morphosyntactic mismatch between constituents belonging to the same language, such as number and verb tense violations. A control condition was created for these sentences by repairing the mismatch. This type of filler was included in order to determine whether both measures of the experimental task, ratings and RTs, were able to elicit the expected distinctions between grammatical and ungrammatical sentences. Similarly, another 15 sentences were created

[18] French and English word frequencies were determined using the Lexique 2 database (New et al. 2004) and the SUBTlexUS database (Brysbaert and New 2009), respectively. As they are very similar in nature—both based on subtitle corpora—it is possible to compare frequencies across languages.

such that they contained a type of language switch that is commonly considered in the literature to be ungrammatical, such as between an auxiliary and a main verb, between a pronoun and a verb, or between a noun and a post-posed adjective (which produces a word-order conflict). Again, a control condition was created for these sentences by repairing the mismatch. This type of filler was included in order to determine whether language switch violations were treated in the same way as the within-language morphosyntactic violations with respect to both dependent measures. Finally, 10 mixed-language sentences were created that were completely grammatical, with language switches occurring at uncontroversial sites, such as between a noun and a verb, or between a noun and a prepositional phrase.

4.3. Procedure

All participants were recruited through the University of Ottawa's Integrated System of Participation in Research (ISPR), and were awarded one point towards their course grade. All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Social Sciences and Humanities Research Ethics Board of the University of Ottawa (REB# 02-16-26). The entire testing session lasted approximately 60 min for each participant. Participants were first asked to complete the consent form, language background questionnaire and language-mixing questionnaire. They were then seated in a sound-attenuated room in front of a computer screen for the experimental task.

Participants were instructed that they would be reading mixed-language sentences one word at a time on the computer screen. They were told that the words of each sentence would be first covered in dashes, and to see each successive word, they were to press the space bar (as per the "moving window paradigm" of SPR). After each sentence, they would be asked to rate how acceptable it was to them on a scale from 1 to 4 by pressing the corresponding number key on the keyboard. They were instructed that a rating of "1" meant that the sentence "sounded terrible, it was not acceptable"; a rating of "2" meant that it "sounded pretty bad"; a rating of "3" meant that it "sounded ok, but not very natural; and a rating of "4" meant that is "sounded good, it was perfectly acceptable". This 4-point scale was visible on the computer screen each time they were asked to make a rating, with the "1" marked as "Terrible" and the "4" marked as "Fine".

Each participant saw a total of 88 experimental sentences, presented using the Linger software program (version 2.94, http://tedlab.mit.edu/~dr/Linger), plus eight practice sentences. All words were displayed in black 20-point font on a light grey background. Spaces between words were not covered in dashes, and a period marked the end of each sentence. They saw each critical (n = 48) and filler (n = 40) sentence once, counterbalanced across conditions such that they saw an equal number of sentences in each condition. Sentences were presented in a semi-randomized order, and separated into six blocks such that each block began with a filler sentence and no more than two sentences of the same condition appeared consecutively. Participants were given a self-regulated break between each block. One third of both critical and filler sentences were followed by a yes/no question that was entirely in English.

4.4. Hypotheses

The first set of hypotheses that were tested pertained to the mixed DP asymmetry. Based on the arguments presented in Section 3.2.1, it was expected that there should be no directional asymmetry for language switches between English and French determiners and nouns. This was expected to be the case at least for the simultaneous bilinguals. In contrast, and as discussed in Section 3.3, late L1-French L2-English bilinguals were expected to possibly show a partially-mixed DP asymmetry; DPs such as *the soleil*$_{.M}$ should be acceptable, but not those such as *the maison*$_{.F}$. Finally, this group of late bilinguals may also show reduced processing costs for mixed DPs consisting of an English D and

a French N, because this language switch goes in the direction of the weak to the dominant language (see Section 3.4).

The second set of hypotheses tested predictions regarding gender assignment for English roots in mixed DPs, and the extent to which the analogical criterion applied. Based on the arguments presented in Section 3.2.2, it was expected that there would be considerable variability for English roots whose translation equivalents are licensed under $n_{u[+FEM]}$ (i.e., sentences containing *le house* and *la house* should have similar acceptability ratings and similar RTs for the critical nouns), but no variability for roots whose translation equivalents are licensed under $n_{u[-FEM]}$ (i.e., only *le soleil* should be acceptable, not *la soleil*).

It was further expected that the late bilinguals would have a greater tendency to apply analogical gender to English nouns whereas the simultaneous bilinguals were expected to possibly show a greater tendency to apply the masculine gender as default. This is because the direct lexical connections between translations in the L2-to-L1 direction translation equivalents were expected to be stronger for the late bilinguals, as per the RHM, and thus easier to activate (see Section 3.4).

5. Results

The accuracy of participants' responses to the comprehension questions was first calculated in order to determine whether they were paying sufficient attention to the reading task. It was determined that all participants responded with high accuracy to the comprehension questions, with a mean of 95.9% (SD = 4.8%), and a range of 77.4–100%.

Acceptability ratings were collected for each sentence, and the mean and standard deviation were calculated for each participant across all conditions in order to identify any outliers. One participant's mean rating was indeed more than 2.5 standard deviations from the group's overall mean, and so was excluded from the data analysis.[19] For the 38 remaining participants, mean ratings were calculated for each participant by condition, and this was used as the first dependent variable.

RTs were time-locked to participants' button press on the critical word in each sentence, that is, the word where the language switch or morphosyntactic violation became apparent (or not, for their respective control conditions). Mean RTs were calculated for each participant, and data points that were 2.5 standard deviations above (4.9% of the data) or below (none) each participant's mean were replaced by their mean value. Mean RTs were then re-calculated based on this trimmed data for each participant across conditions in order to identify outliers; one participant's mean was greater than 2.5 standard deviations from the group mean (a simultaneous bilingual), and so this participant's RT data was excluded from the statistical analyses; however, their acceptability judgment data were still analyzed. For each of the remaining 37 participants, mean RTs were calculated by condition, and this was used as the second dependent variable.

5.1. Grammaticality Violations in Filler Items

The two types of filler items involving within-language morphosyntactic violations and uncontroversial language switch violations were analyzed using a three-way repeated measures ANOVA for each dependent variable (ratings and RTs). Filler Type (language switch violation vs. morphosyntactic violation), Grammaticality (violation vs. control), and Group (simultaneous vs. early vs. late) were the independent variables.

Results for the ratings revealed a significant main effect of Grammaticality ($F(1,35) = 188.437$, $MSE = 32.352$, $p < 0.001$), indicating that both types of violations received significantly lower ratings than their grammatical counterparts. There was also a significant main effect of Filler Type

[19] This participant was the oldest of the group (age at testing = 35), and was also one of the less frequent language mixers (her reported overall tendency to mix language was rated as 5/7, where 7 is "never"). Her low ratings in the experiment (mean = 1.31/4, SD = 0.63) suggest that she did not find the mixed language sentences natural, something that was not typical of the rest of the group.

(F(1,35) = 6.548, MSE = 1.480, p = 0.015), which was likely driven by the significant interaction between Filler Type and Grammaticality (F(1,35) = 7.772, MSE = 1.585, p = 0.009). This interaction indicates that, while both types of violations were rated significantly lower than their grammatical counterparts, as confirmed by planned paired-sample t-tests (t(37) = 11.963, $p < 0.001$; and t(37) = 7.969, $p < 0.001$, respectively), the magnitude of this difference was significantly larger for the within-language morphosyntactic violations (mean difference = 1.25 compared to 0.73). This likely reflects the fact that the participants have more metalinguistic knowledge of grammatical errors than they do of language switch violations. The main effect of Group did not reach significance (p = 0.897), nor was its three-way interaction with Filler Type and Grammaticality, suggesting that all groups showed similar effects.

Results for RTs mirrored those of the ratings. There was a significant main effect of Grammaticality (F(1,34) = 22.750, MSE = 527,233.974, $p < 0.001$), indicating that both language switch violations and morphosyntactic violations were read significantly more slowly than their grammatical counterparts. The main effect of *filler type* was only marginally significant (p = 0.067), as was the interaction between Filler Type and Grammaticality (p = 0.093). This interaction suggests that, while both within-language morphosyntactic violations and language switching violations were read significantly more slowly than their grammatical counterparts, as confirmed by planned paired-sample t-tests (t(36) = 5.590, $p < 0.001$; and t(36) = 3.685, p = 0.001, respectively), the magnitude of this difference was marginally larger for the within-language morphosyntactic violations. There was also a significant main effect of Group (F(2,34) = 5.550, MSE = 787,180.376, p = 0.008), indicating that the simultaneous bilinguals were overall slower than the early and late groups (p = 0.027 and p = 0.032, respectively).[20] There was, however, no three-way interaction between Group, Filler Type, and Grammaticality (p = 0.596), suggesting again that all groups showed similar effects of grammaticality on both types of filler sentence.

5.2. The Mixed Determiner Phrase Asymmetry

A three-way repeated measures ANOVA was conducted for each dependent variable (ratings and RTs) with Determiner Language (French[21] vs. English), DP Type (unmixed vs. mixed) and Group (simultaneous vs. early vs. late) as independent variables. Results for ratings (Figure 2a) revealed a significant main effect of DP Type (F(1,35) = 62.703, MSE = 17.851, $p < 0.001$), indicating that participants preferred unmixed over mixed DPs overall (mean difference = 0.69).[22] The main effect of Determiner Language was not significant (p = 0.343), and crucially, neither was the interaction between Determiner Language and DP Type (p = 0.457), indicating that there was indeed no preference for mixed DPs with French determiners over English determiners. The main effect of Group was not determined to be significant (p = 0.897), nor did it significantly interact with any other variable, suggesting that all groups shared the same pattern.

The patterns for ratings are mirrored in Figure 2b, which shows the mean RTs for critical nouns in the same four conditions. Indeed, statistical analyses confirm these patterns. There was a significant main effect of DP Type (F(1,34) = 22.799, MSE = 486,233.390, $p < 0.001$), indicating that participants preferred unmixed over mixed DPs (mean difference = 135 ms). The main effect of Determiner Language was not significant (p = 0.519), and again, neither was the interaction between Determiner Language and DP Type (p = 0.318). Interestingly, the numerical trend was in the opposite direction than would be expected for the Spanish–English mixed DP asymmetry, with the English nouns following French determiners (e.g., *la house*) being read more slowly than French nouns following English

---

[20] A similar effect was also found in another study conducted with the same population (Sabourin et al. 2014). The authors attributed this effect to the idea that simultaneous bilinguals have a larger, more intimately integrated bilingual mental lexicon from which to activate and retrieve words during processing.

[21] These comparisons were only made with the gender congruent conditions with French determiners. This decision was based on previous results (e.g., Liceras et al. 2008) that indicated that this is the likely overall preference. Results in Section 5.3 support this decision.

[22] This preference could be a consequence of the fact that sentences with mixed DPs had two language switch points, compared to sentences with unmixed DPs, which only had one language switch point.

determiners (e.g., *the maison*; mean difference = 59 ms). There was a significant main effect of Group (F(2,34) = 3.694, MSE = 422,112.890, p = 0.035) suggesting that the simultaneous bilinguals were again the slowest overall. The factor Group did not interact with any other factor, again suggesting that the patterns were similar across all three participant groups. Because, however, the hypotheses were made a priori regarding differences between the simultaneous and late bilingual groups, these were further investigated in order to identify any potential patterns or tendencies.

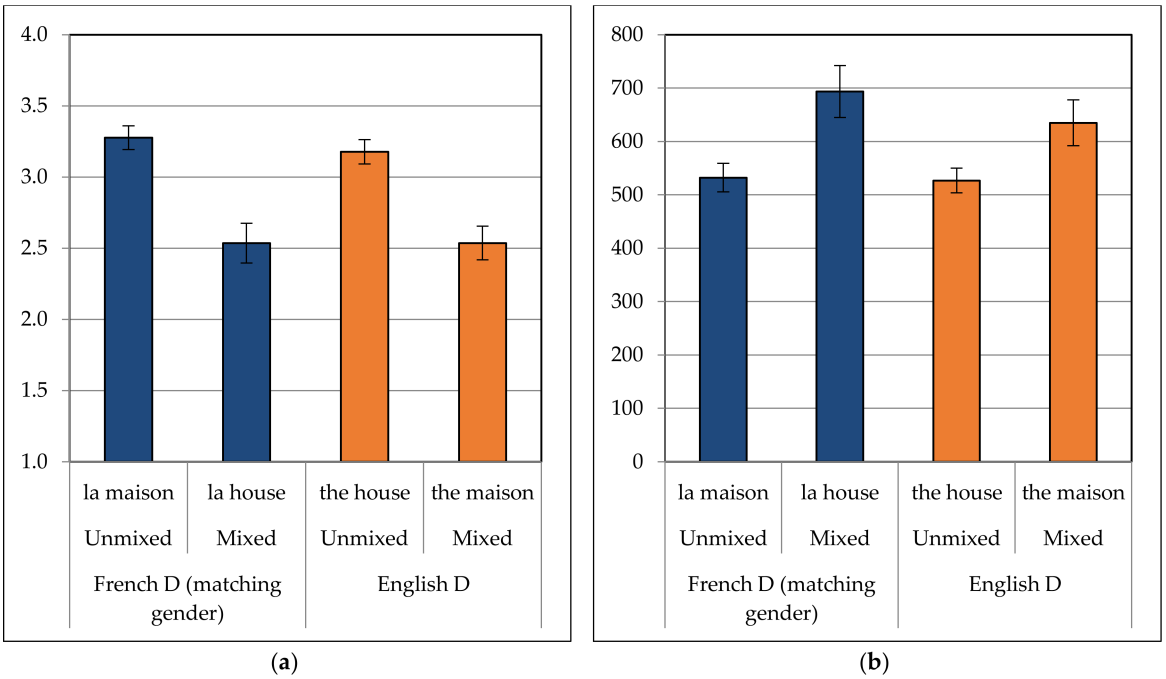


Figure 2. Results for the mixed determiner (D) phrase asymmetry: (**a**) Mean global acceptability ratings (N = 38); (**b**) Mean reading times of the critical noun in milliseconds (N = 37).

5.2.1. Between-Group Differences: *The Maison* vs. *The Soleil*

It was hypothesized that there might be a partial mixed DP asymmetry for the late bilinguals, whereby they might have only a dispreference for mixed DPs consisting of an English D paired with a feminine French N (e.g., *the maison*$_{.F}$) and not one for those paired with a masculine French N (e.g., *the soleil*$_{.M}$). In order to further investigate this possibility, a closer look was taken at the conditions including an English D. The ratings data are shown in Table 5.[23] The raw means alone suggest that both groups seem to have a slight preference for mixed DPs with a masculine N over those with a feminine N, and that the preference is greater for the late bilinguals (0.19 compared to 0.07). This pattern is consistent with the hypothesis, but of course the magnitudes of the differences are very small (though not nearly as small as for their unmixed counterparts, where the differences are very near null).

[23] Note that inspection of the RT data in the same conditions revealed no interpretable patterns whatsoever, and so this was not explored any further.

Table 5. Mean ratings for determiner phrases (DPs) with an English determiner.

	Unmixed DP			Mixed DP		
	Translation = F	Translation = M	Difference	Noun = F	Noun = M	Difference
	"the house"	"the sun"		"the maison"	"the soleil"	
SimBil	3.19	3.18	**0.01**	2.65	2.72	**0.07**
LateBil	3.21	3.21	**0.00**	2.33	2.52	**0.19**

An exploratory two-way ANOVA was conducted just for the late bilinguals, with DP Type (unmixed vs. mixed) and Noun Gender as the independent variables. While there was a marginally significant effect of DP Type (F(1,12) = 4.437, MSE = 165,983.554, $p = 0.057$), neither the effect of Noun Gender nor the interaction were significant ($p = 0.674$ and $p = 0.645$, respectively). Exploratory planned pair-wise comparisons confirmed that the difference in ratings between mixed DPs with English D did not differ significantly depending on the gender of the French N ($p = 0.840$). As such, there is no evidence here that late bilinguals demonstrate any type of mixed DP asymmetry, despite the faint indication of the expected pattern in Table 5.

5.2.2. Between-Group Differences: Processing Costs

It was also hypothesized that the group of late bilinguals (but not the simultaneous bilinguals) would show a greater processing cost for mixed DPs where the switch direction was from the stronger (French) to the weaker (English) language (e.g., *la house*) compared to the opposite direction (e.g., *the maison*). Table 6 presents the relevant data.

Table 6. Language switch costs: Reading times of the critical noun by condition.

	French Determiner			English Determiner		
	Unmixed	Mixed	Difference	Unmixed	Mixed	Difference
	"la maison/ "le soleil"	"la house"/ "le sun"		"the house"/ "the sun"	"the maison"/ "the soleil"	
SimBil	624.4	823.1	**198.7**	576.0	686.5	**110.5**
LateBil	448.8	611.3	**162.5**	491.3	616.3	**125.0**

Looking at the patterns alone, it appears that both groups have a greater language switch cost in the French-to-English direction compared to the English-to-French direction. An exploratory two-way ANOVA was conducted for each of the two participant groups, with DP Type (unmixed vs. mixed) and Determiner Language as the independent variables. While there was a significant effect of DP Type for both the late and simultaneous groups (F(1,12) = 10.073, MSE = 268,640.337, $p = 0.008$; and F(1,16) = 19.891, MSE = 406,207.887, $p < 0.001$, respectively), this factor did not significantly interact with Determiner Language for either group ($p = 0.439$ and $p = 0.227$, respectively). As such, there is no robust evidence of greater processing costs for either group, and contrary to the hypothesis, both groups shared the same pattern. The reason for this similar, but a non-significant pattern will become clearer in the analysis investigating the analogical criterion, presented in the next section.

5.3. The Analogical Criterion in Mixed Determiner Phrases with English Nouns

Looking only at experimental conditions with critical DPs containing a French D, a four-way repeated measures ANOVA was conducted for each dependent variable, with DP Type (unmixed vs. mixed), Gender Congruence (congruent vs. incongruent), Noun Gender (feminine vs. masculine), and Group (simultaneous vs. early vs. late) as independent variables.

Results for ratings (Figure 3) showed a significant main effect of Gender Congruence (F(1,35) = 95.228, MSE = 51.487, $p < 0.001$) as well as of Gender (F(1,35) = 14.794, MSE = 2.230, $p < 0.001$), but not of DP Type ($p = 0.150$). The interaction between DP Type and Gender Congruence was

significant (F(1,35) = 65.262, MSE = 23.218, $p < 0.001$), indicating that the difference between DPs with congruent vs. incongruent gender was much larger in unmixed French DPs compared to mixed DPs. Crucially, there was also a very strong trend towards significance for the three-way interaction between DP Type, Gender Congruence and Noun Gender (F(1,35) = 4.065, MSE = 0.510, $p = 0.051$). Planned pairwise comparisons indicated that for unmixed DPs there was a significant gender congruence effect for nouns with both masculine and feminine gender (t(37) = 9.520, $p < 0.001$; t(37) = 12.166, $p < 0.001$); however, for mixed DPs the congruence effect was significant only when the gender of the translation equivalent was masculine (t(37) = 4.545, $p < 0.001$), and not when it was feminine (t(37) = 1.357, $p = 0.183$). There was no significant main effect of Group ($p = 0.601$), nor did it participate in a significant interaction with the other three variables ($p = 0.138$).

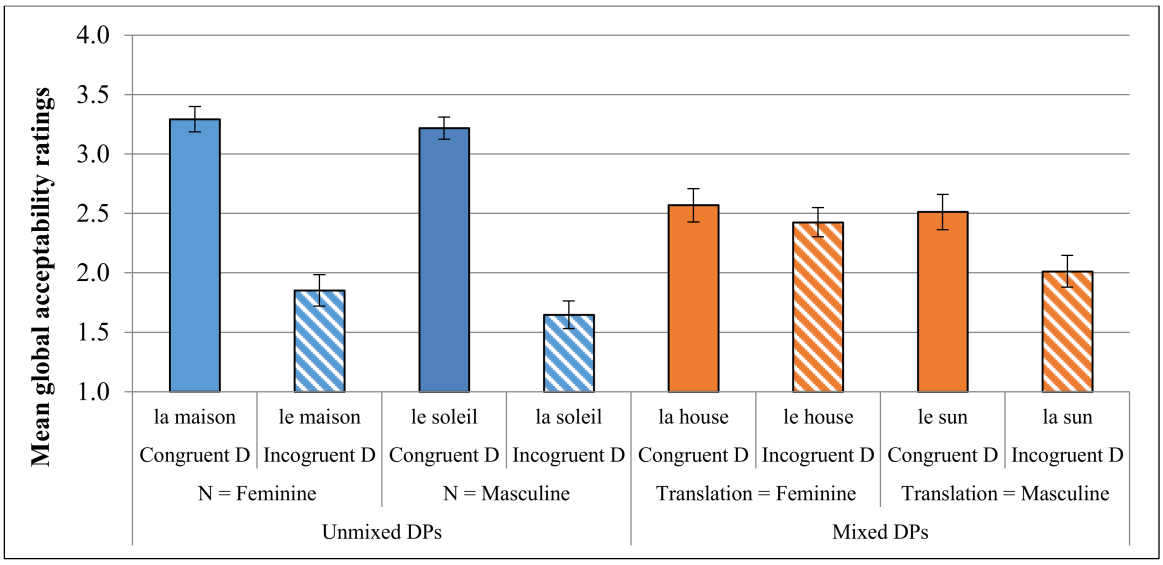


Figure 3. Ratings results for gender congruence in mixed determiner phrases with French determiner ($N = 38$).

Even though the factor Group did not interact with any of the other independent variables, separate two-way repeated measures ANOVA were conducted for the simultaneous and late bilinguals. This was done because it was hypothesized a priori that these two groups might have different preferences with respect to gender assignment for English nouns. Specifically, it was expected that Late L2-English bilinguals would have a greater tendency to apply analogical gender to English nouns in mixed DPs. As such, these two ANOVAs looked only at mixed DPs with French Ds, and the independent variables were Gender and Congruence. For the late bilinguals, the main effect of Congruence was significant (F(1,12) = 12.600, MSE = 1.942, $p = 0.004$), but the interaction between Gender and Gongruence was not ($p = 0.496$), suggesting that the effect of congruence was the same regardless of the translation's gender. Indeed, planned pair-wise comparisons indicated that nouns with both feminine and masculine translations were significantly preferred with analogical gender (t(12) = 2.758, $p = 0.017$ and t(12) = 2.485, $p = 0.029$, respectively). For the simultaneous bilinguals, the main effect of Congruence was significant (F(1,17) = 5.316, MSE = 2.067, $p = 0.034$) and the interaction between Gender and Congruence was marginally significant (F(1,27) = 4.048, MSE = 0.467, $p = 0.060$). Planned pair-wise comparisons indicated that the effect of congruence was only significant when the noun's translation equivalent had masculine gender (t(17) = 2.990, $p = 0.008$), but not when it had feminine gender (t(17) = 1.061, $p = 0.303$). This supports the hypothesis that the late bilinguals have a stronger preference for assigning analogical gender to English nouns in mixed DPs than do simultaneous bilinguals.

In terms of the four-way ANOVA using RT as the dependent variable (Figure 4), there was a significant main effect of DP Type (F(1,34) = 5.783, MSE = 247,550.451, $p = 0.022$), but no significant effect of Gender ($p = 0.386$) or of Gender Congruence ($p = 0.181$). There was, however, a significant interaction between DP Type and Gender Congruence (F(1,34) = 13.904, MSE = 386,115.394, $p = 0.001$). Planned pairwise comparisons suggested that this was again the result of a different pattern for unmixed compared to mixed DPs. For unmixed DPs, gender incongruence between a French N and a French D resulted in significantly slower RTs (t(36) = −4.729, $p < 0.001$), as expected; however, for mixed DPs, the RTs were significantly slower when the French determiner was congruent in gender with the English noun's translation equivalent (t(36) = 2.690, $p = 0.011$). This is the opposite of the expected result. While the three-way interaction between DP Type, Gender Congruence and Noun Gender was not significant ($p = 0.536$), planned pair-wise comparisons were conducted in order to further investigate the surprising pattern. The results of these comparisons suggest that it was only in cases where the English N's translation equivalent was feminine that the RTs were significantly slower in the congruent (e.g., *la house*) compared to the incongruent (e.g., *le house*) condition (t(36) = 2.661, $p = 0.012$). When the English noun's translation equivalent was masculine (e.g., *sun*), there was no significant gender congruency effect (t(36) = 0.147, $p = 0.884$). Further, this effect for English Ns with a feminine translation was statistically significant for the late bilinguals (t(12) = 2.478, $p = 0.029$), but only marginally significant for the simultaneous bilinguals (t(16) = 1.944, $p = 0.070$), despite the greater statistical power of the latter group due to the larger group size. These results will be interpreted and further discussed in the next section.

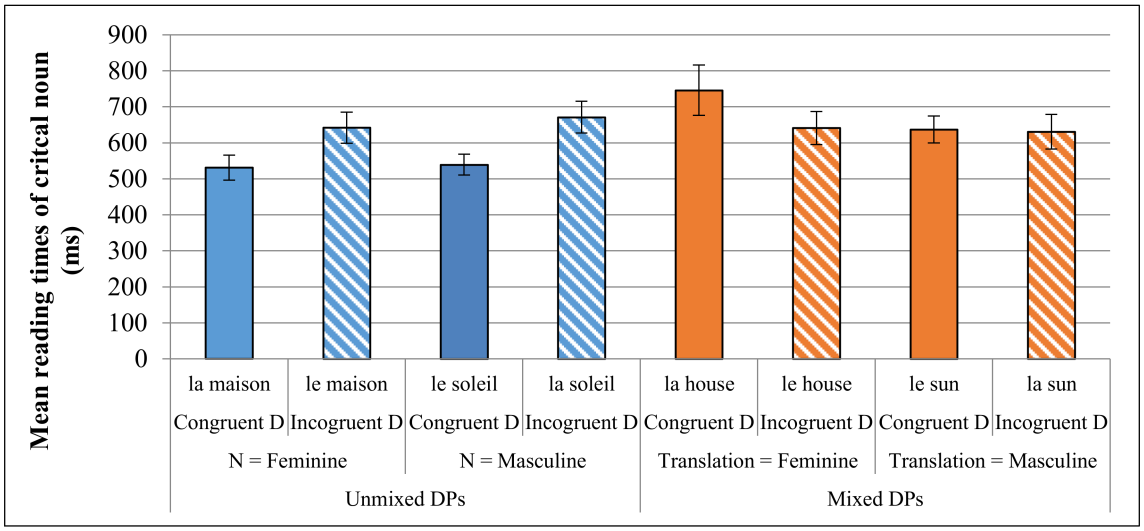


Figure 4. Reading time (RT) results for gender congruence in mixed determiner phrases with French determiner ($N = 37$).

6. Discussion

The two primary research questions that the experimental study sought to address were first, whether there was evidence of a French–English mixed DP asymmetry, and second, how the analogical criterion is applied in mixed DPs with an English lexical root. Secondary goals were to investigate the impact of language background factors such as AoA, and to take into account factors that affect processing, such as L2 proficiency and dominance.

With respect to the mixed DP asymmetry, the results for both dependent measures indicated that, overall, there was no preference for switching between a French D and an English N (e.g., *la house*) compared to the reverse (e.g., *the maison*). These results thus support the theoretically motivated hypothesis that there should be no mixed DP asymmetry for French–English.

It must, of course, be conceded that even in Spanish–English language mixing, where the mixed DP asymmetry is expected based on corpora data, AJTs have not provided corroborative evidence of its existence (Liceras et al. 2008, 2016). The experimental task used in the present study, however, differs from those of previous studies in several ways. First, several things were done in order to decrease the saliency of the mixed DPs being investigated—critical DPs were embedded in sentences that had other language-switch points; no unilingual sentences were included, in either critical or filler conditions; and several other types of within- and between-language morphosyntactic anomalies were included in filler items. Second, combining the AJT with SPR might also have had the benefit of reducing the participants' tendency to be prescriptive or to apply metalinguistic knowledge when making these judgments for two reasons—having comprehension questions after one third of the sentences encouraged participants to focus a significant part of their attention on meaning as opposed to just grammatical structure; and, because the sentences were presented only one word at a time and never in their entirety, participants could not go back and reread sentences. As such, they might not have the same opportunity to deconstruct and consciously analyze their grammatical structure, thus relying more on their intuition as opposed to their metalinguistic knowledge. It is perhaps the application of metalinguistic knowledge that allowed Spanish–English bilinguals to accept DPs such as *the casa* in previous studies; the DM-based analysis in Section 3.1 predicts that such DPs are not ungrammatical per se, but are blocked by the availability of those with a more highly specified determiner (i.e., *la casa*). Intuitively, bilinguals might not like such a DP, but if given time to consider it, they might not know why. Mixed DPs such as *the casa* might even be in their input if there are L1-English L2-Spanish speakers in their environment, for whom the mixed DP asymmetry is not expected. Of course, the only way to determine whether the more implicit methodology used here is capable of detecting a mixed DP asymmetry would be to use it to test a language pair where it is expected, i.e., Spanish–English. One indication that it might be able to do so is that there was robust evidence from the filler items that other uncontroversial language switch violations were indeed rated significantly lower and read significantly more slowly.

It was also hypothesized that late L1-French L2-English bilinguals may show a partial mixed DP asymmetry, specifically a dispreference for mixed DPs consisting of an English D and a feminine French N (e.g., *the maison*$_{F}$). This was hypothesized because such bilinguals may not have a plain *n* in their repertoire of abstract morphemes due to their late AoA, and because the English D *the* is expected to have the same feature specification as the French D *le*. No evidence, however, was found of any asymmetry. This suggests that late bilinguals may have a plain *n*, which could be because plain *n* is a subset of an *n* with a gender feature. Consequently, no "new" morphosyntactic feature actually needs to be acquired by the late learners. It is also possible that such a partial asymmetry could still be found using this methodology with participants having an even later age of L2 immersion, as others assume a later end to the sensitive period for morphosyntax, such as around puberty (e.g., Johnson and Newport 1989; DeKeyser 2000).

The group of late bilinguals was also hypothesized to show increased processing costs for mixed DPs consisting of a French D and an English N because this language switch goes from the dominant to the weak language, following the results of Litcofsky and Hell (2017). Unexpectedly, both the late and simultaneous groups demonstrated this pattern despite the latter group being very balanced. It was later determined that this was because mixed DPs such as *la house* were read significantly more slowly than any other mixed DP with a French D, even those which violated both the analogical criterion and default licensing under the masculine *n* (e.g., *la sun*). This result was also unexpected. It suggests, however, that participants might not have too much difficulty reading/assessing English nouns that appear following the masculine determiner (e.g., *le house*, *le sun*) because any English noun can appear with the default masculine determiner. Likewise, they seem to have little difficulty reading/assessing (and later rejecting) English nouns with masculine translations following a feminine determiner (e.g., *la sun*), suggesting that the issue is not in accessing the licensing conditions of the corresponding French translation equivalent. Rather, it seems that this may be a result of variability

in the input for English nouns with feminine translations (*le/la house*), which is shown to be a robust pattern in Spanish–English corpora (Poplack et al. 1982; Jake et al. 2002; Clegg and Waltermire 2009), and likely also for French (Poplack et al. 1982). Because participants do not have a strong intuition regarding such mixed DPs, this may result in a slower reading time when they encounter one due to the uncertainty. This appears to be particularly so with the late bilinguals, for whom the effect was the most statistically robust. Interestingly, this was also the group who tended to more strongly prefer the analogical criterion, as indicated by their results for acceptability ratings. While this preference was expected for this group, the RT results suggest that its cause is not related to the predictions of the RHM, whereby translations were expected to be more easily accessed by unbalanced sequential bilinguals due to their stronger direct lexical connections. If this were the case, they would have also shown extra processing effort when attempting to access the translation equivalent of masculine nouns following feminine determiners (e.g., *la sun*). Rather, the late bilinguals' preference for analogical gender might reflect a difference in their attitude compared to simultaneous bilinguals, such that the former group attributes greater importance to preserving "French" properties when mixing their languages, thus spending more time deciding what their preference is, and then applying analogical gender preferentially.

These results also suggest that the self-paced reading times of mixed DPs do reflect conscious, strategic preferences to a certain extent. This contrasts with the RTs for unmixed DPs, which really did seem to reflect modulations in processing time related to the grammaticality of particular constructions. Having participants rate the sentences with respect to their acceptability might reduce the benefits of the more implicit RT measure. As such, while it was argued that adding the SPR element to an AJT may decrease the strategic nature of the latter task, this might also increase the strategic nature of the former methodology. It would be interesting to see if the same patterns would persist when the acceptability judgment aspect is removed from the self-paced reading task, rendering it a simple reading task with comprehension questions focusing on meaning.

In a similar vein, the inclusion of within-language morphosyntactic violations may also have affected participants' acceptability ratings of sentences containing mixed DPs. First of all, gender agreement violations in unmixed French DPs are very salient errors to native speakers. As such, drawing participants' conscious attention to such errors may have increased their likelihood of applying the same "rules" to mixed DPs. This could have resulted in an increased tendency to assign analogical gender to English nouns in mixed DPs, particularly for the late bilinguals, who might be more prescriptive. Second, asking participants to use the same acceptability scale to assess within-language violations and potential language switch violations might have also reduced the effectiveness of the task. Bilinguals likely have significant metalinguistic knowledge of within-language morphosyntactic violations as a result of the explicit teaching and correction of grammatical errors, particularly in formal, written (unmixed) language. Language mixing, however, is typically used in informal oral contexts, and bilinguals rarely pay conscious attention to issues of grammaticality. Indeed, the results for the control conditions indicated that participants did have significantly stronger judgments for within-language violations than for language switch violations. This suggests that the scale was more sensitive to the former type of violation, and thus perhaps less able to detect subtle differences in the latter.

7. Conclusions

The mixed DP asymmetry of Spanish–English has been investigated from various perspectives over the years, yet there are still many conflicting views with respect to what it means for the underlying linguistic system acquired by bilinguals. The DM-based theoretical account presented in Section 3 assumes a somewhat different underlying system than do previous accounts.

Generally speaking, there are two important theoretical advantages to a DM approach to language mixing. First, because morphology is taken out of the lexicon, this means that the same syntactic processes apply both above and below the word-level (Grimstad et al. 2014; Alexiadou et al. 2015);

"words" are therefore derived from their morphological sub-components in the same way that sentences are derived from their lexical sub-components. As such, word-internal and word-external language mixing, which are often treated separately (e.g., Poplack and Dion 2012; see also Grimstad et al. 2014), can be accounted for, at least formally, using the same mechanisms. Secondly, because VIs from both languages compete for insertion into fully specified syntactic terminal nodes, this allows for optionality when VIs from the two languages have the same feature specifications. This leads to theoretically motivated predictions with respect to where optionality can be expected in mixed language utterances, and where there should be a categorical language preference. This optionality is one of the reasons why it is challenging to model language mixing, yet it is an aspect that DM is particularly well suited to handle.

Demonstrating that the DM framework can account for mixed language data also provides empirical support for the underlying assumptions of this model. In particular, the assumption that gender is not a root property, but a feature on *n*, is supported by the argument that bilinguals have options as to which *n* a translation can become licenced under. Further, the theoretical analysis of Spanish gender and declension class argued for by Kramer (2015) generated specific predictions regarding how this phenomenon was expected to manifest itself in French–English, and also predictions regarding how different paths of dual-language acquisition may affect the patterns. The study presented addressed some of the resulting hypotheses and provided positive evidence for this account, thus lending empirical support to Kramer's analysis. The two main findings were that there was no evidence for a mixed DP asymmetry for French–English bilinguals, as expected, and that late sequential bilinguals are more likely than simultaneous bilinguals to apply analogical gender to English roots in mixed DPs. The strength of conclusions based on the first result rests on the assumption that such an asymmetry could be found for Spanish–English using this particular experimental methodology. As such, testing this is an important direction for future work. While the second result was predicted based on a model of the bilingual mental lexicon, the reading time data suggest that this pattern might not actually reflect differing types of lexical access across languages for the two groups, but different attitudes and preferences. Conducting experimental studies using even more covert measures, such as event-related potentials and eye-tracking, could help to investigate these questions further.

Supplementary Materials: The following are available online at http://www.mdpi.com/2226-471X/3/2/10/s1, Table S1: Experimental Stimuli, Table S2: Participant data for acceptability ratings, Table S3: Participant data for reading times, Figure S1: Results for the mixed DP asymmetry by group.

Acknowledgments: The author has been funded by a Canada Graduate Scholarship from the Social Sciences and Humanities Research Council, as well as by an Ontario Graduate Scholarship. Many thanks to Éric Mathieu for his valuable guidance regarding theoretical aspects of this paper, as well as to Laura Sabourin for help with the experimental side. Thanks also to Dan Siddiqi for helpful comments on an earlier version of this paper. This paper has also benefited considerably from the comments of two anonymous reviewers and the academic editors. Finally, thanks to members of the Ottawa-Carleton Linguistics Reading Group and the ERP Linguistics Lab at the University of Ottawa for their support and feedback, and of course to all the enthusiastic language mixers who participated in the study.

Conflicts of Interest: The author declares no conflict of interest.

References

Alexiadou, Artemis, Terje Lohndal, Tor A. Åfarli, and Maren Berg Grimstad. 2015. Language Mixing: A Distributed Morphology approach. In *NELS 45: Proceedings of the Forty-Fifth Annual Meeting of the North East Linguistic Society: Volume 1*. Edited by Thuy Bui and Deniz Özyıldız. North Charleston: CreateSpace Independent Publishing Platform.

Belazi, Hedi M., Edward J. Rubin, and Almeida Jacqueline Toribio. 1994. Code switching and X-Bar Theory: The Functional Head Constraint. *Linguistic Inquiry* 25: 221–37.

Brysbaert, Marc, and Boris New. 2009. Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41: 977–90. [CrossRef] [PubMed]

Burkholder, Michele, Éric Mathieu, and Laura Sabourin. 2017. The role of gender in mixed-language nominal phrases: Insights from Distributed Morphology. Paper presented at Generative Approaches to Second Language Acquisition (GASLA) 14. Southampton, UK, April 7–9.

Byers-Heinlein, Krista. 2013. Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition* 16: 32–48. [CrossRef]

Cantone, Katja Francesca, and Natascha Müller. 2008. Un nase or una nase? What gender marking within switched DPs reveals about the architecture of the bilingual language faculty. *Lingua* 118: 810–26. [CrossRef]

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Edited by Kenneth Hale and Samuel Jay Keyser. Cambridge: MIT Press, pp. 1–52.

Chomsky, Noam. 1995. Bare phrase structure. In *Government and Binding Theory and the Minimalist Program*. Edited by Gert Webelhuth. Oxford: Blackwell, pp. 71–132.

Chomsky, Noam. 2005. Three Factors in Language Design. *Linguistic Inquiry* 36: 1–22. [CrossRef]

Clegg, Jens H., and Mark Waltermire. 2009. Gender assignment to English-origin nouns in the Spanish of the Southwestern United States. *Southwest Journal of Linguistics* 28: 1–18.

DeKeyser, Robert M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22: 499–533.

De Vincenzi, Marica, Remo Job, Rosalia Di Matteo, Alessandro Angrilli, Barbara Penolazzi, Laura Ciccarelli, and Francesco Vespignani. 2003. Differences in the perception and time course of syntactic and semantic violations. *Brain and Language* 85: 280–96. [CrossRef]

Deuchar, Margaret. 2006. Welsh-English code-switching and the Matrix Language Frame model. *Lingua* 116: 1986–2011. [CrossRef]

Di Sciullo, Anne-Marie, Pieter Muysken, and Rajendra Singh. 1986. Government and code-switching. *Journal of Linguistics* 22: 1–24. [CrossRef]

Ditman, Tali, Phillip J. Holcomb, and Gina R. Kuperberg. 2007. An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology* 44: 927–35. [CrossRef] [PubMed]

Embick, David, and Rolf Noyer. 2007. Distributed Morphology and the Syntax/Morphology interface. In *The Oxford Handbook of Linguistic Interface*. Edited by Gillian Ramchand and Charles Reiss. Oxford: Oxford University Press, pp. 289–324.

Fairchild, Sarah, and Janet G. Van Hell. 2017. Determiner-noun code-switching in Spanish heritage speakers. *Bilingualism: Language and Cognition* 20: 150–61. [CrossRef]

Fernández Fuertes, Raquel, and Juana M. Liceras. 2009. FerFuLice Corpus: English-Spanish Bilingual Twins. CHILDES. Available online: https://childes.talkbank.org/access/Biling/FerFuLice.html (accessed on 4 April 2018).

Fernández Fuertes, Raquel, Juana M. Liceras, and Aurora Bel. 2011. Functional-Lexical Code-Switching in Simultaneous and Sequential Child Bilingualism. Paper presented at the International Symposium on Bilingualism (ISB) 8. Oslo, Norway: University of Oslo, June 15–18.

González-Vilbazo, Kay, and Luis López. 2011. Some properties of light verbs in code switching. *Lingua* 121: 832–50. [CrossRef]

Grimstad, Maren Berg, Terje Lohndal, and Tor A. Åfarli. 2014. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd* 41: 213–37. [CrossRef]

Haberlandt, Karl. 1994. Methods in reading research. In *Handbook in Psycholinguistics*. Edited by Morton Ann Gernsbacher. San Diego: Academic Press, Inc., pp. 1–31.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In *The View from Building 20*. Edited by Kenneth Hale and Samuel Jay Keyser. Cambridge: MIT Press, pp. 111–76.

Halle, Morris. 1997. Distributed Morphology: Impoverishment and Fission. *MIT Working Papers in Linguistics* 30: 425–49.

Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40: 225–76. [CrossRef]

Hawkins, Roger, and Cecilia Yuet-hung Chan. 1997. The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research* 13: 187–226. [CrossRef]

Jake, Janice L., Carol Myers-Scotton, and Steven Gross. 2002. Making a minimalist approach to code switching work: Adding the matrix language. *Bilingualism: Language and Cognition* 5: 69–91. [CrossRef]

Jake, Janice L., Carol Myers-Scotton, and Steven Gross. 2005. A response to MacSwan (2005): Keeping the matrix language. *Bilingualism: Language and Cognition* 8: 271–76. [CrossRef]

Johnson, Jacqueline S., and Elissa L. Newport. 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60–99. [CrossRef]

Jorschick, Liane, Antje Endesfelder Quick, Dana Glässer, Elena Lieven, and Michael Tomasello. 2011. German–English-speaking children's mixed NPs with 'correct' agreement. *Bilingualism: Language and Cognition* 14: 173–83. [CrossRef]

Joshi, Aravind K. 1985. Processing of sentences with intrasentential code switching. In *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*. Edited by David R. Dowty, Lauri Kattunen and Arnold M. Zwicky. Cambridge: Cambridge University Press, pp. 190–205.

Julien, Marit. 2005. *Nominal Phrases from a Scandinavian Perspective*. Amsterdam: John Benjamins.

Kihm, Alain. 2005. Noun class, gender and the lexicon-syntax interface: A comparative study of Niger-Congo and Romance languages. In *The Oxford Handbook of Comparative Syntax*. Edited by Guglielmo Cinque and Richard S. Kayne. Oxford: Blackwell, pp. 459–512.

Kramer, Ruth. 2015. *The Morphosyntax of Gender*. Oxford: Oxford University Press.

Kroff, Jorge R. Valdés, Paola E. Dussias, Chip Gerfen, Lauren Perrotti, and M. Teresa Bajo. 2016. Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguistic Approaches to Bilingualism* 7: 163–98. [CrossRef] [PubMed]

Kroll, Judith F., and Erika Stewart. 1994. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33: 149–74. [CrossRef]

Kučerová, Ivona. 2017. ϕ-Features at the Syntax-Semantics Interface: Evidence from Nominal Inflection. *Linguistic Inquiry*. Available online: http://ling.auf.net/lingbuzz/003483 (accessed on 4 April 2018).

Liceras, Juana M. 2014. La adquisición de las lenguas segundas aquí y ahora o... cómo abordar la hipótesis de la interlengua en el siglo XXI. In *El español global*. Edited by J. Santiago Guervós and Y. González Plasencia. Valladolid: Fundación Siglo para el Turismo y las Artes de Castilla y León, pp. 124–48.

Liceras, Juana. M., Raquel Fernández-Fuertes, Susana Perales, Rocío Pérez-Tattam, and Cristina Martínez. 2006. L2A and the pidgin/creole *continuum*: The pidginization/nativization hypothesis revisited. Paper presented at EUROSLA. Antalya, Turkey: Boğaziçi University, September 13–16.

Liceras, Juana M., Raquel Fernández Fuertes, Susana Perales, Rocío Pérez-Tattam, and Kenton Todd Spradlin. 2008. Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional–lexical mixings. *Lingua* 118: 827–51. [CrossRef]

Liceras, Juana M., Raquel Fernández Fuertes, and Rachel Klassen. 2016. Language dominance and language nativeness: The view from English-Spanish code-switching. In *Spanish-English Codeswitching in the Caribbean and the U.S.*. Edited by R. E. Guzzardo Tamargo, C. M. Mazak and M. C. Parafita Couto. Amsterdam and Philadelphia: John Benjamins, pp. 107–38.

Lieber, Rochelle. 1992. *Deconstructing Morphology*. Chicago: University of Chicago Press.

Lipski, John. 1978. Code-switching and the problem of bilingual competence. In *Aspects of Bilingualism*. Edited by Michel Paradis. Columbia: Hornbeam Press, pp. 250–64.

Litcofsky, Kaitlyn A., and Janet G. Van Hell. 2017. Switching direction affects switching costs: Behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. *Neuropsychologia* 97: 112–39. [CrossRef] [PubMed]

Lohndal, Terje, and Marit Westergaard. 2016. Grammatical gender in American Norwegian heritage language: Stability or attrition? *Frontiers in Psychology* 7: 344. [CrossRef] [PubMed]

MacSwan, Jeff. 1999. *A Minimalist Approach to Intrasentential Code Switching*. New York: Garland.

MacSwan, Jeff. 2000. The architecture of the bilingual faculty: Evidence from intrasentential code switching. *Bilingualism: Language and Cognition* 3: 37–54. [CrossRef]

MacSwan, Jeff. 2005. Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism". *Bilingualism: Language and Cognition* 8: 1–22. [CrossRef]

MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah: Lawrence Erlbaum Associates, Available online: https://childes.talkbank.org (accessed on 4 April 2018).

Mahootian, Shahrzad. 1993. A Null Theory of Code-Switching. Ph.D. dissertation, Northwestern University, Chicago, IL, USA.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own Lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium: Penn Working Papers in Linguistics 4*. Edited by Alexis Dimitriadis. Philadelphia: Department of Linguistics, University of Pennsylvania, pp. 201–25.

Moro, Monica. 2001. *The Semantic Interpretation and Syntactic Distribution of Determiner Phrases in Spanish/English Codeswitching*, Paper presented at the 3rd International Symposium on Bilingualism; Bristol, UK, April 17–24.

Moro Quintanilla, Monica. 2014. The semantic interpretation and syntactic distribution of determiner phrases in Spanish/English codeswitching. In *Grammatical theory and Bilingual Codeswitching*. Edited by Jeff MacSwan. Cambridge: MIT Press, pp. 213–26.

Moyer, Melissa Greer. 1993. Analysis of Code-Switching in Gibraltar. Ph.D. dissertation, Universidad Autónoma de Barcelona, Barcelona, Spain.

Müller, Stefan, and Stephen Wechsler. 2014. Lexical approaches to argument structure. *Theoretical Linguistics* 40: 1–76. [CrossRef]

New, Boris, Christophe Pallier, Marc Brysbaert, and Ludovic Ferrand. 2004. Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society* 36: 516–24.

Pesetsky, David, and Esther Torrego. 2001. T-to-C Movement: Causes and Consequences. In *Ken Hale: A life in Language*. Edited by Michael Kenstowicz. Cambridge: The MIT Press, pp. 355–426.

Pfaff, Carol W. 1979. Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language* 55: 291–318. [CrossRef]

Poplack, Shana, Alicia Pousada, and David Sankoff. 1982. Competing influences on gender assignment: Variable process, stable outcome. *Lingua* 56: 139–66. [CrossRef]

Poplack, Shana, and Nathalie Dion. 2012. Myths and facts about loanword development. *Language Variation and Change* 24: 279–315. [CrossRef]

Poplack, Shana. 1980. "Sometimes I'll start a sentence in Spanish y termino en Espanol": Toward a typology of code-switching. *Linguistics* 18: 581–618. [CrossRef]

Poplack, Shana. 1981. The syntactic structure and social function of code-switching. In *Latino Language and Communicative Behavior*. Edited by Richard Duran. Norwood: Ablex, pp. 169–84.

Riksem, Brita Ramsevik. 2017. Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now. *Languages* 2: 3. [CrossRef]

Sabourin, Laura, Christie Brien, and Michèle Burkholder. 2014. The effect of age of L2 acquisition on the organization of the bilingual lexicon: Evidence from masked priming. *Bilingualism: Language and Cognition* 17: 542–55. [CrossRef]

Sabourin, Laura, Jean-Christophe Leclerc, Myriam Lapierre, Michèle Burkholder, and Christie Brien. 2016. The Language Background Questionnaire in L2 Research: Teasing Apart the Variables. In *Proceedings of the 2016 Annual Conference of the Canadian Linguistic Association*. Edited by Lindsay Hracs. Available online: http://cla-acl.ca/wp-content/uploads/actes-2016/Sabourin_etal_CLA2016_proceedings.pdf (accessed on 4 April 2018).

Sankoff, David, and Shana Poplack. 1981. A formal grammar for code-switching. *Research on Language and Social Interaction* 14: 3–45.

Siddiqi, Daniel. 2009. *Syntax within the Word: Economy, Allomorphy, and Argument Selection in Distributed Morphology*. Amsterdam: John Benjamins.

Van Dulm, Ondene. 2009. English-Afrikaans intrasentential code-switching: Testing a feature checking account. *Bilingualism: Language and Cognition* 12: 193–212. [CrossRef]

Van Gelderen, Elly, and Jeff MacSwan. 2008. Interface conditions and code-switching: Pronouns, lexical DPs, and checking theory. *Lingua* 118: 765–76. [CrossRef]

Woolford, Ellen. 1983. Bilingual Code-Switching and Syntactic Theory. *Linguistic Inquiry* 14: 520–36.