International Journal of Applied Linguistics * Vol. ** No. ** 2012

Not wrong, yet not quite right: Spanish ESL students' use of gerundial and infinitival complementation

Maria Teresa Martinez-Garcia University of Kansas Stefanie Wulff University of North Texas

This study presents a contrastive analysis of gerundial and infinitival complementation produced by Spanish and German ESL learners and English native speakers. An analysis of more than 1,100 attestations of the target constructions obtained from the *International Corpus of Learner English* reveals that (i) advanced learners' construction choices are not necessarily ungrammatical, yet often non-idiomatic, and (ii) German learners are overall more attuned to native-like choices than Spanish learners. A preliminary case study suggests that Spanish ESL teaching materials may benefit from incorporating results of corpus-based analyses of authentic language such as the present one. The present study seeks to close the gap between learner language research and ESL teaching materials, and to raise awareness about the gradual nature of second language proficiency.

Keywords: Spanish ESL student writing, gerundial vs. infinitival complementation, corpus linguistics, distinctive collexeme analysis, second language proficiency

Este estudiopresenta un análisis contrastivo del uso de completivas de gerundio e infinitivoproducidas por parte de estudiantes españoles y alemanes de inglés como lengua extranjera, así, como hablantes nativos de esta lengua. Un análisis de más de 1,100 ejemplosde las estructuras estudiadas obtenidos a partir del Corpus Internacional de Estudiantesde Inglés revela que (i) las construcciones utilizadas por los estudiantes más avanzadosno son necesariamente agramaticales, aunque a menudo no sean idiomáticas, y (ii)los estudiantes alemanes son, en general, más sensibles a las opciones nativas quelos estudiantes españoles. Un estudio preliminar sugiere que los materiales didácticospara la la enseñanza de inglés para los hablantes nativos de español podrían beneficiarsede los resultados obtenidos mediante el análisis de la lengua en uso que se encuentraen el corpus, como el actual. El presente artículo tiene por objeto cerrar la brechaentre la investigación de la lengua usada por los estudiantes y los materiales didácticospara la enseñanza de inglés como lengua extranjera, y así crear conciencia sobre el carácter gradualdel aprendizaje de una segunda lengua.

14

15

18 19

23

2425

28

30

36

39

Palabrasclave; redacciones en ingles como segunda lengua de parte de estudiantesespañoles, complementos de gerundio e infinitivo, lingüística decorpus, análisis distintivo del collexeme, aprendizaje de segundaslenguas

Introduction

The English language offers a great variety of semantically similar complement patterns that present a challenging puzzle for learners of English as a second language (ESL). In this study, we focus on the ESL students' use of gerundial and infinitival complements as shown in (1):

(1) a. Maria **began to feed** the squirrels. [infinitival construction] b. Maria **began feeding** the squirrels. [gerundial construction]

Gerundial and infinitival complements differ on several semantic dimensions. Biber, Johansson, Leech, Conrad and Finegan (1998) state that the gerundial complementation expresses general events, whereas the 'to'-construction licenses a more specific reading (examples from Biber et al. 1998: 758):

- (2) a I **tried rocking** the baby gently when it cried.
 - b. I **tried to rock** the baby gently when it cried.

Moreover, the gerundial construction tends to denote actuality or *realis*, whereas the infinitival construction tends to denote potentiality or *irrealis*. Accordingly, (3b) likely sounds less acceptable to native speakers than (3a) because the gerundial complement in (3b) implies that the act of bribing was carried out, not just attempted as implied by the use of the infinitival complement in (3a) – this stands at odds with the second half of the sentence "but failed".

- (3) a. Sheila **tried to bribe** the jailor but failed.
 - b. ??Sheila **tried bribing** the jailor but failed.

Another difference concerns the temporal interpretation of the event described: the gerundial construction denotes a simultaneous interpretation related to the utterance, whereas the infinitival construction points to the future (Quirk et al., 1985: 1191–3):

- (4) a. I **remembered filling out** the form.
 - b. I remembered to fill out the form.

Isolated examples like the above suggest that gerundial and infinitival complements are clearly distinguishable constructions. Recent corpus-based research, however, indicates that all these are tendencies at best when one

15

24

27

30

44

45

Spanish ESL students' use of gerundial and infinitival complementation • 3

considers all instances of these constructions and the verbs they instantiate. Gries and Wulff (2009) used the British Component of the International Corpus of English (ICE-GB) to replicate previous claims. With regard to the infinitival construction, they found that associated verbs such as 'try' and 'wish' denote potentiality, and that many of the other verbs are future-oriented (as, for example, 'intend', 'hope', 'learn', and 'aim'). On the other hand, they found that the gerundial construction evokes an interpretation in relation to the frame in which the action denoted by V₁ unfolds ('avoid', 'end', 'imagine', 'hate', etc.) and the most distinctive gerundial verbs, 'keep', 'start', and 'stop', correspondingly denote actual events. However, their study also suggested that neither is the gerundial complementation restricted to follow factive predicates (as claimed by Kiparsky and Kiparsky 1971: 347f.), nor are the verbs in this construction necessarily implicatives (as asserted by Givón 1990: 534). Most importantly for the present study, they showed that many verbs which, strictly speaking, license both constructions, display (varying degrees of) preference towards one complementation pattern in terms of frequency of use (Gries and Wulff 2009).1

1

It is not surprising that choosing between these two constructions presents difficulty even-for advanced ESL learners (Petrovitz 2001; Bourke 2007), and even advanced learners often sound non-idiomatic (Celce-Murcia and Larsen-Freeman 1999). While complementation features in any given ESL textbook or course syllabus, there is surprisingly little research to date on how ESL learners use complementation patterns. Moreover, previous research has mainly focused on learner errors, suggesting that not only do ESL learners make more mistakes with the gerundial construction; they also tend to acquire it at a later stage of learning, and its accurate use tends to lag behind that of the infinitival construction (see Anderson 1976, on Spanish and Persian learners; and Schwartz and Causarano 2007, on Spanish learners of English). An error analysis approach may, however, provide only a partial picture of the learners' proficiency – as Gilquin (2007: 288) points out, target-like language use is much more than the absence of errors: "[A]n error analysis . . . only lifts a corner of the veil. Equally important are indications as to what learners get right, what they underuse and what they overuse".

In other words, L2 accuracy should not be understood as a binary phenomenon such that learners make either correct or incorrect choices – much rather, L2 accuracy is a *gradual* concept in the sense that—at a specific grammatical structure is not acquired fully at once, but learners gradually approach target-like use of that structure. This definition not only requires a theory of second language learning that can accommodate an understanding of L2 accuracy as a gradual concept, it also raises the question how to measure L2 accuracy at any given level of L2 proficiency in a way that displays the probabilistic nature of L2 accuracy at that level. The present paper aims to propose one possible way of doing so by presenting the results of an empirical case study of Spanish L2 English intermediate-level learners' use of complementation in academic writing. We approach the data through the

theoretical lens of construction grammar, and employ a corpus-based method called distinctive collexeme analysis (DCA) that quantifies the association strengths between the matrix clause verbs and the gerundial and infinitival complementation patterns, respectively. By applying this statistic to the data, we can not only see whether a verb is used in either complementation pattern or not – we obtain a value that tells us to what degree a given verb is associated with either pattern when taking its overall distribution into account. Results obtained from native speaker corpus data are compared with the Spanish learner data as well as German ESL learners at a comparable level of proficiency.

This paper is structured as follows: we first position our study in the framework of construction grammar, briefly outlining how the gerundial and infinitival complementation patterns qualify as constructions as defined in this theory. We then give a brief introduction into distinctive collexeme analysis as the measure of choice applied in the present study. The third section describes how the Spanish learner data were extracted. The fourth section presents the results of the corpus-based analyses, as well as findings gathered from a preliminary examination of Spanish ESL textbooks with regard to the recommendations made therein on adequate use of the two complementation constructions. The fifth section closes with a discussion of our main findings and desiderata for future research.

A constructionist approach to accuracy in L2 production

Complements as constructions

The framework that we adopt here is construction grammar (Goldberg 1995; 2006), in which constructions are defined as follows:

Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency (Goldberg 2006: 5):

In other words, morphemes, words, and even abstract syntactic patterns can be defined as constructions. The two complementation patterns of interest here are good candidates for constructions in this sense: as we discussed earlier, each one is associated with some (if only subtle and fuzzy) semantic features such as actuality (gerundial) vs. potentiality (infinitival), etc.

In Goldbergian construction grammar as adopted here, a complex utterance is a combination of several constructions (morphemes, words, syntactic patterns). For instance, we can decompose (1) into the following constructions: the words 'Maria', 'began', 'to', 'feed', 'the', 'squirrels'; a NP

Spanish ESL students' use of gerundial and infinitival complementation • 5

construction; a VP construction; and, last but not least, an infinitival complementation construction. The only difference between (1) and (2) is the choice of complementation construction. Constructions can be combined freely as long as their specifications are compatible with each other. Many verbs such as 'try', 'begin', or 'start', for instance, can occur in both constructions because their semantic specifications are sufficiently compatible with each other. A verb like 'end', on the contrary, is not compatible with the infinitival construction:

- (5) a. ?Steffi ended to feed the squirrels every day.
 - b. Steffi ended feeding the squirrels every day.

From a constructionist perspective we can account for the ill-formedness of (5a) such that 'end' strongly entails that some action was carried out on a regular basis or for an extended period of time, which is not sufficiently compatible with the semantic specifications of the infinitival construction.

Of crucial importance for the present study is that in construction grammar, well-formedness is a matter of degree: it is at least in part correlated with the frequency with which the constituent constructions of an utterance occur together. To illustrate this point, consider the examples in (6):

- (6) a. Maria tried to feed the squirrels.
 - b. Maria attempted to feed the squirrels.

Both 'try' and 'attempt' are semantically compatible with the infinitival construction, and they are near synonyms across many contexts, yet to native speakers of English, (6a) will sound (if only slightly) more well-formed than (6b). (6b) is not ill-formed – it just seems that attempt is not the most *typical* or *idiomatic* choice in this context. As we will show below, the difference in idiomaticity of 'try' and 'attempt' in the infinitival construction is reflected in the frequencies with which the two verbs occur in that construction. In other words, the conditional frequency of 'try' | infinitival construction is much higher than that of 'attempt' | infinitival construction. We outline the method we used to calculate these conditional frequencies for the present study in the next section.

A constructionist perspective as outlined here has several implications for our understanding of second language proficiency and second language teaching. First, the traditional distinction between vocabulary and syntax is discarded with: instead of learning words and grammatical rules to accurately combine these words, knowing a language means to know constructions at different levels of complexity and abstraction, and how often these constructions are combined with each other. Second, learner accuracy is not a binary phenomenon: choosing a verb to go with a particular syntactic construction is not either right or wrong, but can result in anything on a scale ranging from perfectly accurate to entirely ungrammatical, depending on the

6 • Maria Teresa Martinez-Garcia and Stefanie Wulff

match between the conditional frequency with which a learner uses a verb in a given construction compared to the frequency with which native speakers use that verb in the given construction (Wulff and Gries 2011). Several studies have provided experimental and corpus-linguistic evidence that (at least advanced) L2 learners of English in fact have that kind of knowledge for argument structure constructions and complementation constructions (Liang 2002; Gries and Wulff 2005; 2009).

Measuring accuracy in a constructionist framework: collostructional analysis

Collostructional analysis (CA) refers to a family of related corpus-linguistic methods (Stefanowitsch and Gries 2003; Gries and Stefanowitsch 2004), that measure the afore-mentioned association between two constructions, most typically the association between a verb and a construction it occurs in. CA has been applied in various domains, including dialectal variation (Mukherjee and Gries 2009), and diachronic change (Hilpert 2006), and contrastive studies of native speaker and learner data (Gilquin to appear). Wulff and Gries (2011: 71) point out how collostructional analysis can also be applied as a measure of L2 accuracy: "Collostructional analysis is a technical operationalization of accuracy when defined as native-like selection, asking: what is the likelihood of a construction X in the environment of another construction Y?" The present study is the first one to contrast English native speaker and Spanish L2 English learner data using this method.

The kind of collostructional analysis that we employed in the present study is called distinctive collexeme analysis (DCA). A DCA is specifically tailored to contrasting two (or more) constructions (not necessarily related) in terms of the verbs occurring in them (Gries and Stefanowitsch 2004). If a verb exhibits a significant preference for either construction, it is called a *collexeme* of that construction (*collexeme* is a blend of *collocate* and *construction*). For each candidate verb, four frequency values are entered into a 2-by-2 table:

- 1. the token frequency of the verb lemma in the gerundial construction;
- 2. the token frequency of the verb lemma in the infinitival construction;
- 3. the corpus frequency of the gerundial construction;
- 4. the corpus frequency of the infinitival construction.

The table is then submitted to a Fisher Yates exact test (or any other association measure; see Stefanowitsch and Gries 2003: 217–8 for reasons why the Fisher Yates exact test may be the preferred choice). To give one example, consider Table 1, which summarizes the frequency values for the verb 'begin' in the Spanish learner data: out of 1,094 occurrences of the infinitival construction, 51 occur with 'begin', and 'begin' occurs 2 times in the gerundial

Spanish ESL students' use of gerundial and infinitival complementation • 7

Table 1. Frequencies of begin in the gerundial and infinitival construction in SP-ICLE

	40	
Begin	Other verbs	Total
51	1,043	1,094
2	177	179
53	1,220	1,273
	51 2	51 1,043 2 177

construction, which occurs 179 times overall (the remaining cell values are determined through addition and subtraction).

The Fisher Yates exact test provides a p-value which we report in its negative \log_{10} -transformed form for easier interpretation. P_{\log} values that are equal to or larger than 1.3 correspond to a probability of error of 5% or less. The p_{\log} for 'begin' is 1.86, so 'begin' is a distinctive collexeme of the infinitival construction.

Data

We retrieved all instances of words ending in 'ing' and all instances of the preposition 'to' from the Spanish component of the *International Corpus of Learner English* (SPICLE), which contains ~300 essays (~200,000 words) of intermediate to advanced Spanish learners of English (Granger et al. 2009). True hits of either complementation construction were then manually identified, yielding a sample of 179 instances of the gerundial and 1,094 instances of the infinitival construction (out of 4,807 words ending in 'ing' and 6,073 instances containing the preposition 'to'). The principal criterion to identify true hits was that the two verbs had to meet the semantic constraint that the first verb specifies the action that the second one denotes (Langacker 1991: 445). In order to ascertain correct frequency counts of the verbs instantiating the two constructions, misspelled variants as in (7) were corrected:

(7) Most people who find happiness meet it having a mint of money to not limit what they wnt to do each time.

The final data sample was then used to carry out a distinctive collexeme analysis using Gries' (2007) *R*-script.

Results

Corpus analyses

In order to address the question of how target-like the Spanish learners' collexeme preferences are, we can compare the DCA results of the Spanish

8 • Maria Teresa Martinez-Garcia and Stefanie Wulff

learner data with the ones reported in Gries and Wulff (2009) for native speaker data obtained from the British component of the *International Corpus of English* (ICE-GB), and those reported in Wulff and Gries (2011) for German learner data obtained from the German component of the *International Corpus of Learner English*. Table 2 lists the distinctive collexemes for both constructions in all three data sets; overlapping collexemes are highlighted in bold print and with a superscript that indicates which other data set the collexeme also occurs in (^G for German L2 English, ^{NS} for native speaker English, and ^{SP} for Spanish L2 English).

Spanish vs. native speaker data

For the infinitival construction, Table 2 reveals nearly no overlap of distinctive collexemes: the only verb distinctively associated with the infinitival construction is 'try'. Looking at the two constructions in contrast, two semantic groups emerge from Table 2 both in the Spanish learner and in the NS data that resonate with our previous description of the main semantic differences between gerundial and infinitival constructions: most of the verbs distinctively associated with the infinitival construction evoke an interpretation in relation to the future (top three collexemes in the learner data: 'try', 'want', 'begin'; top three collexemes NS data: 'try', 'wish', 'seek'); in contrast, the verbs most distinctively associated with the gerundial complementation are aspectual verbs that relate to an event or action unfolding during speaker time, be it the beginning ('start'), the termination ('stop', 'end', 'finish', and 'get rid of'), and the continuation ('continue', 'go on', 'keep', and 'keep on') of that action or event. A second group of collexemes distinctive for the gerundial construction connotes liking ('enjoy', 'like', (prefer').

While the collexemes that the Spanish learners select for both constructions reflect the general semantic trends associated with these constructions in the native language rather well, they differ considerably in terms of the specific collexemes that instantiate these trends. This holds in particular for the infinitival construction: most of the distinctive collexemes for the infinitival construction in the NS data (such as 'tend', 'manage', 'wish', 'refuse', 'intend', 'plan', etc.), do not yield significant p_{log} values in the Spanish learner data. (One may argue that the corpora that the learner and native speaker data were extracted from are not identical in thematic focus and genre and register make-up, so naturally no complete collexemic overlap should be expectable; while this caveat holds generally true, we will see below that German data obtained from the same learner corpus are considerably more similar to the native speaker data overall.) For the gerundial construction, we find more overlap between the Spanish L2 and the NS data: 6 verbs rank among the top distinctive collexemes in both data sets ('stop', 'start', 'enjoy', 'avoid', 'finish', and 'keep').

Spanish ESL students' use of gerundial and infinitival complementation • 9

					,						
		Infinitival construction	structio	u				Gerundial construction	struction	u	
Spanish L2 English	L2 h	German L2 English	L L2 h	NS English	ish	Spanish L2 English	L2 h	German L2 English	7	NS English	sh
collexeme	p_{\log}	collexeme	p_{\log}	collexeme	p_{\log}	collexeme	p_{\log}	collexeme	$p_{ m log}$	collexeme	p_{\log}
try NS	18.57	try NS, SP	CO	try G, SP	22.44	stop G, NS	13.65	keep NS, SP	13.99	keep G, SP	76.45
need	3.02	manage NS	5 4 74	wish G	5.39	continue c	13.50	gton NS, SP	13.6 9.4	start c, sr ston G, SP	35.23
decide	2.81	tend NS		seek	4.35	enjoy ^{G, NS}	7.74	start NS, SP	8.71	avoid G, SP	11.87
begin	1.86	learn ^{NS}		tend ^G	4.06	go on ^G	88.9	avoid NS, SP	6.2	end	11.87
pretend	1.60	$\mathbf{begin}^{\mathrm{SP}}$		intend	3.67	Keep on G	88.9	enjoy ^{NS, SP}	6.2	enjoy ^{G, SP}	11.87
		dare	1.58	attempt	3.19	mean	4.04	end up SP	3.57	mind	11.87
				hope	3.19	like	3.53	give up	2.38	remember ^G	10.14
				fail	3.09	avoid G, NS	3.38	continue SP	2.3	80 c	7.99
				like ^G	3.03	end up ^G	2.56	hate NS	2.3	consider	5.45
				refuse	2.98	insist on	2.56	remember NS	2.3	envisage	3.38
				learn ^G	2.1	prefer ^G	2.38	finish NS, SP	1.78	finish ^{G, SP}	3.38
				plan	1.89	finish ^{G, NS}	1.71	keep on SP	1.78	carry	2.53
				continue	1.53	get rid of	1.71	go on SP	1.78	fancy	2.53
				afford	1.49	keep ^{G, NS}	1.71	prefer SP	1.36	imagine	2.53
						spend	1.71			resist	2.53
										catch	1.69
										hate ^G	1.38
										bear	1.25
										begin	1.03
										7	0

^{© 2012} Blackwell Publishing Ltd

Also noteworthy are cases where we observe directly opposite constructional associations: 'begin' strongly prefers the gerundial construction in NS English, but the Spanish L2 learners significantly prefer 'begin' in the infinitival construction. Conversely, 'like' and 'continue' are distinctive for the gerundial construction in the Spanish L2 English data, while distinctive for the infinitival construction in the English L1 data.

Finally, it should be noted that several verbs are distinctive for the gerundial construction in the native speaker data, but not attested at all in the Spanish learner data, including 'envisage', 'fancy', 'imagine', 'resist', 'catch', or 'hate'; the only exception here is 'consider', which occurs 5 times in the gerundial construction in the Spanish L2 data, but does not yield significance. The verb 'fail' deserves particular attention in this context because although it ranks among the most distinctive infinitival collexemes in English L1, it does not appear in the Spanish L2 English data.

Spanish vs. German learner data

In order to address the question to what extent the overlap between the Spanish L2 and the English native speaker data may be attributed to transfer effects, we can glean some answers from comparing the Spanish learner data with German learner data.

Turning first to the infinitival construction, we can see in Table 2 that there is just as little overlap between the Spanish and the German data as there is between the Spanish and the native speaker data: the only collexemes shared between both learner groups are 'try', which is in accord with native speaker preferences, and 'begin', which runs counter to native speaker preferences. In German, the translation equivalents for 'begin', 'anfangen' and 'beginnen', are strongly associated with the infinitival construction ('zu' is the German infinitive marker), which may invite transfer:

(8) Es fing an zu regnen. / Es begann zu regnen. It began to rain.

Likewise, the Spanish word for 'begin', 'empezar', is a prepositional verb always followed by an infinitive as in (9). Similar to examples (11) and (12) in note 5, it is conceivable that the Spanish learner falsely aligns the preposition 'a' in Spanish with the infinitival marker 'to' in English.

(9) Completamente extravertido y a voz en grito **empiezo** a **exclamar**.² Completely extrovert and, at the top of my voice, I begin to cry out.

When we compare the German and the native speaker preferences for the infinitival construction, we see that this is largely because the German learners are, in fact, very similar to the native speakers in terms of the collexemes they

Spanish ESL students' use of gerundial and infinitival complementation • 11

prefer to use in the infinitival construction: of the seven collexemes in the German learner data, all but 'dare' are also collexemes of the infinitival construction in the NS data. From this, we may conclude that while both learner groups use verbs in the infinitival construction that reflect its main semantic function (in contrast to the gerundial construction), the German learners are much more attuned to the specific collexemes that occupy the infinitival construction than the Spanish learners are. The one striking exception here is again 'begin': both learner groups strongly prefer to use it in the infinitival construction, which runs counter to native speaker preferences.

As far as the gerundial construction is concerned, the overlap between the two learner groups is much bigger: 'stop', 'continue', 'start', 'enjoy', 'go on', 'keep on', 'avoid', 'prefer', 'finish', and 'keep' are among the top collexemes in both data sets. This also matches native speakers' preferences rather well.

Aside from their non-idiomatic preference of 'begin' in the infinitival construction, another commonality between the Spanish and German learners is the marked preference of phrasal verbs such as 'keep on', and 'go on' in the gerundial construction. Wulff and Gries (2009: 16) speculated about this characteristic in the German data and explained that the possible underlying motivation for the frequent use of these verbs could be the attempt by German learners to transfer a construction in German, 'X ist am $V_{infinitive}$ ' ('X is V_{ing} '):

2

[T]he combination of the preposition *am* with the bare form of a verb is one of the few ways in which progressive aspect can be expressed in German. The semantics of the gerundial complementation construction are sufficiently compatible with a progressive reading, and learners may fill the slot of the German *am* with the particle of the phrasal verb.

In Spanish, there is no such structure that would invite transfer, so we can only speculate that one of the following applies: (i) students memorize lists of verbs used in each construction and they use them systematically; (ii) they associate the gerund with the meaning of the main verb (we also find the main verbs 'keep' and 'go' among the most distinctive ones, although 'go' does not yield statistical significance), disregarding the particle; or (iii) they associate the gerund with the particle, considering that it will work as if it were a preposition; one rule that many Spanish ESL text books point out is that prepositions are frequently followed by a gerund form.

Two other collexemes stick out when comparing the German and Spanish learners with the native speaker data: 'like' and 'continue'. Unlike the Spanish learners, the German learners correctly prefer 'like' in the infinitival construction; both learner groups, however, share the non-target-like preference of 'continue' in the gerundial construction. One way to account for this may be that while the German learners are, overall, more attuned to the collexeme preferences of L1 English, they continue to struggle with collexemes that do not yield particularly high p_{\log} values ('like' yields a p_{\log}

value of 3.03 in the native speaker data, 'continue' a value of 1.53). As one anonymous reviewer points out, this may be related to the fact that the learners receive much less exposure to these examples because they are, as their p_{\log} values in native language indicate, much less frequent in comparison, so they are harder to catch the learners' attention.

Also, both learner groups may have overgeneralized the association of the gerundial construction with aspectual verbs such as 'start', 'stop', and 'end' (and, to a lesser extent, the association of the gerundial construction with verbs of liking: German learners strongly prefer 'hate' in the gerundial construction as well, again going against native speaker preferences). In the case of 'continue', Spanish learners may be invited to negative transfer because its translation equivalent in Spanish is one of the few verbs that take the gerundial construction in Spanish, as shown in (10):

(10) Aún **continúo preguntándome** —había escrito el maestro anterior—cómo aquel fanatismo lúcido y abnegado del joven Luciano ha venido a desembocar en esta mezquina dominación tribal y en esta ciega avidez de sangre... I **continue asking** myself-the former teacher had written-how that lucid, selfless fanaticism of young Luciano has derived in this mean tribal domination and in this blind eagerness of blood . . .

As far is 'like' is concerned, on the contrary, negative L1 transfer can be ruled out: both 'gustar' and 'querer', the two translation equivalents of 'like' in Spanish, only occur with infinitival complementation.

Instruction materials used in Spanish ESL

In order to further investigate why overall, the Spanish learners seem to be less attuned to the idiomatic verb-construction combinations than the German learners, we surveyed eight popular Spanish ESL textbooks and grammars (listed in the Appendix) and took inventory of the rules and recommendations these references make as far as the adequate choice of construction is concerned. Obviously, this analysis can neither lay claim to provide a comprehensive overview of the materials used in Spanish ESL instruction, nor is it a representative impression of Spanish ESL students' input; however, it serves only as a pointer to more in-depth future studies examining the role of input and instruction.

Table 3 lists the verbs (in alphabetical order) for which at least one of the textbooks (titles are abbreviated; see the Appendix) explicitly recommend using it only with the gerundial construction. Verbs that are significantly distinctive for either construction in the native speaker or the Spanish learner data are listed as ING and TO, respectively; verbs at least attested in the data are listed as (ING)/(TO) to indicate their distributional trend. Sixteen verbs

Spanish ESL students' use of gerundial and infinitival complementation • 13

Table 3. Verbs referenced in Spanish ESL instruction materials as taking gerundial complementation

Verb	Textbook(s)	NS preference	Spanish NNS preference
avoid	C, D	ING	ING
bear	B, G	(ING)	NA
carry on	В	ÌNG	(ING)
catch	В	ING	NA
consider	C, D	ING	(TO)
deny	В, Е	NA	(ING)
enjoy	B, E, F, H	ING	ING
finish	C, D	ING	ING
give up	B, C, D	NA	(ING)
go on	В	NA	ING
ĥelp	В	ING	ING
insist on	В	NA	ING
keep	В	ING	ING
keep on	В	ING	ING
look forward to	В	NA	(ING)
mind	C, D	ING	(ING)
miss	В	(ING)	NA
practis ce	C, D	(ING)	NA
prevent	В	NA	(TO)
stand	B, C, D, G	(TO)	NA
stop	В, Е	ING	ING

are referenced in the instruction materials as occurring with gerundial complementation, but are attested neither in the native speaker nor the Spanish learner data ('admit', 'burst out', 'detest', 'feel like', 'find', 'insist', 'mention', 'object to', 'postpone', 'put off', 'put up with', 'risk', 'see to', 'suggest', 'swear', and 'tolerate').

Overall, Table 3 suggests that the ESL instruction materials make adequate suggestions in the sense that the verbs they list as taking predominantly or exclusively the gerundial construction are indeed significantly distinctive of the gerundial construction in authentic native speaker data. Likewise, the Spanish learners seem to use these verbs in a target-like fashion. The only minor exceptions here are 'consider' and 'prevent', which, while not yielding significant *p* values, are attested in the infinitival construction in the learner data. What is striking, however, is that with the exception of textbook B (*Morfosintaxis*), which stands out as the material that lists the majority of verbs actually attested also in native speaker data, the number of verbs suggested by other materials is rather scarce. Also, when we contrast Table 3 with the list of highly distinctive collexemes of the gerundial construction in the native speaker data, we find that some of the most strongly associated collexemes are

14 • Maria Teresa Martinez-Garcia and Stefanie Wulff

Table 4. Verbs referenced in Spanish ESL instruction materials as taking infinitival complementation

Verb	Textbook(s)	NS preference	Spanish NNS preference
afford	B, D	TO	(ING)
agree	B, D	NA	(TO)
ask	B, D	NA	(TO)
attempt	В	TO	(TO)
decide	B, D, E	(TO)	TO
expect	B, D, H	NA	(TO)
fail	В	TO	NA
forget	В	(TO)	(TO)
hesitate	В	(TO)	(TO)
hope	B, D, H	TO	(TO)
intend	В	TO	(TO)
learn	B, D	TO	(TO)
manage	B, D	TO	(TO)
need	B, D	NA	ŤΟ
offer	D	(TO)	(TO)
plan	D	TO	(TO)
prefer	E	(TO)	ÌNĠ
prepare	D	(TO)	(TO)
pretend	В	(TO)	ŤΟ
promise	B, D, H	(TO)	(TO)
refuse	B, D	TO	(TO)
threaten	В	NA	(TO)
try	В	TO	ŤO ´
use	В	NA	(TO)
want	B, C, D, H	TO	ŤO ´
wish	В	TO	(TO)

apparently not mentioned in any of the instruction materials, including highly frequent (and hardly topic-specific) verbs such as 'start', 'end', 'remember', 'imagine', or 'hate'. It seems, then, that most instruction materials do not provide the Spanish learners with wrong advice – they just do not necessarily provide them with the most comprehensive or most idiomatic input.

In an analogy to Table 3, Table 4 lists the verbs referenced in the ESL instruction materials as taking the infinitival construction. Two other verbs are mentioned in the instructional materials, but not attested in either the native speaker or the Spanish learner data ('demand' and 'trouble').

Looking at Table 4, a very similar picture emerges for the infinitival construction as we saw before for the gerundial construction: the textbook recommendations match native speaker preference quite well overall. Similarly, when we compare the verbs in Table 4 with the most distinctive collexemes of the infinitival construction in the native speaker data, we find

Spanish ESL students' use of gerundial and infinitival complementation • 15

Table 5. Verbs referenced in Spanish ESL instruction materials as taking both gerundial and infinitival complementation

Verb	Textbook(s)	NS preference	Spanish NNS preference
begin	В	(ING)	ТО
continue	В	ŤΟ	ING
hate	E	ING	NA
like	E	TO	ING
love	E	(TO)	(TO)
start	E	ÌNĠ	ÌNĠ

that the overlap is much larger than for the gerundial construction: of the top 15 distinctive collexemes in the native speaker data, only four ('seek', 'tend', 'like', and 'continue') are not referenced in any of the ESL instruction materials we examined. Again, however, we see that there are two references in particular that are responsible for the lion's share of this overlap: *Morfosintaxis* (B) and *Bonus* 2° *bachillerato* (D). Furthermore, even these textbooks introduce a series of verbs that are not even attested in the native speaker data with the infinitival construction (e.g. 'agree', 'ask', 'expect', etc.).

Finally, Table 5 lists the verbs that license both constructions according to the ESL instruction materials. What stands out in Table 5 is the fact that one ESL material in particular (*Grammar in gobbets*) lists several verbs as licensing both constructions which in fact show significant constructional preferences in native speaker data: 'continue', 'hate', 'love', and 'start' do in fact all occur in both constructions, but overwhelmingly in just one construction. With the exception of 'start', the Spanish learners are apparently not aware of these probabilistic preferences, as is evidenced by the fact that their significant preferences stand in direct opposition to the native speakers' preferences.

In summary, we can conclude that the ESL instruction materials do not provide wrong advice, but at the same time, the majority of them are far from providing input that would aid the Spanish learners in making truly idiomatic (as opposed to grammatically correct) decisions on which verb to pair with which complementation construction.

Discussion

The main findings of the present study can be summarized as follows:

 Spanish (and German) ESL learners are better attuned to the verb-specific preferences associated with the gerundial construction than with those of the infinitival construction;

16 • Maria Teresa Martinez-Garcia and Stefanie Wulff

- German ESL learners and Spanish ESL learners share certain overgeneralization tendencies errors, including the overuse of phrasal verbs and aspectual verbs in the gerundial construction;
- overall, German learners outperform Spanish ESL learners in terms of the overlap of specific verbs attested in the learner data compared to the native speaker data;
- Spanish (and, to a lesser extent, German) ESL learners' difficulty with the infinitival construction may partially be accounted for by flawed L1-L2 alignments and negative transfer from the L1;
- with few exceptions, Spanish ESL instruction materials do not provide inadequate advice regarding the specific verbs to be used in gerundial/ infinitival complementation, but they largely fail to provide (dense) input on verbs that are particularly highly associated with these two constructions in L1 English; at the same time, they reference verbs that, while principally licensed by the two constructions, are not attested in a large-scale corpus of native English.

In consequence, the Spanish ESL learners represented in ICLE appear less proficient than the German ESL learners, not because they make *ungrammatical* verb choices in the two constructions, but because their choices are not entirely *idiomatic* or target-like. This reinforces Pawley and Syder's (1983) distinction between 'native-like fluency' and 'native-like selection' as two separate components of second language proficiency, and provides an empirical illustration of Howarth's (1998) claim that:

many learners fail to understand the existence of the central area of the phraseological spectrum between free combinations and idioms. It is in handling restricted collocations that errors of both a lexical and grammatical structure constantly occur. Moreover, learners need to understand that restricted collocations make up a significant part of a typical native speaker's production in both speech and writing. (Howarth 1998: 186)

The extent to which ESL teaching materials impact (Spanish) learners' use of complementation constructions is a question that requires further systematic and more comprehensive research into teaching materials and their interaction with other sources of learner input than we could provide here. Similarly, more studies are needed that examine the use of different structures across learner groups from different L1 backgrounds, across different proficiency levels, and across registers and genres. Nonetheless, the results of the preliminary case study presented here suggest that Spanish teaching materials may benefit from presenting this grammar topic in a fashion that more adequately reflects actual native language use. Maybe most importantly, learners (at least at intermediate to advanced levels of proficiency) could potentially benefit hugely from understanding that

40

41 42

43

Spanish ESL students' use of gerundial and infinitival complementation • 17

choosing a verb and the complementation pattern to go with it is not just a matter of right or wrong, but that some verbs are associated with these complementation patterns much more strongly than others. Text books could easily incorporate information as provided in Table 2 to raise learners' awareness for which verbs are most strongly preferred in either complementation construction. Depending on the proficiency level of learners, this list could comprise only the top most significantly distinctive collexemes, and be extended to incorporate, or replaced by, verbs further down the list in order to build a comprehensive vocabulary.

More generally speaking, the present study may be taken as one of many examples of how corpus-based studies of authentic language provide insights into language that would otherwise escape one's attention. While it may be too cumbersome to have students replicate the exact analysis presented here in the classroom, a simple query of the two patterns in a large-scale corpus of English (such as, say, the *Contemporary Corpus of American English*, which is freely available online⁶), may already be a powerful way to raise learners' awareness of which verbs occur in these constructions most frequently. Such data-driven, explorative learning approaches have been gaining attention in corpus linguistics and TESOL as of late, reporting promising results overall (see Braun 2006 and Chambers 2007 for recent overviews).

In summary, we hope to have illustrated the usefulness of corpus-based and quantitative investigations into learner language, and to have provided a starting point for future research.

Acknowledgements

We would like to thank Douglas Biber, Jerid Francom, Stefan Gries, John Swales, Jiyoung Yoon, and two anonymous reviewers for their helpful comments on an earlier version of this paper.

Notes

- 1. To complicate matters even more, many ESL learners struggle with the proper use of the gerundial construction in particular since this construction is much rarer than the infinitival construction (in English and Spanish, for instance; see Schwartz and Causarano 2007), or it is completely absent in many languages (as in German; see Mair 2003). Consequently, depending on their L1 background, ESL learners may not benefit from positive transfer when acquiring the two constructions in English.
- 2. All Spanish examples were taken from the *Real Academia de la Lengua* unless indicated otherwise.
- 3. An alternative explanation offered by one anonymous reviewer is that German speakers may associate the gerund with the nominalized infinitive construction (as opposed to the ordinary form preceded by the infinitival marker *zu* 'to') in these

instances: the particles could be misinterpreted as nominal markers, resulting in a preference for the gerundial construction.

4. Among the ESL teaching materials we surveyed for our analysis, textbook C and textbook H, for instance, mention this rule (see the Appendix for full references):

We use the gerund: (1) as the subject of a sentence. Example: "smoking is bad for you", (2) after prepositions. Example: "the thought of going on holiday was exciting", and (3) after certain verbs. Example: "he admitted stealing the bikes". (textbook C: 104; emphasis added)

The gerund. Uses: (1) subject of a verb. Example: "smoking is bad for you", (2) object of a verb. Example: "do you like cooking?", (3) the complement of to be. Example: "his favourite activity is watching television", (4) after prepositions. Example: "she's very good at listening to people", (5) after determiners. Example: "the bombing of civilians horrified everyone", (6) after possessive adjectives, object pronouns or nouns. Example: "I hope you don't mind his/him/John coming here", (7) after adjectives. Example: "regular swimming is very good for you", (8) in certain expressions. Example: "it's no use trying to help her". (textbook H: 111; emphasis added).

- 5. As in Example (9), Spanish learners' preference for the infinitival construction in these cases may be due to the learners' flawed interlanguage analysis of structural correspondences between Spanish and English. For *consider*, for instance, while the Spanish verb *considerar*, is not followed by an infinitival complement, it frequently occurs with *that*-complements that contain infinitives as in (11). The complementizer *qua* could be falsely aligned with the preposition *to* in English.
 - (11) Y no sólo estoy muy satisfecho de haberlo hecho, sino que considero que defender la democracia es algo por lo que puede merecer la pena llegar a sacrificar la vida. And I am not only proud of having done it, but I also consider that to defend the democracy is something that may be worth sacrificing your life for.

An analogous explanation may hold for *prevenir*, the Spanish word for *prevent*:

- (12) Te **prevengo** que no **soy** una amante cómoda. I **prevent you** that I am not a lazy lover.
- 6. http://corpus.byu.edu/coca/>

References

Anderson, J. I. (1976) A comparison of the order of difficulty of English sentential complements between native speakers of Spanish and native speakers of Persian. Paper presented at the Los Angeles Second Language Research Forum, UCLA.

45

46 47

49

Spanish ESL students' use of gerundial and infinitival complementation • 19

- Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan (1998) *Longman grammar of spoken and written English*. London: Longman.
- Bourke, J. M. (2007) Verbal complementation: a pedagogical challenge. *Reflections on English Language Teaching* 6.1: 35–50.
- Braun, S. (2006) ELISA a pedagogically enriched corpus for language learning purposes. In S. Braun, K. Kohn, and J. Mukherjee (eds.), *Corpus technology and language pedagogy: new resources, new tools, new methods.* Frankfurt/M: Peter Lang. 25–47.
- Chambers, A. (2007) Popularising corpus consultation by language learners and teachers. In E. Hidalgo, L. Quereda, and J. Santana (eds.), *Corpora in the foreign language classroom*. Amsterdam: Rodopi. 3–16.
- Celce-Murcia, M. and D. Larsen-Freeman (1999) *The grammar book: an ESL/EFL teacher's course.* Boston, MA: Heinle and Heinle.
- Gilquin, G. (2007) To err is not all: what corpus and elicitation can reveal about the use of collocations by learners. *Zeitschrift fuer Anglistik und Amerikanistik* 55.3: 273–91.
- (to appear) Lexical infelicity in causative constructions: comparing native and learner constructions. In J. Leino and R. von Waldenfels (eds.), *Analytical Causatives*. Munich: Lincom.
- Givón, T. (1990) Syntax: a functional-typological introduction. Amsterdam: John Benjamins.
- Goldberg, A. E. (1995) Constructions: a construction grammar approach to Argument structure. Chicago, IL: University of Chicago Press.
- (2006) Constructions at work: the nature of generalization in language. Oxford: Oxford University Press.
- Granger, S., E. Dagneaux, F. Meunier, and M. Paquot. (2009) *International corpus of learner English v2*. Louvain-la-Neuve: Presses universitaires de Louvain.
- Gries, S. F. (2007) Coll.analysis 3.2. A program for R for Windows 2.x.
- and A. Stefanowitsch (2004) Extending collostructional analysis: a corpus-based perspectives on 'alternations'. *International Journal of Corpus Linguistics* 9.1: 97–129.
- and S. Wulff (2005) Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics* 3: 182–200.
- (2009) Psycholinguistic and corpus-linguistic evidence for L2 constructions. *Annual Review of Cognitive Linguistics* 7: 163–86.
- Hilpert, M. (2006) Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory* 2.3: 243–57.
- Howarth, P. (1998) The phraseology of learners' academic writing. In A. P. Cowie (ed.), *Phraseology*. Oxford: Clarendon. 161–86.
- Kiparksy, P., and C. Kiparksy (1971) Fact. In D. Steinberg, and L. A. Jakobovitz (eds.), Semantics: an interdisciplinary reader in philosophy, linguistics, and psychology. Cambridge: Cambridge University Press. 345–69.
- Langacker, R. W. (1991) *Foundations of cognitive grammar*. Vol. II. Stanford, CA: Stanford University Press.
- Liang, J. (2002). Sentence comprehension by Chinese learners of English: verb-centered or construction-based. Unpublished M.A. thesis, Guangdong University of Foreign Studies.
- Mair, C. (2003) Gerundial complements after *begin* and *start*: grammatical and sociolinguistic factors, and how they work against each other. In G. Rohdenburg,

41

42

20 • Maria Teresa Martinez-Garcia and Stefanie Wulff

- and B. Mohndorf (eds.), *Determinants of grammatical variation in English*. Berlin: Mouton de Gruyter. 347–77.
- Mukherjee, J., and S. T. Gries (2009) Verb-construction associations in the International Corpus of English. *English World-Wide* 30.1: 27–51.
- Pawley, A., and F. H. Syder (1983) Two puzzles for linguistic theory: native-like selection and native-like fluency. In J. C. Richards and R. W. Schmidt (eds.), *Language and communication*. New York: Longman. 191–226.
- Petrovitz, W. (2001) The sequencing of verbal-complement structures. *English Language Teaching Journal* 55.2: 172–7.
- Quirk, R., S. Greenbaum, G. Leech, and S. Svartvik (1985) A comprehensive grammar of the English language. London: Longman.
- Schwartz, M., and P. L. Causarano (2007) The role of frequency in SLA: an analysis of gerunds and infinitives in ESL written discourse. *Arizona Working Papers in SLA and Teaching* 14.1: 43–57.
- Stefanowitsch, A., and S. T. Gries (2003) Collostructions: investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8.2: 209–43.
- Wulff, S., and S. T. Gries (2011) Corpus-driven methods for assessing complexity and accuracy in learner production. In P. Robinson (ed.), *Researching task complexity: task demands, task-based language learning and performance*. Amsterdam: John Benjamins. 61–87.

Appendix: ESL textbooks and grammar books analysed

- A: Abbs, B., C. Barker, and I. Freebairn (1999) Snapshot for ESO. Student's book 2. Madrid: Longman.
- B: Alcaraz, E., and B. Moddy (1990) Morfosintaxis Inglesa para hispanohablantes. tgoría y práctica. Alcoy: Editorial Marfil.
- C: Bolton, D., and J. Wildman (2001) Bonus 1° bachillerato. Student's book 1. Oxford: Oxford University Press.
- D: —— (2001) Bonus 2° bachillerato. Student's book 1. Oxford: Oxford University Press.
- E: Campos, M.A., A. Lillo, and V.M. Pina (2002) Grammar in gobbets. A guide to English grammar and usage. Alicante: Editorial Agua Clara, S.L.
- F: Jones, R. (2001) Inglés 2° de ESO. Cuaderno de repaso. Madrid: Longman.
- G: Walker, E. and S. Elsworth. (2000) *Grammar practice for intermediate students with key*. Madrid: Longman.
- H: —— (2000) *Grammar practice for upper intermediate students with key*. Madrid: Longman.

email: maria.martinezgarcia@ku.edu Stefanie.Wulff@unt.edu [Received 9 January 2012]

Toppan B	Sest-set Premedia Limited
Journal Code: IJAL	Proofreader: Elsie
Article No: 310	Delivery date: 1 March 2012
Page Extent: 20	

AUTHOR QUERY FORM

Dear Author,

During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof.

Many thanks for your assistance.

Query References	Query	Remark
q1	JTHOR: To match the reference list, should Kiparsky and Kiparsky 1971 be changed to Kiparksy and Kiparksy 1971? Please advise	
q2	UTHOR: Wulff and Gries (2009 has not seen included in the Reference List, please supply full publication details.	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	Textual mark	Marginal mark
Leave unchanged Insert in text the matter indicated in the margin	··· under matter to remain k	New matter followed by k or k
Delete	/ through single character, rule or underline or through all characters to be deleted	6) or 6/20
Substitute character or substitute part of one or more word(s)	/ through letter or through characters	new character / or new characters /
Change to italics	— under matter to be changed	
Change to capitals	under matter to be changed	=
Change to small capitals	= under matter to be changed	_
Change to bold type	under matter to be changed o	~
Change to bold italic	under matter to be changed	***
Change to lower case	Encircle matter to be changed	≢
Change italic to upright type	(As above)	4
Change bold to non-bold type	(As above)	
Insert 'superior' character	/ through character or k where required	y or \\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
Insert 'inferior' character	(As above)	over character e.g. $\frac{1}{2}$
Insert full stop	(As above)	·
Insert comma	(As above)	,
Insert single quotation marks	(As above)	ý or ý and/or ý or ý
Insert double quotation marks	(As above)	y or y and/or y or y
Insert hyphen	(As above)	H
Start new paragraph	_	
No new paragraph	ے	رے
Transpose	ப	ш
Close up	linking characters	
Insert or substitute space between characters or words	/ through character or k where required	Y
Reduce space between characters or words	between characters or words affected	个