

Gendering subtitles? Evidence from a Netflix corpus

¿Género en los subtítulos? Resultados de un corpus de Netflix

MIKOŁAJ DECKERT

Uniwersytet Łódzki / University of Lodz. Pomorska 171/173, 90-236 Łódź, Polska.

Dirección de correo electrónico: mikolaj.deckert@uni.lodz.pl

ORCID: <https://orcid.org/0000-0003-1569-2399>

KRZYSZTOF W. HEJDUK

Uniwersytet Łódzki / University of Lodz. Pomorska 171/173, 90-236 Łódź, Polska.

Dirección de correo electrónico: krzysztof.hejduk@edu.uni.lodz.pl

ORCID: <https://orcid.org/0000-0002-9476-4428>

JAGODA BANASIAK

Uniwersytet Łódzki / University of Lodz. Pomorska 171/173, 90-236 Łódź, Polska.

Dirección de correo electrónico: jagoda.banasiak@edu.uni.lodz.pl

ORCID: <https://orcid.org/0009-0001-2183-9329>

LILIANA ZATORSKA

Uniwersytet Łódzki / University of Lodz. Pomorska 171/173, 90-236 Łódź, Polska.

Dirección de correo electrónico: liliana.zatorska@edu.uni.lodz.pl

ORCID: <https://orcid.org/0009-0006-9993-7929>

Recibido/Received: 11/5/2023. Aceptado/Accepted: 13/9/2023.

Cómo citar/How to cite: Deckert, Mikołaj *et al.*, «Gendering subtitles? Evidence from a Netflix corpus», *Hermēneus. Revista de Traducción e Interpretación*, 25 (2023): pp. 133-167.

DOI: <https://doi.org/10.24197/her.25.2023.133-167>

Artículo de acceso abierto distribuido bajo una [Licencia Creative Commons Atribución 4.0 Internacional \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/). / Open access article under a [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

Abstract: We start from the observation that language pairs like Polish and English display an intriguing asymmetry: Polish has the structural resources to express grammatical gender in nouns while English predominantly leaves these under-determined. Therefore, a speaker of English can conventionally say “I am a nurse” irrespective of the speaker’s gender, while in Polish male/female/non-binary speakers would likely use variable forms. This asymmetry leads to nuanced translation choices which this paper aims to explore. Based on a newly compiled corpus of English-to-Polish Netflix subtitles – amounting to 3.5 million words – we examine how 113 different occupation-related feminatives are deployed by translators. We address three main research questions (RQ1-RQ3). RQ1 is about the frequency of the relevant feminatives in our

corpus. RQ2 deals with how the feminatives are distributed across the 284 productions analysed in the dataset. Finally, RQ3 addresses the relationship between the frequency of feminatives in Netflix subtitles and the acceptability of these feminatives, as judged by speakers.

Keywords: Gender; feminativa; interlingual subtitling; corpus data; streaming platforms.

Resumen: Partimos de la observación de que pares de lenguas como el polaco y el inglés muestran una asimetría muy atractiva: el polaco tiene los recursos estructurales para expresar el género gramatical en los sustantivos, mientras que el inglés los deja predominantemente indeterminados. Por lo tanto, un hablante de inglés puede decir convencionalmente «I am a nurse» (soy enfermera) independientemente del género del hablante, mientras que en polaco los hablantes masculinos/femeninos/no binarios probablemente utilizarían formas variables. Esta asimetría da lugar a opciones de traducción matizadas que este artículo pretende explorar. A partir de un corpus recién compilado de subtítulos de Netflix del inglés al polaco –que asciende a 3,5 millones de palabras–, examinamos cómo los traductores emplean 113 feminativos diferentes relacionados con la ocupación. Abordamos tres preguntas principales de investigación (RQ1-RQ3). La RQ1 se refiere a la frecuencia de los feminativos relevantes en nuestro corpus. La RQ2 trata de cómo se distribuyen los femeninos en las 284 producciones analizadas en el conjunto de datos. Por último, la RQ3 aborda la relación entre la frecuencia de los feminativos en los subtítulos de Netflix y la aceptabilidad de estos feminativos, a juicio de los hablantes.

Palabras clave: Género; feminativa; subtitulación interlingüe; corpus de datos; plataformas de streaming.

Summary: 1. Introduction – asymmetric coding across languages; 2. The study, 2.1. Methodology and corpus composition, 2.1.1. Sourcing Polish target texts [TTs], 2.1.2. Acquiring Polish target texts, 2.1.3. Operationalising feminativa, 2.1.4. Extracting feminativa from target texts, 2.1.5. Annotating the detected feminativa cases; 3. Results, 3.1. Feminative cases/tokens across the relevant word types [RQ1], 3.2. Feminative cases/tokens across target texts [RQ2], 3.3. Feminative cases/tokens across their social acceptability levels [R]; 4. A qualitative perspective; 5. Discussion; 6. Summary of findings; 7. Relevant considerations and further research; Acknowledgements; List of figures and tables; references.

Sumario: 1. Introducción: codificación asimétrica entre lenguas; 2. El estudio, 2.1. Metodología y composición del corpus, 2.1.1. Obtención de textos meta polacos [TT], 2.1.2. Adquisición de textos meta polacos, 2.1.3. Puesta en funcionamiento de los feminativos, 2.1.4. Extracción de los feminativos de los textos meta, 2.1.5. Anotación de los casos de feminativos detectados; 3. Resultados, 3.1. Casos/indicios de feminativos en los tipos de palabras pertinentes [RQ1], 3.2. Casos/indicios de feminativos en los textos meta [RQ2], 3.3. Los casos/indicios y sus niveles de aceptación social; 4. Una perspectiva cualitativa; 5. Discusión; 6. Resumen de los resultados; 7. Consideraciones relevantes y futuras investigaciones; Agradecimientos; Lista de figuras y tablas; referencias.

1. INTRODUCCIÓN – ASYMMETRIC CODING ACROSS LANGUAGES

Languages partition reality in variably commensurable fashions (Lakoff, 1987; Lewandowska-Tomaszczyk, 1987, 2010) and the coding of gender is a notable case of cross-linguistic asymmetry. Some languages have the category of grammatical gender – encompassing a dual masculine/feminine distinction or getting more fine-grained – while others

do not. In this paper we examine a case of English-to-Polish interlingual translation which exposes the asymmetry as a challenge for the translator, on the one hand, and a rich source of possibly far-reaching reception implications, on the other hand.



Figure 1. Screenshot from *History of Swear Words* (2021 ©Netflix); the Polish subtitles contain a feminativum

Let us illustrate the subject of inquiry with an example of Polish subtitles for an English-language documentary show available from Netflix, entitled “History of Swear Words”. The documentary extensively draws on interviews with experts who introduce themselves when first appearing. One such introduction is as follows:

My name is Mellisa Mohr. I have a PhD from Stanford University and I write about swearing.

What might seem relatively straightforward in the terms of the original message includes an interesting translation choice which becomes clear when one reads the Polish subtitles:

Nazywam się Melissa Mohr. Jestem doktorką Uniwersytetu Stanforda i piszę o wulgaryzmach.

[My name is Melissa Mohr. I am a doctor* of Stanford University and I write about swearwords.]

The case in point is the lexeme “doktorka” which codes the person’s feminine gender, as opposed to “doktor” which would code masculine grammatical gender and could likely be used as a generic form to conventionally refer to an individual irrespective of gender – as would be the case with the official Polish term for the academic degree (*e.g.* in legal documents). The target variant proposed by Netflix is notable because the lexeme “doktorka” is not as “conventionalised” (*cf. e.g.* Langacker, 2007, p. 425; Langacker, 2008; Schmid, 2015) as “doktor”, which opens many avenues of inquiry centred on how the viewers perceive the speaker. With variable cognitive, social, professional or interpersonal implications either of the target variants might be bringing, the key point is that in the original utterance the speaker did not encode that choice because she was not required to by the structure of English. In the subtitle, one might argue, it is the translator who decides for the speaker. To make matters more nuanced, the source text did not unconditionally require the subtitler to use a noun phrase and thus mark grammatical gender – the subtitler could have gone for a verbal phrase like “napisałam doktorat” [I wrote a PhD dissertation] or “zrobiłam doktorat” [I did a PhD] which would code the speaker’s gender grammatically in the verb and would be analogously formulaic to the source text expression. It would additionally avoid a technically flawed notion introduced by the Polish translation where “a doctor of Stanford University” suggests that “doctor” is an academic position (like an assistant professor) at a particular institution rather than a degree awarded by that institution.¹

2. THE STUDY

Starting from that lack of cross-linguistic symmetry, the objective of this paper is to provide an insight into the deployment of feminative forms in translation, with Polish subtitling as a case in point. Using a corpus of Netflix subtitles, we extract feminative forms referring to professions. In addition to looking at the frequencies and distribution of the identified

¹ Pertinently, Mohr writes on her personal website (<https://www.melissamohr.com/about.html>) as follows: “I received a PhD from Stanford University in Medieval and Renaissance English Literature, but left academia to write for a wider audience.”

nouns, we relate them to the tripartite reception-based categorisation proposed by Waszyńska *et al.* (2022). The patterns extracted from the quantitative analysis are then supplemented with qualitative insights which bring some methodological considerations.

In sum, we seek to answer the following main research questions:

- (RQ1:) What is the frequency of the relevant feminatives in the corpus?
- (RQ2:) How are the relevant feminatives distributed across audiovisual products in the corpus?
- (RQ3:) How does the frequency of feminatives in the corpus relate to reception?

2.1. Methodology and corpus composition

Methodologically, the study seeks to add to the body of work where corpora have been compiled and employed in audiovisual translation research – in subtitling (Bywood *et al.*, 2013; Rica Peromingo *et al.*, 2014; Sotelo Dios, 2015) as well as in other modes of audiovisual translation and media accessibility like audio description (Salway, 2007; Jiménez Hurtado and Soler Gallego, 2013; Matamala and Villegas, 2016; Reviere, 2018) and dubbing (*cf.* Bruti, 2009; Freddi, 2009; Monti, 2009; Pavesi, 2014; Valentini, 2016, 2018).

Our corpus consisted of target subtitles from the Netflix films and shows that held most popularity² in Poland across the reference period of 70 weeks, between 4 July 2021 and 6 November 2022. The corpus amounts to approximately 3.5-million-word tokens (~150 thousand unlemmatised word types) across 2050 distinct subtitle files. To develop this dataset, we first obtained a sample list of titles available from Netflix. We then acquired their official Polish translations in the form of movie subtitles. This made it possible to extract those subtitles in which feminative forms were used. A resulting “sub-corpus” of feminativa found in Polish subtitles was annotated and analysed quantitatively as well as qualitatively.

2.1.1. Sourcing Polish target texts [TTs]

² We chose to sample Netflix titles using the criterion of their popularity despite our recognition that less popular titles are equally worthy of investigation. Our reasoning behind this choice was that selecting titles with the greatest viewership would provide us with a reflection of subtitles/translational decisions that a substantial portion of the Polish audience might be exposed to.

Each week, Netflix publishes an official list of ten most popular films and TV series by country (Netflix 2022). With that, it was possible to compile a list of the most popular films and shows on Netflix in Poland. We downloaded the official XLSX with per-country collections and extracted the weekly lists pertaining to the Polish viewership.³ After removing duplicate entries (some films and series hit the top 10 many times over), we attained a raw sample of over 500 titles. Importantly, the 500-title list spanned source texts [STs] in 19 different languages (English, Spanish, French, Portuguese, etc.) as well as creations that were natively Polish. While interesting for a potential future study, comparing target texts produced from all of these different languages could affect the internal validity⁴ of our investigation.

We therefore filtered out non-English titles, resulting in a corpus of 284 anglophone productions (98 films, 186 shows). Noteworthy, we kept English-language productions designated by the thematic tags “German”, “Japanese”, and “African”.

2.1.2. Acquiring Polish target texts

With our list of sample films and shows completed, we initiated the process of semi-manual compilation of our subtitles corpus, using a popular 2017 user script by Paweł Wit ([GreasyFork.org/scripts/26654](https://greasyfork.org/scripts/26654)). After manual querying, the script functionally automates the process of requesting locales available to all end-users watching Netflix. It moreover facilitated the creation of individual translation files (with meta-data), from the interface of the Internet browser.

³ Contrary to what has been stated on the Netflix website (Netflix 2022), the Polish lists start on 4 July 2021, not 28 June 2021. This date opens our reference period. The latest weekly list published by Netflix at the time of our investigation commencing was dated 6 November 2022, marking the ending of our reference period.

⁴ While the homogeneity of our final corpus’ files was designed to be high, late into the research we have detected that the subtitles of *Fatal Affair* –a title not among the most popular titles on Netflix in Poland– were acquired instead of *Lakeview Terrace*’s. Moreover, *365 days* and *365 days: This day* were mistakenly considered anglophone, while they were in fact Polish. The impact this had on our results was insignificant, though. *Fatal Affair* had a single feminativum case, “radczyni” (the only token of this type), and the two Polish-source TTs used no feminativa. Disqualifying those TTs at that stage would lead to major recalculations, so considering the overall size of our corpus, we decided to include them as exceptions to our criteria.

The translation data covered full Polish subtitles (“timed texts” in Netflix’s nomenclature), Polish SDH (subtitles for the deaf-and-hard-of-hearing, or “CCs” – closed captions – as Netflix calls them) and localised Forced Narrative subtitles. This last category pertains to “a text overlay that clarifies communications or alternate languages meant to be understood by the viewer” (Netflix Partner Help Center 2023), used for texted graphics (*cf.* Deckert & Augustyn 2022) or extradiegetic location identification, among other things. “Forced Narrative subtitles are only displayed if full Subtitles and CC are set to *off* in the user’s playback settings” (Netflix Partner Help Center, 2023). We decided to include all three variants since SDHs and Forced Narrative may present additional information, not found in the regular captioning, and this information might embed linguistic patterns relevant to the study (*feminativa*).

2.1.3. Operationalising *feminativa*

Since the lexical potential for creating *feminativa* nouns in Polish is plentiful, we needed to restrict the scope of our investigation to a representative sample of *feminativa*. The focus is therefore on nouns denoting occupation since it is a salient domain when it comes to the use of *feminativa* in Polish. For this, we utilised a list of 113 unique occupational *feminativa* presented by Waszyńska *et al.* (2022). That list was compiled based on the official list of professions published by the Polish Ministry of Labour and Social Policy (2014), while removing duplicates/variations of the same profession and creating *feminativa* forms from the masculinative ones. However, Waszyńska *et al.*’s (2022) list consisted of lemmas only, and since Polish is a fusional language, the *feminativa* detectable in our subtitle corpus most certainly occurred as more inflected word forms, too. We therefore expanded the 113-item list by adding every word form declined for the Polish grammatical case (nominative, genitive, dative, accusative, instrumental, locative, and vocative) in singular and plural forms. This was our final lemmatised baseline for corpus search queries.

2.1.4. Extracting *feminativa* from target texts

For the purpose of extracting relevant fragments, we used Laurence Anthony’s toolkit AntConc (64-bit version 3.5.9 on Windows) which is a “[a] freeware corpus analysis toolkit for concordancing and text analysis”

(Anthony, 2019). AntConc can therefore find “sequences of words within a corpus of text” (Gore, 2023; *cf.* Altun, 2019, p. 125), which is how it was utilised in our research. Our corpus, made up of Polish subtitles, was parsed for concordances. The search queries in our case were feminine form declensions based on a lemma list by Waszyńska *et al.* (2022). We utilised AntConc’s Advanced Search functionality which supports searching for multiple terms from a loadable list. The hit results were saved to a TXT file and converted into an Excel-compatible spreadsheet for annotation.

2.1.5. Annotating the detected feminativa cases

Before analysis, the spreadsheet with concordances had to be manually checked for false positive hits and duplicate matches. As was mentioned earlier, our subtitle corpus of Polish translations included up to three types of captions for a single film/show episode: the Polish subtitles, Polish SDHs, and locale’s “Forced Narrative” subtitles (Netflix Partner Help Center, 2023). This required us to process the AntConc results for any duplicate hits as the content of SDHs and Forced Narrative events sometimes overlapped with the regular subtitles. This resulted in the exact same instance of feminativa being found by AntConc twice or even thrice: in the regular captioning file, the SDH file, and possibly in the Forced Subtitles file. Indeed many – but not all – Forced Narrative events containing feminativa appeared in full subtitles or SDHs, since “all Forced Narrative events are also [required to be] included in each full Subtitle and SDH/CC file” (Netflix Partner Help Center, 2023).

Such instances of major overlap had to be filtered out for validity’s sake. This process could not have simply been automated, as it required careful qualitative assessment. Certain duplicates had time codes insignificantly differing between the three caption types (*e.g.* when the same subtitle is displayed for a longer duration in SDHs/Forced Narrative compared to regular subtitles), hence looking for *identical* time codes would not be enough. Another possible automation criterion, *i.e.* the co-text embedding the feminine cases, could also differ with Forced Narrative subtitles and SDH. We therefore considered the same tokens of a feminativum type to be *duplicates* only if they occurred in the same episode/film, ideally within similar co-texts, but importantly with time

codes virtually concurrent to one another.⁵ This meant that flashbacks (or extradiegetic reminders of previous episodes' narration) were not considered duplicates, since they occurred in different episodes and/or with different time codes. Similarly, instances where the same feminativum form (but not the same instance/case of a feminativum) is deployed more than once in the same subtitle did not qualify as duplicates: e.g. *Silverton Siege's* (2022) „Christine. Jestem *kierowniczka* [= match]. *Kierowniczka* [= match] Christine”.

False positives on the other hand were instances where AntConc detected a concordance with a search term from our list of feminativa, but it was not a token of a feminativum after all. This is because Polish declension of certain feminine forms results in forms homonymous to their corresponding masculinativa or other nouns (e.g. “drukarka” = a printing device, also a feminativum for a printer profession). Some of these were completely unrelated homophones – an extreme example of this could be a lexeme “premiera” (= “a prime minister”) which is a nominative form of a feminativum, but also a genitive declension of a masculinativum, as well as a homonym for a word meaning “premiere” or “release”. At times it was possible to check the subtitle's co-text of a matched lexeme to determine if it referred to a male or female entity. For instance, Polish verbs and adjectives are inflected congruently with the inflection of nouns, which makes them a giveaway of the intended grammatical gender category in cases of ambivalent homography. To illustrate this: the phrase “z ministrami” may both mean “with ministers (men)” and “with ministers (women)”, even though they derive from two different root forms (plural, canonical: “ministrowie” and “ministry” respectively). Another giveaway could be the inclusion of the social honorific “pani” (feminative gender) or “pan” (its masculine counterpart) which may appear before nouns indexing people, but not things or concepts. If in some cases it was impossible to assess the gender within the limited co-text of several left-right collocates (some of their spaces being taken up by subtitle timestamp codes), we consulted the audiovisual context of the original movie.

⁵ A noteworthy liminal situation we detected in *The Good Nurse* (2022) where around the first half-hour of the movie, two characters reiterate the same type of feminativum several times (these cases are analysed further on). We did not consider these cases as duplicates, but indeed as individual instances of feminine form use. The Polish SDH here simply states that the two characters repeat one after another, which defused any potential for duplicates between SDH and regular subs.

3. RESULTS

The removal of false positives resulted in a complete output of Polish feminativa detected in the subtitles of most popular Netflix films and shows. We have identified 917 unique cases of feminative forms used in the English-to-Polish translations of 178 unique and *available* productions: 91 (out of in-total 186⁶) films and 87 (out of in-total 98) TV shows (only first season each). By *available* we mean films/shows that could be viewed on Netflix in the Polish locale as of 16 January 2022, out of the titles originally present in the listing of most popular films and shows on Netflix. As many as 55 films and 3 shows were no longer available (*cf.* Section 2.1.5.). Below we present findings corresponding to the RQ1-RQ3 formulated above.

3.1. Feminative cases/tokens across the relevant word types (RQ1)

To answer the first research question, we searched the corpus for 113 lemmatised feminativa word types (Waszyńska *et al.* 2022). We detected 917 cases of matching word tokens.⁷ 41% (46/113) of feminativa types returned 0 tokens in our corpus. As many as 40% (45/113) of feminativa types each returned fewer than 10 tokens. This 40% of types stands for 13% of all tokens in the corpus (121/917). Thus, 19% of types (22/113) stand for the 87% remaining tokens (796/917). Fewer than half (43%) of our feminativa types/lemmas were unrepresented in the corpus. The exact frequencies of the relevant feminativa tokens are presented in Table 1, alongside their lemma types and their English equivalents (our translation).

⁶ *Sic* – the Polish subtitles of 105/186 originally anglophone films had no feminative cases at all. It is worth noting, however, that the span of a single film can be quite a bit shorter compared to a single season of a series, giving it far less material for feminative cases to be even possible to occur. This may account for there conversely being no feminative cases detected in as little as 11/98 of the relevant TV show translations.

⁷ This data is not proportional – future research should investigate how often the translators were given a chance to decide on feminativa, rather than a masculinativa, rephrasings, etc.

Polish feminine lemma	English equivalent	# of cases in our corpus
asystentka	assistant	88
pielęgniarka	nurse	83
agentka	agent	68
nauczycielka	teacher	68
tancerka	dancer	56
opiekunka	babysitter/caretaker	53
aktorka	actress	52
pisarka	writer	41
policjantka	policewoman	40
dziennikarka	journalist	38
trenerka	coach	34
kierownicza	executive/superintendent	24
dyrektorka	executive/superintendent	22
barmanka	bartender	20
detektywka	detective	17
kucharka	cook	17
maszynistka	train driver	14
bibliotekarka	librarian	13
mechaniczka	mechanic	13
burmistrzynie	mayor	12
strażniczka	guard	12
fryzjerka	hairdresser	11
cukierniczka	confectioner	9
kasjerka	cashier	9
malarka	painter	8
projektantka	designer	8
reżyserka	director	8
fotografka	photographer	6
tłumaczka	translator	5
przewodniczka	chairperson	5
doradczyne	adviser	4
choreografka	choreographer	3
piekarka	baker	3
prezydentka	president	3

psycholożka	psychologist	3
redaktorka	editor	3
żołnierka	soldier	3
prezeska	CEO	2
wychowawczyni	educator/form teacher	2
rektorka	rector	2
rektora	rector	0
ekonomistka	economist	2
fizycyżka	physician	2
historyczka	historian	2
inżynierka	engineer	2
operatorka	operator	2
rzeczniczka	representative	2
sędzina	judge	2
urzędniczka	clerk/official	2
adwokatką	barrister	1
analityczka	analyst	1
architektka	architect	1
biolożka	biologist	1
chemiczka	chemist	1
dekoratorka	decorator	1
dziekanka	dean	1
dziekana	dean	0
ekspedientka	sales clerk	1
fizjoterapeutka	physiotherapist	1
instruktorka	instructor	1
kapitanka	captain	1
kurierka	courier	1
menedżerka	manager	1
ministra	minister	1
ministerka	minister	0
ogrodniczka	gardener	1
premiera	prime minister	1
premierka	prime minister	0
radczyni	solicitor	1*
ratownicza	lifeguard	1

stolarzka	carpenter	1
administratorka	administrator	0
blacharka	tinsmith	0
chirurgzka	surgeon	0
dietetyczka	dietician	0
drukarka	printer	0
dyrygentka	conductor	0
elektryczka	electrician	0
farmaceutka	pharmacist	0
filolożka	philosopher	0
górniczka	miner	0
graficzka	graphic designer	0
hodowczyni	breeder	0
hydrauliczka	plumber	0
informatyczka	informatician	0
kanclerka	chancellor	0
kierowczyni	driver	0
komiczka	comedian	0
korektorka	copyreader	0
lektora	voice actress/lector	0
lektorka	voice actress/lector	0
listonoszka	mail carrier	0
logopedka	speech therapist	0
marynarka	sailor/mariner	0
murarka	bricklayer	0
muzyczka	musician	0
nauczycielka akademicka	university professor	0
notariuszka	notary	0
nurkini	diver	0
oficerka	officer	0
optyczka	optician	0
pedagożka	pedagogue	0
pilotka	pilot	0
polityczka	politician	0
rolniczka	farmer	0
spawaczka	welder	0

starościna	prefect	0
strażaczka	firefighter	0
ślusarka	locksmith	0
taksówkarka	cab/taxi driver	0
tatuażystka	tattooist	0
techniczka	technician	0
wójtina	governor	0

Table 1. Feminativa across their relevant lemmas, with translations into English

3.2. Feminative cases/tokens across target texts (RQ2)

On top of determining the frequency of the different feminativa in the corpus (RQ1), we sought to establish how they are distributed across TTs (RQ2), *i.e.* how the different TTs ‘contribute’ to the total count of the relevant expressions in the corpus. Crucially, we aimed to see whether particular TTs feature feminativa more frequently than others.

Feminativa were detected in the translations of 178 out of 284 sources. A detailed breakdown of findings shows that in the case of films, 49% (91/186) of their translations used any (>0) tokens of feminativa, while in the case of series translations, 88% (87/98) used any (>0) feminine form tokens.⁸ Then, 30% of sources (53/178) had exactly 1 case each in their TTs, standing for 6% of all tokens (53/917). Moreover, 84% (150/178) of sources used 10 or fewer cases each in their TTs, corresponding to 50% of all tokens (455/917). This entails that the top 16% (28/178) of productions in our corpus resulted in as many feminativa use-cases (50%; *cf.* Figure 2) as the remainder. Finally, the top 10% (18/178) was responsible for 38% (346/917) of the detected tokens.

⁸ This disproportion of *any* feminativa use between films and series could be explained by TV shows having more relative watchtime compared to films. Consequently, the translators of series may on average have more opportunities to use at least one feminativum during the span of the entire season. Future research should look at both formats separately.

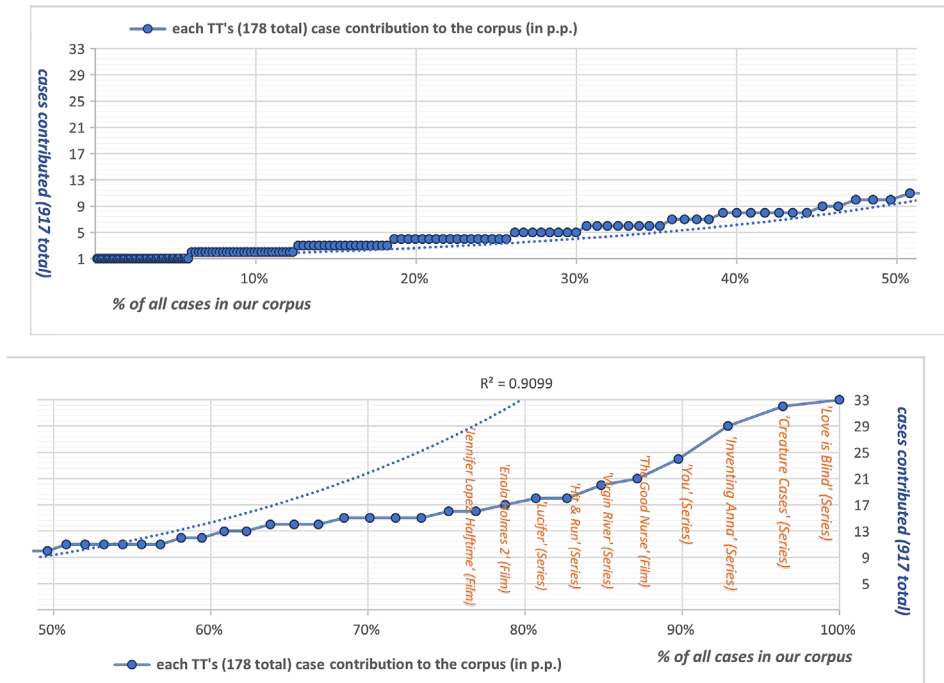


Figure 2. Detected feminativa by their STs, ranked by absolute number of occurrences

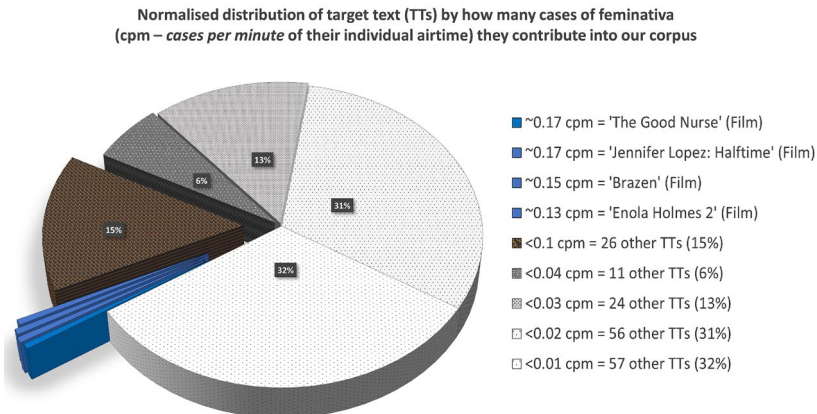


Figure 3. Detected feminativa by their STs, ranked by relative occurrence frequency per watchtime

The distribution of the relevant feminativa across our corpus is illustrated in Figure 2 and Figure 3. All subtitle sets corresponding to the same ST were treated as one TT. Figure 2 ranks feminativa use-cases based on the absolute number of tokens detected, irrespective of factors like the total airtime of a given film/show. Hence, Figure 2 indicates *how many* feminativa occurred in the different TTs in our corpus. Figure 3 shows a normalised ranking – detected case count divided by a given production’s length in minutes. Thus, Figure 3 offers an overview of *how often* feminativa occur in the different TTs in our corpus.

The time frame of the productions in our corpus was vastly different, both between shows and films and within those formats. Therefore, with respect to RQ2, we offer two measures of occurrence-distribution analysis over the corpus files: one is absolute (case frequency), and the other is relative (*i.e.* case frequency divided by minutes of airtime). In the absolute measurement, 10% of all cases in our corpus come from 2% of STs (4/178) – all TV shows. Accounting for the time available for feminativa to occur, a different set of most productive STs emerges: 2% of STs (4/178) – all film translations⁹ – broke the threshold of contributing 1 case on average for every 10 minutes of their airtime (0.1 *cases per minute*). In other words, a feminativum on average occurs in them more frequently than for every 10 minutes of their relative time span. Compared to the rest of our corpus, the STs whose translations had the most feminativa – both in terms of *how many* and *how often* they were used – appear to be *The Good Nurse* (2022), *Enola Holmes 2* (2022) and *Jennifer Lopez: Halftime* (2022).

3.3. Feminative cases/tokens across their social acceptability levels (RQ3)

Waszyńska *et al.* (2022) proposed a three-level categorisation of acceptability (high/moderate/low) of various Polish feminativa. The sample of their study consisted of 300 students who were asked to rate their reception of 113 types of feminatives (*e.g.* “nauczycielka” = female

⁹ This also is explained by films offering less relative watchtime compared to shows, whereby a show translator may not be given the opportunities to frequently use feminativa if relativised against the entire span of the production – unless *e.g.* the entire series revolves around themes of femininity. It follows that looking at the two formats separately in future research will likely offer more detailed conclusions.

teacher) on a 6-point Likert-type scale from 0=*unacceptable* to 5=*fully acceptable*. The types of feminativa that had been assigned levels of acceptability between 4.04–4.94 were considered most positively-received. Types that had been scored between 3.05–3.99 were moderately-received, and the scoring of 2.09–2.98 was deemed least positive (Waszyńska *et al.* 2022).

We utilised the same word types in our study to query the subtitles corpus for any grammatically-declined occurrences (*e.g.* “nauczycielki” = female teachers; “nauczycielkom” = to female teachers, etc.). Figure 4 charts the detected feminativa tokens across the feminativa types and the number of productions where those tokens (and therefore also types) were detected in their Polish subtitles. The Figure omits feminativa types for which no tokens were detected, visualising this omission with bolded black bars (*cf.* Figure 4).

As aforementioned, we found 917 tokens of the 113 lemmas/types proposed by Waszyńska *et al.* (2022). Among these, 91.82% of all tokens¹⁰ (842/917) represented positively received types, although this category constitutes 58% of the studied types. Moderately received types (30% of studied types) were represented by 7.74% (71/917), while the least positively received types (12% of studied types) by 0.44%. In the ‘most positive’ category, the detected tokens stood for 49 types, since 26% (17/66) of this category’s types returned no tokens. In the ‘moderate’ and ‘least positive’ categories, 59% (20/34) and 69% (9/13) of types returned no tokens, respectively.

This suggests that the (vast) majority of tokens detected in the subtitles corpus represent feminativa of the highest socio-cognitive acceptability. They also returned fewest 0-token outputs out of the three reception categories. Less than 1 percent of tokens constitute feminativa types of the lowest acceptability in the subtitles and most of the lowest-acceptability types also went unrepresented (no tokens).

¹⁰ The margin of error is of at least one case/type (“radczyni”) to be disqualified – *cf.* footnote 4.

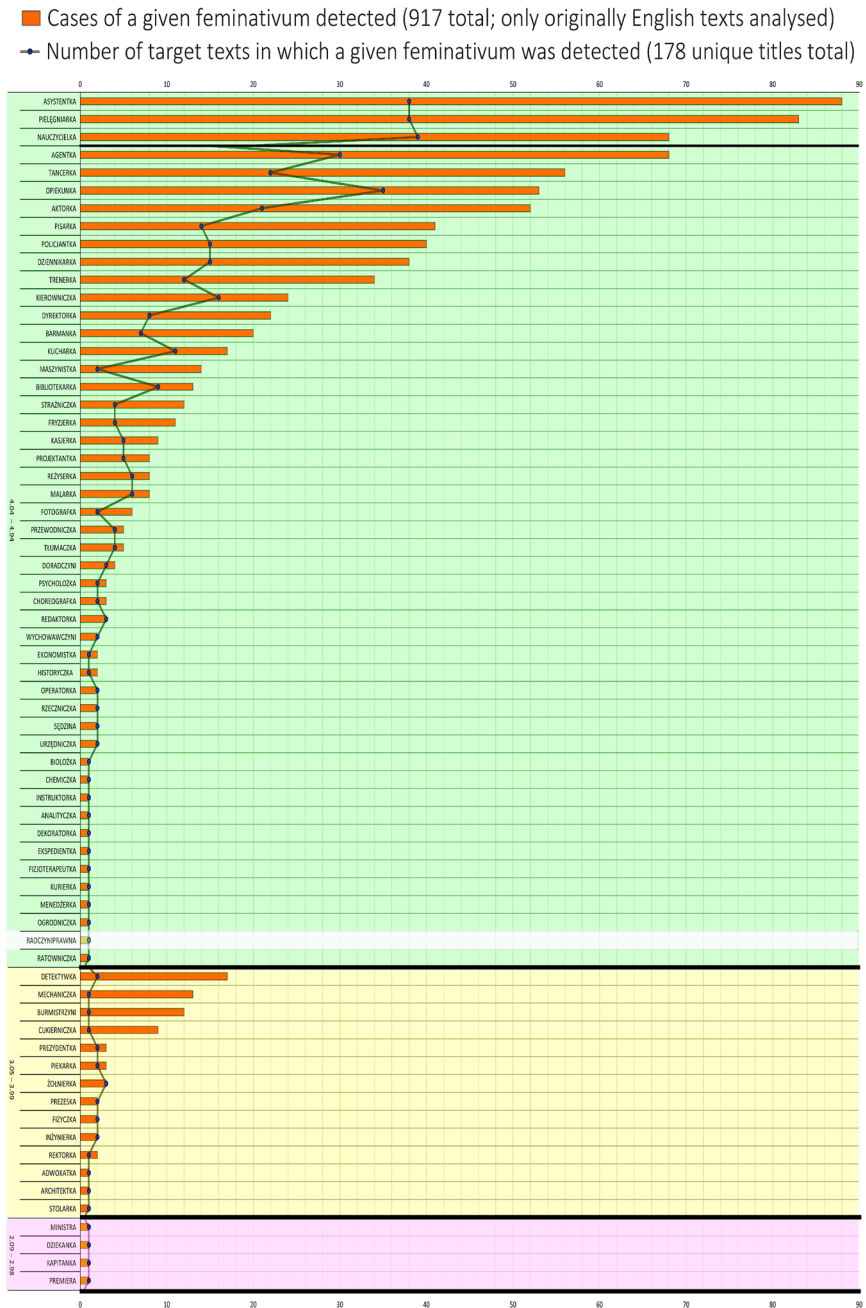



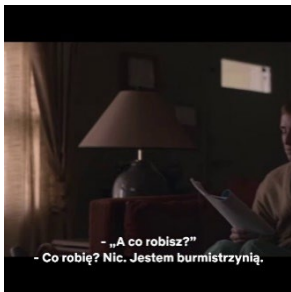


Figure 4. Feminativa across lemmas (0-token types not visualised) and their social acceptability categories (Waszyńska *et al.* 2022)

4. A QUALITATIVE PERSPECTIVE





Let us then supplement the quantitative findings with a qualitative perspective to see how the identified patterns can be contextualised and interpreted to ultimately arrive at a more comprehensive account. To that end we use one of the titles from the corpus – *The Good Nurse* (2022). It displays an interesting pattern; feminative forms occur in this single film multiple times within a very limited fragment. As we outline below the pattern can be easily expounded through qualitative analysis of the relevant scene from *The Good Nurse*, where Alex, a 9-years old girl, is practicing her lines for a school play with her mother’s friend (Charlie), younger sister (Maya) and mother (Amy). At one point Alex fumbles for words and goes on to mention her teacher’s remark regarding her part in the performance as follows: “the teacher said that I should wear a suit and have a mustache. Dress like a man because I’m the mayor. He said a girl mayor would be weird.” In response she receives reassurance from her companions and the scene concludes with her and Charlie holding hands and repeatedly saying one of the lines of Alex’s part – “I am the mayor of Humdrum Falls” (*cf.* Table 2). Remarkably, the male-female mayor opposition – which emerges as a salient element of the narrative – is mirrored in the TT by tapping the structural resources of the Polish language to mark nouns for gender. We observe a shift within the dialogue whereby the first three times Alex utters the relevant lines (“I am the mayor of Humdrum Falls”, “I’m the mayor”, “in my book of *How to be a mayor*”) the Polish variant is “burmistrz” which does not specify the female gender of the speaker. As the scene unfolds, the translator then uses “pani burmistrz” (= “Ms. mayor”) as a translation of “a girl mayor” which is the phrase that the teacher had used according to Alex. Then, when Charlie replies (“Women can be mayors”), a (plural) feminative form is employed in the TT (“Przecież są burmistrzynie”) for the first time. The form (singular) is then instantly used again to translate Amy’s comment where she uses the expression “female mayor” (“[...] the female mayor’s the weird part?”), and then the feminativum occurs repeatedly in the final exchange featuring Charlie and Alex. In sum, it is interesting to note that the line “I am the mayor of Humdrum Falls” remains intact in English throughout the scene while its translation tellingly changes by

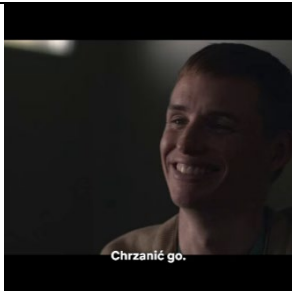



implementing a feminativum. The use of the feminative form appears to reinforce the communicated message, making it arguably even more empowering than the original one. Linked to the title of this article, the qualitative vantage point makes it possible to demonstrate how subtitles are more “gendered” than the original and simultaneously serve as a “gendering” factor, drawing the Polish viewer’s attention to gender more than could be the case for viewers who rely on the original. This example also yields a useful methodological insight into how a simple – and yet not necessarily expected – narrative device can account for accumulation of tokens.

#	Source dialogues	Target text dialogue (SDH)	Target text dialogue (Sub)	Visual context
ALEX	Make way, make way. I am an important person of the town. In fact, I am the most important person of the town, if not the most important person... west of the Pecos.	Z drogi! Jestem ważną osobistością. Najważniejszą w mieście, jeśli nie na... Na zachód od Pecos.	Zróbcie przejście. Jestem ważną osobą w mieście. Właściwie to najważniejszą osobą w mieście, a może nawet najważniejszą osobą na zachód od Pecosu.	
CHARLIE	“Wherever that is!”	„Gdziekolwiek to jest”.	„Gdziekolwiek to jest”.	

ALEX	I am the mayor of Humdrum Falls.	Jestem <u>burmi-strzynia</u> Humdrum Falls.	Jestem burmi-strzem Humdrum Falls.	
CHARLIE	“So, what do you do?”	„A co robisz?”	„Co robisz?”	
ALEX	“Do”? I don't do anything. I'm the mayor .	Co robię? Nic. Jestem <u>burmi-strzynia</u> .	Co robię? Nic nie robię. Jestem burmi-strzem .	
CHARLIE	Okay. “Well, you have to do something now. This may be just a one-horse town without the horse, but it's been invaded by aliens!”	[śmieje się] „Teraz musisz działać. Może to dziura zabita dechami, ale tu wylądowali kosmici”.	„Teraz musisz coś zrobić. Może to zaścianek bez ścian, ale zaatakowali nas kosmici”.	
ALEX	I'll have to look it up in my book of <i>How to be Mayor</i> . One. You get to wear a fancy chain. Two. You get	Zajrzę do poradnika <i>Jak być burmistrzem</i> . Po pierwsze, włóż modny łańcuch. Po drugie, niech szofer wozi cię wielką	Muszę zajrzeć do książki <i>Jak być burmistrzem</i> . Po pierwsze, nosisz drogi łańcuch. Po drugie, masz własnego szofera, który	

	driven around in a big car by your very own chauffeur. Three...	bryką. Po trzecie...	wozi cię wielkim samochodem. Po trzecie...	
CHARLIE	You eat. It's...	–	Jesz.	
ALEX	“Three. You eat and drink for free”. I'm never gonna get this part!	„Po trzecie, jedz i pij za darmo”. Nigdy nie zapamiętam.	„Po trzecie, jesz i pijesz za darmo”. Nie dostanę tej roli.	 „Po trzecie, jedz i pij za darmo”.
AMY	Don't get frustrated, honey.	Nie wściekaj się.	Nie denerwuj się.	 - Nigdy nie zapamiętam. - Nie wściekaj się.

CHARLIE	Are you kidding? This is great, Alex.	Jesteś super.	Żartujesz? Jesteś super.	
ALEX	No, and the teacher said that I should wear a suit and have a mustache. Dress like a man because I'm the mayor . He said a girl mayor would be weird.	Nauczyciel mówi, że powinnam mieć garnitur i wąsy, skoro jestem burmistrzem . Bo dziewczyna burmistrz wygląda dziwnie.	Nie. Nauczyciel powiedział, że powinnam mieć garnitur i wąsy. Ubrać się jak mężczyzna , bo pani burmistrz byłaby dziwna.	
CHARLIE	Women can be mayors .	Kobiety mogą być burmistrzami .	Przecież są <u>burmistrzynie</u> .	
ALEX	I told him that, but...	Mówiłam mu.	Mówiłam mu, ale...	
AMY	It's a play about an alien invasion and the female mayor's the weird part?	W sztuce o inwazji kosmitów dziwna jest kobieta burmistrz?	To sztuka o inwazji kosmitów i <u>burmistrzynie</u> jest dziwna?	

CHARLIE	Screw that guy!	[śmieje się] Chrzanić go.	Walić tego gościa.	
(...)				
CHARLIE	Okay, take my hands. Repeat after me: I am the mayor of Humdrum Falls.	Chodź. Weź mnie za rękę. I powtarzaj za mną. Jestem <u>burmi-strzynia</u> Humdrum Falls.	Weź mnie za rękę. Powtarzaj za mną. Jestem <u>burmi-strzynia</u> .	
ALEX	I am the mayor of Humdrum Falls.	[Alex powtarza]	Jestem <u>burmi-strzynia</u> .	
CHARLIE	I am the mayor of Humdrum Falls.	[powtarzają na przemian]	Jestem <u>burmi-strzynia</u> .	
ALEX	I am the mayor of Humdrum Falls.		Jestem <u>burmi-strzynia</u> .	
CHARLIE	I am the mayor of Humdrum Falls!	[powtarzają energicznie]	Jestem <u>burmi-strzynia</u> .	



ALEX	I am the mayor of Humdrum Falls!		Jestem <u>burmi-strzynią</u> .	 <p>[powtarzają energicznie]</p>
CHARLIE	I am the mayor of Humdrum Falls! Okay. You got this.	Masz to!	Jestem <u>burmi-strzynią</u> . Dasz radę.	 <p>[Charlie] Masz to!</p>

Table 2: Multimodal analysis juxtaposing the ST subtitles against the Polish subtitles with the visual context (*The Good Nurse*, 2022 ©Netflix)

5. DISCUSSION

Instantiations of how audiovisual translation and gender converge have continued to attract scholarly attention. Von Flotow and Josephy-Hernández (2019, p. 300) talk about three main isolatable “approaches to studying questions of gender in audiovisual products.” They observe that these approaches “derive from the critical feminist thinking of the 1970s and display a certain advocacy: the first focuses on feminist materials in Anglo-American audiovisual products and their translation into Romance languages; the second studies the differences between subtitled and dubbed versions of Anglo-American source texts; and the third looks at gay and queer source text materials and their treatment in translation.” As the study discussed in this article could be roughly matched to the second category, we simultaneously tap the affordances that come with electronically-searchable corpora – identifying larger-scale patterns that are inaccessible through manual inspection of small subtitling samples. In doing so, the findings correspond with the postulate of Prewitt-Freilino *et al.* (2012, p. 280) which – although voiced over a decade ago – still holds today: “[t]o truly understand the intersection of gender in language,

cognition, and culture, researchers will ultimately need to draw connections between large-scale cross-cultural trends, cognitive process models, and experimental research on interpersonal behavior.”

Given how much subtitles users consume – and what proportion of an individual’s daily reading this amounts to for certain user profiles – as well as how much subtitling choices could be reinforcing or reshaping linguistic, cognitive and cultural patterns, it seems a necessary endeavour to produce insights into how gender is represented in subtitles. Pertinently, Di Sabato and Perri (2020, p. 369) point to the position of gender in translation by saying that “any form of translation, be it from a diachronic, synchronic, interlingual or intersemiotic perspective, illustrates how the effects of a translator’s work in identifying gender aspects of a source text, and in determining the ideological impact of gender connotations in both the source and target text, are all but neutral”.

The implications of the asymmetry we foreground in this paper have received a fair amount of scholarly attention in the work that probes the relationship of language and cognition. Phillips and Boroditsky (2003; *cf.* Elpers *et al.*, 2020) looked into whether the conceptualisation of inanimate objects which get grammatical gender assigned in languages like Spanish and German differs in native speakers of those languages compared to speakers of English where no such distinctions are grammatically introduced. A range of studies investigated the links between grammatical gender of a speaker’s language and the speaker’s cognitive processes and found confirmatory evidence (Vigliocco *et al.*, 2005; Cubelli *et al.*, 2011; Haertlé, 2017; Saalbach *et al.*, 2017; Imai *et al.*, 2014) while Samuel *et al.* (2019) conclude their review of over forty studies by arguing that “support for an influence of grammatical gender on concepts is strongly task- and context-dependent” and “comes for the most part from tasks that are susceptible to clear alternative explanations”. The account we currently have from research that brings together language and cognition is not conclusive and translation-centred inquiry of the type we propose, looking into the translator’s choices and, by extension, their possible effect on the viewers’ conceptualisations and cognition, can offer a valuable vantage point to supplement the evidence.

6. SUMMARY OF FINDINGS

This study has investigated the utilisation of Polish feminine forms in interlingual subtitles where the source language and target language

code gender asymmetrically. The Polish subtitles of 284 STs (~3.5 million word tokens) were parsed for 113 feminativa types (Waszyńska *et al.*, 2022) lemmatised based on Polish declension. We have detected 917 word tokens of unique feminativa cases across English-to-Polish subtitles of 178 of those STs.

Our investigation more specifically pertained to the quantifiable relations between the distribution of word tokens across word types (RQ1: frequency of studied feminatives), word tokens across the subtitle files (RQ2: use of feminatives in AVT productions), and word tokens viewed against the acceptability of feminativa¹¹ (RQ3: relation to lexeme-reception).

The key takeaways are that the subtitles used only 59% of the studied feminativa types, and the majority of use-cases [87% of tokens] represented a minority of those types [19% of lemmas] (RQ1). The subtitles also predominantly used feminativa that are most socially-acceptable, as indicated by reception evidence (Waszyńska *et al.*, 2022): the category of “positively-received” types constituted the majority of use-cases [~92% of tokens] and was the only category where most types [~74% of lemmas] were case-represented (RQ3). In this sense, the least socially-acceptable feminativa appear underused, while best-received feminativa types are overrepresented in our sample of AVT subtitles.

Moreover, 16% (28/178) of the TTs¹² used as many cases of feminativa as another 84% of the TTs (RQ2: measure of *numerosity*). When adjusting for the relative watchtime of STs, the majority of cases [63%] come from STs where a single feminativum occurred less often than every 50 minutes [<0.02 cases per minute] (RQ2: measure of *frequency*).

These results were then utilised for preliminary triangulation of quantitative measures with qualitative analysis. We examined the case of one of four translations that stood out as most prominently using feminatives. Multimodal analysis of a pertinent scene from *The Good*

¹¹ According to the high/mid/low-acceptability categories resulting from Waszyńska *et al.*'s survey (2022).

¹² We do not extrapolate our findings to the decision of subtitlers, since that would assume that each ST was subtitled by a single subtitler. This might not always be the case, *e.g.* for the entire span of a TV show season. Future research should investigate how many different subtitlers worked on the same ST translation, though, especially in SDH/regular subtitle distinction, as both of these were included in our corpus and analysed as two translations of the same ST.

Nurse (2022) produced useful methodological caveats, demonstrating why tokens may co-occur in (translated) audiovisual material.

7. RELEVANT CONSIDERATIONS AND FURTHER RESEARCH

By way of conclusion, let us return to the case from “History of Swear Words” we briefly discussed in Section 1, to illustrate some other especially notable considerations around feminative translation. The first question is about the translator’s decisions across occurrences of comparable feminatives, as similar decision-making instances are likely to re-appear throughout audiovisual productions. In episode 3 of the documentary an instance largely analogous to “I have PhD from Stanford” is deployed where another expert, Mireille Miller-Young, first offers her commentary and introduces herself as “a professor of feminist studies at UC Santa Barbara”. Here the subtitler’s choice is consistent with the choice in episode 1¹³ as the target text is “(...) jestem profesorką (...)”. Examining patterns like this appears a highly productive research endeavour.



¹³ We confirmed that all the episodes of the show were subtitled into Polish by the same person, Juliusz P. Braun.

Figure 5. Screenshot from the show *History of Swear Words* (2021 ©Netflix) – Polish subtitles with a feminative form

What is additionally remarkable about the example in the screenshot above, Miller-Young characterises herself by a reference to “feminist studies”, adding “And my specialty is sexuality and race”. These details could likely frame the speaker ideologically or at least activate certain – perhaps stereotypical – notions, also when it comes to Miller-Young’s linguistic choices. This might be potentially misleading because it is still far from clear how she would have worded the message in a language like Polish where gender gets assigned. Nonetheless, what the translator knows about speakers will naturally guide translation choices –if not necessarily in straightforward ways – and here another variable can be introduced in the case of films that feature real individuals, especially public ones, as the translator can look for information outside the film. In the case of Miller-Young the subtitler might choose to read (about) her research and make a more informed decision. An option that could be even safer would be to contact Miller-Young directly, explain the relevant translation choice and ask for advice. Needless to say, both the options might (perhaps regrettably) be variably realistic as a routine part of subtitling practice. It would therefore be valuable to obtain translation process data on how decisions are taken in cases of asymmetry like those exemplified by feminatives and how much deliberation on the part of translators goes into such decisions.

In a similar vein – as our findings show that a limited proportion of TTs yielded most of the identified feminatives – it remains to be ascertained in further research to what extent this has to do with the limited use of (profession) terms referring to female characters in some STs, and to what extent this results from decision-making patterns in particular translators or translator profiles.

Another research step is to extend the analysis beyond binary gender distinctions. Translation choices could here be again nuanced as a character that is taken to be female based on physical appearance might not (fully) identify with the female gender. Language forms preferred by individuals in film – whether fictional or not – might be difficult to ascertain just as being a woman does not invariably entail a preference for a feminative form.

ACKNOWLEDGEMENTS

This research received support from the Faculty of Philology Scientific Development Fund (University of Lodz) as part of the project “Dimensions of inclusivity: choices in audiovisual translation” awarded to Mikołaj Deckert.

LIST OF FIGURES AND TABLES

- *Figure 1.* Screenshot from *History of Swear Words* (2021 ©Netflix); Polish subtitles contain a feminativum.
- *Table 1.* Feminativa across their relevant lemmas, with translations into English.
- *Figure 2.* Detected feminativa by their STs, ranked by absolute number of occurrences.
- *Figure 3.* Detected feminativa by their STs, binning by relative occurrence frequency per watchtime.
- *Figure 4.* Feminativa across lemmas (0-token types not visualised) and their social acceptability categories (Waszyńska *et al.* 2022).
- *Table 2.* Multimodal analysis juxtaposing the ST subtitles against the Polish subtitles with the visual context (*The Good Nurse*, 2022 ©Netflix).
- *Figure 5.* Screenshot from the show *History of Swear Words* (2021 ©Netflix) – Polish subtitles with a feminine form.

REFERENCES

- Altun, Lokman (2019). A corpus based study: Analysis of the positive reviews of Amazon.com users. *Advances in Language and Literary Studies*, 10 (1), 123-128. <http://dx.doi.org/10.7575/aiac.all.v.10 n.1 p.123>
- Anthony, Laurence (2019). *Laurence Anthony's AntConc* [concordancer software program]. <https://www.laurenceanthony.net/software/antconc/>

- Bruti, Silvia (2009). Translating Compliments and Insults in the Pavia Corpus of Film Dialogue: Two Sides of the Same Coin? In Maria Freddi, and Maria Pavesi (Eds.), *Analysing Audiovisual Dialogue. Linguistic and Translational Insights* (pp. 143-163). Clueb.
- Boyd, Joel, Schaefer, Sara, Stoddard, Ben (2021). *History of Swear Words*. Netflix. <https://www.netflix.com/title/81305757>
- Bywood, Lindsay, Volk, Martin, Fishel, Mark, and Georgakopoulou, Panayota (2013). Parallel Subtitling Corpora and Their Application in Machine Translation and Translatology. *Perspectives: Studies in Translatology*, 21 (4), 577-594. Special Issue: *Corpus Linguistics and Audiovisual Translation: In Search of an Integrated Approach*. <https://doi.org/10.1080/0907676X.2013.831920>
- Cubelli, Roberto, Paolieri, Daniela, Lotto, Lorella, and Job, Remo (2011). The effect of grammatical gender on object categorization. *J. Exp. Psychol. Learn. Mem. Cogn. (Journal of Experimental Psychology Learning Memory and Cognition)*, 37, 449-460. <https://doi.org/10.1037/a0021965>
- Deckert, Mikołaj, and Augustyn, Rafał (2022). From film reception to translation production: Suboptimal visual-verbal coding. *Translation and Interpreting*, 14 (2), 84-102. <https://doi.org/10.12807/ti.114202.2022.a06>
- Di Sabato, Bruna, and Perri, Antonio (2020) Grammatical gender and translation. A cross-linguistic overview. In Luise von Flotow and Hala Kamal (Eds.), *The Routledge Handbook of Translation, Feminism and Gender* (pp. 363-373). Routledge. <https://doi.org/10.4324/9781315158938-32>
- Freddi, Maria (2009). The Phraseology of Contemporary Film Speech: Formulaic Language and Translation. In Maria Freddi and Maria Pavesi (Eds.), *Analysing Audiovisual Dialogue. Linguistic and Translational Insights* (pp. 101-123). Clueb.

- Gore, Genevieve. (2023) *Guides: Text Mining for Searching and Screening the Literature: AntConc*. McGill University Library Guides. <https://libraryguides.mcgill.ca/text-mining/AntConc>
- Haertlé, Izabella (2017). Does grammatical gender influence perception? A study of polish and French speakers. *Psychol. Lang. Commun.* 21, 386-407. <https://doi.org/10.1515/plc-2017-0019>
- Imai, Mutsumi, Schalk, Lennart, Saalbach, Henrik, and Okada, Hiroyuki (2014). All giraffes have female-specific properties: Influence of grammatical gender on deductive reasoning about sex-specific properties in German speakers. *Cogn. Sci. (Cognitive Science: A multidisciplinary journal)*, 38, 514-536. <https://doi.org/10.1111/cogs.12074>
- Jiménez Hurtado, Catalina, and Soler Gallego, Silvia (2013). Multimodality, Translation and Accessibility: A Corpus-Based Study of Audio Description. *Perspectives: Studies in Translation Theory and Practice*, 21 (4), 577-594. Special Issue: *Corpus Linguistics and Audiovisual Translation: In Search of an Integrated Approach*. <https://doi.org/10.1080/0907676X.2013.831921>
- Lakoff, George (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press. <https://doi.org/10.7208/Chicago/9780226471013.001.0001>
- Langacker, Ronald Wayne (2007). Cognitive Grammar. In Dirk Geeraerts and Hubert Cuyckens (Eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 421-462). Oxford University Press. <https://doi.org/10.1016/B0-08-044854-2/02029-0>
- Langacker, Ronald Wayne (2008). *Cognitive Grammar: A Basic Introduction*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195331967.001.0001>
- Lewandowska-Tomaszczyk, Barbara (1987). *Conceptual Structure, Linguistic Meaning, and Verbal Interaction*. Łódź: Łódź University Press.

- Lewandowska-Tomaszczyk, Barbara (2010). Re-conceptualization and the emergence of discourse meaning as a theory of translation. In Barbara Lewandowska-Tomaszczyk and Marcel Thelen (Eds.), *Meaning in Translation* (pp. 105-147). Peter Lang. <https://doi.org/10.3726/978-3-653-00940-8>
- Lindholm, Tobias, Wilson-Cairns, Krysty Graeber, Carles Franklin, Scott Aronofsky, Darren Jackman, Michael, and Lee Lipes, Jody (2022). *The Good Nurse*. Netflix. <https://www.netflix.com/title/81260083>
- Matamala, Anna, and Villegas, Marta (2016). Building an Audio Description Multilingual Multimodal Corpus: The VIW Project. In Jens Edlund, Dirk Heylen, and Patrizia Paggio (Eds.), *Proceedings of Multimodal Corpora: Computer Vision and Language Processing (MCC 2016)* (pp. 29-32). ELRA Language Resources Association <http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-MCC-2016-proceedings.pdf>
- Ministry of Labour and Social Policy (2014, 7 August). *Rozporządzenie Ministra Pracy i Polityki Społecznej z dnia 7 sierpnia 2014 r. w sprawie klasyfikacji zawodów i specjalności na potrzeby rynku pracy oraz zakresu jej stosowania* (2014) (Poland). Internetowy System Aktów Prawnych. <https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20140001145>
- Monti, Silvia (2009). Codeswitching and Multicultural Identity in Screen Translation. In Maria Freddi, and Maria Pavesi (Eds.), *Analysing Audiovisual Dialogue. Linguistic and Translational Insights* (pp. 165-183). Clueb. <https://doi.org/10.1400/133813>
- Netflix (2022, 6 November). *Top 10 By Country: Weekly Top 10 lists of the most watched TV and films in countries around the world*. Netflix. <https://top10.netflix.com/poland>
- Netflix Partner Help Center (2023, 15 March). *What is a Forced Narrative Subtitle?* Netflix. <https://partnerhelp.netflixstudios.com/hc/en-us/articles/224198488>

- Pavesi, Maria (2014). The Pavia Corpus of Film Dialogue: A Means to Several Ends. In Maria Pavesi, Maicol Formentelli, and Elisa Ghia (Eds.), *The Languages of Dubbing. Mainstream Audiovisual Translation in Italy* (pp. 29-55). Peter Lang. <https://doi.org/10.3726/978-3-0351-0809-5>
- Prewitt-Freilino, Jennifer, Caswell, Andrew, and Laakso, Emmi (2012). The gendering of language: a comparison of gender equality in countries with gendered, neutral gender, and genderless languages. *Sex Roles*, 66, 268-281. <https://doi.org/10.1007/s11199-011-0083-5>
- Reviere, Nina (2018). Tracking multimodal cohesion in Audio Description: Examples from a Dutch audio description corpus. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 14, 22-35. <https://doi.org/10.52034/lanstts.v17i0.477>
- Rica Peromingo, Juan Pedro, Albarrán Martín, Reyes, and García Riaza, Blanca (2014). New Approaches to Audiovisual Translation: The Usefulness of Corpus Based Studies for the Teaching of Dubbing and Subtitling. In Elena Bárcena, Timothy Read, and Jorge Arús Hita (Eds.), *Languages for Specific Purposes in the Digital Area* (pp. 303-322). Springer-Verlag. https://doi.org/10.1007/978-3-319-02222-2_14
- Saalbach, Henrik, Imai, Mutsumi, and Schalk, Lennart (2012). Grammatical genders and inferences about biological properties in German-speaking children. *Cognitive Science*, 36, 1251-1267. <https://doi.org/10.1111/j.1551-6709.2012.01251.x>
- Salway, Andrew (2007). A Corpus-Based Analysis of Audio Description. In Jorge Díaz Cintas, Pilar Orero, and Aline Remael (Eds.), *Media for All: Subtitling for the Deaf, Audio Description and Sign Language* (pp. 151-174). Rodopi.
- Schmid, Hans-Jörg (2015). A blueprint of the Entrenchment-and-Conventionalization Model. *Yearbook of the German Cognitive Linguistics Association*, 3 (1). 3-26. <https://doi.org/10.1515/gcla-2015-0002>

- Sotelo Dios, Patricia (2015). Using a Multimedia Corpus of Subtitles in Translation Training. Design and Application of the Viega Corpus. In Agnieszka Lenko-Szymanska and Alex Boulton (Eds.), *Multiple Affordances of Language Corpora for Data-Driven Learning* (pp. 238-264). John Benjamins.
- Valentini, Cristina (2006). A Multimedia Database for the Training of Audiovisual Translators. *The Journal of Specialised Translation*, 06, 68-84. https://www.jostrans.org/issue06/art_valentini.php
- Valentini, Cristina (2008). Forlì 1 – The Forlì Corpus of Screen Translation: Exploring Macrostructures. In Delia Chiaro, Christine Heiss, and Chiara Bucaria (Eds.), *Between Text and Image. Updating Research in Screen Translation Studies* (pp. 37-50). John Benjamins. <https://doi.org/10.1075/btl.78.06val>
- Vigliocco, Gabriella, Vinson, David, Paganelli, Federica, and Dworzynski, Katharina (2005). Grammatical gender effects on cognition: implications for language learning and language use. *Journal of Experimental Psychology: General*, 134, 501-520. <https://doi.org/10.1037/0096-3445.134.4.501>
- von Flotow, Luise and Josephy-Hernández, Daniel E. (2019). Gender in audiovisual translation studies. Advocating for gender awareness, in Luis Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation* (296-311). Routledge.
- Waszyńska, Katarzyna, Klon, Weronika, and Zachorska, Martyna Faustyna (2022, 21 September). *Studenci wobec feminatywów. (Nie)akceptowana przestrzeń* [Conference aper presentation]. XI Ogólnopolski Zjazd Pedagogiczny: Przesilenie – Budujmy Lepszy Świat w Sobie i Pomiędzy Nami. Poznań, Poland. https://web.archive.org/web/20230414103126/https://zjazdpedagogiczny2022.pl/dokumenty/program_3.pdf