

Youth speech in translated fiction: a corpus-based comparison of selected pragmatic markers in Catalan and Spanish

ADRIANA RAYA PALMER
University College Dublin
adriana.rayapalmer@ucd.ie

Abstract: The aim of this study is to explore fictional youth speech in translated novels in contrast to real youth speech in spoken conversation. Specifically, this analysis focuses on a selection of pragmatic markers as one of the many features of orality that can be found in literary dialogue. These markers function as expressive turn-management units between characters and are particularly prevalent in youth speech. The study examines the frequency and distribution of the pragmatic markers of interest in two languages, Catalan and Spanish, using a parallel corpus of translated dialogues from contemporary novels and two spoken corpora. The results show that, in line with previous descriptions of fictional orality, pragmatic markers are less common in fictional youth speech than they are in real conversation. However, there are some exceptions that highlight the characteristics of translated language and the literary traditions of Catalan and Spanish.

Keywords: translation, fictional orality, pragmatic markers, corpora, youth speech

El lenguaje juvenil en la ficción traducida: una comparación basada en corpus de una selección de marcadores pragmáticos en catalán y español

Resumen: El objetivo de este estudio es explorar el habla juvenil ficticia en novelas traducidas, en comparación con el habla juvenil real en la conversación oral. En concreto, el análisis se centra en una selección de marcadores pragmáticos, uno de los muchos rasgos de la oralidad presentes en los diálogos de la ficción. Estos marcadores, que funcionan como elementos expresivos y de manejo de los turnos de habla, son particularmente frecuentes en el habla juvenil. El presente estudio analiza la frecuencia y distribución de los marcadores pragmáticos en

dos lenguas, el catalán y el español, mediante un corpus paralelo de diálogos traducidos de novelas contemporáneas y dos corpus orales. Los resultados indican que, acorde con las descripciones previas de la oralidad ficticia, los marcadores pragmáticos son menos frecuentes en el habla juvenil ficticia que en las conversaciones reales. Sin embargo, ciertas excepciones destacan las particularidades del lenguaje traducido y las tradiciones literarias del catalán y el español.

Palabras clave: traducción, oralidad ficticia, marcadores pragmáticos, corpus, habla juvenil

1. Introduction

Pragmatic markers (PMs) that are typical of informal speech, such as *well, like, so, yeah*, are commonly used in the creation of fictional dialogue in works of fiction to evoke orality. In everyday conversation, PMs are crucial to guide discourse, negotiate turns, or to signal the stance of the speaker, given that spoken interaction is commonly unplanned and spontaneous. In contrast, fictional texts are carefully planned, yet still authors introduce PMs in their work, especially in the direct speech of characters to make dialogue more life-like.

While this is a stylistic technique used commonly for all types of characters, it is especially productive when representing adolescence. In many societies, teenagers are openly stigmatized by older generations by how they speak, an attitude that is present in popular media and art in the form of stereotypes or gags. An example is young Millat in Zadie Smith's *White Teeth* (2001), who overuses the tag *yeah?*, as can be seen in the example below. Thus, another function of PMs emerges: they are used for discourse and stance purposes, to evoke speech in fiction, but they can also be used to recreate effective stereotypes of speakers, especially teenagers.

- (1) I just say, **yeah?** One for Bradford, **yeah?** You got some problem, **yeah?** Speaka da English? This is King's Cross, **yeah?** One for Bradford, innit? [YouLiL_EN, WT, 74]

Youth speech can be considered a social variety of a language, strongly linked to diatopic, or geographical, varieties. As such, its use in a work of fiction is effective in indexing social and emotional cues for the intended reader in the context where the text is published. Consequently, and as has been reported extensively in the literature, relaying the particularities of youth speech in another language

proves a considerable challenge for the translator, who will have to consider an array of constraints and priorities linked to the target language and the target context (see Van Coillie 2012). In this paper, the focus is on Catalan and Spanish as target languages. These languages are comparable as they are formally, geographically, and socially close. However, there is an unequal influence of one over the other, especially as regards youth language: it has been shown that Catalan youth speech makes use of Spanish loanwords or code-switching into Spanish (Pujolar 1997), while Spanish youth speech, in general, is not influenced by Catalan in the same way.

This paper aims to explore the degree of orality in translated fictional dialogue, with a focus on PMs in youth speech. The method consists of carrying out a quantitative comparison of a selection of PMs in a parallel corpus of translated dialogues in Catalan and Spanish, in comparison to a corpus of real youth speech in Catalan and another in Spanish.

2. Pragmatic markers in fiction and their translation

The term “pragmatic marker” is an umbrella term that covers all those words with little propositional meaning that exist outside of the syntactic structure, that guide discourse, be it in the capacity of organizing or of expressing the speaker’s stance on what is being said, and that are usually short (Aijmer & Simon-Vandenberg 2011). Also known as discourse markers, inserts, or small words, their classification is still a matter of debate among linguists. To contextualize the analysis in this paper, let us examine the role of PMs in youth speech and its representation in fiction, and how they have been studied in the context of translation.

2.1. *PMs in youth speech*

The study of youth speech has been of interest to linguists in the past decades chiefly due to its innovative nature and the social situations that make it arise (e.g., Eckert 1989). Teenagers interact with their peers in structured groups, where they build their identity in relation to each other as they leave childhood. This is the ideal context for new linguistic forms to emerge, although few spread to the speech of other generations.

In their extensive account of the speech of British teenagers, Stenström et al. (2002) establish a set of features that define youth language and that they name “Slanguage” or “slangy language”. It is composed by slang words, dirty words, vague words, among others,

and, lastly, “small words”, i.e., PMs. The authors focus especially on invariant tags (*eh, okay, yeah, right, innit*), due to their high frequency of use in teenage conversation, a feature that they even suggest might be a “universal of teenage talk” (Stenström *et al.* 2002: 166). Another in-depth account of youth speech is carried out by Tagliamonte (2016), who delves into “sentence starters” (*like, well, so*) and “sentence enders” (*whatever, you know*).

Research on youth language in Spanish and Catalan often goes hand in hand with the study of informal or colloquial language and it has focused extensively on lexical aspects, such as word formation or swearing (see Regueiro Rodríguez 2023). In Spanish, there are studies on PMs in youth speech, such as the description of vocatives (Stenström 2020), or innovative features like *en plan* (‘like’) (De Smet & Enghels 2020), to name a few. In Catalan, however, the study of colloquial and youth language takes the form of comprehensive descriptive studies of communicative situations, where PMs are not usually the main concern (e.g., Bernal & Sinner 2009).

2.2. Translating PMs

In fictional orality, PMs convey metalinguistic and turn-management strategies, making them fundamental items in the creation of dialogue. Publitz (2017) defines orality in fiction as “reduced orality”, since the features that index speech—coordination rather than subordination, generalised vocabulary, hesitation and self-correction, among others—are purposefully placed in the dialogue for stylistic effect. Therefore, while PMs are essential, they do not appear at the same frequency as they do in real speech (see also López Serena 2007 on the features of colloquial Spanish in literary fiction). Bednarek (2010) suggests that PMs are less frequent in fiction because, as non-lexical items, they are not productive towards advancing narration. In turn, this also makes them more noticeable when they are used and are thus a productive strategy for characterization.

As with most features of orality, the context-boundness and multiple functions of PMs render them challenging to translate. Practices that are observed in translated target texts at large, such as explicitation, standardization, lexicalization, or omission, are also observed in the translation of PMs in particular (González 2012). On the other hand, the most prevalent markers are usually translated as markers in the target language, but not necessarily in a one-to-one fashion (see González Villar & Arias Badia 2017). Other studies have found that the more stable the function of a marker, the more homogeneous its translation (Brumme & Schmid 2021).

The target language and target context of a translation conditions the translation choices or solutions that are taken by the translator. Translating spoken, colloquial, or youth language leads to a negotiation between what is expected in the tradition of literary translation in a given context and what is closest to communication between real speakers. In the case of Catalan, this distance between fiction and reality is heightened by the presence of Spanish and English code-switching and loanwords in youth speech. Regardless of the presence of foreign languages, translators, writers, and scholars advocate that colloquial Catalan in fiction can be more transgressive with the strategies that spoken Catalan lends itself to, without having to resort to Spanish or English (see Ainaud *et al.* 2020; Cabal Guarro 2024; Gurt 2024).

3. Methods

This study relies on the comparison of a selection of PMs across several corpora. Youth speech in fiction is observed in the YouLiL corpus¹, a parallel corpus of dialogues translated from English into Spanish and Catalan that have been extracted from five contemporary novels: *White Teeth* (Smith 2001), *Middlesex* (Eugenides 2002), *I am the Messenger* (Zusak 2002), *Paper Towns* (Green 2008), and *The Casual Vacancy* (Rowling 2012)². The novels depict young characters in realistic, urban settings and across Western English-speaking regions. The token count in the Catalan subcorpus is 22 307 tokens, while in the Spanish subcorpus it is 20 920 tokens.

The dialogue is compared to reference corpora of real spontaneous conversation between young speakers. For Catalan, the COC corpus is used (initials in Catalan for Spoken Corpus of Colloquial Conversation), a corpus compiled by researchers at Universitat de Barcelona (accessible in Payrató & Alturo 2002). COC consists of conversations in informal settings, namely gatherings between friends or family, that were recorded between 1993 and 1997. For this study, only turns by young people aged 14-29 were considered: a total of five conversations and 18 123 tokens.

The Spanish subcorpus of dialogues is compared to sections in Val. Es.Co 3.0 (Pons Bordería 2024), a corpus of colloquial conversation in Spanish created by Universitat de València. This is a larger corpus, with conversations that span from 1989 to 2022. For this analysis, however, only seven conversations from 1994 and 1996 were chosen, given that they explicitly portrayed young speakers (mostly university students,

¹ See Raya Palmer (2023: 3.2 and 3.3) for a detailed description of the YouLiL corpus.

² These novels will henceforth be abbreviated as WT, MS, IATM, PT, and TCV respectively.

in the age group 18-34), and to ensure comparability with COC. The total token count is 16 673.

The transcriptions of the selected conversations were extracted from COC and Val.Es.Co 3.0 and were introduced into #Lancsbox (Brezina et al. 2020), a language data analysis software that enables the quantification of words and n-grams (i.e., co-occurrences of words). The PMs were chosen, initially, according to the most relevant markers for the creation of youth speech found in the Catalan subcorpus of translated dialogues in Raya Palmer (2023: 4.2.2.3). The markers were contrasted in fictional and real speech, and in Spanish and Catalan, with the aim of observing their frequency and functions across corpora. In sum, this is a form-to-function methodology, which, as Aijmer points out, “has the advantage that linguistic elements (...) can be studied with great precision” (2020: 30).

Comparing the frequencies of lexical items across corpora requires the normalization of frequencies (in this case, per 10k tokens). It is also crucial to consider the distribution of items within the sections of the different corpora; to do so, the percentage coefficient of variation (CV %) is observed, with > 50 % as the threshold for uneven distribution (Brezina 2018: 51). One of the benefits of dealing with small corpora is that the automatic analysis can be complemented with a manual analysis, which is especially valuable to ensure the correct selection of PMs, as they often share forms with lexical words (e.g., *mira* ‘look’, can be used as a marker or as a verb).

4. Analysis

As an introduction to the results, Figure 1 below presents the relative frequencies per 10k tokens across fictional dialogue (FD) and spoken conversation (SC) and in the Spanish and Catalan corpora. The focus is on *doncs*, *pues* and *bueno* (‘well’); *és que* and *es que* (‘it is that’); and the question tags *oi?*, *no?*, *val?*, *eh?* and *¿verdad?*, *¿no?*, *¿vale?*, and *¿eh?*. At first glance, it appears that most items are more frequent in spoken conversation than in fictional dialogue, as is expected. However, there are some units—*doncs*, *oi?*, and *val?*, in Catalan, and *¿verdad?*, and *¿vale?* in Spanish—that are more frequent in fictional dialogue.

The subsections that follow provide a more detailed description of each marker, alongside the values of their frequency and distribution across the corpora.

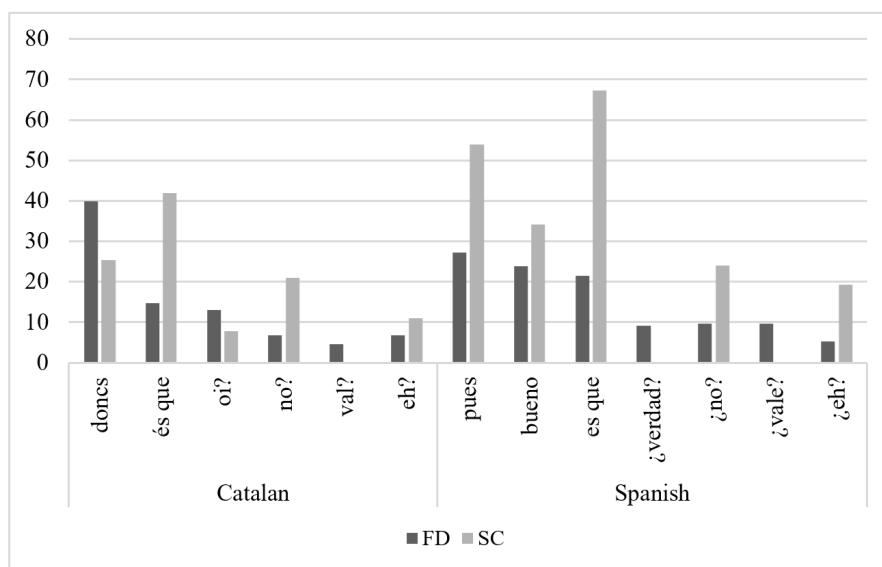


Figure 1. Relative frequencies per 10k tokens of PMs in FD and SC in Catalan and Spanish

4.1. Doncs, pues and bueno

In the Catalan subcorpus of fictional dialogues, *doncs* is one of the most frequent markers, as are *pues* and *bueno* in the Spanish parallel subcorpus. These are often translated from *well* in the source texts (ST) in English, which is used widely as “a discourse boundary marker, a response utterance initiator, a generalized starter, and an attention-getter” (Tagliamonte 2016: 111). Similarly, *doncs* can be used as a connector that signals continuity with what has been said previously (IEC 2023: 26.4.1) and *pues* introduces new information (Briz et al. 2008). *Bueno* can be used both to signal continuity with what has been said previously and to introduce complete or partial agreement (Briz et al. 2008). In the example below, the PMs introduce an explanation with a slight hesitation, both in the ST and the target texts (TT).

- | | | |
|-----|-------|---|
| (2) | ST | What'd you do to her?
Well, when she told me about Jase, I sort of shot the messenger. |
| | TT CA | I què li has fet?
Doncs quan m'ha explicat això d'en Jase, es podria dir que he matat el missatger. |
| | TT SP | ¿Qué le hiciste?
Bueno, cuando me contó lo de Jase, de alguna manera maté al mensajero. [YouLiL, PT, 225-26] |

These markers are present across subsections in the corpora, as shown by the CV% values in Table 1. While *pues* and *bueno* follow the expected pattern, whereby they are more frequent in spoken conversation than in fictional dialogue, *doncs* does not. This is most probably due to the fact that in Catalan spoken conversation the functions of introducing turns and partial agreement are also carried out by Spanish or non-standard *pues* (and variant *pos*) and *bueno*, with absolute token counts of 28 (rel. freq. 15.5) and 75 (rel. freq. 41.4) respectively. Though common in everyday speech, they are not conventionally accepted in literary dialogue. Conversely, the preferred option in literary writing *bé*, equivalent in many contexts to Spanish *bueno*, is used in Catalan fictional dialogue (albeit rarely, with only six occurrences), but does not appear once in spoken conversation.

Language	Marker	Corpus	Abs. freq.	Rel. freq.	CV %
Catalan	<i>doncs</i>	FD	78	40.0	17.0
		SC	46	25.4	11.0
Spanish	<i>pues</i>	FD	57	27.3	18.0
		SC	90	54.0	26.9
Spanish	<i>bueno</i>	FD	50	23.9	27.6
		SC	57	34.2	21.6

Table 1 Frequency and distribution values of *doncs*, *pues*, and *bueno*.

Aside from frequency, patterns in the co-occurrence of markers (see Cuenca & Marín 2009) consistently present differences between fictional dialogue and spoken conversation in both languages. In Catalan spoken conversation, *doncs* co-occurs with *ah* (*ah doncs*, 'oh, well'), *bueno* (*bueno doncs*, 'well, well'), and also non-standard *vale* (*vale doncs*, 'okay, well'). In fictional dialogue it co-occurs with *així* (*així doncs*, 'so, well'), *molt bé* (*molt bé doncs*, 'very good, well'), *mira* (*doncs mira*, 'well, look'), and *sí* (*doncs sí*, 'well yes').

In Spanish, *bueno* and *pues* often co-occur together, both in fictional dialogue and in spoken conversation. In spoken conversation, *pues* also co-occurs with *sí* (*sí pues*, 'yes, well') and *nada* (*pues nada*, 'well, nothing'), while *bueno* co-occurs most commonly with *y* (*bueno y*, 'well, and') and *es que* (*bueno es que*, 'well, the thing is'). In Spanish fictional dialogue, both *pues* and *bueno* co-occur with *sí* ('yes') and *vale* ('okay'). In sum, although these markers are present across registers, patterns of co-occurrence differ considerably. A unifying feature is that the markers are significantly more frequent as single units across all corpora than in pairs of co-occurrence.

4.2. És que and es que

És que/es que ('it is that' or 'the thing is') is used to introduce a justification of what has been said previously, an excuse, an apology or a mitigated objection. It can also introduce a reaction to an implied rejection or a polite excuse (Briz *et al.* 2008) and is sometimes used for emphasis (Marín & Cuenca 2012). This analysis considers the grammaticalized marker alone, but also *és que/es que* pseudo-cleft constructions with a generic noun or clause, such as *el cas és que* or *el caso es que* ('the case is that') (see Marín & Cuenca 2012).

In the corpora of fictional dialogue, characters mostly use this marker to clarify what they are saying, thus avoiding conflict, and as a result of hesitation, as in example 3. *És que/es que* appears in the corpus as a translation from 'I mean', 'just', and 'the thing is', but also as an addition where there is no marker in the ST, as in example 4. In contrast to the markers in the previous section, pairs in the TTs where this marker is used in both Catalan and Spanish are scarce.

- (3) ST **I just** thought... like when she says, here: Then will I swear, beauty herself is black... (...)
 TT CA **És que** em pensava... com això que diu aquí: Llavors juraré que la bellesa és negra... (...)
 TT SP Ø Yo pensé... Como aquí dice: «Entonces juraré que es negra la hermosura misma...». (...)
 (...) [YouLiL, WT, 162]
- (4) ST Okay. Sometimes I think I have bad breath. (...)
 TT CA D'acord. Ø A vegades em penso que em put l'alè. (...)
 TT SP Vale. Es que a veces me parece que me huele el aliento. (...) [YouLiL, MS, 217]

The token counts reveal that this marker is more frequent in Spanish than it is in Catalan, both in fictional dialogue and in spoken conversation; in parallel, it is used much more in spoken conversation than in fictional dialogue in both languages. As Table 2 shows, the distribution is even across corpora.

Language	Marker	Corpus	Abs. freq.	Rel. freq.	CV %
Catalan	<i>és que</i>	FD	33	14.8	9.0
		SC	76	41.9	18.5
Spanish	<i>es que</i>	FD	45	21.5	14.0
		SC	112	67.2	13.5

Table 2. Frequency and distribution values of *és que* and *es que*.

Not only do the overall frequencies reveal differing tendencies in fictional dialogue and spoken conversation, but the type of construction varies, too: in fictional dialogue, pseudo-cleft structures with general nouns make up 55 % of cases in Catalan and 47 % in Spanish; in spoken conversation, they make up only 11 % of cases in Catalan and 13% in Spanish. In other words, fictional dialogue in the corpora of this study prefers the marker *és que/es que* in pseudo-cleft structures, instead of the grammaticalized PM. The most common pseudo-cleft constructions in fictional dialogue, aside from the pair in the example below *el cas és que/lo que pasa es que* ('the thing is that'), are *la verdad es que* ('the truth is') and *lo que quiero decir es que* ('what I want to say is that') in Spanish, and its equivalent *el que vull dir és que* in Catalan. Interestingly, while this final construction does appear in the Catalan corpus of spoken conversation, its equivalent does not appear at all in the Spanish subcorpus, suggesting that it might be a construction reserved for fictional dialogue.

- | | | |
|-----|-------|--|
| (5) | ST | It's just... I wish it was easier, for me, you know? |
| | TT CA | El cas és que... Voldria que fos més fàcil, que em resultés més fàcil, ¿m'entens? |
| | TT SP | Lo que pasa es que me gustaría que fuera más fácil para mí, ¿sabes? [YouLiL, IATM, 283] |

4.3. Oi?, ¿no?, and other question tags

The final set of markers to consider are question tags. In the TTs in YouLiL, they arise when the STs use canonical and non-canonical question tags, such as *isn't he?*, *right?*, *innit?*, etc. The functions of question tags usually involve seeking confirmation from the listener about what the speaker is saying, as in the example below, inviting further information, or as a contact-check to make sure the listener is following the conversation (see Andersen 2001; Cuenca & Castellà 1995).

- | | | |
|-----|-------|---|
| (6) | ST | Cubby upset, was he? |
| | TT CA | En Cubby està molt afectat, oi? |
| | TT SP | Cuby está muy afectado, ¿no?
[YouLiL, TCV, 109] |

In contrast to the previous sets of markers, these are far less frequent overall, and are less evenly distributed, as can be appreciated in Table 3. They are worth contemplating, however, as notable contributors to the portrayal of youth stereotypes; as noted in Raya Palmer (2023) and

in the example in the start of the study, they are highly productive for character delineation.

Language	Marker	Corpus	Abs. freq.	Rel. freq.	CV%
Catalan	oi?	FD	29	13.0	31.5
		SC	14	7.7	58.4*
	no?	FD	15	6.7	55.7*
		SC	38	21.0	39.8
	val?	FD	10	4.5	100.0*
		SC	0	0.0	0.0
eh?	FD	15	6.7	37.5	
	SC	10	11.0	41.0	
Spanish	¿verdad?	FD	19	9.1	30.0
		SC	3	0.0	100.0*
	¿no?	FD	20	9.6	50.7*
		SC	40	24.0	18.6
	¿vale?	FD	20	9.6	46.8
		SC	1	0.0	100.0*
	¿eh?	FD	11	5.3	33.3
		SC	32	19.2	25.8

Table 3. Frequency and distribution values of response elicitors. Note: The asterisks (*) mark the CV% values that present an uneven distribution across subsections.

The preferred marker in Catalan fictional dialogue is *oi?*, which, like *doncs* above, is more common in fictional dialogue than in spoken conversation. It is also the most evenly distributed marker in the corpus, followed by *eh?*, *no?* and *val?*, which are more restricted, with *val?* (non-standard 'okay?') appearing in only one novel. This last marker is an exception, as the translator chooses a non-standard form to add a layer of colloquiality to the dialogue; however, it is not supported by the subcorpus of real spoken conversation in Catalan³.

In the Spanish subcorpus of dialogues, the tags present a wider variety of use, with *¿verdad?*, *¿no?* and *¿vale?* being used at similar relative frequencies across novels. *¿No?*, on the other hand, presents an uneven distribution, like in Catalan. As is the case with Catalan, *¿no?* and *¿eh?* are more frequent in spoken conversation than in fictional dialogue. In fact, *¿verdad?* and *¿vale?* present extremely low frequencies in spoken conversation, with only three and one occurrence respectively.

5. Discussion and final thoughts

This paper set out to establish potential differences between the frequency of PMs in fictional and real youth speech, in Catalan and

³ On the other hand, there are instances of code-switching to Spanish *vale*, although there are only three occurrences.

Spanish. Given that this is a broad topic, the analysis focused on a selection of PMs: the Catalan sentence starters *doncs* and *és que*; the Spanish sentence starters *pues*, *bueno*, and *es que*; the Catalan question tags *oi?*, *no?*, *val?*, and *eh?*; and the Spanish question tags *¿verdad?*, *¿no?*, *¿vale?*, *¿eh?*. These PMs were chosen due to their frequency in the corpora and their significance in portraying youth speech. The contrast between the corpora reveals that, in general, these PMs are more prevalent in the corpora of real youth speech than in the fictional corpora, thus agreeing with the insights of Bublitz (2017) and Bednarek (2010) on the reduction of PMs in fiction.

In contrast, there are two cases in Catalan that present an inverse pattern, that of *doncs* and *oi?*; markers that, though belonging to the structures of spoken interaction, appear to be more typical of fictional dialogue. Comparing dialogues to spoken conversation among different age groups could assert whether this is specific to young speakers or to Catalan speakers in general. If the latter is true, a preference for certain markers of orality in fictional dialogue might imply the existence of an established style in Catalan fictional dialogue, which can be differentiated from both spoken Catalan and from literary Catalan.

The repertoire of question tags in the fiction corpora suggests varied and unfixed translation solutions. There is a deliberate intent of reproducing the wide range of tags in the STs, despite a clear preference in spoken conversation in both languages for *no* and *eh*. Another result that stems from the translation process is the preference for explicitation in the *és que/les que* pseudo-cleft sentences, instead of the more informal grammaticalized marker. These findings lead one to consider other consequences of the translation process: for instance, there seems to be a higher frequency of PMs in Spanish spoken corpora than in Catalan corpora, yet the translated dialogues do not follow these distributions to the same extent. Presumably, this is due to the influence of the STs in English. Further research could study the distribution of these markers in STs in Catalan and Spanish, to assert whether they are closer to patterns in real conversation than the ones in this paper.

This study has contributed to the definition of fictional dialogue, especially in relation to youth speech. However, it is important to note that these results are particular to the data analyzed. While the data in the corpus of fictional dialogues is supra-local and follows the conventions of translated literature across Catalan and Spanish-speaking regions, the COC and Val.Es.Co 3.0 corpora are specific to the regions where they were created. In this sense, the patterns in this study would benefit from being contrasted to larger corpora with conversations by speakers of different language varieties of Catalan and Spanish.

6. Funding

This article is part of the Youth Engagement in European Language Preservation Project, which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 802695).



Funded by
the European Union



European Research Council
Established by the European Commission

REFERENCES

- Aijmer, Karin (2020), "Contrastive Pragmatics and Corpora", *Contrastive Pragmatics* 1 (1): 28-57. DOI: 10.1163/26660393-12340004.
- Aijmer, Karin & Anne-Marie Simon-Vandenberg (2011), "Pragmatic markers", in Jan Zienkowski, Jan-Ola Östman & Jef Versuieren (eds.), *Discursive Pragmatics*, Amsterdam, John Benjamins Publishing Company: 223–247.
- Ainaud, Jordi, Anna Espunya & Dídac Pujol (2020), *Manual de traducció anglès-català*, Barcelona, Universitat Pompeu Fabra.
- Andersen, Gisle (2001), *Pragmatic Markers and Sociolinguistic Variation: A Relevance-Theoretic Approach to the Language of Adolescents*, Amsterdam & New York, John Benjamins Publishing Company.
- Bednarek, Monika (2010), *The Language of Fictional Television: Drama and Identity*, London and New York, Continuum.
- Bernal, Elisenda & Carsten Sinner (2009), "Al seu rotllo: Aproximació al llenguatge juvenil català", *Zeitschrift für Katalanistik*, 22: 7-36. DOI: 10.46586/ZfK.2009.7-36.
- Brezina, Vaclav (2018), *Statistics in Corpus Linguistics*, Cambridge, Cambridge University Press.
- Brezina, Vaclav, Pierre Weill-Tessier & Tony McEnery (2020), #LancsBox 5.x and 6.x. [software package].

- Briz, Antonio, Salvador Pons & José Portolés (coords.) (2008), *Diccionario de partículas discursivas del español*, Available at: www.dpde.es. [Accessed: August 3rd, 2024].
- Brumme, Jenny & Beatrice Schmid, (2021), "Convergències i divergències en l'ús de les partícules discursives", *Actes Del XVIIIè Col·loqui de l'AILLC*, 96–108. DOI: 10.2436/15.8090.02.6.
- Bublitz, Wolfgang, (2017), "Oral features in fiction", in Mariam Locher & Andreas. H. Jucker (eds.), *Pragmatics of Fiction*, Berlin and Boston, De Gruyter: 235-263.
- Cabal Guarro, Miquel (2024), "Tenim de traduir-ho, aixòs: la (re)creació del col·loquial en la traducció literària", in Elisenda Bernal (ed.), *Col·loquial(s): Estudis de l'ús del català actual*, Leipzig, Leipziger Universitätsverlag: 101-118.
- Cuenca, Maria Josep & Josep Maria Castellà (1995), "Una caracterització cognitiva de les preguntes confirmatòries (*question tags*)", *Caplletra*, 18(Spring): 65–84.
- Cuenca, Maria Josep & Maria Josep Marín (2009), "Co-occurrence of discourse markers in Catalan and Spanish oral narrative", *Journal of pragmatics*, 41 (5): 899-914. DOI: 10.1016/j.pragma.2008.08.010.
- De Smet, Emma & Renata Enghels (2020), "Los datos en Twitter como fuente del discurso oral coloquial: Estudio de caso del marcador discursivo *en plan*", *Oralia*, 23 (2): 199-218. DOI: 10.25115/oralia.v23i2.6379.
- Eckert, Penelope (1989), *Jocks and Burnouts: Social categories and identity in the high school*, New York, Teachers College Press.
- González, Montserrat (2012), "Pragmatic markers in translation", in Jenny Brumme & Anna Espunya (eds.), *The Translation of Fictive Dialogue*, Amsterdam and New York, Rodopi: 218-232.
- González Villar, Alejandro & Blanca Arias Badia, (2017), "Marcadors conversacionals en la traducció literària alemany-català: *also i na a Jeder stirbt für sich allein*", *Zeitschrift Für Katalanistik*, 30, 245–267. DOI: <https://doi.org/10.46586/ZfK.2017.245-267>.
- Gurt, Carlota (2024), "El català col·loquial des de la trinxera de l'escriptor: la teoria de l'equalitzador", in Elisenda Bernal (ed.), *Col·loquial(s): Estudis de l'ús del català actual*, Leipzig, Leipziger Universitätsverlag: 87-100.
- Institut d'Estudis Catalans (IEC) (2023), *Gramàtica essencial de la llengua catalana (GEIEC)*, [online ed.], Accessible at: <https://geiec.iec.cat/>.

- López Serena (2007), *Oralidad y escrituralidad en la recreación literaria del español coloquial*. Madrid, Gredos.
- Marín, Maria Josep & Maria Josep Cuenca (2012), “De l’atribució a la modalitat: Construccions amb és que en català oral”, *Caplletra* (Spring): 65–94. DOI: <https://doi.org/10.7203/caplletra.52.4699>.
- Raya Palmer, Adriana (2023), *The Representation of Youth Language in Fictional Dialogue in English and its Translation into Catalan*, tesis doctoral, Universitat Pompeu Fabra.
- Payrató, Lluís & Núria Alturo (2002), *Corpus oral de conversa col·loquial. Materials de treball*, Barcelona, Publicacions de la Universitat de Barcelona.
- Pons Bordería, Salvador (2024), *Corpus Val.Es.Co 3.0.*, Available at: <http://www.valesco.es>. [Accessed: July 21st, 2024].
- Pujolar, Joan (1997), *De què vas, tio?*, Barcelona, Empúries.
- Regueiro Rodríguez, María Luisa (2023), *Diccionario del léxico juvenil en España: del lenguaje juvenil al español coloquial*. Navarra, EUNSA. Ediciones Universidad de Navarra, SA.
- Stenström, Anna-Brita (2020), “English- and Spanish-speaking teenagers’ use of rude vocatives”, in Nico Nassenstein & Anne Storch (eds.), *Swearing and Cursing: Contexts and practices in a critical linguistic perspective*, Berlin and Boston, De Gruyter: 281-302.
- Stenström, Anna-Brita, Gisle Andersen, & Ingrid Kristine Hasund (2002), *Trends in Teenage Talk*, Amsterdam and Philadelphia, John Benjamins Publishing Company.
- Tagliamonte, Sali A. (2016), *Teen Talk: The language of adolescents*, Cambridge, Cambridge University Press.
- Van Coillie, Jan (2012), “Cool, geil, gaaf, chouette or super. The challenges of translating teenage speech”, in Martin Fischer & Maria Wirf Naro (eds.), *Translating Fictional Dialogue for Children and Young People*, Berlin, Frank & Timme: 217-232.