Combinations of discourse markers with repairs and repetitions in English, French and Spanish

Abstract

Discourse markers have a central role in planning and repairing processes of speech production. They relate with fluency and disfluency phenomena such as pauses, repetitions and reformulations. Their polyfunctionality is challenging and few form-function mappings are stable cross-linguistically. This study combines a functional and a structural approach to discourse markers and their combination with and within repetitions and self-repairs in native English, French and Spanish, in order to establish the inter-relation between these three fluency-related devices and to find potentially universal patterns of use. Qualitative coding and quantitative analyses of categories of markers and repairs allowed us to identify discourse markers which are specific to repair sequences and others which are much more pervasive. Combinations with repetitions vary across languages and repair types. Our findings fill a gap in cross-linguistic fluency research, disentangle the overlap between discourse markers, repairs and repetitions, and can be integrated into pedagogical materials.

Keywords: discourse markers, repair, repetitions, disfluency, English, French, Spanish

Highlights:

- Discourse markers and repetitions combine within and outside repairs, although not in the same proportions or with the same formal and functional categories.
- Discourse markers in repair sequences are more varied in Romance languages than in English.
- In each language, some discourse markers are specific to repair sequences (En. *or*, Fr. *ou*, Sp. *digamos*) whereas others are more pervasive within and outside repairs (e.g. En. *well*, Fr. *enfin*, Sp. *o sea*).

Introduction

Discourse markers (henceforth DMs) have been the focus of a strong – and still growing – trend in pragmatics investigating in particular their polyfunctionality and context-sensitivity (e.g. Schiffrin 1987). These frequent expressions (such as *you know*, *well* or *so*) are characterized by their role as structuring, addressee-oriented cues for interpretation and by their flexible syntactic status (formally varied, optional). DMs are quite paradoxical in that, while very frequent and essential for successful communication (e.g. Crystal 1988; House 2013), they often go unnoticed during an interaction or, on the other hand, can be perceived as superfluous and even detrimental if used under the wrong conditions. Several non-academic articles and

online videos provide examples of such a negative attitude towards “cringing verbal tics”, especially when used quite frequently.[1]

Yet, authors in fluency research tend to agree on the positive effects of DMs, especially in second language acquisition where they are associated with naturalness, automaticity and efficient planning strategies (Hasselgren 2002; Götz 2013). DMs are indeed intrinsically linked to fluency: they constitute windows on the cognitive processes behind speech production and perception, with many of their functions being directly connected to (dis)fluency moves (e.g. reformulation, planning).

While these general processes of production might be shared cross-linguistically, the specific linguistic encoding of form-meaning patterns is more likely to differ across languages. Such differences may lead to transfer effects (Odlin 1989, Gilquin 2008), that is, non-native-like uses of DMs in the target language inherited from uses in the mother tongue (e.g. Beeching 2012 on French-English false friends).

In this paper, we propose a fine-grained corpus-based study of the relation between DMs and fluency in three native languages, namely English, French and Spanish. We focus on the functions of DMs and on their relation to repairs and repetitions. By combining a functional and a structural approach to DMs and their context, we aim at filling a gap in cross-linguistic fluency research, especially in Spanish, which is particularly under-studied in this respect (see Pascual 2018). Our goal is to tease out cross-linguistic (potentially universal) from language-specific patterns of combination between DMs, self-repairs and repetitions by native (henceforth L1) speakers, which can further serve as a baseline for non-native (henceforth L2) language learners.

We will start by reviewing previous research on DMs in relation to L1 and L2 fluency (Section 1), as well as existing frameworks on self-repairs and repetitions (Section 2). The research questions and hypotheses will be developed in Section 3. Our corpus data and annotation method will be detailed in Section 4. The results will be discussed in Section 5, starting with an analysis of discourse markers in isolation (5.1), then combined with and within repetitions (5.2) and finally looking at discourse markers and repetitions within repairs (5.3). Lastly, we will conclude in Section 6.

1. Discourse markers

It is beyond the scope of this paper to provide a full list of DM expressions or to review previous definitions of DMs (see Maschler & Schiffrin 2015). In this study, DMs are defined as non-propositional and non-syntactically integrated fixed expressions fulfilling discourse-structuring functions (e.g. Aijmer & Simon-Vandenbergen 2011; Hansen 2006). This definition includes some adverbs (*so*, Fr. *donc*, Sp. *pues*), verb phrases (*you know*, Fr. *tu vois*, Sp. *sabes*) and particles (*yeah*, Fr. *ouais*, Sp. *sí*), among many others. More details on inclusions and exclusions will be provided in the methodology (Section 4).

1.1 The functions of discourse markers

[1] For example, a LanguageLog article is very critical of the use of *you know* and *um* by a US Senator (http://languagelog.ldc.upenn.edu/nll/?p=964, last accessed on 19/02/2018).

DMs are a widely researched object of study since Schiffrin's (1987) seminal work. They are famously challenging to define, categorize and classify consensually. In particular, their polyfunctionality, as a category and as individual expressions, has been conceptualized under a very large number of frameworks, and no agreement has yet been reached on the number and labels of meanings that discourse markers can express. Classification schemes, and the theoretical assumptions behind them, vary with the methodology and particular linguistic discipline used to describe them: from Conversation Analysis (Schiffrin's (1987) five "planes of talk") to more cognitive (Redeker's (1990) three "domains") and computational approaches, among many others. Overall, three main functional categories emerge from existing taxonomies: textual functions related to the coherence and structure of discourse (relations between utterances, topics or turns); modal functions related to the expression of the speaker's attitude or emotion; interpersonal functions, related to the speaker-hearer relationship (e.g. Briz & Pons 2010; Cuenca 2013; González 2005; Maschler 1994).

The field is also flourishing with qualitative, in-depth case studies on the meaning variation of specific discourse markers which display a wide range of meanings. These studies (e.g. Aijmer 2016 on *anyway*; Degand 2014 on French *alors* 'well/then') tend to show how function varies with form and with the communicative context. However, the polysemy of some discourse markers such as English *so* (Buysse 2012) or *well* (Cuenca 2008) cannot always be reduced to clear-cut formal contexts of use, nor is it always possible to narrow a particular instance down to one meaning only (Bunt 2011). Different theoretical views and frameworks on this polyfunctionality (e.g. polysemy vs. monosemy) are presented and discussed in the contributions to Fischer's (2006a) volume.

1.2 The paradox of discourse markers

DMs still fascinate linguists after generations of research. Part of this interest is motivated by their great formal-functional variety and ambivalence, which some might call a paradox (Degand 2016). In fact, DMs are paradoxical in more than one way. First, in spite of their very high frequency in naturally occurring discourse, DMs often go unnoticed: speakers are not always aware of their use of DMs, and listeners do not always perceive them (Götz 2013). At the same time, DMs are often mentioned as verbal tics and "bad habits", especially when someone uses the same DM repeatedly with high frequency: DMs are usually caricatural of a hesitant, "unskilful" or "powerless" speaker (O'Donnell & Todd 1980:67; Ragan 1983:166).

This negative reputation partly results from another paradoxical aspect of DMs, namely their great polyfunctionality, which ranges from speech-specific, sometimes disruptive uses to discourse-structuring or connective functions. While most DM expressions tend to specialize in one of these two extremes (e.g. *however* vs. *you know*), some – among the most frequent – are more variable, depending on the interaction settings or the linguistic context (e.g. syntactic position, prosody). For instance, *well* can be used to take the turn (textual function) or to indicate the speaker's disagreement (interpersonal function), but is also often found in the context of repairs or reformulations of errors (Aijmer 2011; Cuenca 2008; Schourup 2001).

This polyfunctionality leads to a third paradox of DMs: although highly complex and variable, DMs are rarely taught explicitly in the first or second/foreign language (Gilquin 2016; House 1996). DMs are taken as central for the quality, naturalness and overall

fluency of speech, yet somehow native and non-native speakers are supposed to master their use on their own, from natural exposure alone. It might not be necessarily obvious which DMs to use and when to use them appropriately.

1.3 Discourse markers in L1 and L2 (dis)fluency

As mentioned above, the functions of DMs vary strongly – but not always systematically – with linguistic and extra-linguistic factors such as syntactic position, prosody, co-occurrence and communicative context. As a result, while native speakers produce these expressions fairly automatically, learners find them challenging to acquire and use in a "native-like" way, especially because pragmatic elements are rarely taught explicitly and learners' exposure to naturalistic settings is limited (e.g. Bardovi-Harlig 2005). Even in L1 speech, DMs can be used to signal production trouble or to correct previously uttered material. Crible (2018) has shown that the frequency and types of DMs vary with the degree of preparation (spontaneous vs. prepared) and the degree of interactivity (free dialogue vs. monologue or quasi-monologue) of the interaction, and that, in particular, interpersonal DMs such as *you know* are more typical of contexts containing cues of disfluency (long sequences of interruptions).

While research on DMs in L1 speech is rich and diverse, studies on learner language are both scarcer and more restricted in scope (largely focused on English and targeting individual, high-frequency markers), although two research trends have started to develop in recent years. The first of these investigates discourse markers within second language fluency: these studies are mostly quantitative and compare the frequency and variety of discourse markers in native (L1) vs. non-native (L2) corpus data. This line of work relates to general issues of language acquisition such as phraseological competence (De Cock et al. 1998), fossilization (Romero Trillo 2002, Zhao 2013), proficiency levels (Hasselgren 2002) or type of input (Fung & Carter 2007, Gilquin 2016).

The second type of studies strives to identify differences in the number and type of pragmatic functions for selected discourse markers across L1 and L2 speakers. In particular, the marker *well* has been the focus of several studies (Aijmer 2011, Buysse 2015, Li & Xiao 2012). Müller (2005) and Denke (2009) provide important and detailed functional analyses of a selection of items (*so*, *well*, *you know*, *like* and *you know*, *I mean*, *well*, respectively) across German and Swedish learners of English. Most studies report on quantitative differences of some (uses of) discourse markers, identifying "teddy bear" effects (i.e. restricted number of markers; Hasselgren 2002), transfers from the L1 and disfluent uses associated with lexical access trouble or repair.

The interest in studying DMs from the perspective of fluency and disfluency is two-fold. Firstly, analyzing the use of DMs is telling of the speakers' pragmatic competence (House 2013). Secondly, DMs are the results of, on the one hand, cognitive constraints on spoken production which are supposedly shared across languages and, on the other, language-specific preferences for form-function encoding, potentially leading to non-native-like uses and transfer effects from a language to another. This study sets out to disentangle which aspects in the use of DMs in English, French and Spanish are language-specific and which others are part of a cross-linguistic "fluencicon", focusing in particular on the relation between DMs, repairs and repetitions.

2. Repairs and repetitions

Two main approaches can be distinguished in the study of repair and disfluencies in general: structural approaches, where the items are segmented on the basis of their form (e.g. a repair sequence made of a *reparandum*, an editing phase and a *reparans*); componential approaches, which are not so much interested in the overall structure of the sequence but also (or mostly) in the individual elements comprised in the sequence (e.g. a repair containing a filled pause, a truncation and a repetition). These two trends come from different theoretical and methodological frameworks and are rarely combined.

A case in point is the study by Denke (2009), where she analyzes separately the functions of i) pragmatic markers (specifically *y'know*, *I mean* and *well*), ii) repairs (self- and other-initiated) and iii) repetitions, across native and non-native speakers. She makes interesting observations about differences and similarities in the use and functions of these three phenomena across the two speaker populations. However, in her study, there is no attempt at combining or integrating DMs, repairs and repetitions into a single framework of analysis, so that we have no information on their co-occurrence tendencies.

2.1 Levelt's categories of repair

While the notion of repair was first investigated by conversation analysts (Schegloff et al. 1977), it was fully developped by Levelt (1983, 1989) as part of a larger cognitive model of speech production, which remains referential in the domain. In Levelt's (1983:44) terminology, self-repairs comprise four main parts, a *reparandum* ("item to be repaired"), a moment of interruption ("the point at which the flow of speech is interrupted"), an editing phase (also called *interregnum*, e.g. Shriberg 1994) with an optional editing term, and a "repair" (also called *reparans*, i.e. the repairing segment).

Self-repairs are the result of monitoring, the final component of Levelt's (1989) "blueprint" model of speech production, in charge of comparing the linguistic output with language standards and the speaker's intentions. They can take two main forms, namely overt or covert repairs: the former necessarily involves a change, addition or deletion of morpheme, while the latter merely constitutes an interruption point, such as pausing or repeating the same word with no change (*I went to to London*). Focusing on overt self-repairs, he identifies three main functional types:

- delay repairs (henceforth D-repairs) answer the question "do I want to say this now?" and correspond to re-arrangements of messages;
- appropriateness repairs (henceforth A-repairs) answer the question "do I want to say it this way?" and signal a need of qualification for better adequacy;
- error repairs (henceforth E-repairs) answer the question "am I making an error?" and can target lexical, syntactic or phonetic errors.

Levelt successfully showed that different types of repair are expressed by different forms, in meaningful clusters of cues which are designed to help the listener interpret the utterance.

2.2 Repetitions in fluency

Repetition is mentioned in models of reformulation targeting spoken language, including in corrective contexts corresponding to Levelt's error-repairs. For instance, De Gaulmyn (1987: 86) distinguishes between four types of rephrasing which she termed "repetition": repetition (including modifications by partial addition or substraction), delayed restart, repetition of a truncation, and repetition of self-dictation. In their conversation-analytic study, Auer & Pfänder (2007) analyze "multiple retractions" in spoken French and German. This type of structure consists in "re-us[ing] a syntactic position which has already been filled" (2007: 59), either to signal hesitation, turn-holding or list construction. Its relation to repair is made explicit: "Syntactically speaking, retraction is the basis of repair, but not all retractions do repair work, let alone correct a previous item. Retraction is also the basis of list construction, and it is used for numerous other, non-repair functions" (2007: 59). Their results indicate that retraction is used quite similarly in the two languages except for an additional rhetorical function in French that does not appear as frequently in German, a stylistic difference which the authors explain by a higher sensitivity to norms and standards in French.

There is little experimental evidence in the literature of the perceptive or cognitive effect of repetitions (e.g. Fox Tree 1995; MacGregor et al. 2009). These studies have shown that repetitions either speed up comprehension or do not alter it (positively nor negatively). Ejzenberg (2000) corroborates the fluent role of repetitions: "In psycholinguistic terms, redundancy and repetition allow a speaker to set up a paradigm and slot in new information where the frame for the new information stands ready, rather than having to be newly formulated" (2000: 299), observing that such a strategy is not mastered by low-fluency learners, as Rabab'ah & Abuseileek (2012) have also shown. Clark & Wasow's (1998) comprehensive corpus study also suggests that repetitions are the by-products of preliminary commitments the speakers make because of the "temporal imperative" or pressure to keep speaking. According to Fung (2007), this "production-based" function of repetitions, i.e. to fill a pause, is the most frequent one, besides three other categories (semantic, comprehension- and interaction-based).

In sum, authors studying repetitions have shown their ambivalence between discourse-structuring and reformulative uses. The bulk of reformulation studies, however, tends to focus on markers of reformulation (e.g. Cuenca 2003; Rossari 1994) rather than on formal structures, so that an integration of repairs, repetitions and discourse markers should fill this gap in the literature.

3. Research questions and hypotheses

The present study adopts a functional-structural approach to a wide range of DMs, their functions and their combination with repairs and repetitions in English, French and Spanish native conversations. The qualitative analysis of repair types, from Levelt's (1983) typology, and the functional disambiguation of DMs will be combined with the formal identification of repetitions within and outside repair sequences.

The quantitative study will first report on the proportion of different types and functions of discourse markers when they occur in isolation (i.e. without repetitions or repairs). We will then examine different types of combinations of DMs and repetitions outside

repairs. Thirdly, the study will integrate the three phenomena, viz. DMs, repairs and repetitions, in order to identify the specific role of DMs within repairs as well as possible cross-linguistic differences in this respect. It is hypothesized that DMs and repetitions do not often co-occur within repairs because of their functional redundancy: the repair will be either signaled by a dedicated DM (e.g. *well*) or by the repetition of previously uttered material, but not necessarily by both devices simultaneously. We also expect to find DMs and repetitions in high frequency outside repairs, since they both correspond to what Levelt (1983) terms “covert repair”, where nothing is actually repaired. Previous research suggests that the functions and uses of DMs and repetitions go beyond self-repair, and this study intends to disentangle the overlap with precise quantitative and qualitative analyses.

Throughout the analysis, cross-linguistic differences between English, French and Spanish will be systematically examined. The latter two languages belong to the same Romance family, which could suggest a greater similarity than with English. However, Crible (2018) has shown that English and French do not greatly differ on the types and functions of DMs, nor on the use of repairs and repetitions. Other studies focusing on DMs (Cuenca 2003; Cuenca & Bach 2007) have shown that Romance languages tend to make use of more polysemous markers and more complex discursives strategies than English. No further differences are *a priori* expected at this stage.

4. Data and method

4.1 Corpus data

For this study, we used a sample of about 11,000 words from spoken conversations in each of the three languages, as can be seen in Table 1.

Table 1: Corpus size by language

English	10,937 words	5 dialogues
French	11,024 words	4 dialogues
Spanish	11,149 words	3 multilogues

The samples were extracted from the following corpora: the British component of the International Corpus of English (Nelson et al. 2002), the VALIBEL corpus of French (Dister et al. 2009) and the Val.Es.Co 2.0 corpus of Spanish (Cabedo & Pons 2013). All texts are comparable: they correspond to free informal conversations between friends or relatives; the topic is freely determined by the participants, the intervention of the researcher is highly limited. As a result, the data are very natural, interactive and spontaneous.

4.2 Annotation of discourse markers

DM tokens were manually identified according to the following criteria: lexical item (excluding non-lexical fillers such as *uhm*) with a procedural meaning that does not contribute to the propositional content of the utterance and is not syntactically integrated to it. As suggested by Fischer (2006b), any operational definition of DMs

should not only include formal criteria but functional ones as well. We restrict our identification to DMs which either express a discourse relation (e.g. reformulation, contrast), mark relations between hierarchical units (e.g. topic change, turn-taking), express the speaker's modality (e.g. approximation, mitigation) or perform control-of-contact functions (e.g. backchannelling, monitoring). The full list of English discourse markers identified in the sample according to these criteria is the following:

actually, and, anyway, but, first of all, having said that, I mean, in fact, like, look, mind you, no, now, okay, or, plus, right, secondly, so, sort of, then, well, yeah, yes, you know, you see.

Once identified, each item was manually classified in one of three functional categories taken from Briz & Val.Es.Co (2014):

- textual: the marker organizes the speech flow by connecting utterances, turns, topics (e.g. *anyway, but, first of all*);
- modal: the marker expresses the speaker's attitude such as attenuation or uncertainty (e.g. *sort of*, *right*);
- interpersonal: the marker creates or maintains contact with the interlocutor (e.g. *look, you know*).

This functional analysis is part of a segmentation process which takes into account the type of host-unit which contains the discourse marker (e.g. act, turn) and the position of the marker in this unit, in addition to the semantics of the marker itself. With this model, a given marker can vary from one function to another depending on its structural configuration. For instance, when *right* is turn-initial, as a confirmation marker, it is classified as modal, whereas when it is act- or turn-final, as a contact-control marker, the occurrence is interpersonal.

A sample of pilot data in English was annotated by both authors so as to set guidelines and criteria for the different categories and to maximize the comparability of the annotations. As this functional disambiguation is based on the semantics of the discourse markers and on the discourse segmentation (type of unit and position), the annotations were quite straight-forward and no specific problems or disagreements were found.

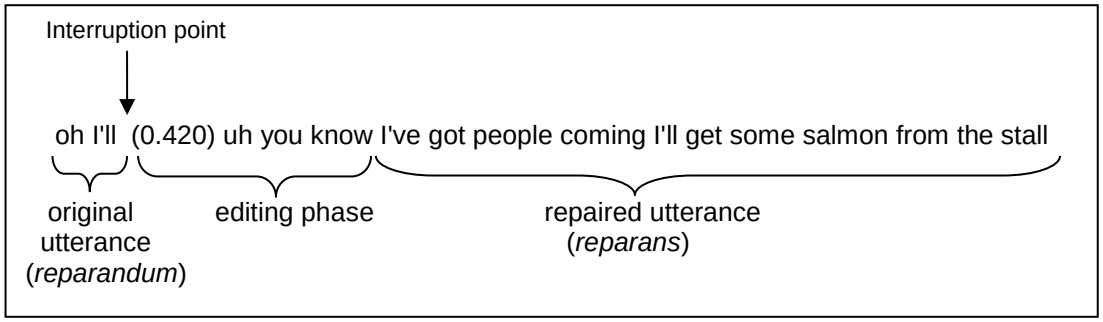
This segmentation-based view of discourse functions is compatible with the present structural approach to repair inherited from Levelt (1983), which is presented in the next section.

4.3 Analysis of repairs

The analysis covers same-turn overt self-repairs, that is, segments where an original utterance (*reparandum*) is interrupted and later repaired by a *reparans*, either immediately or after an editing phase which can contain pauses, fillers, DMs or other non-propositional elements, as shown in Figure 1. We consider cases of repair when there is an identifiable change from the original utterance, either at the formal level (e.g. re-ordering words) or at the semantic level (e.g. word replaced by a more specific term), such that the *reparandum* can be substituted by the *reparans* in order to arrive at the speaker's final intended message. This basic definition implies a discursive move of reformulation and excludes cases of false starts, that is, interruptions of the syntactic

structure with nothing in common formally nor semantically between the interrupted utterance and the following one.

Figure 1: Example of self-repair and its constituent parts



Once a repair sequence is identified, its internal boundaries are drawn by delimiting the end of the original utterance and the beginning of the repaired utterance. There can – but need not – be some non-propositional elements in-between: if there are, these are isolated in the "editing phase".

Then, a qualitative analysis classifies the type of repair, following Levelt's (1983) typology. The analysis uses the three main categories defined in Section 2.1 above: delay repairs (henceforth D-repairs), linked to macro-structure and ordering of ideas; error repairs (henceforth E-repairs), when the speaker feels the need to correct some mistake; appropriateness repairs (henceforth A-repairs), if the material to be repaired is not entirely incorrect but another choice of words would be more appropriate.

The two annotators (the authors) discussed dubious cases, where the repair category was not obvious, and thus resolved any hesitations in the coding. However, much like the functional disambiguation of discourse markers, these three repair categories strongly rely on lexical and syntactic cues in the linguistic context and are therefore straight-forward to disambiguate.

4.4 Analysis of repetitions

Unlike the structural analysis of repairs, the present approach to repetitions is componential (cf. above), inspired by disfluency annotation schemes such as Shriberg (1994) and Crible et al. (in press). Following the latter, we distinguish between i) identical repetitions, where the words are repeated quasi-immediately with no change in form or meaning (e.g. *the the the house is big*) and ii) modified repetitions, where something is changed in the direct context of the repetition. Modified repetitions include, for instance, cases where some lexical items are inserted between the repeated words (e.g. *the house the big house*) or cases where the repeated words are part of an "anchor" structure which is later modified (e.g. *the house is big the house is beautiful*).

Repetitions (identical and modified) are only identified when they are immediately adjacent to a DM and/or when the repetitions themselves are constituted by DMs, either within or outside repair sequences. An isolated repetition that is not combined with a DM or does not occur within a repair will not be identified.

Identical repetitions are expected to be more typical of covert repairs (outside repair sequences), whereas modified repetitions should be quite frequent in overt self-repairs, to build the structure of the reformulation.

All variables for DMs, repairs and repetitions are manually annotated by two expert coders under the open-source EXMARaLDA annotation software (Schmidt & Wörner 2014). This tool allows to add multiple annotation layers and map them onto a sound-aligned transcription. The audio was available during the analysis for all transcriptions.

5. Results

We will present corpus findings of the three phenomena under scrutiny, viz. discourse markers, repetitions and repairs, starting with the features of DMs in English, French and Spanish. Repetitions and repairs will then be progressively integrated in the analysis, in order to draw a full portrait of the co-occurrence tendencies of these devices. Given the very similar size of each subcorpus (11,000 words +/- 100 words), we will present absolute frequencies, as well as percentages, when relevant.

5.1 Isolated DMs

A total of 1,347 DM tokens have been identified that do not combine with a repetition or appear within a repair. The three most frequent discourse markers in each language are: in English *well* (77), *you know* (45), *and* (42); in French, *et* 'and' (107), *mais* 'but' (85), *hein* 'right?' (43); in Spanish, *pero* 'but' (63), *y* 'and' (44) and *bueno* 'well' (38). We see that they are strikingly similar, with basic additive and contrastive conjunctions as well as more pragmatic, speech-specific devices.

In terms of DM diversity, English stands out from the other two languages with its relatively low number of different DM types: only 26 against 44 in French and 42 in Spanish. This lower variety of expressions is also reflected in the distribution of functional categories: English has the highest proportion of textual markers (75%), in addition to 23% of interpersonal markers (e.g. *you know*) and very few modal markers (e.g. *sort of*) in the sample. By contrast, in Spanish, the textual category takes up a smaller proportion of the data (53%) and contains much more modal markers (23%) such as *claro* 'right', *hombre* 'man/well', *bueno* 'well', *vale* 'okay/right' or *a ver* 'well', in equal frequency (24%) with interpersonal markers. French is somewhat intermediate between English and Spanish, with 66% of textual markers, 29% of interpersonal markers and few modal markers.

In all three languages, the most frequent category is the textual function: despite the conversational nature of the data, speakers still mainly resort to discourse markers for structuring purposes when they do not combine with repetitions or repairs. Some textual DMs, however, can also be used for modal or interpersonal functions, such as French *ben* 'well' in (1) or Spanish *pero* 'but' in (2).

(1) <spk1> j'aimerais bien lui téléphoner ce soir mais <spk2> oui **ben** écoute tu téléphoneras tout à l'heure hein

<spk1> *I would like to call her tonight but* <spk2> *yes* **ben** *listen you will call her later okay*

(2) lo peor no es que pisarais los pivotes de la <a>Gran Plaza es que llega a venir un coche (RISAS) [¡y os FOLLA] **pero** vamos!

C: the worst thing isn't that you went over the bollards in the Great Square it's that if a car had come (LAUGH) it would have FUCKED YOU OVER ***pero*** *vamos!*

Other markers are also polyfunctional between the modal and interpersonal categories, such as Sp. *claro* 'yeah/right' or Fr. *quoi* in (3) and (4).

(3) ouais c'est clair c'est le pire au monde **quoi** ils en pouvaient plus

yeah for sure it's the worst **quoi** *they couldn't take it anymore*

(4) c'est ça une sorte de euh d'un magistrat service de presse **quoi** c'est ça

that's it a sort of uh press office judge **quoi** that's it

In (3), *quoi* is modal as it reinforces the speaker's expressivity with evaluative language (emphasis function), whereas in (4) it serves as an interpersonal marker to check that the interlocutor has understood the term which the speaker was looking for (monitoring or control-of-contact function). Such polyfunctional items as illustrated in (1)-(4) were not found in the English sample, which could suggest a higher degree of specificity (stronger form-meaning mapping) than in the Romance languages, as expected from the literature review.

In sum, isolated DMs are quite varied formally and functionally, especially in Spanish and French, even though the top three DMs are highly similar in the three languages.

5.2 DMs and repetitions outside repairs

In addition to the 1,347 isolated DMs, 213 tokens were found to co-occur with a repetition, which shows that DMs are more often isolated than combined with repetitions. Such combinations are much more frequent in Spanish (131 cases) than in English (39) or French (43). This quantitative difference could be related to the presence of more than two speakers in the Spanish data, creating overlaps and thus more reasons to repeat oneself. It can also be explained by the fact that in the Spanish data many of these combinations (62 in total) consist in the repetition of a DM itself, for the purposes of either planning or emphasis, as in (5), where the repeated *vale* 'okay/right' intensifies the agreement:

(5) B: pienso tenerlo terminado el jueves [por la noche]
A: [¡ah! **vale**] **vale**
B: es el objetivo sí sí
A: **vale vale** (1,11) entonces bien

B: my plan is to have it finished [by Thursday night]
A: [ah! ***vale****] **vale***
B: that's the objective yes yes
A: ***vale vale*** *(1,11) so OK*

The lack of repeated DMs in English and French suggests that this is a rather exclusive or at least a recurrent phenomenon in spontaneous Spanish. Textual and interpersonal DMs are the most frequently repeated types.

The (by far) most frequent configuration containing DMs and repetitions across all languages is the co-occurrence between a DM and an identical repetition (187 cases). The DM is usually the first element in the sequence (6), but it can also be in the middle (7) or at the end of the repetition (8).

(6) yes uh **in fact there's a there's a** tendency towards the nice

(7) and Jane yeah **the one that Philip actually (0.160) the one that Philip** got very annoyed

(8) **I know I know but** if you're doing phonetics analysis

DMs more rarely combine with modified repetitions outside repairs. Modified repetitions bring about a change in the linguistic context and are therefore more typical of repair sequences. There are however a few cases where some materials are repeated, not to be changed but to add more information on the basis of the same structure, as in (9).

(9) well you bought some and I bought some

Both types of repetitions can simultaneously combine with DMs, although this pattern is also very rare:

(10) you have a you have a a mallet and you have a a ball

In (10), there are several identical repetitions ("you have a", "a" twice), a modified repetition ("you have a mallet" - "you have a ball") and the DM "and". It remains that the vast majority of combinations of DMs and repetitions across the three languages do not bring any change or add any information apart from a momentary stalling effect, possibly for planning purposes (cf. Clark & Wasow 1998).

Turning to the functions of DMs when combined with repetitions, the proportions of the three functional categories remain roughly the same as for isolated DMs, with the notable exception that textual functions are even more prominent than modal or interpersonal functions in English: up to 87% of DMs are textual in the context of repetitions (against 67% in isolation). This tends to confirm that these combinations serve discourse-structuring and discourse-planning purposes.

5.3 DMs and repetitions within repairs

To fully understand the inter-relation between discourse markers, repetitions and repairs, we need to adopt an integrated approach to all three phenomena and look at their co-occurring behavior. This will allow us to verify the hypothesis according to which DMs and repetitions do not frequently co-occur within repairs because of the redundancy of their signaling function.

A total of 206 repair sequences have been identified in the sample: 49 in English, 75 in French and 82 in Spanish. Contents and features of the repairs themselves will be discussed and interpreted in the following sections.

5.3.1 Frequency and types of elements in repairs

Overall, DMs and/or repetitions are included in about 85% of all repairs across the three languages, less so in French. The data is presented in Table 2.

Table 2: Proportions of repairs involving DMs and repetitions

Configurations	**English**	**French**	**Spanish**	**Total %**	**Total**
Repair alone	14.29%	24.00%	14.63%	17.96%	37
Repair + DM	14.29%	10.67%	29.27%	18.93%	39
Repair + repetition	51.02%	34.67%	29.27%	36.89%	76
Repair + DM + rep.	20.41%	30.67%	26.83%	26.21%	54
Total	**49**	**75**	**82**	**100%**	**206**

In English, the most frequent configuration (51%) includes a repetition (or several) but no DM: DMs are only included in about a third of all repairs (with a repetition or not). No clearly preferred structure can be observed from Table 2 for French or Spanish. However, DMs are more frequently involved in repairs in the Romance languages than in English: about 40% in French and 60% in Spanish. Overall, 136 DM tokens were annotated as included in any position of a repair: 24 in English, 54 in French and 58 in Spanish. This distribution confirms Cuenca's (2003) study of reformulative markers in English, Spanish and Catalan, showing a closer similarity between the two Romance languages than with English. French and Spanish are also closer in terms of the diversity of DM types included in repairs, as can be see in Table 3.

Table 3: DM types included in repairs

English	French	Spanish
and, anyway, but, or, so, well, you know	allez, d'abord, donc, en fait, enfin, en tout cas, et, hein, je vais dire, mais, ou, quoi, si tu veux, tu vois	a ver, bueno, digamos, es que, hombre, mira, no, o sea, pero, pues, tío, y, ya

None of these types are specific to repairs, as they were also found in isolation, with the exceptions of English *or*, French *ou* 'or' and Spanish *digamos* 'let's say'. These three DMs can be considered as unequivocal markers of repairs. All other types are either much more pervasive (*and*, *but*, *so*) or they are semantically related to reformulation but also include other uses, such as *well* (turn-taking), Fr. *enfin* 'I mean' (specification) or Sp. *o sea* 'that is' (consequence, conclusion).

Turning to repetitions, the most frequent type is modified repetitions, with 90 cases out of the 130 repairs that contain a repetition (or several). Only 21 repairs contain identical repetitions, not including the 19 cases where both types are included. This large discrepancy between the two types perfectly mirrors the situation of combinations outside repairs, where we saw in the previous section that identical repetitions are the vast majority. This corroborates the association between, on the one hand, identical repetitions and covert repairs and, on the other, modified repetitions and overt repairs.

5.3.2 Types of repairs

We will now refine these results by looking at the distribution of DMs and repetitions in different repair types. As a reminder, repairs can target three categories of inadequacies in the original utterance: an appropriateness issue (the chosen term is too ambiguous or too imprecise), a delay (some information needs to be added first) or an error (the lexeme, phoneme or syntactic structure is wrong, according to the speaker). We can see in Table 4 that error-repairs are the most frequent category in all three languages, with 65% in English and French and 57% in Spanish. The other two categories are much less frequent, especially in English and French where they each take less than 20%.

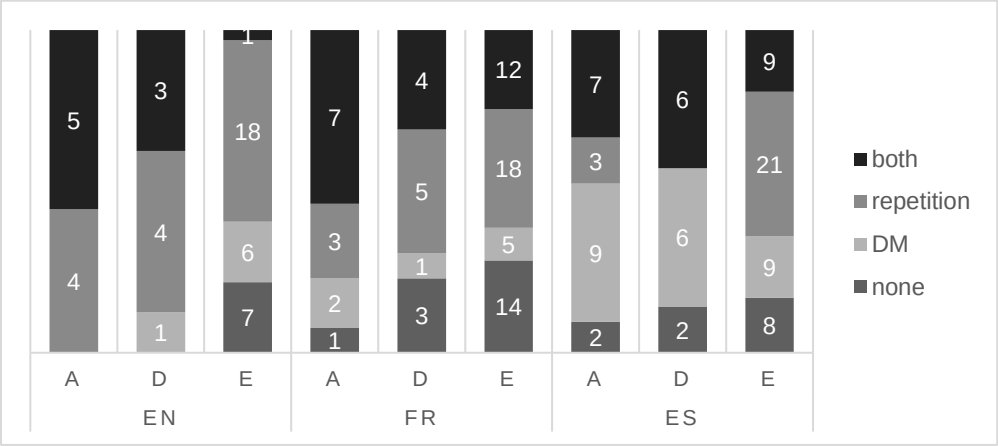
Table 4: Proportions and frequencies of repair (sub)types in the three languages

	English	**French**	**Spanish**	**Total**
Appropriateness	18.37% (9)	17.33% (13)	25.61% (21)	20.87% (43)
Delay	16.33% (8)	17.33% (13)	17.07% (14)	16.99% (35)
Error	65.31% (32)	65.33% (49)	57.32% (47)	62.14% (128)
Total	**49**	**75**	**82**	**206**

A great similarity between English and French can be observed from this table, with almost identical proportions for all three repair categories, while the Spanish data displays a smaller gap between appropriateness and error repairs. This result suggests that the Spanish speakers in the sample notice and correct their inappropriate lexical choices more often than the other speaker populations.

Figure 1 shows the amount of repairs containing a discourse marker, a repetition, both or none of them across languages and repair categories (viz. appropriateness “A”, delay “D” or error “E”).

Figure 1: Presence of discourse markers and repetitions across repair types and languages



We can observe that, for all repair types in the three languages, there is almost always either a DM or a repetition, if not both, except for error repairs (E), which show a substantial proportion of cases of “none”. This finding suggests that, in error repairs,

the relation between the *reparandum* and the *reparans* is sometimes explicit enough and does not need extra signaling by a reformulative DM or by a structuring repetition. In appropriateness (A) and delay repairs (D), on the other hand, the repair is most of the time reinforced by a DM, a repetition or both.

This graph further suggests a difference regarding the presence of DMs between English, French and Spanish: in the former two, repetitions are much more frequent in repairs than DMs, which mainly occur with a repetition ("both") and not so much without; in Spanish, on the contrary, repetitions rarely occur without DMs in A- and D-repairs. In (11), we see that the appropriateness repair is built upon a modified repetition, whereas in (12), the same repair type uses the DMs *bueno* 'well' and *en realidad* 'in truth' to signal the reformulation.

(11) it was too much for her the house was too much for her

(12) la lengua de los gitanos (0.240) mm (0.190) bueno en realidad el tratamiento lexicográfico que se le da a los gitanismos

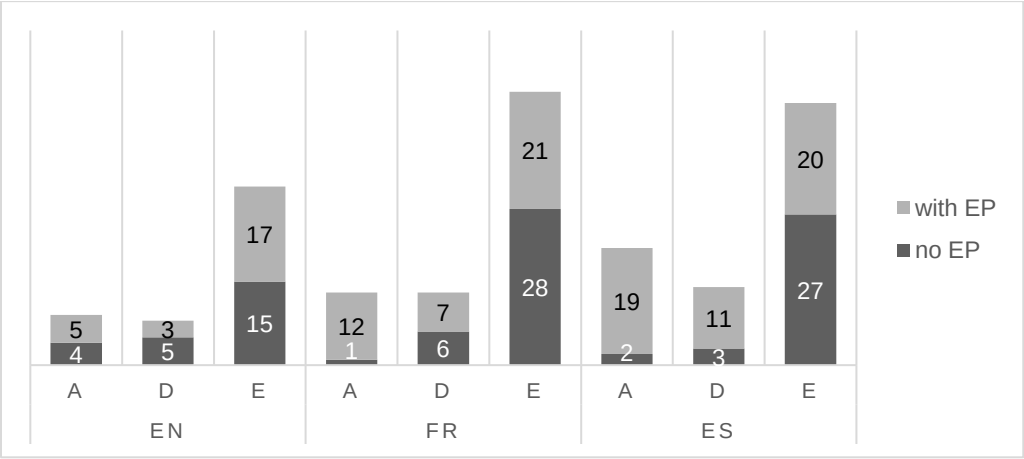
the language of the gypsies (0.240) mm (0.190) well in fact the lexicographic treatment that is given to "gypsyisms"

The number of cases where both DMs and repetitions are included in repairs leads us to reject our hypothesis of their mutual exclusion. However, the data in Figure 1 includes DMs regardless of their position in the repair (either in the *reparandum*, in the editing phase or in the repairing segment). In the next section, we will therefore focus on the DMs in the editing phase of repairs.

5.3.3 The editing phase of repairs

As a reminder, the editing phase is the medial part between the *reparandum* and the repairing segment. It typically contains pauses, filled pauses (e.g. *uh*) and discourse markers. Figure 2 reports the data for this structural feature in the different repair categories.

Figure 2: Presence of an editing phase across repair types and languages



It appears that some repair types typically occur with an editing phase: this concerns appropriateness repairs in French and Spanish (almost 100%) and delay repairs in Spanish. The results are much more balanced for all repair categories in English,

where repairs are often restarted with no interruption (no pause or discourse marker), and for error repairs in all three languages. This corroborates our previous observation regarding the smaller need to signal error repairs.

In the editing phase of repairs, DMs display a particularly coherent pattern as being almost exclusively reformulative markers: *or*, *well*, French *ou* 'or', *enfin* 'well', *en tout cas* 'in any case', Spanish *o sea* 'that is', *bueno* 'well'. The bulk of these DMs in the editing phase express a textual function (47 out of 57), which is consistent with their reformulative use. The few remaining others are more clearly interpersonal (*you know*, Sp. *mira* 'look') or modal (Sp. *a ver*, *hombre*), as in Examples (13)-(14).

(13) you think oh I'll (0.420) uh **you know** I've got people coming I'll get some salmon

(14) te dicen el día de antes (0.195) **hombre** a nosotros nos lo dijeron el día de antes

they tell you the day before (0.195) hombre *'well'* to us they said the day before

These two markers do not simply signal a repair between the two segments they connect (basic reformulation, textual function), they also add another dimension: a call to the addressee in (13), to build complicity or common ground, and an expressive, intensifying value in (14), to convey the speaker's emotion (here, resentment). It remains that the vast majority of discourse markers in the editing phase perform textual functions. Compared with uses of DMs outside repairs, it appears that there is a smaller diversity of DM types and functions in the editing phase of repairs, whereas more pervasive markers such as conjunctions can occur both at the periphery of repairs and outside them.

When there is a DM in the editing phase, repetitions are also included in 28 cases (49%), 15 of which are modified repetitions, nine identical repetitions and four repairs with both types. Given the small number of repairs containing an editing DM in each language, it is difficult to refine these results by repair type. What we can conclude is that, against our hypothesis, DMs and repetitions do not systematically exclude each other, even when the DM is located in the editing phase. The preference for DMs or repetitions seems to vary cross-linguistically and with the repair type, as was discussed in the previous section (e.g. DMs for appropriateness repairs in Spanish, repetitions for error repairs in English).

6. Conclusion

This corpus-based study of discourse markers and repetitions within and outside repairs in native English, French and Spanish has revealed a number of differences and similarities between these three languages. As isolated occurrences, the most frequent markers are very similar across languages (*and*, *you know* and *well* in English) and they mostly perform textual functions, although they are more formally and functionally varied in the Romance languages. Combinations of DMs and repetitions outside repairs are most frequent in Spanish, where repetitions of DMs themselves are common, as opposed to English and French. Combinations of DMs and repetitions mainly involve identical repetitions (no change or addition of information). DMs are more frequent in isolation than combined with repetitions, and their functional distribution is stable with and without repetitions.

The analysis then investigated the types, functions and position of DMs and repetitions within different types of overt repairs (error, appropriateness, delay). Overall, error repairs are most frequent. While most repairs (85%) include at least a DM and/or a repetition in all languages, DMs are more frequent in Spanish and more diverse in both Romance languages. We found one DM in each language which only occurs in repair sequences, viz. English *or*, French *ou* 'or' and Spanish *digamos* 'let's say', which we can consider as typical repair markers. Repetitions in repairs are mainly modified (as opposed to identical repetitions outside repairs), and often combine with DMs, although this combination varies with the repair type and the language. More specifically, English and French seem to favor repetitions in repairs, whereas DMs are more frequent in Spanish. Finally, focusing on the editing phase of repairs, we found that DMs are mostly restricted to reformulative markers with a textual function, and that they co-occur with (mainly modified) repetitions in half of the repairs. DMs in the core part of repairs are therefore much more specific and restricted than in other parts and outside repairs.

While the hypothesis of mutual exclusion between DMs and repetitions was rejected, we were able to confirm that DMs and repetitions are more frequent outside than within repairs, which relates to their association with covert planning processes. Regarding cross-linguistic differences, French and Spanish resemble each other in several aspects, although the similarity between English and French was also observed in our data.

This study has investigated the complex interplay between discourse markers, repetitions and repairs in three languages, thus filling a gap in the study of spoken fluency and in crosslinguistic fluency research. Fluency and disfluency are notions which are relevant to native speakers as well as to learners, who need to acquire a "fluencicon", that is, a repertoire of linguistic devices which will enhance the automaticity and naturalness of their production processes, usually from naturalistic exposure alone. Our results point to particular markers which could be taught more explicitly to learners so that they resemble native speech, even in contexts of hesitation. The distinction between pervasive and repair-specific reformulation markers (e.g. *o sea* 'that is' vs. *digamos* 'let's say' in Spanish) is particularly worthwhile to communicate to language teachers and students. Such applications could be further developed through a more extensive corpus study on larger samples, on more languages and on non-native data as well, in order to compare (and potentially adapt) the repairing strategies of learners and of native speakers of English, French and Spanish.

References

Aijmer, Karin. 2016. Pragmatic markers as constructions. The case of *anyway*. In G. Kaltenböck, E. Keizer & A. Lohmann (eds), *Outside the Clause: Form and Function of Extra-Clausal Constituents*, Amsterdam, John Benjamins: 29-58.

Aijmer, Karin. 2011. "*Well I'm not Sure I think*… The Use of *Well* by Non-native Speakers." *International Journal of Corpus Linguistics* 16: 231-254.

Aijmer, Karin & Simon-Vandenbergen, Anne-Marie. 2011. Pragmatic markers. In J. Zienkowski, J.-O. Östman & J. Verschueren (eds), *Discursive pragmatics*. Amsterdam, John Benjamins: 223-247.

Auer, Peter & Pfänder, Stefan. 2007. Multiple retractions in spoken French and spoken German. A contrastive study in oral performance styles. *Cahiers de Praxématique* 48: 57-84.

Bardovi-Harlig, Kathleen. 2005. Contextualizing interlanguage pragmatics. In A. Tyler (Ed.), *Language in use: Cognitive and discourse perspectives on language and language learning*. Washington, DC, Georgetown UP: 65-84.

Beeching, Kate. 2012. Semantic change: Evidence from false friends. In Lauwers et al. (eds), *Pragmatic Markers and Pragmaticalization: Lessons from false friends*. Amsterdam, John Benjamins: 11-36.

Briz, Antonio & Pons, Salvador. 2010. Unidades, marcadores discursivos y posicion. In O. Loureda & E. Acin (eds), *Los Estudios sobre Marcadores del Discurso*. Madrid: Acro Libros: 327–358.

Briz, Antonio & Val.Es.Co. Group. 2014. Las unidades del discurso oral. La propuesta Val.Es.Co. de segmentación de la conversación (coloquial). *Estudios de Lingüística del Español* 35(1): 11-71.

Bunt, Harry. 2011. Multifunctionality in dialogue. *Computer Speech and Language* 25: 222-245.

Buysse, Lieven. 2012. *So* as a multifunctional discourse marker in native and learner speech. *Journal of Pragmatics* 44(13): 1764-1782.

Buysse, Lieven. 2015. 'Well it's not very ideal…' The pragmatic marker 'well' in learner English. *Intercultural Pragmatics* 12/1: 59-89.

Cabedo, Adrián & Pons, Salvador. 2013. *Corpus Val.Es.Co. 2.0.* http://www.valesco.es/?q=es/corpus.

Clark, Herbert H. & Wasow, Thomas. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37: 201-242.

Crible, Ludivine. 2018. *Discourse Markers and (Dis)fluency. Forms and Functions across Languages and Registers*. Amsterdam: John Benjamins.

Crible, Ludivine, Dumot, Amandine, Grosman, Iulia & Notarrigo, Ingrid. In press. (Dis)fluency across spoken and signed languages: Application of an interoperable annotation scheme. In L. Degand, G. Gilquin, L. Meurant & A.-C. Simon (eds), *Fluency and Disfluency across Languages and Language Varieties*, Louvain-la-Neuve: Presses Universitaires de Louvain.

Crystal, David. 1988. Another look at *well*, *you know*... *English Today* 4(1): 47-49.

Cuenca, Maria Josep. 2013. The fuzzy boundaries boundaries between discourse marking and modal marking. In L. Degand, B. Cornillie & P. Pietrandrea (eds), *Discourse Markers and Modal Particles. Categorization and Description*. Amsterdam, John Benjamins 191-216.

Cuenca, Maria Josep. 2008. Pragmatic markers in contrast: the case of *well*. *Journal of Pragmatics* 40: 1373–1391.

Cuenca, Maria Josep. 2003. Two ways to reformulate : A contrastive analysis of reformulation markers. *Journal of Pragmatics* 35: 1069-1093.

Cuenca, Maria Josep & Bach, Carme. 2007. Contrasting the form and use of reformulation markers. *Discourse Studies* 9(2): 149-175.

De Cock, Sylvie, Granger, Sylviane, Leech, Gregory & McEnery, Tony. 1998. An automated approach to the phrasicon of EFL learners. In S. Granger (Ed.), *Learner English on Computer*. London: Addison Wesley Longman, 67-79.

De Gaulmyn, Marie-Madeline. 1987. Actes de reformulation et processus de reformulation. In P. Bange (Ed.) *L'Analyse des Interactions Verbales. La Dame de Caluire: Une Consultation*, Bern, Peter Lang: 83-98.

Degand, Liesbeth. 2016. "Spoken discourse segmentation and the paradox of discourse markers." L&C talk, MPI Nijmegen, March 23rd 2016.

Degand, Liesbeth. 2014. "So very fast, very fast then" Discourse markers at left and right periphery in spoken French. In K. Beeching & U. Detges (eds), *The Role of the Left and Right Periphery in Semantic Change: Crosslinguistic Investigations of Language and Language Change*. Leiden, Brill: 151-178.

Denke, Anita. 2009. *Nativelike Performance. Pragmatic Markers, Repair and Repetition in Native and Non-native English Speech*. Saarbrücken: Verlag Dr. Müller.

Dister, Anne, Francard, Michel, Hambye, Philippe, & Simon, Anne-Catherine. 2009. Du corpus à la banque de données. Du son, des textes et des métadonnées. L'évolution de la banque de données textuelles orales VALIBEL (1989-2009). *Cahiers de Linguistique* 33(2): 113-129.

Ejzenberg, Roseli. 2000. The juggling act of oral fluency : A psycho-sociolinguistic metaphor. In H. Riggenbach (Ed.), *Perspectives on Fluency*. Ann Arbor, The University of Michigan Press: 288-313.

Fischer, Kerstin. 2006a. *Approaches to Discourse Particles*. Amsterdam: Elsevier.

Fischer, Kerstin. 2006b. Towards an understanding of the spectrum of approaches to discourse particles: Introduction to the volume. In K. Fischer (Ed.), *Approaches to Discourse Particles*, Amsterdam, Elsevier: 1-20.

Fox Tree, Jean E. 1995. The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language* 34: 709-738.

Fung, Lancy. 2007. The communicative role of self-repetition in a specialised corpus of business discourse. *Language Awareness* 16(3): 224-238.

Fung, Lancy & Carter, Ronald. 2007. Discourse markers and spoken English: Native and learner use in pedagogic settings. *Applied Linguistics* 28(3): 410-439.

Gilquin, Gaëtanelle. 2016. Discourse markers in L2 English. From classroom to naturalistic input. In Timofeeva et al. (eds), *New Approaches in English Linguistics*. Amsterdam: John Benjamins, 213-249.

Gilquin, Gaëtanelle. 2008. Combining contrastive and interlanguage analysis to apprehend transfer: detection, explanation, evaluation. In Gilquin et al. (eds), *Linking up Contrastive and Learner Corpus Research*. Amsterdam, Rodopi: 3-33.

González, Montserrat. 2005. Pragmatic markers and discourse coherence relations in English and Catalan oral narrative. *Discourse Studies* 77(1): 53-86.

Götz, Sandra. 2013. *Fluency in Native and Nonnative English Speech*. Amsterdam: John Benjamins.

Hansen, Maj-Britt M. 2006. A dynamic polysemy approach to the lexical semantics of discourse markers (with an exemplary analysis of French *toujours*). In K. Fischer (Ed.), *Approaches to Discourse Particles*. Amsterdam, Elsevier: 21-41.

Hasselgren, A. 2002. Learner corpora and language testing: Smallwords as markers of learner fluency. In S. Granger, J. Hung & S. Petch-Tyson (eds), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, Amsterdam, John Benjamins: 143-173.

House, Juliane. 2013. Developing pragmatic competence in English as a lingua franca: Using discourse markers to express (inter)subjectivity and connectivity. *Journal of Pragmatics* 59: 57-67.

House, Juliane. 1996. Developing pragmatic fluency in English as a Foreign Language: Routines and metapragmatic awareness. *Studies in Second Language Acquisition* 18(2): 225-252.

Levelt, Willem J.M. 1989. *Speaking: From Intention to Articulation*. Cambridge: MIT Press.

Levelt, Willem J.M. 1983. Monitoring and self-repair in speech. *Cognition* 14: 41-104.

Li, Min & Xiao, Yan. 2012. A comparative study on the use of the discourse marker “well” by Chinese learners of English and native English speakers. *International Journal of English Linguistics* 2(5).

MacGregor, L., Corley, Martin & Donaldson, D. 2009. Not all disfluencies are are equal: The effects of disfluent repetitions on language comprehension. *Brain & Language* 111: 36-45.

Maschler, Yael. 1994. *Metalanguage in Interaction. Hebrew Discourse Markers*. Amsterdam: John Benjamins.

Maschler, Yael & Schiffrin, Deborah. 2015. Discourse markers : Language, meaning, and context. In D. Tannen, H. Hamilton & D. Schiffrin (eds), *The Handbook of Discourse Analysis*. Hoboken, NJ, John Wiley & Sons: 189-221.

Müller, Simone. 2005. *Discourse Markers in Native and Non-Native English*. Amsterdam: John Benjamins.

Nelson, Gerard, Wallis, Sean & Aarts, Bas. 2002. *Exploring Natural Language: Working with the British Component of the International Corpus of English*. Amsterdam: John Benjamins.

O’Donnell, W. R. & Todd, Loreto. 1980. *Variety in Contemporary English*. London: Allen and Unwin.

Odlin, T. 1989. *Language Transfer. Cross-linguistic Influence in Language Learning*. Cambridge: CUP.

Pascual, Elena. 2018. Análisis prosódico de las estructuras truncadas en la conversación coloquial española: funciones de formulación y atenuación. In A. García Ramón & M. A. Soler Bonafont (eds), *ELUA: Estudios de atenuación en el discurso, Anexo IV*: 57-84.

Pons, Salvador. 2018. The combination of discourse markers in spontaneous conversations: Keys to untying a Gordian knot. *Revue Romane* 53(1): 121-158.

Rabab'ah, G. & Abuseileek, A. 2012. The pragmatic functions of repetition in TV discourse. *Research in Language* 10(4): 445-460.

Ragan, Sandra L. 1983. Alignment and Conversational Coherence. In R. Craig & K. Tracy (eds), *Conversational Coherence: Form, Structure and Strategy*. Bervely Hills, Sage Publications: 157–171.

Redeker, Gisle. 1990. Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14(3): 367-381.

Romero Trillo, Jesus. 2002. The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics* 34(6): 769-784.

Rossari, Corinne. 1994. *Les Opérations de Reformulation*. Bern: Peter Lang.

Schegloff, Emanuel, Jefferson, Gail & Sacks, Harvey. 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53: 361–382.

Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Schmidt, Thomas & Wörner, Kai. 2014. EXMARaLDA. In J. Durand, U. Gut & G. Kristoffersen (eds), *The Oxford Handbook of Corpus Phonology*. Oxford, Oxford University Press: 402-419.

Schourup, Lawrence. 2001. Rethinking *well*. *Journal of Pragmatics* 33, 1025-1060.

Shriberg, Elizabeth E. 1994. *Preliminaries to a theory of speech disfluencies*. Doctoral dissertation, University of California.

Zhao, Hongwei. 2013. A study on the pragmatic fossilization of discourse markers among Chinese English learners. *Journal of Language Teaching and Research* 4(4), 707-714.