



HAL
open science

Investigating Disfluencies Contribution to Discourse-Prosody Mismatches in French Conversations

Laurent Prevot, Roxane Bertrand, Stéphane Rauzy

► **To cite this version:**

Laurent Prevot, Roxane Bertrand, Stéphane Rauzy. Investigating Disfluencies Contribution to Discourse-Prosody Mismatches in French Conversations. The 10th Workshop on Disfluency in Spontaneous Speech, Aug 2021, Paris, France. hal-03415890

HAL Id: hal-03415890

<https://hal.science/hal-03415890v1>

Submitted on 5 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Investigating disfluencies contribution to discourse-prosody mismatches in French conversations

Laurent Prévot^{1,2} Roxane Bertrand¹ and Stéphane Rauzy¹

¹Aix Marseille Université & CNRS, Aix-en-Provence, France

²Institut Universitaire de France, Paris, France

Abstract

In conversation, discourse and prosodic units association can be articulated through an interesting range of configurations. The situation in which these units are mismatching is the least studied and understood of these configurations. We make the hypothesis in this paper that disfluencies are a major cause for such mismatches. Our quantitative analysis based on a 8 hour corpus of French conversations manually annotated with disfluencies, discourse units (DU) and prosodic units (PU), confirms that disfluencies do play a major role in PU-DU mismatch but also that other sources should be considered. In the analysis, we also provide some insight about the different types of disfluencies and their frequency in the different DU-PU configurations.

Introduction

Discourse, Prosody and Syntax interplay is a crucial aspect of linguistic analysis of conversation. Previous literature had described many aspects of the association between these three levels in terms of boundary alignment (Degand & Simon, 2009; Prévot et al., 2015; Lacheret-Dujour & Kahane, 2020) and explored them in relation to *discourse genre* and *speaking style* (Degand & Simon, 2009b). The configuration least understood remains the case in which major prosodic unit and discourse unit boundaries do not match. In this paper, we hypothesize that a large number of those mismatches are related to disfluency (Shriberg, 1994). We analyse an 8 hours corpus of French conversations (Bertrand et al., 2008) that had been manually annotated with prosodic units, discourse units and disfluencies. After introducing previous work, we present our annotated data and a set of quantitative analyses aiming at better understanding the impact of disfluencies on mismatches between these units.

Related Work

While much work has been done on the link between prosody and syntax, and more particularly intonation and syntax, much remains to be done on the link between prosody and discourse. Some

studies have shown that syntax and prosody play a role in the construction and identification of TCU (Turn-Constructional Units) (Ford & Thompson, 1996; Selting, 1996). In French, some studies attempt to model such a unit at the interface of syntax, prosody and discourse (see Degand & Simon, 2009; Lefeuvre & Moline, 2011 for a review of different approaches; Lacheret-Dujour & Kahane, 2020).

Following Lacheret-Dujour and Kahane (2020) or Degand and Simon (2009) we consider the different levels as autonomous. The basic discursive unit (BDU) in Degand and Simon refers to the “segments that speakers use to build a representation (interpretation) of the discourse. BDUs have a cognitive function since they correspond to steps of production and discourse processing. BDUs require syntax and prosody and their different matching give rise to several types of BDU corresponding to different discursive strategies.

The syntactic and discourse units of Lacheret-Dujour & Kahane (2020) is based on macro syntactic approach (Deulofeu, 2016) taking into account the illocutionary force (Austin, 1962) of the unit.

However, some difficulties remain in segmenting these units due to the specific phenomena frequents natural conversations. Among them, we consider that disfluencies represent a source of confusion for analyzing these levels.

Disfluencies are very frequent in spontaneous speech (about 1 every 15 words in the CID, Pallaud et al., 2019) and can occur at phonetic or morphosyntactic level anywhere in the utterance. They consist in an interruption of the flow that can be repaired or abandoned (Shriberg, 1994).

Concerning more precisely PU-DU mismatches, called *mixed-BDU* in Degand and Simon (2009b), they are not considered crucial in their analysis but they state that this ‘unexpected’ category deserves more attention, at least to understand why it occurs at significant rates. Lacheret-Dujour and Kahane (2020) called them *asynchronous* (12% of their prosody-syntax units) and relate it to difficulties for the speaker to produce and plan the utterance, which indeed suggest to look with more attention at their relationship with disfluencies.

Data

This work is performed on the whole Corpus of Interactional Data (8 conversations of 1 hour each). In this corpus participants have a chat about “unusual situations” or “conflicts at work”. See Bertrand et al. (2008) and Blache et al. (2017) for details on the corpus. The annotations used in this study are coming from three independent annotation campaigns. Overall discourse and prosodic segmentation have been performed through independent annotation campaigns realized by naive annotators trained and equipped with guidelines. Disfluency annotations have also been annotated in this way and an expert (one of the authors of the present paper) manually corrected and enriched the whole dataset.

Compared to earlier work, the amount of units annotated is much larger since the study deals with 17,102 discourse units and 30,970 prosodic units.

Disfluencies annotation

Disfluency phenomena were manually annotated following the guidelines presented in Pallaud et al. (2019). The disfluencies are defined as interruptions of the verbal fluency of the utterance at the morphosyntactic level. Some of these interruptions are characterized by utterances that are simply given up (referenced hereafter as *DISI*), some others correspond to a suspension of the verbal fluency but which continues without any impact on the syntactic structure (*DISS*), and a last kind implies the repair of the morphosyntactic sequence with the presence of a truncated word (*DIST*) and / or of a *break* (*DISB*), for which the annotation scheme of Shriberg (1995) is applied. This scheme proposes a three terms structure composed of the *Reparandum* (the term to be repaired), the *Interregnum* (*Break point*, which can be empty) and the *Reparans* (the repairing term). In case of multiple repairs, the disfluency annotation follows a tree structure which traces the paradigmatic pile. The annotation task does not present any major difficulty except the ambiguity in deciding whether an utterance is marked as abandoned or marked as the repaired term of the next utterance. The categories introduced here are illustrated in examples (1) and (2) below.

Prosodic Units

Prosodic units (*PU*s) are based on the two main consensual units in French (Di Cristo, 1998; Jun & Fougeron, 2000; etc). The Accentual Phrase (AP) is the lowest tonal unit which is the domain of primary and secondary stress. The right boundary of AP is demarcated by a final rise (LH) and the lengthening

of the final syllable. The Intonation Phrase (IP) is higher than AP. It is marked by a major f0 movement on the last or two last syllables of the IP, a large final lengthening and often followed by a pause (Di Cristo, 1998; Fougeron & Jun, 1998). We will only consider the latter here.

The guidelines were simplified to be used by naive annotators (2 annotators for each file). The annotation was conducted manually and the annotators did not have strict instructions regarding silent pauses or hesitations. Thus, as long as disfluency items did not interfere with the prosodic phrasing, the annotators were free to annotate them either independently of the rest of the utterance or by integrating them. The prosodic units then obtained reflect how annotators have treated disfluencies. This first step of non-expert annotation was partly aimed at focusing on true sources of difficulty and then enabled us to better disentangle between the problematic items (Portes & Bertrand, 2011). Also, we hypothesize that the presence of disfluencies could have an impact on the mismatch between discourse and prosodic units.

Manual prosodic segmentation with our guidelines has proven to be relatively reliable with κ -scores (Cohen, 1960) ranging between 0.5 and 0.65 for naive coders and 0.75–0.85 for expert coders.

Discourse Units

Our discourse unit segmentation was inspired by Muller et al. (2012) and corresponds to Elementary Discourse Units used in Afantenos et al. (2012) but adapted to our interactional spoken data and simplified to be used by naive annotators. The guidelines combined semantic (eventualities identification), discourse (discourse markers) and pragmatic (speech acts) instructions. Such a mixture of levels has been made necessary by the nature of the data featuring both rather monologic narrative sequences and highly interactional ones. The annotation was performed on the transcript alone without access to audio files (but including pause and timing information). Manual discourse segmentation with our guidelines has proven to be reliable with κ -scores ranging between 0.8 and 0.85. In this approach *DUs* are semantic counterparts of independent syntactic clauses, at discourse level. They are also closely related to the macro-syntactic *Illocutionary Units* (Lacheret-Dujour & Kahane, 2020) as well as to the Discourse Units of Degand & Simon (2009a).

In the annotation we distinguished between *Discourse Units* (*DU*) and *Abandoned Discourse Units* (*ADU*) (illustrated in (1) below) that correspond to false starts that cannot be easily related

to the material coming after (as illustrated below in example (1)). As a consequence, *ADU* are disfluencies in which there is no *reparans*; and disfluencies should not interfere with *DUs*.

Illustration

Example (1) below illustrates the *ADU* vs. *DU*, as well as interrupted units (*DISI*) at disfluency level. Example (2) illustrates the different disfluency categories : suspensive (*DISS*), with break (*DISB*), Truncated words (*DIST*) as well as the *reparandum* (*REP*) along with *PU* and *DU* structures. Finally, (3) shows a case of *PU* crossing *DU* boundaries.

- (1) <[que j'avais envie (d-) *DIST* enfin bref] *PU* > *ADU*
(#) *DISI* <[et (#) *DISS* on l'a accueillie (b-) *DIST* (a-) *DIST* on lui a rien demandé] *PU* > *DU*
- (2) <[(ou des) *REP* (euh non) *DISB*] *PU* [(pas des) *REP* (f-) *DISB* pas des frustrations] *PU* > *DU* <[(des (#) *DISS* espèces de) *REP* (euh) *DISB*] *PU* (# mhm #) *DISB* [(ouais) *DISB*] *PU* [(des) *REP* des vues] *PU* [différentes] *PU* [sur le boulot] *PU* [quoi] *PU* > *DU*
- (3) < ... [tu as un décalage] *PU* [quand même par rapport à l'âge] *PU* [c'est normal] > *DU* < *surtout*] *PU* [à cet âge -là] *PU* ... > *DU*

Disfluencies and PU-DU congruence

We approach the relationship between disfluencies and *PU-DU* congruence by scrutinizing what happens at *PU* and *DU* boundaries in disfluent vs. fluent sequences. More precisely we start by comparing *PU-DU* matching within *ADUs* and within other *DUs*. Our prosodic units being overall much smaller than our discourse units, we then explore disfluencies when *DU-PU* mismatches, excluding the *ADU* case before looking in detail the different disfluency categories in this context. Finally explore other potential sources of mismatches.

Abandoned Discourse Units

By definition, *ADUs* are disfluent speech. Figure 1 illustrates the difference between *DU* and *ADU*, the latter hosting a much larger proportion of some mismatches between *PU* and *DU*.

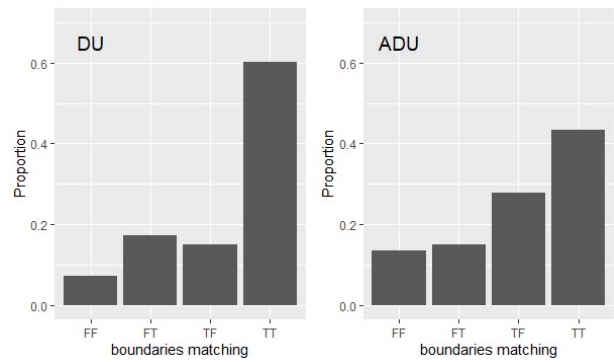


Figure 1. For the two groups *DU* and *ADU*, the proportions of mismatches encoded as *TT* (no mismatch), *TF* (mismatch on the right boundary), *FT* (mismatch on the left boundary and *FF* (mismatch on both boundaries).

PU crossing DU boundaries

We then consider with Figure 2 the *DU* case, excluding *ADU* by looking at whether a given *PU* includes or not a disfluency when it is either internal or matching a *DU* or crossing a *DU* boundary. Given our fine-grained *PU*s, the majority of *DU-PU* relationships are either 1-to-1 mapping or one *DU* including several *PU* (while matching *PU* left and right boundaries). However, 10.8% of our *PU*s are crossing *DUs* boundaries (as in example (3) above). Figure 2 illustrates that disfluencies are more frequent in those mismatch situations than in matching boundaries cases. This explains a major source of mismatches between *PU*s and *DUs*, putting aside *ADU* cases.

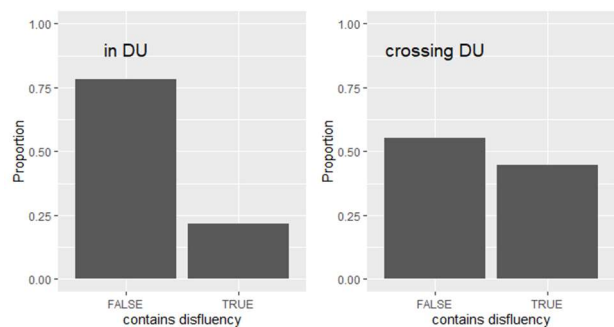


Figure 2. Comparison of *PU* hosting disfluency or not depending on its relation with *DUs*

Disfluency type

Among the *PU*s in *DUs*, we compare *PU*s ending *DUs* vs. non-ending *DUs* according to the type of disfluency. The Figure 3 illustrates that *breaks* tend to terminate a *PU* (but not a *DU* which is not surprising given *DU* definition), and a new *PU* starts with the *reparans*. Suspensive disfluency, that does not alter the syntactic flow but is likely to impact the prosodic flow with empty or filled pauses, is also a phenomena that tends to close *PU*.

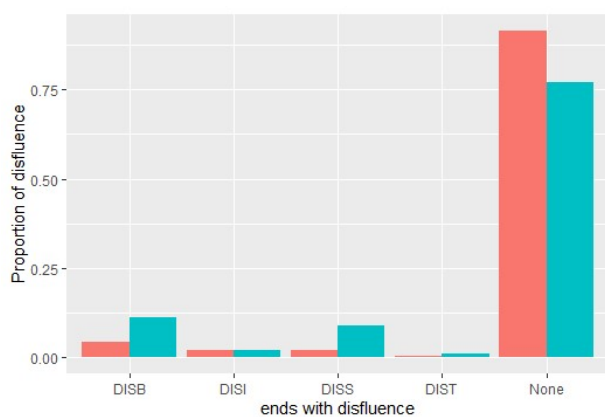


Figure 3. PUs ending DUs (red) vs. non-ending DUs (blue) according to the type of disfluency (with break (DISB), abandoned (DISI), without repair (DISS), truncation (DIST) or not ending with a disfluency (None)).

Other sources of DU-PU mismatch

In order to figure out better what happens in the mismatch zone, we extracted the tokens distribution of such zones and normalized these raw frequencies based on the distribution of DU-final and DU-initial tokens (which seems to be the best candidate for such a normalisation). The tokens over-represented in mismatch areas are Reported Speech (RS) verb introducer (*dire – say*), spoken particles (*quoi / hein / bon / tu_vois / enfin / tu_sais*); filled pauses (*euuh*), and to a less extent, first person pronouns (*je / moi*).

The presence of filled pauses in this short list simply confirms the impact of disfluency on DU-PU alignment. First person pronouns can also be a confirmation in that direction.

When Reported Speech (RS) verb introduces are the lexical items the most associated with PU-DU mismatches. (Lacheret-Dujour & Kahane, 2020) also mentioned RS as a source of asynchronous units. The main cause is that changing perspective through reporting speech clearly initiates a new discourse unit starting right after the verb introducing the RS, but sometimes the initial element of RS is prosodically grouped with the introducer.

Spoken particles create two challenges. At prosodic level, even if they are extremely short, their phrasing can vary a lot from one example to another leading to very different PU segmentations. At discourse level, some of them can be both DU-initial or DU-final (*enfin / bon / tu_vois*). It makes it difficult to decide whether they should be included in the DUs they follow or in the one after.

On the side of the spectrum some lexical items are associated PU-MU matches. This is the case of clearly initial discourse markers such as *et (and) / parce que (because) / donc (so) / mais (but) / alors (then) / ben (well)*. Those items with their clear signal of initiating a new unit could be used as some kind of synchronisation place for the different levels. The

second part of the french negation *pas* falls also in this category, but in this case in final position. There are also some other items for which we do not have a clear explanation : *là / ça / y / ils*

Conclusion

This study allowed us to refine our understanding of the impact of disfluencies on discourse-prosody interfaces. Results largely confirm what is known on this matter, namely that disfluency strongly impacts prosodic units but less discourse ones once false starts are put aside. Disfluencies explain a sizable proportion of such mismatches that are not easy to analyse from discourse-prosodic interface viewpoint. Some other sources of mismatches (such as direct reported speech) could be further investigated in order to cover the whole range of phenomena generating those DU-PU mismatches. In this paper, we pushed the analysis both in terms of scale (8 hours of conversational speech) as well as in terms of granularity specifically with regards to the different types of disfluencies involved.

References

- Afantenos, S., N. Asher, F. Benamara, M. Bras, C. Fabre, M. Ho-dac, A. Le Draoulec, P. Muller, M.-P. Péry-Woodley, L. Prévoit, J. Rebeyrolles, L. Tanguy, M. Vergez-Couret & L. Vieu. 2012. An empirical resource for discovering cognitive principles of discourse organisation: the ANNODIS corpus. In: N. Calzolari, K. Choukri, T. Declerck, M. U. Doğan, B. Maegaard, J. Mariani, A. Moreno, J. Odijk & S. Piperidis (eds.), *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, May 21–27, 2012, Istanbul, Turkey, 2727–2734.
- Austin, J. L. 1962. *How to do things with words*. Oxford: Clarendon Press.
- Bertrand, R., P. Blache, R. Espesser, G. Ferré, C. Meunier, B. Priego-Valverde & S. Rauzy. 2008. Le CID-Corpus of Interactional Data-Annotation et exploitation multimodale de parole conversationnelle. *Traitement automatique des langues*, 49(3): 105–134.
- Blache, P., R. Bertrand, G. Ferré, B. Pallaud, L. Prévoit & S. Rauzy. 2017. The corpus of interactional data: A large multimodal annotated resource. In: N. Ide & J. Pustejovsky (eds.), *Handbook of linguistic annotation*, Dordrecht: Springer, 1323–1356. https://doi.org/10.1007/978-94-024-0881-2_51
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1): 37–46.
- Degand, L. & A.-C. Simon. 2009a. On identifying basic discourse units in speech: theoretical and empirical issues. *Discours*, 4. <https://doi.org/10.4000/discours.5852>
- Degand, L. & A.-C. Simon. 2009b. Mapping prosody and syntax as discourse strategies: How Basic Discourse

- Units vary across genres. In: D. Barth-Weingarten, N. Dehé & A. Wichmann (eds.), *Where prosody meets pragmatics*, Leiden, Netherlands: Brill, 79–105. https://doi.org/10.1163/9789004253223_005
- Deulofeu, J. 2016. La macrosyntaxe comme moyen de tracer la limite entre organisation grammaticale et organisation du discours [Macrosyntax as a means of drawing the line between grammatical organization and the organization of discourse]. *Modèles linguistiques*, 38(74), 135–166. <https://doi.org/10.4000/ml.2040>
- Di Cristo, A. 1998. Intonation in French. In: D. Hirst & A. Di Cristo (eds), *Intonation systems: A survey of Twenty Languages*. Cambridge, UK: Cambridge University Press, 195–218.
- Ford, C. E. & S. A. Thompson. 1996. Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns. In: E. Ochs, E. A. Schegloff & S. A. Thompson (eds.), *Interaction and grammar*, Cambridge, UK: Cambridge University Press, 134–184. <https://doi.org/10.1017/CBO9780511620874.003>
- Fougeron, C. & S-A. Jun. 1998. Rate effects on French intonation: Prosodic organization and phonetic realization. *Journal of Phonetics* 26(1), 45-69. <https://doi.org/10.1006/jpho.1997.0062>
- Jun, S. & C. Fougeron. 2000. A Phonological model of French intonation. In: A. Botinis (ed.) *Intonation: Analysis, Modeling and Technology*, Dordrecht: Kluwer Academic Publisher, 209–242. https://doi.org/10.1007/978-94-011-4317-2_10
- Lacheret-Dujour, A. & S. Kahane. 2020. Unités syntaxiques et unités intonatives majeures en français parlé: inclusion, fragmentation, chevauchement [Syntactic units and major intonation units in spoken French: inclusion, fragmentation, overlap]. In: F. Neveu, B. Harmegnies, L. Hriba, S. Prévost & A. Steuckardt (eds.), *7e Congrès Mondial de Linguistique Française*, July 6–10, 2020, Montpellier, France, Article 14005. <https://doi.org/10.1051/shsconf/20207814005>
- Lefevre, F. & E. Moline. 2011. Unités syntaxiques et unités prosodiques: Bilan des recherches actuelles [Syntactic and prosodic units: Review of current research]. *Langue française*, 170, 143–157. <https://doi.org/10.3917/lf.170.0143>
- Muller, P., M. Vergez-Couret, L. Prévot, N. Asher, B. Farah, M. Bras & L. Vieu. 2012. Manuel d'annotation en relations de discours du projet ANNODIS [Annotation manual in discourse relations of the ANNODIS project]. *Carnets de Grammaire* 21, 1–34.
- Pallaud, B., R. Bertrand, L. Prévot, P. Blache & S. Rauzy. 2019. Suspensive and Disfluent Self Interruptions in French Language Interactions. In: L. Degand, G. Gilquin, L. Meurant, A. C. Simon (eds.), *Fluency and Disfluency across Languages and Language Varieties*, Louvain-la-Neuve, Belgium: Presses universitaires de Louvain, 109–138.
- Portes, C. & R. Bertrand. 2011. Permanence et variation des unités prosodiques dans le discours et l'interaction [Permanence and variation of prosodic units in speech and interaction]. *Journal of French Language Studies* 21(1), 97–110. <https://doi.org/10.1017/S0959269510000499>
- Prévot, L., S.-C. Tseng, K. Peshkov & A. C. H. Chen. 2015. Processing units in conversation: A comparative study of French and Mandarin data. *Language and Linguistics*, 16(1): 69–92. <https://doi.org/10.1177/1606822X14556605>
- Selting, M. 1996. On the interplay of syntax and prosody in the constitution of turn-constructive units and turns in conversation. *Pragmatics*, 6(3): 371–388. <https://doi.org/10.1075/prag.6.3.06sel>
- Shriberg, E. E. 1994. *Preliminaries to a Theory of Speech Disfluencies*. Ph.D. dissertation, University of California at Berkeley.
- Shriberg, E. E. 1995. Acoustic properties of disfluent repetitions. In: K. Elenius & P. Branderud (eds.), *Proceedings of the 13th International Congress of Phonetic Sciences*, August 13–19, 1995, Stockholm, Sweden, 4, 384–387.