Can prosody meet pragmatics?
Case of discourse particles in French

Lou Lee[1,2], Katarina Bartkova[1], Denis Jouvet[2], Mathilde Dargnat[1], Yvon Keromnes[1]

[1] Université de Lorraine, CNRS, ATILF, F-54000 Nancy, France
[2] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
{lou.lee;katarina.bartkova;mathilde.dargnat;yvon.keromnes}@univ-lorraine.fr, denis.jouvet@loria.fr

ABSTRACT

The goal of this study is to investigate how prosodic parameters can contribute to retrieving the pragmatic meaning of words with considerably diminished semantic content. The paper focuses on six French words used as discourse particles (DP): '*alors*' (so), '*bon*' (well), '*donc*' (thus), '*enfin*' (finally/anyway), '*quoi*' (what), '*voilà*' (there you go). Prosodic properties of DPs are analysed with respect to their pragmatic functions, using occurrences of these words extracted from a large speech corpus, and manually annotated with pragmatic labels. It was found that occurrences with the same pragmatic value have a great tendency to share the same prosodic pattern; hence, the question of their commutability arises. An auditory test has thus been conducted to find out whether DPs of uniform pragmatic functions are perceived as commutable and how prosodic features help listeners in identifying considered DPs. The results are analysed and discussed with respect to pragmatic functions and prosodic patterns.

Keywords: prosody, pragmatics, discourse particles, F0 patterns

1. INTRODUCTION

There has been a growing interest in discourse particles (DP) during the past decades and many researchers have been studying this semantico-pragmatic discursive phenomenon ([12], [13], [14], [15], [18]). There are also numerous studies on how prosody can express emotions or attitude of the speaker in utterances, that is, prosody adds non-linguistic load to an utterance conveying other pragmatic functions ([3], [12], [24]).

However, scant attention has been paid to prosodic features of DPs and this is even more so in the case of French DPs. Some of recent studies have focused on how prosodic features of DPs differ from the ones of their non-DP usage ([1], [2], [6], [7], [8], [9], [19]). In these studies, relevant prosodic patterns were found for DP and non-DP usage of the same word as well for different pragmatic functions that DPs obtain not only in their linguistic context but also situational one. Therefore, prosody can be considered as a good indicator to predict DPs' pragmatic functions which are hard to detect by lexical content alone.

In this paper, we present and discuss the results of an experiment that was designed to investigate the prediction of DPs from their contexts. As our goal is to better understand the impact of prosody on the perception of DPs, the prediction has been assessed, on the one hand, using only text data, and on the other hand, using audio data accompanied with text. These experiments give insights on how DPs are perceived as commutable, with respect to their pragmatic functions. For this study, we focused on six French words which are frequently used as DPs: '*alors*' (so), '*bon*' (well), '*donc*' (thus), '*enfin*' (finally/anyway), '*quoi*' (what), '*voilà*' (there you go). The results of our auditory test should also indicate whether DPs of uniform pragmatic functions are perceived as commutable and how prosodic features help listeners in identifying considered DPs.

In addition, an analysis of prosodic patterns of DPs with respect to their pragmatic functions has been conducted. This study relied on an automatic clustering of the F0 (fundamental frequency) values over the DPs (i.e., the DP and preceding and following words). F0 values were extracted from a large number of occurrences of DPs taken from a large speech corpus. Resulting clusters are analysed with respect to the DPs and their pragmatic functions.

2. SPEECH DATA

2.1. Speech data and annotation

DPs in this study were extracted from a large set of French speech corpora comprised of several hundred hours of recording; ORFEO project [23] and ESTER2 evaluation campaign [16]. To avoid being biased by a single type of speech style, these corpora were chosen as they exhibit a large variety of speech styles ranging from storytelling and prepared speech to more spontaneous speech of interviews and interactions. Automatic speech-text alignments were carried out on these speech corpora (see [20] for more detail).

Approximately 1,000 occurrences per studied word were randomly selected and manually annotated after listening to a speech segment spanning each considered occurrence (15 words before and 15 words after). For each word, around 70% of the occurrences were identified and annotated as DP, and then their specific pragmatic value label was attributed. Some

of the pragmatic functions are shared across different DPs, such as introduction, conclusion, reformulation, hesitation, etc. (see Table 2 for details).

Table 1: Annotation agreement rates

Nb. of annotators	Occurrences
Agreement by 3 annotators	44.0%
Agreement by 2 annotators	38.7%
No agreement among 3 annotators	17.3%

A subset of the data has been annotated by three experts. Table 1 reports an analysis of the inter annotator agreement. A total agreement among the three annotators is observed in 44% of the occurrences, and an agreement between two annotators is observed in 38.7%. Therefore, there is an overall agreement between at least two annotators for 82.7% of the occurrences.

2.2. Pragmatic functions

DPs under consideration in this study are as follows: '*alors*' (so), '*bon*' (well), '*donc*' (thus), '*enfin*' (finally/anyway), '*quoi*' (what), '*voilà*' (there you go), which are the most frequent DPs in our corpora. Semantic load of a word when used as a DP is significantly lighter compared to the one when used as non-DP. However, pragmatic meanings remain and this can be multi-categorical for the same DP.

Table 2: Pragmatic functions of DPs

	alors	bon	donc	enfin	quoi	voilà
introduction	X		X			X
reintroduction	X		X			
conclusion	X	X	X	X	X	X
hesitation	X	X		X		
addition	X		X	X		
interaction	X	X	X			X
emotion	X			X	X	
confirmation		X				
parenthetical		X				
interruption		X		X		
dialog transition		X				
reformulation				X	X	
correction				X		
evidence					X	
punctuation					X	

Manually attributed pragmatic labels were defined based on the literature ([5], [10], [12], [22]) and then refined according to their usage in the speech corpora used. Table 2 shows how pragmatic functions are distributed and shared throughout different DPs. Occurrences where a DP was used combined with one or several other DPs were annotated as 'complex-DP' and excluded from our analysis for this paper.

3. DP'S INTERCHANGEABILITY TESTS

As mentioned in the previous section, some DPs share the same pragmatic functions. Moreover, prosodic analysis of DPs (cf. [21]) found that DPs with the same pragmatic functions have also very similar prosodic patterns. Therefore, our hypothesis is that DPs of such occurrences may be interchangeable, which was investigated in the following experiments. In order to also evaluate the contribution of the prosody in DPs identification, two separate experiments were carried out; one relying only on textual information, and the other combining textual and audio data.

3.1. Experimental set-up

The experiments rely on data extracted from the speech corpora described in Section 2. Segments containing 10 to 15 words before and after the DP were presented to participants, after removing the DP; the participants had to indicate what was the missing DP in the given context. Data were presented either as text only, or as text with associated audio. The objective of these experiments was to find out whether DPs with the same pragmatic functions are perceived as commutable; and how prosodic information helps listeners in identifying the appropriate DPs in the context. There were 53 participants, all native French speakers, mainly undergraduate and graduate students in the linguistics department.

20 stimuli of various pragmatic functions for each DP were chosen from randomly extracted data according to the following criteria: good sound quality and articulation clarity. Each DP is situated approximately in the middle of the speech segment, containing left and right contexts sufficiently large to allow a correct interpretation of the conveyed DP's pragmatic function. Using manually checked and corrected segmentations, the speech segments corresponding to the DPs are replaced by a '*hum*' sound that has the same prosodic characteristics as the DP it replaces; this is achieved using an anonymisation script on Praat (cf. [18]). The resynthesised speech signal preserved the DP's original duration for the sake of speech rate coherence of each stimulus, whereas its energy level was set to 6 dB higher than the energy of the DP for a better perception during the tests.

In the text test, only the text transcription of the speech signal was presented. The DPs were erased from the text and their place was indicated by underline characters. In the audio test, the speech signal was presented, accompanied by the text, in order to aid the comprehension since the utterances of DPs are generally very short and it was possible for

the participants to not perceive their place correctly. The tested DPs were the same in the two test sections (text and audio tests). The participants accessed the tests online and they were allowed to listen to the examples as many times as they wished for a better comprehension. Participants were asked to make 3 choices of possible DPs in the context of prompt. The first choice was mandatory while the 2nd and 3rd were optional (see Figure 2).

Figure 1: Auditory test screen for audio tests
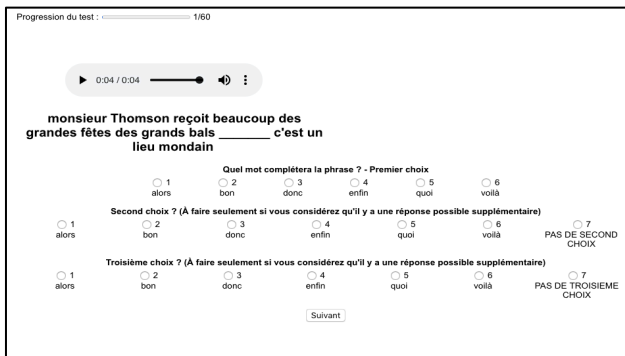
3.2. Results

Table 3 reports the results of the text and audio tests. It indicates the percentage of times the participants answered with the correct missing DP in the first rank answer, or in any of the answers (i.e. including the optional 2nd and 3rd answers), and the average rank of the answers corresponding to the missing DP. An average rank close to 1 indicates that the missing DP was correctly identified mainly as the first choice (first rank answer).

Table 3: Percentage of answers corresponding to the correct DP, and average rank of correct answer

	First rank answer (%)		In one of the answers (%)		Average rank	
	Text	Audio	Text	Audio	Text	Audio
alors	35	51	51	65	1.40	1.28
bon	24	38	39	59	1.49	1.42
donc	40	51	62	68	1.41	1.31
enfin	31	36	52	49	1.51	1.34
quoi	46	69	60	77	1.28	1.15
voilà	29	44	52	61	1.59	1.35

Results show that the presence of audio data facilitates effectively the DP identification compared to when only text data is given; this reveals that prosody has an important role in correct DP identification. For each DP, the audio test has better rates of correct DP identification compared to those of the text test. For instance, in the case of '*quoi*', the correct choice is present in 69% of the rank one answers in the audio test compared to 46% in in the text test.

In addition, the average rank is lower (closer to 1) in the audio tests than in the text tests. This indicates that participants placed the correct DP as their first choice, more frequently when the audio signal is provided, than when the context was presented only in text form. This demonstrates that DP identification was facilitated by audio signals, that is, by the presence of prosodic information.

Table 4: Most frequent first rank answers for various DPs and various pragmatic functions, when predicting DP from Text or from Audio

	Text	Audio
addition		
alors	41% *alors*, 24% *donc*	51% *alors*, 21% *donc*
donc	48% *donc*, 19% *bon*, 15% *alors*	52% *donc*, 18% *alors*
conclusion		
alors	29% *donc*, 27% *alors*, 14% *bon*	31% *alors*, 27% *donc*
bon	30% *voilà*, 21% *alors*, 15% *bon*	51% *bon*, 20% *voilà*
donc	42% *donc*, 19% *enfin*, 18% *alors*	53% *donc*, 15% *alors*
quoi	46% *quoi*, 17% *alors*, 11% *enfin*	66% *quoi*, 13% *voilà*
hesitation		
alors	80% *alors*, 8% enfin, 8% *donc*	87% *alors*, 13% *bon*
bon	29% *bon*, 25% voilà, 25% *alors*	62% *voilà*, 38% *bon*
introduction		
alors	39% *alors*, 23% *donc*, 13% *bon*	57% *alors*, 18% *donc*
donc	51% *alors*, 37% *donc*, 7% *bon*	65% *donc*, 31% *alors*
reintroduction		
alors	30% *donc*, 20% *alors*, 20% *enfin*	35% *alors*, 35% *donc*
donc	32% *alors*, 26% *donc*, 15% *quoi*	31% *donc*, 26% *bon*, 24% *alors*
interruption		
bon	38% *bon*, 26% *enfin*, 15% *alors*	30% *enfin*, 30% *bon*

Table 4 presents more detailed results for a few pragmatic functions that are common to different DPs and the most frequent first rank answers. For example, considering the items corresponding to the DP '*alors*' with pragmatic function 'addition', in the text test, 41% of the first rank answers were '*alors*', and 24% '*donc*'; in the audio test, the answer '*alors*' is more frequent (51%). As predicted, DPs sharing the same pragmatic functions received significantly more of confusion in the choice. As can be seen in Table 4, '*alors*' and '*donc*', which share almost all pragmatic functions, occur quasi-systematically in each other's occurrences as an alternative option, either as the most selected or the second.

Even though better correct identification scores were observed in the audio test, that is, participants were aided by prosody in their choice of DP, some DPs still show a certain level of confusion especially when they share the same pragmatic function. For instance, for the 'reintroduction' pragmatic function, '*alors*' was confused with '*donc*' in the text test (30% '*donc*', 20% '*alors*'). Providing prosody helped improve the results but the confusion was still existent (both '*alors*' and '*donc*' reached 35%).

For the 'interruption' function, '*bon*' in the audio test showed even lower rates of correct identification,

it was confused with '*enfin*', indicating that, in this case, prosody confused listeners to an even greater extent. This could be explained by the similar prosodic patterns shared among different DPs with the same pragmatic functions.

4. ANALYSIS OF F0 PATTERNS

Our tests above show that some DPs, sharing the same pragmatic function, are interchangeable, even when audio is provided. Therefore, one can assume that these DPs should also have similar prosodic markings when sharing the same pragmatic function.

4.1. Prosodic articulation

F0 movements between DPs and their immediate contexts (preceding and following words) were studied to investigate DPs' prosodic articulation according to their pragmatic function. F0 movements measured on DPs and their contexts were classified into three classes according to the F0 slope directions: falling, rising, and plateau (Figure 2).

Figure 2: Example of F0 patterns for DP '*alors*'

falling_plateau   falling_rising
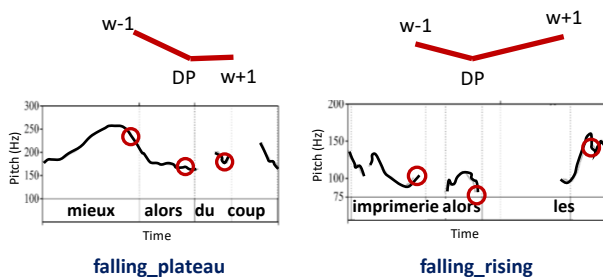
4.2. Prosodic marking of pragmatic functions

In this section, we consider only the F0 patterns to make the study of the DPs prosodic markings easier, although we are aware that listeners use all available prosodic parameters, such as pause occurrences and their durations, vowel energy, syllabic duration, and so on. However, one can consider that F0 pattern is one of the most important prosodic parameters and it should provide valuable information about the prosodic articulation of the DPs with their left and right immediate contexts.

To compare F0 patterns, F0 values are calculated on three places: on the last syllable of the DP's left context (w-1), on the last syllable of the DP, and on the first syllable of the DP's right context (w+1). The F0 values in semi-tones are further normalised according to each speaker's F0 range, each F0 value is thus expressed as a specific level (in %) inside the speaker's F0 range.

In order to retrieve representative F0 patterns for each pragmatic function, a vector quantisation procedure is used. The representative F0 patterns correspond to the centroids of the classes. It was observed that DPs sharing the same pragmatic functions are frequently in the same classes. This grouping obtained on the F0 patterns can further support the idea of the DPs' commutability.

The most frequent F0 patterns of the pragmatic functions shared by several DPs are described in Table 5. Two F0 patterns presented in the table for each pragmatic function are the two most frequently obtained by the vector quantisation technique for each pragmatic function. The F0 level 'low' corresponds to a level lower than or equal to 20% of the speaker's F0 range and the 'high' to higher than 75%.

Table 5: Representative F0 patterns and F0 levels of each pragmatic function

Pragmatic function	F0 Pattern	F0 Level
conclusion	plateau	low
(all DPs)	falling	mid-low
addition	plateau	mid
('donc','alors')	plateau	high
parenthetical	plateau	low
('enfin','bon')	rising-falling	low-mid
reformulation	plateau	low
('quoi','enfin')	falling	mid-low
introduction	rising	low-mid
('alors','donc')	rising	low-high
reintroduction	plateau	mid
('alors','donc')	rising	mid-high

The classes obtained by the vector quantisation technique grouped several pragmatic functions together, though some of them were more frequent than others and thus they strongly influenced the representative centroid values. However, this indicates that F0 patterns of different pragmatic functions can also be similar and this can be why some DPs not sharing the same pragmatic function were also confused during the auditory tests.

5. CONCLUSION

The main goal of this study was to investigate how prosodic information can be of use on the perception and prediction of DPs and whether DPs are interchangeable when sharing the same pragmatic functions, thus, sharing similar prosodic patterns. Results show that prosodic cues do help listeners correctly identify the DPs. Answers show higher confusion rates in the case of DPs sharing the same pragmatic functions, sometimes even when the prosodic information is provided. To confirm this, a prosodic analysis was conducted and found out that different DPs of the same pragmatic functions do share the same or similar prosodic markings.

6. ACKNOWLEDGEMENT

7. REFERENCES

[1] Bartkova, K., Bastien, A., & Dargnat, M. (2016). How to be a Discourse Particle?. In *Speech Prosody 2016*.

[2] Bartkova, K., Dargnat, M., Jouvet, D., & Lee, L. (2017). Annotations de particules de discours en français sur une large variété de corpus. In *ACor4French-Les corpus annotés du français, TALN'2017-Traitement Automatique des Langues Naturelles*.

[3] Barth-Weingarten, D. (2009). *Where prosody meets pragmatics*(Vol. 8). BRILL.

[4] Bartkova, K., & Jouvet, D. (2013). Automatic detection of the prosodic structures of speech utterances. In *International Conference on Speech and Computer* (pp. 1-8). Springer, Cham.

[5] Beeching, K. (2007). La co-variation des marqueurs discursifs bon, c'est-à-dire, enfin, hein, quand même, quoi et si vous voulez : une question d'identité ?. *Langue française*, (2), 78-93.

[6] Bertrand, R., & Chanet, C. (2005). Fonctions pragmatiques et prosodie de enfin en français spontané. *Revue de sémantique et pragmatique*, (17), 41-68.

[7] Cabarrão, V., Moniz, H., Ferreira, J., Batista, F., Trancoso, I., Mata, A. I., & Curto, S. (2015). Prosodic classification of discourse markers. In *International Congress of Phonetic Sciences (ICPhS 2015)*. International Phonetic Association.

[8] Cresti, E., & Firenzuolil, V. (1999). Illocution and intonational contours in Italian. *Revue française de linguistique appliquée*, *4*(2), 77-98.

[9] Dargnat, M., Bartkova, K., & Jouvet, D. (2015). Discourse particles in French: prosodic parameters extraction and analysis. In *International Conference on Statistical Language and Speech Processing* (pp. 39-49). Springer, Cham.

[10] Degand, L., & Fagard, B. (2011). Alors between discourse and grammar: the role of syntactic position. *Functions of Language*, *18*(1), 29-56.

[11] Delattre, P. (1966). Les dix intonations de base du français. *French review*, 1-14.

[12] Denturck, E. (2008). Étude des marqueurs discursifs, l'exemple de quoi. *Master in de Taal-en Letterkunde, Universiteit Gent*.

[13] Dostie, G. (2004) *Pragmaticalisation et marqueurs discursifs. Analyse sémantique et traitement lexicographique*, Bruxelles: DeBoeck/Duculot.

[14] Fischer, K. (Ed.). (2006). *Approaches to discourse particles* (pp. viii+-498). Amsterdam: Elsevier.

[15] Fraser, B. (1999). What are discourse markers?. *Journal of pragmatics*, *31*(7), 931-952.

[16] Galliano, S., Gravier, G., & Chaubard, L. (2009). The ESTER 2 evaluation campaign for the rich transcription of French radio broadcasts. In *Tenth Annual Conference of the International Speech Communication Association*.

[17] Hansen, M. B. M. (1998). *The function of discourse particles: A study with special reference to spoken standard French* (Vol. 53). John Benjamins Publishing.

[18] Hirst, D. (2013). Anonymising long sounds for prosodic research. *TRASP 2013*, 36.

[19] Horne, M., Hansson, P., Bruce, G., Frid, J., & Filipsson, M. (2009). Discourse markers and the segmentation of spontaneous speech–The case of Swedish men 'but/and/so'. *Working Papers in Linguistics*, *47*, 123-139.

[20] Jouvet, D., Bartkova, K., Dargnat, M., & Lee, L. (2017). Analysis and Automatic Classification of Some Discourse Particles on a Large Set of French Spoken Corpora. In *International Conference on Statistical Language and Speech Processing* (pp. 32-43). Springer, Cham.

[21] Lee, L., Bartkova, K., Dargnat, M., & Jouvet, D. (2018). Prosodic and pragmatic values of discourse particles in French. In *ExLing 2018-9th Tutorial and Research Workshop on Experimental Linguistics*.

[22] Lefeuvre, F. (2011). Bon et quoi à l'oral: marqueurs d'ouverture et de fermeture d'unités syntaxiques à l'oral. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (64-65), 223-240.

[23] ORFEO project: http://www.projet-orfeo.fr/

[24] Wichmann, A. (2005). Prosody and discourse: a diachronic approach. In *Actes de la conférence Interfaces Discours Prosodie (IDP), Aix en Provence* (pp. 1-11).