The Prosody of Discourse Makers *alors* and *et* in French: A Corpus-Based Study on Multiple Speaking Styles

George Christodoulides(✉)

Language Sciences and Metrology Unit, Université de Mons, Place du Parc 18, 7000 Mons, Belgium
george@mycontent.gr

Abstract. In this study, we investigate the prosodic characteristics of two French discourse markers (DMs), *alors* and *et*. Our study is based on a 8-h corpus covering 8 different speaking styles, with an average of 10 speakers per communicative situation. The tokens were classified depending on whether they are being used as discourse markers (DMs) or not; additionally in the case of *et* used as a conjunction, the type of the co-ordinated syntactic elements was identified. An automated prosodic analysis of all occurrences was performed. Results show that the use of *et* as a DM was more prevalent in non-planned speech; silent pauses preceded occurrences of *alors* and *et*, both as DMs and as non-DMs; the difference in silent pause duration, in the DM uses vs in the non-DM uses, was not statistically significant for *alors* and was statistically significant for *et*; DMs did not systematically constitute a separate prosodic unit; a strong prosodic boundary differentiates between the use of *et* as a DM or as a co-ordinating conjunction between verb phrases and subordinate clauses, and its other non-DM uses.

Keywords: Prosody · Discourse markers · Corpus linguistics · French

1 Introduction

Spoken language comprehension entails multiple tasks for the listener, such as segmenting the incoming stream of speech, lexical access, syntactic parsing, integration of information into some form of cognitive representation, and understanding of discourse relations. Prosody plays an important role in all these steps, by guiding the listener's comprehension (for a review, see [1,5,7]). The relationship between prosody and information structure, whether specific prosodic structures cue specific discourse relations, and whether prosody can facilitate the processing of discourse relations are research questions whose importance is increasingly recognised.

Fraser defines discourse markers as "a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs, and prepositional phrases [that] with certain exceptions, signal a relationship between the

A. Karpov et al. (Eds.): SPECOM 2018, LNAI 11096, pp. 94–102, 2018.
https://doi.org/10.1007/978-3-319-99579-3_11

interpretation of the segment they introduce, S2, and the prior segment, S1. They have a core meaning, which is procedural, not conceptual, and their more specific interpretation is 'negotiated' by the context, both linguistic and conceptual" [8]. Discourse markers aid in the segmentation of speech (similarly to punctuation marks in written language), and Schriffin defines them as "sequentially dependent elements which bracket units of talk" [15].

In this study, we investigate the prosodic characteristics associated with the use of two discourse markers in French: *alors* (then) and *et* (and). Are there specific prosodic features that can distinguish between the use of these words as a discourse marker, and their use as an adverb or a conjunction (respectively)? When used as conjunction, the token *et* may link (co-ordinate) two segments at different syntactical levels (e.g. two noun phrases, two adjectives). When used as a discourse marker, *et* may convey several discourse relations; this is also the case for *alors* [16]. In this study, we will investigate whether there are prosodic characteristics that distinguish between these uses, on the basis of the C-Phonogenre corpus [10], an 8-h corpus covering 8 different speaking styles.

2 Related Work

Studies have attempted to investigate the phonetic and prosodic properties of discourse markers in speech, using both experimental and corpus-based approaches. For example, [11] confirm the importance of intonation in interpreting the Swedish DM *men* (but/and/so), and in choosing between its sentential interpretation and its interpretation as a DM. They show that when the token *men* is used as a discourse marker, it has a positive f_0 reset, with a mean value of 13.8 ST when preceded by a glottalisation, and of 5.7 ST without glottalisation; whereas in the case of sentential tokens, the mean value of the f_0 reset was 2.2 ST. In English, it has been claimed that DMs constitute a separate prosodic unit surrounded by brief pauses, and that this configuration helps distinguish between DMs and other uses of the same token. However, [12] show that DMs only form a separate intonation unit when opening/closing a conversation or when marking transitions from one topic to another. [12] postulate that the intonation of DMs depends on the speaker's perception of how important a particular marker is, and therefore the relationship between the function of a DM, its prosodic characteristics and its position in the utterance is arbitrary. It has to be noted that studies on the subject are scarce, and therefore it is not yet possible to draw clear conclusions (also given the large number of different discourse markers, and the fact that few language have been studied).

The present study should be read in conjunction with [6], which is a speech elicitation experiment on the use of the DMs *alors* and *et* in French. In this experiment, twenty adult native speakers of French were asked to prepare and to read aloud 64 sequences consisting of a first segment, the discourse marker *alors* or *et*, and a second segment; all first segments were extracted from a speech corpus. The sequences were constructed in order to convey one of six predefined discourse relations. The prosodic characteristics of the resulting recorded utterances were analysed, and results suggest that the silent pause duration before the

DM, as well as the absolute duration of the DM itself are used by the speaker to differentiate between the core meaning of the DM and its less predictable meanings; and that DMs did not systematically constitute a separate prosodic unit. Our study will try to re-evaluate these findings by analysing the occurrences of the tokens *alors* and *et* in a corpus that better represents natural and contextualised language use.

3 Corpus and Methodology

3.1 The C-PhonoGenre Corpus

The corpus used in this study is C-PhonoGenre [10], which was compiled to study situation-dependent speaking styles in French and the associated prosodic variation. It contains data from 8 speaking styles: instructional speech [DIDA]; spontaneous narration [NARR]; speeches during "Question Time" at the French parliament [PARL]; religious sermons [RELG]; radio press reviews [RPRW]; three kinds of sports commentary [SPOR]: rugby, basketball and football; presidential New Years wishes [WISH] and weather forecasts [MET]. The average sample duration per speaker is 5:30 min. The corpus composition is presented in Table 1.

Table 1. Composition of the C-PhonoGenre corpus.

Genre	Sub-Genre	Nb	Dur (min)	Syll	Tokens	Audience	Media	Prepared	Interactive
DIDA	Radio	17	100	26304	18 717	1	2	2	2
	TV					0	2	2	0
	Lecture					2	0	1	0
NARR	Narration	10	44	11396	9 546	1	0	0	2
PARL	Question	10	20	5710	3 613	2	1	2	1
	Answer					2	1	1	1
RELG	Mass on the Internet	7	54	8726	6 141	0	1	2	0
	Sermon on TV					2	1	2	0
RPRW	Radio press review	15	95	26359	17 531	0	2	2	0
SPOR	Basket	5	35	7601	5 305	0	2	0	0
	Rugby/football					1	2	0	2
MET	Weather forecast	10	9	2861	1 947	0	2	2	0
WISH	Pres. New Year	15	98	18614	12 578	0	1	2	0
Total		89	455	107571	75 378				

The corpus samples were selected using the methodology detailed in [10]. The corpus contains recording of both female and male speakers, originating from 3 different French-speaking areas: Metropolitan France, Belgium and Switzerland. Speaking situations were described by features on four dimensions: audience, media, preparation and interactivity; each dimension had 3 different states: 0 indicates absence of a feature (e.g. Preparation = 0 for spontaneous speech) and 2 the full presence of a feature (e.g. Media = 2 for broadcasts), while the value of 1 indicates intermediate situations. For example, Media = 1 indicates

speech directed to an individual or a small group, yet in front of a microphone or camera (indirect audience), and Preparation = 1 indicates semi-prepared speech, situated between spontaneous and read speech. In the case of parliamentary debates, a question is prepared, while the answer is semi-prepared. Interactivity indicates whether the main speaker may be interrupted. The values for each dimension and each speaking style in the C-PhonoGenre corpus are also indicated in Table 1.

3.2 Annotation Methodology and Feature Extraction

The C-PhonoGenre corpus has been manually transcribed orthographically and a phonetic transcription and segmentation was obtained using EasyAlign [9]. The alignment was manually corrected. A single annotator added speech delivery information: (i) disfluencies, articulation and phonological phenomena (schwa, vowel lengthening whether associated to hesitation or not, creaky voice, liaison and elision) (ii) symbols to distinguish between complete silence, audible and less audible breaths, and mouth noises; (iii) indices of paralinguistic phenomena (laugh, cough) and external sounds; (iv) overlapping segments and syntactic interruptions.

The C-PhonoGenre corpus has been processed using the annotation pipeline for French in Praaline [2]. The DisMo annotator [3] was applied to the entire corpus, providing part-of-speech and disfluency annotations. Pitch stylisation was performed using Prosogram [13]. An automatic annotation of prosodic prominence and prosodic boundaries was performed using Promise [4]. Features extracted using these plug-ins are stored in an SQL database, and include durations (of pauses, segments, syllables etc.), pitch information (e.g. intonation contour descriptors), and symbolic annotations (e.g. prominences and boundaries). The database from Praaline was linked to the R statistical software [14] for analysis.

Finally, all occurrences of the tokens *alors* and *et* were identified using Praaline's concordancer, and they were manually annotated depending on whether the token is being used as a discourse marker (cf. the definition given in the Introduction). Additionally, in the case of *et* used as a conjunction, we have annotated the type of the co-ordinated syntactic elements as follows (Table 2):

Table 2. Annotation scheme for *et* when used as a conjunction and not a DM.

Code	Co-ordinated elements	Example
np_np	Noun phrase	ses idées et ses valeurs
pp_pp	Prepositional phrase	dans l' hôpital et dans la médecine
adj_adj	Adjective/Complement	fort et cohérent
vp_vp	Verb phrase	consommons et rejetons
sub_sub	Subordinate clauses	qui se diront et qui se souviendront
num	Number	vingt et un
other	Other cases	

4 Results and Discussion

4.1 Discourse Markers and Speaking Style

In the following we will present the main results of the analysis of the corpus. There were 1944 occurrences of *et* and 177 occurrences of *alors* in all samples. In the case of *alors*, it was used as a discourse marker in 138 (77.9%) of the cases; in the conjunction *alors que* (while) in 35 of the cases and as an adverb in 4 cases. The distribution of the different uses of *et*, normalised by the number of tokens, by speaking style is given in Fig. 1.

Genre	PARL	DIDA	RELG	MET	NARR	RPRW	SPOR	WISH	Total
Total tokens	**3613**	**18717**	**6141**	**1947**	**9546**	**17531**	**5305**	**12578**	**75378**
Conjunction	**1.63%**	**1.06%**	**1.87%**	**1.64%**	**0.68%**	**1.19%**	**0.55%**	**2.50%**	**1.35%**
np_np	0.69%	0.41%	0.47%	0.72%	0.08%	0.55%	0.25%	0.87%	0.49%
pp_pp	0.42%	0.24%	0.47%	0.41%	0.10%	0.25%	0.09%	0.79%	0.34%
vp_vp	0.14%	0.10%	0.67%	0.21%	0.14%	0.15%	0.04%	0.25%	0.19%
adj_adj	0.11%	0.07%	0.11%	0.21%	0.04%	0.11%	0.00%	0.29%	0.12%
sub_sub	0.08%	0.12%	0.10%	0.00%	0.13%	0.07%	0.09%	0.20%	0.11%
locution	0.03%	0.06%	0.03%	0.05%	0.15%	0.03%	0.00%	0.04%	0.05%
num	0.03%	0.05%	0.02%	0.05%	0.04%	0.02%	0.08%	0.04%	0.04%
other	0.14%	0.01%	0.00%	0.00%	0.00%	0.03%	0.00%	0.03%	0.02%
Discourse Marker	**1.00%**	**1.46%**	**0.70%**	**0.67%**	**2.46%**	**0.79%**	**2.21%**	**0.53%**	**1.22%**
Total	**2.63%**	**2.52%**	**2.57%**	**2.31%**	**3.14%**	**1.98%**	**2.75%**	**3.03%**	**2.58%**

Fig. 1. Distribution of different uses of *et*, by speaking style (normalised by the number of tokens).

We observe that in communicative situations where we have spontaneous, non-planned speech (e.g. NARR, SPOR) the majority of the occurrences of *et* were discourse markers, while in the more planned speaking styles (e.g. WISH, PARL, RELG), *et* is used primarily as a conjunction.

4.2 Temporal and Intonational Properties

We then examined the prosodic characteristics of the different uses of *alors* and *et* in our corpus. Figure 2 shows the distribution of the length of silent pauses before DM and non-DM uses of the two tokens. We observe that DM are often preceded by silent pauses; we observe that this is also the case for occurrences of *et* used as a conjunction between verb phrases and subordinate clauses. Furthermore, both DM and non-DM uses of the two tokens were almost never followed by a silent pause. Articulation rate did not significantly vary depending on the DM or non-DM use of the two tokens.

A pitch reset is a prosodic signal for segmentation between the end of a discourse segment and a discourse marker introducing the next discourse segment. Figure 3a shows the pitch movement between the last syllable of the segment between the token *alors* or *et*, by its use (as a discourse marker or not). We observe that DM uses of *alors* tend to have a flat contour, but there is no other

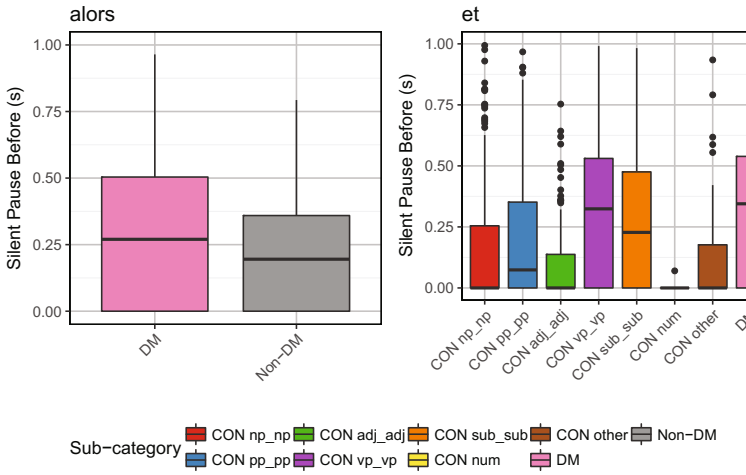
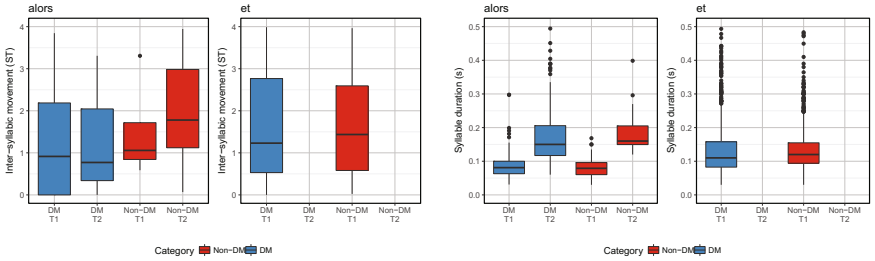


Fig. 2. Pause duration before the token, for DM and non-DM uses of *alors* (left) and *et* (right).

significant use of prosodic cues to differentiate between DM and non-DM uses of *alors* and *et*. With respect to the duration of the two tokens, we do not observe a significant difference between DM and non-DM uses, as can be seen on Fig. 3b.

4.3 Prosodic Prominence and Boundaries

We have also examined the percentage of prosodically prominent syllables, and syllables carrying a prosodic boundary, immediately preceding the tokens *alors* and *et*. The results for prosodic prominence are shown in Fig. 4, and for prosodic boundaries in Fig. 5. We can observe that uses of *alors* as a DM are preceded by a strong prosodic boundary in 54% of the occurrences, compared to 38% of



(a) Pitch movement between the end of S1 and the DM *alors* (left) or *et* (right).

(b) Duration of the DM, for disyllabic *alors* (left) and monosyllabic *et* (right).

Fig. 3. Duration and Pitch reset for DM and non-DM uses of the tokens. T1 and T2 are the first and second syllables of the target DM respectively.

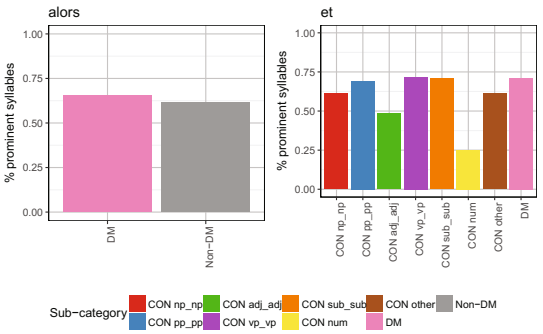


Fig. 4. Prominent syllables (percentage) at the last syllable before the token *alors* (left) or *et* (right).

the occurrences (there is no significant difference for prominence though). We also observe that uses of *et* as a discourse marker are also preceded by a strong prosodic boundary in 48% of the cases. This finding would not be enough to distinguish between DM and non-DM uses of *et*, as a strong prosodic boundary is present in 46% of its uses as a conjunction between verb phrases and 40% of its uses as a conjunction between two subordinate clauses.

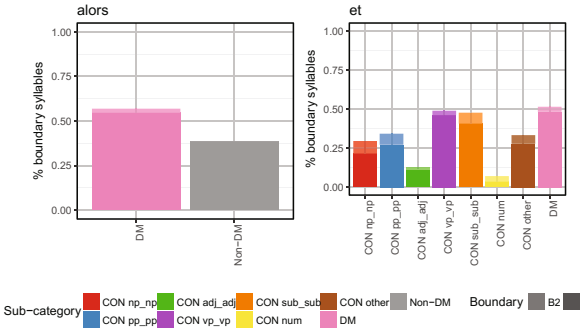


Fig. 5. Prosodic boundaries (percentage) at the last syllable before the token *alors* (left) or *et* (right). B3 = major prosodic boundary and B2 = medium prosodic boundary.

5 Conclusion and Perspectives

In this study, we investigated the prosodic characteristics of *alors* and *et*, two words that are often used as discourse markers in French. We conducted a corpus-based study, based on an 8-h corpus covering 8 different speaking styles, and the results can be summarised as follows:

– The use of *et* as a discourse marker was more prevalent in non-planned speech.

- Silent pauses preceded occurrences of *alors* and *et*, both as DMs and as non-DMs. The Mann-Whitney U non-parametric test shows that the difference between the preceding pause length in the DM uses vs in the non-DM uses was not statistically significant for *alors* and was statistically significant for *et*. In this respect, our corpus study only partly confirms the results of the speech elicitation experiment in [6].
- DMs did not systematically constitute a separate prosodic unit, and both DM and non-DM uses of the two tokens were almost never followed by a silent pause. However, in the case of *et*, a strong prosodic boundary differentiates its use as a discourse marker or as a co-ordinating conjunction between verb phrases and subordinate clauses, and its other non-DM uses.
- There were no statistically significant differences in the articulation rate and in token duration, between the DM and non-DM use of *alors* and *et*.

We plan to expand this study in two directions. First, an annotation of discourse relations expressed by the 138 uses of *alors* and the 922 uses of *et* as a discourse marker, in order to further investigate whether specific prosodic cues are linked to specific prosodic relations. Secondly, we plan to replicate this corpus study on a corpus with longer recordings, so that we can test the effects of individual variation (by examining more occurrences of each token produced by the same speaker). An application of the results of the present study is also envisaged. While prosodic cues seem not to be sufficient to distinguish between DM and non-DM uses of *et*, we would like to test whether the prosodic information identified as pertinent by the present study (i.e. preceding silent pause length and preceding prosodic boundary) can be used to improve the accuracy of statistical parsing of transcriptions.

The prosody associated with the expression of discourse relations, or with the use of certain discourse markers, is highly variable. If such an association does indeed exist, for some specific discourse markers, or in some specific cases of discourse relations (e.g. for the purposes of disambiguation), studies on very large corpora will be needed before we are able to extract meaningful patterns from the data. This is because the prosody of an utterance is influenced by multiple factors, including several factors that are totally unrelated to discourse structure, and because the observed individual variation in the prosodic realisation of discourse relations is fairly high. While experimental studies may indicate relevant acoustic correlates, they are not enough and should be reviewed in light of corpus data, to avoid conclusions based on spurious correlations. More studies, on larger corpora and controlling for individual variation, are needed.

References

1. Carlson, K.: How prosody influences sentence comprehension. Lang. Linguist. Compass **3**(5), 1188–1200 (2009)
2. Christodoulides, G.: Praaline: integrating tools for speech corpus research. In: Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC), May 26–31, Reykjavik, Iceland, pp. 31–34 (2014). http://www.praaline.org

3. Christodoulides, G., Avanzi, M., Goldman, J.P.: DisMo: a morphosyntactic, disfluency and multi-word unit annotator. In: Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC), May 26–31, Reykjavik, Iceland, pp. 3902–3907 (2014)
4. Christodoulides, G., Avanzi, M., Simon, A.C.: Automatic labelling of prosodic prominence, phrasing and disfluencies in French speech by simulating the perception of naïve and expert listeners. In: Proceedings of the 18th Annual Conference of the International Speech Communication Association, Interspeech 2017, 20–24 August 2017, Stockholm, pp. 3936–3940 (2017)
5. Cutler, A., Dahan, D., van Donselaar, W.: Prosody in the comprehension of spoken language: a literature review. Lang. Speech **40**(2), 141–201 (1997)
6. Didirková, I., Christodoulides, G., Simon, A.C.: The prosody of discourse markers alors and et in French. A speech production study. In: Proceedings of Speech Prosody 2018, Poznan (2018)
7. Féry, C.: Intonation and Prosodic Structure. Key Topics in Phonology. Cambridge University Press, Cambridge (2017)
8. Fraser, B.: What are discourse markers? J. Pragmat. **31**, 931–952 (1999)
9. Goldman, J.P.: EasyAlign: an automatic phonetic alignment tool under Praat. In: Proceedings of the 12th Annual Conference of the International Speech Communication Association, Interspeech 2011, 27–31 August 2011, Florence, pp. 3233–3236 (2011)
10. Goldman, J.P., Prsir, T., Christodoulides, G., Auchlin, A.: Speaking style prosodic variation: an 8-hour 9-style corpus study. In: Campbell, N., Gibbons, Hirst, D. (eds.) Proceedings of Speech Prosody 2014, pp. 105–109 (2014)
11. Horne, M., Hansson, P., Bruce, G., Frid, J., Filipsson, M.: Discourse markers and the segmentation of spontaneous speech: the case of Swedish men 'but/and/so'. Working Papers, Lund University, Department of Linguistics, vol. 47, pp. 123–139 (1999)
12. Komar, S.: The interface between intonation and function of discourse markers in English. Engl. Lang. Overseas Perspect. Enq. (ELOPE) **4**(1–2), 43 (2007). https://doi.org/10.4312/elope.4.1-2.43-55
13. Mertens, P.: The Prosogram: semi-automatic transcription of prosody based on a tonal perception model. In: Proceedings of Speech Prosody 2004, 23–26 March 2004, Nara, pp. 549–552 (2004)
14. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2017). https://www.R-project.org/
15. Schiffrin, D.: Discourse Markers. Cambridge University Press, Cambridge (1987)
16. Uygur-Distexhe, D.: Right peripheral discourse markers in SMS: the case of alors, donc and quoi. Papers from the Lancaster University Postgraduate Conference in Linguistics and Language Teaching (2010)