

Prosodic Correlates of Discourse Structure and Emotion in Discourse Markers that Preface Announcements of News

Émilie Marty, Roxane Bertrand, Caterina Petrone, James German

▶ To cite this version:

Émilie Marty, Roxane Bertrand, Caterina Petrone, James German. Prosodic Correlates of Discourse Structure and Emotion in Discourse Markers that Preface Announcements of News. Speech Prosody 2022, May 2022, Lisbonne, France. pp.605-609, 10.21437/speechprosody.2022-123. hal-03773567

HAL Id: hal-03773567

https://hal.science/hal-03773567

Submitted on 21 Sep 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.



Prosodic Correlates of Discourse Structure and Emotion in Discourse Markers that Preface Announcements of News

Emilie Marty ¹, Roxane Bertrand ¹, Caterina Petrone ¹, James S. German ¹

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

emilie.marty@univ-amu.fr, roxane.bertrand@univ-amu.fr, caterina.petrone@univ-amu.fr, james.german@univ-amu.fr

Abstract

Discourse markers serve important structuring functions such as concluding a contribution or resuming a topic. We address whether, along with their role in structuring discourse, discourse markers carry prosodic cues to the emotional valence of upcoming news, perhaps to prepare the listener's emotional reaction. Specifically, we explored the realization of French voilà donc (yeah so) when occurring between an announcement of news and its preface: Je vous appelle au sujet de votre chat qui était malade [preface], voilà donc [discourse marker] il est désormais guéri [announcement] ("I'm calling about your sick cat, yeah so he's now cured"). We recorded 15 speakers reading voicemail messages announcing negative, positive or neutral (e.g., factual) news. We found that the intonation patterns produced with voilà donc correspond to its discursive functions, in line with existing findings, though the choice of pattern did not depend on the emotional valence of the news. Valence was, however, associated with phonetic variation, in that high f0 targets were higher for positive and neutral valence and pitch range was larger for positive valence. This finding suggests that phonetic variation projects the emotional valence of upcoming news even though discourse function primarily determines the choice of intonation pattern.

Index Terms: prosody, intonation, interaction, discourse markers, emotional valence, French

1. Introduction

During conversational interactions, listeners may have the impression that the speaker's voice betrays whether they are about to say something positive or negative. This impression is the consequence of *projection*. Projection is the fact that "an individual action or part of it foreshadows another" [1]. Projection allows participants in a conversation to collaborate gradually towards a successful interaction, thanks to the display of a conversational trajectory [2].

A specific type of projection is the *preface*: "It's about your camera" could preface the negative news "it's broken", or the positive news "I found it". Prefaces are useful especially in the case of negative news announcements, as this type of action is commonly considered a "dispreferred action" which could be rejected or badly received by the listener. Hence, using a preface may help to mitigate the telling of the negative news and to avoid a potential failure of the interaction [3]. Speakers may use prosody to signal in advance of upcoming emotional information [4]. In [4], an acoustic experiment tested whether speakers use prosody in prefaces ("We have had interviews last week") to signal the valence of upcoming news ("and I would like to inform you that we want to offer you the job"). The announcements were produced as if left on an answering

machine. Prefaces to announcements of negative news were characterized by a lower f0 range, lower f0 maximum and faster articulation rate than prefaces to announcements of positive news. These results are in line with previous literature on the vocal expression of emotions, which reported lower f0 means, f0 maxima, and smaller pitch range for negative emotions with low arousal (such as sadness) than for positive emotions with high arousal (such as happiness) [5]. A limitation of the studies reported in [4] and [5] is the fact that they are primarily based on global phonetic measurements, while it is still unclear to what extent phonological aspects of intonation in prefaces can be used to signal the valence of upcoming news [6].

In conversational interactions, speakers may add discourse markers (like 'yeah', 'so' or 'well') to signal other information: "it's about your camera, *yeah so*, it's broken". They are used to build the coherence of the discourse [7][8][9] and, as such, they help to manage conversation by signaling the transition from one topic to another. Transitions can have a conclusive or resuming purpose, for example.

Our study focuses on two French discourse markers: voilà and donc when placed immediately before an announcement. In a corpus-based study, [10] found that when signaling topic conclusion, voilà is often produced with a falling f0. When signaling a strong topic conclusion (i.e., the speaker emphasizes the change to a new topic), voilà is characterized by a "bellshaped" f0 pattern (whose phonology will be discussed further along). A third function of voilà is persuasion, where the speaker wants to convince the listener to agree. Authors found it to be linked to an f0 rise. The discourse marker donc ("so", "consequently") is associated with similar functions [11]. When used as a conclusive, it tends to be realized with a flat contour. Another study [12] focusing on voilà accounts for the use of "voilà donc" used in a row. It determined that voilà functions also to manage the organization of a sequence of actions, namely it marks the transition between the end of the first one and the beginning of the second. When paired with donc, voilà still highlights the upcoming sequence, while donc performs the action of transition. For both voilà and donc, it is not clear to what extent the projection of valence in upcoming news might affect these discursive functions and/or influence their prosody. Our study aims at assessing how the phonological and phonetic prosodic characteristics of voilà and donc depend on whether they precede the announcement of negative, positive, or neutral news. In a laboratory-recorded corpus of news announcements, we observed the realization of voilà and donc when combined into a larger two-part discourse marker. Based on [10], we hypothesize that speakers will more often use a bell-shaped pattern to signal emotional involvement on the matter discussed, i.e., when about to deliver emotional information (negative or positive news) compared to neutral news (i.e. "it's on the table"). A falling tonal pattern is expected when voilà donc has a conclusive function, and its presence should be independent of the valence of the upcoming news.

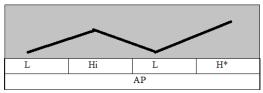


Figure 1: Schematic contour of the basic French tonal pattern For the analysis of intonational phonology, our study follows the model for French developed in [13][14][15] and based on Autosegmental Theory. They characterize the basic underlying tonal pattern as the series of tones /LHiLH*/ (fig.1) which are associated with an Accentual Phrase (AP). The LHi is a bitonal phrase accent marking the left edge of the AP and its realization depends on many factors such as the length of the accentual phrase or the speaking rate. Its L tone is associated to the left edge of the first syllable of the first content word in the AP while Hi is more variable as it can be realized in either of the first two syllables. The LH* is considered an obligatory pitch accent. The L tone is realized in either of the two last syllables of the AP, and the H* tone is associated to the right edge of the AP. This basic underlying pattern can be effectively realized in various ways: LLH* (without Hi) or LH* (without LHi), HiLH* (without the initial L). Patterns ending with H* are the default but [14] assessed that when the H* part of the pitch accent cannot be realized due to following constraints, it is realized as L*, resulting in a LHiL* pattern. Thus, we assume the following patterns to be the transcriptions of the shapes described by [10]: HiL* (the falling pattern for topic conclusion), LHiL*(the bell-shaped pattern for stronger topic conclusion) and LH*(the rising pattern for persuasion). In our transcription, the marking of low tones differentiates between L1, the phrase-initial tone, and L2 the tone that precedes H*. The high target Hi of the LHi rise was annotated in the first two syllables of the AP. The phrasal L1 tone of the LHi rise was marked at the left boundary and left edge of the first content word, and the second L2 was placed within the second part of the AP. The pitch accent LH* was realized on the last syllable, or on the prenuclear syllable.

Within this context, the central questions addressed by our study are the following: (1) Is the valence of the news announcement anticipated in the phonological or phonetic characteristics of the voila donc? We expect that the patterns associated with the functions of the discourse marker will be widely present, especially those associated with topic conclusion (HL*) and strong topic conclusion (LHiL*). If the valence of the news does influence the choice of pattern, we should find more instances of LHiL* on voilà donc preceding emotional news (either negative or positive) than neutral ones, as this pattern has been linked in the literature with the idea of "emotional implication". Patterns could be part of an AP, ip or IP. We also consider the possibility that voilà donc will be considered as two different elements needing two different prosodic units. As for phonetic correlates, we expect higher f0 values for high targets and a larger pitch range when the speaker announces positive news[16]. In contrast we expect speakers forecasting negative news to use a lower pitch range and produce high targets with a lower peak f0 as compared to positive news.

2. Method

Participants

Fifteen native French female speakers (mean age = 38 y.o., SD = 6.5) were recruited for the main study. Each participant provided written consent and was compensated 10 euros. The local ethics committee at Aix-Marseille University approved the study.

Materials and procedure

The materials for the study included 36 written scenarios involving news announcements. The scenarios were grouped into 12 contexts, with three matched valence types for each context: positive, negative, or neutral news (12 contexts x 3 valences). The condition for neutral valence was included for comparison purpose. All scenarios concerned everyday news drawn from professional and institutional contexts. The announcer's field of work varied (veterinarian, steward, teacher...). An example of scenario is illustrated in Table 1.

PREFACE	Discourse Markers	ANNOUNCEMENT	
I'm calling about your flight tonight (Je vous appelle au sujet de votre vol de ce soir)	yeah so (voilà donc)	Boarding has been moved from Gate 1 to Gate 4 (L'embarquement aura lieu en porte 1 et non en porte 4)	Neutral news
		We can offer to upgrade you to first class, with the same price (Nous vous offrons une place en première classe, au même prix)	Positive news
		It's canceled and we won't find a new one today (II a été annulé et nous ne pouvons pas vous en proposer un autre aujourd'hui)	Negative news

Table 1: Example of a recorded scenario with news announcements displayed by valence type.

Participants were recorded in an anechoic chamber at the Laboratoire Parole et Langage. They were asked to read aloud the announcements as if left on an answering machine [4]. This avoids any influence of the recipient's reaction to the announcement and allows for greater control over the manner of delivery across valence conditions. Each speaker read all the scenarios in every valence condition, and these were organized in three blocks of neutral, positive and negative valence, with the order of blocks counterbalanced across participants. Before the reading of the scenarios with news announcement (Table 1), they were provided with a small written context: "You are a flight attendant, and you call a client. You offer an upgrade from economy to first class for the same price. You call her and leave a message on the answering machine".

The 36 test items were drawn from a previous study based on a larger set of scenarios (63 scenarios). Prior to the acoustic recordings, the scenarios were validated in their written forms for intended emotional valence using an online rating task. 30 participants between age 18 and 30 were recruited. They read the 63 scenarios and then assessed their valence using a sliding analog scale (bad news -100 to +100 good news). The 36 contexts with the optimal valence characteristics were retained for the current study.

Analyses

In total, 540 observations (12 scenarios x 3 valences x 15 speakers) were obtained. Two items were excluded due to mispronunciation of the item texts. The items were transcribed semi-automatically at the word and phone level using SPPAS [17]. Subsequently, we annotated manually the phonological patterns of voilà donc in Praat [18]. Our transcription protocol is close to French ToBI [19] but adheres more closely to the inventory proposed by [13] and [14]. The prosodic annotation included two tiers, one for the tones and one for the boundaries. Boundaries were annotated at three phrase levels: the Intonational Phrase (IP), the intermediate phrase (ip) and the Accentual Phrase (AP). Occasionally, voilà donc was realized in two prosodic units (one on voilà and one on donc) The most recurrent pattern was voilà donc uttered in a single unit (either one AP, or one IP containing one AP), and these are the ones we focus on in this article.

For the phonological analyses, we considered the four patterns that reached the threshold of making up at least 5% of all collected trials. We ran one logit model (GLM) for each of the four most frequent patterns. We tested for presence vs. absence of the patterns in the items [20]. For each phonological pattern, we tested the effect of the VALENCE of the upcoming news (positive, negative, neutral). The model includes, in addition to a random intercept for SPEAKER, a random slope in VALENCE. For the phonetic analyses, we ran linear mixed models. The models were used to analyze the effects of the valence of the news on the f0 maximum of the H targets (either Hi or H*) and on the f0 range (f0 max – f0 min) calculated on the time span of the voilà donc. The linear mixed models tested the effect of VALENCE (negative/positive /neutral) on phonetic correlates (f0 at the high targets and pitch range). SPEAKER and ITEM were included as random intercepts, with by-speaker random slopes for VALENCE. Statistical analyses were realized using R [21] and the lme4 package[22].

3. Results

Phonological data

Most of the *voila donc* in the corpus were uttered as a single prosodic phrase (426 out of 538 items). Table 1 shows the repartition of those most frequent patterns across all items.

Table 1: Presence of the patterns in the corpus.

Pattern	Percentage	Count
HiL*	19%	100
L1HiL*	14%	78
HiL2H*	9%	48
HiL%	8%	44
L1L2H*	5%	28
L1HiL2H*	4%	22
other	41%	218

Table 2 shows the occurrence of the four most recurrent phonological patterns across the valence of the upcoming news. As it can be seen, they are distributed evenly across the three valence conditions. Neither the difference between the negative and the positive valence ($\beta = 0.0$ SE = 0.0 t = 0.26, p = 0.79), nor the one between the negative and the neutral valence (= 0.2, SE = 0, t = 1.17, p = 0.24) were significant. Thus, phonology doesn't seem to be used in

discourse markers preceding announcements to convey emotional valence.

Table 2: Occurrence of the more frequent patterns of voila donc across the three news' valences.

Pattern	Negative	Positive	Neutral
HiL*	17%	18%	21%
L1HiL*	14%	12%	19%
HiL2H*	9%	9%	8%
HiL%	7%	7%	10%

Phonetic data

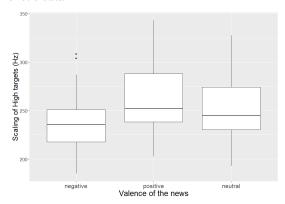


Figure 3: High Targets (Hz) by valence of the news

As shown in Figure 3, valence has an effect on the f0 of high targets realized within the *voila donc*, in that f0 is higher overall in the positive (β = 26.5, SE = 5, t = 5.36, p < 0.001) and neutral (β = 17.2, SE = 5, t = 3.46, p < 0.01) conditions as compared to the negative condition. There was no difference between the height of high targets in positive *versus* negative news.

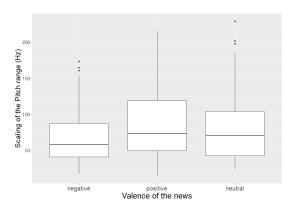


Figure 4: Pitch range (Hz) by valence

As shown in Figure 4, valence has a significant effect on the f0 of pitch range in that f0 is larger in the positive (β = 17.2, SE = 5, t = 3.4, p < 0.01) condition as compared to the negative.

We also observed phonetic variation in the production of *voilà donc* when looking at the phonetic variation for each of the four most frequent phonological patterns. We found that the direction of the effects is confirmed for the four patterns, but the phonetic variable affected by valence differ depending on the specific pattern. For phonetic data in HiL* and HiL% patterns, valence has a significant effect on the f0 of high targets

in that f0 is higher in the positive news as compared to the negative news ($\beta=0.0$, SE=0.0 t = 2.43, p<0.05; $\beta=0.2$, SE=0.0 t = 6.91, p<0.001), and has also a significant effect in the neutral news compared to negative news for HiL% ($\beta=0.1$, SE=0.0 t = 4.92, p<0.001). In L1HiL* pattern, valence has a significant effect on the f0 of pitch range, which is wider in the positive and neutral news conditions as compared to the negative news ($\beta=0.2$, SE=0.1, t = 2.37, p < 0.05; $\beta=0.3$, SE=0.1, t = 3.32, p < 0.01). These observations are similar to the ones found across all patterns.

4. Discussion

The goal of this study was to investigate whether *voila donc* carries prosodic cues to the emotional valence of upcoming news, along with its usual discourse functions, when placed between the preface and the announcement of the news. To answer this question, we created and recorded a set of 36 voicemail-like messages of news announcements. We found that the phonetic realization of *voilà donc* is different depending on the valence of the upcoming news. At the phonological level, there was substantial variability in the phonological patterns employed on *voilà donc*, but their use did not relate to the emotional valence of the news. In particular, the four most recurrent patterns produced on *voila donc were* HiL*, LHiL*, HiLH* and HiL%.

The variability in the choice of the phonological patterns might be linked to the multiple discourse functions conveyed by voilà donc. HiL* and HiL% could be linked to the function of 'topic conclusion' already found for the discourse marker voilà [10]; similarly, LHiL* could be linked to stronger topic conclusion [10][11][12]. Indeed, topic conclusion is coherent with the placement of voilà donc at the end of the preface. HiLH* pattern was not expected, because it didn't correspond to any of the previous phonological descriptions associated to the functions of *voila* and *donc*. A possible explanation for the occurrence of the early rise (L)Hi in the HiLH* pattern could be due to a need of highlighting the discourse marker [10]. The late rise LH* in the HiLH* pattern corresponds to the continuation rise described in the literature on French intonation [23][24]. The use of a LH* continuation rise appears coherent with the action organization function found by [12] in a corpus study of mediatic interactions. It allows transition from the previous action of prefacing to the following action of announcing the

LH*, which has been traditionally linked in the literature to persuasion function [10], was found in very small quantities in our corpus. At the first impressionistic observations, this LH* does not seem to be linked to persuasion, but rather to the continuation rise.

The essential discourse structuring function of discourse markers described above could explain the absence of a link between emotional valence and phonology, as it seems that structuring discourse has been preferred over the projection of emotional cues. However, our results show phonetic variations in the production of *voilà donc* depending on the valence of the upcoming news. High targets (Hi or H*) had higher f0 values when the *voilà donc* preceded the announcement of a positive or neutral news as compared to negative news. Similarly, the pitch range on *voila donc* was larger before positive news. This is in line with literature on vocal expressions of emotions [4][5] showing that emotions with higher arousal and positive valence (such as joy) are expressed with higher f0 mean and larger pitch

range than emotions with lower arousal and negative valence (such as sadness).

We also looked at phonetic variation on a subset of the corpus, within each of the four most frequent phonological patterns. This analysis was done to evaluate whether the valence effects on f0 varied depending on the phonological pattern employed. We found differences in the sensitivity of the phonetic correlates to the valence of the news depending on the pattern. For two patterns (HiL* and HiL%), the height of f0 in the high targets on *voila donc* was more sensitive to valence conditions. On the other hand, for LHiL*, pitch range was more sensitive, in that it was larger in contexts of positive and neutral news compared to negative news. It would be interesting in future analysis to explore the magnitude and direction of these effects as a function of the choice of a specific phonological pattern.

Results concerning the scaling of the high targets produced on *voilà donc* before neutral news were unexpected. Previous studies have shown that neutral prosody can be confused with sad prosody, which is a negative valence emotion[25]. It has also been shown that emotions with low arousal are more likely to be confused with neutral prosody [26][27]. However, in our corpus the characteristics of the prosody of *voilà donc* in neutral valence news are closer to the ones found in positive valence news. A reason for this divergence could be that neither of the neutral or positive news can be considered a dispreferred action. The need to project emotional valence in those contexts could be less important for speakers than in negative news announcements.

5. Conclusion

Our study focused on the discourse markers placed between the announcement of news and their preface. Our aim was to assess whether phonology and phonetics convey emotional cues on tiny parts of the discourse usually studied for their structural use. We found that discourse markers do not seem to bear emotional cues in the construction of the phonological patterns, but phonetic variations inside of the patterns add nuance to their meaning. In the future, looking at the prefaces could allow to verify whether the absence of link between phonology and discourse markers was inherent to their function.

6. Acknowledgements

This work was done as part of a thesis project funded by Aix-Marseille University and supported by the research group Physiology and Speech Organization of Laboratoire Parole et Langage, France. We thank Sabrina Kadri for helping with the recording of the corpus.

7. References

- [1] Auer, P. (2005). Projection in interaction and projection in grammar. Text-Interdisciplinary journal for the study of discourse, 25(1), 7-36.
- [2] Levinson, S.C. (1983). Pragmatics. Cambridge, Cambridge University Press
- [3] Schegloff, E. (2007). Sequence Organization in Interaction: A Primer in Conversation Analysis. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511791208
- [4] Swerts, M., & Hirschberg, J. (2010). Prosodic predictors of upcoming positive or negative content in spoken messages. The Journal of the Acoustical Society of America, 128(3), 1337-1345. https://doi.org/10.1121/1.3466875
- [5] Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms. Speech Communication, 40(1-2), 227-256. https://doi.org/10.1016/S0167-6393(02)00084-5.
- [6] Cao, H., Beňuš, Š., Gur, R. C., Verma, R., & Nenkova, A. (2014). Prosodic cues for emotion: analysis with discrete characterization of intonation. Speech prosody (Urbana, Ill.), 2014, 130.
- [7] Schiffrin, D. (1987). Discourse markers (No. 5). Cambridge University Press.
- [8] Schiffrin, D. (2001). Discourse markers: Language, meaning, and context. The handbook of discourse analysis, 1, 54-75.
- [9] Fraser, B. (1999). What are discourse markers?. Journal of pragmatics, 31(7), 931-952.
- [10] Petit, M. (2020). La prosodie de voilà en français dans le discours médiatique. De Gruyter. https://www.degruyter.com/document/doi/10.1515/97831106224 54-005/html.
- [11] Ferre, G. (2011). Multimodal Analysis of Discourse Markers' donc', 'alors' and'en fait'in Conversational French. In ICPhS VXII (pp. 671-674).
- [12] Haileselassie, A. (2015). Voilà, an orientation shift marker in modern French discourse: A conversation analytic perspective. University of Illinois at Urbana-Champaign.
- [13] Jun, S. A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), Intonation. Text, Speech and Language Technology, 15. https://doi.org/10.1007/978-94-011-4317-2_10.
- [14] Jun, S. A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. Probus, 14(1), 147–172. https://doi.org/10.1515/prbs.2002.00
- [15] Welby, P. (2006). French intonational structure: Evidence from tonal alignment. Journal of Phonetics, 34(3), 343–371. https://doi.org/10.1016/j.wocn.2005.09.001.
- [16] Amir, N., & Globerson, E. (2014). On the Role of Pitch Perception of Emotional Speech. In *Speech Prosody*, 2014 (pp. 154-158).
- [17] Brigitte Bigi (2015). SPPAS Multi-lingual Approaches to the Automatic Annotation of Speech. In "the Phonetician" -International Society of Phonetic Sciences, ISSN 0741-6164, Number 111-112 / 2015-I-II, pages 54-69.
- [18] Boersma, Paul & Weenink, David (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.56, retrieved 3 November 2021 from http://www.praat.org/
- [19] Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Cristo, A.D., Feldhausen, I., Jun, S., Martin, P., Meisenburg, T., Rialland, A., Sichel-Bazin, R., & Yoo, H. (2015). Intonational phonology of French: Developing a ToBI system for French.
- [20] Larson-Hall, J. (2008). Analyzing linguistic data: a practical introduction to statistics using R. Harald Baayen (2008). Sociolinguistic Studies, 2(3), 471-487.
- [21] R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/
- [22] Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using Ime4." Journal of Statistical Software, 67(1), 1–48. doi: 10.18637/jss.v067.i01.
- [23] Delattre, P. (1966). Les dix intonations de base du français. French review, 1-14.

- [24] Martin, P. (2002). Regional Variations of Sentence Intonation in French-The Continuation Contour in Parisian French. In Speech Prosody 2002, International Conference.
- [25] Audibert, N., Aubergé, V., & Rilliard, A. (2005). The prosodic dimensions of emotion in speech: the relative weights of parameters. In Ninth European Conference on Speech Communication and Technology.
- [26] Audibert, N. (2008). Prosodie de la parole expressive: dimensionnalité d'énoncés méthodologiquement contrôlés authentiques et actés (Doctoral dissertation, Institut National Polytechnique de Grenoble-INPG).
- [27] Laukka, P., Audibert, N., & Aubergé, V. (2012). Exploring the determinants of the graded structure of vocal emotion expressions. Cognition & Emotion, 26(4), 710-719.