

The Discourse Function of Final Rises in French Dialogues

1 Introduction

In this paper, we report the results of an empirical study which aims at describing the discourse meaning of rises at right edge intonation boundaries in French dialogues. According to most French speakers, it is possible to turn an assertion into a question in French solely by pronouncing it with a rising intonation. While existing empirical studies [8, 7] confirm that there is some correlation between rising and falling contours, and questions and assertions, respectively, they also show that rising intonation does not always go hand in hand with question intonation. Leaving aside the problem of question identification, one can thus legitimately raise the issue: What is the meaning of final rises in French?

Clearly, an answer to this question cannot be given without a proper empirical study of the use of rises in natural speech but a corpus study of this sort is currently lacking for French (for other languages, see [13], for Glasgow English, or [6] for Australian English). An initial study has been performed on Post's Map Task corpus [16] with two speakers and two dialogues (for a total of 301 speech turns); its goal was to resolve a number of annotation issues, such as the definition of intonation phrase boundary, its automatic assignment, the reliability of the employed algorithm and alphabet for intonation transcription, as well as the contribution of different kinds of dialog act and discourse structure taxonomies. The tested methodology is used for the study of the Caelen corpus [2], which is currently in progress.

In the following sections of this paper, we first describe in detail the basic theoretical issues pertaining to the methodology of analyzing the discourse function of intonation. In section 2., we focus on the definition of rises and on assignment of intonational boundaries. In section 3., we discuss the annotation of dialog acts and dialog structure. In the final section, we present the results of Post's Map Task corpus study.

2 Annotating Intonation: rises and intonation boundaries

2.1 Definition of a rise

As noted by [16], there exists no consensus in French intonation studies about which changes in contours are categorical and whether one should take contours as holistic units or as a composition of individual tones, anchored on stressed syllables and intonation unit boundaries. Also with respect to rising intonation, a number of proposals can be found in the literature [8, 16, 9]. Leaving aside direct perceptual distinctions, one can either opt for a phonetic description – with direct reference to the F0 contour, or a phonological one, which presupposes the adoption of an intonational grammar (a theoretical construct). A phonetic description (e.g., points of maximum and minimum pitch) has the advantage that its result is annotator-independent; it can be done automatically and its quality depends only on the algorithm used to calculate f0. The disadvantage is that there may be no linguistic reality corresponding to the phonetic information and generalizing over phonetic parameters may be difficult in a larger corpus study (with (semi-)free speech and many speakers). A phonological description (like ToBI) is/should be by definition linguistically relevant but it is time-consuming, costly (several professional annotators have to be employed), not quite reliable and with respect to intonation meaning probably both too fine-grained and not powerful enough.

In our present study, we have made use of the INTSINT annotation system which, in our view, circumvents some of the difficulties mentioned above. INTSINT (INternational Transcription System for INTonation) [10] is a language-independent intonation transcription system developed in Aix-en-Provence. It is phonetic to the extent that it makes use of automatically calculated macro-prosodic component of the F0 done by the accompanying MOMEL (MODélisation de MELodie) algorithm; at the same time it is phonological in that

it only labels certain target points on the MOMEL curve which are assumed to carry linguistic information. Basically, the MOMEL algorithm [11] provides an automatic stylization of the F0 contour, detected from the acoustic signal with the comb algorithm [4] (see also [14]). INTSINT covers both absolute prosodic events (**T** – Top; **M** – Mid; **B** – Bottom) and relative ones (**H** – Higher; **S** – Same; **L** – Lower; **U** – Upstep; **D**). The results still have to be checked manually but in general, the process is less time consuming than ToBI labeling.

2.2 Intonation Boundaries

When studying the role of rising intonation, it is not enough to focus on utterance boundaries given that all boundary tones associated with right edges of intonational phrases are assumed to be meaningful (viz [3]). Intonational phrases in French are optionally associated with acoustically and perceptually identifiable events of both rhythmical and tonal nature, such as pauses, drop in amplitude, final syllable lengthening, pitch resetting (on the first syllable of the subsequent phrase), and lack of some segmental assimilation processes (viz [12], [16], [5], among others). Nevertheless, they also appear to be related to information structure articulation (e.g., [3] define them only as a reflection of the information structure of an utterance) and it is normally assumed that there is some correlation between prosodic phrasing and syntactic boundaries. Taking these observations into account, the following rules were used in the process of intonational boundary annotation, together with the perceptual impression of the speech signal:

1. Every completed turn boundary is a right edge IP boundary.
2. Phonologically, an IP boundary is often (i) indicated by a pause, (ii) accompanied by syllable lengthening of the preceding syllable, (iii) followed by pitch resetting and (iv) accompanied by a drop in amplitude.
3. An IP boundary often coincides with a major syntactic boundary (e.g., a finite clause boundary).
4. An information structure constituent (topic, focus) can be followed by an IP boundary.

The results were evaluated for inter-annotator agreement with the average κ for the three annotators being .718 ('good'). Evaluation of problematic examples showed that short phrases like '*oui*' (yes) were often a source of disagreement, cf., e.g., *Bonjour, oui...* Note that a general rule is impossible, since in some cases, *oui* is clearly parenthetical, identifiable by lower intensity than the rest of the unit, and should be treated as a separate IP, while other examples are more arguable. Short phrases such as '*bon*', as in '*Bon, d'accord*', particles and adverbial phrases like '*alors*', '*donc*' or '*parcontre*' and the utterance final '*quoi*' raised a similar problem. The annotators also disagreed at hesitation points (often filled with '*eah*') and interruptions and self-corrections,¹ and at events which normally imply an intonational phrase boundary, such as pauses and vowel lengthening. Given that in case of disagreement, it was usually difficult to decide for or against a label, all the proposed intonational phrase boundaries were merged together in the final annotation.

The results of manual intonational phrase boundary annotation were compared to a semi-automatic method of boundary assignment, based on the automatic determination of pauses (with minimal length of 15 ms and maximal intensity of 40 dB) and a manual assignment of boundaries to all points of speaker switch. The semi-automatic method gave 2/3 and 3/4 of the manually assigned intonational phrase boundaries; only in a small number of cases did the pause not coincide to the original IP label, mostly due to long pre-plosive silences (some of them longer than 350ms). Also in view of the fact that some of the original manually assigned boundaries were quite likely just boundaries of smaller phrases (i.e., the accentual phrases), the result of the semi-automatic method was judged suitable to replace the manual method in the Caelen corpus study.

3 Annotating dialogue acts and dialogue structure

3.1 Dialogue acts

As noted above, rising intonation is often assumed to be a marker of questions. One problem with testing this intuition empirically is that many utterances are ambiguous between questioning and asserting. [8]:26

¹These have been found to be problematic also in MAE-ToBI (the American English standard for prosody labelling, viz [1]) for the same reason.

lists the following cases as typically posing a problem to a clear question/non-question classification: (i) the speaker wants a simple confirmation from the addressee; (ii) the speaker is making a supposition which is only partially interrogative; (iii) the speaker is suggesting some word to the addressee to complete his utterance; (iv) the speaker pronounces only a part of his utterance which would have been a question if completed. [7] considered as questions all utterances that received a *oui/non* reply. This definition is most likely too strong, given that many assertions receive an acknowledgment synonymous with the *oui*-reply, as well as too weak: some utterances which function as replies only contextually entail a *yes/no* response or express speaker's ignorance with respect to an issue. Questions are also often defined with reference to their intonation but for the purposes of the current study, it was necessary to identify them independently of their prosodic properties.

Assuming that the annotation of wh-questions is unproblematic, we made use of the following definition to identify polar questions (PQ), originally developed for English. The definition takes into consideration segments larger than single utterances and/or turns.

Polar Questions [Def.] A polar question is an utterance that satisfies the following properties:

- it is turn-final
 - it is followed by a reply from the addressee that contextually entails *yes/no/I don't know*
 - if the utterance is of a declarative form, it can in the context be turned into a corresponding interrogative by inverting the subject and the finite verb, without resulting in an infelicitous discourse.
-

While the first two conditions are more or less straightforward, the third one actually relies on the intuitions of the annotator; nevertheless, it seems that a good inter-annotator agreement can be achieved. Since in French, syntactic inversion is a rather obsolete way of forming questions, the inversion test can be replaced by a similar one using the *est-ce que* phrase. For example, in the context bellow (taken from Post's Map Task corpus), the declarative (I_{106}) satisfies the first two conditions of the definition above and can also be felicitously turned into an *est-ce que* question in its context. While the definition of questions proposed above may not identify *all* utterances intended as questions, it rarely overgeneralizes (though it is sometimes difficult to distinguish between the *assertion - acknowledgment* and *question - answer* sequences).

- (1) (I_{103}) est-ce que tu as IP tu as le profond étang H IP
 (K_{104}) oui, H sur la gauche IP
 (I_{105}) oui, tout à gauche. IP
 (I_{106}) et tu as la grande plaine H IP DA
 (K_{107}) non IP DA *no*

Apart from questions, the corpus was also annotated for other types of dialog acts, partly based on an existing taxonomy for route description dialogues. For the dialogue act annotations, the annotators had no access to the recordings and to the original punctuation signs in the transcript to avoid bias. Because of the difficulty of the segmenting and labeling task, the final annotation was mainly based on a post-hoc discussion. As in the case of intonational phrase boundary assignment, a number of problematic cases was identified, e.g.,

- (i) It was sometimes difficult to determine whether an utterance was an alternative bipolar question with an ellipsis of the second constituent, or whether the *ou* connective merely served to indicate speaker's uncertainty (viz 2), especially if the question was responded to with a *'yes/no'* answer.
 (ii) The *'est-ce que'*-test for questionhood sometimes gave unnatural renderings of the original declaratives, given that this form of questioning is rarely used in spoken French.
 (iii) It was also difficult to decide whether a sequence of *'d'accord'* - *'d'accord'* represented a feedback elicitation (ALIGN in the MAPTASK schema [13]) and its answers or only two acknowledgments.
 (iv) With respect to wh-questions, there is a potential difficulty with interrogative utterances with an ellipsis, as in some cases (viz (3) below), the *wh* constituent may be missing.
 (v) It appeared desirable to classify also utterances of the sort exemplified in (4) as questions.

- (2) (I_{198}) et, IP à beaucoup de centimètre du pic IP ou IP (*and at many centimeters from the peak or*)
 (K_{199}) euh IP oui, H au moins, euh IP au moins cinq, hein IP (*err yes at least err at least five*)

- (3) (*I*₁₈₇) et alors l'hôtel T IP par rapport aux torrents et l'océan IP (*and so the hotel with respect to the torrents and the ocean*)
 (*K*₁₈₈) alors, IP il est IP euh U IP presque euh, IP le, le, T IP la dernière dune des torrents IP (*so it is almost err the the last dune of the torrents*)
- (4) (*I*₁₂) il y a un petit pin T IP tout à droite de la feuille T IP (*there is a little pine tree at the extreme right of the sheet*)
 (*K*₁₃) je ne sais pas si tu le vois H T IP (*I don't know if you can see it*)
 (*I*₁₄) oui H IP (*yes*)

The presence/absence of a rise against the main act categories (acknowledge, instruct, inform, question, answer, with or without new referent introduction) was statistically evaluated. There was no convincing correlation between the act labels and the presence of a rise, except for instruction (with new referent) associated with a rise (chi-square, $p=0.006$), answers to a question with absence of a rise ($p=0.001$), and polar question (with new referent) with a rise ($p=0.03$).

3.2 Annotating Discourse Structure

The main aim of the discourse structure annotation task was to test for a possible correlation between discourse opening/closing and rises. However, it appears that at this level of discourse organization, two organizational principles are in competition with each other – i.e., game and topic structures. While game and topic openings are often realized through the same move (e.g., a question), new games do not necessarily bring new topics into discussion (e.g., simple checks or verification questions) and topic shifts can occur within one game (e.g. a long speech-turn introducing a complex discourse structure made of several discourse topics). The clues for recognizing these structures are also very different: While topic structures may require a deep semantic understanding of the conversation (or at least keeping track of new/old information), game structures can be determined from the series of individual moves types.

For our purposes, the discourse structure was partly determined on the basis of the dialogue act annotation. The targets of each dialogue act were systematically identified (including "backward-looking" acts, such as acknowledgment or answer, and others) and discourse relation (such as Elaboration, Background, Narration) between the act and its target were annotated. The resulting discourse structure provided a hierarchy of sub-dialogues, including cases of discourse popping (attachment of a new constituent higher in the hierarchy than the previous utterance). *Openings* were identified with the following clues: **(i) discourse pop-ups, (ii) activity changes (from landmark management to instruction, and the opposite), (iii) new landmarks, (iv) clarification and feedback requests.** Additional clues were provided by some discourse markers such as *donc* and *alors*. The clues for *closings* relied more directly on the dialogue act annotation and included: **(i) double acknowledgments, (ii) acknowledgment following answers, (iii) answers to feedback request, (iv) strong acceptances** and specific discourse markers such as *voilà* and *bon* (see [15] for more details).

In the resulting annotation, the number of openings was significantly higher than the number of closings which can be surprising (75 vs 52). In fact, in many cases, it was not possible to identify a closing using the rules summarized above because some closings seem to be implicitly determined by the presence of a new opening which is not a sub-structure.

Rise was found to be correlated with the open/close distinction ($p<0.001$), rises being associated with openings and rise absences with closings. The corpus size was not sufficient to analyze the link between intonation and speaker roles, but there was no apparent bias due to specificities of the speakers. More work is needed to investigate the 'local roles' of speakers (associated with competence with respect to the current topic), which seems to be closely related to Kowtko dialogue game definitions [13].

4 Conclusion and Future work

The results of the study of Post's Map Task corpus showed that with respect to dialog acts, a positive correlation can be found between rises and (polar) questions, thus confirming earlier observations in the literature, and between rises and prescriptions using landmarks. On the other hand, answers to questions

were more likely to appear without a rise. Mirroring similar results for English, we found that rises were significantly correlated to topic openings and rise absences with closings. The rise/openings correlation was stronger than the correlation rise/questions, suggesting that the first association was not simply due to the question effect of introducing new discourse topics. Finally, speaker variation was observed, especially in the use of rises on acknowledgments which could, however, be due to their distinct dialogue roles (instruction giver vs. instruction follower), given that one of the dialogues was substantially shorter than the other.

Although the results of the Map Task corpus study are promising, they need to be tested on a corpus of a larger size and containing free conversations. A study of the Caelen corpus of tourist office dialogues is currently in progress. In order to describe the role of intonation in discourse in more detail, it may also turn out to be necessary to use a more fine-grained intonational transcription; alternatives to the MOMEL-based INTSINT alphabet are being investigated.

References

- [1] M. Beckman and G. Ayers. Guidelines for ToBI labeling, version 3.0, 1997. Unpublished manuscript.
- [2] M. Bessac and G. Caelen-Haumont. Analyses pragmatiques, prosodiques et lexicales d'un corpus de dialogue oral, homme-homme. In *Proceedings of the 3rd International Conference on Statistical Analysis of Textual Data*, pages 363–370, Rome, 1995.
- [3] C. Beyssade, E. Delais-Roussarie, J. Doetjes, J.-M. Marandin, and A. Rialland. Prosodic, syntactic and pragmatic aspects of information structure. In F. Corblin and H. de Swart, editors, *Handbook of French Semantics*, pages 455–475. CSLI Publications, 2004.
- [4] R. Espesser. Un système de détection du voisement et de F0. In *TIPA8*, pages 241–261, 1982.
- [5] C. Féry. Gradient prosodic correlates of phrasing in French. *Nouveaux départs en phonologie*, to appear.
- [6] J. Fletcher, R. Wales, L. Stirling, and I. Mushin. A dialogue act analysis of rises in australian english map task dialogues. In *Proceedings of Speech and Prosody '02*, Aix-en-Provence, 2002.
- [7] I. Fónagy and E. Bérard. Questions totales simples et implicatives en français parisien. In A. Grundstrom and P. Léon, editors, *Interrogation et Intonation*, number 8, pages 53–98. Didier, Paris, 1973.
- [8] A. Grundstrom. L'intonation des questions en français standard. In A. Grundstrom and P. Léon, editors, *Interrogation et Intonation*, number 8, pages 19–51. Didier, Paris, 1973.
- [9] C. Gunlogson. *True to Form: Rising and Falling Declaratives as Questions in English*. PhD thesis, UCSC, 2001.
- [10] D. Hirst and A. Di Christo, editors. *Intonation systems: a survey of twenty languages*. Cambridge University Press., 1998.
- [11] D. Hirst and R. Espesser. Automatic modelling of fundamental frequency using a quadratic spline function. *Travaux de l'Institut de Phonétique d'Aix*, 15:71–85, 1993.
- [12] S-A. Jun and C. Fougeron. The realizations of the accentual phrase in French intonation. *Probus*, 14:147–172, 2002.
- [13] J. Kowtko. *The function of intonation in task-oriented dialogues*. PhD thesis, University of Edimburgh, 1996.
- [14] J.A. Louw and E. Barnard. Automatic intonation modeling with INTSINT. In *Proceedings of the Pattern Recognition Association of South Africa*, pages 107–111, 2004.
- [15] Philippe Muller and Laurent Prévot. An empirical study of acknowledgement structures. In *Proceedings of Diabrock, 7th workshop on semantics and pragmatics of dialogue, Saarbrücken*, 2003.
- [16] B. Post. *Tonal and Phrasal Structures in French Intonation*. PhD thesis, University of Nijmegen, 2000.