

The Discourse Function of Final Rises in French dialogues

Marie Šafářová

ILLC

University of Amsterdam
m.safarova@uva.nl

Philippe Muller

IRIT

Université Paul Sabatier
muller@irit.fr

Laurent Prévot

Laboratory for Applied Ontology
ISTC-CNR, Trento
prevot@loa-cnr.it

Abstract

We report the results of an empirical study which aims to describe the discourse function of rises at right edge intonation boundaries in French. A Map-Task corpus containing two dialogues was annotated for IP boundaries and pitch transition points with the INTSINT international alphabet. The transcripts of the dialogues were labeled for dialogue structure and dialogue acts, using form and function tags. The relation between rises at IP boundaries with types of dialogue acts and topic shifts was statistically evaluated. As expected, the results show a positive correlation between rises and polar questions and between rises and discourse topic openings. Interestingly, the second correlation was stronger than the first, suggesting that the association of rises with topic openings is not simply due to the effect of questions as introducing new topics.

1 Introduction

In this paper, we report the results of an empirical study which aims to describe the discourse meaning of rises at right edge into-

nation 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 (Grundstrom, 1973; Fónagy and Bérard, 1973) 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 (Kowtko, 1996), for Glasgow English, or (Fletcher et al., 2002) for Australian English). An initial study has been performed on Post's Map Task corpus (Post, 2000) 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 dialogue acts and discourse structure taxonomies. The tested methodology is used

for the study of the Caelen corpus (Bessac and Caelen-Haumont, 1995), 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 dialogue acts and dialogue 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 (Post, 2000), 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 (Grundstrom, 1973; Post, 2000; Gunlogson, 2001). Leaving aside the option of direct perceptual distinctions, which may be highly unreliable, 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 (necessarily a theoretical construct). A phonetic description (e.g., using points of maximum and minimum pitch) has the advantage that its result is annotator-independent; it can be done automatically and its quality depends solely 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 speak-

ers). 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) is a language-independent intonation transcription system developed in Aix-en-Provence (Hirst and Christo, 1998). 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 (Hirst and Espesser, 1993) provides an automatic stylization of the F0 contour, detected from the acoustic signal with the comb algorithm (Espesser, 1982) (see also (Louw and Barnard, 2004)). INTSINT covers both absolute prosodic events (**T** – Top; **M** – Mid; **B** – Bottom) and relative ones (**H** – Higher; **S** – Same; **L** – Lower; **U** – Up-step; **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 ends of utterances. All boundary tones associated with right edges of intonational phrases are assumed to be meaningful (Beysade et al., 2004); moreover, important intonational events (e.g., encoding the difference between questions and non-questions) may not be aligned with right utterance edges. 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 (Jun and Fougeron, 2002), (Post, 2000), (Féry, to appear), among others). Nevertheless, they also appear to be related to information structure articulation and (Beysade et al., 2004) define them only as a reflection of the information structure of an utterance. It is also normally assumed that there is some correlation between prosodic phrasing and syntactic boundaries. Taking these observations into account, prior to the annotation process, the following rules were proposed to serve as a guidance to the annotators, together with their perceptual impression of the speech signal:

Intonation Boundary [Def.]

- Every completed turn boundary is a right edge IP boundary.
- 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.
- An IP boundary often coincides with a major syntactic boundary (e.g., a finite clause boundary).
- An information structure constituent (topic, focus) can be followed by an IP boundary.

The three authors of the study (two native speakers of French, one non-native) served as annotators of the corpus. The results were evaluated for inter-annotator agreement using the *kappa*-statistics 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. 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 '*par contre*' and the utterance final '*quoi*' raised a similar problem. The annotators also disagreed at hesitation points (often filled with '*euh*') 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 intonational phrase boundaries proposed by the three annotators were merged together in the final annotation.

Because the manual annotation of IP boundaries was judged to be rather time-consuming and thus unsuitable for the subsequent large corpus study, its results 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 $\frac{2}{3}$ and $\frac{3}{4}$ of the manually assigned intonational phrase boundaries for the two dialogues, respectively;² 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 re-

¹These have been found to be problematic also in MAE-ToBI (the American English standard for prosody labelling, viz (Beckman and Ayers, 1997)) for the same reason.

²The fact that the results were better for the second dialogue than for the first is probably due to the fact that the manual annotations were better in the second dialogue due to improved annotating skills of the labelers.

place 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. (Grundstrom, 1973):26 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. The ambiguity between questions and assertions is one of the reasons why finding an objective procedure for identifying questions in a corpus is problematic. (Fónagy and Bérard, 1973) propose a simple solution by considering as questions all utterances that received a *oui/non* reply. This definition is too strong, however, given that many assertions receive an acknowledgment synonymous with the *oui*-reply. It is also too weak because 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.³ The defini-

³The procedure was tested for inter-rater agreement and against native speaker judgments, for results and discussion,

tion 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. 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 instance, in example 1 (taken from Post's Map Task corpus), the declarative (G_{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 was designed to avoid cases of overgeneralization (though it is sometimes difficult to distinguish between the *assertion – acknowledgment* and *question – answer* sequences).

- (1) (G_{103}) *est-ce que tu as IP tu as le profond étang H IP*⁴
[do you have the deep pond]
(F_{104}) *oui, H sur la gauche IP*
[yes, on the left]
(G_{105}) *oui, tout à gauche. IP*
[yes, all the way on the left]

see (Šafářová, in prep).

⁴Our examples will be presented with numbered utterances annotated with the speaker information (Giver and Follower). **H** here stands for 'High' in the INTSINT international alphabet, **IP** stands for 'intonation phrase'.

(G₁₀₆) et tu as la grande plaine H IP
[and you have the big plain]

(F₁₀₇) non IP [no]

Apart from questions, the corpus was also annotated for other types of dialogue acts, partly based on an existing annotation scheme (Prévot, 2004) 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 dialogue act segmenting and labeling task, the final annotation was mainly based on a post-hoc discussion (rather than a majority decision). 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 if an utterance was an alternative bipolar question with an ellipsis of the second constituent, or if the utterance-final *ou* connective merely served to indicate speaker's uncertainty (as in "*et à beaucoup de centimètre du pic ou...?*" - "and at many centimeters from the peak or...?"), 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 (Kowtko, 1996)) 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, the *wh* constituent may be missing, as in "*et alors l'hôtel par rapport aux torrents et l'océan?*" - "and so the hotel with respect to the torrents and the ocean?";
- (v) it appeared desirable to classify also utterances like *je ne sais pas si tu le vois* - "I don't

know if you can see it" as questions. The presence/absence of a rise against the main act categories (acknowledge, instruct, inform, question, answer, with or without new landmark introduction) was statistically evaluated. We considered as instances of IP-final rises those intonational events that were aligned with the manually assigned right-edge IP boundaries and labeled as **T**, **H** or **U** in INTSINT.⁵ There was no convincing correlation between the act labels and the presence of a rise, except for instruction (using a landmark) associated with a rise (χ^2 , $p=0.006$), 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 aim of the discourse structure annotation task is to test for a possible correlation between discourse opening/closing and rises. At this level of discourse organization, however, two organizational principles are competing: game and topic structures. Game initiations and topic openings are often realized through the same move. These moves are utterances whose discourse functions are primarily "forward-looking" rather than "backward-looking" in the DAMSL terminology (Core and Allen, 1997). Despite this vicinity new games do not necessarily bring new topics into discussion (e.g., simple checks or verification questions), nor do topic shifts always initiate a new game (e.g. a long speech-turn introducing a complex discourse structure made of several discourse topics). One possible explanation for these discrepancies is the very purpose of the dialogue game account which is purely to describe dialogues. Topics, on the other hand, concern any kind of discourse and in particular monologue stretches which are not interesting for dialogue games. The clues for recognizing these structures are also very

⁵In theory, the upstepped (U) rise should be a rather small pitch movement upwards but in practice, the U-rises were often as big as the H- or T-rises, which is why we included them in the evaluation.

different: While topic structures may require a deep semantic understanding of the conversation, game structures might be determined more directly from move types and move type sequences.

The notions of dialogue games and discourse topics have been discussed at length in the past and many proposals already exist in the literature. With respect to dialogue game definition, we opted for the MAPTASK schema detailed in (Kowtko, 1996). In this framework, dialogue games are sequences of potential moves initiated by a particular move (instruct, check, queries, explain, align)⁶. Regarding the notion of discourse topic, we rely on the account of (Asher, 2004), who recalls discourse topics can be either explicit and introduced by a specific utterance, or implicit and inferred from discourse content. In practice, it is thus difficult to identify topic openings in a systematic way. Finally, both game and topic structures admit sub-structures like embedded games and sub-topics.

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 discourse relations (such as Elaboration, Background, Narration) 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).

Game and topic openings and closings should be derived without much difficulties from discourse structure. However, for succeeding in this task we need a rich discourse structure of our dialogues. The discourse relations involved in direction-giving dialogues have been studied in details in (Prévoit, 2004) and we present a rough sum-up below.

⁶In DAMSL (Core and Allen, 1997) these functions are classified under the forward-looking function.

Successive *instructions* (e.g 2: $G_{125-126}$) are related by the coherence relation of *Narration* and constitute a topic.⁷ Therefore, a sequence of instructions without landmark explanation constitutes only one discourse topic. Similarly in the dialogue game framework, an INSTRUCTING-game is possibly made of sequences of acknowledged instructions.

Landmark introductions (e.g 1: G_{103}) are related to *background* and are explicit new topics that can be elaborated with landmark descriptions. Similarly, in the MAPTASK they corresponds to the EXPLAINING-game which often appears embedded in the INSTRUCTING game.

We treat landmark descriptions and localizations (e.g 2: G_{130}) as *elaborations* of the constituent in which the landmark has been introduced. This could have been tackled in MAPTASK dialogue definitions by allowing the EXPLAINING-game to be recursive just like the INSTRUCTING ones.

Openings were identified with the following clues: **(i) discourse pop-ups, (ii) 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) specific discourse markers** such as *voilà* and *bon* (see (Prévoit, 2004) for more details).

The examples 2 and 3 illustrate the opening/closing annotation. In the bracketed text are given some of the tags we used: SURFACE-FORM, FUNCTION, DISCOURSE-STRUCTURE, DISCOURSE-TOPICS. Surface forms included assertions (ASS), yes-no questions (QYN), wh-questions (QWH), alternative questions (QAL) and indeterminate

⁷We do not develop this point here but see (Asher, 2004) for more details on the nature of discourse relation and their consequences for discourse topic.

forms (IND). Functions included instructions (PAR, PSR), landmark introduction (IR), question-answer pair (QAP) and acknowledgment (ACK). Discourse structure tags give information about discourse relations and targets. Finally discourse topic is added in case of an opening or a closing and discourse topics are numbered.

- (2) (F_{124}) euh tu fais une boucle autour du deuxième petit pin [*err you do a loop around the second small pine tree*]
[ASS PAR NARR-119 OPEN-28]
(G_{125}) c'est à dire que tu passes par derrière [*so that means that you pass behind*]
[ASS PSR ELA-124]
(G_{126}) et tu reviens devant.
[*and you come back in front*]
[ASS PSR NARR-125,ELA-124]
(F_{127}) mm [*mm*] [ASS ACK ACK-126]
(G_{128}) est-ce que tu as une colline [*do you have a hill*]
[QYN IR PELAQ-0 OPEN-29]
(F_{129}) non, j'ai pas de colline [*no I don't have a hill*]
[ASS QAP QAP-128]
(G_{130}) à côté du petit pin [*near the small pine tree*]
[IND DR ELAQ-128]
(F_{131}) j'ai rien à côté du petit pin [*I have nothing near the small pine tree*]
[ASS QAP QAP-130]
- (3) (G_{103}) est-ce que tu as tu as le profond etang [*do you have the deep pond*]
[QYN IR PELAQ-0 OPEN-25]
(F_{104}) oui, sur la gauche [*yes on the left*]
[ASS QAP QAP-103]
(G_{105}) oui, tout à gauche [*yes completely on the left*]
[ASS ACK ACK-104 CLOSE-25]

In the resulting annotation, the number of openings was significantly higher than the

number of closings (75 vs 52). It was sometimes difficult to identify closings by using the rules summarized above because some of them are implicit.

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 (Kowtko, 1996).

It became clear that once the discourse structure is established, openings and closings are easier to determine and this can be done in a general way. However, building discourse structure was possible only with the input of a careful analysis of direction giving dialogues. Though we would like to abstract as much as possible from dialogue genre specificities, it did not appear to be feasible in practice. The reason is that dialogue game rules are usually defined for a particular dialogue game and discourse relation inference rules are established for a given discourse genre (e.g narrative, argumentative).

4 Conclusion and Future work

The results of the study of Post's Map Task corpus showed that with respect to dialogue 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.

Acknowledgments

The authors would like to thank the Elisabeth Delais-Roussarie, Jean-Marie Marandin and Claire Beyssade for their comments on earlier versions of this manuscript, Brechtje Post for allowing us to use her data, and the DI-ALOR reviewers for their comments and suggestions.

References

- N. Asher. 2004. Discourse topic. *Theoretical Linguistics*, (30):161–201.
- M. Beckman and G. Ayers. 1997. Guidelines for ToBI labeling, version 3.0. Technical report, The Ohio State University.
- M. Bessac and G. Caelen-Haumont. 1995. 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.
- C. Beyssade, E. Delais-Roussarie, J. Doetjes, J.-M. Marandin, and A. Rialland. 2004. Prosodic, syntactic and pragmatic aspects of information structure. an introduction. In F. Corblin and H. de Swart, editors, *Handbook of French Semantics*, pages 455–475. CSLI Publications.
- M. Core and J. Allen. 1997. Coding dialogs with the DAMSL annotation scheme. In *Working Notes of the AAAI Fall Symposium on Communicative actions in Humans and Machines*, pages 28–35, Cambridge, MA.
- R. Espesser. 1982. Un système de détection du voisement et de f0. In *TIPA8*, pages 241–261.
- C. Féry. to appear. Gradient prosodic correlates of phrasing in French. *Nouveaux départs en phonologie*.
- J. Fletcher, R. Wales, L. Stirling, and I. Mushin. 2002. A dialogue act analysis of rises in Australian English Map Task dialogues. In *Proceedings of Speech and Prosody '02*, Aix-en-Provence.
- I. Fónagy and E. Bérard. 1973. 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.
- A. Grundstrom. 1973. L'intonation des questions en Français Standard. In A. Grundstrom and P. Léon, editors, *Interrogation et Intonation*, number 8, pages 19–51. Didier, Paris.
- C. Gunlogson. 2001. *True to Form: Rising and Falling Declaratives as Questions in English*. Ph.D. thesis, UCSC.
- D. Hirst and A. Di Christo, editors. 1998. *Intonation systems: a survey of twenty languages*. Cambridge University Press.
- D. Hirst and R. Espesser. 1993. Automatic modelling of fundamental frequency using a quadratic spline function. *Travaux de l'Institut de Phonétique d'Aix*, 15:71–85.
- S.-A. Jun and C. Fougeron. 2002. The realizations of the accentual phrase in French intonation. *Probus*, 14:147–172.
- J. Kowtko. 1996. *The function of intonation in task-oriented dialogues*. Ph.D. thesis, University of Edinburgh.
- J.A. Louw and E. Barnard. 2004. Automatic intonation modeling with INTSINT. In *Proceedings of the Pattern Recognition Association of South Africa*, pages 107–111.
- B. Post. 2000. *Tonal and Phrasal Structures in French Intonation*. Ph.D. thesis, University of Nijmegen.
- L. Prévot. 2004. *Structure sémantique et pragmatique pour la modélisation de la cohérence dans des dialogues finalisés*. Ph.D. thesis, Université Paul Sabatier.
- M. Šafářová. in prep. *Rises and Falls. Studies in the semantics and pragmatics of intonation*. Ph.D. thesis, University of Amsterdam.