

Rising intonation in spontaneous French: how well can continuation statements and polar questions be distinguished?

Emma Valtersson¹, Francisco Torreira¹

¹Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Emma.Valtersson@mpi.nl, Francisco.Torreira@mpi.nl

Abstract

This study investigates whether a clear distinction can be made between the prosody of continuation statements and polar questions in conversational French, which are both typically produced with final rising intonation. We show that the two utterance types can be distinguished over chance level by several pitch, duration, and intensity cues. However, given the substantial amount of phonetic overlap and the nature of the observed differences between the two utterance types (i.e. overall F0 scaling, final intensity drop and degree of final lengthening), we propose that variability in the phonetic detail of intonation rises in French is due to the effects of interactional factors (e.g. turn-taking context, type of speech act) rather than to the existence of two distinct rising intonation contour types in this language.

Index Terms: rising intonation, question intonation, continuation intonation, polar questions, French

1. Introduction

One of the main challenges to the study of human communication is the lack of a one-to-one mapping between pragmatic function and linguistic form. In this study, we tackle this issue by investigating the prosody of continuation statements and polar questions in French, two utterance types that often exhibit a similar form (i.e. SVO word order and final rising intonation) in spite of their markedly different functions (e.g. keeping vs. yielding the floor, conveying vs. requesting information). Despite the syntactic and intonational similarities of these two utterance types, it is possible that their prosodic characteristics can distinguish them robustly, and that participants in a conversation do not need to rely solely on context in order to interpret their pragmatic function. Our aim in this study is to examine whether a clear distinction can be made between the prosody of the two kinds of utterances in a corpus of spontaneous French.

Several claims regarding the prosodic distinction between rising continuation statements and rising polar questions have been made in the French intonation literature. Based on introspection, [1] claimed that these two utterance types both have final rising pitch, but also that they are differentiated by the scaling and shape of the final pitch rise, with questions reaching a higher pitch maximum and exhibiting a more concave final pitch contour than continuation statements. Two later studies based on spontaneous and read speech [2, 3] supported this distinction in the scaling of the final pitch maximum, but did not find evidence for a distinction in terms of contour shape. Moreover, [3] found a different intensity profile and a shorter duration of the final vowel for continuation statements than for questions.

However, [4], based on the variability observed for both continuation statements and polar questions, claimed that a single underlying rising intonation pattern is compatible with

both functions. The possibility that rising polar questions and continuation statements share the same phonological pattern has also been reflected in more recent accounts of French intonation within the Autosegmental-Metrical framework. It has been suggested that the final rise found in both major continuations and polar questions coincides with the end of the highest level of phrasing (i.e. the Intonation Phrase) [5], and that they both consist of the same tonal elements (H*H%) [6].

In summary, there is disagreement in the French intonation literature about the exact prosodic characteristics of rising continuation statements and rising polar questions, and about their possible phonological distinction. In this study, we assess the extent to which these two kinds of utterances can be distinguished phonetically on the basis of their prosody using data from a corpus of French spontaneous conversation.

2. Method

Our data come from the Nijmegen Corpus of Casual French [7], which consists of 23 casual conversations among groups of friends recorded with head-mounted microphones in a sound-attenuated room. We used the initial dyadic part of each conversation, which had a duration ranging from 4 to 40 minutes, and provided a total of around seven and a half hours of spontaneous conversation. Twelve dyads were composed by female speakers, and eleven by male speakers. All speakers originated from Central or Northern France, and were mostly university students, with ages from 18 to 27.

The two authors independently inspected all the conversations in the data and identified cases of rising intonation using Elan and Praat software [8, 9]. We only considered utterances with rising intonation, which were immediately followed by a pause. The reasons for this were to avoid the effects of tonal and segmental coarticulation in the utterance-final syllable, and to make sure that the final rises in our data were all produced in a comparable context in terms of prosodic phrasing. It should also be noted that we did not mark cases with final plateaus, which we observed in tag questions [5] and enumerations [10]. After finishing a first annotation pass, we created a final dataset including all cases of rising pre-pausal intonation that were identified by the two annotators ($n = 320$). Then, for each utterance, we annotated whether it was a question or a continuation statement in the context of the conversation. Questions typically involved knowledge within the listener's epistemic domain [11] (e.g. 'You went to Paris yesterday?') and were followed by a turn transition and an answer, whereas continuation statements involved knowledge within the speaker's epistemic domain (e.g. 'I went to Paris yesterday') and were usually not followed by a turn-transition. No ambiguous cases were observed when the conversational context was taken into account. In total, we included a total of 126 questions and 190 continuation statements in our dataset.

Several acoustic measurements were performed with Praat [9] in the final syllables of the utterances, where differences in

the final prosody of continuation statements and polar questions were likely to be observed according to previous studies [1,2,3]. We measured the minimum and maximum pitch values (semitones re 100 Hz) in the final vowel of the utterance, which was always present regardless of the structure of the final syllable (e.g. open or closed, with or without an onset). These pitch minimum and maximum values corresponded to the start and end of the intonation rise throughout the final vowel and can be used to investigate differences in scaling between the intonation rises of continuation statements and polar questions. As an estimate of the pitch register used prior to the rise, we measured the median pitch in the interval from the start of the utterance up to the start of the final syllable. In order to control for speaker-based variation in pitch, we normalised pitch values with reference to the median value calculated in an excerpt of ten minutes for every speaker. The shape of the pitch rise has been proposed as being important [1] as well as irrelevant [2, 3] for distinguishing between continuation statements and polar questions in French. In order to examine this feature, we took pitch measures every ten ms throughout the final vowel, and approximated these values with a quadratic equation using least-squares linear regression as in [12, 13, 14]. For each pitch rise, we fitted a linear model $y = a + bx + cx^2$, where a represents the initial pitch value, b the initial slope and c the curvature of the rise. Convex pitch rises are fitted by models with negative quadratic coefficients, whereas concave pitch rises are fitted by models with positive quadratic coefficients. In addition to the pitch measurements above, we calculated the final drop in intensity between the intensity peaks in the last two vowels. Furthermore, we also measured the duration of the final vowel, which was always present regardless of the structure of the final syllable (e.g. open or closed, with or without an onset). In order to control for variability in final vowel duration due to differences in speech rate across tokens, we also calculated the speech rate over the last intonational phrase up to the final syllable, so that it could be included as a covariate in our statistical analyses.

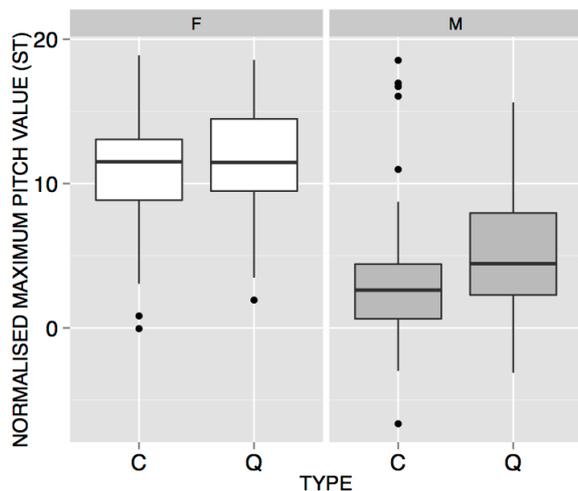


Figure 1: *Speaker-normalised maximum pitch in final vowel (semitones, re 100 Hz) for the two kinds of utterances (C = Continuation statement, Q = Question) for each gender (F = Female, M = Male).*

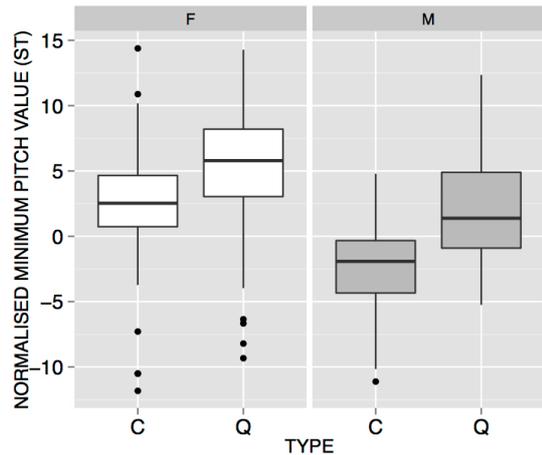


Figure 2: *Speaker-normalised minimum pitch in final vowel (semitones, re 100 Hz) for the two kinds of utterances (C = Continuation statement, Q = Question) for each gender (F = Female, M = Male).*

3. Results

We investigate whether the final prosody of continuation statements and rising polar questions can be clearly distinguished in spontaneous French. Using regression modelling, we first compare the pitch, duration and intensity characteristics between the two kinds of utterances. We focus on the final syllables of the utterances, where relevant differences are expected to occur according to previous literature. Following this, we evaluate the degree of overlap and separation between the two kinds of utterances using a leave-one-out cross-validation procedure.

3.1. Phonetic comparison

Figure 1 shows boxplots of maximum pitch values (i.e. the end of the pitch rise) as a function of utterance type, separated by speaker gender. This figure shows that, excepting a small number of outliers, questions tend to end in higher pitch compared to continuations for males, but not for females. It also shows a clear gender difference, with females having higher values in general. These differences were confirmed by a statistically significant interaction between utterance type and speaker gender in a regression model with maximum pitch as the dependent variable ($\beta = 1.7$, $t = 2.24$, $p < .05$).

Figure 2 shows boxplots of minimum pitch values in the final intonation rise as a function of utterance type, separated by speaker gender. As for maximum pitch, females appeared to have higher values than males in general. Regarding utterance type, questions appear to exhibit a higher minimum pitch value for both females and males. These differences were confirmed by a regression model with minimum pitch as the dependent variable, and utterance type and gender as predictors (utterance type: $\beta = 3.65$, $t = 7.52$, $p < .0001$; gender: $\beta = -4.47$, $t = -9.36$, $p < .0001$). Similarly to minimum pitch, the pitch median up to the final syllable of the utterance also exhibited a higher value for questions. A regression model with the same predictors as for minimum pitch supported this observation (utterance type: $\beta = 2.05$, $t = 6.50$, $p < .0001$; gender: $\beta = -4.39$, $t = -14.12$, $p < .0001$).

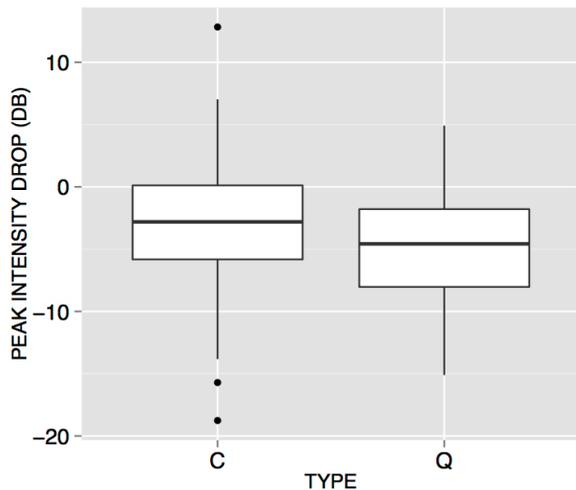


Figure 3: Intensity drop in final vowel in seconds (dB) for the two kinds of utterances (C = Continuation statement, Q = Question).

In order to investigate the shape of the rise (i.e. convex vs. concave) we used quadratic equations modelling its trajectory as explained in the Methods section. We fitted a regression model with the quadratic coefficients of these equations, which capture the degree of curvature of pitch rises, as the dependent variable, and utterance type as predictor. The statistical analysis revealed that polar questions and continuation statements did not differ statistically in the curvature of their rises ($p = .33$). Inspection of the data revealed that the two utterance types could exhibit negative and positive curvature values, indicating that they could both have convex or concave final pitch rises.

Figure 3 displays boxplots of peak intensity in final syllable relative to the peak intensity in the previous syllable. The figure shows that questions tend to exhibit a lower final intensity peak compared to continuation statements. A regression analysis confirmed this difference ($\beta = -1.89$, $t = -3.71$, $p < .001$).

Finally, we investigated if the duration of the final vowel varies with utterance type. Figure 4 shows boxplots of final vowel duration for the two utterance types. We can see in this figure that questions tend to be produced with a slightly shorter final vowel in comparison to continuation statements. We fitted a regression model with duration as the response, utterance type as the main predictor, and also speech rate as a covariate, since we wanted to control for duration variability related to this factor. This model yielded a statistical difference of roughly 10 ms between continuation statements and polar questions ($\beta = -0.012$, $t = -2.36$, $p < 0.05$).

In summary, we have found differences between continuation statements and polar questions in terms of several pitch, duration and intensity measures. However, our data also indicate that these two utterance types overlap considerably in these prosodic features, in particular in the case of maximum pitch for females, and intensity and duration for both genders.

The question arises therefore how well the two utterance types can be distinguished on the basis of the several relevant phonetic features in the comparison. We address this question in the following subsection.

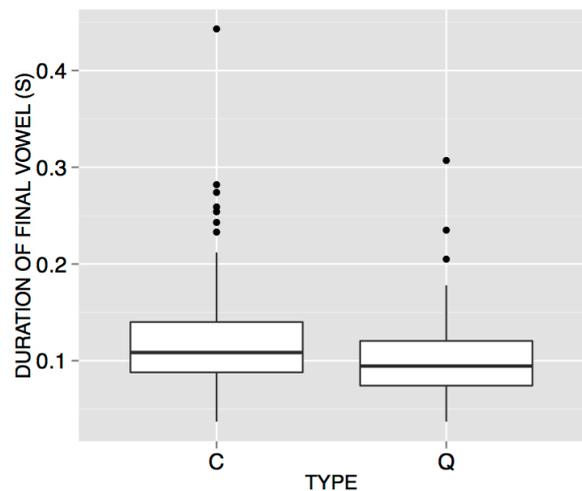


Figure 4: Duration of the final vowel in seconds (s) for the two kinds of utterances (C = Continuation statement, Q = Question).

3.2. Cross-validation procedure

In order to estimate the degree to which continuation statements and polar questions can be separated on the basis of the cues identified in the previous subsection, we performed a leave-one-out cross-validation of the phonetic cues that appeared to be relevant with our data. This procedure simulated predicting new unseen data using our dataset in the following way: utterance type was predicted for each token in our dataset by means of logistic regression models trained on the rest of the dataset including several features, yielding a percentage of correct classifications for each different model. We used regression models that included several combinations distinguishing the two utterance types in the previous subsection. Models with pitch cues as predictors also included speaker gender as a covariate, whereas models including final vowel duration also included speech rate.

Our automatic classifications yielded different accuracy levels dependent on the model used. The minimum value of the pitch rise offered an accuracy of 72%, close to that of a model comprising all features (76%). The accuracy afforded by maximum pitch, on the other hand, was significantly lower (60%). Combining the two final pitch cues (i.e. minimum and maximum pitch) allowed for a marginal increase in accuracy over the model with minimum pitch alone (73%). Interestingly, the pitch median before the last syllable, which we used as an estimate of the pitch register of the utterance prior to the final syllable, was almost as good a predictor of utterance type as the final pitch cues, with an accuracy of 70%. Regarding non-pitch cues, intensity and duration both provided moderate improvements over chance level (61% for intensity, 67% for duration, and 67% for a model combining duration and intensity).

These results show that not all phonetic features are equally useful for distinguishing between the two utterance types. In particular, the minimum pitch value at the beginning of the pitch rise was the best cue for the contrast between continuation statements and polar questions, performing almost as well as the full model with all cues, whereas the final pitch maximum only led to a moderate gain over chance level.

4. Discussion and conclusion

This study has investigated whether a clear distinction can be made between the prosody of French continuation statements and polar questions, which both exhibit rising intonation patterns previously claimed to be distinct or similar by different authors [1, 4]. Our comparative analysis in Section 3.1 has shown phonetic differences between the prosody of the two kinds of utterances in terms of several prosodic features. However, a great amount of phonetic overlap between the phonetic realizations of the two utterance types has been observed. This has been evidenced in Section 3.2 by the fact that roughly a quarter of the data was wrongly classified by a logistic regression model containing all relevant prosodic features identified in Section 3.1.

Regarding the phonetic differences, we have observed that the maximum pitch value in the final intonation rise tended to be higher for questions than for continuation statements, but only for male speakers. A more consistent difference was observed between the two utterance types in the minimum pitch value at the beginning of the rise, and in the pitch register prior to the final syllable of the utterance. As for the shape of the rise, we did not observe any differences between continuations and questions. On the other hand, the duration and peak drop intensity of the final vowel showed differences between the two utterance types. Questions tended to have a shorter final vowel and end in a lower peak intensity target than continuations.

Our findings therefore only partially agree with those from past studies. The higher maximum pitch in the rise for questions is in line with the findings of [1, 3], and the more pronounced final drop in intensity for questions agrees with the observations of [3]. Regarding the shape of the final pitch rise, which had been suggested as a relevant feature by [1], we did not observe any consistent differences between the two utterance types, in line with [2, 3]. In our data, final intonation rises in both continuation statements and polar questions could adopt slightly concave or convex shape.

On the other hand, the minimum pitch value in the intonation rise, a feature judged to be irrelevant by [1] and not investigated in detail by [3], clearly offered the best cue for distinguishing between continuation statements and polar questions. Our findings regarding the duration of the final vowel, which was the second-best cue to the contrast between continuations and questions, were also opposed to the previous literature [3]. In our study, polar questions exhibited shorter, not longer, final vowels than continuation statements.

These discrepancies between our findings and those of previous studies may be due to the fact that studies in the past were mostly based on introspection and read speech data, whereas ours used utterances extracted from spontaneous conversations more directly affected by the interaction between speakers. In our study, for instance, the fact that polar questions tended to have less final lengthening than continuation statements may be due to the fact that questions typically yield the floor to the interlocutor, and that their final part usually marks the end of the speaker's turn. This would be in line with observations from studies on English dialogue, which have shown that turn-final utterances tend to exhibit less final lengthening and a more marked final drop in intensity than turn-medial utterances [15, 16]. Since these differences in duration and intensity were observed in utterances with different intonation patterns, it can be concluded that they were not directly related to differences in

intonation contour choice, but rather to differences in turn-taking actions (i.e. turn-yielding vs. turn-keeping). It is therefore likely that the observed differences in duration and intensity between continuation statements and polar questions in our French data are due to the different turn-taking contexts in which these utterances tend to occur.

Interactional factors may also be the reason why the minimum pitch value at the start of the rise and the pitch register at which the utterance was spoken before the final syllable, features neglected in previous research, provide the best cues in distinguishing continuation statements from polar questions in our corpus. It has often been proposed that questioning utterances in general, not only those with final rising intonation tend to exhibit higher pitch registers compared to assertive utterances [e.g. 17, 18, 19 among others]. Our finding that the minimum pitch at the start of the final rise and the pitch median before the final syllable were consistently different between continuation statements and polar questions is most likely due to the fact that the pitch register throughout the utterances is generally higher in questions than in continuations.

Our findings therefore appear to be in line with previous proposals [4, 6] that a single rising intonation pattern (e.g. H*H%) is compatible with both continuation statements and polar questions in French. We have observed an important amount of overlap in our data, showing that a clear phonetic distinction cannot be drawn between the prosody of rising continuation statements and polar questions in French. Moreover, we have argued that the phonetic differences observed between the two utterance types (i.e. overall F0 scaling, final intensity drop, degree of final lengthening), are likely to be due to the effects of interactional factors (i.e. turn-taking context). For these reasons, we conclude that continuation statements and polar questions in French both make use of the same rising contour type, and that this contour type is subject to contextual variation which may help cue different speech acts during conversation (signalling that one's turn is not complete vs. asking a question).

5. Acknowledgements

The contribution of the second author was made possible thanks to the financial support of the Language and Cognition Department at the Max Planck Institute for Psycholinguistics, Max-Planck Gesellschaft, and a European Research Council's Advanced Grant (269484 "INTERACT") to Steve Levinson.

6. References

- [1] Delattre, P., "Les dix intonations de base du français", *French Review*, 40(1):1-14, 1966.
- [2] Grundstrom, A., "L'intonation des questions en français standard", in *Interrogation and Intonation*, *Studia Phonetica*, 8, 19-49, 1973.
- [3] Rossi, M., Di Cristo, A., Hirst, D., Martin, P. and Nishinuma, Y., "L'intonation: de l'acoustique à la sémantique.", 1981.
- [4] Di Cristo, A., "Intonation in French", in D. J. Hirst & A. Di Cristo [Ed], *Intonation Systems. A Survey of Twenty Languages*. 195-218, Cambridge University Press, 1998.
- [5] Jun, S. A. and Fougeron, C., "Realizations of accentual phrase in French intonation", *Probus*, 14(1):147-172. 2002.
- [6] Post, B., "French tonal structures". *Speech Prosody*, 2002.
- [7] Torreira, F., Adda-Decker, M. and Ernestus, M., "The Nijmegen corpus of casual French", *Speech Communication*, 52(3):201-212. 2010.
- [8] Brugman, H. and Russel, A., "Annotating Multimedia/ Multimodal resources with ELAN. Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation. 2004.
- [9] Boersma, P. and Weenink, D., "Praat: doing phonetics by computer" (Version 5.3. 51). <http://www.praat.org>. 2013.
- [10] Portes, C., Bertrand, R. and Espesser, R. "Contribution to a grammar of intonation in French. Form and function of three rising patterns", *Nouveaux cahiers de linguistique française*, 28:155-162. 2007.
- [11] Heritage, J., "Epistemics in action: Action formation and territories of knowledge", *Research on Language & Social Interaction*, 45(1):1-29. 2012.
- [12] Andruski, J. E. and Costello, J., "Using polynomial equations to model pitch contour shape in lexical tones: an example from Green Mong", *Journal of the International Phonetic Association*, 34(2):125-140. 2004.
- [13] Grabe, E., Kochanski, G. and Coleman, J., "Connecting intonation labels to mathematical descriptions of pitch", *Language and speech*, 50(3):281-310. 2007.
- [14] Torreira, F., "Tonal realization of syllabic affiliation in Spanish", *ICPhS XVI, Saarbrücken*, 6-10. 2007.
- [15] Gravano, A. and Hirschberg, J., "Turn-taking cues in task-oriented dialogue", *Computer Speech & Language*, 25(3):601-634. 2011.
- [16] Duncan, S., "Some signals and rules for taking speaking turns in conversations", *Journal of personality and social psychology*, 23(2):283-292. 1972.
- [17] Bolinger, D., *Intonation and its uses: Melody in grammar and discourse*. Stanford University Press, 1989.
- [18] Ohala, J. J., "An ethological perspective on common cross-language utilization of F0 of voice", *Phonetica*, 41(1):1-16. 1984.
- [19] Gussenhoven, C., "Intonation and interpretation: phonetics and phonology", *Speech Prosody*. 2002.