



On the intonation of French *wh*-in-situ questions:

What happens before the *wh*-word is reached?

*Stella Gryllia*¹, *Lisa Lai-Shen Cheng*^{1,2}, *Jenny Doetjes*¹

¹ Leiden University Centre for Linguistics (LUCL)

² Leiden Institute for Brain and Cognition (LIBC)

s.gryllia@hum.leidenuniv.nl, l.l.cheng@hum.leidenuniv.nl, j.doetjes@hum.leidenuniv.nl

Abstract

Previous studies on the intonation of *wh*-in-situ questions in French have focused mainly on the utterance final movement and ignored the prosodic properties of the region preceding the *wh*-constituent. Yet, these latter properties are particularly interesting from a processing perspective, as they may help the listener anticipate a question before the *wh*-word is reached, which might facilitate parsing. In this paper we present the results of a production experiment testing the hypothesis that the prosodic properties of *wh*-in-situ questions differ from the prosodic properties of their declarative counterparts. The results of the production experiment verified this hypothesis. The subject and the first syllables of the verb are significantly shorter in *wh*-in-situ questions than in declaratives, while the last syllable of the verb is lengthened in questions. Moreover, we found that the first syllable of the *wh*-word systematically bears an emphatic accent or C-accent (Rossi 1985, Beyssade et al. 2004). This leads to the hypothesis that the prosodic differences on the syllables preceding the *wh*-word could well be a side effect of the presence of a C-accent on the *wh*-word.

Index Terms: intonation, *wh*-in-situ questions, clause type, F0, emphatic accent, duration, French

1. Introduction

French employs a variety of structures for forming *wh*-questions. In one of these structures, normal word order is used, and the *wh*-constituent is left *in situ*, as illustrated in (1). Note that the question in (1) can be used as a real new information-seeking question.

- (1) Jean a acheté quoi ?
Jean has eaten what
'What did Jean eat?'

Previous studies on the prosodic properties of *wh*-in-situ questions in French have mainly focused on the utterance final pitch movement. For instance, Cheng and Rooryck (2000) examine the licensing conditions of French *wh*-in-situ questions and argue for an intonation morpheme which licenses *wh*-in-situ and *yes-no* questions. They also argue that *wh*-in-situ questions in French are realized with a sentence final rising intonation. Contra Cheng and Rooryck, Adli (2006) argues that French *wh*-in-situ questions can be realized

with sentence final rising or falling intonation. Déprez, Syrett and Kawahara (2013) ran a production experiment to test Cheng and Rooryck's proposal that *wh*-in-situ questions end with a rising intonation, and concluded that the majority of their speakers realized the *wh*-in-situ questions with a sentence final rising intonation, as predicted. They also compared this final pitch rise with the pitch rise exhibited in *yes-no* questions, and reported that the two differ. All these studies are concerned with the utterance final pitch movement (rising vs. falling, similar or different from the pitch rise of *yes-no* questions) and do not consider the prosodic properties of the region preceding the *wh*-word. However, these prosodic properties are particularly interesting from a processing perspective, as prosodic cues may help the listener anticipate a question before the *wh*-word is reached. To determine these prosodic properties in detail we ran a production experiment.

2. Production Experiment

The aim of this production experiment was to investigate whether the prosodic properties of *wh*-in-situ questions differ from the prosodic properties of their declarative counterparts when looking at the region preceding the *wh*-constituent. Our working hypothesis is that the two differ in duration and F0.

2.1. Participants

Forty monolingual native speakers of French (30 female and 10 male) participated in the experiment. At the time of recording they were graduate and postgraduate students at the University of Nantes and they ranged in age from 18 to 27 years old. None of them reported any speech or hearing disorders. All participants spoke Standard French and originated from the region Pays de la Loire. One speaker was not considered for further analysis, as he suffered from a cold.

2.2. Stimuli

A set of 16 utterances of subject-verb-direct object-indirect object was designed. Each of these syntactic constituents could in principle form its own accentual phrase (AP) (for a definition of the accentual phrase in French, see Jun and Fougeron 2002). All the utterances were further manipulated with respect to the length of the subject (short: one-syllable first name, vs. long: three-syllable first name) and the type of the direct object (a noun phrase composed of an indefinite determiner and a two-syllable noun, vs. a *wh*-phrase composed of a *wh*-word and a two-syllable noun). All utterances contained a ditransitive verb composed of an inflected

auxiliary verb (*a* ‘has’) followed by a three-syllable past participle and a four-syllable indirect object which was composed of a one-syllable preposition, a one-syllable possessive adjective and a two-syllable noun. Sonorants were used as much as possible to minimize the disruption of the F0 and to make easier the analysis of F0. This resulted in a total of 64 target stimuli (16 items × 2 subject length × 2 clause types). The 64 target stimuli were intermingled with fillers and were presented to each participant. Two pseudo-randomized lists were used to avoid any ordering effect. This production experiment was part of a bigger production experiment which investigated the prosody of questions in French. An example of a stimulus is given in (2). (2a) exemplifies a short subject in declarative, while a *wh*-in-situ question is exemplified in (2b).

- (2) a. Anne a proposé une sortie à ses élèves.
 Anne has suggested an outing to her students
 ‘Anne suggested an outing to her students.’
 b. Anne a proposé quelle sortie à ses élèves ?
 Anne has suggested which outing to her students
 ‘Which outing did Anne suggest to her students?’

2.3. Task and procedure

Participants were tested individually in a sound booth at Pôle Audiovisuel and Multimédia (PAM) at the University of Nantes. They were seated in front of a computer screen and wore headphones. In the beginning of the recording session participants received written instructions that were displayed on the computer screen. They were instructed to read the sentence that appeared on their computer screen first silently and then utter it. The sentences were presented without any preceding context. Participants were also asked to utter the sentence as naturally as possible and at a normal speech pace. Once they had uttered the sentence, the next sentence appeared on the screen. Participants were directly recorded on a computer disk using Audacity Software at 44.1 kHz. The experiment lasted approximately 50 minutes and participants were reimbursed for their cooperation.

2.4. Data analysis

The productions of 39 participants, 2496 utterances in total (39 participants × 64 targets) were inspected for speech errors, hesitations or unnatural pausing. After inspection, 380 utterances were excluded from any further analysis due to disfluencies (15.2% of the total data). Each of the remaining 2116 utterances was segmented in phones, syllables and words with EasyAlign scripts (Goldman 2011, see also Goslin, Content, Goldman and Frauenfelder 1999 for the syllabification rules). Segmental and syllable boundaries were then manually checked and corrected where necessary by close inspection of waveforms and spectrograms. We then extracted the duration of all segments and syllables with the help of a Praat script (Boersma and Weenink 2015). We inspected the data again to decide on F0 measurement points taking into consideration the six surface realizations of the accentual phrase (AP) that have been proposed by Jun and Fougeron (2002). Given the aim of this study, we examined the F0 pattern of the subject and the verbal complex (the auxiliary verb and the past participle). The subject was always realized with a single rise, irrespective of its length (one- or three-syllables long) and of the condition (declarative or *wh*-

in-situ question). With the help of a Praat script we set F0 measurement points for the beginning of the rise L (F0 minimum) and the end of the rise H (F0 maximum); see Figure 1 and 2.

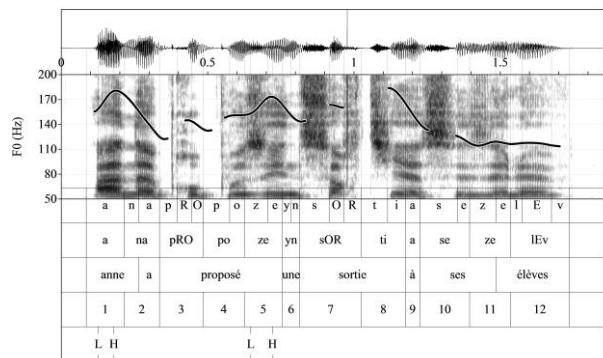


Figure 1: Waveform, spectrogram and F0 of a declarative sentence with short subject uttered by a male speaker. The second tier indicates the syllabification.

The realization of the past participle varied across conditions; in declaratives the past participle ended with a rise, while in *wh*-in-situ questions it was flat and remained low. With the help of a Praat script we set F0 measurement points for the beginning of the rise L (F0 minimum) and the end of the rise H (F0 maximum); see Figure 1. When the F0 was flat, we set only one measurement point, namely, the minimum F0 value of the final vowel of the verb participle; see Figure 2. F0 values were also extracted using a Praat script.

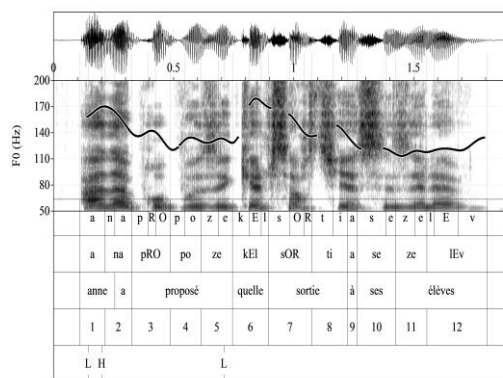


Figure 2: Waveform, spectrogram and F0 of a *wh*-in-situ question with short subject uttered by a male speaker. The second tier indicates the syllabification.

2.5. Results

2.5.1. Duration results

Figure 3 presents syllable durations before the direct object for the short subject (one-syllable long), while Figure 4 presents syllable durations before the direct object for the long subject (three-syllables long). We examined whether the clause type has an effect on syllable duration, running a series of linear mixed effects models (REML) and using the R package lme4 (version 1.1-8). In the models, we included the relevant

syllable duration as a dependent variable, the clause type as a fixed factor, and items and participants as random factors. Here, we report only the durational differences that are statistically significant and are long enough to be relevant for perception. Particularly, when the subject was short, syllable 1 (s1) in declarative was significantly longer (10 ms) than syllable 1 in *wh*-in-situ question ($\beta = 0.007$, $se = 0.002$, $t = 3.193$). The penultimate syllable of the past participle (s4) in declarative was also significantly longer (9 ms) than its counterpart in *wh*-in-situ question ($\beta = 0.006$, $se = 0.0008$, $t = 7.674$).

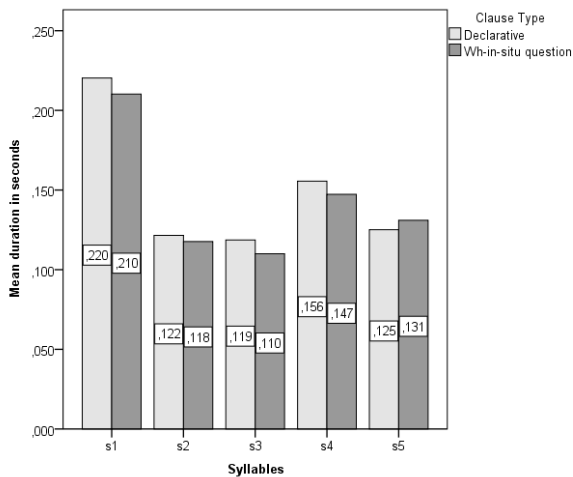


Figure 3: Mean duration of syllables 1-5 measured in seconds for short subject broken down in clause type.

When the subject was long, the penultimate syllable of the past participle (s5) in declarative was significantly longer (9 ms) than its counterpart in *wh*-in-situ question ($\beta = 0.010$, $se = 0.001$, $t = 7.803$).

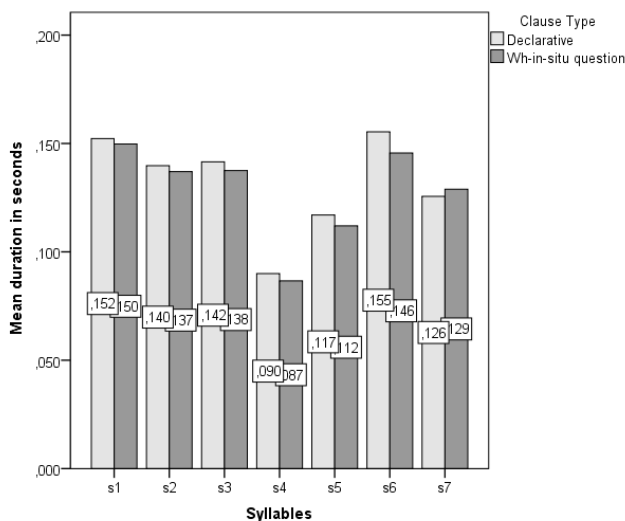


Figure 4: Mean duration of syllables 1-7 measured in seconds for long subject broken down by clause type.

2.5.2. F0 results

Figure 5 shows F0 measurement points before the direct object for the short subject (one-syllable long), while Figure 6 shows F0 measurement points before the direct object for the long subject (three-syllables long). As shown in Figures 5 and 6, the main difference between declaratives and *wh*-in-situ questions is the lack of a F0 maximum point (H) in the verb region in *wh*-in-situ questions. Considering the remaining three F0 measurement points, we examined whether the clause type has an effect on F0, running a series of linear mixed effects models (REML) in R. In the models, we included the relevant F0 point as a dependent variable, the clause type as a fixed factor, and items and participants as random factors. When the subject was short, the L point in the subject did not differ between declaratives (192.0 Hz) and *wh*-in-situ questions (193.6 Hz), ($\beta = 2.670$, $se = 1.674$, $t = 1.595$). Moreover, there was no significant difference between declaratives (250.5 Hz) and *wh*-in-situ questions (252.5 Hz) at the H point in the subject, ($\beta = 4.286$, $se = 2.484$, $t = 1.725$). The two clause types differed significantly at the L point of the verb participle. The L in declarative (189.3 Hz) was located at a mean distance of 21ms from the syllable onset of the ultimate syllable of the verb participle and it was significantly higher than the L in *wh*-in-situ questions (183.5 Hz), ($\beta = 4.565$, $se = 1.121$, $t = 4.074$).

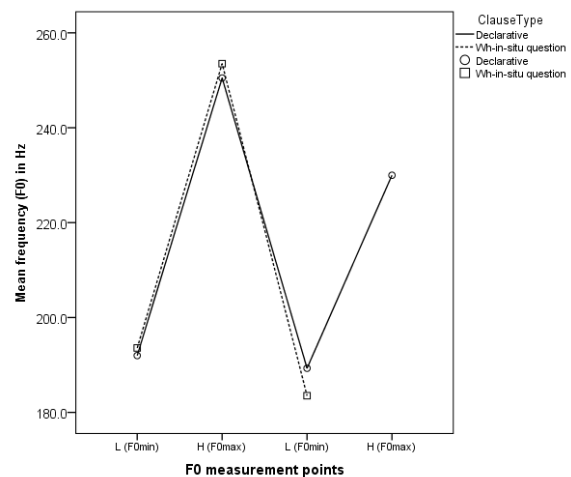


Figure 5: Mean F0 of Lows and Highs for short subject.

In the long subject condition, the L point in the subject did not differ between declaratives (189.2 Hz) and *wh*-in-situ questions (186.2 Hz), ($\beta = 2.339$, $se = 1.633$, $t = 1.432$). In declaratives the L point was located at a mean distance of 112 ms from the first syllable onset. In *wh*-in-situ questions the L point was located at a mean distance of 108 ms from the first syllable onset. At the H point of the subject, there was also no significant difference between declaratives (256.7 Hz) and *wh*-in-situ questions (260.9 Hz), ($\beta = 3.339$, $se = 1.825$, $t = 1.859$). The L in the ultimate syllable in declaratives (189.0 Hz) was significantly higher than the L in *wh*-in-situ questions (184 Hz), ($\beta = 5.213$, $se = 0.973$, $t = 5.357$).

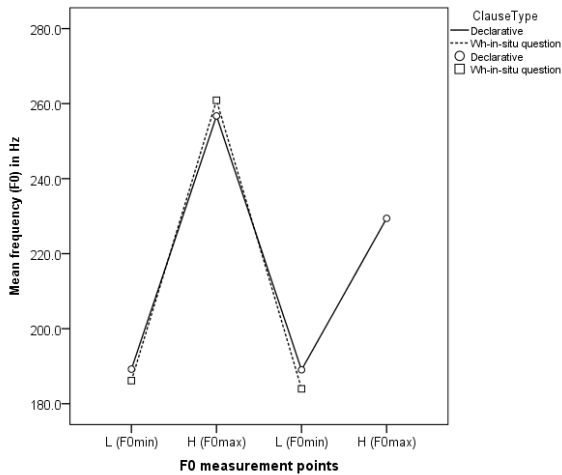


Figure 6: Mean F0 in Lows and Highs for long subject.

2.5.3. Connecting with the C accent

In our data we found that the *wh*-word *quel* ‘which’ systematically bears an emphatic or C-accent. Beyssade et al. argue that C accents are characterized by a longer initial consonant and a high tone [11]. In our data, the initial /k/ of *quel* has a mean duration of 80 milliseconds. Indicatively, we mention that preliminary results of a second production experiment that examined fronted *wh*-questions of the type *Quel billet Elodie a procuré à son amie ?* ‘Which ticket did Elodie give to her friend?’ show that the mean duration of /k/ in fronted *wh*-questions was much shorter (47 milliseconds). Furthermore, Beyssade et al. [11] argue that compulsory C accents occur on contrastive topics and foci (“they occur in utterances performing a complex discourse strategy”, p. 191). Their presence could be related to the strong presuppositional nature that French *wh*-in-situ questions typically have (Cheng & Rooryck 2000). In our data, the presence of the C-accent correlates with the lack of a rise on the preceding final syllable of the verb participle. The lack of this rise suggests the lack of an accentual phrase boundary between the verb and the *wh*-phrase (see Jun and Fougeron 2002 on accentual phrases and their prosodic properties).

3. Conclusions

Most of the previous studies on the prosodic properties of *wh*-in-situ questions have focused on the final pitch movement ignoring the prosodic properties of the region preceding the *wh*-word. However, these prosodic properties are particularly interesting from a processing point of view, as they may help the listener anticipate a question before the *wh*-word is reached. In this paper we presented the results of a production experiment that tackled this issue and examined the prosodic properties of French *wh*-in-situ questions in the region before the *wh*-word comparing them with the prosodic properties of their declarative counterparts. The results of the production experiment can be summarized as follows: the *wh*-word bears systematically a C-accent. In the pre-*wh*-word region questions differ from their counterpart declaratives in duration and F0. In particular, the duration of the penultimate syllable of the subject is significantly longer in declaratives than in *wh*-in-situ questions. The final syllable of the verb participle is realized with a pitch rise in declaratives, while the verb

participle remains flat and stays low in *wh*-in-situ questions. The systematic presence of the C-accent, the duration and F0 data seem to suggest that the verb and the *wh*-in-situ phrase is not separated by a prosodic boundary, in contrast with the verb and its direct object in a declarative sentence. Whether listeners make use of the durational and F0 acoustic cues to anticipate an in-situ *wh*-question is an issue for further research by means of a perception experiment. A second issue for further research concerns the C-accent. Are the durational and F0 differences between *wh*-in-situ-questions and declaratives in the pre-*wh*-word region due to the clause type (question vs. declarative) or are they due to the C-accent and the prosodic marking of declaratives and questions as such in the pre-*wh*-word region is less direct ?

4. Acknowledgements

This research was funded by the NWO project *Understanding Questions*. We would like to thank the members of our research team Leticia Pablos, Yang Yang and Aliza Glasbergen-Plas for their comments. We would also like to thank Sylvie Thomeer for constructing the experimental stimuli, Lucas Tual for helping with running the experiment and Thomas Jansen for general assistance with the experiment. Finally, we would like to thank Hamida Dermidache and particularly Eric Quézin at the University of Nantes.

5. References

- [1] M. Rossi, “L’intonation et l’organisation de l’énoncé”, *Phonetica*, vol.42, pp. 135-153, 1985.
- [2] C. Beyssade, E. Delais-Roussarie, J. Doetjes, J.M. Marandin and A. Rialland “Prosody and Information in French” in F. Corblin and H. de Swart (eds.) *Handbook of French Semantics*. CSLI, pp. 477-499, 2004.
- [3] L. Cheng and J. Rooryck, “Licensing *wh*-in-situ”, *Syntax*, vol.3, pp. 1-19, 2000.
- [4] A. Adli, “French *wh*-in-situ, Questions and Syntactic Optionality: Evidence from Three Data Types”, *Zeitschrift für Sprachwissenschaft*, vol.25, pp. 163-203.
- [5] V. Déprez, K. Syrett and S. Kawahara, “The interaction of syntax, prosody, and discourse in licensing *wh*-in-situ questions”, *Lingua*, vol. 124, pp. 4-19, 2013.
- [6] S.A. Jun and C. Fougeron “Realizations of accentual phrase in French” *Probus*, vol. 14, pp. 147-172, 2002.
- [7] J.-P. Goldman “EasyAlign: An automatic phonetic alignment tool under Praat”, *Proceedings of Interspeech*, pp. 3233-3236, 2011.
- [8] J. Goslin, A. Content, J.-P. Goldman and U.H. Frauenfelder “Syllable boundary placement in French; man and machine; a comparison” *Actes des 2èmes Journées d’ Etudes Linguistiques*, Nantes, 1999.
- [9] P. Boersma and D. Weenink Praat: doing phonetics by computer [Computer program]. Version 6.0.05, retrieved from “<http://www.praat.org/>”, 8 November 2015.
- [10] D. Bates, M. Maechler, B. Bolker, S. Walker “Fitting Linear Mixed-Effects Models Using lme4” *Journal of Statistical Software*, vol. 67(1), pp. 1-48. 2015.
- [11] C. Beyssade, E. Delais-Roussarie, J. Doetjes, J.M. Marandin and A. Rialland “Prosodic, syntactic and pragmatic aspects of information structure. An introduction” in F. Corblin and H. de Swart (eds.) *Handbook of French Semantics*. CSLI, pp. 455-476, 2004.