


ARTICLE

# Prosody and speech act interpretation: The case of French indirect requests<sup>†</sup>

Nicolas Ruytenbeek<sup>1\*</sup> , Benjamin Bergen<sup>2</sup> and Sean Trott<sup>3</sup>

<sup>1</sup>Department of Translation, Interpreting and Communication, Ghent University, Gent, Belgium; visiting scholar at the Leiden University Centre for Linguistics (LUCL), Leiden, The Netherlands; the Université de Lille, Lille, France, <sup>2</sup>Department of Cognitive Science; University of California, San Diego, USA and <sup>3</sup>University of California at San Diego, San Diego, USA

\*Corresponding author. Email: [nicolasruytenbeek@gmail.com](mailto:nicolasruytenbeek@gmail.com)

(Received 08 October 2021; revised 24 June 2022; accepted 14 July 2022; first published online 06 September 2022)

## Abstract

Some utterances are pragmatically ambiguous. For instance, *Tu peux fermer la fenêtre ?* (“Can you close the window?”) can be a request for information or an “indirect request” (IR) to close the window. A possible way for speakers to make it clear whether they intend these expressions as a direct or indirect speech act is to use cues such as gestures or prosody. It has been shown for English that participants’ identifications of IRs are predicted by  $f_0$  slope, mean  $f_0$ , and  $f_0$  duration. However, the extent to which these findings extend to other languages remains unknown. In this article, we explore the prosodic features associated with French IRs, a language poorly documented from that perspective. We address two research questions: Are listeners’ pragmatic interpretations of French IR constructions predicted by speaker’s original intent? Do prosodic cues play the same role in French modal interrogatives as in declarative remarks? We find, first, that remarks with more positive  $f_0$  slope are more likely to be interpreted as requests, but modal interrogatives with more positive  $f_0$  slope are more likely to be taken as questions. Second, while longer remarks were more likely to be interpreted as requests, longer modal interrogatives were more likely to be interpreted as questions.

**Keywords:** prosody; pragmatics; French; requests; indirect speech acts; indirect requests; modals; negative state remarks

## 1. INTRODUCTION

Addressees might face difficulties when interpreting ambiguous expressions in context. It is therefore possible that, in the case of spoken communication, speakers will try to avoid misunderstandings by exploiting prosodic features to clarify the meaning they want to get across. Indeed, experimental evidence demonstrates that speakers use prosodic cues to aid in the disambiguation of syntactically ambiguous utterances (Price et al. 1991; Beach 1991; Snedeker & Trueswell

<sup>†</sup>The online version of this article has been updated since original publication. A notice detailing the changes has also been published at <https://doi.org/10.1017/S0959269522000308>

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

2003) and in the recognition of irony (e.g., Cheang & Pell 2008; Bryant & Fox Tree 2002, 2005). Previous research also indicates that prosody enables speech act (SA) disambiguation. For instance, intonational contours can be reliably used to interpret one-word utterances, such as *beer*, as the SAs of criticism, doubt, warning, suggestion, naming, and wish (Hellbernd & Sammler 2016). Does prosody also play a role in the performance and the interpretation of indirect speech acts (ISAs), for example “indirect requests” (Ruytenbeek 2017)?

The findings reported on by Nickerson & Chu-Carroll (1999) provide evidence for a rising intonational pattern associated with polar questions in English. These authors collected data about the intonational patterns of the utterance’s nucleus (the accented syllable of the intonation unit), the maximal  $f_0$  at the final high boundary tone (corresponding to a rise or fall in  $f_0$ ), and the  $f_0$  mean of the unstressed vowel in the intonational phrase. In their experiment, they asked participants to read aloud parts of small dialogues including the utterance of an indirect request (IR) construction consisting in one of the following (the remainder of the dialogues were played by a recorded voice):

- (1) Can you move the couch?
- (2) Would you move the couch?
- (3) Would you be willing to move the couch?

Each utterance was produced in each condition: direct speech act (DSA) condition, where the utterance had the meaning of a polar question (direct request for information), and indirect speech act (ISA) condition, where it was meant as an IR. Nickerson & Chu-Carroll found that the majority of all utterances were produced with an intonational contour typical of polar questions (“L\*HH%” in autosegmental terms, i.e., an intonational contour consisting in a low plateau of  $f_0$  followed by a rise). However, utterances intended as requests were more likely to have low boundary tones (falling  $f_0$  curve at the end of the utterance): 19 utterances with this intonational contour were IRs vs. seven direct requests for information. In addition, while DSAs and ISAs were similar in terms of the mean  $f_0$  of the vowel carrying the nuclear pitch accent, they differed in the  $f_0$  values of their boundary tones.

Concerning the relationship between prosody and SAs, Ward (2019: 79–82) notes the absence of a final rise in intonation for polar (or “yes-no”) information questions. He also discusses intonational differences for an English utterance such as (4) meant as a request for information (“Is it the case that you like tea?” asked as part of a survey) or as an invitation (offering one’s guest a cup of tea).

- (4) Do you like tea?

While, according to Ward, a flat intonational contour is associated with the force of a request for information, the presence of a late increase in fundamental frequency ( $f_0$ ) (*late pitch peak*) would be expected in the case of an invitation. He also comments that utterances that could otherwise be taken as invitations would be perceived as rude when they lack a late rise in  $f_0$ .

More recently, Trott et al. (2019) asked, like Nickerson & Chu-Carroll (1999), whether speakers use prosodic information to differentiate indirect requests and

their literal counterparts, and whether listeners use the same information to distinguish between these two sorts of SA interpretations. They considered seven possible prosodic cues: mean  $f_0$ ,  $f_0$  range (maximum  $f_0$  minus minimum  $f_0$ ), standard deviation of  $f_0$ , duration (i.e., number of voiced frames: this amounts to the duration of the utterance for which there was voicing, i.e., no pauses or non-voiced consonants), mean intensity, standard deviation of intensity, and slope of  $f_0$  (i.e., the coefficient obtained by regressing  $f_0$  against time). They also considered two distinct grammatical forms: modal interrogatives (e.g., *Can you open the window?*) and remarks such as (5) intended as a request e.g., to open the window.

(5) My office is really hot.

To do this, they recorded native speakers uttering sentences formatted in both kinds of grammatical constructions (i.e., both modal interrogatives and declarative remarks describing negative states of affairs – “negative state remarks”). Each speaker produced two versions of each utterance: once as a request, and once as a literal or direct speech act – i.e., a question for the modal interrogatives, and a complaint for the negative state remarks. Trott et al. (2019) found that the intent of modal interrogatives was predicted by their duration (longer utterances were more likely to be questions), their variation in intensity (utterances with more variation were more likely to be questions), and their  $f_0$  slope (utterances with a more positive  $f_0$  slope were more likely to be questions). In contrast, longer negative state remarks were more likely to be intended as requests. If one considers that a request performed with a modal interrogative such as *Can you...?* is a conventionalized request, it makes sense to say that the less predictable speech act is marked with duration in both modal interrogatives and negative state remarks.<sup>1</sup> For both kinds of constructions, utterances intended as requests had a higher mean for intensity.

In a separate experiment, Trott et al. (2019) also asked whether listeners could identify – in the absence of any other contextual information – the speaker’s intended interpretation at a rate above chance, and whether the same acoustic features predicted their accuracy. In their study, participants listened to a series of paired utterances of the same speaker producing the same sentence – once as a request, and once as a non-request (i.e., either a question or direct statement). Participants were instructed to identify which of the two was meant as a request for action (IR). They were presented with a total of 12 pairs of utterances (six for modal interrogatives and six for the negative state remarks) spoken by five different speakers.

They found that listeners could discriminate which utterance was intended as a request, and that listener accuracy was predicted by  $f_0$  slope, mean  $f_0$ , mean intensity and standard deviation of intensity. Modal interrogative pairs with a larger difference in mean  $f_0$ , as well as a larger difference in their  $f_0$  slope, were more likely to be answered correctly. Utterance pairs with a larger difference in mean intensity were also more likely to be answered correctly.

The results of the task designed by Trott et al. (2019) provide important insights into the processing of pragmatically ambiguous constructions. They demonstrate

<sup>1</sup>We are grateful to an anonymous reviewer for bringing this to our attention.

that speakers produce reliable prosodic cues to enable interpreters to arrive at the intended meaning, and that interpreters effectively use these cues to disambiguate between alternative readings. They also demonstrate that, at least in some cases, different grammatical constructions (i.e., modal interrogatives vs. declaratives) may be disambiguated by distinct grammatical cues.

Interestingly, Trott et al.'s (2019) study differs from previous work on prosody and SA interpretation in that these authors examined two types of constructions that differ in terms of degree of conventionalization qua requests: *Can you VP?*, which has a high degree of conventionalization, and negative state remarks, which are associated with a lower degree of conventionalization. It should be borne in mind that, first, these English constructions, and their French counterparts, might be associated with different cues to their IR interpretation. Second, it is possible that these two types of constructions have different default interpretations: while the request meaning has become the more salient (or the more easily accessible) interpretation of *Can you VP?* interrogatives (see Ruytenbeek 2021 for a discussion), this is less likely to be the case for negative state remarks such as *My soup is cold* (or maybe in very specific contexts, such as addressed to a waiter in a restaurant). Both kinds of differences might interact with the prosodic cues used by speakers to signal their pragmatic intent. The question whether prosody should play a more important role for one type of construction or for the other is a difficult one to answer. It is possible that prosody plays a stronger role when the “prior distribution” (i.e., the relative likelihood of each interpretation) over interpretations is very flat, e.g., [.5, .5]. Alternatively, prosody could play a stronger role when the prior distribution is very skewed, and speakers need to signal the non-default interpretation, e.g., [.1, .9]. At this stage, it is unclear which hypothesis is correct. In addition, we do not believe that Trott et al.'s data can resolve this issue because the range of “priors” is very small, i.e., only 12 utterances in Trott et al.'s (2019) study, and these utterances are mostly defined in terms of their grammatical form.

The results of Trott et al.'s (2019) study were obtained using English stimuli with a population of native English speakers. Importantly, it remains unknown whether similar findings would be obtained in other languages. This article makes a first step in that direction by exploring the prosodic features associated with IRs in French, a language closely related to English but for which the prosodic patterns associated with different sentence-types and SAs types are rarely substantiated with empirical data. In Section 2, we address available research on the relationship between intonation and SAs in French, focusing on the prosody of polar interrogatives (Section 2.1) and the prosody of non-inverted interrogatives in particular (Section 2.2). Section 3 is devoted to a production study in which we collected spoken utterances of remarks either used as direct speech acts (question/statement) or as indirect requests. These utterances were used as experimental stimuli for a perception study reported on in Section 4, in which we tested whether native speakers of French are able to identify the request and non-request uses of the same remarks. Section 5 offers a general discussion of our results and outlines directions for future research on SAs and prosody.

## 2. INTONATION AND SPEECH ACT DISAMBIGUATION IN FRENCH

In French as in other languages, intonation plays different roles, the two most relevant ones for our present purposes being the expression of attitudes and emotions and an illocutionary function (for an overview, see Di Cristo 2016). To the best of our knowledge, the prosody of negative state remarks and that of ‘conventional’ IRs such as modal polar interrogatives has not been investigated yet in the French language. We will therefore concentrate, in Section 2.1, on French interrogative sentence-types, which are ambiguous between a direct, question interpretation, and an IR interpretation.

### 2.1 The prosody of polar interrogatives

As in languages such as English, Dutch, and Italian (for a review, see Shriberg et al. 1998), French polar questions correlate with rising intonation (Safarova et al. 2005), and questions are shorter than statements because of their shorter final vowels in particular (Torreira & Valtersson 2015; but see Smith 2002). In addition, French questions have a more pronounced final intensity drop compared to statements (Rossi et al. 1981; Valtersson & Torreira 2014; Torreira & Valtersson 2015) and their pitch contours start at a higher register than those of statements (Torreira & Valtersson 2015). The automatic classification of questions and statements carried out by Torreira & Valtersson (2015) reveals an increase in accuracy from 60–65% for most individual cues to almost 80% if all the phonetic cues we have just mentioned are present. Concerning the rising intonation associated with polar questions and continuation statements (H\*H% in autosegmental terms), Valtersson & Torreira (2014) propose that, while both SAs make use of a pattern of rising intonation, the phonetic realization of this pattern is conditioned by the interactive and communicative parameters of the contexts in which these utterances are produced.

French interrogatives with different pragmatic interpretations can be identified on the basis of their prosodic properties. For instance, Michelas et al. (2015) have shown that, unlike unbiased questions such as in (6), questions that are uttered in a negatively biasing context have a  $f_0$  peak in the penultimate syllable of the final word of the utterance (a noun, here *guéridon*).

- (6) Yves a revendu le guéridon ? Ou devons-nous le remettre en vente ?  
‘Has Yves resold the pedestal table? Or do we need to place it on sale again?’
- (7) Yves a revendu le guéridon ? Mais je viens de le voir dans le garage !  
‘Yves resold the pedestal table? But I have seen it in the garage a few minutes ago.’

Another interesting case study concerns rhetorical questions (Delais-Roussarie & Beysade 2019). These authors demonstrate that polar and Wh- interrogatives such as (8)–(9) respectively more often have a rising intonational contour and are spoken faster when they are used as information-seeking questions.<sup>2</sup>

<sup>2</sup>For a prosodic analysis of French wh-questions that are ambiguous between an information-seeking and a non-canonical meaning of surprise (or expectation denial), see also Brunetti et al. (2021), as well as Celle and Pélissier (2021).

- (8) Est-ce que quelqu'un mange des épinards ?  
'Does anyone eat spinach?'
- (9) Qui mange des épinards ?  
'Who eats spinach?'

When they are used as rhetorical questions, these interrogatives more frequently exhibit a falling intonational contour.<sup>3</sup> These authors remark, however, that not all utterances of rhetorical vs. information-seeking questions contain these prosodic cues. According to them, the rise-fall contour typical of rhetorical questions is used by speakers to signal an implicit meaning (Beyssade & Marandin 2006), i.e., that nobody eats spinach in the case of (8)–(9).

## 2.2 The prosody of French non-inverted interrogatives

Delattre (1966) identifies ten different patterns for French prosodic units, based on a combination of three acoustic features, i.e.,  $f_0$ , intensity and duration. With respect to interrogative sentences, he distinguishes between two intonational profiles corresponding, on the one hand, to non-inverted interrogatives (what he ambiguously calls *questions*, which also seem to include polar interrogatives), such as (10), and Wh- interrogatives (what he calls *interrogations*) such as (11).

- (10) C'est bien toi ?  
'It's you, right?'
- (11) Qui les vend ?  
'Who sells them/is selling them?'

The intonational pattern for polar interrogatives is a rising  $f_0$  curve (Fig. 1), while for Wh- interrogatives, it is characterized by a falling  $f_0$  curve (Fig. 2).

Even though sentences produced with these intonational contours will certainly sound familiar to native speakers of French, their accuracy and representativeness have been criticized. For example, Di Cristo (2016: 294) points out that a distinction between different intonational patterns in terms of morphosyntactic criteria is missing from Delattre's approach, and that his repertoire of interrogative sentence-types is incomplete. Following Di Cristo, polar questions can be asked using declaratives with a rising intonation such as (12), subject-verb inverted interrogatives such as (13), interrogatives initiated by the *Est-ce que* ("is it the case that") such as (14), with or without a linguistic negation.

- (12) Tu (ne) fermes (pas) la fenêtre ?  
'Are(n't) you closing the door?'
- (13) (Ne) fermes-tu (pas) la fenêtre ?  
'Are(n't) you closing the window?'

<sup>3</sup>Falling intonation is not uncommon in French wh-questions more generally (see e.g., Delais-Roussarie et al. 2015).

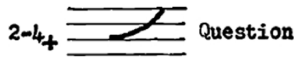


Figure 1. Intonational pattern for polar interrogatives (Delattre 1966: 4)

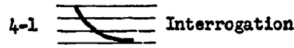


Figure 2. Intonational pattern for Wh- interrogatives (Delattre 1966: 4)

- (14) Est-ce que tu (ne) fermes (pas) la fenêtre ?  
 ‘Are(n’t) you closing the window?’

He observes that, while some utterances can be understood as questions without having a final rising  $f_0$  curve, the majority of questions in contemporary French are asked using non-inverted interrogatives, typically uttered with rising intonation. In fact, non-inverted interrogatives have overwhelmingly replaced inverted polar interrogatives in casual conversation (this is discussed by Coveney 2011, see e.g., Wall 1985 for empirical evidence that inverted interrogatives are rare in casual conversation).

Di Cristo (2016) does not discuss possible prosodic differences in the final contour between the use of a non-inverted modal interrogative such as (15) as a direct question (“I am asking you whether it is the case that you are passing me the keys”) or as an indirect request for action (“I am requesting that you pass me the keys”).

- (15) Tu me passes les clés ? (Di Cristo 2016: 330–331)  
 ‘Are you passing me the keys?’

Some non-inverted modal interrogatives can also be used with the illocutionary force of a threat, as (16) illustrates, in which case, again, Di Cristo proposes that no departure from the final  $f_0$  curve is necessary for the threat to be distinguished from the SA of request for information.

- (16) Tu veux que je t’apprenne la politesse ? (Di Cristo 2016: 331)  
 ‘Do you want me to teach you politeness?’

This example is particularly interesting, as it suggests that other prosodic features would be used by the speaker to indicate that the utterance is to be interpreted as a threat, e.g., a higher intensity and louder speech.

### 3. PRODUCTION STUDY

To address the question whether native speakers of French use prosodic information to differentiate indirect requests and their literal counterparts, and, if so, which prosodic cues they produce when uttering these sentences, we carried out a production study.

### 3.1 Method

For the data collection phase of our study, we recruited 14 French native speakers after posting an announcement on Facebook. Of these participants, 13 self-identified as female (one male). The average self-reported age was 23 (median = 23). Ten speakers originated from Belgium, and the remaining four from France. Each speaker received 7,5€ for participating.

Each speaker was given 24 utterances to read aloud and record (see Appendix 1 for a complete list).<sup>4</sup> Twelve of these had the modal interrogative form “Tu peux + verbal phrase (VP) ?” (“Can you (2<sup>nd</sup> person singular) VP?”), e.g., “Tu peux ouvrir cette fenêtre ?” (“Can you open that window?”); these will be referred to as modal interrogatives. Twelve took the form of “My X is Y”, e.g., “Mon verre est vide” (“My glass is empty”); these utterances will be referred to as negative state remarks. Speakers were instructed to say each utterance twice – once as a request and once as a literal question or statement (counterbalanced for order). In order to make the recording process as naturalistic as possible, speakers were given a few words of explanation about a possible situation in which the utterance was produced. We provided them with (minimal) contextual information about the situations in which they could imagine producing the stimuli. Here is an example of contexts for the *Tu peux ouvrir la fenêtre?* (“Can you open the window?”) stimuli: a) As a question: Your friend broke his arm and you want to know whether he is able to open the window. (*Votre ami s’est cassé le bras et vous voulez savoir s’il est capable d’ouvrir la fenêtre.*) b) As a request: You feel too warm and you ask your friend to open the window. (*Vous avez trop chaud et vous demandez à votre ami d’ouvrir la fenêtre.*) They were also given examples at the beginning of the instruction sheet demonstrating how the same sentence could be used either as a request or as a statement (or question). The example sentences did not appear in the target stimuli, and the different versions of each utterance were not spoken aloud to the participants (so that they could not imitate the prosody of the experimenter). Two complementary lists of stimuli were created: in one list, participants were instructed to utter the first sentence as a request first, the second one as a literal SA first, and so on; in the other list, participants were instructed to utter the first sentence as a literal SA first, the second one as a request first, and so on. Speakers were randomly assigned to either list.

Due to the COVID-19 health situation at the time, the speakers could not perform any audio recordings at the university lab, and they were instead asked to make their recordings at home using the microphone of their laptop or phone. Detailed instructions regarding the recording process, including three training items, were provided to them, along with a list of utterances to produce and self-record. Each speaker was instructed to make their recordings in a quiet room with closed windows and door, sitting down, while remaining at a distance of approximately 30cm from their laptop microphone. They were also told to

<sup>4</sup>Empirical evidence shows that the intonation of reading aloud differs from spontaneous speech production (see e.g., Blaauw 1994; Guaitella 1999; Nakamura et al. 2008). Future research will therefore be necessary to see if the results of our production study are replicated in spontaneous conditions of speech production.



keep their voice intensity constant across the recording process. The recorded data was listened to and verified by the first author, and three speakers were asked to record a few utterances again.

### 3.2 Predictions

Based on available experimental data for indirect requests in English (Banuazizi & Creswell 1999; Hedberg et al. 2014; Trott et al. 2019, 2022), we focused on the seven utterance-level acoustic features identified by Trott et al. (2019), i.e., mean  $f_0$ ,  $f_0$  range, standard deviation of  $f_0$ , duration, mean intensity, standard deviation of intensity, and slope of  $f_0$ . We extracted the mean  $f_0$ , the range of  $f_0$ , mean intensity, and the number of voiced frames (as a proxy for duration). We also included measures of dispersion for both pitch (standard deviation of  $f_0$ ) and intensity (standard deviation of intensity).<sup>5</sup> Finally, because the non-request interpretation of our modal interrogatives is a yes-no question about the addressee's ability, and because prototypical French questions have a low-rise pitch contour (e.g., Di Cristo 2016), we used the slope of the  $f_0$  component (slope of regressing  $f_0$  against time) to measure the degree to which an utterance exhibits a rising or falling contour (Roche et al. 2019; Trott et al. 2019).

Following the results of Trott et al.'s (2019; 2022) studies, we predicted, across all utterances, i.e., for both interrogatives and declaratives, that items with higher mean  $f_0$  would be more likely to be requests. We also predicted, across all utterances, that items with more positive  $f_0$  slopes would be less likely to be requests. As we explained in the introduction, it is very difficult, on the basis of available experimental evidence, to determine whether prosody should play a more important role for one type of construction (e.g., the more conventionalized modal interrogatives) or for the other (the less conventionalized negative state remarks). Therefore, we will not make any specific predictions regarding the possible role of conventionalization on the production and interpretation of prosodic cues disambiguating between the IR and the 'literal' uses of modal interrogatives and negative state remarks.

We predicted that, for non-inverted modal interrogatives, requests would have a less positive  $f_0$  slope than their question counterparts, given that the hallmark of information seeking questions in French is the rising intonation (low-rise pitch contour) (Di Cristo 2016). We also predicted, as Trott et al. (2019; 2022) did for English, that negative state remarks in French should have a longer duration (i.e., contain more voiced frames) when intended as requests than when intended as literal statements. That is, we expected speakers to "signal" a deviation from the default interpretation of these remarks as statements by putting the emphasis on specific words or syntactic constituents, which should be detectable in the number of voiced frames. By contrast, we approached the remaining acoustic features that could help distinguish between the literal and IR uses of our stimuli from a more exploratory perspective.

<sup>5</sup>We initially measured intensity in order to be consistent with Trott et al.'s (2019) study. However, we decided not to report on intensity data because of methodological issues related to the recording conditions, i.e., the recordings were done in different places, with different tools (laptop/phone) and without the direct supervision of the experimenter.

### 3.3 Data processing

For each of the 336 recordings (14 speakers producing 12 utterances with two versions each), we used Parselmouth (Jadoul et al. 2018), a Python interface to Praat, to extract the seven acoustic features described above. We then z-scored each of these features with respect to each speaker's mean and standard deviation for that particular feature, to account for considerable inter-speaker variability. The acoustic features and the original audio recordings will be made available on the Open Science Framework.<sup>6</sup>

### 3.4 Results

#### 3.4.1 Analysis of individual acoustic features

To analyse individual acoustic features, we first asked how much independent variance in Meaning (request vs. literal) was explained by each individual feature. To do this, we first constructed a full generalized linear mixed model with a logit link, with Meaning as a dependent variable, and seven acoustic features as predictors (along with random intercepts for each item). We compared this full model to a series of reduced models omitting each feature in turn; model comparisons were performed using log-likelihood ratio tests, and the p-values obtained from each comparison were adjusted for multiple comparisons using Holm-Bonferroni correction (Holm 1979).

In each case, a positive coefficient represents a higher likelihood of a request meaning, while a negative coefficient represents a higher likelihood of a literal meaning. We first conducted these analyses across all utterances, then conducted subset analyses focusing on interrogatives and declaratives independently.

When predicting Meaning across interrogatives and declaratives (i.e., all utterances), mean  $f_0$  and  $f_0$  slope emerged as significant predictors after controlling for multiple comparisons [ $X^2(1) = 35.697$ ,  $p < .001$ ]. Specifically, items with higher mean  $f_0$  were more likely to be requests ( $\beta = 0.92$ ,  $SE = 0.17$ ), and items with more positive  $f_0$  slopes were less likely to be requests ( $\beta = -1.1$ ,  $SE = 0.17$ ).

For interrogative items only, model fit was improved by the inclusion of only  $f_0$  slope [ $X^2(1) = 68.07$ ,  $p < .001$ ]. Consistent with our predictions and with past work (Trott et al. 2019), items with more positive  $f_0$  slopes were less likely to be requests [ $\beta = -1.57$ ,  $SE = 0.25$ ] (see the example sentences in Figure 3).

Finally, for declarative items only, model fit was improved by mean  $f_0$  ( $X^2(1) = 28.36$ ,  $p < .001$ ). As predicted, utterances with higher mean  $f_0$  were more likely to be requests [ $\beta = 1.62$ ,  $SE = 0.35$ ] (see the example sentence in Fig. 4).

Model fit was also improved by the number of voiced frames ( $X^2(1) = 7.53$ ,  $p = .04$ ). Consistent with Trott et al. (2019), longer utterances were more likely to be requests [ $\beta = 0.56$ ,  $SE = 0.21$ ].

#### 3.4.2 Machine learning classifier

While the analysis above indicates which features are informative about the meaning of the recorded utterances, it does not specify how much information a particular feature contains, in particular when all seven features are combined.

<sup>6</sup>[https://osf.io/5t3v4/?view\\_only=d4fe73716dfe4ec285842ba71a1c97b3](https://osf.io/5t3v4/?view_only=d4fe73716dfe4ec285842ba71a1c97b3).

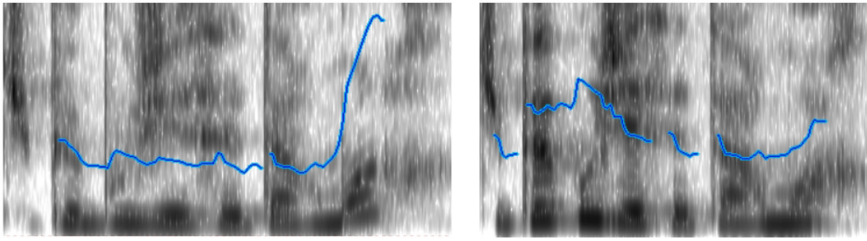


Figure 3.  $F_0$  curves for *Tu peux bouger cette armoire ?* as a literal question (left) and as a request (right)

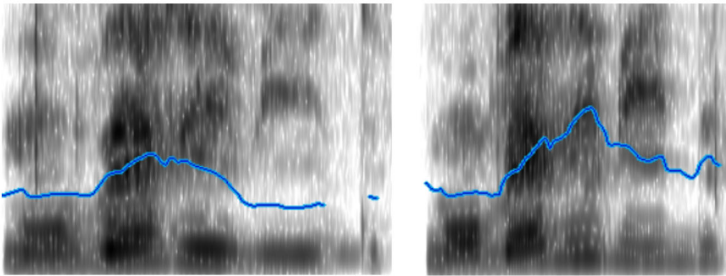


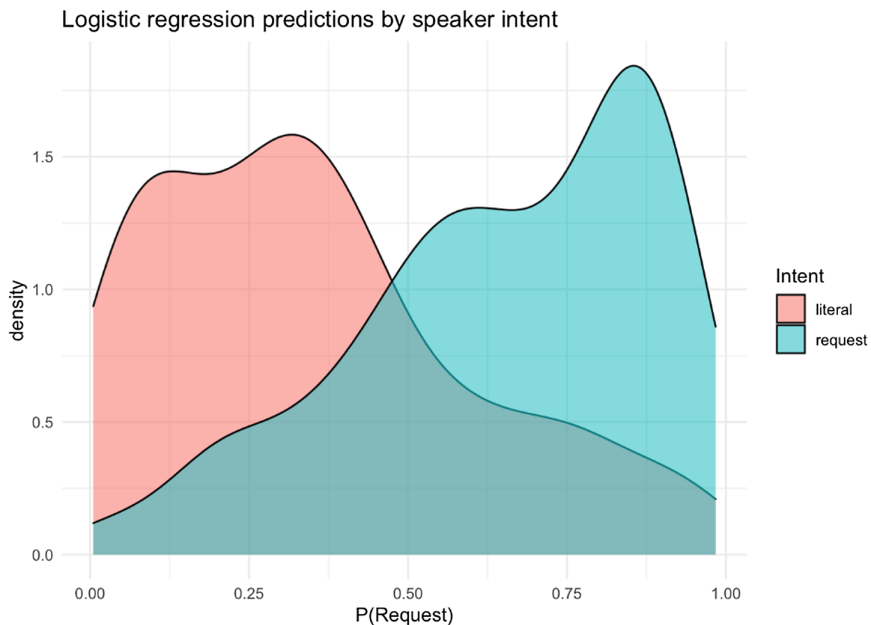
Figure 4.  $F_0$  curves for *Mon verre est vide* as a literal statement (left) and as a request (right)

Building on Trott et al. (2019), one way to address this question is to ask how accurately a classifier can predict Meaning from all seven features combined. We used leave-one-out cross-validation (LOOCV) to quantify the ability of a classifier to generalize from its training set to novel samples. That is, a model was fitted to every item in the dataset but one; the model was then used to classify the held-out test item, enabling us to determine whether the model successfully generalized across to other items (i.e., whether the predicted label from the model matches the actual label for the held-out test item). This leave-one-out procedure was performed for every item in the dataset, ultimately providing an accuracy score, i.e., the percentage of correctly classified held-out items. For each of the 336 splits of the data (336 recorded utterances in total), we thus fitted a logistic regression classifier to the 335 training utterances. The classifier was trained to predict an utterance's original Meaning from all seven acoustic features and their interaction with Form (interrogative or declarative). This classifier was then used to predict the Meaning of the held-out test item.

The classifier successfully predicted Meaning on 76% of held-out test items, a rate substantially above chance. As shown in Figure 5, held-out test items that were estimated as being more likely to be requests (i.e., a larger value for  $p(\text{Request})$ ) were more likely to have originally been meant as requests.

### 3.5 Interim conclusion

To summarize the results of our production study, our data confirm our prediction that, for both modal interrogatives and negative state remarks, the higher the mean



**Figure 5.** A classifier equipped with the seven acoustic features correctly predicted a held-out test item's Meaning 76% of the time; the figure illustrates the distribution of classifier probabilities over classes (e.g., literal, request), coloured by the original Meaning of a given item.

$f_0$  of the utterance, the higher the likelihood that it was intended as a request. In addition, our second prediction according to which both types of utterances with more a positive  $f_0$  slope should be less likely to have been meant as requests was borne out. We also confirmed our third prediction: for negative state remarks stimuli, longer utterances were more likely to be requests. Furthermore, a classifier equipped with seven acoustic features correctly predicted the meaning of the stimuli 76% of the time.

#### 4. PERCEPTION STUDY

The results of our production study indicate that mean  $f_0$ ,  $f_0$  slope, and duration play a role in the disambiguation of IR expressions by speakers producing these utterances. However, at this stage, it remains unknown whether listeners actually use the cues produced by speakers to interpret ambiguous utterances such as IRs. To answer this question, we designed a perception study in which we presented native speakers of French with a sample of utterances recorded in the production study, and asked them to identify which of these utterances have a request meaning.

##### 4.1 Method

###### *Participants*

We recruited 80 participants on the online testing platform Prolific; 39 self-identified as female, 39 as male, and two as non-binary. The average self-reported age was 27.8

(SD = 9.6; median = 24; range 18–59). A number of 74 participants originated from France, and the remaining six participants from Belgium. Each participant received £2.50 for participating (Prolific is a UK-based platform).

### Materials

Four complementary lists of 84 stimuli (six sentences x two intents x seven speakers) were created. The first and second lists contained all declarative utterances from speakers 1–7 or 8–14 (“Mon X est Y”, e.g., Mon verre est vide – *My glass is empty*), respectively. The third and fourth lists contained the interrogative utterances from speakers 1–7 or 8–14 (“Tu peux VP ?”, e.g., Tu peux ouvrir cette fenêtre ? – *Can you open that window?*), respectively.

### Procedure

Each participant was presented with a total of 84 utterances to discriminate. Participants were randomly assigned to a list. Within each list, the utterances were blocked by the speaker, and the order of blocks was randomized across participants.

In order to help participants perform the task successfully, they were informed that many French sentences can be interpreted in different ways, e.g., the sentence *Tu peux passer le sel?* can be a request (as in saying *Passe le sel s’il te plait*), or a question about the physical ability to do some action (as in asking whether or not the addressee would be able to pass the salt). Similarly, the sentence *Mon steak est trop cuit* could be a complaint about the quality of the meal, or it could be a request for the waiter to bring another steak. They were then told that they would listen to recorded utterances and that, after listening to a particular utterance, they would have to indicate whether they believe the utterance is a request or not (decide whether the speaker had the intention that the utterance be interpreted as a request) by responding to a question just below the audio file, and that, if necessary, they could replay the file several times.

## 4.2 Predictions

We first predicted that, as in English, participants’ pragmatic interpretations would be predicted by the speaker’s original intent – that is, participants should interpret requests as requests, and questions (or negative state remarks) as non-requests.

We also had predictions about which acoustic features ought to predict participants’ pragmatic interpretations. For both types of constructions, we predicted that longer duration would increase the likelihood of a request interpretation (Trott et al. 2019). In addition, we expected non-inverted modal interrogatives with less positive f0 slopes to be more likely to be interpreted as requests.

## 4.3 Results

All analyses were conducted in R (R Core Team 2020). Mixed effects models were constructed using the *lme4* package (Bates et al. 2015), and log likelihood ratio tests were compared using the *anova* function.

Our first question was whether participants' pragmatic interpretations were reliably predicted by the speaker's original Intent. To address this question, we constructed a generalized linear mixed effects model with Interpretation (Request vs. Non-Request) as a dependent variable, fixed effects of Intent (Request vs. Non-Request) and Form (Interrogative vs. Declarative), by-subject random slopes<sup>7</sup> for the effects of Intent and Form, by-item random slopes for the effect of Intent, and random intercepts for items, subjects, lists, and speakers. We compared this model to a model omitting only the effect of Intent, and found that the full model exhibited a better fit, as demonstrated by a log likelihood ratio test [ $X^2(1) = 17.78$ ,  $p < .001$ ]. As predicted, Request interpretations were more likely for utterances originally intended as Requests [ $\beta = 1.18$ ,  $SE = 0.2$ ]; participants responded correctly for 61.38% of items. Altogether, these results indicate that our participants successfully made use of prosodic information to infer a speaker's intent. Participants' responses were also influenced by the form of the utterance: the log-odds of Request interpretations was higher for interrogative utterances than declaratives [ $\beta = 1.36$ ,  $SE = 0.41$ ]

Second, we asked whether participants' pragmatic interpretations could be predicted by any of four acoustic features (Mean  $f_0$ ,  $f_0$  Slope, and Number of Voiced Frames), and their interaction with Form. Consistent with past work (Trott et al. 2019), we constructed a full model with Interpretation as a dependent variable, fixed effects for each of the seven features mentioned earlier, as well as their interaction with Form, and random intercepts for subjects, items, lists, and speakers. We then compared this full model to a series of reduced models, each of which omitted an interaction term between Form and one of the critical four acoustic features; we also compared each of those reduced models to a model omitting the acoustic feature entirely (i.e., to ask about a main effect of that acoustic feature). After correcting for multiple comparisons, we found that model fit was improved by the interactions between Form and  $f_0$  Slope [ $X^2(1) = 45.38$ ,  $p < .001$ ], as well as a main effect of  $f_0$  Slope [ $X^2(1) = 153.06$ ,  $p < .001$ ]. The log odds of a request interpretation increased for utterances with a more positive  $f_0$  Slope [ $\beta = 0.12$ ,  $SE = 0.08$ ], but this effect reversed for interrogative utterances [ $\beta = -0.62$ ,  $SE = 0.09$ ] (Fig. 6). In other words, declarative utterances with a more positive  $f_0$  slope were more likely to be interpreted as requests, while interrogatives with a more positive  $f_0$  slope were *less* likely to be interpreted as requests (i.e., they were more likely to be interpreted as yes/no questions).

Additionally, we found a main effect of Mean  $f_0$  [ $X^2(1) = 141.91$ ,  $p < .001$ ] (Fig. 7). Utterances with a higher Mean  $f_0$  were more likely to be interpreted as requests, regardless of Form [ $\beta = 0.47$ ,  $SE = 0.07$ ]. Finally, we found an interaction between Number of Voiced Frames and Form [ $X^2(1) = 14.43$ ,  $p < .001$ ], and a main effect of Number of Voiced Frames [ $X^2(1) = 29.5$ ,  $p < .001$ ]. Declarative utterances with a larger number of voiced frames were more likely to be interpreted as requests [ $\beta = 0.42$ ,  $SE = 0.06$ ], but this effect reversed for interrogatives [ $-0.37$ ,  $SE = 0.1$ ].

<sup>7</sup>By-subject random slopes for a given predictor mean that the statistical model calculates within-subject relationships between that predictor and the dependent variable.

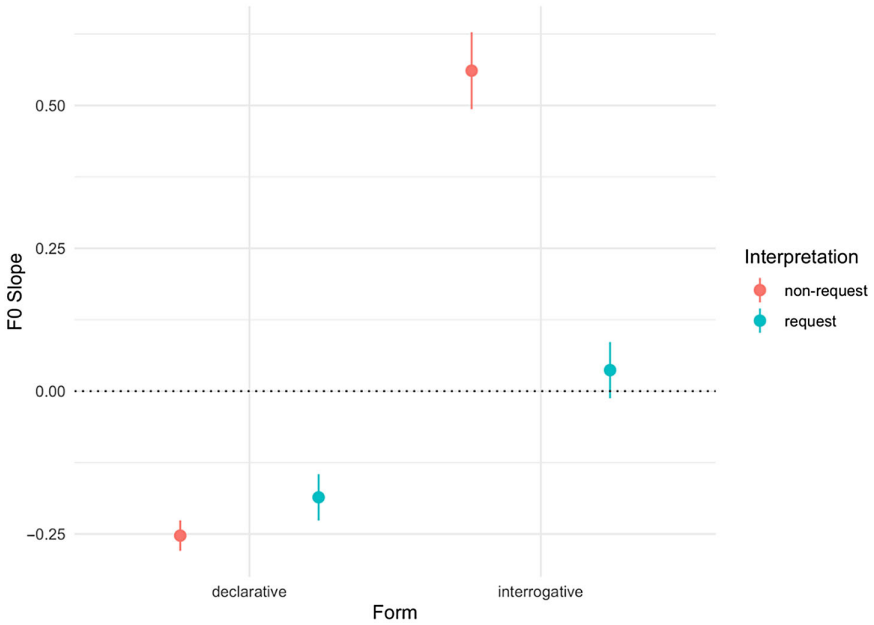


Figure 6. Z-scored  $f_0$  slope by Form and Interpretation.

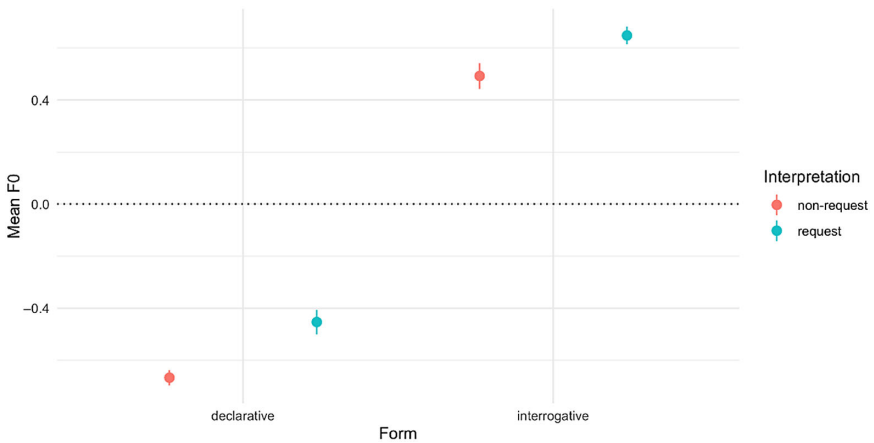


Figure 7. Z-scored mean  $f_0$  slope by Form and Interpretation.

### 5. GENERAL DISCUSSION AND CONCLUSION

In this research, we asked whether listeners use the prosodic cues produced by speakers to disambiguate the meaning of French IR expressions. To answer this question, we collected native speakers' spoken utterances and presented these to participants in a request classification task.

We found, first, that, several acoustic features reliably predicted a speaker's original intent. For modal interrogatives, intent was predicted by  $f_0$  slope (requests had a less positive  $f_0$  slope). For negative state remarks, intent was predicted by mean  $f_0$  (requests had a higher mean  $f_0$ ) and number of voiced frames (requests were longer on average). Second, in the perception experiment, listeners were able to detect a speaker's original intent at a rate above chance, indicating that they relied on some form of prosodic information to make their decision. Specifically, listeners' decisions were predicted by mean  $f_0$ ,  $f_0$  slope, and number of voiced frames. In each case, their pragmatic interpretations mirrored the relationship between prosodic features and intent. For example, modal interrogatives with a more positive  $f_0$  slope were more likely to be interpreted as requests (see Figure 4).

Importantly, these results bear several similarities to the findings of Trott et al.'s (2019) study conducted on English. For example, in both studies, the intent and interpretation of modal interrogatives was predicted by  $f_0$  slope: utterances with a more positive  $f_0$  slope – consistent with a rising intonation – were less likely to be *intended* and *interpreted* as requests. Similarly, in both studies, the intent of negative state remarks was predicted by the number of voiced frames: longer utterances were more likely to be intended as requests. Of course, English and French are similar languages along a number of dimensions, such as, to name but a few, their Indo-European origin and their lexicon. An interesting question for future work would be to ask whether these consistent relationships between intent (or interpretation) and specific prosodic features are obtained in other, more distant languages. We would assume that, if prosody and grammar are linked in some way, then the extent to which languages share grammatical constructions might influence the degree to which we observe similar prosodic outcomes. That is, assuming a given language has a modal interrogative construction that can be used as a request or a yes/no question, do speakers consistently use more positive  $f_0$  slopes when producing the utterance as a yes/no question? More generally, researchers could consider the systematicity with which three crucial factors correlate across languages: the *form* of an utterance (e.g., modal interrogatives), the *prosodic contours* with which it is produced (e.g., rising intonation), and its *pragmatic intent* (e.g., a yes/no question vs. a request).

An important issue that cannot be settled on the basis of the present results alone is whether the specific prosody we found for sentences performing a request are related to the request act in itself, or to the more general fact that the sentence has a noncanonical or indirect interpretation (for a discussion, see e.g., Brunetti et al. 2021). This question is difficult to answer. It is possible that the answer is “both, and it depends on the feature and the construction”. For example, in the English data that serve as a starting point for the present work, modal interrogatives have the default interpretation as a request – so the “prosodic marking” would actually serve as a deviation from a request and towards a yes/no question. And, critically, in this case, those prosodic markings are actually consistent with yes/no question prosody. We observe a similar phenomenon in our French data, in the sense that the  $f_0$  slope is more exaggerated for the yes/no question interpretation (See e.g., Figures 3 and 4). Furthermore, rising



intonation is there for both literal/IR uses. It is possible that there is an interaction between the default pragmatic interpretation of a particular form (e.g., modal interrogatives = request) and the degree to which particular prosodic features are emphasized. A related issue is that the prosodic features analysed in the present study may not necessarily correspond to what listeners actually attend to when interpreting pragmatically ambiguous utterances (Trott et al. 2022). Rather, it is possible that these features correlate with other, finer-grained cues to pragmatic meaning. For example,  $F_0$  slope and duration might be a proxy for the focal stress put on the verb *pouvoir* in non-inverted modal interrogatives. It will therefore be important to adopt, as Trott et al. (2022) suggest, a more fine-grained approach to the prosody of ISAs in future research.

A limitation of the present study is that, in our production task, participants were asked to read aloud pairs of utterances, one of which as a request and the other one e.g., as a statement/request for information. It remains to be investigated whether participants would produce similar prosodic cues when asked to read ambiguous utterances randomly instead of one pair after the other. It is difficult, of the basis of the present production data, to generalize our results to real life interactions and claim that people use prosodic information to differentiate indirect requests and their literal counterparts in authentic situations of speech. We thus believe that our study constitutes the starting point of a wider research program that first aims to establish the viability of an effect before assessing the extent to which this effect is reliable different conditions.

SA types and (im)politeness are not completely independent. For instance, insofar as a command is more face-threatening than a request (Brown & Levinson 1987), it is also more likely to be perceived as (more) impolite than a request. To ensure that the respective effects of (im)politeness differences and SA interpretations on prosody can be disentangled, we did not provide any (social) information about face-threat or politeness in the instructions of our production study and of our experiment. Of course, it is possible that the degree of face-threat associated with each utterance and the implied request did vary across stimuli, but we did not manipulate or otherwise confound the context of our stimuli. Further work will be necessary to document the prosodic features that signal the degree of (im)politeness associated with different request utterances.

Despite the insights contributed by our findings, very little is known about the exact circumstances under which speakers would use these prosodic cues, such as the social aspects of conversational interaction. It also remains to be seen whether these cues signal indirect interpretations or are simply contrastive markers that receive their SA interpretation in context. In addition, with the exception of Hellbernd & Sammler's (2016) work on one-word utterances, little empirical research has been devoted to the prosodic features specific to less grammaticalized SA types, such as advices, warnings, and offers. This is problematic, as the prosodic features observed in our study and in Trott et al.'s (2019) are not necessarily specific to indirect communication nor to the SA of requesting, as distinct from the DSAs of questioning or stating. In addition, we are aware that the prosodic cues associated with the indirect interpretations of IR constructions

have an emotional component that is not necessarily unique to these requests, but can also be found in other SA types. Future research should also address the role of prosody on the interpretation of imperative sentences, which can be used in the performance of a variety of SAs.

**Data Availability.** The present study has been preregistered on the Open Science Framework platform. Our results will be available upon request.

**Competing Interests.** The authors declare that they have no competing interests of any kind associated with the present research manuscript.

## References

- Banuazizi, A., & C. Creswell.** 1999. Is that a real question? final rises, final falls, and discourse function in yes-no question intonation. *CLS* 35: 1–14.
- Bates, D., Maechler, M., Bolker, B., & S. Walker.** 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1): 1–48. DOI: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)
- Beach, C. M.** 1991. The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language* 30(6): 644–663. DOI: [10.1016/0749-596X\(91\)90030-N](https://doi.org/10.1016/0749-596X(91)90030-N)
- Beysade, C., & J.-M. Marandin.** 2006. *From Complex to Simple Speech Acts: a Bidimensional Analysis of Illocutionary Forces*. Brandial, Postdam.
- Blaauw, E.** 1994. The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech. *Speech Communication* 14 (4): 359–375.
- Brown, P., & S. Levinson.** 1987. Politeness. *Universals in language usage*. Cambridge: Cambridge University Press.
- Brunetti, L., Yoo, H., Tovena, M. L., and R. Albar,** 2021. French reason-comment (how) questions: A view from prosody. In A. Trotzke and X. Villalba (eds.) *Expressive Meaning Across Linguistic Levels and Frameworks*, 248–278. Oxford University Press.
- Bryant, G. A., & J. E. Fox Tree.** 2002. Recognizing verbal irony in spontaneous speech. *Metaphor and Symbol* 17(2): 99–117. DOI: [10.1207/S15327868MS1702\\_2](https://doi.org/10.1207/S15327868MS1702_2)
- Bryant G. A., & J. E. Fox Tree.** 2005. Is there an Ironic Tone of Voice? *Language and Speech* 48(3): 257–277. DOI: [10.1177/00238309050480030101](https://doi.org/10.1177/00238309050480030101)
- Celle, A. & M. Péliissier.** 2021. Surprise questions in spoken French. To appear in *Linguistics Vanguard*.
- Cheang, H. S., & M. D. Pell.** 2008. The sound of sarcasm. *Speech Communication* 50(5): 366–381. DOI: [10.1016/j.specom.2007.11.003](https://doi.org/10.1016/j.specom.2007.11.003)
- Coveney, A.** 2011. L'interrogation directe. *Travaux de linguistique* 63 : 112–145. <https://doi.org/10.3917/tl.063.0112>
- Delais-Roussarie, E., & C. Beysade.** 2019. Information seeking vs. rhetorical questions: from gradience to categoricity. *Phonetics and Phonology in Europe (PaPE)*, 17–19 juin, University of Salento, Lecce, Italie.
- Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., Jun, S.-A., Martin, P., Meisenburg, T., Rialland, A., Sichel-Bazin, R., & Yoo, H.-Y.** 2015. Intonational phonology of French: Developing a ToBI system for French. In Sonia Frota and Pilar Prieto (Eds.), *Intonation in Romance*. Oxford Scholarship Online.
- Delattre, P.** 1966. Les dix intonations de base du français. *The French Review* 40 (1): 1–14.
- Di Cristo, A.** 2016. *Les musiques du français parlé: Essais sur l'accentuation, la métrique, le rythme, le phrasé prosodique et l'intonation du français contemporain*. Berlin & Boston: De Gruyter Mouton.
- Guaïtella, I.** 1999. Rhythm in speech: What rhythmic organizations reveal about cognitive processes in spontaneous speech production versus reading aloud. *Journal of Pragmatics* 31 (4): 509–523.
- Gunlogson, C.** 2001. *True to Form: Rising and Falling Declaratives as Questions in English*. New York: Routledge.
- Gunlogson, C.** 2008. A question of commitment. *Belgian Journal of Linguistics* 22 (1): 101–136. DOI: [10.1075/bjl.22.06gun](https://doi.org/10.1075/bjl.22.06gun)

- Hedberg, N., Sosa, J. M., & E. Görgülü. 2014. The meaning of intonation in yes-no questions in American English: A corpus study. *Corpus Linguistics and Linguistic Theory* 13(2): 321–368.
- Hellbernd, N., & D. Sammler. 2016. Prosody conveys speaker's intentions: Acoustic cues for speech perception. *Journal of Memory and Language* 88: 70–86. DOI: [10.1016/j.jml.2016.01.001](https://doi.org/10.1016/j.jml.2016.01.001)
- Holm, S. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of statistics* 6(2): 65–70.
- Jadoul, Y., Thompson, B., & De Boer, B. 2018. Introducing Parselmouth: a Python interface to Praat. *Journal of Phonetics* 71: 1–15.
- Michelas, A., Portes C., & M. Champagne-Lavau. 2015. When pitch Accents Encode Speaker Commitment: Evidence from French Intonation. *Language and Speech* 59(2): 266–293. DOI: [10.1177/0023830915587337](https://doi.org/10.1177/0023830915587337)
- Nakamura, M., Iwano, K., & Furu, S. 2008. Differences between acoustic characteristics of spontaneous and read speech and their effects on speech recognition performance. *Computer Speech & Language* 22 (2): 171–184.
- Nickerson, J. S., & J. Chu-Carroll. 1999. Acoustic-prosodic disambiguation of direct and indirect speech acts. Proceedings of the 14<sup>th</sup> International Congress of Phonetic Sciences, San Francisco. International Phonetic Association, 1309–1312.
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & C. Fong. 1991. The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 90: 2956–2970. DOI: [10.1121/1.401770](https://doi.org/10.1121/1.401770)
- R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Roche, J. M., Morgan, S. D., Fissel Brannick, S., & K. Bryndel. 2019. Acoustic correlates of female confidence: A production and comprehension study. *The Journal of the Acoustical Society of America* 145(6): 3410–3426. DOI: [10.1121/1.5109661](https://doi.org/10.1121/1.5109661)
- Rossi, M., Hirst, D., & A. Di Cristo. 1981. Continuation and question. In Rossi, M., Di Cristo, A., Hirst, D., Martin, P., & Y. Nishinuma (Eds.): *L'intonation: de l'acoustique à la sémantique*. Paris, Klincksieck, 149–177.
- Ruytenbeek, N. 2017. The comprehension of indirect requests: Previous work and future directions. In Depraetere, I. & R. Salkie (Eds.). *Semantics and Pragmatics. Drawing a Line*, Cham: Springer, 293–322.
- Ruytenbeek, N. 2021. *Indirect speech acts*. Cambridge: Cambridge University Press.
- Safarova, M., Muller, P., & L. Prevot. The Discourse Function of Final Rises in French dialogues. *Ninth workshop on the semantics and pragmatics of dialogue*, June 2005, Nancy, France.
- Shriberg, E., Stolcke A., Jurafsky, D. et al. 1998. Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech? *Language and Speech* 41(3–4): 443–492. DOI: [10.1177/002383099804100410](https://doi.org/10.1177/002383099804100410)
- Smith, C. L. 2002. Prosodic finality and sentence type in French. *Language and Speech* 45: 141–178. DOI: [10.1177/00238309020450020301](https://doi.org/10.1177/00238309020450020301)
- Snedeker, J., & J. Trueswell. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 48 (1): 103–130. DOI: [10.1016/S0749-596X\(02\)00519-3](https://doi.org/10.1016/S0749-596X(02)00519-3)
- Torreira, F., & E. Valtersson 2015. Phonetic and Visual Cues to Questionhood in French Conversation. *Phonetica* 72: 20–42.
- Trott, S., Reed, S., Kaliblotzky, D., Ferreira, V., & Bergen, B. 2022. The role of prosody in disambiguating English indirect requests. *Language and Speech*, April 2022. DOI: [10.1177/00238309221087715](https://doi.org/10.1177/00238309221087715).
- Trott, S. T., S. Reed, V. Ferreira, & B. Bergen. 2019. Prosodic cues signal the intent of potential indirect requests. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*.
- Valtersson, E., & Torreira, F. 2014. Rising intonation in spontaneous French: how well can continuation statements and polar questions be distinguished? 7<sup>th</sup> International Conference on Speech Prosody, *Speech Prosody* 7, 785–789.
- Wall, K. 1985. Les questions dans la langue parlée. Usage actuel et problèmes pédagogiques, I et II. *Moderna språk* 85 (79): 233–238 and 309–317.
- Ward, N. G. 2019. *Prosodic Patterns in English Conversation*. Cambridge: Cambridge University Press. DOI: [10.1017/9781316848265](https://doi.org/10.1017/9781316848265)

## Appendices

### Appendix 1. List of recorded utterances

Recorded utterances (French)	English translation
<i>Tu peux ouvrir cette fenêtre ?</i>	Can you open this window?
<i>Tu peux porter cette valise ?</i>	Can you carry this suitcase?
<i>Tu peux lire cette facture ?</i>	Can you read this invoice?
<i>Tu peux bouger cette armoire ?</i>	Can you move this cabinet?
<i>Tu peux ranger ces bouquins ?</i>	Can you store these books?
<i>Tu peux suivre cet itinéraire ?</i>	Can you follow this itinerary?
<i>Ma soupe est froide.</i>	My soup is cold.
<i>Mon téléphone est cassé.</i>	My phone is broken.
<i>Mon verre est vide.</i>	My glass is empty.
<i>Mon ordi est hors service.</i>	My laptop is out of order.
<i>Mes mains sont gercées.</i>	My hands are cracked.
<i>Mes cheveux sont trempés.</i>	My hair is soaked.

### Appendix 2. Instructions for the production task

#### Consignes

- **Utilisez le micro de votre ordinateur portable (PAS de votre téléphone).**
- **Réalisez vos enregistrements au calme,** dans un local dont vous aurez fermé la porte et les fenêtres, en vous installant confortablement en position assise.
- **Gardez la même distance** (environ **30 centimètres**) entre votre bouche et votre micro tout au long des enregistrements.
- **Gardez l'intensité de votre voix** constante d'une phrase à l'autre.

#### Tâche

- Vous allez prononcer au total 12 phrases, avec à chaque fois 2 enregistrements par phrase. Cela fait donc 24 fichiers audio au total (.wav ou.mp3).
- Pour chacune des phrases à enregistrer, nous vous donnerons un contexte afin que vous puissiez mieux vous imaginer dans quelles circonstances ces phrases pourraient être dites.

Exemple 1 : Prononcez la phrase : *Tu peux éteindre la télé ?*

- a) Comme une **question** : vous posez à un ami une question sur sa capacité à éteindre la télévision, ce qui correspond au sens : « Es-tu capable d'éteindre la télé ? »

- b) Comme une **demande** : vous demandez à un ami d'éteindre la télévision, ce qui correspond au sens de : « S'il te plait, éteins la télé ».

Exemple 2 : Prononcez la phrase : *Il fait froid ici.*

- a) Comme une **affirmation** : vous constatez simplement qu'il fait froid dans la pièce, vous ne demandez pas à quelqu'un d'agir en conséquence. Ce qui correspond au sens : « Je trouve qu'il fait froid ici ».
- b) Comme une **demande** : vous dites cela pour que votre ami agisse en allumant le chauffage, ce qui correspond au sens de : « S'il te plait, allume le chauffage ».
- **Attention à l'ordre** : l'ordre de la prononciation comme question/demande ou affirmation change à chaque nouvelle phrase.
  - **Renommez vos fichiers un par un comme indiqué en vert** : [nom du fichier : ...]
  - **N'hésitez pas à faire plusieurs essais** avant d'enregistrer vos phrases.
  - En cas de nécessité, vous pouvez me contacter: [nicolasruytenbeek@gmail.com](mailto:nicolasruytenbeek@gmail.com).

Voici les phrases à prononcer par deux fois (ordre à respecter)

1) *Tu peux ouvrir cette fenêtre ?*

Contextes :

- a) Comme une **question** : votre ami s'est cassé le bras et vous voulez savoir s'il est capable d'ouvrir la fenêtre. [nom du fichier : 1question]
- b) Comme une **demande** : vous avez trop chaud et vous demandez à votre ami d'ouvrir la fenêtre. [nom du fichier : 1demande]

2) *Tu peux porter cette valise ?*

Contextes :

- a) Comme une **demande** : vous avez une volée d'escaliers à franchir avec vos deux valises et vous demandez à votre ami d'en porter une. [nom du fichier : 2demande]
- b) Comme une **question** : la valise de votre ami est très lourde et vous voulez savoir s'il est capable de la porter. [nom du fichier : 2question]

3) *Tu peux lire cette facture ?*

Contextes :

- a) Comme une **question** : le texte de la facture de votre ami est presque effacé ; vous voulez savoir s'il est capable de le lire. [nom du fichier : 3question]
- b) Comme une **demande** : vous avez reçu une facture, et vous demandez à votre ami de vous en lire le contenu. [nom du fichier : 3demande]

4) *Tu peux bouger cette armoire ?*

Contextes :

- a) Comme une **demande** : vous préparez votre déménagement et vous demandez à votre ami (qui est costaud) de bouger l'armoire. [nom du fichier : 4demande]
- b) Comme une **question** : votre ami prépare son déménagement et vous voulez savoir s'il est capable de bouger son armoire. [nom du fichier : 4question]

5) *Tu peux ranger ces bouquins ?*

Contextes :

- a) Comme une **question** : votre ami s'est foulé le poignet et vous voulez savoir s'il est capable de ranger ses bouquins. [nom du fichier : 5question]
- b) Comme une **demande** : votre ami a fini de lire plusieurs de vos bouquins et vous lui demandez de les ranger. [nom du fichier : 5demande]

6) *Tu peux suivre cet itinéraire ?*

Contextes :

- a) Comme une **demande** : votre ami vous reconduit chez vous et vous lui demandez de suivre l'itinéraire indiqué par le GPS. [nom du fichier : 6demande]
- b) Comme une **question** : votre ami n'est pas encore habitué à son nouveau GPS et vous voulez savoir s'il est capable de suivre l'itinéraire proposé. [nom du fichier : 6question]

7) *Ma soupe est froide.*

Contextes :

- a) Comme une **affirmation** : vous êtes au restaurant avec votre ami et vous lui faites savoir que votre soupe est froide. [nom du fichier : 7affirmation]
- b) Comme une **demande** : vous êtes au restaurant avec votre ami et vous dites cela pour que le serveur vous apporte une soupe chaude. [nom du fichier : 7demande]

8) *Mon téléphone est cassé.*

Contextes :

- a) Comme une **demande** : vous êtes en rue avec votre ami et vous dites cela pour qu'il vous prête son téléphone. [nom du fichier : 8demande]
- b) Comme une **affirmation** : vous appelez votre ami depuis un téléphone fixe et vous lui faites savoir que votre téléphone est cassé. [nom du fichier : 8affirmation]

9) *Mon verre est vide.*

Contextes :

- a) Comme une **affirmation** : vous êtes au restaurant avec votre ami et vous lui faites savoir que votre verre de vin est vide. [nom du fichier : 9affirmation]
- b) Comme une **demande** : vous êtes au restaurant avec votre ami et vous dites cela pour que le serveur remplisse votre verre. [nom du fichier : 9demande]

10) *Mon ordi est hors service.*

Contextes :

- a) Comme une **demande** : vous vous rendez chez votre ami et vous dites cela pour qu'il vous prête son ordinateur. [nom du fichier : 10demande]
- b) Comme une **affirmation** : vous vous promenez avec votre ami et vous lui faites savoir que votre ordinateur est hors service.[nom du fichier : 10affirmation]

11) *Mes mains sont gercées.*

Contextes :

- a) Comme une **affirmation** : vous avez jardiné toute la matinée, et en téléphonant à votre ami vous lui faites savoir que vos mains sont gercées. [nom du fichier : 11affirmation]
- b) Comme une **demande** : vous avez jardiné toute la matinée, et vous dites cela pour que votre ami vous passe la crème pour les mains. [nom du fichier : 11demande]

12) *Mes cheveux sont trempés.*

Contextes :

- a) Comme une **demande** : vous avez marché sous la pluie, et en rentrant chez vous, vous dites cela pour que votre ami vous passe le sèche-cheveux. [nom du fichier : 12demande]
- b) Comme une **affirmation** : vous avez marché sous la pluie et, en arrivant dans le local de cours, vous faites savoir à votre ami que vos cheveux sont trempés. [nom du fichier : 12affirmation]

Nicolas Ruytenbeek gratefully acknowledges the research grant BOF.PDO. 2019.0010.01 from the Special Research Fund at Ghent University (2019–2022). The authors are very grateful to the anonymous reviewers and to the editor for their useful comments and suggestions.

---

**Cite this article:** Ruytenbeek N, Bergen B, and Trott S (2023). Prosody and speech act interpretation: The case of French indirect requests. *Journal of French Language Studies* 33, 103–125. <https://doi.org/10.1017/S0959269522000254>