

REGISTERED REPORT

Do prosodic cues convey intent directly or through contrastive marking? A study of French indirect requests

Nicolas Ruytenbeek, Department of Linguistics, Faculty of Arts, KU Leuven, Sint-Jacobsmarkt 49–51, Antwerp, Belgium; Laboratoire Savoirs, Textes, Langage (STL), University of Lille, UMR 8163 – CNRS, Lille, France, Nicolas.Ruytenbeek@kuleuven.be

Sean T. Trott, Department of Cognitive Science, UC San Diego, 9500 Gilman Dr., La Jolla, CA 92093, United States of America, trott.sean@gmail.com

This study investigates how prosody contributes to the interpretation of French indirect requests. We ask whether prosodic cues directly map onto speech acts (Direct Mapping view) or whether they primarily serve as contrastive markers, signaling a departure from the most likely interpretation (Contrastive Marking view). Four interrogative constructions were examined, each compatible with both a request and a yes/no question reading: modal interrogatives (*Tu peux fermer la fenêtre ?* ‘Can you close the window?’), non-modal interrogatives (*Tu fermes la fenêtre ?* ‘Are you closing the window?’), and their counterparts with *est-ce que* (*Est-ce que tu peux fermer la fenêtre ?* ‘Can you close the window?’; *Est-ce que tu fermes la fenêtre ?* ‘Are you closing the window?’). Norming studies with 320 French speakers established baseline request probabilities for these forms, confirming that modal interrogatives strongly favor a request interpretation, while non-modal forms – especially with *est-ce que* – were less likely to be interpreted as such. A production study with 8 native speakers elicited utterances intended as requests or questions, and acoustic features (mean F0, F0 slope, duration) were analyzed. A perception study with 280 listeners then tested whether these prosodic cues guided interpretation. Our results reveal a dual pattern. F0 slope consistently distinguished requests from questions across constructions, supporting the Direct Mapping view. By contrast, mean F0 and duration interacted with constructional features: their predictive value varied, depending on the presence of a modal or *est-ce que*, in line with the Contrastive Marking view. Together, these findings suggest that prosody plays a non-uniform role in speech act recognition. Some features act as stable signals of intent, while others are sensitive to the construction’s default interpretation. This work advances models of prosody and pragmatics by showing that prosodic cues can simultaneously function as direct markers of meaning and as signals of deviation from canonical interpretations.



1. Introduction

Linguistic ambiguity is pervasive and can hinder successful communication, especially in settings where facial and gestural cues are missing, such as online messages and phone conversations. In the latter case, a possible way for speakers to avoid being misunderstood is to use prosodic features to clarify how they intend their utterances to be taken by their addressees. In line with this idea, experimental research has shown that speakers indeed make use of prosodic cues to enable their addressees to disambiguate syntactically ambiguous utterances (Beach, 1991; Price et al., 1991; Snedeker & Trueswell, 2003) and potentially ironic messages (e.g., Bryant & Fox Tree, 2002, 2005; Cheang & Pell 2008). Previous work also revealed that prosody enables the disambiguation of speech acts. For example, intonational contours are reliable cues when it comes to interpreting one-word utterances, such as *beer*, as the speech acts of criticism, doubt, naming, suggestion, warning and wish (Hellbernd & Sammler, 2016). Specific acoustic features also appear to aid hearers in detecting whether a speaker is making an *indirect request* or not (Ruytenbeek et al., 2023; Trott et al., 2023).

In the present study, we address the general question of *whether* and *how* prosody plays a role in the production and interpretation of indirect requests in French. In particular, we focus on whether prosodic cues convey speech act information directly, or whether their primary function is as contrastive markers – signaling a deviation from the canonical interpretation of an utterance. We specifically investigate this research question in the case of requests, a speech act type that is routinely performed in an indirect manner. Towards this end, we have selected four French interrogative constructions that can be used both as literal questions and as indirect requests, i.e., non-inverted interrogatives with or without the modal *pouvoir* ('can') and with or without the interrogative particle *est-ce que* ('is it the case that').

2. Literature review

2.1 Past work on prosody in pragmatics

The role of prosody in pragmatic disambiguation has been explored across various types of utterances and speech acts. In this section, we describe research focusing on indirect requests and irony specifically, as both are cases of pragmatic ambiguity. In each case, the same construction can be used for distinct meanings, and the question of *which* meaning a speaker intended may be informed by the prosodic cues present in the utterance. We therefore believe that such work can inform the hypotheses that we specifically test in our present research.

Past work on both English (Trott et al., 2019; Trott et al., 2023) and French (Ruytenbeek et al., 2023) has identified several acoustic cues that reliably predict whether or not a speaker is making a request. In each language, two sets of grammatical constructions were tested: modal interrogatives (e.g., *Can you open that window?* in English and the corresponding *Tu peux ouvrir*

cette fenêtre? in French) and negative state remarks (e.g., *My office is really hot*). An important difference between the English modal interrogatives and the French modal interrogatives in these two studies is that the latter are characterized by non-inverted word order. This is why they were called “non-inverted interrogatives” by Ruytenbeek et al. (2023, p. 108). These non-inverted interrogatives have replaced, to a large extent, inverted interrogatives in contemporary spoken French (see Wall, 1985 for empirical evidence).

Modal interrogatives and negative state remarks differ in two important ways. First, although both constructions license a request interpretation, modal interrogatives are more likely to be interpreted as requests than negative state remarks, overall (Trott et al., 2023). Second, the *literal* speech act conveyed by each construction is also different. A negative state remark, interpreted literally, is an assertion; a modal interrogative, interpreted literally, is a yes/no question. That is, each construction has a different set of *candidate interpretations*: a negative state remark can be interpreted as a request or an assertion, while a modal interrogative can be interpreted as a request or a yes/no question about ability.

Accordingly, past work also identified distinct acoustic cues predicting the original intent of an utterance, as well as how it was interpreted.¹ Modal interrogatives were disambiguated by F0 slope, i.e., the degree to which pitch rose, fell, or stayed constant over time; this was true in both English (Trott et al., 2023) and French (Ruytenbeek et al., 2023). Specifically, a more positive F0 slope was associated with utterances intended (and interpreted) as yes/no questions, and less associated with the request interpretation. This is consistent with past work on English, suggesting that yes/no questions have a rising pitch contour (Banuazizi & Creswell, 1999; Hedberg et al., 2014). In contrast, negative state remarks were disambiguated by the number of voiced frames, i.e., the duration of the utterance. In both English (Trott et al., 2023) and French (Ruytenbeek et al., 2023), longer utterances were more likely to be intended (and interpreted) as requests.

Regarding the relationship between speech act information and the prosody of French interrogatives, Beyssade and Delais-Roussarie’s (2022) comparison between the information seeking and rhetorical question uses of polar interrogatives with *est-ce que* showed that the former reading can be associated with a rising, a high plateau or a rising-falling contour. In another study, Delais-Roussarie & Beyssade (2019) had provided evidence that the rhetorical question use of *wh-* (e.g., *Qui mange des épinards?* ‘Who eats spinach?’) and polar interrogatives (e.g., *Il pleut?* ‘It’s raining?’) results in increased utterance duration. In the same vein, Celle and Pélissier (2022) documented the prosody of the information seeking reading and the surprise question reading of *Qu’est-ce que...?* interrogatives, a construction that can be used to express criticism (Coveney and Dekhissi, 2017). They found that, as surprise questions, *Qu’est-ce que...?*

¹ Or, in some cases, a distinct *relationship* between a particular acoustic cue and intent across different grammatical constructions.

interrogatives are prosodically differentiated using slower speech rate (i.e., longer mean duration of the first syllable); they are also less frequently produced with a rising final intonational contour.

There is also evidence that the comprehension of irony depends, in part, upon prosodic information. Several studies have shown that speakers of different languages, i.e., English, German and Spanish, use prosodic cues to signal ironic meanings and that comprehenders rely on the same cues when interpreting ironic utterances (e.g., Bryant, 2010; Cheang and Pell, 2008; Padilla, 2012; Rockwell, 2000; Scharrer et al., 2011). When it comes to French utterances, the cues that have been found to be reliably associated with ironic meanings are pitch raising or lowering, higher pitch variability, higher intensity, and increased utterance duration (for a discussion, see González-Fuente et al., 2016; Loevenbruck et al., 2013). For instance, Loevenbruck et al. (2013) investigated the prosody of ironic vs. literal utterances of one and the same sentence. They found that mean F0 and average duration are higher for ironic utterances of e.g., *Il est en pleine forme* ('He is in great shape') compared to their literal counterparts. They also reported on a slight final rise in F0 for ironic utterances. Their findings for mean F0 and average duration were replicated in the first experiment by González-Fuente et al. (2016), who aimed to identify the individual contribution of the features of pitch range expansion, syllable lengthening and intonational contours to the overall ironic meaning of French utterances of the type *Nous sommes allées voir un film formidable* ('We went to see a fantastic film'). González-Fuente and colleagues also showed that, in the ironic utterances, the final word of the sentence was characterized by higher F0 variability.

Returning to the case of indirect requests, previous empirical work on the prosody of indirect requests has a major limitation: it cannot address the question of whether these distinct acoustic cues correspond directly to distinct *speech acts*, or whether the acoustic cues serve primarily to *mark* a deviation from the *default* (i.e., canonical) interpretation of an utterance.² This issue is discussed in greater detail in the section below.

2.2 How does prosody communicate intent?

As Trott et al. (2023, p. 137) note, a long-standing question in research on prosody and pragmatics is “whether particular prosodic features convey intent directly, or whether they function primarily as contrastive markers, which invite the listener to perform additional inference”. This question motivated past work on irony, e.g., on whether there is an “ironic tone of voice” (Attardo et al., 2003; Bryant and Fox Tree, 2005). It is also central to the question of prosody and indirect requests: does prosody convey a request meaning directly – i.e., is there a “requesting tone of voice” – or does it signal that a non-canonical speech act is intended? Here, we refer to these distinct hypotheses as the Direct Mapping view and the Non-canonical Marking view, respectively.

² For a negative state remark, the “default” interpretation may be a complaint or assertion; for some kinds of modal interrogative, the “default” interpretation may, in fact, be a request (Trott et al., 2023).

One preliminary answer to this question comes from observing the fact that distinct prosodic cues were predictive of the request interpretation across distinct grammatical constructions, i.e., modal interrogatives vs. negative state remarks (Ruytenbeek et al., 2023). This fact is inconsistent with the strongest form of the Direct Mapping view. However, this interpretation is complicated by the fact that the grammatical constructions tested (Ruytenbeek et al., 2023) differed along two dimensions. First, modal interrogatives are more likely to be interpreted as requests than negative state remarks; and second, modal interrogatives and negative state remarks have a distinct (though partially overlapping) set of *candidate interpretations*. Modal interrogatives are ambiguous between requests and yes/no questions, while negative state remarks are ambiguous between requests and assertions.

Thus, a weaker form of the Direct Mapping view might postulate that controlling for the set of candidate speech acts in question – e.g., a *request* vs. a *yes/no question* – the same set of acoustic cues will reliably and predictably discriminate between those speech acts. As noted in the section above, this weaker account is also partially consistent with past evidence, at least when it comes to modal interrogatives: the specific prosodic cues that predicted intent were also cues previously identified as conveying the *yes/no question* speech act, i.e., rising pitch contour (Ruytenbeek et al., 2023; Trott et al., 2023). A proponent of the Direct Mapping view might, therefore, claim that this weak version holds for modal interrogatives specifically.

Yet proponents of the Non-canonical Marking view might (justifiably) object that even this weaker conclusion is premature on the basis of the evidence. Namely, past work (Ruytenbeek et al., 2023; Trott et al., 2023) suffered from a confound between the salience of a request and which prosodic features were most plausibly associated with the candidate speech acts. For example, modal interrogatives have a higher probability of being interpreted as requests than as yes/no questions (Ruytenbeek et al., 2023; Trott et al., 2023). Thus, even though the prosodic features associated with the non-request speech act – a rising pitch contour – are, plausibly, meaningful, they might also merely signal a deviation from the canonical interpretation. Additionally, and unlike yes/no questions, the acoustic cues associated with the request speech act have less grounding in previous theoretical literature; this problem is further confounded by the fact that the request interpretation is the default for modal interrogatives. Evidence for even the weaker Direct Mapping view would need to demonstrate that those same features predict the intended speech act, regardless of whether the request or yes/no question meaning is the canonical one.

Here is a brief summary of the contrasting views and hypotheses outlined thus far:

- A. H_1 (Direct Mapping view): For interrogatives specifically, there is a predictable, consistent mapping between sets of acoustic features and intended speech act and interpretation (i.e., request vs. yes/no question).

B. H_2 (Non-canonical Marking view): Even among interrogatives, the mapping between acoustic features and intended speech act will depend not only on the intended speech act, but also on which candidate interpretation is most salient (i.e., request vs. yes/no request).

Of course, it is important to note that additional nuance exists within this hypothesis space, particularly in terms of how prosodic cues might influence pragmatic inference “in the wild”. For example, it is possible that there is *not* a direct mapping between acoustic features and a comprehender’s interpretation (thereby negating H_1), and *also* that comprehenders do not merely rely on prosodic marking to search for plausible alternatives to a salient interpretation (thereby negating H_2); instead, or in addition, comprehenders might also deploy context-specific cues that raise or lower the probability of a given interpretation. For the purposes of the current work, we focus on the narrower question of which cues signal which intents for a controlled set of utterances without explicit discourse context (see the norming studies and explanation below).

It is unclear which hypothesis is best supported by existing evidence. Answering this question is crucial for theories of speech production and the role that prosody plays in signaling intent more generally: does the production process “select” a prosodic contour directly on the basis of the intended speech act (as implied by the Direct Mapping view), or does it generate marked features dependent on whether the intended speech act is the canonical one (as implied by the Non-canonical Marking view)? Notably, in the latter case, the question of what is “canonical” is by no means trivial; the most salient interpretation of a given construction could vary, depending on the situational context, the relationship between speaker and addressee, previous turns in the discourse, and more. We operationalize *canonical* in an admittedly narrower sense, in terms of the interpretation that comes to mind most easily in the absence of specific contextual cues supporting any particular interpretation (see the norming studies below).³

These possibilities may also have distinct implications for the role of audience design in the speech production process (Ferreira, 2019). Specifically, the Direct Mapping view raises the possibility that speakers could generate a disambiguating prosodic pattern more “automatically”, without having to weigh the possible interpretations of the utterance. The Non-canonical Marking view presents a more complicated situation. On one account, speakers may need to first identify the default interpretation of their utterance, then produce a prosodic pattern that would convey a departure from that interpretation, i.e., engaging in what Ferreira (2019) would describe as *recurrent audience design*. Another possibility, however, is that speakers

³ In turn, this process of “calling to mind” various interpretations of an utterance could engage multiple mechanisms. For example, comprehenders might imagine scenarios in which that utterance is likely to have been produced, then reason about the likely intents of that utterance in those contexts. Alternatively, they might engage a more compositional, bottom-up process that involves inferring the likely meaning of an utterance from its constituent parts.

still engage disambiguating prosodic cues relatively automatically, by relying on contextually-appropriate linguistic knowledge that identifies the relative salience of different speech acts for a given utterance in a given context (thus eliminating the need to “simulate” a comprehender’s likely interpretation).

The hypotheses also have important implications for language processing and comprehension. If acoustic cues convey speech acts directly, it suggests that – in principle – pragmatic inference could be performed by mapping specific cues onto stored representations, i.e., constructions that map between prosodic features and speech acts. But if acoustic cues primarily serve to mark a deviation from the default, it suggests that the role of prosody is primarily as the catalyst to a more complex inferencing process. These possibilities have different theoretical implications for how speech act information is stored and accessed, as well as for the time course of pragmatic inference.

2.3 Present work

The present work aims to adjudicate between these hypotheses for a restricted set of French constructions – namely, grammatical constructions whose candidate interpretations are either request or yes/no question. This includes modal interrogatives, which have already been tested in previous work on French (Ruytenbeek et al., 2023), and which were also included in the current work. It also includes non-modal interrogatives (*Tu fermes cette fenêtre?*), as well as a modified version of both types of interrogatives with the interrogative particle *est-ce que* (ESQ) (*Est-ce que tu peux fermer cette fenêtre?* and *Est-ce que tu fermes cette fenêtre?*). These are described in **Table 1**; note that the values in the $p(\text{request})$ column were obtained from the norming studies, which we describe below.

Table 1: Overview of the constructions investigated in the present research, including request priors from the norming studies.

Construction	ESQ?	Modal?	Example	English Translation	P(request)
Modal interrogative	No	Yes	Tu peux porter cette valise ?	Can you carry this suitcase?	0.94
Non-modal interrogative	No	No	Tu portes cette valise ?	Are you carrying this suitcase?	0.56
Modal interrogative (<i>est-ce que</i>)	Yes	Yes	Est-ce que tu peux porter cette valise ?	Can you carry this suitcase?	0.98
Non-modal interrogative (<i>est-ce que</i>)	Yes	No	Est-ce que tu portes cette valise ?	Are you carrying this suitcase?	0.30

Importantly, each of these constructions plausibly licenses both a request and a yes/no question interpretation. But, equally importantly, the relative salience of those interpretations differs across the constructions. Past work (Ruytenbeek et al., 2023) has already shown that a request interpretation is most salient for the French modal interrogative construction. Further, the norming studies described below have confirmed that a request interpretation is less likely for constructions without a modal (e.g., *Tu portes cette valise?*); within the category of non-inverted interrogatives, the presence of *est-ce que* was associated with a lower probability of a request interpretation.

Thus, these constructions provide an ideal contrast for addressing the outstanding question from above, as H_1 and H_2 make distinct predictions.

- C. H_1 predicts that the same acoustic cues will predict speech act (request vs. yes/no question), regardless of the construction used.
- D. H_2 predicts that different acoustic cues will predict speech act (request vs. yes/no question), depending on which construction is used.

Put another way: if H_1 is correct, then the results for the new constructions tested should resemble past findings for the non-inverted modal interrogative construction, i.e., *Tu peux VP?* (Ruytenbeek et al., 2023). When intended as a yes/no question, non-modal interrogatives (with and without *est-ce que*) should have a rising pitch contour; when intended as a request, they should not.

But if H_2 is correct, then the results for non-modal interrogatives will look different from those found in past work (Ruytenbeek et al., 2023). Critically, there are multiple ways in which they could be different. One possibility is that non-modal interrogatives are not discriminated by any of the acoustic cues tested. Another possibility is that the same acoustic cues are predictive, but in different or even opposite ways: for example, it is possible that non-modal interrogatives intended as yes/no questions will not have a rising pitch contour, but that non-modal interrogatives intended as requests will (i.e., the opposite of the findings for modal interrogatives).

To sum up: the hypotheses make distinct empirical predictions about which prosodic cues will predict intent (and interpretation) across the grammatical constructions tested. This will allow us to adjudicate between the hypotheses on the basis of the study's results. While the constructions tested represent only a narrow slice of the broader prosody/pragmatics interface, resolving this debate has important theoretical implications not only for speech production, but also for language processing and pragmatic inference.

2.3.1 Est-ce que interrogatives

The results of the present study will also expand our knowledge of the prosody of French polar questions prefixed by the interrogative particle or marker *est-ce que* (cf. Behnstedt, 1973; Rooryck, 1994). The phrase *est-ce que*⁴ has sometimes been called an interrogative morpheme, the phrase

⁴ See Dendale (2025) for an overview of the studies addressing the phrase *est-ce que*.

having lost its compositional meaning in the course of grammaticalization (Druetta, 2003, p. 27) and it has been considered as a morphological marker with the force of a request for information (see Dupras, 2018, pp. 108–109 for a synthesis). However, there is no one-to-one relationship between polar interrogatives with *est-ce que* and the speech act of requesting information. To begin with, an utterance of the interrogative polar construction prefixed by *est-ce que*, illustrated in (1), has different implications than its counterpart lacking *est-ce que* (2). For instance, (1) cannot be used for seeking confirmation (Déprez et al., 2012).

(1) #Est-ce que tu sors ce soir, je présume ?
 ESQ 2.NOM get.out this evening I.NOM suppose
 ‘Are you going out tonight, I suppose?’

(2) Tu sors ce soir, je présume ?
 2-SG.NOM get.out this evening I.NOM suppose
 ‘You’re going out tonight, I suppose?’

In addition, scholars’ intuitions regarding the directive uses of this type of construction are not unanimous. For instance, while Coveney (2011, p. 139) claims that there is a pragmatic constraint that prevents *Est-ce que...?* questions from being used as requests for action, Reinhardt (2019, pp. 198–199) provides corpus-based evidence such as (3).

(3) Anthony, est-ce que tu peux me prêter tes lunettes, s’il te plaît ?
 Anthony, ESQ you.NOM can.2SG.PST I.DAT lean 2.POSS.PL glasses-PL please
 ‘Anthony, can I borrow your glasses, please?’

Reinhardt’s data also point towards a correlation between the use of *Est-ce que tu peux* + infinitive and the addition of the politeness marker *s’il te plaît* (‘please’), which suggests that polar interrogatives with *est-ce que* would convey a higher degree of politeness compared to their counterparts without *est-ce que*. To this day, however, despite several studies documenting the use of prosodic cues to disambiguate the meaning of indirect request constructions, ironical utterances and specific types of French interrogative constructions, such as *Qu’est-ce que P* (See 2.1), the prosodic features of *Est-ce que...?* polar constructions used in the speech act of requesting remain unexplored.

3. Experimental study

3.1 Materials

The utterances of French sentences used as experimental stimuli in the perception study will focus on the various kinds of interrogatives illustrated in Appendix A. All code and data required to reproduce the analyses and figures in this paper can be found on GitHub.⁵ The design and

⁵ Link: https://github.com/seantrott/french_prosody_analysis.

analysis (except for the norming studies) were pre-registered ahead of data collection; the pre-registration can be found on OSF (<https://osf.io/gb3fn>).

3.2 Norming Studies

The goal of the norming studies was to establish the salience of distinct candidate interpretations for each utterance in our set of stimuli – independent of its prosody. This is crucial for testing the hypothesis outlined in the Non-canonical Marking view: namely, that the salience of distinct interpretations affects which cues are used to signal a speaker’s intent. Importantly, we operationalize *salience* here as how easily human comprehenders can bring a given interpretation to mind in the absence of explicit contextual cues favoring one interpretation or the other. (As noted earlier, participants may still construct plausible contextual scenarios when evaluating the salience of different interpretations, or engage more bottom-up, compositional processes; in either case, our goal is to establish the prior probability of a given sentence being interpreted as a request in the absence of prosodic cues.)

We conducted four separate norming studies, i.e., one for each construction illustrated in Appendix A. Following Trott et al. (2023), we expected to find a main effect of Grammatical Form: specifically, we expected that written versions of modal interrogatives would be more likely to be interpreted as requests than written versions of negative state remarks. More relevantly to the current work, we predicted that modal interrogatives would also be more likely to be interpreted as requests than non-modal interrogatives. Within each construction type, we did not have a specific prediction about the impact of *est-ce que*.

In each of the four norming studies, we recruited 320 native speakers of French (4×80) from France, Belgium or Switzerland, using Prolific as a recruiting platform. Of these participants, 145 self-identified as female (172 male; 1 other). The average self-reported age was 32 (standard deviation = 10; median = 31; range = 18–69). Each speaker received 2€ for participating.

In these norming studies, participants read a series of written sentences. After reading each sentence, they were asked to determine whether or not the sentence was a request by selecting the “yes” or “no” response option.

The goal of each norming study was to establish request probabilities for each of the four construction types in question. Each study included, in addition to the 48 sentences corresponding to one of the four construction types, 16 negative state remarks (*My phone is broken*, cf. Appendix B), and, as controls, 16 imperative sentences (*Listen to this song*), and 16 declarative sentences such as *Paris is the capital of France*. The key differences between the norming study were whether the target construction was a modal interrogative or a non-modal interrogative, and whether the target construction contained *est-ce que*.

3.2.1 Results

As depicted in **Figure 1**, there was a clear difference in the proportion of request interpretations across constructions containing a modal (*peux*) and those that did not contain a modal. Further, among constructions that did not contain a modal, the presence of *est-ce que* (ESQ) was associated with a lower proportion of request interpretations ($M = 0.30$). See **Table 1** for the by-condition means.

To analyze these results, we compared nested statistical models with log-likelihood ratio tests. Each model was a generalized linear mixed-effects model with a logit link, built using the *lme4* package (Bates et al., 2015) in R (R Development Core Team, 2019). First, we constructed the “full” model, which had Request Interpretation (Yes vs. No) as a dependent variable, fixed effects of Modal (Yes vs. No) and ESQ (Yes vs. No), an interaction between Modal and ESQ, and random intercepts for Item and Participant (random slopes were not included).

The full model, including the ESQ:Modal interaction, explained significantly more variance than a model omitting only this interaction [$\chi^2(1) = 14.5, p < .001$]. Inspecting the parameter estimates of this model, we found that Request interpretations were more likely overall for Modal constructions [$B = 9.1, SE = 1.15, p < .001$], less likely for constructions with ESQ [$B = -6.33, SE = 1.73, p < .001$], and relatively more likely for constructions with both a Modal and ESQ [$B = 7.44, SE = 1.99, p < .001$]. These results are illustrated in **Figure 1** below, and the coefficients are plotted in **Figure 2**.

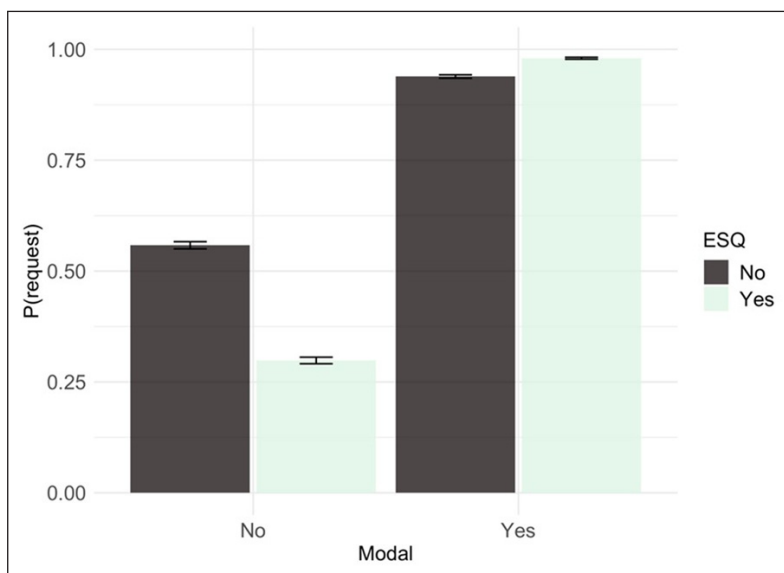


Figure 1: Proportion of request interpretations across conditions. Participants were more likely to interpret items containing a Modal as requests. Further, for items without a Modal, the presence of ESQ decreased the probability of request interpretations, though the inverse pattern was observed for items with a Modal.

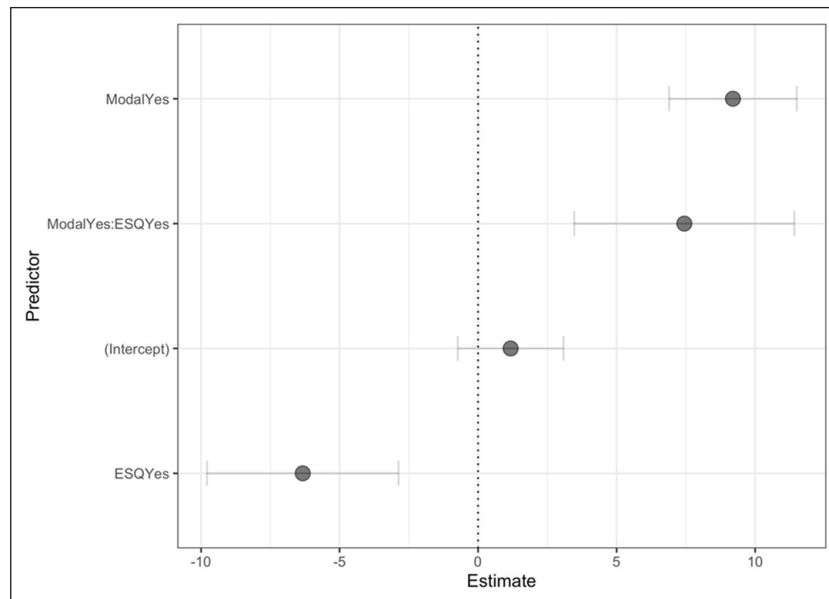


Figure 2: Parameter estimates for a generalized linear mixed-effects model predicting Request Interpretation (Yes vs. No) as a function of whether the construction contains a Modal or an ESQ, as well as the interaction between the two; the full model also included random intercepts for individual Participants and Items.

3.2.2 Discussion

Written sentences with a modal (*peux*) were more likely to be interpreted as requests than written sentences without a modal. Further, the presence of *est-ce que* decreased the likelihood of request interpretations for sentences without a modal.

Critically, these results are promising for the use of these stimuli in a production study examining the role of prosody in disambiguating intent. Because $p(\text{request})$ varies by construction, these stimuli allow us to adjudicate between the competing hypotheses raised at the beginning of this article. If prosody conveys speech act information directly, then the same prosodic cues should be predictive of intent (and in the same way) for each type of construction – regardless of its prior probability of being interpreted as a request. But if prosody serves primarily as a contrastive marker, then we should identify distinct prosodic markers (or distinct manifestations of the same prosodic cues) that signal intent for each kind of construction.

3.3 Procedure

3.3.1 Production study

We recruited 8 native speakers of French (university students with France or Belgium as country of origin) after posting an announcement on LinkedIn and via the mailing list of the Faculty of Arts of the host institution. Information about the age and gender of the speakers was

collected to describe the sample of participants involved in the production study. Each speaker received 12€ for participating (the whole recording process, including participant briefing and debriefing, took about 45 minutes to complete). Speakers were asked to read and sign an informed consent form, after which they were brought to a small testing booth containing a laptop including an internal microphone. They were instructed to sit down and remain at a constant distance from the microphone.

Before the recording began, speakers were given examples demonstrating how the same sentence can be used as either a request for action or a question. These example sentences were different from those used as target stimuli, and, to keep participants from imitating the prosody of the experimenter, these example sentences were not spoken aloud to them. We explained the following as part of the instructions before the study (English translation of the original French instructions): “Many French sentences can be interpreted in different ways. For example, the phrase *Can you pass the salt?* could be a request (as in saying *Please pass the salt*), or a question about the addressee’s physical ability to perform an action (as in asking whether or not the recipient would be able to pass the salt). Similarly, the phrase *My steak is overcooked* could be a complaint about the quality of the dish, or it could be a request for the waiter to bring a new steak.”

Each speaker was asked to record 80 utterances. Twelve of these were non-modal interrogatives without *est-ce que* (*Tu verbal phrase (VP)?*), twelve were modal interrogatives also without *est-ce que* (*Tu peux VP?*), twelve utterances were non-modal interrogatives with *est-ce que* (*Est-ce que tu VP?*), and twelve were modal interrogatives with *est-ce que* (*Est-ce que tu peux VP?*). The 48 polar interrogative stimuli were administered using a counterbalanced Latin Square design, according to which a participant would not be presented with a given propositional content in more than one construction. Participants were randomly assigned to one of four lists of stimuli. Stimuli were randomized for each participant. Speakers were allowed to speak each item aloud multiple times before recording the version they were satisfied with. Each speaker also recorded sixteen imperative sentences that are obviously requests (e.g., *Listen to this song*) and sixteen declarative sentences that are obviously not requests (e.g., *Paris is the capital of France*); these utterances served as control items for the perception study.

3.3.2 Perception study

The perception study was created on the Psytoolkit platform (Stoet, 2017). It consisted of a request identification task. The stimuli consisted of a subset of the spoken utterances recorded in the production study (see 3.4 below for more details). 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 is able to pass the salt). 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. Responses were collected in the same way for the target stimuli and the controls (imperative requests and declarative assertions).

The main dependent variable of the perception study is the yes/no answer to the question “Is it a request?”. The participants were asked to answer that question by clicking either the “yes” or “no” button. These judgments were recorded in scores of 0 (“no”) and 1 (“yes”). After each response, the next item was automatically displayed. The other variables are the answers to demographic questions about age and gender. These answers were used to describe the sample of participants involved in the perception study.

3.4 Methods

The perception study had a within-participant design, in which all participants are subjected to all manipulations (Modal \times ESQ \times Intent). Each participant was presented with all the indirect request utterances recorded by one of the 8 speakers. Thus, each participant heard 48 utterances in total (plus 20 control utterances); 35 participants were assigned to each of the 8 lists of stimuli corresponding to the 8 speakers in the production task.

3.4.1 Sample size calculation

We recruited 8 native speakers of French for the production study. For the perception study, we recruited 280 participants, for a total of 13,440 trials (280 \times 48). This corresponds to 1,680 trials per cell (13,440 / 8). Past work (Ruytenbeek et al., 2023) identified a significant effect of Intent on Interpretation, with 1,680 trials per cell. Thus, our sample size is based on the sample used in past work. For the perception study, we recruited participants aged between 18–60 (and with France as country of residence to minimize variation within the sample⁶) on the testing platform Prolific.

3.4.2 Methods for data analysis: Data processing

For each of the recorded utterances, we used Parselmouth (Jadoul et al., 2018), a Python interface to Praat, to extract the three acoustic features that differed significantly across the conditions

⁶ To ensure that we would eventually reach the figure of 280 participants for the perception study, we had to adjust our inclusion criteria slightly by allowing native speakers of French located not only in France, but also potentially in Belgium and Switzerland, to participate on Prolific.

in Ruytenbeek et al.'s (2023) study: mean F0, duration and slope of F0.⁷ 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 inter-speaker variability.

3.4.3 Outliers and exclusion criteria

Two inclusion criteria were set for the perception study:

- Since the study was about the interpretation of utterances in the French language, only participants with French as their mother tongue were recruited;
- A minimum age of 18 and a maximum of 60 were set for the study.

As for the exclusion criterion, participants who provided less than 9/10 correct responses to the clearly non-request items (e.g., *Paris is the capital of France*) were rejected. The strict application of this criterion resulted in the dismissal of only two participants.

3.5 Results

Here, we report the results for each study in turn. First, we focus on the analysis of acoustic features for the production study, which includes investigating which features are predictive of Meaning and quantifying the performance of a machine learning classifier equipped with each feature. Then, we ask how successfully humans identify the intended meaning of an utterance, and whether there is any overlap in the acoustic features that predict a speaker's original intent and those that predict a comprehender's interpretation. All analyses were pre-registered and carried out in R, unless otherwise indicated.

3.5.1 Production study: Analysis of individual acoustic features

First, we asked how much independent variance in Meaning (request or non-request) was explained by each acoustic feature. To do this, we constructed a full generalized linear mixed model with a logit link, with Meaning as a dependent variable, and the three acoustic features (mean F0, F0 Duration, and F0 Slope) as predictors (along with random intercepts for each item). We also included interactions between each acoustic feature and the presence of a Modal (yes vs. no), as well as interactions between each acoustic feature and the presence of ESQ (yes vs. no). We compared this full model to a series of reduced models, omitting each feature and its interaction with Modal or ESQ in turn. Model comparisons were carried out using log-likelihood ratio tests,

⁷ In Ruytenbeek et al. (2023), mean F0 and duration were significant predictors of Intent in the production study (for all items and for negative state remarks specifically, but not for modal interrogatives only); F0 slope was a significant predictor of Intent for all items and for modal interrogatives specifically. In the perception study, mean F0, duration and slope of F0 were significant predictors of Meaning.

and the p-values obtained from each comparison were adjusted for multiple comparisons using Holm-Bonferroni correction (Holm, 1979). (All p-values reported below represent the corrected values, unless indicated otherwise.)

In each case, a positive coefficient represents a higher likelihood of a request meaning, while a negative coefficient represents a higher likelihood of a non-request meaning. Altogether, the analyses of feature:Modal and feature:ESQ interactions inform the question of whether a given feature's association with a particular Meaning varies by specific constructional features. In contrast, the analysis of the main effect of a given feature informs the question of whether a particular feature exhibits consistent associations with a particular Meaning across different constructions.

We found that model fit was improved by a main effect of mean F0 [$\chi^2(1) = 13.4, p < .001$] and a main effect of F0 Duration, i.e., the number of voiced frames [$\chi^2(1) = 9.3, p = .02$]. The parameter estimates for both variables in the full model (i.e., a model including all interaction terms) were negative – that is, increases in both mean F0 [$B = -0.71, SE = 0.24$] and F0 duration [$B = -0.67, SE = 0.23$] were associated with a lower probability of an item being a request, holding other variables constant. Additionally, a mean F0:Modal interaction was trending towards significance after correcting for multiple comparisons [$\chi^2(1) = 5.76, p = 0.1$], as was a main effect of F0 Slope [$\chi^2(1) = 6.8, p = .06$]. The interaction term was positive, suggesting that a higher mean F0 was (relatively) more associated with request meanings specifically for items without a modal; the trending main effect of F0 Slope was negative, suggesting that items with a more positive F0 Slope were less likely to be requests.⁸

The trending effect of F0 Slope was consistent with past work on both French and English (Ruytenbeek et al., 2023; Trott et al., 2023), which found that positive F0 slopes were associated with the literal intent of modal interrogatives (i.e., as a yes/no question). The main effect of mean F0 was in the opposite direction of past work, which found that items produced as requests tended to have higher pitch on average (Ruytenbeek et al., 2023; Trott et al., 2023); notably, however, the effect in past work was identified for negative state remarks specifically (e.g., *My soup is cold*), not for modal interrogatives – further, we found that the direction of the effect inverted for items without a modal.

3.5.2 Production study: Machine learning classifier

We then 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. This model was then used to classify the held-out test item, enabling us to determine whether the label predicted from the model matches the actual label for the held-out

⁸ Additionally, in an exploratory analysis including *only* the main effects of each acoustic feature, all three emerged as significant predictors (with the same direction of effect in each case).

test item. This leave-one-out procedure was performed for every item in the dataset, ultimately providing an accuracy score (percentage of correctly classified held-out items). For each of the 381 splits of the data (8×48 recorded interrogative utterances in total, with two missing items), we thus fitted a logistic regression classifier to the 380 training utterances. The classifier was trained to predict an utterance’s original Meaning from all three acoustic features. In addition to main effects of each acoustic feature, the classifier included interactions between each acoustic feature and the presence of a Modal, as well as interactions between each acoustic feature and the presence of ESQ; i.e., the classifier was the “full model” from the section above. This classifier was then used to predict the Meaning of the held-out test item.

The classifier achieved 58% accuracy on held-out items. That is, using 0.5 as a decision threshold (such that output probabilities > 0.5 were labeled as requests, and all others were labeled as non-requests), 58% of the items were correctly assigned the “ground truth” label. In an exploratory analysis, we found that the accuracy of the classifier was considerably higher for items containing a modal (66.5%) than those without a modal (49.7%). Accuracy was also higher for some speakers than others, ranging from a minimum of 41.7% to a maximum of 83.3%.⁹

3.5.3 Perception Study

Following previous work (Ruytenbeek et al., 2023), we conducted two sets of analyses for the perception study. These analyses were designed to answer three separate, but related, questions: (i) Can listeners determine the intended interpretation of an utterance using its prosody?, (ii) Does their success depend on the type of construction in question?, and (iii) Which acoustic features are most associated with a listener’s pragmatic interpretation, and does the effect of each feature depend on the type of construction?

For the first question, we constructed a generalized linear mixed-effects model with a logit link, predicting a participant’s Interpretation (Request vs. Non-Request). We included a fixed effect of the speaker’s original Intent (Request vs. Non-Request), as well as random intercepts for item, speaker, and participant. As pre-registered, we also added by-participant random slopes for the effect of Intent. We compared this model to a model omitting only Intent and found that the full model did indeed explain additional variance [$\chi^2(1) = 49.21, p < .001$]. Specifically, items intended as requests were associated with a higher likelihood of being interpreted as a request [$B = 0.35, SE = 0.04, p < .001$]. Overall, 60.5% of Literal items were interpreted as requests, while 68.5% of Request items were interpreted as requests (see also **Figure 3** below). Thus,

⁹ Notably, similar speaker-wise variation was observed in a further exploratory analysis using a “leave-one-speaker-out” method, in which the classifier was fit on data from all speakers but one and tested on each held-out speaker in turn. Mean accuracy was 57.6%, and speaker-wise accuracy again ranged considerably, from 37.5% to 79.2%; these values were highly correlated with the measure of speaker-wise accuracy using LOOCV ($r = 0.95$), as were the underlying classifier output probabilities ($r = 0.97$).

regardless of the construction in question, a speaker's original intent was predictive of a listener's pragmatic interpretation, though the effect was small (mean accuracy was 54%). Consistent with the classifier results above, some speakers also produced more discriminable cues than others: speaker-wise accuracy ranged from 49.5% to 59.6%. Participants also varied considerably in their performance on the task, from 34.7% to 85.4% (median = 52.2%).

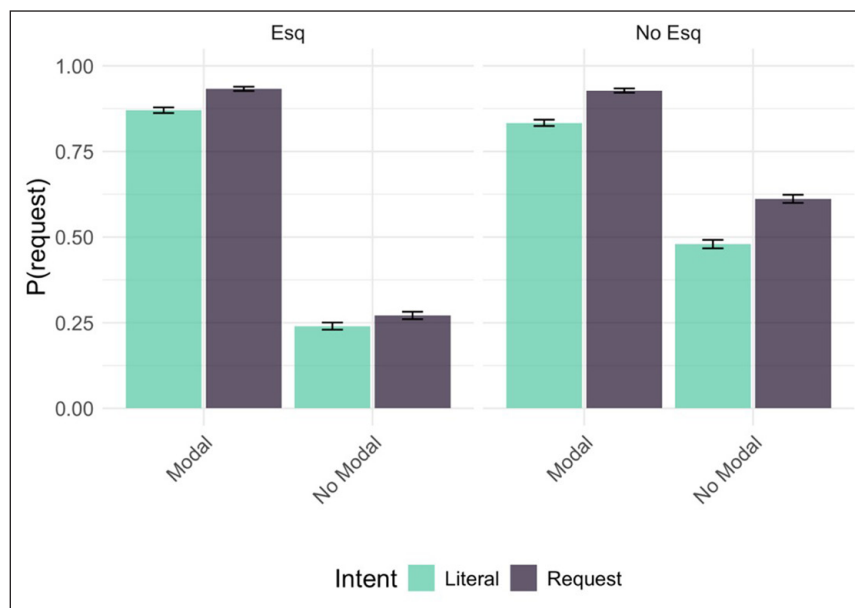


Figure 3: Rate of request interpretations by original speaker Intent (Literal vs. Request), as well as the constructional features of the item, i.e., whether it contained a modal or *est-ce que*.

We then added fixed effects of Modal and ESQ, as well as interactions between Modal:Intent and ESQ:Intent. Using a series of nested model comparisons, we asked whether each of the interactions explains more variance than a model containing only the main effects: namely, does the success with which a speaker conveys their intended interpretation through prosody depend on whether the construction contains a modal or *est-ce que*? A model containing both interaction terms explained more variance than a model omitting only the interaction between Modal:Intent [$\chi^2(1) = 24.97, p < .001$], as well as a model omitting only the interaction between ESQ:Intent [$\chi^2(1) = 47.9, p < .001$]. Specifically, items intended as Requests were less likely to be interpreted as such when they did not contain a modal [$B = -0.64, SE = 0.13, p < .001$], and items intended as Requests were also relatively less likely to be interpreted as such when they contained *est-ce que* [$B = 0.77, SE = 0.07, p < .001$]. Parameter estimates for each of the main effects and interaction terms from the full model are depicted in **Figure 4**.

Further model comparisons allowed us to identify the independent effects of Modal and ESQ, i.e., whether listeners were more or less likely to interpret an utterance with a modal (or *est-ce*

que) as a request. A model containing main effects of Intent, Modal, and ESQ explained more variance than a model omitting only Modal [$\chi^2(1) = 4722.1, p < .001$] and a model omitting only ESQ [$\chi^2(1) = 482.4, p < .001$]. Notably, and consistent with past work (Trott et al., 2023), the presence or absence of a modal was one of the strongest predictors of a participant's interpretation: 89.1% of trials with a modal were interpreted as requests (vs. 40% without).

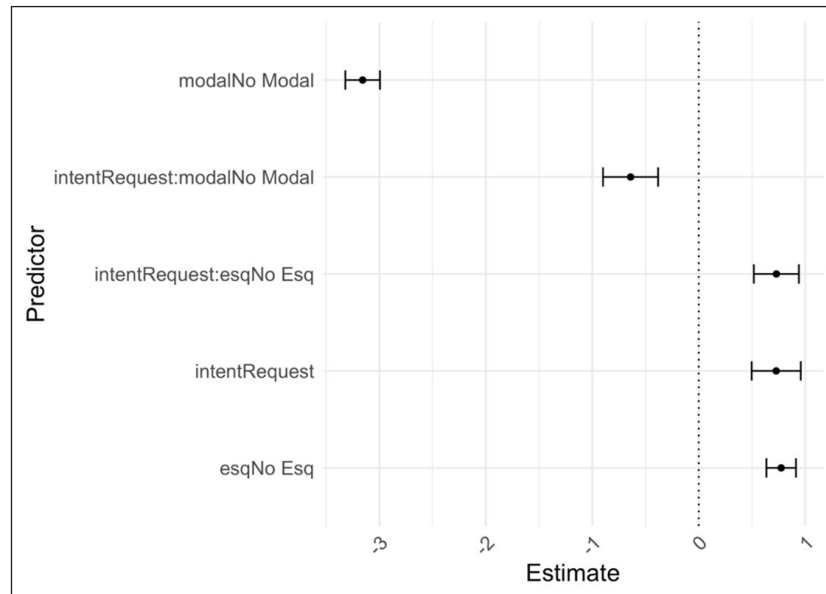


Figure 4: Coefficient estimates for each of the main effects and interaction terms in the full model predicting Interpretation.

We then asked about the relationship between each acoustic feature and pragmatic interpretations. To test this, we built a model predicting Interpretation, containing fixed effects of each of the acoustic features in question (mean F0, F0 slope, and F0 Duration), as well as an interaction between each acoustic feature and Modal, and an interaction between each acoustic feature and ESQ. The model also contained random intercepts for Participant, Item, and Speaker. This model mirrored the model used in the production study, in that the goal was to ask: (i) whether interpretations varied by the presence of certain features (i.e., a main effect of each acoustic feature), and (ii) whether the relationship between a given acoustic feature and interpretations varied according to constructional features. We answered (ii) by comparing the full model to a model omitting only each interaction in question. We addressed (i) by comparing a model including the main effect of a given feature (as well as all other acoustic features and their interactions with modal and ESQ) to a model omitting that feature entirely. There were a total of nine likelihood ratio tests (six unique interaction terms, three main effects); we adjusted for multiple comparisons using Holm-Bonferroni corrections across all nine tests.

After adjusting for multiple comparisons, three of the six interactions tested remained significant: Mean F0:Modal [$\chi^2(1) = 47.7, p < .001$], Mean F0:ESQ [$\chi^2(1) = 16.88, p < .001$], and F0 Duration:ESQ [$\chi^2(1) = 65.58, p < .001$]. For items without ESQ, a higher Mean F0 was associated with a lower likelihood of a request interpretation [B = -0.27, SE = 0.07]; for items without a modal, a higher mean F0 was associated with a higher likelihood of a request interpretation [B = 0.48, SE = 0.07]. Finally, longer utterances were associated with a particularly reduced likelihood of request interpretations when they did not contain an ESQ [B = -0.59, SE = 0.07].

Further, two of the three main effects remained significant: F0 Slope [$\chi^2(1) = 71.1, p < .001$] and F0 Duration [$\chi^2(1) = 16.98, p < .001$]. To analyze the direction of these effects, we considered the coefficients from a model with just the main effects (as well as main effects of Modal and ESQ).¹⁰ Items with a more positive F0 Slope were less likely to be interpreted as requests overall [B = -0.14, SE = 0.04], and items with a longer duration were also less likely to be interpreted as requests overall [B = -0.17, SE = 0.04]. Differences in F0 slope are depicted in **Figure 5**. Both results were consistent with past work on the prosody of interrogative requests (Ruytenbeek et al., 2023; Trott et al., 2023), which found that both duration and F0 slope were negatively associated with request interpretations for yes/no interrogatives.

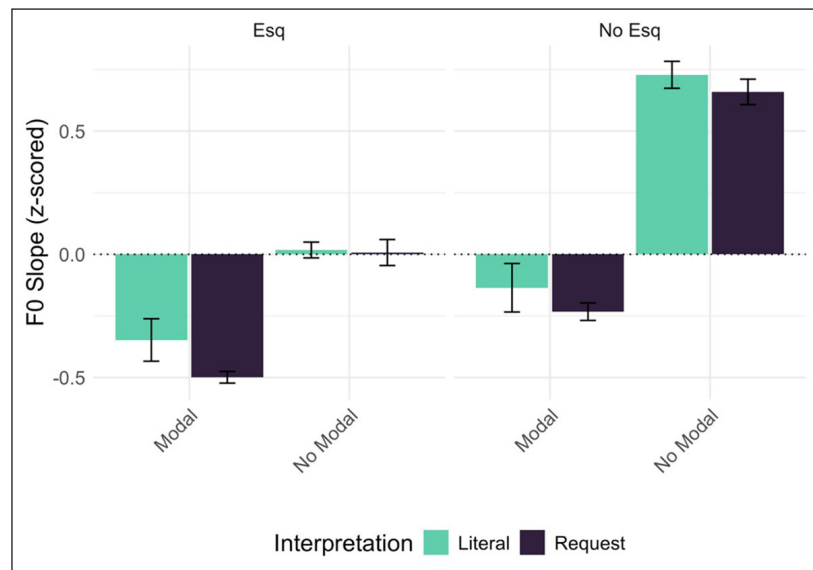


Figure 5: Mean z-scored F0 slope for utterances interpreted as either literal questions or requests. There were substantive main effects of both the presence of a modal and *est-ce que*. Beyond these effects, items interpreted as requests had a reduced F0 slope.

¹⁰ Because the statistical models used indicator coding, the coefficients from a full model (i.e., one including interactions between each acoustic feature and Modal/ESQ) would instead characterize the direction of effect for a specific subset of the data, e.g., for items including a Modal and ESQ. Thus, the sign of the coefficient for F0 Duration in the full model is positive, reflecting the effect of F0 Duration for items with a Modal and ESQ.

4. General discussion

A central aim of the current work was to adjudicate between two competing hypotheses about the role of prosody in disambiguating French indirect requests: do prosodic cues convey speech act information directly (the Direct Mapping view) or do they signal a deviation from the expected interpretation of an utterance (the Contrastive Marking view)? As noted in Section 1, these hypotheses make distinct predictions with respect to our experimental design. The Direct Mapping view predicts that the same acoustic features are a predictor of intent (or interpretation) in similar ways, regardless of the default interpretation of an utterance: that is, there should be a main effect of each acoustic feature on intent (or interpretation), but not an interaction with the constructional features that raise or lower the prior probability of an utterance being a request (i.e., a modal or *est-ce que*; see also the norming studies). The Contrastive Marking view makes the opposite prediction: namely, that the relationship between acoustic features and intent (or interpretation) should depend on the construction used.

An additional aim of this work was to replicate past work on both French (Ruytenbeek et al., 2023) and English (Trott et al., 2023) using an expanded set of stimuli that also controlled for the candidate speech acts of an utterance. Because we manipulated specific constructional features (like the presence or absence of a modal and of *est-ce que*), this design further allowed us to compare the relative impact of a speaker's intended interpretation (i.e., as a request or a yes/no question) to features known to raise or lower the probability of an utterance being interpreted as a request.

As described in **Table 2** below, we found mixed results. One cue (F0 Slope) exhibited consistent relationships with both intent and interpretation, regardless of the construction used (i.e., main effects). The specific direction of this effect – a more positive F0 slope predicting fewer request interpretations – was also consistent with past work on both French (Ruytenbeek et al., 2023) and English (Trott et al., 2023), and also matches work on the prosody of yes/no interrogatives more generally. Taken alone, these results provide support for the Direct Mapping view: the same acoustic feature was predictive of pragmatic interpretations regardless of whether an utterance contained a modal or *est-ce que*. The relationship between F0 Slope and a speaker's original intent trended in the same direction, but was not significant after correcting for multiple comparisons.

At the same time, other cues (Mean F0 and F0 Duration) exhibited interactions with the constructional features manipulated. Although both features were negatively associated with request interpretations on average, these effects reversed in the presence of *est-ce que*, a feature found independently to decrease the probability of a request interpretation (see the norming studies). Moreover, the presence of a modal (a constructional feature positively associated with request interpretations) further decreased the likelihood of request interpretations as Mean F0 increased. These results are more consistent with the Contrastive Marking view: the effect of a particular feature (e.g., Mean F0) on pragmatic interpretations depended crucially on

Table 2: Summary of the main effects and interactions for each acoustic feature in both the production and perception studies. A positive effect indicates that increases in that feature’s value were correlated with increases in the likelihood of a request meaning or interpretation. For interactions, the direction of the effect is indicated for the *on-class* (i.e., for items containing Modals or ESQ). Because the on-class was used as the reference class in the main analyses, this flips the sign from the coefficients reported above; thus, if a positive coefficient is found for Mean F0:NoModal (as was reported above), the interaction is interpreted as negative here, i.e., Mean F0 is negatively correlated with request interpretations for items with a modal. Items marked with ~ indicate that the effect was trending, i.e., it did not survive correction for multiple comparisons, but was significant prior correction.

Feature	Study	Main effect	Modal interaction	ESQ interaction
Mean F0	Production	✓ Negative	~ Negative	n.s.
Mean F0	Perception	n.s.	✓ Negative	✓ Positive
F0 Duration	Production	✓ Negative	n.s.	n.s.
F0 Duration	Perception	✓ Negative	n.s.	✓ Positive
F0 Slope	Production	~ Negative	n.s.	n.s.
F0 Slope	Perception	✓ Negative	n.s.	n.s.

constructional features that modulated the utterance’s prior request probability. Specifically, for the constructional feature that decreased the prior probability (i.e., *est-ce que*), the effect of both acoustic features reversed; for the constructional feature that increased the prior probability (i.e., modal), the effect of Mean F0 was exacerbated. This also might help explain why these relationships were not always consistent with our past work (Trott et al., 2023; Ruytenbeek et al., 2023), which tested different constructions that had different prior request probabilities: if the primary role of acoustic features is to signal a deviation from the default interpretation, then the direction of each effect will depend on the stimuli used.

Together, these results suggest that both hypotheses might, in fact, be true when it comes to French polar interrogatives. It simply depends on the feature in question. Of course, as noted above, it is difficult to know whether these results would generalize to a broader class of stimuli with various constructional features, or to other languages. That said, we might expect the F0 Slope results to generalize to other polar interrogatives, while the Mean F0 and F0 Duration results might be more dependent on the constructions (and candidate interpretations) in question. It is also notable that the effects of each acoustic feature (while significant) on both intent and interpretation were considerably weaker than that of the constructional features. One way to gain insight into this difference is to compare the effect of a speaker’s original intent to the effect of having a modal. In total, 68% of utterances intended as requests were interpreted as such, while only 60% of utterances intended as questions were interpreted as requests – a difference of 8% across conditions. In contrast, the gap in request interpretations across items containing

a modal (89%) and those without (40%) was 49%. Concretely, the effect of having a modal was about 6 times as large as the effect of a speaker’s original intent.

These results also inform research on the prosody (and pragmatics) of the *est-ce que* phrase. While Beyssade and Delais-Roussarie (2022) demonstrated that information-seeking *est-ce que* polar questions are compatible with a range of intonational contours, and Celle and Pélissier (2022) showed that *qu’est-ce que* surprise questions are prosodically marked by slower speech rate and less frequent final rises, our findings highlight a different dimension of the contribution of *est-ce que* to speech act understanding. In our norming studies, the presence of *est-ce que* decreased the likelihood of a request interpretation. Our perception data refined this observation: items prefixed with *est-ce que* and intended as requests were more likely to be interpreted as such. Moreover, the prosodic effects of Mean F0 and F0 Duration, which generally decreased the likelihood of a request interpretation, reversed in the presence of *est-ce que*. These findings expand the empirical picture of French polar interrogatives by showing that *est-ce que* interacts with prosodic features, modulating how listeners distinguish between information-seeking and request readings.

4.1 Limitations and future work

One limitation of the current work is that the analyses focused on coarse-grained acoustic features measured across the entirety of an utterance (e.g., Mean F0). It is plausible that a finer-grained analysis would reveal additional differences or nuances, perhaps with respect to the *est-ce que* phrase (cf. Druetta, 2002, 2003): for instance, past work on modal interrogatives (Trott et al., 2023) revealed increased duration of the modal verb (‘can’) when an utterance was intended as a yes/no question. Future work would also benefit from the use of established prosodic analysis frameworks, such as ToBI (Beckman et al., 2005). In addition to identifying subtle prosodic differences, such analyses would be critical for building a more precise mechanistic model of how listeners integrate prosodic cues in the act of language comprehension. Researchers might also make use of the growing body of machine learning toolkits and audio-language models (Baeovski et al., 2020) to extract informative cues in a more “bottom-up” fashion.

Another limitation is that the elicitation of stimuli was not done in a naturalistic setting: although speakers were naïve and not given instructions about how to produce an utterance, they were explicitly told to produce an utterance as a request or a question. This raises the question of whether the same cues (or any cues at all) would be produced in more naturalistic valid settings – and, if such cues are produced, whether listeners benefit from them in similar ways. Work on prosody (and language production and comprehension more generally) necessarily involves trade-offs between ecological validity and experimental control. In our case, we emphasized control over the specific utterances produced and the context in which they were interpreted. Future work could explore more contextualized designs, ranging from embedding the existing recordings in a discourse context to analyzing the prosody of utterances produced “in the wild”.

Parametric manipulation of the context in which utterances are perceived also opens the door to additional theoretical questions. In many cases, listeners might need to navigate a variety of potential cues to an utterance's meaning, such as constructional features (e.g., the presence of a modal), gesture (Kelly, 2001; Kelly et al., 1999; Kirk et al., 2011), or a speaker's knowledge state (Trott and Bergen, 2020). These cues may not always point in the same direction. Which factors influence the privileging of certain cues over others, as well as the *timing* with which they are integrated into a model of an utterance's meaning (Deliens et al., 2017)? Moreover, requests are, of course, a *social action*: beyond conveying a particular propositional meaning and illocutionary force, indirect requests likely serve other purposes as well, such as saving face (Brown and Levinson, 1987) or avoiding impoliteness (Ruytenbeek, 2019, 2020). Researchers could investigate how the prosodic features identified here influence the perceived (im)politeness of potential indirect requests, and how these judgments interact with situational features, such as the power differential between interlocutors (Brown and Levinson, 1987; Vergis et al., 2020) or a speaker's entitlement to make the request (Craven and Potter, 2010; Curl and Drew, 2008; Trott and Rossano, 2020).

5. Conclusion

Indirect requests are a pervasive form of linguistic ambiguity, raising the question of how listeners interpret a speaker's intended meaning. Focusing on French polar interrogatives, we asked whether the same acoustic features consistently convey specific speech acts, or whether the relationship between these features and a speaker's intended meaning (or a listener's interpretation) depends on constructional features that raise or lower the likelihood of an utterance being interpreted as a request (e.g., the presence of a modal and of the phrase *est-ce que*). We found evidence for both hypotheses: the F0 slope across an utterance was predictive of listeners' interpretations, regardless of the construction used, while the mean F0 and duration of the utterance led to different interpretations, depending on the inclusion of a modal or *est-ce que*. This suggests that, at least for polar interrogatives, the role of prosodic features is not monolithic: in some cases, they might function more like grammatical constructions, i.e., they are "stored" with the speech act; in others, their use might not signal a specific speech act, but rather mark a deviation from the expected interpretation of a given sentence. Understanding these distinct mechanisms is crucial for developing more nuanced models of pragmatic inference and for understanding how speakers and listeners navigate the interplay between linguistic form, context, and meaning.

Appendix A. List of interrogative sentences used as stimuli for the present study

Modal interrogatives	Non-modal interrogatives	ESQ modal interrogatives	ESQ non-modal interrogatives
Tu peux ouvrir cette fenêtre ? 'Can you open this window?'	Tu ouvres cette fenêtre ? 'Are you opening this window?'	Est-ce que tu peux ouvrir cette fenêtre ? 'Can you open this window?'	Est-ce que tu ouvres cette fenêtre ? 'Are you opening this window?'
Tu peux porter cette valise ? 'Can you carry this suitcase?'	Tu portes cette valise ? 'Are you carrying this suitcase?'	Est-ce que tu peux porter cette valise ? 'Can you carry this suitcase?'	Est-ce que tu portes cette valise ? 'Are you carrying this suitcase?'
Tu peux lire cette facture ? 'Can you read this invoice?'	Tu lis cette facture ? 'Are you reading this invoice?'	Est-ce que tu peux lire cette facture ? 'Can you read this invoice?'	Est-ce que tu lis cette facture ? 'Are you reading this invoice?'
Tu peux bouger cette armoire ? 'Can you move this cabinet?'	Tu bouges cette armoire ? 'Are you moving this cabinet?'	Est-ce que tu peux bouger cette armoire ? 'Can you move this cabinet?'	Est-ce que tu bouges cette armoire ? 'Are you moving this cabinet?'
Tu peux ranger ces bouquins ? 'Can you store these books?'	Tu ranges ces bouquins ? 'Are you storing these books?'	Est-ce que tu peux ranger ces bouquins ? 'Can you store these books?'	Est-ce que tu ranges ces bouquins ? 'Are you storing these books?'
Tu peux suivre cet itinéraire ? 'Can you follow this itinerary?'	Tu suis cet itinéraire ? 'Are you following this itinerary?'	Est-ce que tu peux suivre cet itinéraire ? 'Can you follow this itinerary?'	Est-ce que tu suis cet itinéraire ? 'Are you following this itinerary?'
Tu peux attraper cette boîte ? 'Can you grab that box?'	Tu attrapes cette boîte ? 'Are you grabbing that box?'	Est-ce que tu peux attraper cette boîte ? 'Can you grab that box?'	Est-ce que tu attrapes cette boîte ? 'Are you grabbing that box?'
Tu peux fermer cette porte ? 'Can you close that door?'	Tu fermes cette porte ? 'Are you closing that door?'	Est-ce que tu peux fermer cette porte ? 'Can you close that door?'	Est-ce que tu fermes cette porte ? 'Are you closing that door?'
Tu peux goûter ce toast ? 'Can/you taste this toast?'	Tu goûtes ce toast ? 'Are you tasting this toast?'	Est-ce que tu peux goûter ce toast ? 'Can you taste this toast?'	Est-ce que tu goûtes ce toast ? 'Are you tasting this toast?'

(Contd.)

Modal interrogatives	Non-modal interrogatives	ESQ modal interrogatives	ESQ non-modal interrogatives
Tu peux noter ce numéro ? 'Can you write down this number?'	Tu notes ce numéro ? 'Are you writing down this number?'	Est-ce que tu peux noter ce numéro ? 'Can you write down this number?'	Est-ce que tu notes ce numéro ? 'Are you writing down this number?'
Tu peux déplacer cette table ? 'Can you move this table?'	Tu déplaces cette table ? 'Are you moving this table?'	Est-ce que tu peux déplacer cette table ? 'Can you move this table?'	Est-ce que tu déplaces cette table ? 'Are you moving this table?'
Tu peux garer cette voiture ? 'Can you park this car?'	Tu gares cette voiture ? 'Are you parking this car?'	Est-ce que tu peux garer cette voiture ? 'Can you park this car?'	Est-ce que tu gares cette voiture ? 'Are you parking this car?'
Tu peux démonter cette armoire ? 'Can you dismantle this wardrobe?'	Tu démontes cette armoire ? 'Are you dismantling this wardrobe?'	Est-ce que tu peux démonter cette armoire ? 'Can you dismantle this wardrobe?'	Est-ce que tu démontes cette armoire ? 'Are you dismantling this wardrobe?'
Tu peux recopier cette phrase ? 'Can you copy this sentence?'	Tu recopies cette phrase ? 'Are you copying this sentence?'	Est-ce que tu peux recopier cette phrase ? 'Can you copy this sentence?'	Est-ce que tu recopies cette phrase ? 'Are you copying this sentence?'
Tu peux dessiner ce bâtiment ? 'Can you draw this building?'	Tu dessines ce bâtiment ? 'Are you drawing this building?'	Est-ce que tu peux dessiner ce bâtiment ? 'Can you draw this building?'	Est-ce que tu dessines ce bâtiment ? 'Are you drawing this building?'
Tu peux monter ce meuble ? 'Can you assemble this furniture?'	Tu montes ce meuble ? 'Are you assembling this furniture?'	Est-ce que tu peux monter ce meuble ? 'Can you assemble this furniture?'	Est-ce que tu montes ce meuble ? 'Are you assembling this furniture?'
Tu peux enlever cet autocollant ? 'Can you remove this sticker?'	Tu enlèves cet autocollant ? 'Are you removing this sticker?'	Est-ce que tu peux enlever cet autocollant ? 'Can you remove this sticker?'	Est-ce que tu enlèves cet autocollant ? 'Are you removing this sticker?'
Tu peux nettoyer cette tache ? 'Can you clean this stain?'	Tu nettoies cette tache ? 'Are you cleaning this stain?'	Est-ce que tu peux nettoyer cette tache ? 'Can you clean this stain?'	Est-ce que tu nettoies cette tache ? 'Are you cleaning this stain?'

(Contd.)

Modal interrogatives	Non-modal interrogatives	ESQ modal interrogatives	ESQ non-modal interrogatives
Tu peux défaire ce nœud ? 'Can you untie this knot?'	Tu défais ce nœud ? 'Are you untying this knot?'	Est-ce que tu peux défaire ce nœud ? 'Can you untie this knot?'	Est-ce que tu défais ce nœud ? 'Are you untying this knot?'
Tu peux laver ces fruits ? 'Can you wash these fruits?'	Tu laves ces fruits ? 'Are you washing these fruits?'	Est-ce que tu peux laver ces fruits ? 'Can you wash these fruits?'	Est-ce que tu laves ces fruits ? 'Are you washing these fruits?'
Tu peux tailler cet arbuste ? 'Can you trim this shrub?'	Tu tailles cet arbuste ? 'Are you trimming this shrub?'	Est-ce que tu peux tailler cet arbuste ? 'Can you trim this shrub?'	Est-ce que tu tailles cet arbuste ? 'Are you trimming this shrub?'
Tu peux remplacer cette ampoule ? 'Can you replace this light bulb?'	Tu remplaces cette ampoule ? 'Are you replacing this light bulb?'	Est-ce que tu peux remplacer cette ampoule ? 'Can you replace this light bulb?'	Est-ce que tu remplaces cette ampoule ? 'Are you replacing this light bulb?'
Tu peux décoller cette affiche ? 'Can you take down this poster?'	Tu décolles cette affiche ? 'Are you taking down this poster?'	Est-ce que tu peux décoller cette affiche ? 'Can you take down this poster?'	Est-ce que tu décolles cette affiche ? 'Are you taking down this poster?'
Tu peux détacher ces rideaux ? 'Can you take down these curtains?'	Tu détaches ces rideaux ? 'Are you taking down these curtains?'	Est-ce que tu peux détacher ces rideaux ? 'Can you take down these curtains?'	Est-ce que tu détaches ces rideaux ? 'Are you taking down these curtains?'
Tu peux mélanger ces ingrédients ? 'Can you mix these ingredients?'	Tu mélanges ces ingrédients ? 'Are you mixing these ingredients?'	Est-ce que tu peux mélanger ces ingrédients ? 'Can you mix these ingredients?'	Est-ce que tu mélanges ces ingrédients ? 'Are you mixing these ingredients?'
Tu peux débrancher ce câble ? 'Can you unplug this cable?'	Tu débranches ce câble ? 'Are you unplugging this cable?'	Est-ce que tu peux débrancher ce câble ? 'Can you unplug this cable?'	Est-ce que tu débranches ce câble ? 'Are you unplugging this cable?'
Tu peux gonfler ce matelas ? 'Can you inflate this mattress?'	Tu gonfles ce matelas ? 'Are you inflating this mattress?'	Est-ce que tu peux gonfler ce matelas ? 'Can you inflate this mattress?'	Est-ce que tu gonfles ce matelas ? 'Are you inflating this mattress?'

(Contd.)

Modal interrogatives	Non-modal interrogatives	ESQ modal interrogatives	ESQ non-modal interrogatives
Tu peux remplir ce vase ? 'Can you fill this vase?'	Tu remplis ce vase ? 'Are you filling this vase?'	Est-ce que tu peux remplir ce vase ? 'Can you fill this vase?'	Est-ce que tu remplis ce vase ? 'Are you filling this vase?'
Tu peux plier cette chemise ? 'Can you fold this shirt?'	Tu plies cette chemise ? 'Are you folding this shirt?'	Est-ce que tu peux plier cette chemise ? 'Can you fold this shirt?'	Est-ce que tu plies cette chemise ? 'Are you folding this shirt?'
Tu peux dresser la table ? 'Can you set the table?'	Tu dresses la table ? 'Are you setting the table?'	Est-ce que tu peux dresser la table ? 'Can you set the table?'	Est-ce que tu dresses la table ? 'Are you setting the table?'
Tu peux allumer le barbecue ? 'Can you light the barbecue?'	Tu allumes le barbecue ? 'Are you lighting the barbecue?'	Est-ce que tu peux allumer le barbecue ? 'Can you light the barbecue?'	Est-ce que tu allumes le barbecue ? 'Are you lighting the barbecue?'
Tu peux classer ces magazines ? 'Can you sort these magazines?'	Tu classes ces magazines ? 'Are you sorting these magazines?'	Est-ce que tu peux classer ces magazines ? 'Can you sort these magazines?'	Est-ce que tu classes ces magazines ? 'Are you sorting these magazines?'
Tu peux replier cette couverture ? 'Can you fold up this blanket?'	Tu replies cette couverture ? 'Are you folding up this blanket?'	Est-ce que tu peux replier cette couverture ? 'Can you fold up this blanket?'	Est-ce que tu replies cette couverture ? 'Are you folding up this blanket?'
Tu peux ouvrir cette bouteille ? 'Can you open this bottle?'	Tu ouvres cette bouteille ? 'Are you opening this bottle?'	Est-ce que tu peux ouvrir cette bouteille ? 'Can you open this bottle?'	Est-ce que tu ouvres cette bouteille ? 'Are you opening this bottle?'
Tu peux fermer ce récipient ? 'Can you close this container?'	Tu fermes ce récipient ? 'Are you closing this container?'	Est-ce que tu peux fermer ce récipient ? 'Can you close this container?'	Est-ce que tu fermes ce récipient ? 'Are you closing this container?'
Tu peux attraper ce bocal ? 'Can you grab that jar?'	Tu attrapes ce bocal ? 'Are you grabbing that jar?'	Est-ce que tu peux attraper ce bocal ? 'Can you grab that jar?'	Est-ce que tu attrapes ce bocal ? 'Are you grabbing that jar?'

(Contd.)

Modal interrogatives	Non-modal interrogatives	ESQ modal interrogatives	ESQ non-modal interrogatives
Tu peux emballer ce jouet ? 'Can you wrap this toy?'	Tu emballes ce jouet ? 'Are you wrapping this toy?'	Est-ce que tu peux emballer ce jouet ? 'Can you wrap this toy?'	Est-ce que tu emballes ce jouet ? 'Are you wrapping this toy?'
Tu peux refaire ce bandage ? 'Can you redo that bandage?'	Tu refais ce bandage ? 'Are you redoing that bandage?'	Est-ce que tu peux refaire ce bandage ? 'Can you redo that bandage?'	Est-ce que tu refais ce bandage ? 'Are you redoing that bandage?'
Tu peux capturer cet insecte ? 'Can you catch that insect?'	Tu captures cet insecte ? 'Are you catching that insect?'	Est-ce que tu peux capturer cet insecte ? 'Can you catch that insect?'	Est-ce que tu captures cet insecte ? 'Are you catching that insect?'
Tu peux égoutter ces légumes ? 'Can you drain these vegetables?'	Tu égouttes ces légumes ? 'Are you draining these vegetables?'	Est-ce que tu peux égoutter ces légumes ? 'Can you drain these vegetables?'	Est-ce que tu égouttes ces légumes ? 'Are you draining these vegetables?'
Tu peux préparer du café ? 'Can you make some coffee?'	Tu prépares du café ? 'Are you making some coffee?'	Est-ce que tu peux préparer du café ? 'Can you make some coffee?'	Est-ce que tu prépares du café ? 'Are you making some coffee?'
Tu peux chanter cette mélodie ? 'Can you sing this tune?'	Tu chantes cette mélodie ? 'Are you singing this tune?'	Est-ce que tu peux chanter cette mélodie ? 'Can you sing this tune?'	Est-ce que tu chantes cette mélodie ? 'Are you singing this tune?'
Tu peux vider ce coffre ? 'Can you empty that trunk?'	Tu vides ce coffre ? 'Are you emptying that trunk?'	Est-ce que tu peux vider ce coffre ? 'Can you empty that trunk?'	Est-ce que tu vides ce coffre ? 'Are you emptying that trunk?'
Tu peux mémoriser cette adresse ? 'Can you memorize this address?'	Tu mémorises cette adresse ? 'Are you memorizing this address?'	Est-ce que tu peux mémoriser cette adresse ? 'Can you memorize this address?'	Est-ce que tu mémorises cette adresse ? 'Are you memorizing this address?'
Tu peux rincer cette casserole ? 'Can you rinse that pan?'	Tu rinces cette casserole ? 'Are you rinsing that pan?'	Est-ce que tu peux rincer cette casserole ? 'Can you rinse that pan?'	Est-ce que tu rinces cette casserole ? 'Are you rinsing that pan?'

(Contd.)

Modal interrogatives	Non-modal interrogatives	ESQ modal interrogatives	ESQ non-modal interrogatives
Tu peux déboucher ce flacon ? 'Can you uncork that flask?'	Tu débouches ce flacon ? 'Are you uncorking that flask?'	Est-ce que tu peux déboucher ce flacon ? 'Can you uncork that flask?'	Est-ce que tu débouches ce flacon ? 'Are you uncorking that flask?'
Tu peux retirer ce clou ? 'Can you pull out that nail?'	Tu retires ce clou ? 'Are you pulling out that nail?'	Est-ce que tu peux retirer ce clou ? 'Can you pull out that nail?'	Est-ce que tu retires ce clou ? 'Are you pulling out that nail?'
Tu peux déterrer cette plante ? 'Can you dig up that plant?'	Tu déterres cette plante ? 'Are you digging up that plant?'	Est-ce que tu peux déterrer cette plante ? 'Can you dig up that plant?'	Est-ce que tu déterres cette plante ? 'Are you digging up that plant?'

Appendix B. List of negative state remarks stimuli used in the norming studies

Negative state remarks
Ma soupe est froide. 'My soup is cold.'
Mon téléphone est cassé. 'My phone is broken.'
Mon verre est vide. 'My glass is empty.'
Mon ordi est hors service. 'My laptop is out of order.'
Mes mains sont gercées. 'My hands are cracked.'
Mes cheveux sont trempés. 'My hair is soaked.'
Mon mouchoir est déchiré. 'My handkerchief is torn.'
Ma chaise est mouillée. 'My chair is wet.'
Ma voiture est en panne. 'My car has broken down.'
Ma carte est illisible. 'My card is illegible.'
Mes lunettes sont embuées. 'My glasses are fogged up.'
Mon vélo est inutilisable. 'My bike is out of action.'
Ma gorge est sèche. 'My throat is dry.'
Ma fourchette est sale. 'My fork is dirty.'
Mon chien est affamé. 'My dog is very hungry.'
Ma valise est lourde. 'My suitcase is heavy.'

Data accessibility statement

The design and analysis (except for the norming studies) were pre-registered ahead of data collection; the details of the pre-registration can be found on OSF (<https://osf.io/gb3fn>). All code and data required to reproduce the analyses and figures in this article can be found on GitHub (https://github.com/seantrott/french_prosody_analysis).

Ethics and consent

The experiments reported in this article are covered under protocol number *G-2025-9287* submitted to the SMEC ethical committee of the KU Leuven. The protocol was approved on March 25, 2025.

Acknowledgements

We are very grateful to the anonymous reviewers of this article, to Sharon Peperkamp, the editor who handled our manuscript, for their constructive feedback and suggestions, and to Fernanda Ferreira for the follow-up. Nicolas Ruytenbeek gratefully acknowledges support from the grant STG/22/010 (KU Leuven internal funds).

Competing interests

The authors have no competing interests to declare.

Authors contributions

Ruytenbeek and Trott were in charge of data and stimuli curation, formal (theoretical) analyses, and methodology, as well as the process of writing the original draft. Ruytenbeek was in charge of project administration. Ruytenbeek and Trott contributed reviews and editing to the original draft. Ruytenbeek collected and pre-processed the data for the norming studies and production and perception experiments. Trott conducted the statistical analyses and created the data visualizations. Ruytenbeek acquired the necessary funding for this project. The contributor roles are based on the CRediT system taxonomy.

ORCID IDs

Nicolas Ruytenbeek: <https://orcid.org/0000-0002-1651-9777>

Sean Trott: <https://orcid.org/0000-0002-6003-3731>

References

- Attardo, S., Eisterhold, J., Hay, J., & Poggi, I. (2003). Multimodal markers of irony and sarcasm. *Humor*, 16(2), 243–260. <https://doi.org/10.1515/humr.2003.012>
- Baevski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). wav2vec 2.0: A framework for self-supervised learning of speech representations. *Advances in Neural Information Processing Systems*, 33, 12449–12460. <https://doi.org/10.48550/arXiv.2006.11477>
- Banuazizi, A., & Creswell, C. (1999). Is that a real question? Final rises, final falls, and discourse function in yes-no question intonation. *CLS*, 35, 1–14.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <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. [https://doi.org/10.1016/0749-596X\(91\)90030-N](https://doi.org/10.1016/0749-596X(91)90030-N)
- Beckman, M. E., Hirschberg, J. B., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S. Jun (Ed.), *Prosodic models and transcription: Towards prosodic typology* (pp. 9–54). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199249633.003.0002>
- Behnstedt, P. (1973). *Viens-tu ? Est-ce que tu viens ? Tu viens ? Formen und Strukturen des direkten Fragesatzes im Französischen*. Gunter Narr.
- Beyssade, C., & Delais-Roussarie, E. (2022). The prosody of French rhetorical questions. *Linguistics Vanguard*, 8(2), 277–286. <https://doi.org/10.1515/lingvan-2020-0121>
- Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511813085>
- Bryant, G. A. (2010). Prosodic contrasts in ironic speech. *Discourse Processes*, 47(7), 545–566. <https://doi.org/10.1080/01638530903531972>
- Bryant, G. A., & Fox Tree, J. E. (2002). Recognizing verbal irony in spontaneous speech. *Metaphor and Symbol*, 17(2), 99–117. https://doi.org/10.1207/S15327868MS1702_2
- Bryant, G. A., & Fox Tree, J. E. (2005). Is there an ironic tone of voice? *Language and Speech*, 48(3), 257–277. <https://doi.org/10.1177/00238309050480030101>
- Celle, A., & Péliissier, M. (2022). Surprise questions in spoken French. *Linguistics Vanguard*, 8(2), 287–302. <https://doi.org/10.1515/lingvan-2020-0109>
- Cheang, H. S., & Pell, M. D. (2008). The sound of sarcasm. *Speech Communication*, 50(5), 366–381. <https://doi.org/10.1016/j.specom.2007.11.003>
- Coveney, A. (2011). L'interrogation directe. *Travaux de Linguistique*, 2, 112–145. <https://doi.org/10.3917/tl.063.0112>
- Coveney, A., & Dekhissi, L. (2017). Parallel innovations in conflictual rhetorical questions in the multicultural vernaculars of London English and Parisian French. In H. Tyne, M. Bilger, P. Cappeau, & E. Guerin (Eds.), *La variation en question(s), hommages à Françoise Gadet* (pp. 181–208). Peter Lang. <https://doi.org/10.3726/b13079>

- Craven, A., & Potter, J. (2010). Directives: Entitlement and contingency in action. *Discourse Studies*, 12(4), 419–442. <https://doi.org/10.1177/1461445610370126>
- Curl, T. S., & Drew, P. (2008). Contingency and action: A comparison of two forms of requesting. *Research on Language and Social Interaction*, 41(2), 129–153. <https://doi.org/10.1080/08351810802028613>
- Delais-Roussarie, E., & Beyssade, C. (2019, June 17–19). Information seeking vs. rhetorical questions: From gradience to categoricity. 3rd PaPE Conference, *Phonetics and Phonology in Europe*. University of Lecce, Italy.
- Deliens, G., Antoniou, K., Clin, E., & Kissine, M. (2017). Perspective-taking and frugal strategies: Evidence from sarcasm detection. *Journal of Pragmatics*, 119, 33–45. <https://doi.org/10.1016/j.pragma.2017.08.002>
- Dendale, P. (2025). *Lexicales* (edition 23.1). *Bibliographie en ligne d'études linguistiques portant sur des unités lexicales et grammaticales du français*, (7350 + références). <https://www.uantwerpen.be/lexicales>
- Déprez, V., Syrett, K., & S. Kawahara. (2012). Interfacing information and prosody: French wh-in-situ questions. In *Romance Languages and Linguistic Theory 2010: Selected papers from "Going Romance" Leiden 2010*, 135–154. John Benjamins. <https://doi.org/10.1075/rllt.4.07dep>
- Druetta, R. (2002). *Qu'est-ce tu fais ?* État d'avancement de la grammaticalisation de *est-ce que*. Première partie. *Linguae & Rivista di Lingue e Culture Moderne*, 2, 67–88.
- Druetta, R. (2003). *Qu'est-ce tu fais ?* État d'avancement de la grammaticalisation de *est-ce que*. Deuxième partie. *Linguae & Rivista di Lingue e Culture Moderne*, 1, 21–35.
- Dupras, S. (2018). *Est-ce que, c'est libre ou non ?* *Langues et Linguistique*, 37, 108–111.
- Ferreira, V. S. (2019). A mechanistic framework for explaining audience design in language production. *Annual Review of Psychology* 70, 29–51. <https://doi.org/10.1146/annurev-psych-122216-011653>
- González-Fuente, S., Prieto, P., & Noveck, I. (2016). A fine-grained analysis of the acoustic cues involved in verbal irony recognition in French. *Proceedings of Speech Prosody 2016*, 902–906. <https://doi.org/10.21437/SpeechProsody.2016-185>
- Hedberg, N., Sosa, J. M., & Görgülü, E. (2014). The meaning of intonation in yes-no questions in American English: A corpus study. *Corpus Linguistics and Linguistic Theory*, 13(2), 321–368. <https://doi.org/10.1515/cllt-2014-0020>
- Hellbernd, N., & Sammler, D. (2016). Prosody conveys speaker's intentions: Acoustic cues for speech perception. *Journal of Memory and Language*, 88, 70–86. <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. <https://doi.org/10.1016/j.wocn.2018.07.001>
- Kelly, S. D. (2001). Broadening the units of analysis in communication: Speech and nonverbal behaviours in pragmatic comprehension. *Journal of Child Language*, 28(2), 325–349. <https://doi.org/10.1017/S0305000901004664>

- Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language*, 40(4), 577–592. <https://doi.org/10.1006/jmla.1999.2634>
- Kirk, E., Pine, K. J., & Ryder, N. (2011). I hear what you say but I see what you mean: The role of gestures in children's pragmatic comprehension. *Language and Cognitive Processes*, 26(2), 149–170. <https://doi.org/10.1080/01690961003752348>
- Loevenbruck, H., Ben Jannet, M., D'Imperio, M., Spini, M., & Champagne-Lavau, M. (2013). Prosodic cues of sarcasm speech in French: Slower, higher, wider. *Proceedings of INTERSPEECH 2013 – 14th Annual Conference of the International Speech Communication Association, August 2013, Lyon, France, 2013*, 3537–3541. <https://doi.org/10.21437/Interspeech.2013-761>
- Padilla, X. (2012). ¿Existen rasgos prosódicos objetivos en los enunciados irónicos? *Oralia*, 14, 203–224. <https://doi.org/10.25115/oralia.v14i1.8187>
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America*, 90(6), 2956–2970. <https://doi.org/10.1121/1.401770>
- R Development Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- Reinhardt, J. (2019). *Regularity and variation in French direct interrogatives: The morphosyntax and intonation of question forms in reality TV shows, audio books and teaching materials* [Unpublished doctoral dissertation]. University of Konstanz.
- Rockwell, P. (2000). Lower, slower, louder: Vocal cues of sarcasm. *Journal of Psycholinguistic Research*, 29, 483–495. <https://doi.org/10.1023/A:1005120109296>
- Rooryck, J. (1994). On two types of underspecification: Towards a feature theory shared by syntax and phonology. *Probus*, 6, 207–233. <https://doi.org/10.1515/prbs.1994.6.2-3.207>
- Ruytenbeek, N. (2019). Indirect requests, relevance, and politeness. *Journal of Pragmatics*, 142, 78–89. <https://doi.org/10.1016/j.pragma.2019.01.007>
- Ruytenbeek, N. (2020). Do indirect requests communicate politeness? An experimental study of conventionalized indirect requests in French email communication. *Journal of Politeness Research*, 16(1), 111–142. <https://doi.org/10.1515/pr-2017-0026>
- Ruytenbeek, N., Bergen, B., & Trott, S. (2023). Prosody and speech act interpretation: The case of French indirect requests. *Journal of French Language Studies*, 33(1), 103–125. <https://doi.org/10.1017/S0959269522000254>
- Scharrer, L., Christmann, U., & Knoll, M. (2011). Voice modulations in German ironic speech. *Language and Speech*, 54(4), 435–465. <https://doi.org/10.1177/0023830911402608>
- Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, 48(1), 103–130. [https://doi.org/10.1016/S0749-596X\(02\)00519-3](https://doi.org/10.1016/S0749-596X(02)00519-3)
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24–31. <https://doi.org/10.1177/0098628316677643>

- Trott, S. T., & Bergen, B. (2020). When do comprehenders mentalize for pragmatic inference? *Discourse Processes*, 57(10), 900–920. <https://doi.org/10.1080/0163853X.2020.1822709>
- Trott, S. T., Reed, S., Ferreira, V., & Bergen, B. (2019). Prosodic cues signal the intent of potential indirect requests. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*.
- Trott, S. T., Reed, S., Kaliblotzky, D., Ferreira, V., & Bergen, B. (2023). The role of prosody in disambiguating English indirect requests. *Language and Speech*, 66(1), 118–142. <https://doi.org/10.1177/00238309221087715>
- Trott, S. T., & Rossano, F. (2020). The role of entitlement in formatting preferences across requesters and recipients. *Discourse Processes*, 57(7), 551–572. <https://doi.org/10.1080/0163853X.2020.1719796>
- Vergis, N., Jiang, X., & Pell, M. D. (2020). Neural responses to interpersonal requests: Effects of imposition and vocally-expressed stance. *Brain Research*, Article 1740. <https://doi.org/10.1016/j.brainres.2020.146855>
- Wall, K. (1985). Les questions dans la langue parlée. Usage actuel et problèmes pédagogiques, I et II. *Moderna Språk*, 85(79), 233–238; 309–317.

