

EFFECTS OF TUNE AND INTERSPEAKER DIFFERENCES ON THE INTERPRETATION OF REQUESTS AND OFFERS

Caterina Petrone¹, Elisa Sneed German², James Sneed German¹ and Kiwako Ito^{3,4}

¹Laboratoire Parole et Langage, CNRS & Aix Marseille Université, France ; ²Université Paul Valéry Montpellier 3, EMMA, Montpellier, France ; ³University of Newcastle, AUS ; ⁴Ohio State University, US
caterina.petrone@lpl-aix.fr, elisa.sneed-german@univ-montp3.fr, james.german@lpl-aix.fr, kiwako.ito@newcastle.edu.au

ABSTRACT

A growing body of evidence reveals that tune meaning is multidimensional and flexible, with the choice of a tune depending both on linguistic and metalinguistic purposes. This study explores how perlocutionary meaning is influenced by tune for requests and offers. Two female speakers of American English produced 96 request-offer pairs in the form of polar questions with both rising and falling tunes. Using an online survey system, participants' ratings of speaker authority were elicited. Falling tunes raised speaker authority to a greater degree for requests than for offers. Speaker 2, who had generally larger f0 movements than Speaker 1, was rated as more authoritative. Hence, different intonational tunes are assessed along with their metalinguistic and social dimensions, with individual differences in tune implementation also modulating listeners' judgments.

Keywords: rising/falling tunes, perlocutionary effects, online survey, request and offer.

1. INTRODUCTION

Traditional accounts of the semantics of intonational contours assume compositionality, such that the meaning of a given contour depends on the combined functions of pitch accents and boundary tones [7]. This framework, however, has yet to incorporate recent research showing that affective meaning may influence the judgement of speech act (e.g., statement vs. question [8]), that the speaker may choose different tunes (e.g., for requests and offers) according to their familiarity with the listener [1, 5], or that perlocutionary meaning is a function of both sentence type and tune [6].

In the current study, we investigate the interplay of tune and illocutionary force on perlocutionary effects with the ultimate goal of better defining intonational meaning by looking at the multidimensional interpretations that intonation can evoke. More specifically, we explore how perlocutionary meaning is influenced by tune (rising vs. falling) for two distinct, yet comparable illocutionary acts: requests and offers (e.g., *Can [you/I] check the weather for [me/you]?*).

In our study, an interpretational rating task elicited participants' responses along three scales, including speaker AUTHORITY (cf. [1, 6, 10, 5]). In line with [6], we expected the combination of a falling contour and the use of a polar question to evoke a perception of higher speaker authority than the same sentence type with a rising tune. We also expected a possible asymmetry between requests and offers with respect to the effects of falling tune on perceived speaker authority. Because requests are highly face-threatening [2], the use of a falling contour might evoke increased speaker authority.

2. BACKGROUND

In the Autosegmental-Metrical (AM) framework, [7] propose that speakers choose a particular tune to specify a particular relationship between the propositional content realized in the intonational phrase over which the tune is employed, the mutual beliefs of participants in that discourse, and presumed subsequent contributions to the discourse.

For [7], tunes bear meaning and the meaning is composed of three different types of tones (pitch accents, phrasal accents and boundary tones), which convey information about the discourse function of different parts of the utterance. The smallest tone unit, the pitch accent, conveys information about discourse referents, predicates, and the relationships between them, while phrasal accents convey the degree of relatedness to the immediately preceding and following phrases. Finally, boundary tones determine whether a given phrase is interpreted with the speaker's subsequent discourse contribution or not.

However, this tonal inventory is insufficient to describe certain types of variability in interpretation that have been observed. It is also necessary to take into account the fact that these tunes are linked to both illocutionary force (e.g., polar questions in English are often assumed to have a rising tune by default, but this association might vary with syntax [6]) and perlocutionary force (e.g., emotional state [10], power relations [4], sincerity [9], politeness [1, 5]).

Previous research on AE [6] has shown that across different sentence types such as polar questions, imperatives, WH-questions and

declaratives, the use of a level tune consistently conveyed more speaker annoyance than other tunes. On the other hand, a falling tune conveyed authority, and a rising tune conveyed politeness and favourable stance to the speaker. However, this effect of tune on perlocutionary force was affected by sentence type. Sentences that were biased towards an invitation illocution were interpreted to convey much more annoyance when they were polar interrogatives than when they were declaratives (e.g., *Do you wanna go to the movies?* vs. *We can go dancing.* (actual stimuli from [6])), but using a declarative for a request conveyed a greater perception of annoyance and diminished perception of politeness.

This previous research has thus examined the role of sentence type and tune on perlocutionary force, as well as the effect of sentence type and illocutionary force on perlocutionary force. In the present study, in order to examine the effects of illocutionary force and tune on perlocutionary force, we fix the tune to rising (L*L-H%) or falling (H*L-L%), and the illocutionary force to offers and requests. Specifically, we consider the contrast between requests and offers with polar question syntax, which sets aside the issue of sentence type and allows us to have a controlled manipulation of illocutionary force.

We believe that requests and offers are apt for this experiment because they are socially relevant illocutionary acts. They are charged exchanges with consequences for perceived politeness, sincerity, power and entitlement [3]. Additionally, both may be explicitly marked as polar questions with virtually identical surface syntax (*Can [you/I] bring [me/you] some water?*), and they have relatively transparent surface illocutionary force. They are conversational, and this particular form of question has been shown to be used in situations where the speaker believes the conditions for her/his request to be carried out are met [3].

Moreover, such questions should normally be produced with a rising tune, but can also be produced with a falling tune. As noted previously, in AE, tune has been shown to modulate stance, mood, authority, and politeness for requests [6]. In Catalan, tune for both requests and offers is also associated with social factors [1, 5]. [1] and [5] have shown that a greater social distance (e.g., strangers vs. siblings) between two participants in an exchange leads to the use of more falling tunes and lower pitch, while a higher cost of action (e.g., borrowing a car vs. asking for directions) leads to the use of more rising tunes.

The present paper reports the interim findings of a larger scale project that investigates the effect of tunes on multidimensional perlocutionary forces in

and out of context. Here, we focus on the effect of rising/falling tunes on the interpretation of speaker authority without contextual information.

3. METHODS

2.1. Corpus

A corpus of 96 request-offer pairs was created, in which all sentences were in the form of polar (yes/no) questions. Given that the cost of the action might impact tune choices [1, 5], task difficulty was rated on a written form of each offer and request sentence using Mechanical Turk. Fifty-two participants who self-reported as native speakers of AE evaluated the difficulty of completing a task (for requests) or asking someone else to complete a task (for offers) on their behalf on a continuous scale ranging from -100 (very easy) to 100 (very difficult). All items were judged as relatively easy (mean score: -29.5, SD = 4.3), with requests being slightly more difficult than offers ($t=2.8$, $p < .001$).

For the perlocutionary ratings, the request-offer pairs were produced with rising (L* L-H%) and falling (H* L-L%) contours by two female speakers of Midwestern AE who are well-trained on prosody and intonation phonology. Figs. 1 and 2 show an example of pitch tracks and ToBI annotation for f0 rises and falls for both speakers.

Acoustic analyses of the stimuli showed similar speech rates for the two speakers, while individual differences were found in the phonetic implementation of the f0 contours. In particular, for Speaker 2, the nuclear pitch accent was higher before falling contours and lower before rising contours (interaction between speaker and tune: $t=11.83$, $p < .001$). Furthermore, both the rising and falling contours had larger f0 movements for Speaker 2 than for Speaker 1 (interaction between speaker and tune: $t=3.26$, $p < .05$).

Figure 1. Two renderings of the request-offer pair *Can [you/I] check the weather for [me/you]?* produced by Speaker 1.

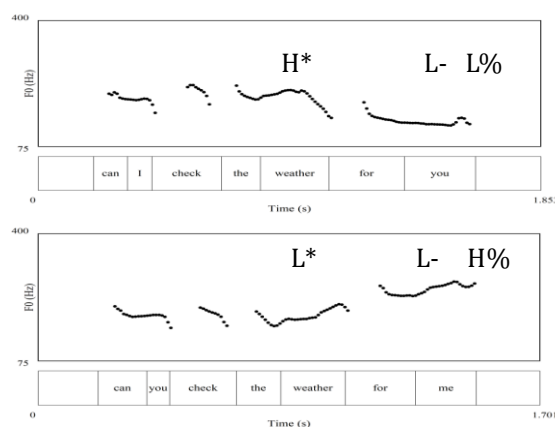
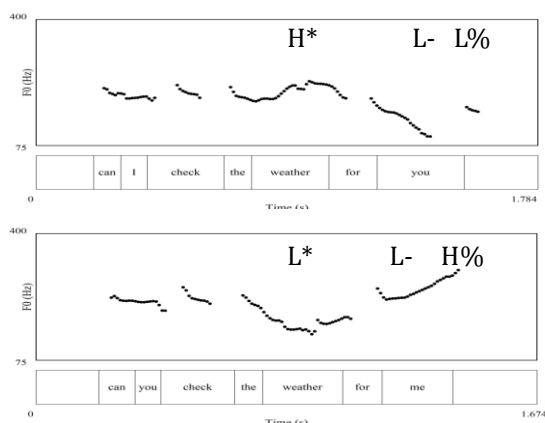


Figure 2. Two renderings of the request-offer pair *Can [you/I] check the weather for [me/you]?* produced by Speaker 2.



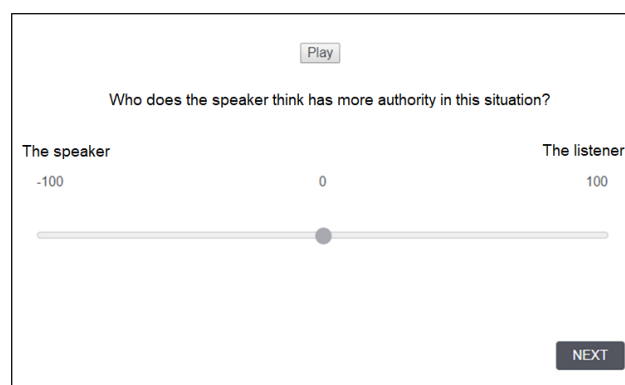
2.1. Participants and procedure

A Mechanical Turk online survey collected responses from 240 American English native speakers. Each participant was randomly assigned to one of 24 lists of 96 items. The lists were created by combining all experimental factors using a Latin Square design: 2 utterance types (request/offer) X 3 question types (authority/mood/sincerity) X 2 tunes (rising/falling) X 2 speakers (Speaker 1/Speaker 2). Each participant received only one of the three scales per item. Here we will focus on results on the ‘authority’ scale (N = 7559).

Participants were asked to wear headphones or earbuds and to sit in a quiet room with no background noise. Prior to the task, participants responded to a short demographic questionnaire including questions about their language background and language use, age, educational level and current occupation. Before starting the rating task, they listened to an unrelated item in order to adjust the volume to a comfortable level.

Each trial presented an audio file, which could be heard twice by pressing a play button. After the audio file was played, participants saw a written question (*Who does the speaker think has more authority in this situation?*), which was presented along with a continuous sliding scale. The poles of the scales were labeled *The speaker* and *The listener* at their right and left extremes, with the position of the labels being flipped in a half of trials to prevent bias from scale order. The sliding scale had an arbitrary range from -100 to 100. Listeners were required to adjust the slider bar left or right to reflect their own answer. A screenshot of a test item is illustrated in Fig. 3.

Figure 3. Screenshot of test item asking participants to judge who has more authority in the situation.



4. RESULTS

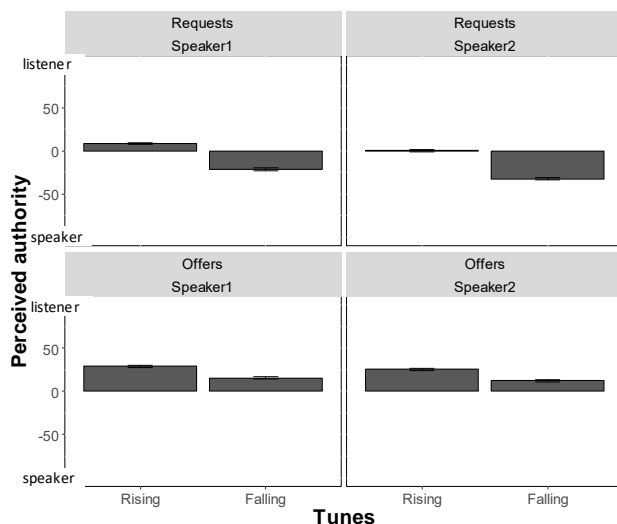
We included in the analysis only data from participants who completed more than 80% of the survey (as typical in standard perception experiment).

Fig. 4 shows the mean score for participants’ judgments along the scale of authority. Scores were converted such that positive values in the y-axes indicate that the listener was judged as more authoritative than the speaker while negative values indicate that the speaker was judged as more authoritative than the listener. The mean scores are -11.1 for requests and 20.5 for offers across speakers and tunes, indicating that requests conveyed more speaker authority while offers conveyed more listener authority. Furthermore, perceived speaker authority increased when requests were uttered with a falling (Speaker 1: -21.3; for Speaker 2: -32.2) than a rising tune (Speaker 1: 8.5; for Speaker 2: 0.3).

A mixed effects model tested the effects of three predictor factors: UTTERANCE TYPE (request/offer), TUNE (rising/falling) and SPEAKER (Speaker 1/Speaker 2) on the ratings for the authority scale. LISTENERS and ITEMS were included as random intercepts with a maximal random slope structure.

Results confirmed that requests led to higher speaker authority than offers ($t=7.23$, $p<.001$). Falling tune generally reduced listener authority (i.e., increased speaker authority) ($t=-8.82$, $p<.001$), with a larger impact of the tune for requests than offers ($t=3.61$, $p<.01$). Speaker 2 was rated more authoritative than Speaker 1 regardless of utterance type and tune ($t=4.78$, $p<.01$). Also, Speaker 2 sounded more authoritative than Speaker 1 for requests ($t=2.28$, $p<.05$), irrespective of the tune.

Figure 4. Means and standard error for perceived authority for requests (top) and offers (bottom), split by tunes and speakers.



5. DISCUSSION

The present results reinforce findings that intonational tune is a fundamental cue for perlocutionary/affective meaning. We focused on two distinct, yet comparable illocutionary acts (requests and offers) and we looked at their interaction with tune (rising and falling) on perceived authority.

All request-offer pairs referred to relatively easy tasks, which minimizes the impact of the cost of the action on tune interpretation. We found that requests evoke higher degrees of speaker authority regardless of intonation. Requests are speech acts with higher costs than offers for the listener/addressee, i.e., they are usually made to get the addressee to perform an action for the benefit of the speaker. Hence, they might indicate an authority imbalance in favour of the speaker. On the other side, offers inherently favour the listener, thus evoking higher listener authority. Furthermore, the falling tune led to a relatively stronger perception of speaker authority for the requests than for the offers. This may imply that participants considered the falling tune as more deviant from the social norm for making a request than for making an offer. Given that, in AE, such questions are usually produced with a rising tune, the use of a falling tune might be associated with the idea that the speaker is more confident regarding the potential for the interlocutor to accept the proposition of the question [1].

Finally, individual differences across speakers modulated listeners' judgments. Since both speakers produced the same intonation contour ($L^*L-H\%$ for rises and $H^*L-L\%$ for falls), we interpret this effect

as resulting from variability in tune implementation. Speaker 2 (who was judged as more authoritative than Speaker 1) realized larger f_0 dynamic changes, with the nuclear accents having more extreme melodic values. It is possible that the lower L^* and the steeper fall after H^* for Speaker 2 might have increased perceived finality, thus resulting in higher speaker authority in requests. This would be also in line with the 'frequency code' by which a low f_0 /falling f_0 would paralinguistically convey more 'assertiveness, dominance and authority' [11]. Additionally, various prosodic cues (e.g., voice quality) might have enhanced the interspeaker differences.

6. CONCLUSIONS

Contrary to the traditional distinction between linguistic and paralinguistic meaning of intonation, our work suggests that the different social ramifications of different illocutionary acts can influence how tune maps onto meaning. We aim to further investigate the correlations among different interpretational dimensions, and test how the presence of the discourse background or knowledge of speaker-listener power relationships influences utterance assessments.

7. ACKNOWLEDGEMENTS

This study was made possible through support by the A*MIDEX project (n° ANR-11-IDEX-0001-02) funded by the Investissements d'Avenir French Government program, managed by the French National Research Agency (ANR); and by an ANR grant to C. Petrone for the project 'Representation and Planning of Prosody' (n° ANR-14-CE30-0005-01). The research was also supported by grants ANR-16-CONV-0002 (ILCB) and ANR-11-LABX-0036 (BLRI).

8. REFERENCES

- [1] Astruc-Aguilera, L., del Mar Vanrell, M., Prieto, P. 2016. Cost of the action and social distance affect the selection of question intonation in Catalan. In Armstrong, M.E., Henricksen, M., del Mar Vanrell, M. (eds), *Intonational Grammar in Ibero-Romance: Approaches across linguistic subfields*. Amsterdam: John Benjamins, 91-114.
- [2] Brown, P., Levinson, S. (1987). *Politeness: Some universals in language use*. Cambridge: Cambridge University Press.

- [3] Curl, T., Drew, P. 2008. Contingency and Action: A Comparison of Two Forms of Requesting, *Research on Language and Social Interaction* 41, 129-153.
- [4] Gussenhoven, C. 2002. Intonation and interpretation: phonetics and phonology. In Bel, B., Marlien, I. (eds), *Proc. 1st Speech Prosody Aix-en-Provence*, 47-57.
- [5] Hübscher, I., Borràs-Comes, J., Prieto, P. 2017a. Prosodic mitigation characterizes Catalan formal speech: The Frequency Code reassessed. *J. of Phonetics* 65, 145-159.
- [6] Jeong, S., Potts, C. 2016. Intonational sentence-type conventions for perlocutionary effects: An experimental investigation. *Proc. of SALT 26* Austin, 1-22.
- [7] Pierrehumbert, J., Hirschberg, J. 1990. The meaning of intonational contours in the interpretation of discourse. In Cohen, P., Morgan, J., Pollack, M. (eds), *Intentions in communication*. Cambridge: MIT Press, 342-365.
- [8] Pihan, H., Tabert, M., Assuras, S., Borod, J. 2008. Unattended emotional intonations modulate linguistic prosody processing. *Brain and Language* 105, 141-147.
- [9] Roberts, F., Francis, A., Morgan, M. 2006. The interaction of inter-turn silence with prosodic cues in listener perceptions of “trouble” in conversation. *Speech Communication* 48, 1079-1093.
- [10] Scherer, K. 2003. Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40, 227-256.
- [11] Ohala, John J., 1984. An ethological perspective on common cross-language utilization of f0 of voice. *Phonetica* 41, 1-16.