Intonational sentence-type conventions for perlocutionary effects: An experimental investigation*

Sunwoo Jeong
Stanford Linguistics

Christopher Potts
Stanford Linguistics

Abstract One of the major open issues in semantics and pragmatics concerns the role of convention in relating sentence types with illocutionary acts and per-locutionary effects. For the type-to-illocution connection, some degree of *force conventionalism* seems to be widely accepted. In contrast, Austin (1962) and many subsequent researchers have assumed that perlocution is not a matter of convention, but rather arises inexorably from illocution, content, and context. In this paper, we challenge this fundamental assumption about perlocution with evidence from a new perception experiment focused on perlocutionary effects relating to the listener's conception of the speaker as a social actor. We find that these effects are predictable from sentence type plus intonation ('type + tune'), that they vary by type + tune, and that they are consistent across a wide range of sentence contents, contexts, and illocutionary inferences. We argue that these conventions are naturally incorporated into existing work on sentence-type conventions.

Keywords: illocutionary acts, perlocutionary effects, sentence types, intonational meaning, experimental pragmatics

1 Introduction

One of the major open issues in semantics and pragmatics concerns the role of convention in relating sentence types (e.g., declarative, interrogative, imperative) with illocutionary acts (e.g., assert, question, request) and subsequent perlocutionary effects (e.g., persuading; appearing authoritative or genial). For the type-to-illocution connection, some degree of *force conventionalism* seems to be widely accepted (Austin 1962; Searle 1969; Green 2007; Lauer 2013). In contrast, Austin (1962) and many subsequent researchers have assumed that perlocution is not a matter of convention, but rather arises inexorably from illocution, content, and context.

In this paper, we challenge this fundamental assumption about perlocutionary effects with evidence from perception experiments. The experiments focus on the

ways in which sentence-level intonation interacts with sentence type to produce stable perlocutionary effects relating to the listener's conception of the speaker as a social actor. We find that these effects are predictable from sentence type plus intonation ('type + tune'), that they vary by type + tune, and that they are consistent across a wide range of sentence contents, contexts, and illocutionary inferences. Together with prior work on cross-linguistic type + tune variation (Jun & Foreman 1996; Gordon 1999; Gussenhoven 2002; cf. Ohala 1983), these findings point to irreducible perlocutionary conventions. We argue that these conventions are naturally assimilated to the *sentence-type conventions* of Condoravdi & Lauer (2011, 2012) and Lauer (2013).

2 Conventions for illocution and perlocution

The connection between sentence types and illocutionary acts is one-to-many and highly uncertain in usage. For instance, interrogatives can be information-seeking, but they can also be used to quiz, to invite, to request, to accuse, and so forth. Declaratives standardly assert information, but they too can be intended and perceived as requests, commands, accusations, threats, and others. Imperatives are stereotypically used to command, but commanding is often not even indirectly part of the act performed, as in invitations and well-wishes (Condoravdi & Lauer 2012; Schmerling 1982).

Condoravdi & Lauer (2011, 2012) and Lauer (2013) propose to understand this constrained variation in terms of conventions governing the use of sentence types (for related approaches, see also Farkas & Roelofsen 2016; Malamud & Stephenson 2015; Portner 2007). Informal statements of such conventions are given in (1):

(1)   a.   A speaker who utters a declarative with content p thereby (in virtue of the act alone) commits to acting as though she believes p.

       b.   A speaker who utters an interrogative with content Q thereby commits to a preference for having the addressee commit to acting as though he believes an answer to Q.

       c.   A speaker who utters an imperative with content p thereby commits to acting in accord with having an effective preference for p.

These conventions circumscribe the range of felicitous uses for their associated sentence types, thereby allowing for the attested variation while still explaining why one's choice of sentence type is pragmatically meaningful. Illocutionary force inferences are then highly variable, context-dependent inferences that emerge in part from these conventional effects.

For English, these conventions often seem to make reference only to sentence type. However, phenomena like rising declaratives (*That's a persimmon?*; Gunlogson

2001) indicate that the conventions can actually be proper type + tune conventions. Rising declaratives share a sentence type with declaratives but associate with a very different convention. Roughly, whereas falling declaratives indicate speaker commitment, rising declaratives generally presume something more like addressee commitment (Gunlogson 2001), conditional commitment (Farkas & Roelofsen 2016), or projected commitment (Malamud & Stephenson 2015). The interactions between type and tune might be predictable in a roughly compositional fashion (Bartels 2014) or arbitrary. In either case, the convention itself crucially depends on both (see also Portner forthcoming for a related approach to rising imperatives). Outside of English, the case for type + tune conventions seems even stronger. For example, it is common for languages to distinguish declaratives from polar interrogatives only intonationally. Thus, although the burden may fall primarily on type or tune, the general characterization of these conventions in terms of type + tune seems justified.

Conventions like those in (1) are not about illocution per se, but they help explain our intentions and perceptions for illocution. The declarative-type convention ensures public commitment, which naturally supports acts of assertion. The interrogative-type convention indicates a preference for addressee commitment, which aligns with inquiry. And the imperative-type convention indicates a preference for a future outcome, which can interact with information in the context to create expectations for acts ranging from well-wishes to commands. Thus, it's not that the form-to-force relationship is conventionalized, but rather that it is constrained by convention.

In speech-act theory, the perlocutionary effects of an utterance are, according to Austin (1962: 101), "certain consequential effects upon the feelings, thoughts, or actions of the audience, or of the speaker, or of other persons". If we assert something to you, it might have the perlocutionary effect of persuading you. Then again, and contrary to our intentions, the effect might instead be to make you feel skeptical or confused. If you issue a gruff command to us, the perlocutionary effects might include making us feel socially diminished, or offended, or motivated. Though perlocution has received much less attention within speech-act theory than illocution, it is essential for understanding why speakers perform certain acts and for characterizing the effects of those acts (Solan & Tiersma 2005).

Are there irreducible sentence-type conventions relevant for how speakers instill perlocutionary effects in listeners? As with the sentence-type conventions discussed above, we wouldn't expect these conventions to *determine* the perlocutionary effects, but rather just to constrain them in predictable ways. Austin (1962) states clearly that perlocutionary effects are not conventional (as have many subsequent researchers; van Dijk 1977; Sadock 1974), but he seems open to having convention play this constraining role: "perlocutionary acts are not conventional, though conventional acts may be made use of in order to bring off the perlocutionary act" (Austin 1962: 121; see also Cohen 1973). To this end, consider the effects of hearing each of

the polar interrogatives in the examples below with falling, level (flat), and rising sentence-level intonation:

(2) a. Do manatees have molars?   (information-seeking bias)

   b. Do you want to go for a run?   (invitation bias)

   c. Can you close the window?   (request bias)

   d. Do you have a problem?   (ambiguous)

These sentences vary widely in their illocutionary biases. But our intuition is that the speaker's choice of intonation has a stable perlocutionary effect across all of these illocutions. For instance, a falling contour projects authority, a level contour annoyance, and a rising contour politeness. These effects of course interact with the illocutionary inferences made by the listener — an accusation can be only so polite, and an invitation can be only so annoyed — but their constancy across these different contexts suggests a role for convention. The case for convention is made stronger by the observation that these inferences are likely to be particular to polar interrogatives. For instance, falling contours in imperatives or declaratives seem to sound relatively more polite than falling contours in polar interrogatives. Similarly, rising contours in declaratives seem to sound less authoritative than rising contours in polar interrogatives, independent of whatever illocution they may carry. Thus, these seem to be conventional effects attached to specific type + tunes, guiding perlocutionary effects relating to the listener's view of the speaker's social persona.

3 Perception experiments

To validate and quantify the above intuitions about the stable effects of intonation on perlocution, we conducted two perception experiments in which we systematically manipulated the terminal contour intonation of a variety of sentence types, probing participants' judgments of the illocutionary acts and perlocutionary effects associated with the utterances. To keep the experiments to a manageable size, we focused on polar interrogatives (experiment 1) and wh-interrogatives (experiment 2), but included other sentence types (declaratives and imperatives) as well. Polar and wh-interrogatives are a useful minimal pair, since they associate with the same sentence-type convention for illocution but differ in their canonical intonational profiles (falling for wh-interrogatives, rising for polar interrogatives).

3.1 Participants

240 native speakers of American English (120 for each experiment) were recruited via Amazon Mechanical Turk as participants for the two experiments, which lasted 20–40 minutes.

3.2 Materials

Experiment 1 consisted of 16 polar interrogatives, 5 declaratives, 6 imperatives, and 4 wh-interrogatives. Example (2) above provides some actual polar-interrogative stimuli. The non-polar-interrogative sentences acted as fillers in experiment 1, as they were not systematically controlled for in their content and intonation. Hence, only the polar interrogative data were extracted from experiment 1 for further analysis. Experiment 2 consisted of 16 wh-interrogatives, 7 declaratives, and 8 imperatives. In addition to the main wh-stimuli, the declarative and imperative stimuli in experiment 2 were adequately controlled and were thus included in the data analysis. Experiment 2 also included two additional intonations for declaratives (low rising intonation) and imperatives (high level intonation), to ensure a natural range of stimuli and also to test a separate pilot hypothesis, but these additional intonational data will not be mentioned further in this paper.

In both experiments, the sentences were designed to systematically vary not only by sentence type but also by the illocutionary biases we expected them to have, which included not only canonical type-to-force associations (e.g., interrogatives used to obtain information) but also less canonical ones (e.g., interrogatives used to make requests). We also sought, as much as possible, to maximize the amount of thematic and illocutionary overlap between the sentences, to facilitate comparison between sentence types. A preference for sonorants (and avoidance of obstruents) was also a consideration in designing the sentences, to make the prosodic manipulations easier, but priority was given to natural-sounding sentences with more controlled illocutionary biases. The full list of sentences used in the experiment is given in appendix A.

The auditory stimuli were recorded in a sound-attenuated room. The sentences were produced by 6 native speakers of American English (2 male and 2 female for each experiment; a different set of female speakers was used for experiment 2). Each speaker participated in two recording sessions. During the first recording session, the speaker read through the target sentences as naturally as possible without any prior instructions. During the second, the speaker was asked to produce the sentences in level (flat) terminal contours, aided by a sample recording. The utterances with level terminal contours from the second session were used as the base materials for subsequent acoustic manipulations, after checking that they were comparable to the recordings produced in the first session in terms of their naturalness.

There were several reasons for using the recordings ending in level contours as baselines for further manipulations. First, this ensured that the subtle prosodic cues from the non-manipulated portion of the stimuli were not biased towards either rising or falling terminal contours. Second, the nuclear pitch accents of the utterances with level contours were ambiguous enough to be neither too high nor too low in

terms of their F0, such that they could be more easily interpreted as either *!H or L*, depending on the subsequent trajectories. This enabled the creation of maximally canonical 'nuclear pitch accent + terminal contour' units, i.e., high pitch accents followed by falling contours, and low pitch accents followed by rising contours. Finally, acoustic manipulations derived from mid-range F0 values produced more natural-sounding stimuli, compared to manipulations derived from high or low F0 values, which frequently involved creaks or other voice qualities that were less amenable to prosodic manipulations.

The level contours produced by our speakers were then manipulated to create falling and rising variants. Our procedure for this was as follows. First, the nuclear pitch accents (the last stressed syllable of a content word) and the endpoints of the utterances were located. Second, three new pitch values were assigned to the endpoints of the utterances. Finally, the pitch values of the nuclear pitch accent and the new endpoints were interpolated in a linear fashion to produce three distinct stimuli (rising, level, and falling) for each recording. The new pitch values at the endpoints were 10 st. higher than the nuclear pitch accent for the rising contour, 10 st. lower than the nuclear pitch accent for the falling contour, and 0.5 st. higher than the nuclear pitch accent for the level contour. Following ToBI transcription conventions (Beckman & Ayers 1997), the three contours corresponded to L* H-H%, !H* L-L%, and !H* H-L%, respectively. All manipulations were done in Praat (Boersma & Weenink 2015), using the built-in PSOLA pitch manipulation program. As an added precautionary measure, all manipulated stimuli were checked for naturalness by 3–4 native speakers of American English.

3.3 Procedure

Our two experiments were identical in their design and questions, differing only in their stimuli. In both experiments, each participant listened to all 31 sentences, each presented in a randomly chosen intonation among the three intonational patterns available. After listening to each sentence, six questions were posed:

(3)   Q1: Asked the participant to type in the sentence they heard, for verification.

        Q2: Probed the participant's understanding of the speaker's intended illocutionary force, by asking them to choose the most likely interpretation of the utterances from 4–6 different illocutions (the response choices reflected the range of pre-existing illocutionary biases of the sentences).

        Q3: A graded response (0–100) to the question 'How annoyed does the speaker sound?' (the annoyance rating).

        Q4: A graded response (0–100) to the question 'How authoritative does the speaker sound?' (the authority rating).

Q5: A graded response (0–100) to the question 'How polite does the speaker sound?' (the politeness rating).

Q6: A graded response (0–100, negative to positive) to the question 'What kind of attitude does the speaker have towards the listener?' (the stance rating).

Q7: An optional free format response about other possible interpretations of the utterance (other than the options given in Q2).

Q8: An optional free format response about other impressions of the speaker (other than the scales given in Q3–Q6).

At the end of the experiment, participants provided basic demographic information (gender, age, and ethnicity).

4 Results

We now review the primary results of our experiments, for illocutionary inferences (sec. 4.1), for perlocutionary effects (sec. 4.2) and their interactions with sentence types (sec. 4.3), and for the interactions between illocution and sentence types (secs. 4.4–4.5).
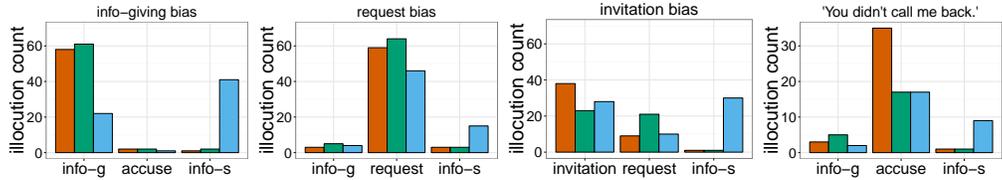
4.1 Participants' illocutionary inferences

In response to our illocutionary-oriented question (Q2), we saw the expected range of variation in participants' responses. In particular, our stimuli were pre-coded for what kind of bias we expected them to convey (e.g., 'information', 'invitation', 'request', or ambiguous, for polar interrogatives), and these categorizations proved to be significant predictors of participants' judgments. Fig. 1 summarizes the results for polar interrogatives (top) and declaratives (bottom). The categories on the x-axis represent the three most frequently chosen illocutionary responses (among the 4–6 choices available) for each group of examples, and the y-axis represents the raw counts of participants' choices. The intonation of the stimuli are color-coded to represent the three types of terminal contour.

The two left panels in fig. 1a represent polar interrogatives that have been pre-coded to represent an 'information-seeking' bias (e.g., *Are armadillos mammals?*) and a 'request' bias (e.g., *Can you open the door?*). These nearly always elicited the expected illocutionary responses. More generally, we found that intonation had almost no effect on participants' illocutionary judgments for sentences having strong, pre-existing illocutionary biases. In contrast, the two right panels in fig. 1a depict two specific examples of sentences that were precoded as having uncertain illocutionary biases. These left much more room for intonation to play a role in shaping
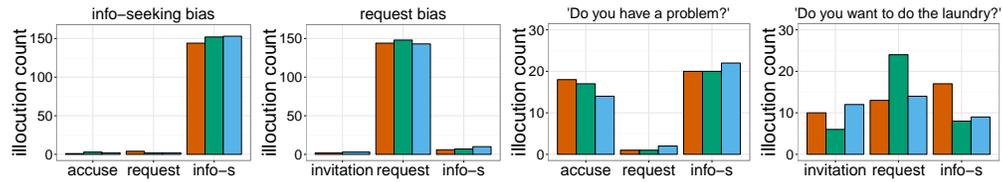
(a) Polar interrogatives.

(b) Declaratives.

Figure 1: Participants' illocutionary inferences. **Legend**: falling intonation: ■ level intonation: ■ rising intonation: ■

illocution. For *Do you have a problem?*, we see a split between information-seeking and accusation. Here, rising intonation (blue) elicits more 'information seeking' interpretations, whereas falling (red) and level (green) elicit more 'accusation' interpretations. Similarly, *Do you want to do the laundry?* produced a tripartite division (information seeking, invitation, request) in illocutionary responses, with falling intonation eliciting more 'information seeking', rising eliciting more 'invitation', and level more 'request'.

The pattern for declaratives (fig. 1b) is more complex. Rising intonation consistently elicited some 'information-seeking' responses across declarative sentences with diverse illocutionary biases; this corresponds to raised blue bars under 'info-s' in all the panels in fig. 1b. This choice was most dominant for sentences precoded as having an 'information-giving' bias (leftmost panel), suggesting that intonation is in fact driving a switch from giving to seeking information. Conversely, neither level nor falling ever elicited the 'information-seeking' illocution for declaratives whereas non-rising intonation still elicited the 'information-seeking' illocution for polar interrogatives, due to their interrogative syntax.

These findings broadly confirm arguments for the existence of separate illocution-oriented conventions for rising declaratives (Gunlogson 2001; Malamud & Stephenson 2015; Farkas & Roelofsen 2016), though the patterns also suggest that rising intonation does not always elicit illocutionary inferences that are categorically distinct from falling declaratives. For instance, for the declarative sentences with information-giving bias (leftmost panel in fig. 1b), rising intonation is consistent with an assertive, information-giving illocution (blue bar under 'info-g'), suggesting

a persistent ambiguity between rising declaratives as biased questions and rising declaratives as essentially assertions with added perlocutionary connotations in the sense of the often-discussed 'uptalk' or High Rising Terminal (HRT) contour phenomenon (Hirschberg & Ward 1995; McLemore 1991).

To summarize, it seems that participants' illocutionary inferences are driven primarily by biases stemming from sentence content (which in turn supports contextual inferences). Intonation has a qualitative effect on these choices primarily for sentences with ambiguous illocutionary biases, suggesting that the intonational effects were dominated by the sentences' content-related biases. These insights seem valuable in their own right, and they reassure us that participants made their perlocutionary judgments given a diversity of illocutions, helping us to separate these judgments from perlocution analytically.

4.2   Core tune conventions for perlocution

Our central hypothesis is that there are conventions oriented toward specific perlocutionary effects that are not predictable from content, context, and illocution alone. We thus predict that the answers to our perlocutionary-oriented Q3–Q6 will be consistent across different sentence contents and illocutionary force inferences. And this is what we generally find.

Fig. 2 summarizes the responses for these questions, broken down by sentence type and intonational contour (collapsing across different illocutionary inferences for now). The categories on the x-axis represent the four sentence types included in the experiments, and the y-axis measures the mean perlocutionary ratings for each sentence type, with error bars giving standard errors. From left to right, the panels represent annoyance, authority, politeness, and stance ratings. The picture is one of remarkable consistency: speaker annoyance is signaled by a level (green) contour, speaker authority by a falling (red) contour, and speaker politeness and positive stance by a rising (blue) contour, across all four sentence types. The results can also be summarized by their characteristic shapes across the intonational levels: 'annoyance' ∧, 'authority' ╲, 'politeness' ⌟, and 'stance' ⌟. (As we show in sec. 4.4, these consistent intonational effects hold not just across sentence types, but also across diverse illocutionary inferences.)

To explore these connections in greater detail, we fit a series of linear mixed-effects models to the combined data set from experiments 1 and 2, using the *lmerTest* package (Kuznetsova, Brockhoff & Christensen 2016) in *R* (R Core Team 2015). In each of these models, a single category of perlocutionary rating (annoyance, authority, etc.) provides the dependent (outcome) variable. The independent (predictor) variables are intonation, illocutionary choice, and sentence type. All the possible two-way and three-way interactions between intonation, illocution, and sentence
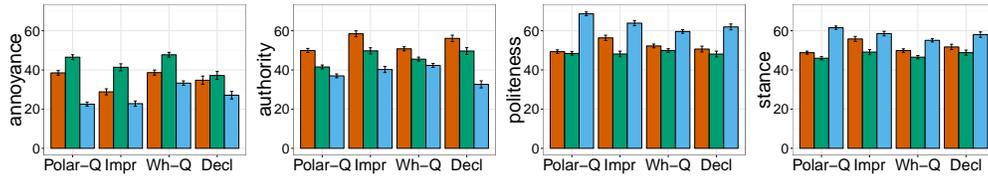
Figure 2: Means and standards errors of participants' perlocutionary ratings, across the four sentence types. Each perlocutionary category has a characteristic shape across the intonational levels, across sentence types: 'annoyance' ∧, 'authority' ╲, 'politeness' ⌐, and 'stance' ⌐. **Legend**: falling intonation: ▮ level intonation: ▮ rising intonation: ▮

type were included as well. (We expect these interactions to quantify the secondary effects discussed in the next two subsections.) As reference levels for each of the independent variables, falling intonation, polar interrogatives, and request/command illocution were set as default baselines, but were occasionally reset to facilitate comparison between all the factors involved. The request/command illocution was chosen as the default baseline because this was the only commonly chosen illocution shared by all four sentence types, which enabled a more straightforward interpretation of the sentence types as a separate predictor that can operate independently from their associated illocutions. As random effects, we initially tried to fit models with the maximal random effects structure for speakers and participants given our design (Barr, Levy, Scheepers & Tily 2011), but we were forced by convergence considerations to simplify the models to just random intercepts for speakers and participants. Full model summaries can be found in appendix B.

These models further confirmed our central hypothesis about the relationships between perlocution and intonation. We find a three-way distinction between the intonational contours for annoyance (level > falling > rising; $p < .001$ for all pairs), authority (falling > level > rising; $p < .05$ for all), and stance (rising > falling > level; $p < .05$ for all), and a two-way distinction for politeness (rising > falling, level; $p < .001$ for both), when polar interrogatives are set as the reference sentence type. Essentially the same significant intonational orders are supported when other sentence types are set as references, albeit with less granularity (e.g., a two-way instead of a three-way distinction for annoyance in wh-interrogatives). Crucially, these intonational effects hold for all illocutionary inferences for a given sentence type.

4.3   Secondary type + tune conventions for perlocution

In the previous section, we identified systematic connections between intonation and perlocution. In broad terms, these connections seem independent of sentence type. The consistency of these effects across different sentence types suggests that the 'tune' part of 'type + tune' is likely to be the primary element in English for perlocutionary conventions. However, in light of the discussion in sec. 2, we expect sentence type to play a role, though perhaps a secondary one. We now show that our experimental results support the existence of such properly 'type + tune' conventions. Our statistical models again help us quantify these findings. Here we review a few highlights; see appendix B for further details.

In some cases, sentence type and intonation seem to come together to amplify an expected perlocutionary effect. For instance, as is evident in fig. 2, panel 2, rising intonation (blue bars) systematically corresponds to lower authority ratings. The breakdown by sentence type shows that this is even more pronounced for declaratives than for other sentence types. Our statistical model for authority substantiates this observation via a significant interaction between the declarative sentence type and rising intonation (Coef. -13.04; $p < .01$). This effect might trace to the 'uptalk' phenomena mentioned briefly in sec. 4.1 above.

Conversely, the combined effect of sentence type and intonation can be to reduce an expected perlocutionary signal. For instance, one of the strongest primary effects in our model is that a level contour signals speaker annoyance (fig. 2, panel 1, green bars). Declaratives reduce this effect, though; for them, a level contour sounds less annoyed than for other sentence types. In our annoyance model, this corresponds to a significant negative interaction between declarative and level contour (Coef. -18.97, $p < 0.01$). Similarly, whereas the association between a rising contour and politeness is very robust overall (fig. 2, panel 3, blue bars), this effect is reduced for imperatives and wh-interrogatives, again as measured by interaction terms in the model (rise * Wh: Coef. -13.59, $p < .01$; rise * Impr: Coef. -6.81; $p < 0.01$).

It is tempting to think about these type + tune interactions as effects driven by departures from the norm. Each of the sentence types has a canonical contour, and we might expect that choosing a different contour will have a special pragmatic effect. This would relate to the general principle that marked forms tend to associate with marked meanings, and unmarked forms with unmarked meanings (Horn 1984; Levinson 2000). Might an account along these lines remove the need for separate type + tune conventions? We think the answer is no. While this principle is certainly relevant for understanding these effects, it does not explain their nature. For instance, given that imperatives and wh-interrogatives have a canonical falling contour, what should we expect from their uses with a rising contour? This type + tune combination is especially marked, so we might expect the (observed) reduction in politeness.

Then again, we might expect a *boost* in politeness, since the rise is so far from the fall. Similarly, no sentences have a default level contour, so such contours should be pragmatically weighty, but their mitigating effect on annoyance is seen only for declaratives. It thus seems inevitable that we will require special conventions to understand these nuances. For additional discussion, see sec. 5.

4.4 Consistency across illocutions

In fig. 2, we collapsed together all the participants' illocutionary inferences (responses to Q2 in (3)) in studying the perlocutionary effects. In fig. 3, we reveal the further numerical breakdown in perlocutionary choices by both sentence type (the rows) and by participants' subjective illocutionary inferences (along the x-axis of each panel). To simplify the displays, only the top three illocutionary choices for each sentence type are included. The raw counts for participants' top three illocutionary choices have also been provided for each sentence type.

The patterns we observed across different sentence types in fig. 2 remain evident: across all the possible choices of illocution, level contour (green) consistently signals annoyance, falling (red) signals authority, and rising (blue) signals politeness and positive stance. In a similar vein, looking column-by-column, we still see the characteristic shapes for 'annoyance' $\wedge$, 'authority' $\searrow$, 'politeness' $\lrcorner$, and 'stance' $\lrcorner$. The baseline level for these effects is influenced by perceived illocution, as expected given the nature of pragmatic inference, but the ordering remains stable. The consistency of these effects creates a striking contrast with the inconsistent and largely content-dependent intonational effects for illocution seen in fig. 1.

4.5 Interactions with illocution

Our models further corroborate the impressions above by explicitly controlling for illocution, sentence types, and their interactions (among other sources of information). There are, however, some clear illocution-dependent differences, which we can bring out by studying the illocution terms and type–illocution interaction terms.

First, the models show significant effects of illocutionary force on perlocutionary effects. All such effects run in intuitive, predictable directions. For instance, invitation illocution elicited lower annoyance ratings than request illocution (Coef. -7.09; $p < .05$), whereas request illocution elicited higher authority rating than invitation illocution (Coef. 5.92; $p < .05$). Effects like these stem from the underlying social factors that come along with performing different speech-acts; a command can only be so polite, an invitation only so authoritative.

Second, there are robust interactions between illocution and sentence type in predicting perlocutionary effects. We again refer to appendix B for a full summary.

(a) Polar interrogatives: info-seeking (834), invitation (393), request/command (543).

(b) Imperatives: request/command (667), wish (328), suggestion (298).

(c) Wh-interrogatives: info-seeking (1458), invitation (224), request/command (225).

(d) Declaratives: info-giving (319), invitation (117), request/command (266). We additionally had 139 responses for info-seeking, all but one with rising intonation.
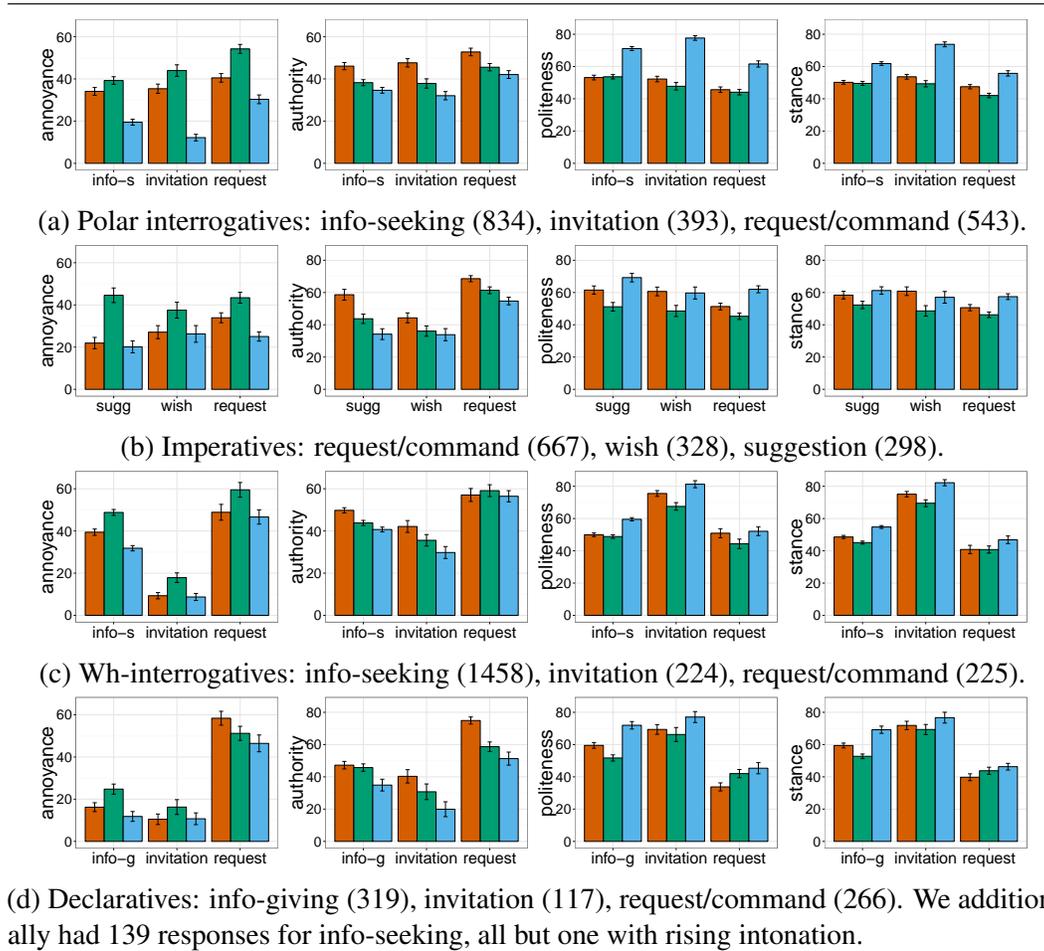
Figure 3: Mean ratings on perlocutionary effects, depending on illocution. **Legend**: falling intonation: ▮ level intonation: ▮ rising intonation: ▮

Here, we highlight just a few important examples.

There are significant interactions between invitation illocution and both declarative and wh-interrogative sentence types for annoyance, politeness, and stance ratings ($p < .001$ for all three). The same invitation illocution carries much less annoyance and much more positive stance and politeness when occurring with the declarative or the wh-interrogative than with the polar interrogative or the imperative. The two sentence types (declarative and wh-interrogative) show significant interactions with the request illocution as well, but the effects of these interactions run in the opposite direction. Here, the same request illocution carries much *more* annoyance and much less positive stance and politeness ($p < .05$ for all three) than the request illocution conveyed by polar interrogatives and imperatives.

Overall, then, it seems that illocution has a much greater effect on perlocution for declaratives and wh-interrogatives than it does for polar interrogatives and imperatives. This is immediately evident in fig. 3, where the first two rows show much less variation (with respect to bar heights across illocutions) within panels than do the bottom two rows. These effects arguably trace to the way these sentence types are used to perform various illocutionary acts. Polar interrogatives and imperatives are widely used to convey a diverse range of heterogeneous illocutions. In contrast, declaratives and wh-interrogatives seem invariably to involve assertion and inquiry, respectively, even when they are used to convey a range of other illocutions. Thus, where such sentence types do manage to achieve non-assertive, non-inquisitive illocutions like invitations and requests, they tend to amplify the potential perlocutionary effects that are associated with these illocutions.

Such an account also explains the few non-conforming patterns of declarative requests: as can be seen in the last rows of fig. 3, these were the only cases where the intonational effects didn't conform to the general pattern. In the mixed-effects models, this is seen in significant three-way interactions between declaratives, falling intonation, and request illocution. In the spirit of the explanation mentioned above, one might conjecture that there is something especially marked about declarative requests (usually arising via the use of strong necessity modals such as *need to* or *got to*), and that they are thus subject to slightly different intonational conventions.

5   Discussion

The results of sec. 4.2 and sec. 4.4 indicate that listeners make consistent social perlocutionary inferences across different sentence types and illocutionary inferences. These findings are strongly reminiscent of the consistency motivating sentence-type conventions oriented towards illocution like those in (1), and they lead us to conclude that similar conventions are guiding these perlocutionary phenomena. There seems to be a difference in the extent to which these different conventions mention type or tune in English, with illocutionary conventions being more sentence-type dependent and perlocutionary conventions more tune dependent, but these differences seem to be a matter of degree rather than kind; the full theory of both must be stated in terms of type + tune conventions.

We therefore propose the following conventions, parallel to those in (1) but oriented towards achieving certain perlocutionary effects in the listener:

(4)   Core tune conventions for perlocution

   a.   A speaker using falling intonation thereby signals that she seeks to be authoritative.

   b.   A speaker using level intonation thereby signals that she is annoyed.

    c.   A speaker using rising intonation thereby signals that she is polite and has a positive stance towards the listener.

These conventions play off each other as well. For instance, in opting to signal politeness and positivity with a rising tune, the speaker will in turn be perceived as avoiding a falling intonation, which will in turn support inferences that she is not authoritative, or does not wish to invoke her authority at that point.

We have formulated these core conventions for perlocution with respect to the core associations between specific tunes and specific effects. More precise statements referring to the orderings between the three types of terminal contours for a given perlocutionary effect could be made (sec. 4.2).

Our proposed conventions refer to the signal itself, rather than to the intentions of the speaker. This is an important feature of perlocution: it is not entirely under the speaker's control. A speaker using level intonation might not intend to sound annoyed, but that is likely to be what the listener perceives in many situations. Likewise, one's use of a rising tone might be accidental but nonetheless sound polite and positive. And so forth for the full range of effects we're studying. The primary issue is how the listener feels about the speaker as a social actor. (Misunderstandings about these issues do arise and can be resolved, just as with other acts, communicative and otherwise.)

Of course, speakers will (more often than not) intentionally employ these conventions to achieve their desired effects. As we noted above, Austin (1962: 121) anticipated that "conventional acts may be made use of in order to bring off the perlocutionary act". If a speaker *wishes* to sound authoritative, she may use falling intonation with the goal of having this perlocutionary effect on the listener. Conversely, she may use rising contours to help establish a positive, polite tone.

Although our perlocution-oriented conventions explicitly mention only a few perlocutionary effects (annoyance, authoritativeness, politeness, and stance towards the listener), we assume that each stands in for a family of social and interactional meanings. For example, rising contours that signal positive stance likely also signal geniality and friendliness. Although there is no easy way of definitively fleshing out or quantifying such associations, participants' free-form answers to Q7 and Q8 in (3) provide a window into these potential clusters. For instance, frequent answers to Q8 for rising polar interrogatives involved assessments such as "the speaker sounds optimistic", "friendly", and "encouraging".

In addition to these core tune conventions, our results also support a series of secondary conventions that reference both sentence type and tune (sec. 4.3):

(5)   Secondary type + tune conventions for perlocution

    a.   A speaker who utters a declarative with a level tune signals annoyance to a lesser degree than for other sentence types.

    b.   A speaker who utters an imperative or wh-interrogative with a rising tune signals politeness to a lesser degree than for other sentence types.

    c.   A speaker who utters a declarative with a rising tune signals even lower authority than for other sentence types.

We expect similar statement to be made in the illocutionary realm to capture, for example, the effects of rising declaratives. These secondary type + tune conventions are defined with explicit reference to both the intonation and the sentence type. They are derived directly from the significant interactions between sentence type and intonation mentioned earlier, and are predicted to overlay on top of the core tune conventions. The first two mitigate main effects of intonation, and the third boosts a main effect of intonation.

The question naturally arises where these type + tune conventions come from. Are they completely arbitrary, or can they be grounded, in whole or in part, in independent features of communication? For instance, Ohala (1983) finds cross-linguistically robust correlations between questions and rising intonation, and between assertions and falling intonation. Based partly on these cross-linguistic tendencies in canonical type + tune associations, the association between tunes and perlocutions might also be sound-symbolic in nature. However, even if sound-symbolism does play a role, the connections mentioned above cannot provide a full explanation of how the perlocutionary conventions came to be the way they are. For instance, it would be very hard to explain why rising intonation, in addition to being standardly associated with questions or inquisitiveness, signals politeness in English, and not alertness, say. Thus, we follow Gussenhoven (2002) in assuming that, though the type + tune associations may have sound-symbolic origins, they are conventionalized in specific ways across different languages.

Another potential source of perlocution-oriented conventions lies in the power of departing from a norm. People's knowledge of the default, canonical type + tune associations of their language may induce them to infer additional meaning upon encountering non-canonical associations. Although such an explanation seems relevant for explaining the secondary type + tune perlocutionary conventions we identified (as noted in sec. 4.3), it is not sufficient. It can act as a trigger for additional pragmatic inferences, but it has little power to explain the precise nature of those inferences, which are highly particularized and vary in more ways than this kind of account would seem to predict. We also note that this kind of account would predict that cross-linguistic differences in canonical type + tune associations (attested in Jun & Foreman 1996; Gordon 1999, contra the universality observed in Ohala 1983) will directly impact how type + tune combinations relate to perlocutionary effects. A convention-based approach is less committed to this broad claim.

While we cannot give a definitive answer here, it seems most likely to us that

the type + tune perlocutionary conventions argued for in this paper derive from a complex mix of arbitrary associations, sound symbolism, and pragmatic expectations relating to using marked forms.

6 Conclusion

We presented two perception experiments whose results suggest that there are separate, context-independent conventions for perlocution that are signaled by specific type + tunes, and distinct from illocutionary acts and inferences. Such results suggest that the conventions of language extend beyond meanings that are relevant to the immediate functional effects of communication to include more subtle interactional information relating to style, stance, and other kinds of social meaning.

Furthermore, the parallels between illocution-oriented and perlocution-oriented conventions suggest that the relationship between illocution and perlocution may need to be seen in a new light. In traditional speech-act theory, illocution is established and then, together with context, determines perlocution. Our proposal seeks to connect perlocution more directly to language. More precisely, we have argued that neither illocution-oriented nor perlocution-oriented conventions directly prescribe illocution or perlocution, but rather constrain them in predictable ways by changing the discourse context. Since contexts shape both illocution and perlocution, we can now envisage much richer interactions between illocution and perlocution, with each influencing the other. Adopting this perspective, one might speculate that the distinction between illocution-oriented and perlocution-oriented conventions may ultimately collapse; for us, it imposes a useful broad distinction, but it need not mark an ineluctable boundary.

We also see the potential for rich connections between our results and those for discourse particles, expressives, and other non-truth-conditional meanings relating to epistemic or evidential bias (see also Ward & Hirschberg 1985; Davis 2008). All these phenomena involve meanings that are separate from the core descriptive content of the utterances involved but cannot be predicted from automatic pragmatic inferences, thus necessitating a role for convention.

A Sentences used in the experiments

Experiment 1: polar interrogatives

Info-seeking bias

Are armadillos mammals?
Do manatees have molars?
Did Maria bring those bananas?
Do ants sleep?

Request bias

Can you open the door?
Can you close the window?
Can you carry this box?
Can you lend me some money?

Invitation bias

Do you wanna go to the movies?
Do you wanna grab a bite?
Do you wanna play games?
Do you wanna go for a run?

Ambiguous bias

Do you wanna do the laundry?
Do you have a problem?
Did you do the dishes?
Did you file the report?

Experiment 2: wh-interrogatives, declaratives, and imperatives

Info-seeking bias (wh)

Where do armadillos live?
How do manatees swim?
Who gave Maria those bananas?
When do ants sleep?

Request bias (wh)

Who has a pen?
Where's my boarding pass?
When are you going to clean your room?
How about turning down the volume on the TV?

Info-giving bias (dec)

Hippos are predators.
Manatees have molars.

Request bias (dec)

You need to help me carry this box.
You gotta close the window.

Advice bias (imp)

Take these pills for a week.
Avoid the highway.

Request bias (imp)

Have the report on my desk by noon.
Hand in the assignment by Friday.

Invitation bias (wh)

Where do you wanna go for the movies?
What do you say we go grab a bite?
When do you wanna play games?
How about we go for a run?

Ambiguous bias (wh)

What are you worried about?
Where's your bag?
Who ate my cupcake?
Where did you put my stuff?

Invitation bias (dec)

We should go get beer.
We can go dancing.

Ambiguous bias (dec)

You didn't call me back.

Well-wish bias (imp)

Get well soon.
Enjoy your dinner.

Offer, invitation bias (imp)

Let's have dinner.
Take a cookie.

B Summary of the mixed-effects models

All models use polar interrogative, request illocution, and falling intonation as the reference levels for each of the three independent variables.

	Coef.	S.E.	$p<$
intercept	41.57	3.47	.001
level	12.31	2.64	.001
rise	−11.87	2.67	.001
info-s	−8.50	2.42	.001
invitation	−7.09	2.77	.05
(Imp) sugg	−17.55	3.64	.001
(Imp) wish	−16.53	3.27	.001
Wh	7.60	3.63	.05
Decl	13.54	3.84	.001
rise * inv	−11.16	4.01	.01
level * info-s	−7.56	3.40	.05
level * sugg	14.66	4.98	.01
rise * wish	12.68	4.99	.05
level * Decl	−18.97	4.97	.001
invitation * Wh	−24.32	4.92	.001
info-g * Decl	−41.60	15.20	.01
inv * Decl	−29.08	5.92	.001
rise * inv * Decl	25.77	8.80	.01

Annoyance

	Coef.	S.E.	$p<$
intercept	47.49	2.23	.001
level	−4.64	1.95	.05
rise	8.97	1.97	.001
invitation	6.99	2.04	.001
suggestion	11.50	2.68	.001
wish	15.86	2.41	.001
Decl	−6.01	2.82	.05
Wh	−6.52	2.66	.05
rise * inv	11.16	2.96	.001
level * wish	−9.71	3.51	.01
rise * wish	−12.48	3.68	.001
level * Decl	7.53	3.66	.05
inv * Wh	23.28	3.63	.001
inv * Decl	20.08	4.37	.001
inv * Imp	12.97	4.54	.01
rise * inv * Wh	−12.17	5.25	.001
rise * inv * Decl	−13.88	6.49	.001
rise * inv * Imp	−16.17	6.29	.001

Stance

	Coef.	S.E.	$p<$
intercept	51.84	2.32	.001
level	−7.65	2.33	.01
rise	−11.88	2.36	.001
info-s	−7.40	2.14	.001
invitation	−5.92	2.44	.05
(Imp) sugg	−14.63	3.21	.001
(Imp) wish	−29.22	2.88	.001
Decl	23.81	3.40	.001
Impr	10.08	1.94	.001
rise * sugg	−9.17	4.50	.05
rise * Wh	9.30	4.25	.05
level * Decl	−11.60	4.38	.01
rise * Decl	−13.04	4.74	.01
info-s * Decl	−19.56	4.42	.001
inv * Decl	−25.00	5.22	.001
inv * Impr	−22.69	5.43	.001

Authority

	Coef.	S.E.	$p<$
intercept	45.60	2.94	.001
rise	16.99	2.21	.001
accusation	−14.99	3.98	.001
info-s	8.98	2.01	.001
invitation	7.49	2.29	.01
suggestion	13.80	3.01	.001
wish	15.72	2.71	.001
Decl	−9.22	3.18	.01
rise * invitation	8.92	3.32	.01
rise * Wh	−13.59	3.99	.001
rise * Impr	−6.81	3.18	.05
info-s * Wh	−9.41	3.28	.01
inv * Wh	12.52	4.07	.01
inv * Decl	19.69	4.90	.001

Politeness

References

Austin, John Langshaw. 1962. *How to do Things with Words*. Cambridge, MA: Harvard University Press 2nd edn.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2011. Random effects structure in mixed-effects models: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. doi:10.1016/j.jml.2012.11.001.

Bartels, Christine. 2014. *The Intonation of English Statements and Questions: A Compositional Interpretation*. New York: Routledge.

Beckman, Mary E. & Gayle Ayers. 1997. Guidelines for ToBI labelling, version 3. The Ohio State University Research Foundation.

Boersma, Paul & David Weenink. 2015. Praat: Doing phonetics by computer (version 5.3.51). http://www.praat.org.

Cohen, Ted. 1973. Illocutions and perlocutions. *Foundations of Language* 9(4). 492–503.

Condoravdi, Cleo & Sven Lauer. 2011. Performative verbs and performative acts. In Ingo Reich (ed.), *Sinn und Bedeutung 15 (SuB 15)*, 1–15. Saarbrücken, Germany: Saarland University Press.

Condoravdi, Cleo & Sven Lauer. 2012. Imperatives: Meaning and illocutionary force. In Christopher Piñón (ed.), *Empirical Issues in Syntax and Semantics* 9, 37–58.

Davis, Christopher. 2008. Decisions, dynamics, and the Japanese particle *yo*. *Journal of Semantics* 26(4). 329–366. doi:10.1093/jos/ffp007.

van Dijk, Teun A. 1977. *Text and Context: Explorations in the Semantics and Pragmatics of Discourse*. London and New York: Longman.

Farkas, Donka & Floris Roelofsen. 2016. Division of labor in the interpretation of declaratives and interrogatives. Ms., UC Santa Cruz and University of Amsterdam.

Gordon, Matthew K. 1999. The intonational structure of Chickasaw. In *14th International Congress of Phonetic Sciences (ICPhS 14)* 3, 1993–1996.

Green, Mitchell. 2007. Speech acts. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/speech-acts/.

Gunlogson, Christine. 2001. *True to form: Rising and falling declaratives as questions in English*: UC Santa Cruz PhD dissertation.

Gussenhoven, Carlos. 2002. Intonation and interpretation: Phonetics and phonology. In Bernard Bel & Isabelle Marlien (eds.), *Speech Prosody 2002*, 47–57.

Hirschberg, Julia & Gregory Ward. 1995. The interpretation of the high-rise question contour in English. *Journal of Pragmatics* 24(4). 407–412. doi:10.1016/0378-2166(94)00056-K.

Horn, Laurence R. 1984. Toward a new taxonomy for pragmatic inference: Q-based

and R-based implicature. In Deborah Schiffrin (ed.), *Meaning, Form, and Use in Context: Linguistic Applications*, 11–42. Washington, D.C.: Georgetown University Press.

Jun, Sun-Ah & Christina Foreman. 1996. Boundary tones and focus realization in African-American intonation. Paper presented at the 3rd joint meeting of the Acoustical Society of America and the Acoustical Society of Japan.

Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2016. lmertest package (version 2.0.30).

Lauer, Sven. 2013. *Towards a dynamic pragmatics*: Stanford University PhD dissertation.

Levinson, Stephen C. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

Malamud, Sophia A & Tamina Stephenson. 2015. Three ways to avoid commitments: Declarative force modifiers in the conversational scoreboard. *Journal of Semantics* 32(2). 275–311. doi:10.1093/jos/ffu002.

McLemore, Cynthia Ann. 1991. *The pragmatic interpretation of English intonation: Sorority speech*: University of Texas at Austin PhD dissertation.

Ohala, John J. 1983. Cross-language use of pitch: An ethological view. *Phonetica* 40(1). 1–18. doi:10.1159/000261678.

Portner, Paul. 2007. Imperatives and modals. *Natural Language Semantics* 15(4). 351–383. doi:10.1007/s11050-007-9022-y.

Portner, Paul. forthcoming. Commitment to priorities. In Daniel Fogel, Daniel Harris & Matt Moss (eds.), *New Work on Speech Acts*, 1–26. Oxford University Press.

R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Sadock, Jerrold M. 1974. *Toward a Linguistic Theory of Speech Acts*. New York: Academic Press.

Schmerling, Susan F. 1982. How imperatives are special and how they aren't. In Robinson Schneider, Kevin Tuttle & Robert Chametsky (eds.), *Chicago Linguistic Society (CLS) Parasession on Nondeclaratives*, 202–218. Chicago Linguistic Society.

Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press. doi:10.1017/CBO9781139173438.

Solan, Lawrence M. & Peter M. Tiersma. 2005. *Speaking of Crime: The Language of Criminal Justice*. University of Chicago Press.

Ward, Gregory & Julia Hirschberg. 1985. Implicating uncertainty: The pragmatics of fall-rise intonation. *Language* 61(4). 747–776. doi:10.2307/414489.

Sunwoo Jeong
Department of Linguistics
Building 460
Stanford University
sunwooj@stanford.edu

Christopher Potts
Department of Linguistics
Building 460
Stanford University
cgpotts@stanford.edu