

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

The Sound of Pedagogical Questions

Permalink

<https://escholarship.org/uc/item/08g0s986>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN

1069-7977

Authors

Bascandziev, Igor
Shafto, Patrick
Bonawitz, Elizabeth

Publication Date

2021

Peer reviewed

The Sound of Pedagogical Questions

Igor Bascandziev^{1,2}

Patrick Shafto¹

Elizabeth Bonawitz²

1. Rutgers University -- Newark
2. Harvard Graduate School of Education

Abstract

Questions are prevalent in everyday speech and they are often used to teach (Newport, Gleitman, & Gleitman, 1977; Siraj-Blatchford & Manni, 2008). When learners receive explicit cues that the intent of a question is pedagogical, they draw inferences that lead to superior learning (Yu, Landrum, Bonawitz, & Shafto, 2018). Although the ability to infer pedagogical intent is critical, very little is known about the mechanisms that support the inference that any particular statement is pedagogical or not. We tested the hypothesis that the prosody of speech marks the intent of pedagogical and information-seeking questions. In Studies 1 and 2, 256 naïve participants rated 100 pedagogical and information-seeking questions, spoken in child- or adult-directed speech. We found that naïve listeners can accurately infer pedagogical intent on the basis of prosody alone. In Study 3, we begin charting the acoustic features that differentiate pedagogical from information-seeking questions. We found that pedagogical questions are longer in duration, have lower F0 variability, and are characterized with a non-canonical pitch contour compared to information-seeking questions. These findings provide a window into the mechanisms that allow learners to infer pedagogical intent in otherwise ambiguous situations.

Keywords: pedagogical questions; information seeking questions; prosody; child-directed speech

Introduction

Questions are ubiquitous in everyday speech. Infants in the US begin receiving questions as early as 5 months-of-age (Bornstein, et al., 1992), and almost half of the utterances that children hear when they are between 12 and 27 months-of-age are questions (Newport, Gleitman, & Gleitman, 1977). Receiving questions is ubiquitous in later childhood and adulthood as well. At school, students hear 300-400 questions per day (Leven & Long, 1981).

Importantly, however, not all questions are the same. Sometimes the intent of a question is to serve the asker (information-seeking) and sometimes the intent of a question is to serve the listener (pedagogical). The recipient's inference about the goal of the question has important consequences for learning. For example, when learners receive *explicit* cues that the intent of a question is pedagogical (e.g., when the knowledge state of the asker is known), they draw different inferences, and they exhibit superior learning outcomes (Yu, Landrum, Bonawitz, & Shafto, 2018; Jean, Daubert, Yu, Shafto, & Bonawitz, 2019). Despite the well documented role that people's inferences about the pedagogical or non-pedagogical goal of a question

plays in learning, very little is known about how learners infer pedagogical intent in everyday situations.

We propose here that a plausible mechanism for communicating pedagogical intent of questions is the prosody of speech. Such a proposal is plausible for multiple reasons. Prosody is a reliable signal that is always present in speech. It can carry information about intent. And, there is evidence that sensitivity to prosody exists even in infancy. For example, even very young infants are sensitive to different registers of speech, such as child-directed speech (Saint-Georges et al., 2013). However, it remains an open question whether the two primary goals of questions, teaching and information-seeking, are also differentiated by strong, reliable prosodic cues, to which listeners are sensitive to and use to inform their inferences from queried content.

There are at least three steps to testing the hypothesis that learners infer pedagogical intent from prosody. The first is to test whether naïve listeners can accurately *detect* differences in prosodic cues that were intended to be pedagogical or information-seeking. The second step would be to show that the prosodic cues that are *recognized* by naïve listeners as pedagogical are also present in speech in naturalistic contexts. The third step would be to show that learners draw different inferences and exhibit different learning outcomes on the basis of prosody alone. In the present study, we take only the first step. We test whether naïve listeners are sensitive to prosodic cues that the asker intended to be pedagogical or information-seeking, and we begin to chart the prosodic features of each.

Pedagogical Communication

Pedagogical communication may be a particularly powerful tool to support human learning (Carey, 2009; Csibra & Gergely, 2009; Harris, 2012; Tomasello, 2008). Learning from teaching differs from standard learning in two critical ways. First, the data that the learner observes are chosen by a knowledgeable and helpful person, not selected at random. Second, the learner infers that the teacher is choosing these data as a knowledgeable and helpful person. In this way, pedagogical communication differs from other forms of communication in that the speaker's goal is for the learner to learn. Critically, the inference by the learner (that the teacher is knowledgeable, helpful, and intends to teach) is what supports such rapid inferences from very few samples (Bonawitz & Shafto, 2016).

Pedagogical Questions. Questions vary according to whether they serve the questioner (as in information-seeking questions, ISQs) or the listener (as in pedagogical questions, PQs). Information seeking questions from a parent to a child include clarifying children’s utterance (“What?”), checking children’s status (“Are you sleepy?”), asking for permission (“Can I see this toy?”) and offering help (“Do you need help?”). These questions are asked because the questioner does not have the knowledge and is seeking information from the child. However, parents also use questions for pedagogical purposes, providing queries despite having knowledge of the correct answer. For example, questions are asked to elicit children’s attention (“See?”), to test children’s knowledge (“Is this green or red?”), to encourage children to keep talking (“Oh really?”), to interpret children’s actions (“You putting the doggie in a chair?”), and to report a fact (“That car doesn’t fit?”) (Olsen-Fulero & Conforti, 1983).

A large survey of the CHILDES database found that pedagogical questions are common but used with different frequencies in different daily contexts and across SES (Yu, Bonawitz, & Shafto, 2017). A follow-up study found that the frequency with which pedagogical questions are used in everyday parent-child interactions predicts children’s exploration following an experimenter’s pedagogical question on a new task (Yu, et al., 2020). Furthermore, when children are given explicit cues that the question is pedagogical, they engage in more exploration and discovery of new features of a novel toy, and they are more persistent at exploring a causal function that is not working (Jean et al., 2018; Yu, et al., 2020).

Unlike in the studies cited above, learners in everyday situations rarely receive explicit cues about the intent of the questions they receive. Most conversations are rife with ambiguity regarding the content, context, knowledge, and intentions of the speaker. That is, when a speaker asks a pedagogical question, she does not necessarily reveal that she already knows the answer, and she does not necessarily state that her intention is to teach. Because of this absence of explicit cues to pedagogy in everyday situations, it may be quite difficult for a learner to infer that a particular question is a pedagogical question. One possibility, however, is that in addition to using the content, context, or knowledge-state of the speaker, learners also use prosodic cues to infer that a particular question is a pedagogical question. Unlike explicit cues to pedagogy (e.g., “*I know the answer to this, but I am asking you because I want to teach you...*”) that are typically absent from speech, prosodic cues can mark pedagogical questions without interrupting the natural flow of the conversation. This raises the question of whether and how naïve listeners can recognize questions as pedagogical on the basis of prosody alone?

Although there is a large literature on the prosody of different kinds of questions (e.g., rhetorical, information-seeking, indirect requests, etc. (See, Bartels, 1999; Banuazizi & Cresswell, 1999; Braun, et al., 2019; Han, 2002; Trott, Reed, Ferreira, & Bergen, 2019), to our knowledge, there are no studies that have documented the prosody of pedagogical

questions. One particular kind of prosody, however, that has been investigated extensively in the context of how the prosody of speech can facilitate learning, is the prosody of child-directed speech (Eaves, Feldman, Griffiths, & Shafto, 2016; Fernald, 1985; Fernald & Simon, 1984; Nencheva, Piazza, & Lew-Williams, 2020; Rowe, 2008; Thiessen, Hill, & Saffran, 2005). Although several studies have shown that the prosody of child-directed speech is also interpreted as an ostensive signal by learners (Gergely, Egyed, & Kirali, 2007; Senju & Csibra, 2008), we note that child-directed speech *cannot* be a reliable signal of pedagogy. This is so because adults use child-directed speech across various situations that do not entail teaching about the world. Also, when children are older, adults no longer use child-directed speech even when they teach (Liu, Tsao, & Kuhl, 2009). We note, therefore, that the main goal of the present study is *not* to extend the child-directed speech research program. Rather, we use child-directed speech as an important analogous domain from which our questions about the novel domain of pedagogical prosody are partly inspired.

Present Studies

Studies 1 and 2 test the hypothesis that naïve adult listeners can differentiate between pedagogical and information-seeking questions on the basis of prosody alone, both within adult-directed and within child-directed speech. Study 3 begins to model the prosodic differences between pedagogical and information-seeking questions at a word level and at a level of an utterance.

To achieve these goals, we created stimuli in English (Study 1) and Macedonian (Study 2) language, where 10 speakers recorded 100 questions in 4 different ways: as a child-directed pedagogical or information-seeking question; and as an adult-directed pedagogical or information-seeking question. Next, we asked naïve MTurk listeners to rate the questions on both dimensions (i.e., pedagogical/ information-seeking; and child/ adult-directed). In Study 3, we analyzed the stimuli created for Studies 1 and 2. In addition, we selected a subsample of accurately rated questions for further acoustic analyses. All three studies were approved by the Institutional Review Board at Rutgers University -- Newark.

Study 1

Participants

A sample of 128 participants was recruited from the US on Amazon’s Mechanical Turk platform. Two participants were excluded and replaced because they completed the experiment in an extremely short timeframe (338 and 270 seconds respectively), which is incommensurable with listening to the audio clips. The participants’ average age was 39 years (range = 19–72; SD = 12.73). Sixty-five participants identified as male, sixty-two as female, and one participant identified as non-binary genderqueer. Sixty-five participants identified as parents and forty said that they had worked with children in some professional capacity (e.g., a teacher, a nanny, etc.).

Stimuli

The stimuli were recorded by 8 speakers (four females and four males) who are native speakers of American English, who come from diverse ethnic and cultural backgrounds, and have extensive experience interacting with children (e.g., as parents or developmentalists). The stimuli consisted of 100 short questions covering the physical, biological, and social domains (e.g., “Is the speed of light the same in all mediums?”; “How do dolphins sleep?”; “Why do people gossip?”). The list of all questions and the context that was given to differentiate between PQs and ISQs can be found at osf.io/kh7xn

Each speaker was asked to read and record all 100 questions, 4 different times (a total of 400 recorded questions per speaker): 1) as an adult-directed pedagogical question, 2) as an adult-directed information seeking question, 3) as a child-directed pedagogical question, and 4) as a child-directed information seeking question. Audio examples can be found at osf.io/kh7xn. Before the recording, the speakers were told descriptively what a pedagogical question is (i.e., a question where the asker knows the answer and asks the question with an intention to teach), what an information-seeking question is (i.e., a question where the asker does not know the answer and asks the question with an intention to learn), and what child-directed and adult-directed speech is. An important aspect of this design is that it allowed a comparison of the exact *same* questions that were recorded as pedagogical and as information-seeking questions. Importantly, the speakers were not told how the stimuli should sound. Each speaker was left on their own to decide how child-directed and adult-directed PQs and ISQs should sound. To aid the recording of the stimuli, each question was preceded by a context sentence providing a pedagogical or an information-seeking context. For example, the information-seeking context sentence that preceded the question “How do dolphins sleep” was “I have no idea how animals that live in the sea sleep. Do you know...” The pedagogical context sentence for the same question was: “I learned about this at the aquarium a few days ago. Do you know...” Importantly, the ending of the provided context (e.g., “Do you know...”) was the same across pedagogical questions and information seeking questions. The speakers read the questions in blocks. The order of blocks was counterbalanced across speakers and the order of questions within each block was randomized.

Design

We created 4 different surveys from the 400 questions recorded by each speaker. Each of the 4 surveys included 100 questions, 25 of each question type. Together, the 4 surveys contained all 400 questions generated by the speaker. Each survey was presented to 4 participants on Mechanical Turk (4 participants per survey x 4 surveys per speaker x 8 speakers = 128 Mechanical Turk participants).

Procedure

The surveys were prepared and administered via Qualtrics. After answering a few demographic questions, participants

read a description of what pedagogical questions are and what information seeking questions are, and also a description of what is meant by child-directed and adult-directed questions. Following that short training, participants were tasked with listening to the audio clips (a total of 100) and answering two questions about each audio clip: 1) Is the question you heard: a) A Pedagogical Question; b) An Information Seeking Question; and 2) Is the question you heard: a) A Question Directed to a Child; b) A Question Directed to an Adult. The 100 audio clips were randomized within each survey. The order in which the two questions appeared (PQ vs ISQ or CD vs AD) and the order in which the two alternative answers appeared after each audio clip were counterbalanced.

Results and Discussion

The main goal of this study was to establish whether naïve listeners can decode the pedagogical and the information-seeking intent of a question. Recall that for each of the 100 audios, participants answered two questions: i) Is the question you heard a) A pedagogical question or b) An information-seeking question; and ii) Is the question you heard a) directed to a child or b) directed to an adult. Each incorrect response received a 0 and each correct response received a score of 1. This allowed us to compute an accuracy score (proportion correct) for each participant (e.g., the participant X correctly identified the pedagogical and information-seeking questions X% of the time (out of 100 questions) and correctly identified that it was directed to a child or an adult X% of the time). We found that the average accuracy score (across all participants) for discriminating pedagogical from information seeking questions was $M_{PQ/ISQ} = .58$, and it was significantly above chance level performance ($t(127) = 6.97, p < .001, d = .62$).

We also computed an accuracy score (proportion correct) for discriminating child-directed from adult-directed speech. Replicating prior research, we found that the average accuracy score $M_{CD/AD} = .56$ was also significantly different from chance performance ($t(127) = 7.42, p < .001, d = .67$).

Table 1 presents the accuracy scores for discriminating pedagogical from information-seeking questions within child- and adult-directed speech.

Table 1. Accuracy for discriminating pedagogical questions from information seeking questions

	ADPQ	ADISQ	CDPQ	CDISQ
Accuracy	.56**	.59***	.65***	.53

Note: One-sample t-test comparisons against chance.

Key: ~ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Inspection of Table 1 reveals some asymmetry in the accuracy of detecting pedagogical or information-seeking questions within different categories of speech (child vs. adult directed). For example, the accuracy score for information-seeking questions spoken in child-directed speech was not different from chance. There are at least two

possible interpretations of this result. One is that some (but not all) prosodic features of pedagogical questions overlap with the prosody of child-directed speech, while some (but not all) features of information seeking questions overlap with the prosody of adult-directed speech. Another, not mutually exclusive possibility, is that listeners might have an expectation that pedagogical questions are typically directed to children and information seeking questions are typically directed to adults. We will return to this issue in the General Discussion.

In conclusion, Study 1 shows that the communicative intent to pose a pedagogical question or an information seeking question can be encoded in the prosody of the question and naïve listeners can correctly decode the intent of the speaker by relying on the prosody alone. Are these cues present at the word level, or does the general “shape” and duration of a question inform this judgement? To explore whether speakers could discern pedagogical from information-seeking questions without relying on any syntactic or semantic cues, we repeated the same study but with audio stimuli recorded in Macedonian language.

Study 2

Participants

A sample of 128 participants was recruited from the US on Amazon’s Mechanical Turk platform. Six participants were excluded and replaced because they reported that they could understand Macedonian language. The participants’ average age was 39 years (range = 24–74; SD = 10.61). Sixty-eight participants identified as male, fifty-nine as female, and one participant identified as non-binary genderqueer. Sixty-seven participants identified as parents and forty-six said that they had worked with children in some professional capacity (e.g., a teacher, a nanny, etc.).

Stimuli

The stimuli were the same as in Study 1, except they were translated and recorded in Macedonian language by 2 speakers (1 male and 1 female). The stimuli were recorded by native speakers of Macedonian language. The two speakers, however, live and work in the US. Just like the English language stimuli, these stimuli were recorded by providing context sentences and imagining that the question is pedagogical or information-seeking and that the question is directed to an adult or a child.

Design

The design was the same as that of Study 1. Four different surveys were created from the 400 questions recorded by the speakers. Each survey contained 100 questions, 25 of each question type. Each survey was presented to 16 participants on Mechanical Turk (16 participants per survey x 4 surveys per speaker x 2 speakers = 128 Mechanical Turk participants).

Results and Discussion

The average accuracy score for discriminating pedagogical from information-seeking questions was $M_{PQ/ISQ} = .66$, and it was significantly different from chance performance, $t(127) = 11.7, p < .001, d = 1$.

Similarly, the average accuracy score for discriminating child-directed from adult-directed speech was $M_{CD/AD} = .65$ and it was also significantly different from chance performance ($t(127) = 13.92, p < .001, d = 1.25$). Table 2 presents the accuracy scores within child- and adult-directed speech.

Table 2. Accuracy for discriminating pedagogical questions from information seeking questions

	ADPQ	ADISQ	CDPQ	CDISQ
Accuracy	.66***	.67***	.68***	.63***

Note: One-sample t-test comparisons against chance.

Key: ~ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

These results confirm that naïve listeners can discriminate between pedagogical and information-seeking questions both within child-directed and within adult-directed speech on the basis of prosody alone, and in the absence of any syntactic or semantic cues. Which prosodic features differentiate the two types of questions is an issue that we begin addressing in Study 3.

Study 3

We investigated acoustic features that have been shown to be important in prior research on child-directed speech, such as the average fundamental frequency (F0) perceived as pitch (typically higher in CD speech), F0 variability (typically higher in CD speech), duration (also typically higher in CD speech), and pitch contour (typically exaggerated in CD speech) (e.g., Fernald & Kuhl, 1987; Fernald & Simon, 1984). In addition to research on child-directed speech, prior research on the prosody of questions has shown that questions’ pitch contours vary systematically and can be separated into canonical and non-canonical contours (Bartels, 1999). Importantly, research on inquisitive semantics has shown that the non-canonical pitch contours of wh- questions are associated with episodes where the asker has partial knowledge of the answer (e.g., Hedberg & Sosa, 2011), suggesting an additional promising factor to explore.

Data

We analyzed the total of 3200 audio files of the 100 matched questions recorded in English language by 8 speakers (800 ADPQs, 800 ADISQs, 800 CDPQs, and 800 CDISQs) and also the total of 800 audio files recorded in Macedonian language. In addition, in order to confirm that the average acoustic differences between PQs and ISQs across all audio files are in fact the acoustic cues that participants use to make judgments, we selected a subsample of audio files from each

question type on the basis of the accuracy ratings from Study 1. Recall that in Study 1, four MTurk participants rated each question, which allowed us to select questions that were accurately rated by 3 or 4 participants on the pedagogical/information-seeking dimension. This resulted in 143 pairs of perfectly matched adult-directed utterances (i.e., 143 accurately rated PQs and the same 143 accurately rated ISQs) and 180 pairs of perfectly matched child-directed PQs and ISQs. The audios from the subsample were then annotated, which allowed us to do an analysis at a word level. All analyses were done separately for child-directed and adult-directed questions.

Results and Discussion

Analysis of the full sample of audios¹. Table 3 summarizes the comparison of pedagogical and information-seeking questions on three dimensions (Mean F0, F0 Range, and Duration), separately for adult-directed and child-directed speech.

Table 3. Summary statistics of three acoustic features of PQs and ISQs within adult- and child-directed speech (Mean F0, F0 Range in HZ, and Duration in seconds)

<i>Adult Directed Questions</i>					
Mean F0		F0 Range		Duration	
PQ	ISQ	PQ	ISQ	PQ	ISQ
193	203	228	247	1.72	1.57
<i>Child Directed Questions</i>					
Mean F0		F0 Range		Duration	
PQ	ISQ	PQ	ISQ	PQ	ISQ
211	217	253	277	1.82	1.71

Inspection of Table 3 shows that the Mean F0 of pedagogical questions is lower than that of information-seeking questions, both within adult-directed ($t(799) = 8.57, p < .001$) and child directed speech ($t(799) = 4.23, p < .001$). The F0 Range is also lower in pedagogical than in information-seeking questions both within adult-directed ($t(799) = 5.47, p < .001$) and child-directed speech ($t(799) = 9.13, p < .001$). Finally, pedagogical questions are with longer duration than information-seeking both in adult-directed ($t(799) = 11.1, p < .001$) and child-directed speech ($t(799) = 7.94, p < .001$). Unlike the other cues, the longer duration of pedagogical compared to information-seeking questions was also true of the stimuli recorded in Macedonian

¹ To avoid making a Type I error, we applied conservative Bonferroni corrections across all analyses (18 comparisons = $p < .003$).

language with adult-directed ($M_{PQ} = 2.9$ and $M_{ISQ} = 1.9, t(199) = 22.85, p < .001$) and child-directed speech ($M_{PQ} = 3.04$ and $M_{ISQ} = 2, t(199) = 24.7, p < .001$). Indeed, the difference was quite prominent in the Macedonian questions, suggesting that this is likely the cue that was used by participants to correctly differentiate between different types of questions.

Analysis of the subsample of accurately rated audios. At the level of an utterance, we confirmed the difference in duration between pedagogical and information-seeking questions. Pedagogical questions were longer duration (in seconds) than information-seeking questions both within adult-directed speech ($M_{PQ} = 1.83$ s., $M_{ISQ} = 1.60; t(142) = 7.69, p < .001$) and within child-directed speech ($M_{PQ} = 1.96, M_{ISQ} = 1.75; t(179) = 10.17, p < .001$). Similarly, we confirmed that the F0 range was lower in pedagogical than in information seeking questions, both within adult-directed ($M_{PQ} = 277, M_{ISQ} = 319; t(142) = 2.99, p = .003$) and within child-directed speech ($M_{PQ} = 309, M_{ISQ} = 379; t(179) = 6.26, p < .001$), suggesting that participants relied on these cues to classify the different question types. The differences between PQs and ISQs in Mean F0 were not significant.

Word level analysis. We also explored how pedagogical questions differed from information-seeking questions at the word level. Here, we present an analysis of adult-directed questions only.² We present an analysis of the first word (a wh- question word), the second word (typically an auxiliary verb, e.g., is/are), and the last word in the utterance. The analysis of the first two and the last word can give us a glimpse into the pitch contours of utterances of variable length and it can also give us a glimpse into the potentially important differences between pedagogical and information-seeking questions in terms of how the wh- question words are stressed. For this analysis, we excluded the yes/no questions, because they do not have wh- question words and because they typically have a different pitch contour than wh- questions (Bartels, 1999). This exclusion resulted in a total of 130 adult-directed pairs of perfectly matched wh- PQs and ISQs.

Duration. The duration of the first and the second word of pedagogical questions was longer than the duration of the first and the second word of information-seeking questions ($M_{PQ1} = .34$ sec. vs. $M_{ISQ1} = .22$ sec., $t(129) = 9.51, p < .001$) and ($M_{PQ2} = .21$ sec. vs. $M_{ISQ2} = .17$ sec., $t(129) = 5.94, p < .001$). Given that pedagogical utterances were overall longer than information-seeking utterances, does this mean that all words in the utterance were longer in duration? Our analysis revealed that the average duration of the last word in the utterance in pedagogical questions was not significantly different from the duration of the last word in information-seeking questions ($M_{PQL} = .55$ sec. vs. $M_{ISQL} = .54$ sec., $p > .05$).

F0 (Pitch). The average pitch of the first and the second word of pedagogical questions was lower than the average

² The analysis of child-directed questions is ongoing.

pitch of the first and second word in information-seeking questions ($M_{PQ1} = 203.09$ HZ vs. $M_{ISQ1} = 223.54$ HZ, $t(129) = 4.00$, $p < .001$ and $M_{PQ2} = 227.49$ HZ vs. $M_{ISQ2} = 288.44$ HZ, $t(129) = 9.43$, $p < .001$). However, this was not true for the last word in the utterance. In fact, the relationship was reversed: the average pitch of the last word in pedagogical questions ($M_{PQL} = 204.17$ HZ) was higher than that of information-seeking questions ($M_{ISQL} = 152.9$ HZ), $t(129) = 7.27$, $p < .001$. This suggests that information-seeking questions follow a canonical pitch contour characteristic of wh- questions that tend to fall over time, whereas pedagogical questions follow a non-canonical contour. Indeed, the average pitch of the first word in pedagogical questions ($M_{PQ1} = 203.09$ HZ) was not significantly different from the average pitch of the last word in pedagogical questions ($M_{PQL} = 204.17$ HZ), whereas the average pitch of the first word in information-seeking questions ($M_{ISQ1} = 223.54$ HZ) was significantly higher than the last word in information-seeking questions ($M_{ISQL} = 152.9$ HZ) ($p < .001$). This finding is consistent with the finding at the utterance level that the F0 range is lower in PQs than in ISQs. Figure 1 represents the pitch contours of a single exemplar of a pedagogical and an information-seeking question that captures the reported differences.

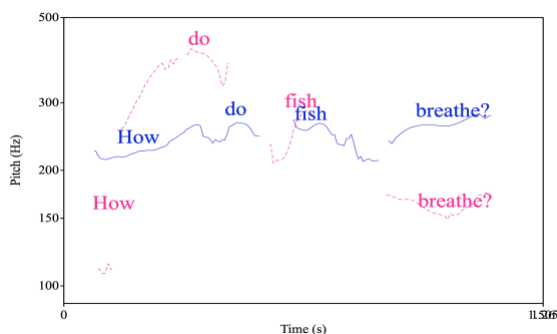


Figure 1. Pitch contour of an exemplar of a pedagogical (solid blue line) and an information-seeking (dashed magenta line) question from a single speaker (logarithmic scale).

In summary, the average pitch of the first two words of pedagogical wh- questions is lower than the pitch of information-seeking questions. In addition, the first two words of pedagogical questions are with longer duration than the first two words of information-seeking questions. Whereas the pitch contour of information-seeking questions is canonical and tends to fall from high to low pitch, the pitch contour of pedagogical questions is non-canonical. As a result, the pitch of the first and the last word in pedagogical questions is not significantly different, whereas the pitch of the last word is significantly lower than the pitch of the first word in information-seeking questions.

Discussion

The present studies demonstrate that English speaking naïve listeners can use the prosody of questions in their native, as

well as a foreign, language in order to accurately detect the pedagogical or the information-seeking intent of a question. Naïve listeners draw accurate inferences about the intent of the question, both when the question is embedded in adult-directed speech and when it is embedded in child-directed speech. Moreover, naïve listeners make these judgments even when they have no syntactic or semantic cues whatsoever.

Compared to information-seeking questions, some of the prosodic features that characterize pedagogical questions are longer duration of the question word and the verb that follows it, lower pitch of those two words, and non-canonical pitch contour also captured by the lower F0 range in PQs at the utterance level. The lower pitch and the longer duration of the first two words in pedagogical questions might play a functional role, which is to mark the question as a non-canonical inquiry, and to emphasize the focus on the wh-question word. In a similar vein, the non-canonical shape of the contour of pedagogical questions, where the pitch at the end of the utterance either raises or stays flat (as opposed to falling down as in canonical wh- questions), is important in light of arguments that non-canonical question contours mark partial knowledge of the answer being asked (Hedberg & Mamani, 2010; Hedberg & Sosa, 2011). This claim is consistent with the finding in the present study showing that pedagogical questions have a non-canonical contour. Namely, pedagogical questions are asked by knowledgeable teachers and it is therefore expected that they will have non-canonical contours that signal the asker's knowledge state. Future research should further investigate this possibility.

Unlike the lower pitch of the first two words and the non-canonical shape of the contour of pedagogical questions, the longer duration of pedagogical questions is a feature that is shared with child-directed speech (Fernald & Kuhl, 1987; Fernald & Simon, 1984). This may explain why the naïve listeners in the present study were less likely to accurately rate child-directed information-seeking questions as information-seeking questions. It is still possible, however, that participants have preconceptions that pedagogical questions are more likely to be child-directed and that information-seeking questions are more likely to be adult-directed. Future research should explore this question.

Another important step for future research is to investigate how the prosody of questions produced in naturalistic settings marks pedagogical intent, and how the learners' interpretation of those prosodic cues shapes their learning. We are actively investigating both issues.

In conclusion, although questions are ubiquitous in everyday experience and they are important for learning, very little is known about how learners identify questions that are intended for learning. The present study has demonstrated for the first time that the prosody of language can encode pedagogical intent and naïve listeners can decode it. This important finding opens up many new important questions about the role of prosody in learning. Might pedagogical questions hold the key?

Acknowledgments

This study was supported by an NSF DCR-EHR DRL grant (#1660885) to Bonawitz and Shafto.

References

- Banuazizi, A., & Cresswell, C. (1999). *Is that a real question? Final rises, final falls, and discourse function in yes-no question intonation*. Paper presented at the 35th Annual Meeting of the Chicago Linguistics Society.
- Bartels, C. (1999). *The intonation of English statements and questions*. A compositional interpretation. New York & London, Garland Publishing.
- Bonawitz, E. B. & Shafto, P. (2016). Computational models of development, Social influences. *Current Opinion in Behavioral Sciences*, 7, 95-100
- Bornstein, M. H., Tal, J., Rahn, C., Galperin, C. Z., Pêcheux, M.-G., Lamour, M., . . . Tamis-LeMonda, C. S. (1992). Functional analysis of the contents of maternal speech to infants of 5 and 13 months in four cultures: Argentina, France, Japan, and the United States. *Developmental Psychology*, 28(4), 593. doi:10.1037/0012-1649.28.4.593
- Braun, B., Dehe, N., Neitsch, J., Wochner, D., & Zahner, K. (2019). The prosody of rhetorical and information-seeking questions in German. *Language and Speech*, 62, 779-807.
- Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13, 148-153.
- Eaves, B.S., Feldman, N., Griffiths, T. & Shafto, P. (2016). Infant-directed speech is consistent with teaching. *Psychological Review*, 123(6), 758-771.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior & Development*, 8(2), 181-195.
- Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279-293.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20(1), 104-113.
- Gergely, G., Egyed, K., & Kiraly, I. (2007). On pedagogy. *Developmental Science*. 10(1), 139-146.
- Han, C. H. (2002). Interpreting interrogatives as rhetorical questions. *Lingua*, 112, 201-229.
- Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Harvard University Press.
- Hedberg, N., & Mameni, M. (2010). The semantic function of rising wh-questions. *Proceedings of the MOSAIC-2 (Meeting of Semanticists Active in Canada)*.
- Hedberg, N., & Sosa, J. M. (2011). A unified account of the meaning of English questions with non-canonical intonation. *Proceedings of the International Seminar on Prosodic Inferences*.
- Jean, A., Daubert, E., Yu, Y., Shafto, P., & Bonawitz, E. B. (2019). Pedagogical questions empower exploration. *Proc. 41st Ann. Conf. Cog. Sci. Soc.*
- Leven, T., & Long, R. (1981). *Effective instruction*. Washington DC: Association for Supervision and Curriculum Development.
- Liu, H.-M., Tsao, F.-M., Kuhl, P. K. (2009). Age-related Changes in Acoustic Modifications of Mandarin Maternal Speech to Preverbal Infants and Five-Year-Old Children: A Longitudinal Study. *J. of Child Lang.*, 36(4), 909-922.
- Nencheva, M. L., Piazza, E. A., & Lew-Williams, C. (2020). The moment-to-moment pitch dynamics of child-directed speech shape toddlers' attention and learning. *Developmental Science*.
- Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition*. Cambridge: Cambridge University Press.
- Olsen-Fulero, L., & Conforti, J. (1983). Child responsiveness to mother questions of varying type and presentation. *Journal of Child Language*, 10(3), 495-520.
- Rowe, M. L. (2008). Child-Directed Speech: Relation to Socioeconomic Status, Knowledge of Child Development and Child Vocabulary Skill. *J. Child Lang.*, 35, 185-205.
- Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., Laznik, M.C., & Cohen, D. (2013). Motherese in interaction: At the cross-road of emotion and cognition? (A systematic review). *PLOS ONE*, 8(10), e78103.
- Senju A, & Csibra G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18(9), 668-671.
- Siraj-Blatchford, I., & Manni, L. (2008). 'Would you like to tidy up now?' An analysis of adult questioning in the English Foundation Stage. *Early Years*, 28(1), 5-22
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7, 53-71.
- Tomasello, M. (2008). *Origins of human communication*. MIT Press.
- Trott, S., Ferreira, V., & Bergen, B. (2019). Prosodics cues signal the intent of potential indirect requests. *Proceedings of the 41st Annual Conference of the Cognitive Science Society*.
- Yu, Y., Bonawitz, E. B., Shafto, P. (2017). Pedagogical questions in mother-child conversations. *Child Development*, 90, -147-160.
- Yu, Y., Laundrum, A. R., Bonawitz, E. B., Shafto, P. (2018). Questioning supports effective transmission of knowledge and increased exploratory learning in pre-kindergarten children. *Developmental Science*, 21.
- Yu, Y., Shafto, P., & Bonawitz, E. (2020). Inconvenient samples: Modeling the effects of non-consent by coupling observational and experimental results. *Open Mind*, 4, 13-24.