When is a Question a Question for Children and Adults?

Mathieu R. Saindon, Sandra E. Trehub, E. Glenn Schellenberg & Pascal H. H. M. van Lieshout

Routledge
Taylor & Francis Group

When is a Question a Question for Children and Adults?

Mathieu R. Saindon[a], Sandra E. Trehub[a], E. Glenn Schellenberg[a],
and Pascal H. H. M. van Lieshout[b]

[a]Department of Psychology, University of Toronto Mississauga; [b]Department of Speech-Language Pathology,
University of Toronto

ABSTRACT

Terminal changes in fundamental frequency provide the most salient acoustic cues to declarative questions, but adults sometimes identify such questions from pre-terminal cues. In the present study, adults and 7- to 10-year-old children judged a single speaker's adult- and child-directed utterances as questions or statements in a gating task with word-length increments. Listeners of all ages successfully used pre-terminal cues to identify utterance type, often only the initial word, and they were more accurate for child-directed than adult-directed utterances. There were age-related differences in identification accuracy and number of words required for correct identification. Age differences were already apparent on the initial (first five) utterances, confirming adults' superior explicit knowledge of intonation patterns that signify questions and statements. Adults' performance improved over the course of the test session, reflecting taker-specific learning, but children exhibited no such learning.

The most salient acoustic differences between declarative questions and statements are the terminal fundamental frequency (F_0) contours, which rise in yes/no questions and fall in statements (Bolinger, 1986; Ladd, 2008). Artificially increasing the pitch excursion of utterance-final syllables increases adults' likelihood of identifying those utterances as questions (Gårding & Abramson, 1965). Even sine-wave patterns with rising pitch contours are identified as questions and those with falling pitch contours as statements (Studdert-Kennedy & Harding, 1973).

Production studies confirm the distinctiveness of terminal pitch contours, but they also reveal differences in pre-terminal positions. In English, for example, successive content words often exhibit rising pitch in questions but not statements, and stressed syllables are typically followed by a pitch decrease in statements but not questions (O'Shaughnessy, 1979). In the present study, we sought to document age-related changes in the ability to differentiate declarative questions from statements on the basis of pre-terminal cues, and to ascertain whether speech register—child- or adult-directed—affects such differentiation.

Gating tasks, which were developed to assess the speed of word identification (Grosjean, 1980), provide a means of determining the minimum information necessary for identifying statements and declarative questions. Listeners judge whether utterances differing only in prosodic cues are questions or statements from incremental increases in syllables or words. Findings from these tasks reveal a different time course of identification across languages. For example, Dutch listeners identify declarative questions after hearing the second accented syllable in an utterance (Van Heuven & Haan, 2000). Portuguese listeners identify statements after the first stressed vowel, but they need the final stressed vowel to identify questions (Falé & Faria, 2006). Spanish listeners identify questions

---

after hearing the first content word (Face, 2005), but French listeners do so after hearing the falling F_0 contour that precedes the final rise (Vion & Colas, 2006). Surprisingly, there are no comparable studies with English-speaking children or adults.

Only one study used a gating task to examine children's identification of statements and declarative questions (Gérard & Clément, 1998). French 5-, 7-, and 9-year-olds judged utterance type from increasing numbers of words (i.e., one word, two words, three, etc.). Even after hearing complete utterances, 5- and 7-year-olds failed to differentiate questions from statements, but 9-year-olds were able to do so. Although children's performance was well below adult levels, adult gating studies suggest that there are fewer pre-terminal cues to question identity in French (Vion & Colas, 2006) than in Dutch (Van Heuven & Haan, 2000) or Spanish (Face, 2005). Moreover, English-speaking children as young as 5 (Saindon, Trehub, Schellenberg, & Van Lieshout, 2016) and Spanish- and Mandarin-speaking 4-year-olds (Armstrong, 2012; Zhou, Crain, & Zan, 2012) can differentiate questions from statements on the basis of prosody alone, although they are much less accurate than adults. Note, however, that Mandarin questions are cued by duration and intensity contrasts rather than pitch contrasts (Zhou et al., 2012).

There are also notable cross-language differences in young children's perception and production of the relevant acoustic distinctions. For example, 21-month-old Spanish-speaking toddlers mark their declarative questions with rising intonation contours (Armstrong, 2012), but English-speaking preschoolers are unlikely to do so (Loeb & Allen, 1993; Patel & Grigos, 2006; Snow, 1994, 1998). Children's difficulties, when evident, do not stem from perceptual limitations. For example, infants differentiate rising from falling pitch contours (Frota, Butler, & Vigario, 2014; Karzon & Nicholas, 1989; Nazzi, Floccia, & Bertoncini, 1998; Soderstrom, Ko, & Nevzorova, 2011), and young children can learn to identify small directional differences (up, down) in pitch (Stalinski, Schellenberg, & Trehub, 2008). Instead, their difficulties seem to arise from interpretive issues with sentence-level prosody, which are resolved progressively with increasing age and cognitive maturity (Cutler & Swinney, 1987; Tyler & Marslen-Wilson, 1978; Wells, Peppé, & Goulandris, 2004). It is also notable that young children perform much better in conversational contexts that provide contextual support (Doherty-Sneddon & Kent, 1996) than in laboratory contexts in which isolated utterances are presented or elicited (e.g., Patel & Grigos, 2006).

In line with our goal of ascertaining the minimum information required to differentiate question from statement intentions, we asked English-speaking adults (Experiment 1) and 7- to 10-year-old children (Experiment 2) to identify stimuli as questions or statements from the initial word of five-word utterances and from incremental additions of words. In contrast to previous gating studies, which used conventional or adult-directed speech, we used child- *and* adult-directed speech. Child-directed speech (or "clear" adult-directed speech) features wider F_0 range, slower speaking rate, increased pitch and intensity accents, and longer pauses than conventional adult-directed speech (Foulkes, Docherty, & Watt, 2005; Jacobson, Boersma, Fields, & Olson, 1983; Smiljanić & Bradlow, 2009). These features enhance the intelligibility of messages for children (Fernald & Mazzie, 1991; Liu, Tsao, & Kuhl, 2009; Syrett & Kawahara, 2014), older adults (Caporael, 1981; Masataka, 2002), and non-native speakers (Uther, Knoll, & Burnham, 2007). Because suprasegmental aspects of child-directed speech typically exaggerate cues to the speaker's intentions, this speaking style should highlight the distinctions between statements and questions. Accordingly, we expected listeners of all ages to differentiate questions from statements earlier (i.e., from fewer words) with child-directed than with adult-directed utterances. We also predicted age-related changes, in line with French-speaking children's performance on a gating task (Gérard & Clément, 1998) and English-speaking children's performance on full utterances presented in isolation (Saindon et al., 2016). The performance of adults was examined in Experiment 1 and that of children in Experiment 2.

Our use of a single talker provided an opportunity to examine attunement to her intonation patterns over the course of the test session. Adults often exhibit adaptation to individual phonetic signatures (Theodore, Myers, & Lomibao, 2015) and conversational style (Pogue, Kurumada, & Tanenhaus, 2016). Children's ability to perceive accented or atypical speech implies similar

attunement, but younger children experience greater difficulty in this regard (Creel, Rojo, & Paullada, 2016). In the present study, adaptation to the speaker's questioning and stating style would be reflected in progressive improvement in the use of pre-terminal cues.

Experiment 1

Method

Participants

Participants were 45 college students (36 women, 9 men; mean age = 18.82 years, SD = 2.85), who received partial course credit for their participation. Inclusion criteria were Canadian birth or arrival in Canada before 8 years of age and absence of hearing difficulty, according to self-report. Three additional adults were tested but excluded from the final data set because of inattentiveness, as reflected in their failure to identify questions and statements above chance levels (14 or more incorrect out of 40) after hearing the complete sentence.

Stimuli and apparatus

A woman with experience in vocal performance and early-childhood education produced 4 versions of 10 5-word utterances (see Table 1). The sentence-length utterances were selected on the basis of their prosodic form rather than their syntax or content. Specifically, each sentence began with a stressed syllable, and words in the same position across sentences had the same number of syllables (first word: two syllables; second word: one syllable; third word: two syllables; fourth word: one syllable; fifth word: two syllables). Sentences were produced as declarative questions or statements in an adult-directed or child-directed style. On average, child-directed versions had higher F_0, greater F_0 range and variability, and longer duration than adult-directed versions (see Table 2). Each sentence was amplitude-normalized at 75 dB SPL using Sound Forge Pro (Version 10.0; Sony, Tokyo, Japan).

The stimuli were recorded in a double-walled, sound-attenuating chamber

(Industrial Acoustics Corporation Co., Bronx, NY) with a microphone (Sony T) connected directly to a Windows 7 workstation. Testing was conducted in the same sound-attenuating chamber. A computer workstation and amplifier (Harmon-Kardon 3380, Stamford, CT) outside the chamber interfaced with a 17-in touch-screen monitor (Elo LCD TouchSystems, Berwyn, PA) and two wall-mounted loudspeakers (Electro-Medical Instrument Co., Mississauga, ON) inside the chamber. The touchscreen monitor facing the participant was used to present instructions and record responses. The loudspeakers were mounted at the corners of the sound chamber, each located

Table 1. Stimulus sentences.

Judy knows Michael is going.
Lisa seems happy for Alex.
Maybe one apple is enough.
Someone went shopping for balloons.
People played music from Japan.
Puddles make rainy days better.
Donuts with sprinkles are tasty.
Brian is going home today.
Brendan is leaving town again.
Roses are growing on bushes.

Table 2. Mean duration and fundamental frequency of adult- and child-directed utterances.

Speech register	Duration (ms)	F_0 mean (Hz)	F_0 min (Hz)	F_0 max (Hz)	F_0 SD (Hz)
Adult-directed	2798.00	258.01	138.74	522.60	82.90
Child-directed	3074.48	276.61	164.08	588.00	103.83

at a distance of .76 m and 45° azimuth from the participant. A customized program created with Affect4 software (Spruyt, Clarysse, Vansteenwegen, Baeyens, & Hermans, 2010) presented the auditory stimuli and recorded responses. All stimuli were played at a comfortable listening level of approximately 65 dB SPL.

Procedure

Listeners heard all 40 stimulus sentences presented in random order, constrained so that different versions of the same utterance (i.e., question or statement, adult- or child-directed) could not appear consecutively. (Due to computer malfunction, one listener heard 39 sentences.) Each sentence was presented in a block of five trials, beginning with the one-word stimulus, with successive trials adding a word, ending with the complete five-word sentence on the fifth trial. After each trial, listeners were required to select one of four options: *question, maybe a question, maybe a statement*, or *statement*, which combined choice of utterance type with confidence level. The task began with a short familiarization phase featuring one trial block with an adult-directed declarative question and another with an adult-directed statement, neither of which appeared in the test phase. The experimenter provided feedback at the end of each trial block (i.e., complete sentence) only during the familiarization phase.

Results and discussion

Preliminary analyses revealed no meaningful differences as a function of whether responses indicated more confidence (i.e., *question*) or less confidence (i.e., *maybe a question*). Accordingly, each response was considered dichotomously as *statement* or *question*, which allowed us to transform these responses into d' scores. For each participant, d' scores were computed separately for each number of words and for adult- and child-directed sentences. Thus, each participant had 10 d' scores. A hit was a response in the question category (*question* or *maybe a question*) when the stimulus was a question. A false-alarm was a response in the question category when the stimulus was a statement. In terms of signal-detection theory, the "signal" consisted of the acoustic cues that distinguish questions from statements, and d' scores reflected sensitivity to these cues.

The stimulus set consisted of 20 adult-directed and 20 child-directed utterances, half questions and half statements. Thus, the maximum number of successes for adult- and child-directed questions was 10 for each number of words. Because d' scores cannot be computed when the hit rate equals 1.0 or the false-alarm rate equals 0 (i.e., infinite scores), proportions of correct responses were adjusted by using the following formula: (hit OR false-alarm rate + .5) /(maximum number of successes + 1), as in previous developmental studies (e.g., Thorpe, Trehub, Morrongiello, & Bull, 1988). These proportions were then converted to z-scores and subsequently to d' scores ($d' = z$[hit rate]—z[false-alarm rate]). The maximum d' score, indicating perfect performance (i.e., 10 hits, 0 false alarms), was 3.38, and chance responding (i.e., equal number of hits and false alarms) was 0. Descriptive statistics are illustrated in Figure 1 as a function of age group, speaking style, and number of words.

Initial analyses compared performance with chance levels using one-sample t-tests. Because there were 10 tests (two speaking styles X five different numbers of words), the alpha-level was lowered to .005. Nevertheless, performance exceeded chance levels in all cases, ps < .001, indicating that cues to statements or questions were discernible even from the first word, regardless of speaking style.

When the entire five-word utterance was presented, performance was at ceiling, as expected, with mean d' scores of 3.20 (SD = 0.36) and 3.12 (SD = 0.38) for adult- and child-directed speech, respectively, including perfect performance by 34 of 45 listeners for adult-directed speech and 29 of 45 for child-directed speech. Moreover, performance on five-word utterances vastly exceeded performance on one-, two-, three-, and four-word stimuli for both speaking styles, ps < .001, yet there was no difference between speaking styles for five-word stimuli, p < .2. Thus, further analyses were limited to stimuli with one to four words.
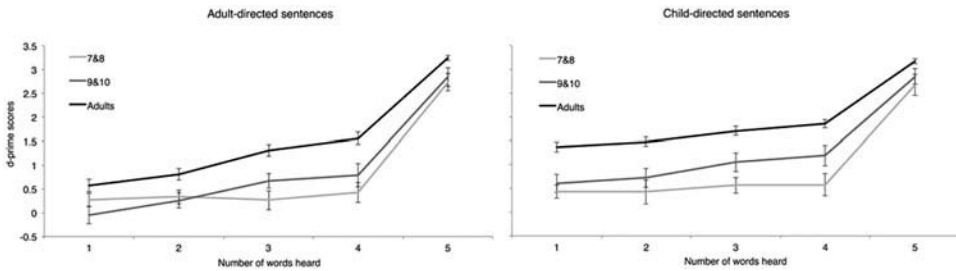
Figure 1. Performance in Experiment 1 (adults) and Experiment 2 (children) as a function of speech register and utterance increments.

Response patterns were analyzed with an analysis of variance (ANOVA), with speaking style (adult- or child-directed speech) and number of words (1–4) as repeated measures and d' scores as the dependent variable. A main effect of speaking style revealed that participants more readily differentiated questions from statements in child-directed than in adult-directed speech, $F(1, 44) = 24.43$, $p < .001$, $\eta_p^2 = .357$. A main effect of number of words stemmed from the fact that performance improved monotonically as the number of words increased, $F(3, 132) = 58.41$, $p < .001$, $\eta_p^2 = .570$. Both main effects were qualified, however, by a significant two-way interaction, $F(3, 132) = 9.29$, $p < .001$, $\eta_p^2 = .174$. Separate analyses of the two speaking styles revealed a linear improvement in performance for adult-directed speech, $F(1, 44) = 97.90$, $p < .001$, $\eta_p^2 = .690$, and for child-directed speech, $F(1, 44) = 27.59$, $p < .001$, $\eta_p^2 = .385$, with the interaction highlighting more dramatic improvement for adult-directed speech. As shown in Figure 1, performance on child-directed speech was relatively good even with one-word stimuli, which left less room for improvement as the number of words increased.

The present results confirmed the perceptibility of acoustic differences between statements and declarative questions that featured identical words. When the pitch contour of the penultimate and final words, whether rising or falling, was present, the identification of declarative questions and statements was at ceiling. Nevertheless, cues earlier in the utterances permitted successful identification, especially when the speaker used a child-directed style. Experiment 2 asked whether children are comparably successful at discriminating statements from questions when terminal pitch cues are present, whether they capitalize on subtle acoustic cues in pre-terminal position, and whether they benefit comparably from the child-directed speaking style.

Experiment 2

Although English-speaking children as young as 5 can differentiate questions from statements on the basis of intonation alone, 5- and 6-year-olds have much greater difficulty compared to older children, even after considerable training (Saindon et al., 2016). Accordingly, the present study was limited to children between 7 and 10 years of age.

Method

Participants

The participants, 32 children from the local community, included 15 7- and 8-year-olds (designated "younger children": 7 girls, 8 boys; mean age = 7 years, 9 months, range = 7;2–8;10), and 17 9- and 10-year-olds (designated "older children": 9 girls, 8 boys; mean age = 10 years, 1 month, range = 9;1–10;11). Inclusion criteria were Canadian birth and no personal or family history of hearing loss, according to parental report. An additional two 7-year-olds were tested but excluded from the final sample because they failed to identify questions and statements above chance levels (14 or more incorrect out of 40) after hearing the complete sentence, indicating poor understanding or attention. Children received a small gift for their participation.

Apparatus and stimuli

The apparatus and stimuli were the same as in Experiment 1.

Procedure

The adults' task from Experiment 1 was modified to make it more engaging for children. On the basis of previous research indicating young children's confusion with the terms *statement* and *question* (Saindon et al., 2016), the response options from Experiment 1 were simplified by changing them to *asking, maybe asking, maybe telling*, and *telling*. In addition, non-contingent feedback was provided during the test phase by presenting the image of one of several cartoon characters after completion of each trial block.

Results and discussion

Responses were converted to *d'* scores, as in Experiment 1, with *telling* and *maybe telling* considered as *statement* responses, and *asking* and *maybe asking* considered as *question* responses. Figure 1 illustrates descriptive statistics as a function of age group, speaking style, and number of words.

Initial analyses examined whether performance exceeded chance levels separately for both age groups, both speaking styles, and different numbers of words, with correction for multiple tests, as in Experiment 1. For the older children, performance exceeded chance levels for three- and five-word stimuli in adult-directed style, $ps < .005$, and was at the cusp of significance for four-word stimuli, $p = .005$. For all stimuli in child-directed style, performance exceeded chance levels, $ps < .005$. For the younger children and adult-directed utterances, performance exceeded chance only for the five-word stimuli, $p < .001$. For child-directed utterances, performance exceeded chance for three- and five-word stimuli, $ps < .005$.

For both speaking styles and age groups, performance for five-word stimuli was much better than performance for utterances with fewer words, $ps < .001$, as it was for adults in Experiment 1 (see Figure 1). Moreover, for both groups of children, performance with five-word stimuli did not vary as a function of speaking style, $ps > .6$. As in Experiment 1, the main analysis focused on stimuli with one to four words.

A mixed-design ANOVA was used to analyze *d'* scores as a function of two repeated measures (speaking style, number of words) and one between-subjects factor (age group). A main effect of speaking style confirmed that children performed better with child-directed than with adult-directed speech, $F(1, 30) = 7.10$, $p = .012$, $\eta_p^2 = .191$. There was also a main effect of number of words, $F(3, 90) = 10.65$, $p < .001$, $\eta_p^2 = .262$, which was qualified by a two-way interaction between age group and number of words, $F(3, 90) = 5.19$, $p = .002$, $\eta_p^2 = .147$. There were no other main effects or interactions, $ps > .2$. Separate analyses of the older and younger children revealed that although both age groups exhibited a monotonic increase in mean *d'* scores as the stimuli increased from one to four words, the linear trend was significant for the older children, $F(1, 16) = 28.65$, $p < .001$, $\eta_p^2 = .642$, but not for the younger children, $p > .3$.

Subsequent analyses considered the present samples of children jointly with the adults tested in Experiment 1. We first examined how many individual participants could identify questions at above-chance levels. We conducted a series of 3 (age group) by 2 (above or at chance) chi-square tests of independence (one for each number of words) separately for adult- and child-directed utterances. According to the normal approximation to the binomial test (one-tailed, correcting for continuity), participants were significantly above chance if they had at least 15 out of 20 correct responses. The results are summarized in Table 3. For adult-directed utterances, age-related differences in the proportion of participants who successfully differentiated questions from statements were evident at three and four words. For child-directed utterances, age-related differences were evident at one, two, three, and four words. In each instance, adults had the greatest proportion of participants performing at above chance levels, followed by older children, and then younger children.

Table 3. Proportions of participants identifying questions above chance after 1–5 words.

	Adult-directed					Child-directed				
Words	7–8	9–10	Adults	χ^2	p	7–8	9–10	Adults	χ^2	P
1	.000	.059	.178	4.157	.125	.133	.353	.511	6.899	.032*
2	.067	.059	.244	4.415	.110	.200	.235	.667	15.071	.001*
3	.067	.235	.489	10.083	.006*	.200	.412	.778	18.141	.001*
4	.200	.471	.600	7.247	.027*	.267	.529	.867	20.590	.001*
5	1.000	1.000	1.000	n/a	n/a	.933	1.000	1.000	4.188	.123

Note. Each chi-square test of independence had $df = 2$ and $n = 77$.

We also found a bias for responding *statement* rather than *question* for adult-directed utterances by way of one-sample *t*-tests, which examined whether the mean number of *statement* responses differed from the number expected in the absence of bias (50 out of 100), separately for each speaking style and age group (see Table 4). The results revealed a significant bias for *statement* responses for adult-directed (*p*s < .005) but not child-directed utterances (*p*s > .06) across age groups. Perhaps listeners expect more pitch variability in questions than in statements, choosing the latter category in the absence of obvious cues, as suggested by Vion and Colas (2006). This would also explain the absence of a comparable bias for child-directed utterances, which had more pitch variability and were therefore judged as questions more often than adult-directed utterances. Moreover, because statements occur more commonly than questions in everyday discourse, a *statement* response is a reasonable default option in the context of ambiguous cues.

Because the stimuli were limited to a single speaker, listeners had an opportunity to profit from speaker-specific cues over the course of the test session. We tested this hypothesis with a multilevel model that had position of correct judgment (1–5 words) as the dependent variable,[1] group (younger children, older children, adults) and speaking style (adult- or child-directed) as fixed factors, and utterance number (1–40, centered) as a fixed, continuous variable. A random intercept was included for each participant. Descriptive statistics are illustrated in Figure 2. There was a main effect of speaking style, $F(1, 2993.045) = 16.01$, $p < .001$, because performance on child-directed utterances exceeded that on adult-directed utterances, as noted above. Main effects of age group, $F(2, 74.006) = 22.92$, $p < .001$, and utterance number, $F(1, 2993.255) = 10.69$, $p = .001$, were qualified by a two-way interaction between age group and utterance number, $F(2, 2993.237) = 4.61$, $p = .010$ (see Figure 2). No other interactions were evident, $Fs < 1$. Separate analyses of the three age groups revealed that performance improved with increasing exposure to the speaker (i.e., fewer words required for correct identification) for adults, $p < .001$, but not for older, $p > .2$, or younger, $F < 1$, children (see Figure 2). Adults also outperformed children on the initial utterances (first 5 of 40), $t(75) = 2.07$, $p = .042$.

Descriptive analyses provided insight into the pre-terminal cues that enabled listeners to distinguish questions from statements. Qualitative annotation conventions from Tonal Breaks and Indices (ToBI) transcriptions (Pierrehumbert, 1980) depict pitch events associated with intonational boundaries and accented syllables. The transcriptions include two tone levels—high (H) and low (L). The *

Table 4. Mean "statement" responses and age.

	Adult-directed			Child-directed		
Age group	Mean	t	p	Mean	t	P
7–8	58.6	3.503	.004*	54.7	1.889	.080
9–10	58.6	3.476	.003*	49.2	−.294	.773
Adults	62.2	8.022	.001*	48.1	−1.890	.065

Note. * indicates values that significantly exceeded chance levels (50).

---

[1]Position of correct judgment represented the smallest number of words required for a correct and consistent judgment.
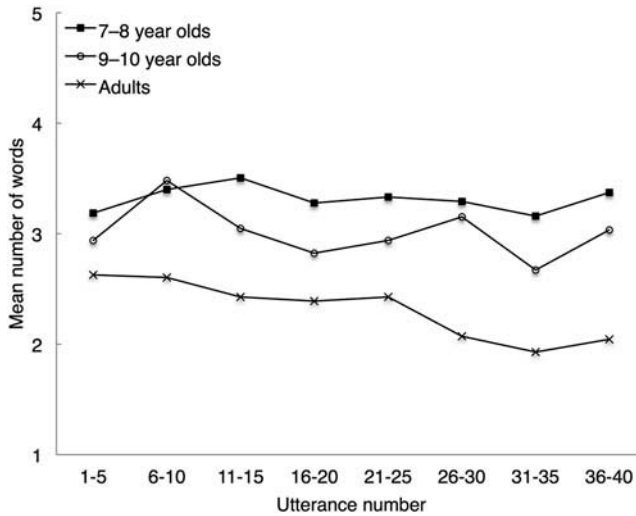
Figure 2. Mean number of words required to identify questions and statements as a function of age and utterance number.

symbol denotes the location of an accented syllable, and the % symbol denotes the end of an intonational phrase. The + symbol is used when a nuclear accent is marked by a bitonal pitch accent, or an accent comprising a fall-rise (L + H) or rise-fall (H + L) contour.

Sample question and statement contours are depicted in Figure 3. All questions in the set began with a low pitch accent followed by a rise (L*+H). By contrast, most statements began with a pitch
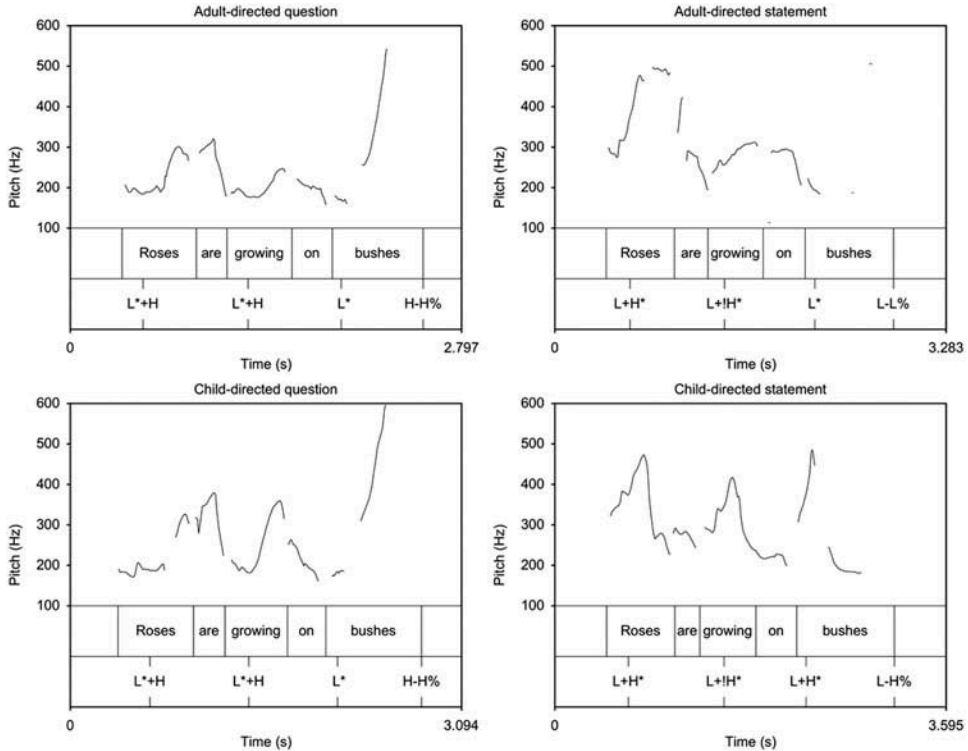

Figure 3. F0 contours and ToBI annotations from sample adult- and child-directed utterances.

rise followed by high pitch accent (L + H*), with the remaining statements having an initial high pitch accent without the preceding rise (H*). Statements tended to end with a low phrase accent and low boundary tone (L-L%), and questions tended to end with a high phrase accent and high boundary tone (H-H%). Some child-directed statements, however, ended with a low phrase accent and high boundary tone (L-H%). The boundary tone of questions was often preceded by a low pitch accent (L*), whereas the boundary tone of statements was often preceded by a high pitch accent (H*). Moreover, down-stepped tones—a sequence of high pitch accents that are not as high as the preceding high ("H") accent—were more common in statements than in questions, reflecting the tendency for pitch to fall gradually over the course of the utterance. Finally, differences between adult- and child-directed utterances were inconsistent, with the patterns being sentence-dependent, presumably because of the location of nuclear tones.

General discussion

The present study is the first to use a gating task to examine the identification of adult- and child-directed questions and statements by English-speaking children and adults. We established that children as young as 7 could use pre-terminal cues to identify questions and statements. Adults differentiated questions from statements after hearing the first word regardless of speech register. Older children (9- to 10-year-olds) performed better than younger children (7- to 8-year-olds), with both groups identifying child-directed utterances more readily than adult-directed utterances. Register also had more dramatic consequences for children than for adults. For example, older children identified child-directed utterances after hearing the first word, but they identified adult-directed utterances only after the first three words. Younger children correctly identified utterance type after hearing the first three words of child-directed utterances, but they required the complete adult-directed utterances to differentiate questions from statements. Finally, repeated exposure to the speaker's voice enhanced performance for adults but not for children.

Performance differences between children and adults were evident early in the test session (the initial 5 of 40 utterances), which indicates that adults' explicit knowledge of the intonation patterns that signify questions and statements exceeds that of children. In previous research that required listeners to judge full sentences, 7- and 8-year-olds identified questions and statements more poorly than 9- and 10-year-olds, who performed at adult levels (Saindon et al., 2016). The findings of both studies are consistent with the notion that pitch contour differences, although discriminable, are less memorable for children than for adults (Creel, 2014).

It is likely that adaptation to the demands of the gating task occurred relatively quickly. By contrast, adaptation to speaker-specific cues would be expected to unfold more slowly after exposure to a range of utterances from the single speaker. In fact, improvement in adults' use of pre-terminal cues became evident at utterances 26–30, with no further improvement thereafter (see Figure 2). Children, by contrast, derived no benefit from such exposure within the time frame of the test session. Talker familiarity advantages are likely to be achieved more readily for word recognition than for intonation recognition. At times, however, familiarity effects in school-age children who receive protracted exposure to talkers are restricted to highly familiar words (Levi, 2015).

For questions in the present study, the pre-nuclear pitch accent was low and followed rapidly by a rising contour (L*+H). For statements, the contour tended to rise during the pre-nuclear pitch accent (L + H*). Although little is known about pre-terminal cues to questions and statements, there is evidence that intonation contours tend to fall during stressed syllables in statements but not in questions (O'Shaugnessy, 1979). This was not the case for the questions and statements in the present study, both of which featured a rising pitch contour during the first word. The main difference between our questions and statements involved the timing of the rise, which occurred during the primary stress for statements but after the primary stress for questions.

Our findings are consistent with adults' ability to differentiate Spanish questions from statements after hearing the first content word, which features higher F_0 peaks in questions than in statements

(Face, 2005). There are notable differences, however, in cues to utterance type across languages and speakers. For example, adults correctly identify Dutch questions and statements after hearing the second F_0 accent, which is larger in questions than in statements for female speakers but smaller for male speakers (Van Heuven & Haan, 2000). Question identification requires information in the penultimate syllable for French questions (Vion & Colas, 2006) and in the final syllable for Portuguese questions (Falé & Faria, 2006).

In principle, intensity and duration could influence the perception of utterance type. Although duration is not a reliable cue to question identity (Peng, Chatterjee, & Lu, 2012), longer utterances and those with higher terminal intensity are more likely to be identified as questions than as statements (Ma, Whitehill, & So, 2010; Patel, 2003; Peng et al., 2012). Intensity is highly correlated with F_0 (Bolinger, 1986), but listeners are thought to use intensity and duration cues for question/statement identification only when F_0 cues are unavailable (Ma et al., 2010; Patel, 2003; Peng & Chatterjee, 2009).

Our study is also the first to examine the impact of child-directed speech on listeners' differentiation of questions and statements. Child-directed speech, which is characterized by slower speaking rate, more careful pronunciation, and wider pitch range relative to adult-directed speech (Foulkes et al., 2005; Jacobson et al., 1983), emphasizes the speaker's affective intentions and focus of interest. The present findings indicate that it also emphasizes the speaker's questioning or declarative intent, as one might expect. Acoustic analyses confirmed that the differences between questions and statements in adult- and child-directed speech were similar except for the greater magnitude of differences in child-directed speech. As can be seen in Figure 3, pitch accents were more prominent in child-directed than in adult-directed utterances.

The finding of age-related differences in overall question and statement identification is consistent with age-related differences in French-speaking children's identification of utterance type (Gérard & Clément, 1998). In contrast to French 7-year-olds' inability to distinguish declarative questions from statements and 9-year-olds' ability to do so only after hearing the entire utterance, English-speaking children in the present study correctly identified statements and questions early in the utterance (except for 7- and 8-year-olds' judgments of adult-directed utterances). As noted, French includes fewer pre-terminal cues to questions and statements (Vion & Colas, 2006), whereas questions and statements in our English stimuli included distinct pre-terminal differences (L*+H tone in questions and L + H* tone in statements). It is unlikely, however, that the cross-language differences are attributable solely to differential cue distinctiveness or issues of incidence. For example, although the French study also featured a single talker, it was considerably more challenging because of five response alternatives—*asking, telling, happy, sad*, and *ironic*—rather the two general response classes—*asking* and *telling* (each qualified by *maybe*)—in the present study.

It is important to note the limitations of the gating paradigm for assessing listeners' identification of the speaker's pragmatic intentions based on intonation alone. In principle, the units of interest are intonation phrases rather than words, and studies with adults have used intonation rather than word gates (e.g., Face, 2005; Van Heuven & Haan, 2000; Vion & Colas, 2006). Young children's difficulty with the identification of complete declarative questions and statements in the absence of conversational context (Saindon et al., 2016) makes it unlikely that that they would be able to interpret the speaker's intentions from partial words (i.e., gates based on intonation contour rather than words). Eye-tracking measures (e.g., Zhou et al., 2012) could provide a viable means of assessing the impact of intonation contour on question and statement identification, but the use of such measures necessitates utterances with clear visual referents, unlike those in the present study.

Although 9- and 10-year-olds exhibited considerable success in question identification, especially for child-directed utterances, their failure to capitalize on speaker-specific intonation patterns prevented them from achieving adult performance levels. Similar challenges may underlie the failure of 13-year-olds to match adults' perception of emotional prosody (Aguert, Laval, Lacroix, Gil, & Le Bigot, 2013). Future research could examine parallels and differences in children's understanding of prosodic cues that signal questions and emotional tone by means of gating tasks involving familiar talkers.

Acknowledgment

Funding

References

Aguert, M., Laval, V., Lacroix, A., Gil, S., & Le Bigot, L. (2013). Inferring emotions from speech prosody: Not so easy at age five. *PLOS One*, 8, e83657. doi:10.1371/journal.pone.0083657

Armstrong, M. (2012). *The development of yes-no question intonation in Puerto Rican Spanish* (Doctoral dissertation). Ohio State University, Columbus, OH.

Bolinger, D. L. M. (1986). *Intonation and its parts: Melody in spoken English*. Stanford, CA: Stanford University Press.

Caporael, L. R. (1981). The paralanguage of caregiving: Baby talk to the institutionalized aged. *Journal of Personality and Social Psychology*, 40, 876–884. doi:10.1016/j.cogsci.2003.06.001

Creel, S. C. (2014). Tipping the scales: Auditory cue weighting changes over development. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 1146–1160.

Creel, S. C., Rojo, D. P., & Pallauda, A. N. (2016). Effects of contextual support on preschoolers' accented speech comprehension. *Journal of Experimental Child Psychology*, 146, 156–180. doi:10.1016/j.jecp.2016.01.018

Cutler, A., & Swinney, D. A. (1987). Prosody and the development of comprehension. *Journal of Child Language*, 14, 145–167. doi:10.1017/S0305000900012782

Doherty-Sneddon, G., & Kent, G. (1996). Visual signals and the communication abilities of children. *Journal of Child Psychology and Psychiatry*, 37, 949–959. doi:10.1111/j.1469-7610.1996.tb01492.x

Face, T. L. (2005). F0 peak height and the perception of sentence type in Castilian Spanish. *Revista Internacional de Lingüística Iberoamericana*. 3, 49–65. Retrieved from http://www.jstor.org/stable/41678113

Falé, I., & Faria, I. H. (2006). Time course of intonation processing in European Portuguese: A gating paradigm approach. *IV Jornadas En Tecnologia Del Habla*, 265–270. Retrieved from http://www.researchgate.net/profile/Isabel_Fale/publication/237244913_Time_course_of_Intonation_Processing_in_European_Portuguese_a_Gating_paradigm_approach/links/0c96053bdee6897003000000.pdf

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, 27, 209–221. doi:10.1037//0012-1649.27.2.209

Foulkes, P., Docherty, G., & Watt, D. (2005). Phonological variation in child-directed speech. *Language*, 81, 177–206. doi:10.1353/lan.2005.0018

Frota, S., Butler, J., & Vigário, M. (2014). Infants' perception of intonation: Is it a statement or a question? *Infancy*, 19, 194–213. doi:10.1111/infa.12037

Gårding, E., & Abramson, A. S. (1965). A study of the perception of some American English intonation contours. *Studia Linguistica*, 19, 61–79. doi:10.1111/j.1467-9582.1965.tb00527.x

Gérard, C., & Clément, J. (1998). The structure and development of French prosodic representations. *Language and Speech*, 41, 117–142. doi:10.1177/002383099804100201

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, 28, 267–283. doi:10.3758/BF03204386

Jacobson, J. L., Boersma, D. C., Fields, R. B., & Olson, K. L. (1983). Paralinguistic features of adult speech to infants and small children. *Child Development*, 54, 436–442. doi:10.2307/1129704

Karzon, R. G., & Nicholas, J. G. (1989). Syllabic pitch perception in 2- to 3-month-old infants. *Perception & Psychophysics*, 45, 10–14. doi:10.3758/bf03208026

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge, UK: Cambridge University Press.

Levi, S. V. (2015). Talker familiarity and spoken word recognition in school-age children. *Journal of Child Language*, 42, 843–872. doi:10.1017/S0305000914000506

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. *Journal of Child Language*, 36, 909–922. doi:10.1017/S030500090800929X

Loeb, D. F., & Allen, G. D. (1993). Preschoolers' imitation of intonation contours. *Journal of Speech Language and Hearing Research*, 36, 4–13. doi:10.1044/jshr.3601.04

Ma, J. K. Y., Whitehill, T. L., & So, S. Y. S. (2010). Intonation contrast in Cantonese speakers with hypokinetic dysarthria associated with Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, 53, 836–849. doi:10.1044/1092-4388(2009/08-0216)

Masataka, N. (2002). Pitch modification when interacting with elders: Japanese women with and without experience with infants. *Journal of Child Language*, 29, 939–951. doi:10.1017/S0305000902005378

Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior & Development*, 21, 779–784. doi:10.1016/s0163-6383(98)90044-3

O'Shaughnessy, D. (1979). Linguistic features in fundamental frequency patterns. *Journal of Phonetics*, 7, 119–145.

Patel, R. (2003). Acoustic characteristics of the question-statement contrast in severe dysarthria due to cerebral palsy. *Journal of Speech, Language and Hearing Research*, 46, 1401–1415. doi:10.1044/1092-4388(2003/109)

Patel, R., & Grigos, M. (2006). Acoustic characterization of the question-statement contrast in 4, 7, and 11 year-old children. *Speech Communication*, 48, 1308–1318. doi:10.1016/j.specom.2006.06.007

Peng, S. C., Chatterjee, M., & Lu, N. (2012). Acoustic cue integration in speech intonation recognition with cochlear implants. *Trends in Amplification*, 16, 67–82. doi:10.1177/1084713812451159

Peng, S. C., Lu, N., & Chatterjee, M. (2009). Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners. *Audiology and Neurotology*, 14, 327–337. doi:10.1159/000212112

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Pogue, A., Kurumada, C., & Tanenhaus, M. K. (2016). Talker-specific generalization of pragmatic inferences based on under- and over-informative prenominal adjective use. *Frontiers in Psychology*, 6, 2035. doi:10.3389/fpsyg.2015.02035

Saindon, J. R., Trehub, S. E., Schellenberg, E. G., & Van Lieshout, P. (2016). Children's identification of questions from rising terminal pitch. *Journal of Child Language*, 43, 1174–1191. doi:10.1017/s0305000915000458

Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compass*, 3, 236–264. doi:10.1111/j.1749-818X.2008.00112.x

Snow, D. (1994). Phrase-Final Syllable Lengthening and Intonation in Early Child Speech. *Journal of Speech, Language and Hearing Research*, 37, 831–840. doi:10.1044/jshr.3704.831

Snow, D. (1998). Children's imitations of intonation contours: Are rising tones more difficult than falling tones? *Journal of Speech, Language and Hearing Research*, 41, 576–587. doi:10.1044/jslhr.4103.576

Soderstrom, M., Ko, E.-S., & Nevzorova, U. (2011). It's a question? Infants attend differently to yes/no questions and declaratives. *Infant Behavior and Development*, 34, 107–110. doi:10.1016/j.infbeh.2010.10.003

Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0: A free software package for implementing psychological and psychophysiological experiments. *Experimental Psychology*, 57, 36–45. doi:10.1027/1618-3169/a000005

Stalinski, S. M., Schellenberg, E. G., & Trehub, S. E. (2008). Developmental changes in the perception of pitch contour: Distinguishing up from down. *The Journal of the Acoustical Society of America*, 124, 1759–1763. doi:10.1121/1.2956470

Studdert-Kennedy, M., & Hadding, K. (1973). Auditory and linguistic processes in the perception of intonation contours. *Language and Speech*, 16, 293–313.

Syrett, K., & Kawahara, S. (2014). Production and perception of listener-oriented clear speech in child language. *Journal of Child Language*, 1373–1389. doi:10.1017/S0305000913000482

Theodore, R. M., Myers, E. B., & Lomibao, J. A. (2015). Talker-specific influences on phonetic category structurea). *The Journal of the Acoustical Society of America*, 138, 1068–1078. doi:10.1121/1.4927489

Thorpe, L. A., Trehub, S. E., Morrongiello, B. A., & Bull, D. (1988). Perceptual grouping by infants and preschool children. *Developmental Psychology*, 24, 484–491. doi:10.1037/0012-1649.24.4.484

Tyler, L. K., & Marslen-Wilson, W. (1978). Some developmental aspects of sentence processing and memory. *Journal of Child Language*, 5, 113–129. doi:10.1017/s0305000900001975

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-LI-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, 49, 2–7. doi:10.1016/j.specom.2006.10.003

Van Heuven, V. J., & Haan, J. (2000). Phonetic correlates of statement versus question intonation in Dutch. In A. Botinis (Eds.), *Intonation: Analysis, modelling, and technology* (pp. 119–144). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Vion, M., & Colas, A. (2006). Pitch cues for the recognition of yes-no questions in French. *Journal of Psycholinguistic Research*, 35, 427–445. doi:10.1007/s10936-006-9023-x

Wells, B., Peppé, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, 31, 749–778. doi:10.1017/s030500090400652x

Zhou, P., Crain, S., & Zhan, L. (2012). Sometimes children are as good as adults: The pragmatic use of prosody in children's on-line sentence processing. *Journal of Memory and Language*, 67, 149–164. doi:10.1016/j.jml.2012.03.005