J. Child Lang. **44** (2017), 968–994. © Cambridge University Press 2016
doi:10.1017/S0305000916000349

Prosodic differences between declaratives and interrogatives in infant-directed speech*

SUSAN GEFFEN
Department of Psychology, University of Southern California

AND

TOBEN H. MINTZ
Department of Psychology, University of Southern California, and Department of Linguistics, University of Southern California

(*Received* 17 *April* 2015 – *Revised* 8 *April* 2016 – *Accepted* 20 *June* 2016 – *First published online 18 July 2016*)

ABSTRACT

In many languages, declaratives and interrogatives differ in word order properties, and in syntactic organization more broadly. Thus, in order to learn the distinct syntactic properties of the two sentence types, learners must first be able to distinguish them using non-syntactic information. Prosodic information is often assumed to be a useful basis for this type of discrimination, although no systematic studies of the prosodic cues available to infants have been reported. Analysis of maternal speech in three Standard American English-speaking mother–infant dyads found that polar interrogatives differed from declaratives on the patterning of pitch and duration on the final two syllables, but *wh*-questions did not. Thus, while prosody is unlikely to aid discrimination of declaratives from *wh*-questions, infant-directed speech provides prosodic information that infants could use to distinguish declaratives and polar interrogatives. We discuss how learners could leverage this information to identify all question forms, in the context of syntax acquisition.

[*] We would like to thank the members of the USC Language Development Lab for their help with this research as well as the families that participated in this study. This material is based upon work supported by the National Science Foundation under Grant No. BCS-1227074 awarded to T. Mintz and S. Geffen. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Address for correspondence: Susan Geffen, PhD, Department of Psychology, University of Calgary, Calgary, AB T2N 1N4, Canada. e-mail: susan.geffen@ucalgary.ca

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

INTRODUCTION

Distinguishing between declaratives and interrogatives is an important ability for language users. In English and many other languages, interrogatives generally involve syntactic structures and word order patterns that are different from declaratives. For example, in English the canonical ordering in simple transitive declarative sentences is subject–verb–object (Slobin & Bever, 1982), as in *Anna likes chocolate*. However, this ordering is disrupted in the associated polar interrogative, which inverts the subject and auxiliary verb, and inserts an utterance-initial *do*, since no auxiliary is present. Thus, *Anna likes chocolate* becomes *Does Anna like chocolate?* Moreover, in so-called *wh*-questions, a grammatical phrase that normally occurs in the declarative form is absent, and its content is queried by a *wh*-word – e.g. *who, what, where, how*, etc. – that occurs at the beginning of the sentence (*What$_t$ does Anna like t_i?*). (The *t* stands for 'trace', and represents the structural location of the constituent that is related to the *wh*-word that shares its index – in this case, the direct object of *like*.) Similar kinds of word order variations occur in many other (primarily Indo-European) languages (e.g. Dryer, 2011; König & Siemund, 2007).

In many prominent theories of syntax acquisition, children learn basic word order (i.e. the relative placement of subject nouns, object nouns, and verbs) and morphosyntax by bootstrapping off of declarative sentences (typically transitive declaratives) and typical linkages between syntactic positions in simple declarative sentences and elements of the scenes they describe (Fisher, Hall, Rakowitz & Gleitman, 1994; Pinker, 1984). Thus, given the differences in word order and morphosyntactic patterning of declaratives and interrogatives, distinguishing the two sentence types is crucial for correctly learning the syntax of both types of utterances. Learning mechanisms that analyzed declaratives and interrogatives as the same type would result in cascading errors in syntax acquisition, since fundamental misanalysis early on would result in incorrect representations in learning as subsequent utterances are processed (Pinker, 1984). Since interrogatives constitute close to half of the utterances in young children's input, in roughly equal proportions of polar interrogatives and *wh*-questions (23% and 21%, respectively; Newport, 1977), understanding when and how learners start to differentiate interrogatives and declaratives is important for understanding grammatical acquisition more broadly.

One way of distinguishing declaratives and interrogatives would simply be to use the word order properties just described as a basis for segregating sentence types. Indeed, 12-month-olds appear to be able to use word order information in just this way (Geffen & Mintz, 2015). However, there is reason to doubt that infants could use distributional information of this

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

type as the INITIAL source of information for discriminating sentences. To see why, consider the kind of information infants would have to rely on if they did not already know the distributional properties of declarative syntax. The distributional information involves a small set of highly frequent closed-class words (auxiliary verbs and *wh*-words) occurring in sentence-initial position. While infants could classify sentences as interrogatives based on this property, it is hard to see how infants could discover this property a priori. For one thing, if such a segregation is based on words rather than categories – i.e. *who, what, which, does*, etc., rather than *wh*-word, AUX – then it is not clear how to avoid segregating other sentences that are syntactically similar but differ in their initial words. For example, such a procedure would create distinctions between utterances that start with determiners and those that start with pronouns, despite their shared underlying structure (e.g. sentences like *The boy saw the girl* vs. *You saw the girl*). This problem might be avoided with knowledge of the syntactic categories (e.g. *wh*-operator, AUX, NP), but it is not clear that infants have the requisite category knowledge at a point in acquisition where making sentence-type discrimination would be important. Furthermore, even if infants did have that category knowledge, it is not clear how they would know to use those category distinctions – i.e. *wh*-word and AUX vs. everything else – and not other distinctions as a basis for segregating declaratives and interrogatives. Indeed, being able to distinguish declaratives from interrogatives might be a prerequisite for successfully learning at least some of those categories in the first place, as well as their correlation with sentence type. Thus, another source of information is likely to be necessary to make initial sentence-type distinctions. There is reason to believe that sentence-level prosody provides just such an information source.

Prosody as a cue to sentence type

Prosody consists of pitch, duration, and intensity. Here we outline the potential role of these components in cueing sentence type.

Pitch. In many languages, the final pitch contour is an important perceptual cue for sentence-type discrimination (American English: Săfárŏvá & Swerts, 2004; Castilian Spanish: Face, 2007; French: Vion & Colas, 2006). In adult-directed speech, polar interrogatives typically end with a final rise in pitch, while declaratives end with final flat or falling intonation (Bartels, 1999; Hedberg, Sosa & Fadden, 2004; Ladd, 2008), although there are some exceptions (e.g. Belfast English: Jarman & Cruttenden, 1976; Chickasaw: Gordon, 1999). The same is true of declarative questions, which differ from declarative statements only in prosody (e.g. *Anna likes chocolate?*). If speech to infants is similarly

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

organized, prosody could provide a robust tool for initially distinguishing sentence types. Perceptual studies have demonstrated that, while the exact cues adults use to distinguish between declaratives and interrogatives may vary by language, most languages rely on some form of pitch cue (e.g. Gussenhoven & Chen, 2000). For example, although Russian has *wh*-words, sentences often do not include lexical or syntactic/grammatical indicators of sentence type (Makarova, 2007). Russian has relatively free word order and no auxiliary verbs (Rojina, 2004), so listeners rely on pitch peak alignment rather than lexical cues to perceive distinctions between declaratives and polar interrogatives (Makarova, 2007; Svetozarova, 1998). Declaratives and exclamations have earlier accent pitch peaks (here defined as THE F0 MAXIMUM ALONG THE SEGMENTAL STRING OF A THREE-SYLLABLE WORD; Makarova, 2007) than interrogatives (this is also seen in Dutch: van Heuven & Haan, 2002). There is also an effect of pitch peak height, where interrogatives have higher pitch peaks than declaratives (Makarova, 2007). In languages like English and French, which do use lexical markings (e.g. auxiliary verbs) to distinguish sentence types, adults primarily rely on fundamental frequency for perceiving distinctions between sentence types (English: Cruttenden, 1986; Lieberman, 1967; French: Vion & Colas, 2006).

In many languages, the final pitch contour is a critical (if not the most important) region for discriminating between interrogatives and declaratives. A gating study with French noun phrases found that listeners' accuracy in making a sentence-type judgment improved nearly 40% (61% to 97%) when final pitch information was introduced at the end of an utterance. In the absence of this final prosodic information, subjects defaulted to a declarative judgment (Vion & Colas, 2006). Various dialects of English also rely on final pitch contours as an important perceptual cue for sentence-type discrimination (British English [e.g. London, Cambridge]: Grabe, 2004; although see Jarman & Cruttenden, 1976). For example, Šafárŏvá and Swerts (2004) found that American English speakers were more likely to identify utterances as declarative questions if the utterances ended with a final rise.

Duration. Duration is a secondary prosodic correlate of sentence type in other Germanic languages (e.g. Dutch, Orkney English: van Heuven & van Zanten, 2005) in which pitch is the primary prosodic cue. For instance, in Dutch, interrogatives typically have a faster speech rate than declaratives, although this difference is more apparent in declarative questions than polar interrogatives. In addition, perceptual studies have demonstrated that 9-month-olds can use syllable duration differences to detect syntactic boundaries, but only in conjunction with other converging cues such as pitch changes (e.g. Hirsh-Pasek, Kemler Nelson, Jusczyk, Cassidy, Druss & Kennedy, 1987; Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward & Piwoz, 1992). Thus, to the degree that duration

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

covaries with pitch in relation to sentence type, infants could potentially use duration in combination with pitch to discriminate declaratives and interrogatives.

Intensity. Intensity correlates with the other dimensions of prosody, especially pitch, and infants appear to be sensitive to this correlation (Fernald, 1984). For example, in English, stressed syllables are marked by higher pitch and intensity compared to non-stressed syllables (Hay & Saffran, 2012). Thus, as with duration, to the extent that intensity correlates with other prosodic cues that differentiate declaratives from interrogatives, intensity could contribute to sentence-type discrimination.

Summary of prosodic cues with respect to sentence type. To summarize the research we have just discussed, in English and many other languages, declaratives and interrogatives give rise to different prosodic patterns, particularly in the final pitch contour. Moreover, adult listeners rely on prosodic information to make sentence-type distinctions in perception. However, with respect to English, the distinguishing prosodic difference between declaratives and interrogatives seems to hold only with respect to polar interrogatives (i.e. *yes/no* questions), and not *wh*-questions. *Wh*-questions usually have a falling or level pitch, and thus have final contours similar to declaratives (Hedberg *et al*., 2004; Ladd, 2008; Ladefoged & Johnson, 2010), at least in adult-directed speech. One important question we address in this paper is whether similar prosodic patterns across sentence types are evident in infant-directed speech, and how infants might use prosody to discriminate sentence types. Establishing the prosodic characteristics of *wh*-questions, polar interrogatives, and declaratives is critical for understanding how prosody could be used by infants to classify utterances in syntactically meaningful ways. However, before discussing our analyses, we briefly discuss prior research on infants' sensitivity to prosody.

Infants' sensitivity to prosody

There is some evidence that young infants might start out using prosody as a perceptual cue to sentence-type distinctions. Frota, Butler, and Vigário (2014) tested infants learning European Portuguese, a language which distinguishes between sentence types using ONLY prosodic information, and found that infants as young as five to six months can distinguish between declaratives and polar interrogatives. Soderstrom, Ko, and Nevzorova (2011) tested English-learning infants between 0;4·5 and 2;0 years of age and found that they could distinguish declarative statements from declarative questions, which only differ in their prosodic contours, suggesting some ability on the part of infants to discriminate sentence types based on prosody. However, the wide age range of the subjects

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

(between 0;4·5 and 2;0) makes it difficult to determine when this ability begins. Moreover, speakers may modify prosody more sharply to mark declarative questions as questions, compared to more typical interrogative types (e.g. polar interrogatives).

While it remains unclear when English-learning infants begin to use prosodic information to distinguish between sentence types, there is a wealth of evidence demonstrating infants' general sensitivity to prosodic information from an early age. Infants can use prosody to make distinctions between broad rhythmic classes of languages (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini & Amiel-Tison, 1988; Moon, Cooper & Fifer, 1993; Nazzi, Bertoncini & Mehler, 1998; see Nespor, Shukla & Mehler, 2011, for an overview of rhythmic classification). In addition, infants can use their sensitivity to prosodic cues to identify syntactic unit boundaries including clauses (e.g. Hirsh-Pasek *et al*., 1987; Kemler Nelson, Hirsh-Pasek, Jusczyk, and Cassidy, 1989) and phrases (Gerken, Jusczyk & Mandel, 1994; Jusczyk *et al*., 1992). Thus, there is ample evidence that infants are sensitive to prosody, suggesting that they could, in principle, use prosodic information to make initial sentence-type distinctions.

OUTLINE OF THE STUDY

The goal of the present study was to evaluate the prosodic properties of declaratives and interrogatives in infant-directed speech in American English, with the aim of understanding whether prosodic information would be useful, in principle, for the initial discrimination of sentence types by infants. There has not, to our knowledge, been any thorough analysis of infant-directed speech that establishes the availability of prosodic cues to sentence type. As we indicated earlier, such information could play an important role in early syntactic acquisition.

Recall that, in adult-directed speech, polar interrogatives and *wh*-questions have different prosodic properties (Hedberg *et al*., 2004), with *wh*-questions patterning similarly to declaratives (Ladd, 2008). Therefore, a particular question we address is whether the two question types have similar prosodic properties in infant-directed speech, as well as how they differ from declaratives. To address this question, we examined typical prosodic contours in the speech of three mothers to their preverbal infants. We evaluated the prosodic characteristics of declaratives, polar interrogatives, and *wh*-questions to determine whether there is a common set of prosodic features that distinguish declaratives and either type of interrogative. Based on the findings with adult-directed speech, we hypothesized that prosodic information would distinguish between declaratives and polar interrogatives, but that prosody would not be a

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

viable source of information for distinguishing between declaratives and *wh-questions*.

In Analysis 1, we describe the prosodic measures we examined across sentence types, and present results organized as contrasting pairs of sentence types (declaratives vs. polar interrogatives, declaratives vs. *wh*-questions, and polar interrogatives vs. *wh*-questions). Broadly, we measured the distribution of pitch (F0), duration, and intensity over the final two syllables of declaratives, *wh*-questions, and polar interrogatives, in speech addressed to infants, and we present two sets of analyses based on these measurements. To preview our results, we do indeed find that prosodic measures differ between declaratives and polar interrogatives, but not *wh*-questions. In Analysis 2, we present a linear classification model that, when trained on the prosodic characteristics of a subset of declaratives and polar interrogatives, can classify untrained utterances at above-chance levels. Finally, we examine the implications of our results for infants' sentence-type discrimination ability.

ANALYSIS 1: PROSODIC DIFFERENCES BETWEEN DECLARATIVES AND INTERROGATIVES

METHODS

Input corpora

We evaluated input to three English-learning children from the Brent corpus (Brent & Siskind, 2001, files s1, s2, and v1) from the CHILDES database (MacWhinney, 2000) for the analysis procedure. We selected the Brent corpus because it contains speech to preverbal infants, and provides the largest number of audio-recordings of Standard American English-speaking mother–child dyads of any corpora in the database. We only analyzed utterances of the mother directed towards her preverbal infant, taken from two to three consecutive sessions spanning a 2–4 week period depending on the subject. Infants ages ranged from 0;8·28 to 0;10·04. Recordings were made in the subjects' homes. Mothers were fitted with a small waist pack and a lapel microphone located just below their mouths, and instructed to go about their daily routines while home alone with the child (Brent & Siskind, 2001).

Selection of utterances

We analyzed approximately thirty utterances per sentence type (declarative, *wh*-question, and polar interrogative) for each speaker, for a total of roughly ninety utterances per speaker. Sentences had to meet several criteria for inclusion. We based our initial selection criteria on structural properties. Polar interrogatives were characterized by subject and auxiliary inversion,

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

possibly in combination with *do*-support. *Wh*-questions were characterized by the typical auxiliary inversion structure of polar interrogatives with a *wh*-word such as *who, what, where*, or *how* in the utterance-initial position. Declaratives followed the canonical transitive word order. There were several exceptions to these criteria. We allowed three declarative sentence fragments – grammatical phrases that were not complete sentences, e.g. *some more carrot*). We referred to the context and punctuation coded in the transcript to determine that these were indeed declaratives. We also included one polar interrogative fragment: *Are you?*

Our analysis primarily focused on the initial instance of a given utterance. (Approximately 11% of utterances (35 out of 311) occurred two or more times in a single session, with the majority of repetitions being questions (19 repeated utterances were *wh*-questions, 13 were polar interrogatives). Of those thirty-five utterances, we analyzed repetitions of nineteen utterances. The remaining sixteen utterances contained repetitions with background vocalizations that made them difficult or impossible to analyze. Results did not change with the addition of the duplicate utterances.) We narrowed our selection further by listening to the accompanying waveform. Utterances were eliminated if they included vocalization or background noise, making analysis difficult or impossible: For example, laughing, crying, blowing raspberries. We eliminated unintelligible sentences because we used the mother's pronunciation to help define syllable boundaries. We analyzed a total of 311 sentences (See Table 1).

Location of analyses within sentences

Separate analyses of each acoustic measure are detailed below. We analyzed the penultimate and final syllables (following the precedent of Fisher & Tokura, 1996) given that, in many languages, the final pitch contour is a critical (if not the most important) region in the utterance where prosodic differences between interrogatives and declaratives occur. We reasoned that the final two syllables were likely to contain most of the variability in final pitch contour. For each utterance, coders marked the boundaries for the penultimate and final syllables for analysis by hand. Boundaries were marked by one or two coders, then verified by the first author. Coders initially marked syllable boundaries by examining the spectrogram and listening to the corresponding waveform. Coders also used regions of silence to mark the end boundary of the final syllable. Disagreements were resolved by having all the coders listen to the sentences, followed by a group discussion, and if there was still a disagreement, the first author decided on the appropriate marking. (Discrepancies occurred in approximately 3% of sentences, and were generally off by only a few

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

TABLE 1. *Frequency of utterance types*

Utterance types	Examples	Frequency (%)[a]
A. Speaker 1 (s1)		
Declarative	That's your mouth. I'm gonna get it.	31 (34·4)
Wh-question	Where did he go? How big is he?	28 (31·1)
Polar question	Are you gonna crawl? Is that your spoon?	31 (34·4)
Total		90 (28·9)
B. Speaker 2 (s2)		
Declarative	Now you closed it. I'll bet you can.	36 (37·1)
Wh-question	What did you think? Who is that?	30 (30·9)
Polar interrogative	Can you find it? Is that Piglet?	31 (32·0)
Total		97 (31·2)
C. Speaker 3 (v1)		
Declarative	You can look at it. They're getting squished.	44 (35·5)
Wh-question	How do you get down? Where did it go?	39 (31·5)
Polar interrogative	Do you see the kitty? Can you get that?	41 (33·1)
Total		124 (39·9)
Total utterances. . . .		311

[a] Frequency (%) for each sentence type is the percentage of utterances for a given speaker (e.g. declaratives account for 34·4% of the total utterances for Speaker 1). The % total is the percentage out of the total number of utterances analyzed (e.g. Speaker 1's utterances account for 28·9% of the total utterances in the corpus analysis).

milliseconds.) Coders labeled each section with a rough English transcription of the syllable. This was a necessary step for information extraction using the ProsodyPro Praat script (Xu, 2012), which carried out batch extraction of MeanFo, MaxFo, MinFo, duration, and intensity from each labeled interval within all 311 files (using the default settings, *Fo range*: 30–400 Hz; *Fo sample rate*: 100 Hz; *Npoints per interval*: 10).

Pitch variables

To assess the availability of pitch as a cue to sentence type in infant-directed speech, we measured average pitch (MeanFo) as well as lowest and highest pitch (MinFo and MaxFo, respectively).

For MeanFo and MaxFo, we were interested primarily in the change of these values across the final two syllables (see 'Methods' section). MeanFo provides a simple, coarse measure that can be used to identify the general degree of sentence-final pitch rise or fall. MaxFo, while clearly related to MeanFo, provides specific information about pitch peak. Recall that some languages, including Russian (Makarova, 2007; Svetozarova, 1998), and Hungarian (Gósy & Terken, 1994), rely on pitch peak alignment rather than lexical cues to mark distinctions between declaratives and polar interrogatives. Declaratives and exclamations have earlier peaks than interrogatives. There is also an effect of pitch peak height, where interrogatives have higher pitch peaks than declaratives. Even languages

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

like English, which do use lexical markings (e.g. auxiliary verbs) to distinguish sentence types, demonstrate differences in pitch peak location between sentence types. For example, English *wh*-questions show an earlier pitch rise and final Fo decline (Best, Levitt & McRoberts, 1991). Thus, the location and degree of MaxFo could be a useful cue for distinguishing between sentence types.

We used MinFo primarily as a way of calculating the derived measure of pitch excursion across the final two syllables that is based on the most extreme values (see 'Methods' section) rather than on averages.

Since infant-directed speech is characterized by exaggerated prosodic contours, including higher mean pitch and expanded pitch excursion (Fernald & Kuhl, 1987), it is an open question which properties of pitch contours that distinguish between sentence types in adult-directed speech are maintained in speech to infants.

For analyzing PITCH EXCURSION, we identified the lowest Fo and highest Fo when considering the penultimate and final syllables together. We then subtracted whichever extreme value was in the penultimate syllable from the extreme value that was in the final syllable. Thus, positive pitch excursions indicated rising pitch, and negative pitch excursions indicated falling pitch. For example, if the lowest pitch was 207·9 Hz in the penultimate syllable and the highest Fo was 256·8 Hz in the final syllable, the pitch excursion would be a pitch rise of 48·9 Hz (256·8–207·9). When both extreme values came from the same syllable, we subtracted the extreme that came first from the extreme that came last, resulting in the same correspondence of the sign of the excursion with excursion direction.

MinFo and MaxFo were not available for twenty-one sentences (12 declaratives, 1 polar interrogative, and 8 *wh*-questions) (7%) due to incomplete pitch information extracted by the Praat script (Xu, 2012), so they were excluded from the pitch excursion analysis.

Duration and intensity

Duration and average intensity were extracted from the marked syllables in Praat using the ProsodyPro Praat script (Xu, 2012) (see section, 'Location of analyses within sentences', above).

RESULTS

Summaries of all direct measurements are given in Table 2. Summaries of the derived measures are given in Table 3.

Results are presented as contrasting pairs of sentence types: declaratives versus polar interrogatives, declaratives versus *wh*-questions, and polar interrogatives versus *wh*-questions by measure. With the exception of the pitch excursion analysis, which only focused on differences between

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

TABLE 2. *Mean and standard deviation by position in utterance*

Sentence type	Penultimate	Final
MEANFo (Hz)		
Declaratives ($n = 96$)	230·96 (*59·60*)	221·83 (*53·00*)
Polar interrogatives ($n = 100$)	242·44 (*57·81*)	264·87 (*54·23*)
Wh-questions ($n = 88$)	238·58 (*49·75*)	229·53 (*57·26*)
MAXFo (Hz)		
Declaratives ($n = 111$)	262·61 (*69·36*)	251·58 (*66·96*)
Polar interrogatives ($n = 103$)	268·94 (*62·75*)	308·49 (*74·43*)
Wh-questions ($n = 97$)	267·39 (*67·44*)	265·94 (*75·94*)
DURATION (ms)		
Declaratives ($n = 111$)	170·25 (*83·62*)	310·72 (*181·73*)
Polar interrogatives ($n = 103$)	168·75 (*86·55*)	267·75 (*144·14*)
Wh-questions ($n = 97$)	170·27 (*79·31*)	297·52 (*149·69*)
INTENSITY (dB)		
Declaratives ($n = 110$)	61·60 (*8·67*)	61·75 (*8·29*)
Polar interrogatives ($n = 103$)	63·42 (*6·70*)	65·05 (*6·62*)
Wh-questions ($n = 97$)	59·51 (*6·78*)	59·15 (*7·56*)

sentence types, not syllable positions, data were subjected to a mixed 2 × 2 ANOVA with Sentence Type as a fixed effect and Syllable Position (penultimate or final) as a repeated measure. (Adding Speaker as an additional variable did not change the pattern of results.) We were primarily interested in an interaction of Syllable Position with Sentence Type.

MeanFo

Analyses revealed a significant main effect of Sentence Type and a significant Sentence Type by Syllable Position interaction for declaratives compared to polar interrogatives and for polar interrogatives compared to *wh*-questions. In contrast, analyses revealed a main effect of Syllable Position only for declaratives compared to *wh*-questions (see Tables 4–6 for details of the ANOVAs). Polar interrogatives had significantly higher mean pitch (M = 264·87, SD = 54·23) for the final syllable compared to declaratives (M = 221·83, SD = 53·0; $t(199) = 5{\cdot}88$, $p < {\cdot}001$) and *wh*-questions ($M = 229{\cdot}01$, $SD = 57{\cdot}14$; $t(189) = 4{\cdot}46$, $p < {\cdot}001$). However, consistent with the significant interaction, there was no significant difference between sentence types in the penultimate syllable. In contrast, there were no significant differences in MeanFo values for the penultimate or the final syllable in *wh*-questions compared to declaratives (ps > ·1). There was a fall in mean pitch from the penultimate to the final syllable in declaratives ($M = -9{\cdot}13$ Hz, $SD = 67{\cdot}05$ Hz) and *wh*-questions ($M = -9{\cdot}05$ Hz, $SD = 53{\cdot}38$ Hz), while polar interrogatives showed a rise in mean pitch ($M = 22{\cdot}43$ Hz,

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

TABLE 3. *Mean difference scores (standard error) and pitch excursion (standard error)*

Sentence type	Final–Penultimate	Highest pitch–Lowest pitch
MEANFo (Hz)		
Declaratives ($n = 96$)	−9·13 (*6·84*)	
Polar interrogatives ($n = 100$)	22·43 (*6·04*)	
Wh-questions ($n = 88$)	−9·05 (*5·69*)	
DURATION (ms)		
Declaratives ($n = 111$)	140·47 (*19·64*)	
Polar interrogatives ($n = 103$)	99·00 (*18·72*)	
Wh-questions ($n = 97$)	127·26 (*16·31*)	
PITCH EXCURSION (Hz)		
Declaratives ($n = 111$)		−102·27 (*354·17*)
Polar interrogatives ($n = 103$)		109·81 (*420·82*)
Wh-questions ($n = 97$)		−100·64 (*355·43*)

TABLE 4. *Statistical comparisons for declaratives vs. polar interrogatives*

Measure	Main effect of sentence type	Main effect of syllable position	Sentence type × Syllable position interaction
Mean Fo	**$F(1,194) = 16·97$***, partial $\eta^2 = ·080$**	$F(1,194) = 2·129$	**$F(1,194) = 12·001$***, partial $\eta^2 = ·058$**
Max Fo	**$F(1,212) = 15·830$***, partial $\eta^2 = ·069$**	**$F(1,212) = 8·245$*, partial $\eta^2 = ·037$**	**$F(1,212) = 25·956$***, partial $\eta^2 = ·109$**
Duration	$F(1,212) = 3·623$	**$F(1,212) = 77·411$***, partial $\eta^2 = ·267$**	$F(1,212) = 2·322$
Intensity	**$F(1,211) = 6·723^{\dagger}$, partial $\eta^2 = ·031$**	**$F(1,211) = 6·502^{\dagger}$, partial $\eta^2 = ·030$**	**$F(1,211) = 4·443^{\dagger}$, partial $\eta^2 = ·021$**

NOTES: † $p < ·05$; * $p < ·01$; ** $p < ·005$; *** $p < ·001$. Statistically significant statistics are in **bold**.

TABLE 5. *Statistical comparisons for declaratives vs.* wh*-questions*

Measure	Main effect of sentence type	Main effect of syllable position	Sentence type × Syllable position interaction
Mean Fo	$F(1,182) < ·001$	**$F(1,182) = 4·095^{\dagger}$, partial $\eta^2 = ·022$**	$F(1,182) = 1·275$
Max Fo	$F(1,206) = 1·306$	$F(1,206) = 1·607$	$F(1,206) = 0·949$
Duration	$F(1,206) = 0·260$	**$F(1,206) = 106·353$***, partial $\eta^2 = ·340$**	$F(1,206) = 0·259$
Intensity	**$F(1,205) = 5·089^{\dagger}$, partial $\eta^2 = ·024$**	$F(1,205) = 0·079$	$F(1,205) = 0·504$

NOTES: † $p < ·05$; * $p < ·01$; ** $p < ·005$; *** $p < ·001$. Statistically significant statistics are in **bold**.

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

TABLE 6. *Statistical comparisons for polar interrogatives vs.* wh*-questions*

Measure	Main effect of sentence type	Main effect of syllable position	Sentence type × Syllable position interaction
Mean Fo	**$F(1,186) = 8{\cdot}177$*, partial $\eta^2 = {\cdot}042$**	$F(1,186) = 2{\cdot}559$	**$F(1,186) = 14{\cdot}161$***, partial $\eta^2 = {\cdot}071$**
Max Fo	**$F(1,198) = 6{\cdot}814^{\dagger}$, partial $\eta^2 = {\cdot}033$**	**$F(1,198) = 13{\cdot}17$ ***, partial $\eta^2 = {\cdot}062$**	**$F(1,198) = 15{\cdot}246$***, partial $\eta^2 = {\cdot}071$**
Duration	$F(1,198) = 1{\cdot}892$	**$F(1,198) = 82{\cdot}209$***, partial $\eta^2 = {\cdot}293$**	$F(1,198) = 1{\cdot}283$
Intensity	**$F(1,198) = 29{\cdot}301$***, partial $\eta^2 = {\cdot}129$**	$F(1,198) = 3{\cdot}002$	**$F(1,198) = 7{\cdot}332$*, partial $\eta^2 = {\cdot}036$**

NOTES: † $p < {\cdot}05$; * $p < {\cdot}01$; ** $p < {\cdot}005$; *** $p < {\cdot}001$. Statistically significant statistics are in **bold**.

$SD = 60{\cdot}44$ Hz). In adult-directed speech, one of the primary distinctions between declaratives and polar interrogatives is the final prosodic contour – i.e. interrogatives have a final rise and declaratives have a final fall. *Wh*-questions also typically end with final falling or flat intonation (Bartels, 1999; Hedberg *et al.*, 2004). Larger differences in mean pitch between penultimate and final syllables could make it easier for infants to acoustically distinguish between sentence types. These results are consistent with adult-directed speech and demonstrate that MeanFo can be useful for distinguishing polar interrogatives from declaratives and *wh*-questions, but not declaratives from *wh*-questions.

MaxFo

Analyses revealed a significant main effect of Sentence Type, a significant main effect of Syllable Position, and a significant Sentence Type by Syllable Position interaction for polar interrogatives compared to declaratives and *wh*-questions. However, there were no main effects or interactions for declaratives compared to *wh*-questions (see Tables 4–6 for details of the ANOVAs). Declaratives showed a decline in MaxFo from the penultimate syllable ($M = 262{\cdot}61$, $SD = 69{\cdot}36$) to the final syllable ($M = 251{\cdot}58$, $SD = 66{\cdot}96$), consistent with the prosodic contour of declaratives in adult-directed speech. Polar interrogatives demonstrated a rise from the penultimate syllable ($M = 268{\cdot}94$, $SD = 62{\cdot}75$) to the final syllable ($M = 308{\cdot}49$, $SD = 74{\cdot}43$), consistent with the final rise pitch contour in adult-directed speech. *Wh*-questions did not appreciably differ in MaxFo between the penultimate syllable ($M = 267{\cdot}39$, $SD = 67{\cdot}44$) and the final syllable ($M = 265{\cdot}94$, $SD = 75{\cdot}94$), consistent with the prosodic contour of *wh*-questions in adult-directed speech. This demonstrates that MaxFo provides information infants could use to distinguish polar interrogatives from declaratives and *wh*-questions, but not declaratives from *wh*-questions.

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

Pitch excursion

Independent samples *t*-tests revealed a significant difference in pitch excursion between polar interrogatives and declaratives ($t(212) = 4{\cdot}00$, $p < {\cdot}01$) and *wh*-questions ($t(198) = 3{\cdot}81$, $p < {\cdot}01$). However, there was no significant difference in pitch excursion between declaratives and *wh*-questions ($t(206) = 0{\cdot}033$, $p = {\cdot}97$). Polar interrogatives had an average pitch excursion of 109·81 Hz ($SD = 420{\cdot}82$ Hz), declaratives had an average pitch excursion of −102·27 Hz ($SD = 354{\cdot}62$ Hz), and *wh*-questions had an average pitch excursion of −100·64 Hz ($SD = 355{\cdot}43$ Hz) across the final two syllables. The positive excursion for polar interrogatives indicates a pitch rise, whereas the negative excursion for declaratives indicates a pitch fall.

Duration

Analyses revealed a significant main effect of Syllable Position for all three comparisons. There was also a non-significant trend for Sentence Type for declaratives compared to polar interrogatives ($p = {\cdot}058$) (see Tables 4–6 for details of the ANOVAs). The penultimate syllables did not appreciably differ in duration for polar interrogatives ($M = 168{\cdot}75$ ms, $SD = 86{\cdot}55$ ms) compared to declaratives ($M = 170{\cdot}25$ ms, $SD = 83{\cdot}62$ ms; $t(212) = 0{\cdot}128$, $p = {\cdot}898$) or *wh*-questions ($M = 170{\cdot}27$ ms, $SD = 79{\cdot}31$ ms; $t(198) = 0{\cdot}128$, $p = {\cdot}898$). Final syllable durations in declaratives ($M = 310{\cdot}72$, $SD = 181{\cdot}73$) tended to be longer than in polar interrogatives ($M = 267{\cdot}75$, $SD = 181{\cdot}73$), although this difference was a non-significant trend ($t(207{\cdot}071) = 1{\cdot}923$, $p = {\cdot}056$). This difference could be due to the fact that the last word was one syllable in most declaratives but two syllables in at least half the polar interrogatives. These results are consistent with adult-directed speech, where declaratives and polar interrogatives also differ in overall duration, with polar interrogatives typically a little shorter than declaratives. This has been shown in a variety of languages (e.g. Manado Malay, Orkney English, and Dutch: van Heuven & van Zanten, 2005). Final syllable durations in *wh*-questions ($M = 297{\cdot}52$, $SD = 149{\cdot}69$) also tended to be longer than in polar interrogatives ($M = 267{\cdot}75$, $SD = 144{\cdot}14$), although this difference was not significant ($t(196{\cdot}12) = 1{\cdot}431$, $p = {\cdot}154$). In contrast, final syllables ($M = 304{\cdot}57$ ms, $SD = 167{\cdot}29$ ms) were significantly longer than penultimate syllables ($M = 170{\cdot}26$ ms, $SD = 81{\cdot}44$ ms) in both declaratives and polar interrogatives, although this durational pattern did not significantly differ across sentence types. These results suggest that duration would not be useful for distinguishing declaratives and *wh*-questions. However, larger differences in duration, particularly in the final syllable, could make it easier for infants to acoustically distinguish between sentence types. All three sentence types had positive difference

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

scores (see Table 3), indicating that the final syllable was longer for all three sentence types, although the difference scores were not significantly different between sentence types. These results suggest that duration would not be useful for distinguishing declaratives and *wh*-questions.

Intensity

Analyses revealed a significant main effect of Sentence Type for all three comparisons. There was a significant Sentence Type by Syllable Position interaction for polar interrogatives compared to declaratives and *wh*-questions. There was also a significant main effect of Syllable Position for polar interrogatives compared to declaratives and a non-significant trend toward a main effect of Syllable Position for polar interrogatives compared to *wh*-questions (see Tables 4–6 for details of the ANOVAs). Polar interrogatives had slightly higher intensity than declaratives and *wh*-questions, across positions. Intensity was significantly higher for polar interrogatives ($M = 65{\cdot}05$, $SD = 6{\cdot}62$) than for declaratives ($M = 61{\cdot}75$, $SD = 8{\cdot}29$; $t(205{\cdot}905) = 3{\cdot}225$, $p = {\cdot}001$) and *wh*-questions ($M = 59{\cdot}15$, $SD = 7{\cdot}56$; $t(190{\cdot}942) = 5{\cdot}864$, $p < {\cdot}001$) in the final syllable, and also in the penultimate syllable for polar interrogatives ($M = 63{\cdot}42$, $SD = 6{\cdot}70$) compared to *wh*-questions ($M = 59{\cdot}51$, $SD = 6{\cdot}78$; $t(196{\cdot}964) = 4{\cdot}104$, $p < {\cdot}001$), as well as a non-significant trend towards a difference from declaratives ($M = 61{\cdot}60$, $SD = 8{\cdot}63$; $t(205{\cdot}718) = 1{\cdot}733$, $p = {\cdot}085$). In addition, declaratives had higher intensity than *wh*-questions in the final and the penultimate syllable: declaratives (Final: $M = 61{\cdot}60$, $SD = 8{\cdot}18$; Penultimate: $M = 61{\cdot}75$, $SD = 8{\cdot}29$); *wh*-questions (Final: $M = 59{\cdot}15$, $SD = 7{\cdot}56$; Penultimate: $M = 59{\cdot}51$, $SD = 6{\cdot}78$). Polar interrogatives show corresponding rises in intensity and pitch consistent with findings on the correlation between intensity and other prosodic measures (Fernald, 1984). Therefore, intensity appears to be a reliable cue for making sentence-type distinctions in conjunction with other prosodic cues.

Summary

Similar to adult-directed speech, in infant-directed speech polar interrogatives differed from declaratives and *wh*-questions in pitch on the final two syllables, and showed a non-significant trend towards a difference in duration. Polar interrogatives were characterized by higher mean and max pitch, and higher intensity on the final syllable. In contrast, declaratives and *wh*-questions tended to demonstrate longer duration on the final syllable, although this difference was not statistically significant. In addition, the only prosodic cue that distinguished between declaratives and *wh*-questions was intensity, with *wh*-questions displaying a slightly lower intensity than declaratives. Polar interrogatives do not significantly differ

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

from declaratives or *wh*-questions in the penultimate syllable, and both declaratives and *wh*-questions had significantly longer final syllables compared to penultimate syllables. These results suggest that prosodic differences between polar interrogatives and both declaratives and *wh*-questions are controlled largely by the final syllable and its relation to the penultimate syllable. This suggests that there is insufficient prosodic information to distinguish between declaratives and *wh*-questions in infant-directed speech. However, sentence-final prosody does provide information that distinguishes polar interrogatives from declaratives and *wh*-questions in infant-directed speech, a point we return to below.

DISCUSSION

This analysis revealed significant differences between declaratives and polar interrogatives for MeanF0, MaxF0, pitch excursion, and intensity, and a non-significant trend between sentence types for differences in duration. Conversely, the only significant difference between declaratives and *wh*-questions was intensity. However, intensity patterned differently in polar interrogatives and *wh*-questions in comparison to declaratives, with declaratives and polar interrogatives showing similar intensity profiles. Collectively, these results demonstrate that prosodic information is in principle available to distinguish between declaratives and polar interrogatives. While intensity could be a basis for also discriminating *wh*-questions and declaratives, it is not obvious how intensity would be used more broadly to distinguish declaratives from interrogatives since it patterns differently across *wh*-questions and polar interrogatives.

The situation we have just described highlights the challenges in developing a comprehensive account of how infants could use prosody to make useful sentence-type discriminations. Although Analysis 1 shows that a variety of prosodic cues provide information that infants could use to distinguish between declaratives and polar interrogatives, it remains unclear how useful these combinations of cues are for actually categorizing individual utterances. We address this question in Analysis 2.

ANALYSIS 2: SENTENCE CLASSIFICATION FROM PROSODIC INFORMATION

Analysis 1 revealed that there are different prosodic patterns between declaratives and polar interrogatives when considering groups of sentences, but the analyses were not designed to address to what degree any of the cues individually, or in combination, are useful for categorizing individual sentences. To investigate this, in Analysis 2 we used a simple linear discrimination model, using half the sentences as a training set and

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

evaluating categorization success with the other half. To reduce the dimensionality of the data, we first carried out a Principal Components Analysis (PCA) on the acoustic measures. This is a fairly simple classification technique that provides a first attempt to assess how well these acoustic measures classify actual sentences into declaratives and interrogatives.

PCA ANALYSIS AND RESULTS

We used PCA to standardize and assess thirteen variables which comprise acoustic properties of declaratives and interrogatives: Penultimate Syllable MinFo, Final Syllable MinFo, Penultimate Syllable MaxFo, Final Syllable MaxFo, Penultimate Syllable MeanFo, Final Syllable MeanFo, Penultimate Syllable Duration, Final Syllable Duration, Penultimate Syllable Intensity, Final Syllable Intensity, Overall Minimum Pitch (across penultimate and final syllables), Overall Maximum Pitch (across penultimate and final syllables), and Mean Fo Difference Score (Final syllable MeanFo – Penultimate syllable MeanFo). PCA is a procedure analogous to factor analysis with orthogonal rotation; however, devoid of theory, PCA seeks to create latent scores that maximize the variance explained between the variables. Since PCA creates uncorrelated (orthogonal) scores, it can play an important role in reducing the dimensionality of many correlated variables that if modeled individually would suffer from multicollinearity. In this analysis, we used PCA so as to incorporate elements of all the variables into the eventual logistic regression model without concern for the multicollinearity of these related variables. The number of component scores was selected based on the scree plot of the eigenvalues. Four components were considered, comprising approximately 87% of the variance in the variables. Variable loadings on each component were examined to determine the make-up of each score. Component 1 was primarily a function of Final Syllable Duration. Component 2 was a moderate loading on all the variables excluding Mean Fo Difference Score, Penultimate Syllable Duration, Penultimate and Final Syllable Intensity. Component 3 was predominantly composed of Difference Score, and had moderate loadings on all variables excluding Overall Minimum Pitch, Overall Maximum Pitch, Final Syllable MinFo, Final Syllable Duration, Penultimate and Final Syllable Intensity. Component 4 was primarily a function of Penultimate Syllable Duration and had moderate loadings on Mean Fo Difference Score and Penultimate MinFo. See Table 7 for the four composite scores used in Analysis 2 (and see Table A1 in the 'Appendix' for the complete table of loadings).

Before presenting the linear classifier results that use the principal components (PCs), we first present results that verify that these principal

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

TABLE 7. *Loadings of individual factors for composites 1–4 for principal components analysis*

	Comp. 1	Comp. 2	Comp. 3	Comp. 4
Min	0·04	−0·23	−0·05	−0·13
Max	0·15	−0·42	0·00	0·08
Difference score	0·01	−0·05	0·52	0·30
PUMinfo	0·09	−0·23	−0·17	−0·31
FNMinfo	0·02	−0·28	0·07	0·04
PUMaxfo	0·08	−0·33	−0·38	−0·07
FNMaxfo	0·18	−0·41	0·27	0·12
PUDuration	−0·07	0·00	−0·48	0·84
FNDuration	0·93	0·33	−0·06	0·04
PUPitchmean	0·11	−0·34	−0·37	−0·15
FNPitchmean	0·19	−0·38	0·32	0·21
PUIntensity	0·01	−0·01	−0·01	0·00
FNIntensity	0·02	−0·01	0·01	0·00

components do indeed correlate with sentence type. We conducted two separate linear regressions using the R 'glm' function with the binomial link function. In all analyses the first independent variables were the first four PCs, and the dependent variable was a binary contrast of two sentence types. The analyses varied in terms of which two sentence types we analyzed: declarative vs. polar interrogative, or declarative vs. *wh*-question.

With the dependent variable contrasting declaratives with polar interrogatives (with declaratives as the reference level), PCs 2 and 3 contributed significantly to the model fit ($Z = -5{\cdot}25$, $p < {\cdot}0001$, and $Z = -3{\cdot}428$, $p = {\cdot}0006$, respectively) whereas PCs 1 and 4 did not ($Z = -0{\cdot}078$, $p = {\cdot}938$, and $Z = 1{\cdot}45$, $p = {\cdot}148$, respectively). Thus, the PCs that loaded primarily on duration measures did not contribute significantly to the model fit. Nevertheless, in the classification analysis we include all four PCs, since they represent the most significant dimensions of correlated variability across the utterances.

With the dependent variable contrasting declaratives with *wh*-questions (with declaratives as the reference level), none of the PCs were predictive (for PCs 1–4: Zs $= -0{\cdot}098$, $-1{\cdot}56$, $-0{\cdot}122$, and $-0{\cdot}057$, respectively; all ps n.s.). This result is not surprising given the results involving the individual measures in Analysis 1.

Potential confounds with measured variables

Because prosody is controlled by a variety of linguistic domains and levels – e.g. lexical, phonological, syntactic, pragmatic – we wanted to rule out some potential confounds that could be related to our prosodic measures (and the

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

PC variables we derived from them) as well as sentence type. Following the recommendation of an anonymous reviewer, we further coded each utterance for the number of syllables in the final word of the utterance, as well as the location of the nuclear pitch accent. Both variables can affect utterance-final prosodic features, and both could reasonably be assumed to vary with sentence type. The existence of such a confound would not necessarily affect conclusions about the availability of the prosodic cues for sentence-type discrimination, but it might affect the broader theoretical interpretation regarding what properties of utterances infants attend to and use in early language acquisition.

Number of syllables in final word

We added an additional variable to each utterance representing the number of syllables in the final word. In our sample, utterance-final words had one, two, or three syllables. The results are summarized in Table 8.

Pearson's chi-squared tests revealed a non-significant but marginal difference in the pattern of number of syllables comparing declaratives and polar interrogatives ($\chi^2(2) = 5{\cdot}0$, $p = {\cdot}082$), no difference between declaratives and *wh*-questions ($\chi^2(2) = 1{\cdot}64$, $p = .44$), and a significant difference between *wh*-questions and polar interrogatives ($\chi^2(2) = 10{\cdot}32$, $p = {\cdot}006$). Thus, the number of syllables in the utterance-final word does seem to correspond to sentence type to some degree, and could contribute to the correlation of our PCs with sentence type.

Location of nuclear pitch accent

We calculated nuclear pitch accent in each utterance in our sample by first locating the highest pitch peak, and coding location as penultimate syllable, final syllable, or elsewhere (since we had only marked the location of the final two syllables in Analysis 1). The results are summarized in Table 9.

Pearson's chi-squared tests revealed a significant difference in the pattern of highest pitch placement comparing declaratives and polar interrogatives ($\chi^2(2) = 35{\cdot}10$, $p < {\cdot}0001$), no difference between declaratives and *wh*-questions ($\chi^2(2) = 0{\cdot}16$, $p = {\cdot}92$), and a significant difference between *wh*-questions and polar interrogatives ($\chi^2(2) = 32{\cdot}08$, $p < {\cdot}0001$). Thus, the location of highest pitch, considering the whole utterance, is significantly correlated with sentence type.

Relationship of PCs to sentence type, controlling for confounds

Our analysis of the number of syllables in the utterance-final word and the placement of highest pitch peak showed that both variables correlated with

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

TABLE 8. *Counts of sentences with one, two, or three syllables in the utterance-final word, organized by sentence type and number of syllables*

	Number of syllables		
Utterance type	One	Two	Three
Declarative	78	32	1
Polar interr.	58	42	3
Wh-question	74	23	0

sentence type. Moreover, those variables patterned similarly with our measures in Analysis 1 and the results of our linear regression with PCs: they differed when contrasting declaratives and polar interrogatives as well as *wh*-questions and polar interrogatives, but not *wh*-questions and declaratives. Therefore, it is possible that these two simple variables (together or independently) are responsible for the relationship between sentence type and the prosodic measures (and composites) we reported so far. If so, that could impact empirical and theoretical investigations with respect to prosody and sentence type in early language acquisition.

We tested this in the following way. First, we conducted a logistic regression, as described in the previous section, but with number of syllables in the final word (NumSyll) and the placement of highest pitch peak (PitchPeak), as coded above, as predictors. We only analyzed a model with the dependent variable coding for declarative vs. polar interrogative, as this is the contrast that all our prior analyses indicated would be the most promising for distinguishing declaratives and interrogatives. Next, we added the first four PCs to the model, as a group, to see if the composite variables improved model fit. Finally, we removed NumSyll and PitchPeak from the model, to see if they made a significant independent contribution to the fit of the larger model.

Both number of syllables in the utterance-final word and highest pitch peak location indeed accounted for variance in sentence type when they were the only variables in the model (NumSyll: $Z = 2{\cdot}5$, $p = {\cdot}012$; PitchPeak: $Z = 4{\cdot}03$, $p < {\cdot}001$). However, adding the first four PCs significantly improved model fit ($\chi^2(4) = 26{\cdot}68$, $p < {\cdot}0001$). When the PCs were entered, NumSyll and PitchPeak no longer significantly predicted sentence type ($Z = 1{\cdot}58$, $p = {\cdot}114$; $Z = 1{\cdot}51$, $p = {\cdot}131$, respectively). Finally, removing NumSyll and PitchPeak from the full model did not significantly decrease model fit ($\chi^2(2) = 4{\cdot}35$, $p = {\cdot}114$). Thus, these two variables do not appear to account for any independent variance in sentence type, whereas the PC variables do. Thus, in the linear classifier we now present, we consider only the composite PC variables.

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

TABLE 9. *Counts of sentences with highest pitch peak in penultimate syllable, final syllable, or elsewhere, organized by sentence type and highest pitch peak location*

Utterance type	Pitch peak location		
	Penultimate	Final	Elsewhere
Sentence	24	29	58
Polar interr.	2	63	38
Wh-quest.	22	27	48

SENTENCE TYPE PREDICTION USING PRINCIPAL COMPONENTS

We used the first four principal components that we calculated in the previous section to predict sentence type in a logistic regression model, within a testing database that was a subset of our total data sample. We used half the declaratives ($N = 55$) and half the polar interrogatives ($N = 51$), selected randomly from all three speakers, to set the coefficients. The coefficients from this model were then applied to the remainder of the dataset, with the predicted probabilities of sentence type compared to the observed classification. For Declaratives versus Polar Interrogatives discrimination, this model showed predictive accuracy for approximately 67·5% of the observations. Of the errors, approximately 57% were incorrectly classifying polar interrogatives as statements. Table 10 shows the regression coefficients for the two logistic regressions. For the declarative vs. polar interrogative discrimination, the component that most strongly predicts a question is Component 3. Recall that Component 3 loads heavily on the pitch difference across syllables, as well as other pitch measures, but it also loads negatively on final syllable duration. Thus, in accord with what we would expect from the results of Analysis 1, pitch appears to be the primary determinant of discriminability, but shorter final syllables also predict polar interrogatives. However, it should be noted that the values of coefficients, and thus their respective influences on utterance categorization, do not differ substantially.

Without a specific theory about how the utterances would be used for further grammatical learning and how robust those learning procedures would be to error in sentence-type discrimination, it is difficult to assess the adequacy of 68% categorization accuracy. Using more components could add to predictive accuracy, although there are additional problems of overfitting as well as multicollinearity when using the raw variables themselves. The main contribution of this analysis, then, is as a benchmark for the kind of information sentence-final prosody provides.

For Declaratives ($N = 55$) versus *Wh*-questions ($N = 48$) discrimination, this model showed predictive accuracy for approximately 51% of the

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

TABLE 10. *Regression coefficients for the logistic regressions on principal components in Analysis 2. Coefficients are in logit units (log odds).*

Discrimination	Component 1	Component 2	Component 3	Component 4
Declaratives vs. polar interrogatives	−0·0016	−0·0065	0·0053	0·0042
Declaratives vs. *wh*-questions	−0·0005	−0·003	0·0002	−0·0033

observations. Thus, this model is basically at chance for categorizing declaratives and *wh*-questions. These results are not surprising given our analysis of the independent variables in Analysis 1, and the logistic regression we reported in the previous section.

These analyses show that a simple linear classifier could use prosodic information that is realized on the last two syllables to discriminate the majority of polar interrogatives and declaratives, at least to some degree. It should be noted that we tested a supervised classification model that was trained on labeled sentences in order to set the coefficient weights to optimize the categorization results. Thus, we do not propose this as a model of sentence discrimination. Indeed, if a learner knew the sentence type of a number of sentences a priori, we suspect that distributional information would be a much more robust cue to sentence type than prosodic information. Rather, in this analysis we show how well a simple linear classifier could perform when considering the kind of prosodic information we targeted.

GENERAL DISCUSSION

The results of this corpus study demonstrate that there is sufficient prosodic information to distinguish declaratives from polar interrogatives in infant-directed speech at above-chance levels, but not from *wh*-questions. Four of the five prosodic measures and derivatives – MeanF0, MaxF0, pitch excursion, and intensity – significantly differed between declaratives and polar interrogatives, while the durational difference between the two sentence types was a statistical trend. In contrast, intensity was the only cue that distinguished between declaratives and *wh*-questions, with *wh*-questions demonstrating lower intensity than declaratives, which demonstrated lower intensity than polar interrogatives. This provides preliminary evidence that infant-directed speech contours follow the same patterns as adult-directed speech contours (Hedberg *et al.*, 2004; Ladd, 2008). Specifically, declaratives and *wh*-questions end with a flat or falling pitch, and polar interrogatives end with a final rise. This suggests that while parents may exaggerate typical speech contours in daily speech to infants, they do not change the shape of the contour. However, further

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

research is needed to determine whether infants are using these prosodic cues and, if so, their relative value for sentence-type discrimination.

The logistic regression model also suggests a multi-cue approach to classifying sentence types. The regression model was intended to narrow down and evaluate what variables these components had in common, to determine which acoustic variables are best at classifying sentence types. However, no single cue turned out to be exclusive for classification purposes. Component 1 was predominated by Final Syllable Duration. Component 4 was primarily a function of Penultimate Syllable Duration and had moderate loadings on two additional variables. Component 3 was predominantly composed of the Pitch Difference Score, but had moderate loadings of six additional variables. Component 2 had moderate loadings of all but four variables. There was no overlap between the predominant variables in Components 1, 3, and 4, and little overlap in moderate loadings of variables either.

These results suggest that sentence-type classification would be most robust when considering a combination of prosodic cues. Further research is needed to determine how well the model performs relative to human ability. Combined with the corpus analysis, these results suggest that infants may rely on multiple prosodic cues to distinguish between sentence types.

Preliminary work evaluating infants' ability to make sentence-type distinctions based on prosody alone failed to find evidence that 7-month-olds use this information to distinguish between declaratives and polar interrogatives (Geffen, 2014). This is not to say that infants CANNOT use prosody in this way, at this age, but rather that they did not use prosody in that experiment. A potential confound in the study was the greater acoustic variability in declarative stimuli compared to interrogatives, for both familiarization and test items. Further research is needed to address this confound and to further address the question of when infants begin using prosodic information to make sentence-type distinctions.

One factor that showed more variability in declaratives than polar interrogatives was MaxF0. While there is a significant difference between polar interrogatives and declaratives, there is no significant difference in MaxF0 between the FINAL syllable of *wh*-questions and declaratives. However, the MaxF0 is not limited to the end of an utterance. In languages like Russian, which have relatively free word order and no auxiliary verbs (Rojina, 2004), pitch peak is an important perceptual cue for distinguishing between declaratives and polar interrogatives (Makarova, 2007; Svetozarova, 1998). Thus, the location of the pitch peak may be another cue infants could use to make sentence-type distinctions. Perhaps a more fine-grained analysis of pitch peak location than we performed in Analysis 2 would reveal more robust independent effects in English. This may be particularly important for distinguishing between polar

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

interrogatives and *wh*-questions, given that *wh*-questions are more prosodically similar to declaratives than polar interrogatives (Bartels, 1999; Hedberg *et al*., 2004; Ladd, 2008). Perhaps *wh*-questions carry their emphasis, and thus their pitch peak, on the *wh*-word. For example, infant-directed *wh*-questions in English show an earlier pitch peak and final F0 decline (Best *et al*., 1991). It would be fruitful in further research to investigate the informativeness of pitch peak location in greater detail.

Although sentence-final prosody is insufficient to distinguish between declaratives and *wh*-questions, the ability to distinguish between declaratives and polar interrogatives provides valuable information for making distinctions between other types of interrogatives. As previously discussed, polar interrogatives and *wh*-questions occur in roughly equal proportions in infant-directed speech (23% and 21%, respectively; Newport, 1977), so polar interrogatives are a significant part of learners' input. Moreover, the characteristic inversion of auxiliaries and subjects in polar interrogatives also occurs in *wh*-questions, after the *wh*-word (e.g. *What did you have for lunch?*). Infants could use prosodic information to differentiate polar interrogatives from declaratives, which would then allow them to notice the characteristic subject–auxiliary inversion pattern. They could then detect the same patterns in *wh*-questions and use the distributional similarity as a basis for distinguishing *wh*-questions from declaratives, and for treating *wh*-questions and polar interrogatives as similar. In this way, the learned associations between word order and prosodic patterns in polar interrogatives may allow infants to generalize to *wh*-questions, given their prosodic similarity to declaratives (Bartels, 1999; Hedberg *et al*., 2004).

As we mentioned earlier, recent work suggests that infants can use distributional cues to discriminate declaratives and polar interrogatives in the absence of prosodic information by age 1;0 (Geffen & Mintz, 2015). It remains to be seen whether infants can make similar distinctions between declaratives and *wh*-questions at the same age.

CONCLUSION

In summary, the current experiment provides evidence that there is sufficient prosodic information to distinguish declaratives from polar interrogatives with above-chance accuracy. Making this distinction could then, perhaps, provide a foundation for distinguishing *wh*-questions from declaratives on distributional grounds, as we have found that those sentence types are prosodically similar in infant-directed speech, as they are in adult-directed speech. While infants can make use of prosody to discriminate sentence types in some languages (e.g. Frota *et al*., 2014), it currently remains unclear when English-learning infants can do the same (see our discussion in the 'Introduction' regarding Soderstrom *et al*., 2011). This ability could

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

provide an important foundation for acquiring syntactic knowledge. Future research is needed to further evaluate the degree to which prosody is sufficient for distinguishing between declaratives and interrogatives in the absence of lexical/semantic information. Finally, by mapping the genesis and development of sentence discrimination abilities in typically developing children, these future studies could lead to the development of a method for assessing individual differences in early language perception for individuals at risk for language impairments.

REFERENCES

Bartels, C. (1999). *The intonation of English statements and questions: a compositional interpretation*. New York / London: Garland Publishing Inc.

Best, C., Levitt, A. & McRoberts, G. (1991). Examination of language-specific influences in infants' discrimination of prosodic categories. In *Proceedings of the 12th International Congress of Phonetic Sciences* (Vol. 4, 162–165). Aix-en-Provence: Universite de Provence Service des Publications.

Brent, M. & Siskind, J. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition* **81**, 31–44.

Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.

Dryer, M. S. (2011). Polar questions. In M. S. Dryer & M. Haspelmath (eds), *The world atlas of language structures online* (chapter 116). Munich: Max Plank Digital Library. Online: <http://wals.info/chapter/116> (last accessed 12 November 2013).

Face, T. (2007). The role of intonation cues in the perception of declaratives and absolute interrogatives in Castilian Spanish. *Estudios de fonética experimental* **16**, 185–225.

Fernald, A. (1984). The perceptual and affective salience of mothers' speech to infants. In L. Feagans, C. Garvey & R. Golinkoff (eds), *The origins and growth of communication*, 5–29. Norwood, NJ: Ablex.

Fernald, A. & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development* **10**, 279–93.

Fisher, C., Hall, D., Rakowitz, S. & Gleitman, L. (1994). When it is better to receive than to give: syntactic and conceptual constraints on vocabulary growth. *Lingua* **92**, 333–75.

Fisher, C. & Tokura, H. (1996). Acoustic cues to grammatical structure in infant directed speech: cross linguistic evidence. *Child Development* **67**, 3192–218.

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

Geffen, S. (2014). *When and how infants discriminate between declaratives and interrogatives*. Unpublished doctoral dissertation, University of Southern California.

Geffen, S. & Mintz, T. (2015). Can you believe it? 12-month-olds use word order to distinguish between declaratives and polar interrogatives. *Language Learning and Development* **11**, 270–84.

Gerken, L., Jusczyk, P. & Mandel, D. (1994). When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition* **51**, 237–65.

Gordon, M. (1999). The intonational structure of Chickasaw. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences 3* (pp. 1993–6). Berkeley: University of California.

Gósy, M. & Terken, J. (1994). Question marking in Hungarian: timing and height of pitch peaks. *Journal of Phonetics* **22**, 269–81.

Grabe, E. 2004. Intonational variation in urban dialects of English spoken in the British Isles. In P. Gilles & J. Peters (eds), *Regional variation in intonation* 9–31. Tubingen: Niemeyer.

Gussenhoven, C. & Chen, A. (2000). Universal and language-specific effects in the perception of question intonation. In B. Yuan, T. Huang, & X. Tang (Eds.), *6th International Conference on*

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

Spoken Language Processing 2000 Vol. 1 (pp. 91–94). Beijing: China Military Friendship Publish.
Hay, J. & Saffran, J. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy* **17**, 610–41.
Hedberg, N., Sosa, J. & Fadden, L. (2004). Meanings and configurations of questions in English. In *Proceedings of Speech Prosody 2004*, Nara, Japan 309–312.
Hirsh-Pasek, K., Kemler Nelson, D., Jusczyk, P., Cassidy, K., Druss, B. & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition* **26**, 269–86.
Jarman, E. & Cruttenden, A. (1976). Belfast intonation and the myth of the fall. *Journal of the International Phonetic Association* **6**, 4–12.
Jusczyk, P., Hirsh-Pasek, K., Kemler Nelson, D., Kennedy, L., Woodward, A. & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology* **24**, 252–93.
Kemler Nelson, D., Hirsh-Pasek, K., Jusczyk, P. & Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language* **16**, 55–68.
König, E. & Siemund, P. (2007). Speech act distinctions in grammar. In T. Shopen (ed.), *Language typology and syntactic description* 276–324. Cambridge: Cambridge University Press.
Ladd, D. (2008). *Intonational phonology*. Cambridge: Cambridge University Press.
Ladefoged, P. & Johnson, K. (2010). *A course in phonetics*, 6th ed. Boston, MA: Cengage.
Lieberman, P. (1967). *Intonation, perception, and language*. Cambridge, MA: MIT Press.
MacWhinney, B. (2000). *The CHILDES Project: tools for analyzing talk*, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.
Makarova, V. (2007). The effect of pitch peak alignment on sentence type identification in Russian. *Language and Speech* **50**, 385–422.
Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* **29**, 143–78.
Moon, C., Cooper, R. & Fifer, W. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, **16** 495–500.
Nazzi, T., Bertoncini, J. & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology* **24**, 756–66.
Nespor, M., Shukla, M. & Mehler, J. (2011). Stress-timed vs. syllable-timed languages. In M. van Oostendorp, C. J. Ewan, E. Hume & K. Rice (eds), *The Blackwell companion to phonology*, Vol. 2, 1147–59. Malden, MA / Oxford: Wiley-Blackwell.
Newport, E. (1977). Motherese: the speech of mothers to young children. In N. J. Castellan, D. B. Pisoni & G. R. Potts (eds), *Cognitive theory*, Vol. 2, Hillsdale, NJ: Lawrence Erlbaum Associates.
Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
Rojina, N. (2004). The acquisition of wh- questions in Russian. *Nordlyd* **32**, 68–87.
Šafářová, M. & Swerts, M. (2004). On recognition of declarative questions in English. In *Proceedings of Speech Prosody 2004*, Nara, Japan.
Slobin, D. & Bever, T. (1982). Children use canonical sentence schemas: a crosslinguistic study of word order and inflections. *Cognition* **12**, 229–65.
Soderstrom, M., Ko, E. & Nevzorova, U. (2011). It's a question? Infants attend differently to polar questions and declaratives. *Infant Behavior and Development* **34**, 107–10.
Svetozarova, N. (1998). Intonation in Russian. In D. Hirst & A. Di Cristo (eds), *Intonation systems: a survey of 20 languages*. Cambridge: Cambridge University Press.
van Heuven, V. & Haan, J. (2002). Temporal distribution of interrogativity markers in Dutch: a perceptual study. In C. Gussenhoven & N. Warner (eds), *Papers in laboratory phonology 7*, 61–86. Berlin: Mouton de Gruyter.
van Heuven, V. & van Zanten, E. (2005). Speech rate as a secondary prosodic characteristic of polarity questions in three languages. *Speech Communication* **47**, 87–99.
Vion, M. & Colas, A. (2006). Pitch cues for the recognition of yes – no questions in French. *Journal of Psycholinguistic Research* **35**, 427–45.
Xu, Y. (2012). ProsodyPro.praat. Online: <http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/>.

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

https://doi.org/10.1017/S0305000916000349 Published online by Cambridge University Press

Appendix

TABLE A1. *Full loadings of individual factors for principal components analysis*

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5	Comp. 6	Comp. 7	Comp. 8	Comp. 9	Comp. 10	Comp. 11	Comp. 12	Comp. 13
Min	0·04	−0·23	−0·05	−0·13	−0·59	−0·10	0·04	−0·10	0·51	0·17	0·52	−0·01	0·01
Max	0·15	−0·42	0·00	0·08	0·23	0·49	−0·08	0·04	0·46	0·37	−0·39	0·07	0·00
Difference score	0·01	−0·05	0·52	0·30	−0·19	0·07	−0·31	−0·11	−0·43	0·53	0·14	0·01	0·01
PUMinfo	0·09	−0·23	−0·17	−0·31	−0·26	0·00	0·54	−0·40	−0·38	0·24	−0·31	−0·01	0·00
FNMinfo	0·02	−0·28	0·07	0·04	−0·58	−0·09	−0·26	0·38	−0·09	−0·35	−0·48	0·01	−0·02
PUMaxfo	0·08	−0·33	−0·38	−0·07	0·09	0·23	0·07	0·59	−0·39	0·12	0·40	−0·02	0·01
FNMaxfo	0·18	−0·41	0·27	0·12	0·09	0·35	0·06	−0·36	−0·12	−0·60	0·27	−0·06	0·00
PUDuration	−0·07	0·00	−0·48	0·84	−0·14	−0·02	0·13	−0·16	0·00	−0·01	−0·03	0·00	0·01
FNDuration	0·93	0·33	−0·06	0·04	−0·09	0·01	−0·03	0·03	−0·01	0·01	−0·01	−0·01	0·00
PUPitchmean	0·11	−0·34	−0·37	−0·15	0·21	−0·40	−0·61	−0·35	−0·10	0·03	0·00	−0·02	0·01
FNPitchmean	0·19	−0·38	0·32	0·21	0·28	−0·64	0·36	0·22	0·09	0·05	−0·01	−0·03	−0·02
PUIntensity	0·01	−0·01	−0·01	0·00	0·00	−0·02	0·00	−0·02	−0·03	−0·02	0·05	0·67	−0·74
FNIntensity	0·02	−0·01	0·01	0·00	0·01	−0·03	0·02	0·00	−0·03	−0·05	0·03	0·74	0·67