Prosody in First Language Acquisition – Acquiring Intonation as a Tool to Organize Information in Conversation

Shari R. Speer* and Kiwako Ito
Ohio State University

Abstract

Recent research on children's acquisition of prosody, or the rhythm and melody in language, demonstrates that young children use prosody in their comprehension and production of utterances to a greater extent than was previously documented. Spoken language, structured by prosodic form, is the primary input on which the mental representations and processes that comprise language use are built. Understanding how children acquire prosody and develop the mapping between prosody and other aspects of language is crucial to any effort to model the role of prosody in the processing system. We focus on two aspects of prosody that have been shown to play a primary role in its use as an organizational device in human languages, prosodic phrasal grouping, and intonational prominence.

Introduction

Prosody, the rhythm and melody of spoken language, has been called the 'organizational structure of speech' (Beckman 1996). In psycholinguistic research, the term refers to patterns of timing, tune, and emphasis that are used to convey a wide range of information from speaker to hearer, including the speaker's affect, illocutionary force, linguistic pragmatic intent (such as emphatic or corrective contrast), and grammatical structure, such as the location of syntactic phrasal boundaries and word boundaries. Prosody provides 'a rhythmic scaffolding' (Arbisi–Kelm and Beckman forthcoming) that highlights important temporal locations in the speech stream, such as the location of words that convey central aspects of an utterance's message, and the locations where critical information about the phonological, syntactic, and semantic content are aligned in time. The production and comprehension of prosody is by necessity examined from a variety of disciplinary perspectives, including phonetics, phonology, speech perception, psycholinguistics, and neurolinguistics. Formal linguistic theories of autosegmental phonology and intonation are used to characterize the possible forms of both lexical and sentential prosody as they vary within and across languages. Acoustic phonetics is used to measure the physical

correlates of spoken prosody in the speech signal, most commonly by examining the shape and height of the fundamental frequency contour, the relative duration of words and silences, and local aspects of spectral information and intensity. Intonation annotation systems[1] are available for many languages, and are used to label intonational events such as local pitch prominences and the location and melody of phrases. Annotation is important to the description of prosody used in experiments, because it provides a description that is not dependent (as duration and fundamental frequency are) on a single spoken language event. For example, a low, phrase-final tone can be given a common annotation (such as the symbol L in ToBI) across utterances and speakers. In contrast, the absolute duration of an utterance-final word is highly dependent on factors such as word identity and the rate of speech, while fundamental frequency varies with such factors as gender, age, and emotional state.

A substantial literature documents the role of prosody in language acquisition. Infants acquire language from input that is almost entirely auditory, and have been shown to prefer the sound of their native language over others as early as 3 days of age, an effect attributed to their ability to recognize its prosodic form (Mehler et al. 1988; Christophe et al. 2001). The acquisition of native language is accelerated in the first 3 years of life, a time when the bulk of adult–child linguistic interaction is spoken, and thus primarily structured by prosody. Although prosody differs substantially across languages, a variety of language-specific rhythmic and/or tonal patterns have a demonstrated impact on the identification of speech sounds and the segmentation of the spoken signal into identifiable words, allowing early word learning (cf. Jusczyk et al. 1993; Morgan and Saffran 1995; Johnson and Jusczyk 2001).[2] As older children's knowledge of their native language becomes more sophisticated and adult-like, they must also eventually develop the use of the full complement of prosodic functions, including the ability to felicitously phrase multi-clausal sentences and to produce and respond to jokes and sarcasm. Some more complex structures are not fully mastered until the pre-teen years. [One such late-occurring skill is the distinction between compound minimal-stress pairs in English like 'HOTdog' (a sausage) vs. 'hot DOG' (a canine).] Although these aspects of prosodic acquisition are fascinating topics in their own right, our discussion in this article will not extend to the entire range of prosodic development. Instead, we hope to offer a current view of the role of two particular aspects of prosody in children's language acquisition, highlighting some more recent findings as well as some well-established results that provide a base for future research. Our focus includes:

(1) Intonational phrasing, or the 'chunking' of words together into interpretable units. Such prosodic phrases often correspond to discourse propositions, or syntactic phrases and clauses. For example, in English, prosodic phrase ends are marked by lengthening of the final word and the presence of phrasal tones such as a fall or fall-rise, while phrase beginnings are distinguished with initial

strengthening and or with a resetting of pitch range. In spoken conversation, children encounter many syntactically incomplete utterances and other partial forms. Thus, they must learn to use the correspondences between intonational phrasing and other types of linguistic units in order to know which words speakers intend to 'go together', and which they intend to separate.

(2) Intonational prominence, or the signaling of the relative importance or salience of discourse entities. Prominence also provides a coherent informational structure across the utterances in a conversation. To illustrate, again with an English example, when an interlocutor has been misunderstood, she can highlight the relevant portions of a correcting utterance by increasing their prominence with localized pitch peaks: 'I didn't say BEAT Tommy, I said MEET Tommy.'

Acquisition of Intonational Phrasing

Infants' sensitivity to prosody's rhythmic and melodic patterns is evident from the earliest days of life. Newborns (as early as 3 days old) can discriminate between two spoken languages on the basis of their prosody (Mehler et al. 1988; Jusczyk et al. 1993), and 6-month-olds use various aspects of prosody to determine the location of words in the stream of running speech (Jusczyk et al. 1993; Morgan and Saffran 1995; Morgan 1996; Johnson and Jusczyk 2001). Consistent with these early abilities, young children are demonstrably sensitive to correspondences among the acoustic aspects of prosodic phrasing when they listen to sentences, and they tend to pronounce their own utterances with appropriate affective and phrasal prosody.

Young infants have been repeatedly shown to perceive prosodic phrasing, and to prefer speech that conforms to the phrasing pattern of their native language. Preferences are most often established using sucking habituation and/or head-turn listening procedures. In sucking habituation, babies listen to a repeating sound while sucking on a non-nutritive nipple that contains a transducer that allows the researcher to measure the rate and duration of sucking. This rate drops off when babies lose interest in the repeating sound, and increases when they recognize a novel one. In head-turn procedures, babies learn an association between where they look and the sounds they hear. For example, they may learn that looking to a light or video on one side will cause a particular sound to play, while looking to a similar visual object on the other side of the room will cause a different sound to play. These procedures allow researchers to determine whether babies can distinguish between two sounds, and when they prefer to listen to one sound over another. Listening to words grouped together in a prosodic phrase can improve 2-month-olds' memory for the words, and they prefer listening to words in coherent prosodic phrases over words pronounced in a list intonation or words presented in incomplete adjacent prosodic phrases (Mandel et al. 1994, 1995). Infants at 4.5 months show sensitivity to whether prosodic phrases in the speech they hear are well-formed, and prefer to listen to passages with artificial pauses inserted to increase the length of silence at prosodic boundaries rather than to the

same passages with silences inserted in the middle of prosodic phrases (Jusczyk et al. 1995; see also Hirsh–Pasek et al. 1987). Although these studies document that infants are sensitive to the acoustic properties that should coincide for the boundaries in a spoken utterance, they do not show that prosodic phrases function to group words together into interpretable units. Additional evidence is needed to show whether and when very young listeners begin to use prosodic phrasing as a signal to interpret the elements of a phrase together in order to understand the sentence. Such a skill is necessary for development of the grammatical interpretation of speech.

Researchers have long been interested to see if young infants are sensitive to the correspondence between prosodic phrases and syntactic constituents, such as a noun phrase (NP) and verb phrase (VP) in a spoken sentence. Evidence about this ability in young infants remains mixed. In Jusczyk et al. (1992), 6-month-olds did not distinguish whether a prosodic boundary coincided with the end of a syntactic phrase or instead interrupted the syntactic phrase, while 9-month-olds preferred passages where the prosodic break divided the sentence into NP and VP (e.g., 'The small boy / wore a red jacket') over sentences where the major prosodic break was elsewhere in the sentence ('The small / boy wore a red jacket.'). However, in Nazzi et al. (2000), 6-month-olds better recognized a word sequence when it was presented as a coherent prosodic phrase in running speech than when it was formed from a word ending a phrase and a word beginning the next phrase, suggesting that the babies might have integrated phrase-internal material. Finally, Soderstrom et al. (2003) report that both 6- and 9-month-olds preferred to listen to the syllable strings forming coherent NPs (e.g., 'At the discount store, **new watches for men**') over the identical syllable strings forming 'syntactic non-units' (e.g., 'The old frightened **gnu # watches for men** and women . . .). Prosodic cues that differed between the strings were primarily in the tune rather than the timing. Thus, the results suggest melody-based integration of the phrase-internal material. (However, the results held only for spoken strings with relatively salient prosodic cues to breaks between NP and VP in the non-units – items with more subtle cues showed no differences).

Children's productions also show use of prosodic phrasal grouping. Snow (1994) provides a longitudinal observation of children's spontaneous speech as they began to produce multi-word utterances 16–25 months. He looked at whether declarative utterances ended with a falling pitch contour, and whether words that fell at the ends of prosodic phrasing showed lengthening of the final stressed syllable (as compared to other syllables in the utterances). He found that these two aspects of prosodic phrasing gained more consistency as the number of words in children's utterances increased, and argued that children are able to signal the end of a prosodic phrase by controlling the fundamental frequency contour and the phrase-final segmental duration by the age of two. He also suggested

that this control of prosodic phrasing corresponds to initial aspects of syntactic acquisition, such as acquisition of verb argument structures, since it appears around the same time in development.

In summary, prelinguistic infants and young children exhibit certain facility with and sensitivity to prosodic phrasal grouping in both comprehension and production. However, the precise timing and mechanisms of its acquisition have yet to be established. To determine how intonational cues contribute to the early parsing of spoken sentences, what type of intonational material is most influential, and whether such development may be language-dependent, cross-linguistic investigation is needed. One of the widely attested characteristics of infant-directed speech is the exaggerated pitch range, which is believed to assist speech segmentation in young infants (e.g., Fernald 1985; Fernald and Kuhl 1987; Fernald et al. 1989). Although investigation on language-specific boundary pitch phenomena has made substantial progress thanks to the development and use of annotation systems, the functions of language-specific boundary tunes in infant-directed speech are yet to be specified.

Even if young children use prosodic phrases to group together words that should be comprehended together as coherent units, they may not be using prosodic phrasing to recover syntactic relationships among those units. For adults, the correspondence between prosodic and syntactic phrase boundaries is an established and powerful factor in the resolution of syntactic ambiguity (cf. Kjelgaard and Speer 1999; Schafer et al. 2003; Snedeker and Trueswell 2004). For some sentence types, prosody is the only information that can resolve the meaning of a spoken sentence. A frequently studied example is the ambiguous attachment of a prepositional phrase in V-NP-PP sequences, as in *The wizard **zapped the witch with the wand***. When this sentence is pronounced as two prosodic phrases with a break following the verb, [*The wizard zapped*][*the witch with a wand*], listeners attach the PP as a modifier of the second NP *witch*, and understand that she had a wand. In contrast, when the break follows the second NP [*The wizard zapped the witch*][*with a wand*], listeners syntactically attach the PP as a sister to the verb, and recover a sentence meaning that the wizard had a wand. Given children's early sensitivity to prosodic phrasing, and especially if they use prosodic phrasal structure as a sort of 'proto-syntax' during development (see work on the controversial notion of 'prosodic bootstrapping', for example, Soderstrom et al. 2003), we would expect them to behave same way that adults do, and thus to recover the intended meaning of sentences like these when they are spoken with felicitous prosody. In addition, children should be able to produce prosody that indicates the meaning they intend to convey for syntactically ambiguous sentences.

However, evidence of children's use of prosody for parsing ambiguous sentences has been decidedly mixed. An early study (Beach et al. 1996) showed that 5- and 7-year-olds were capable of using prosody to choose between two meanings of the phrase '*pink and green and white*' in a way

similar to adults. Spatially grouped pictures of pink-, green- and white-colored animals were shown, while a re-synthesized male voice 'spoke' one of two prosodic phrasings, [[*pink and green*] *and white*] or [*pink and* [*green and white*]]. The relative durations and the fundamental frequency contour of the words were manipulated to create prosodic phrasing. Children correctly pointed to the row of colored animals that were spaced to match the prosodic phrasing. In addition, the longer the pauses were between phrases, the better the children performed. However, the design of this experiment gave the children every chance to succeed at their task. The same set of color terms were presented on every trial, and the pictures were always in the same order on the page as the color terms were in the phrase. Children were instructed about how the sound of language can be used to understand grouping, and were given feedback about the goodness of their answers. This teaching component to the task leaves us to wonder whether the experiment shows only that children can learn to match the grouping of words with a perfectly parallel grouping of pictures. That is, children could have done this task as though solving a puzzle rather than as though comprehending language. Indeed, a companion production study (Katz et al. 1996) suggested that children could not use prosody to produce an unambiguous description of groupings of colored blocks when instructed to say 'which blocks go together.' While adults used both word duration and pitch contour to convey the grouping of color terms, 5- and 7-year-old children used neither of these variables consistently.

Later researchers had difficulty generalizing this finding to other ambiguous sentence types and visual objects. Snedeker and Trueswell (2001, 2004) conducted a series of toy-moving studies investigating the V NP PP attachment ambiguity in English, using sentences such as '*Tap the frog with the flower*.' The task in these studies was more complex than that used by Beach et al., with participants hearing a new sentence and seeing new toys on each trial. An example set included five toys: (i) an instrument – for example, a large flower, (ii) a plain animal – for example, an empty-handed stuffed frog, (iii) the same animal with an instrument – for example, a frog holding a little flower, (iv) a different animal with another instrument – for example, a giraffe holding a little candle, and (v) another potential instrument. Children heard one of two different prosodic versions of the instruction, either [*Tap*][*the frog with the flower*] or [*Tap the frog*][*with the flower*.] If children could use prosodic cues to understand the sentences, they should have used their hands to tap the frog holding a flower when the prosodic boundary followed *tap*, but used the large flower instrument to tap the plain stuffed frog when the prosodic boundary falls between *frog* and *with*. In one experiment, the mothers of 4- and 5-year-olds gave the instructions to their children. Although the mothers were not instructed explicitly to use prosodic cues to disambiguate their instructions, they did so reliably for the instrumental utterances [*Tap the frog*][*with the flower*.], and less reliably so for the companion utterance. However, the

listening children did not respond differently to the two instruction types – they tended to choose one of the possible types of responding and stick with it, either using an instrument or their hand on the majority of trials, regardless of the prosody of the instruction they heard. This effect did not seem to be due to the children's inability to follow the instructions, as they succeeded in following unambiguous instructions with similar meaning, like '*Tap the frog that has the flower.*' Children were also successful in following instructions in a second study, where they heard instructions that did not include prosodic disambiguation, but instead had biased word combinations like '*Tickle the pig with the feather*' (instrument bias) or '*Choose the pig with the stick*' (modifier bias).

Choi and Mazuka (2003) suspected that researchers' failure to find that children can use prosody to understand syntactically ambiguous sentences might have to do with the complexity of the syntactic ambiguity encountered rather than with children's ability to use prosody to group together items in an utterance that should be interpreted together. They devised a clever pair of experiments to test the comprehension of two kinds of phrasing-based prosodic disambiguation in Korean 3- and 4-year-olds, a word segmentation ambiguity and a syntactic phrasal ambiguity. Phrasal grouping in Korean involves a low-to-high phrasing pattern within a phrase, and a silent duration between phrases, so that both types of prosodic information were available to help distinguish between the utterance types. In the word segmentation study, children heard sentences such as '[*Kipper-ka*][*pang-e tilegayo*]' (Kipper enters a room) or '[*Kipper*][*kapang-e tilegayo*]' (Kipper enters a bag), where the grouping of the syllables *ka* and *pang* determines the identity of the direct object noun, but does not change the syntax of the sentence. Twenty-one of 23 children made use of prosodic phrasing to understand the sentences and correctly choose between pictures of the cartoon character Kipper entering either a room or a bag (this was comparable performance to that of a comparison group of adult listeners in the same task). In contrast to the previous work by Beech et al., no feedback was given as to whether the child had chosen the correct picture (children were rewarded regardless of the correctness of their responses). Thus, children as young as age 3 were sensitive to prosody, and could use it to choose appropriately between two sentences that differed only in their phrasing. However, children in the syntactic ambiguity experiment were unable to use prosodic phrasal grouping in this same way. They heard sentences such as '[*Kirin*][*kwaja*][*megeyo*]' ['(A) giraffe eats (a) cookie'] or '[*Kirin kwaja*][*megeyo*]' ['Someone (Null subject) eats (a) giraffe-shaped cookie'], and were asked to choose the correct picture. Prosodic phrasing manipulations and picture illustrations were directly comparable to those used in the word segmentation experiment. With such syntactic ambiguities, 3- and 4-year-old children barely performed at chance level, and a comparison group of 5- and 6-year-olds were correct only 60% of the time – just slightly above chance, and well below the performance of adults.

Interestingly, an analysis of individual children showed a wide range of performance, with two of the 26 3- and 4-year-olds and one of the 21 older children performing perfectly in the task, but nine of the 3- and 4-year-olds and one of the older group with fewer than a quarter of the trials correct. These results indicate that preschoolers are sensitive to prosodic cues to word boundaries, but the use of prosodic phrasing to disambiguate syntactic form develops more gradually. The study cannot distinguish whether the locus of the difficulty was in the resolution of the syntactic ambiguities themselves, in the use of the mapping between prosodic and syntactic structure to resolve the ambiguity, or both. The wide range of children's abilities in the syntactic disambiguation study presents a further puzzle – what aspects of the experiments were responsible for such a wide range of success and failure to use the prosody/syntax correspondence?

A possible answer to this puzzle can be found in a recent study by Snedeker and Yuan (2008). Extending the previous toy-moving studies of (Snedeker and Trueswell above), they presented 4- to 6-year-old children with sentences like '*You can tap the frog with the flower*' in with two disambiguating prosodic phrasings.[3] The researchers had noticed an interesting pattern in the previous work – individual children seemed to favor a particular response type, with some preferring, for example, to use the large instrument to interact with the toy, while others preferred to use their hands. To investigate whether this perseveration was interfering with the experimental results, Snedeker and Yuan used a blocked experimental design where half of subjects heard the instrumental versions, for example, '[*You can tap*][*the frog with the flower*]' for the first half of the experiment and the modifier versions, for example, '[*You can tap the frog*][*with the flower*]' in the second half, and the other half of subjects heard the sentences in the opposite blocked order. Results showed that perseveration did indeed strongly influence the results – and also that young children were able to use prosody to select the intended syntactic meaning. Regardless of which prosody they heard in the first block, instrument, or modifier, children overwhelmingly used the location of the prosodic boundary to correctly interpret the syntax of the sentences. In the second block, however, children who first heard instrument sentences like [*You can tap the frog*][*with the flower*] continued to use the instrument to act on the unmodified toy, regardless of the prosody of the utterances they heard. In contrast, children who first heard modifier sentences shifted over to the instrumental actions. This pattern was found again in a second experiment, where lexical biases in the verb–instrument combinations were manipulated in addition to the location of the prosodic boundaries. Even in biased sentences where the prosodic boundaries were inconsistent with the semantic bias in the words, for example, [*You can choose the pig*][*with the stick*] (not a very likely verb–instrument combination), children continued to be able to use the prosodic phrasing of the sentence to interpret the

syntax 'correctly' and perform an instrumental action. Prosodic phrasing was also used to interpret modifier sentences appropriately in the first block, but less so in the second, confirming the bias toward the instrumental interpretation shown in the first experiment. Both experiments demonstrated that children can indeed use prosodic phrasing to disambiguate syntactic structure, and suggest that their failure to do so in at least some of the previous experiments may have been an artifact of experimental design (see also a similar finding with Japanese 5-year-olds by Mazuka and Tanaka 2006). Still, what could have caused the asymmetry between the two prosodies? One possibility in this instance is not a very interesting one – the researchers note that they needed to prevent children from using the small instrument to perform the action (e.g., they did not want children to use the small flower on the frog-with-the-flower to tap the empty-handed frog). Thus, they discouraged children from manipulating the miniature objects during demonstration and filler trials, but not on critical trials. It may have been that this discouragement made the children who were using the large instrument in the first block of trials less likely to switch over to touching the modified objects in the second trial block. Other researchers have also found that task variables have a strong influence on children's behavior in this type of toy-moving task. For example, Meroni and Crain (2003) showed that children performed better (made fewer syntactic errors) with sentences such as 'put the frog on the red napkin into the box' when they heard the sentence before they were allowed to turn around to view the objects to manipulate. It is possible that children's limited memory capacity, and less controlled inhibition, prompts them to execute actions incrementally, without waiting for a full interpretation of an instruction. Therefore, the above instrument bias may also have come from the availability of readily interactive objects.

In summary, future investigation of children's ability to make use of the correspondence between intonational phrasing and syntactic constituency during language acquisition must consider how young children react to the experimental task and environment, as well as to whether they are sensitive to prosodic contrasts. As researchers continue to develop more sophisticated means of observing and measuring children's ability to use prosodic structure to understand syntactic grouping, this work may be extended to a wider set of prosodic and syntactic forms, and to additional languages. Cross-linguistic work is necessary, if we are to understand which aspects of prosodic systems are most useful for acquisition, and so that we can begin to abstract across prosodic and syntactic forms to determine the basic mechanisms of language acquisition.

Acquisition of Intonational Prominence

Investigating children's pragmatic use of prosody is challenging, as its mastery is tightly linked to the general cognitive development involving

attention, memory, and other executive functions. (See Davidson et al. 2006 for multi-task investigation of the development of working memory, inhibition, and cognitive flexibility. Also see Mazuka et al. forthcoming for a review of the executive function development and its effect on language processing in children). In a conversation among adults, speakers seem to effortlessly express what part of a message is more important than the other parts by using proper sets of words in proper syntactic forms. In spoken communication, intonation accompanies such illocutionary acts and exhibits a large impact on how the listener interprets the message. Imagine an utterance 'It was Ross who contacted the embassy.' Despite the lack of preceding discourse context, the majority of listeners would understand that the speaker assumes that the listeners share the knowledge that somebody contacted the embassy, and that the speaker wants to emphasize that it was Ross and not any other person who completed the action. To highlight 'Ross' and to signal that 'contacted the embassy' is part of the background knowledge or the common belief, the speaker would utter 'Ross' with a strong pitch prominence and 'contacted' and 'embassy' with a much less salient pitch excursion. The prosody of the utterance allows the listener to interpret the relative importance of each word and represent the informational status of discourse entities such as 'Ross' and 'embassy' accordingly. Thus, intonation expresses the informational weight of utterance components, and shapes the focus (the primary information locally under discussion) of the utterance as a whole (Bolinger 1961; Halliday 1967; Pierrehumbert 1980; Gussenhoven 1983a,b, 1994; 1999; Selkirk 1984; Cruttenden 1986; Pierrehumbert and Hirschberg 1990; Needham 1990). As a discourse develops over utterances, the conversants must constantly update the informational status of the entities to which they refer – the discourse referents. To do this, they track what has been brought up as a topic and what has been emphasized, swiftly shift their attention to a new topic or focus when necessary, and suppress or maintain already-discussed issues to avoid redundancy as the conversation continues. Thus, interpreting intonational cues to the informational status of words requires substantial memory resources as well as high-level executive skills. Plausible tight links between memory, executive function, and language skills have been suggested mostly in the literature on second language acquisition and bilingualism. Several studies have shown that phonological memory may predict grammar and vocabulary development (e.g., Ellis 1996; Hu 2003; O'Brien et al. 2006; French and O'Brien 2008). In addition, bilinguals may develop better control of certain cognitive skills than monolinguals in tasks involving attention control and inhibition (e.g., Bialystok and Codd 1997; Bialystok 1999), and theory of mind (e.g., Goetz 2003). Meanwhile, investigation on how memory and executive function affect the use of prosodic prominence for online discourse comprehension is yet to be conducted. Nonetheless, testing the effect of intonation on discourse comprehension is relatively easy with adults, for

whom the experimenters can assume mature working memory and well-developed attention shifting and inhibition skills. In fact, numerous past studies have shown robust effects of intonation on discourse comprehension in adults with tasks, such as phoneme detection, discourse verification, and speeded utterance acceptability judgments (Bock and Mazzella 1983; Birch and Clifton 1995; Cutler 1976; Cutler and Foss 1977; Terken and Nooteboom 1987; Davidson 2001; Ito 2002).

While dynamic rhythm and expanded melodic range are major characteristics of infant/child-directed speech (I/CDS) across languages and cultures (Fernald 1984; Fernald and Kuhl 1987; Fernald et al. 1989; Lieven 1994; Mazuka et al. 2006), it is by no means easy to investigate the impact of such prosodic prominence on a child's discourse representation. We may be able to observe that prosodic prominence mediates the here-and-now communication between the caregiver and the infant, but it is much more challenging to detect when and how pre-linguistic infants might update the information status of a discourse referent according to the intonational cues. While numerous studies on the mother–infant interaction have shown the general effect of I/CDS on the mean length of utterances, turn-taking skill, and syntactic development (e.g., Furrow et al. 1979; Kaye and Charney 1981; Huttenlocher et al. 1991), relatively few report on the early development of the use of prosody to highlight the focus of an utterance. Many caregivers report anecdotes of their babies' surprisingly appropriate responses to their speech or of their babies' imitation of adult-like intonation, but unfortunately, scientific devices for interpreting the baby's mind behind such responses or productions are yet to be invented. The fundamental problems for monitoring intonation-driven informational updates in infants reside not simply in assessing the development of their memory and executive functions, but also in the complex nature of discourse structuring. That is, the prosodic cues that mark the relative importance of words are interpreted meaningfully only in the discourse context in which they are situated (Cruttenden 1985 calls this 'LOCAL meaning' as opposed to 'abstract meaning' that tones express), and it seems unfeasible to describe what infants represent as the common ground and how their representations change as the discourse continues. Even for adults, examining the structure of a natural discourse is somewhat speculative and highly laborious, as spontaneous conversation contains frequent alteration of topics and purposes (Ito and Speer 2006). It is all the more complex to detect how prosody affects discourse reference representation in young children.

Despite the general challenge in observing children's processing of prosody, recent research with both traditional head-turning method and progressive brain-imaging technique have provided some evidence for when infants begin to attend to language-specific prosodic patterns or intonational prominence. Schmitz et al. (2006) tested whether 4-, 6-, 8-, and 14-month-old German learning infants could distinguish between a non-canonical

focus prosody (with an accent on the argument *in situ*) from a default focus prosody (in German, the rightmost argument in the prosodic phrase is typically accented). Although the average orientation time was longer for the 'marked' prosody than for the default prosody in all age groups, only 8-month-olds showed a statistically significant difference in the comparison. Schmitz et al. claimed that infants develop sensitivity to the location of the accent in prosodic phrases by 8 months of age, but then learn that German allows *in situ* prosodic emphasis in the following 6 months and lose their sensitivity to the marked status of the *in situ* intonation. This hypothesis needs to be further tested cross-linguistically to establish when babies become sensitive to language-specific focus-marking prosody and when they start using this information to understand spoken messages.

In another recent study, ERP (event-related potentials) data showed differences in electrophysiological brain responses to native vs. non-native stress patterns in 4-month-old German and French infants (Friederici et al. 2007). In both language groups, non-native stress patterns (e.g., /baba:/ for German and /ba:ba/ for French) evoked larger positive mismatch responses than reversed stress patterns that are more frequent in their respective native languages. (The mismatch responses component was distributed across the comparable central frontal electrodes for both language groups.) If infants as young as 4 months old show sensitivity to language-specific lexical prosody, which may not always exhibit consistent phonetic clarity in running speech, it is not implausible that infants develop sensitivity to the language-specific focus prosody, which is generally accompanied by dynamic boost in intensity and pitch excursion, by 8 months of age. A question remains as to when such cues become meaningful to infants.

Although the traditional head-turning preference paradigm and the ERP technique are both applied to obtain online responses to prosodic input from pre-linguistic infants, such passive measures have limited value for the investigation of focus prosody, which is typically used in context-dependent interactions, such as conversation. Once children gain verbal communication skill, however, simple interactive tasks can be powerful tools for testing their use of intonation in reference resolution. For example, another recent study with young German-learning children (Grassmann and Tomasello 2007) indicates that by the age of 2, children are capable of interpreting the accented nouns as referring to novel objects. Novel toys and novel actions were used to test how prosodic prominence assisted word learning in 2-year-olds. While introducing a novel action with an accent on the novel verb did not lead to above-chance correct performances of the action, an accent on a novel noun did lead to the above-chance correct choices of novel toys. Grassmann and Tomasello argue that infants infer that adults want to direct their attention to a particular part of a scene with prosodic prominence, and this motivates the search for a novel

object upon hearing an accent. The lack of the accentual effect on verb learning is attributed to the task difficulty and the uncertainty as to whether infants were representing each novel action with or without the associating novel object. Grassmann and Tomasello recall a study by Naigles (1998) who reports that a novel word was interpreted as a novel action by 15-month-olds only when it appeared with familiar object nouns and was uttered in an intonational unit separating it from the nouns, and thus predict the successful learning of novel verbs with accents were it be tested with familiar objects. While the effect of accentual prominence on verb learning remains to be confirmed across languages, the universality in the general use of prosodic prominence for word learning is suggested by an earlier finding that caregivers produce prosodic prominence when referring to novel elements more consistently in CDS than in ADS (adult-directed speech; Fernald and Mazzie 1991). If CDS has abundant reliable prosodic cues to novelty across languages and cultures, and if young children can tune themselves to the association between the prosodic alteration and the novelty or predictability of events and objects, the dynamics in rhythm and pitch may play an essential role in their vocabulary growth.

In addition to novelty marking, accentual prominence is known to evoke contrastive interpretation of discourse entities (Bolinger 1961, 1983, 1985; Halliday 1967; Hornby and Hass 1970; Hornby 1971; Chafe 1974, 1976; Solan 1980; Cruttenden 1985, 1986; Pierrehumbert and Hirschberg 1990). For adults, studies have demonstrated that a prominent accent projects a contrastive relation between the accented discourse entity and its alternatives, and thus the information related to the evoked alternatives becomes more accessible. For example, detection of a phoneme (e.g., /k/) is facilitated at the contrastive locus (e.g., 'LAURIE doesn't have a dog. **Kathy** has a dog.') as compared to non–contrastive locus (e.g., 'Laurie doesn't have a DOG. **Kathy** has a dog.' Davidson 2001). Also, brief discourses are judged as acceptable faster when the accentual locus of the target utterance properly corresponds to the contrast locus elicited by the preceding utterance (e.g., 'DORIS didn't fix the radio. → **ARNOLD** fixed the radio. vs. Arnold **FIXED** the radio.' Bock and Mazzella 1983. See also Birch and Clifton 1995, 2002; Ito 2002). Studies with eye-tracking methodology have also demonstrated the immediate effect of intonational cues on reference resolution in adults. For example, Dahan et al. (2002) showed that a prominent accent on a noun leads to immediate fixations to a previously unmentioned visual object, whereas the lack of accent prompts fixations to an already-mentioned object. Later studies have shown that a prominent accent on a prenominal modifier leads to anticipatory fixations to contrastive object in English (Ito and Speer 2008), as well as in German (Weber et al. 2006).

While the impact of contrastive intonation on discourse interpretation and reference resolution is clearly established for adults, much less is

known about the process by which children develop such skills. In fact, a perplexing paradox has been identified for children's production and comprehension of contrast-marking intonation. That is, children who seem to have no difficulty in producing contrastive intonation in their utterances often fail to demonstrate correct interpretation of contrastive prosodic cues in simple oral interactions. Numerous studies report that preschoolers can produce prominent accent properly to express contrast (Hornby and Hass 1970; Hornby 1971; Wieman 1976; Macwhinney and Bates 1978; Culter and Swinney 1987; Wells et al. 2004), while some also detected that preschoolers and even older children often perform much worse than adults in simple comprehension tasks after listening to an utterance involving contrastive intonation (Solan 1980; Cruttenden 1985; Culter and Swinney 1987; Wells et al. 2004). In an early study of this phenomenon, Cruttenden tested 10-year-olds' understanding of focus prosody using a picture-matching task, where participants listened to an utterance and selected a picture that best matched it. He found that children performed significantly worse than adults in choosing a correct picture out of the three options for utterances, such as 'John's got FOUR oranges.' vs. 'John's got four ORANGES.' [Three pictures included either (i) a boy with four oranges and a girl with two oranges, (ii) a boy with four oranges and a girl with four bananas, and (iii) a boy with three oranges and a girl with four oranges.] Most recently, Wells et al. reported that British English speaking children aged between 5 and 13 years all performed above chance when producing a prominent accent on the color modifier to request matching pictures [e.g., 'I want a WHITE bicycle (instead of black one)']. However, when children heard an utterance with contrastive intonation (e.g., 'I wanted CHOCOLATE and honey'), their choices of correct picture when asked to point to a picture of the object that the speaker did not receive varied gradually with age (from a chance level for 5-year-olds to above 90% accuracy for 13-year-olds).

The relatively poor comprehension of contrast-marking prosody by children is somewhat puzzling given the level of appropriate use in their own intonation. Most of the above-mentioned studies confirm that children have interpretation of affect-related prosody, and Nakassis and Snedeker (2002) report that 6-year-olds are sensitive even to ironic prosody. Thus, it is rather inconceivable that preschoolers and older children are not capable of processing prosodic prominence that express focus or contrast in a discourse. We suspect that the past findings of children's inaccurate interpretation of prosody may have been the artifact of the context-free picture selection task. Note that in both Cruttenden and Wells et al., children had to select a picture after listening to an isolated utterance. Without an explicit discourse context that specifies the relations among discourse references, participants had to covertly grasp the contrastive relations among the referents within the visual input and map the information from the auditory input onto them. In such tasks, the observed

responses may reflect the children's inability to detect the contrastive relations in the visual stimuli or to link the linguistic input to the visual referent, rather than their inability to use contrastive intonation to correctly represent the status of referents in a natural discourse.

A recent eye-tracking study with Japanese children suggests that discourse context is crucial for the proper use of contrastive intonation in reference resolution. Ito et al. (2007, 2008, forthcoming) monitored eye movements of 6-year-old and adult Japanese listeners, while they engaged in simple object-detection tasks. In the first experiment, target displays contained a twin (e.g., green cat & pink cat), a competitor that shared the color with one of the twin (e.g., green monkey), and a distracter in another color (e.g., orange turtle). In their pilot study, participants were asked to name each animal as soon as each display appeared. Once all animals were named, they listened to the question such as '*MIDORI-no/midori-no neko-wa doko?*' (Where is the GREEN/green cat?) and identified the location of the questioned animal. With this procedure involving no discourse context, neither adults nor children exhibited the predicted facilitative effect of contrastive intonation in their fixation patterns. Ito et al. suspected that the questions with pitch prominence might have sounded 'out of the blue', instead of like a continuation of the naming task. They replaced the naming task with a prompt leading to a contrast (e.g., *pinku-no neko-wa doko?* 'Where is the pink cat?'). For adults, this change in the procedure led to a faster increase in fixations to the target with the pitch prominence than without, replicating the effect of contrastive accent in English (Ito and Speer 2008) and in German (Weber et al. 2006). As fixation biases were identified across the simple quadrant displays in both adults and children, Ito et al. used more complex displays for the second experiment, mimicking the visual environment in Ito and Speer (2008) where the informativeness of color adjectives was controlled. Using the same discourse structures (e.g., pink cat → green cat), they found the facilitative effect of contrastive intonation in both adults and children. Therefore, 6-year-olds could demonstrate the online use of contrastive intonation when the preceding discourse context licensed the use of contrastive intonation in the following utterance and when the visual stimuli were free of biases. The importance of discourse context is also supported by the findings of Culter and Swinney (1987), who found that children younger than 6 years could not detect accented words any faster than unaccented words when they were embedded in sentences presented in isolation, while the same age groups exhibited the advantage for accented words when they were presented in stories.

In summary, future investigation of the role of novelty- or contrast-marking prosodic prominence in language acquisition must consider how young children represent the discourse environment in question given their developmental level of memory and attention control. Also, experimental tasks and measures must be carefully designed to situate the prosodic tools

to be readily available for meaningful communication. With the participants' memory size and attention control in mind, researchers would be able to better hypothesize about the representational statuses of discourse entities and how they are updated or modified due to prosodic prominence during dynamic discourse progression. Failure to consider participants' cognitive capacity and possible misanalyses of discourse structure may lead to misinterpretation of data and false conclusions, as represented by the past understanding of preschoolers' comprehension of contrastive accents.

Of course, research on the pragmatic use of prosody is by no means limited to the topics mentioned above. The overall pitch range and the utterance-final tonal movements both convey affect and turn-taking cues (Kaye and Charney 1981; Wells et al. 2004; Salerni et al. 2007), which are uncontrovertibly important for the development of general communication skills. Research is progressing on the development of prosody and its impact on the acquisition of communication skills in autistic and mentally retarded children (e.g., Peppe et al. 2003, 2006). Evidence from both clinical and non-clinical work will lead us to better understand the relationship between the general cognitive skills and the processing of prosodic information during natural discourse. In the pursuit of such directions for future research, we should address questions from at least two perspectives. First, we need to advance our knowledge about what executive functions and memory capacities are required to develop the notion of 'focus' or 'contrast' in oral communication. Second, we should also test how prosody assists memory (in both adults and children) and how it influences the development of other cognitive functions necessary for successful oral communication, such as the ability to shift attention and inhibit spurious analyses.

Conclusion

The structure and function of prosody in the comprehension and production of language is an area of research that has drawn increasing attention from researchers in the many fields that study spoken communication. An understanding of the child's early acquisition of prosodic competence is critical to this effort. We have focused in this article on the acquisition of two aspects of prosody that have been shown to play a primary role in its use as an organizational device in human languages, intonational phrasing, and intonational prominence. We suggest that children's mastery of these two basic components of prosodic processing may be the basis for later acquisition of more complex interpretive and expressive uses of intonation. In reviewing the progress of recent research in both areas, we have noted the importance of cross-linguistic contributions to increase the generality of claims for the basic functions of intonation in language acquisition and processing, and to increase the specificity of claims about the particular acoustic cues to which young children respond as they learn to use spoken

language. Finally, we emphasize the importance of the structure of tasks used to evaluate children's use of prosody, particularly the referential structure of the context in which children are tested. As experimental techniques continue to be refined over time, we may continue to find evidence for children's surprisingly sophisticated use of intonation at younger and younger ages.

Short Biographies

Kiwako Ito is Senior Researcher in the Department of Linguistics at Ohio State University and Visiting Researcher at the Laboratory for Language Development at the Riken Brain Science Institute, Japan. She received her PhD in Linguistics from the University of Illinois at Urbana–Champaign, and works in the area of perception and production of intonation in discourse. She has co-authored papers in this area for *Journal of Memory and Language*, *Proceedings of the Annual Conference on Speech Prosody*, and *Proceedings of the International Congress of Phonetic Sciences*. She is currently writing a manuscript on the processing of lexical pitch accent vs. segmental mismatches in Japanese using ERP methodology with colleagues at Riken, and a manuscript on lexical semantics, visual context, and contrastive interpretation of accentual prominence with Shari R. Speer.

Shari Speer is Professor of Linguistics at Ohio State University. She received her PhD in Human Experimental Psychology from the University of Texas at Austin, and works in the area of perception and production of prosodic structure and its interface with other levels of linguistic structure. She has co-authored papers in this area for *Journal of Memory and Language*, *Language and Speech*, Procee*dings of the Annual Conference on Speech Prosody*, and *Proceedings of the International Congress of Phonetic Sciences*. She is currently at work on the manuscript 'Situationally Independent Prosody' with Paul Warren and Amy Schafer, and a manuscript on lexical semantics, visual context, and contrastive interpretation of accentual prominence with Kiwako Ito.

Notes

* Correspondence address: Shari Speer, Department of Linguistics, The Ohio State University, 1712 Neil Ave, Oxley Hall 222, Columbus, OH 43210, USA. E-mail: speer@ling.ohio-state.edu.

[1] Annotation systems currently in use include: ToBI [Tone and Break Indices, Beckman and Ayers 1997; Beckman et al. 2005. See Jun's volume (2005) for extension to German, Greek, Japanese, Korean, Mandarin, etc.], ToDI (Gussenhoven et al. 2003), RaP (Rhythm and Pitch: Breen et al. 2006), and IViE (Intonational Variation in English; Grabe 2004; Grabe et al. 2005).
[2] We focus on acquisition of prosody in spoken language for the purposes of this brief article. However, we note that a more accurate term to use here might be 'articulated language', as the world's sign languages also include prosodic structure. We distinguish the processing of articulated languages from text processing.
[3] This slightly longer version of the prepositional phrase ambiguity has the advantage of having balanced prosodic phrasing in both disambiguated forms, '[*You can tap*][*the frog with the flower*]' and '[*You can tap the frog*][*with the flower*]' (as compared to [*Tap*][*the frog with the flower.*]).

Works Cited

Arbisi-Kelm, T., and M. E. Beckman. forthcoming. Prosodic structure and consonant development across languages. Interactions in phonetics and phonology, ed. by M. Vigario, S. Frota, M. J. Freitas.

Beach, C., W. F. Katz, and A. Skowronski. 1996. Children's processing of prosodic cues for phrasal interpretation. Journal of the Acoustical Society of America 99(2).1148–60.

Beckman, M. E. 1996. The parsing of prosody. Language and Cognitive Processes 11.17–67.

Beckman, M. E., and G. M. Ayers. 1997. Guidelines for ToBI labelling, vers 3.0 [manuscript]: Ohio State University.

Beckman, M., J. Hirschberg, S. Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. Prosodic typology: the phonology of intonation and phrasing, ed. by S. Jun. Oxford, Oxford University Press.

Bialystok, E. 1999. Cognitive complexity and attentional control in the bilingual mind. Child Development 70.636–644.

Bialystok, E., and J. Codd. 1997. Cardinal limits: evidence from language awareness and bilingualism for developing concepts of number. Cognitive Development 12.85–106.

Birch, S. L., and C. Clifton, Jr. 1995. Focus, accent, and argument structure: effects on language comprehension. Language and Speech 38.365–91.

——. 2002. Effects of varying focus and accenting of adjuncts on the comprehension of utterances. Journal of Memory and Language 47.571–88.

Bock, J. K., and J. R. Mazzella. 1983. Intonational marking of given and new information: Some consequences for comprehension. Memory & Cognition 11.64–76.

Bolinger, D. L. 1961. Contrastive accent and contrastive stress. Language 37.83–96.

——. 1983. Two views of accent. Journal of Linguistics 21.79–123.

——. 1985. Intonation and its parts. London: Edward Arnold.

Breen, M., L. Dilley, E. Gibson, M. Bolivar, and J. Kraemer. 2006. Advances in prosodic annotation: a test of inter-coder reliability for the (RaP Rhythm and Pitch) and ToBI (Tones and Break Indices) transcription systems. New York, NY: 19th Annual CUNY Conference on Human Sentence Processing.

Chafe, W. 1974. Language and consciousness. Language. 50.111–33.

——. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. Subject and topic, ed. by C. Li, 122–31. New York, NY: Academic Press.

Choi, Y., and R. Mazuka. 2003. Young children's use of prosody in sentence parsing. Journal of Psycholinguistic Reasearch 32(2).197–217.

Christophe, A., J. Mehler, and N. Sebastian-Galles. 2001. Perception of prosodic boundary correlates by newborn infants. Infancy 2.385–94.

Cruttenden, A. 1985. Intonation comprehension in ten-year-olds. Journal of Child Language 12.643–61.

——. 1986. Intonation. New York, NY: Cambridge University Press.

Cutler, A. 1976. Phoneme-monitoring reaction time s a function of preceding intonation contour. Perception and Psychophysics 20.55–60.

Culter, A., and David A. Swinney. 1987. Prosody and the development of comprehension. Journal of Child Language 14.145–67.

Cutler, A., and D. J. Foss. 1977. On the role of sentence stress in sentence processing. Language and Speech 20.1–10.

Dahan, D., M. K. Tanenhaus, and C. G. Chambers. 2002. Accent and reference resolution in spoken-language comprehension. Journal of Memory and Language 47.292–314.

Davidson, D. 2001. Association with focus in denials. PhD dissertation. Michigan State University.

Davidson, M. C., D. Amso, L. C. Anderson, and A. Diamond. 2006. Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition and task switching. Neuropsychologia 44.2037–78.

Ellis, Nick. 1996. Sequencing in SLA: phonological memory, chunking and points of order. Studies in Second Language Acquisition 18.91–126.

Fernald, A. 1985. Four-month-old infants prefer to listen to motherese. Infant Behavior & Development 8.181–95.

Fernald. A., and P. K. Kuhl. 1987. Acoustic determinants of infant preference for motherese speech. Infant Behavior and Development 10.279–293.

Fernald, A., and C. Mazzie. 1991. Prosody and Focus in Speech to Infants and Adults. Developmental Psychology 27(2).209–21.

Fernald, A., T. Taeschner, J. Dunn, M. Papousek, B. de Boysson-Bardies, and I. Fukuki. 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. Journal of Child Language 16.477–501.

French, L. M., and I. O'Brien. 2008. Phonological memory and children's second language grammar learning. Applied Psycholinguistics 29.463–87.

Friederici, A., M. Friedrich, and A. Christophe. 2007. Brain responses in 4-month-old infants are already language specific. Current Biology 17.1208–11.

Furrow, D., K. Nelson, and H. Benedict. 1979. Mothers speech to children and syntactic development: Some simple relationships. Journal of Child Language 6.423–42.

Goetz, P. J. 2003. The effects of bilingualism on theory of mind development. Bilingualism: Language and Cognition 6.1–15.

Grabe, E. 2004. Intonational variation in urban dialects of English spoken in the British Isles. Regional variation in intonation, ed. by P. Gilles and J. Peters, 9–31. Tuebingen, Niemeyer: Linguistische Arbeiten.

Grabe, E., G. Kochanski, and J. Coleman. 2005. The intonation of native accent varieties in the British Isles – potential for miscommunication? English pronunciation models: a changing scene, ed. by Katarzyna Dziubalska-Kołaczyk and Joanna Przedlacka, 311–37. Bern, Switzerland: Peter Lang, Linguistic Insights Series.

Grassmann, S., and M. Tomasello. 2007. Two-year-olds use primary sentence accent to learn new words. Journal of Child Language 34.677–87.

Gussenhoven, C. 1994. Focus and sentence accents in English. Focus and natural language processing, Vol. 3, ed. by P. Bosch and R. van der Sandt, 83–92. Heidelberg, Germany: IBM Deutschland.

——. 1999. On the limits of focus projection in English. Focus: linguistic, cognitive, and computational perspectives, ed. by P. Bosch and R. van der Sandt, 43–55. Cambridge, UK: Cambridge University Press.

Gussenhoven, C., T. Rietveld, J. Kerkhoff, and J. M. B. Terken. 2003. ToDI: transcription of Dutch intonation, 2nd edn. <todi.let.kun.nl/todi/home.htm>.

Gussenhoven, C. G. 1983a. Focus, mode, and the nucleus. Journal of Linguistics 19.377–417.

——. 1983b. A semantic analysis of the nuclear tones of English. Bloomington, IN: Indiana University Linguistics Club.

Halliday, M. A. K. 1967. Notes on transitivity and theme in English, part 2. Journal of Linguistics 3.199–244.

Hirsh-Pasek, K., D. G. Kemler-Nelson, P. W. Jusczyk, K. Wright-Cassidy, B. Druss, and L. Kennedy. 1987. Clauses are perceptual units for young infants. Cognition 26.269–86.

Hornby, P. A. 1971. Surface structure and topic-comment distinction: a developmental study. Child Development 42.1975–88.

Hornby, P. A., and W. A. Hass. 1970. Use of contrastive stress by preschool children. Journal of Speech and Hearing Research 13.395–9.

Hu, C. F. 2003. Phonological memory, phonological awareness, and foreign language word learning. Language Learning 53.429–462.

Huttenlocher, J., W. Haight, A. Bryk, M. Seltzer, and T. Lyons. 1991. Early vocabulary growth: relation to language input and gender. Developmental Psychology 27(2).236–48.

Ito, K. 2002. The interaction of focus and lexical pitch accent in speech production and dialogue comprehension: evidence from Japanese and Basque. PhD dissertation, University of Illinois at Urbana-Champaign.

Ito, K., and S. R. Speer. 2006. Using interactive tasks to elicit natural dialogue. Methods in empirical prosody research, ed. by P. Augurzky and D. Lenertova, 229–57. Leipzig, Germany: Mouton de Gruyter.

——. 2008. Anticipatory effect of intonation: eye movements during instructed visual search. Journal of Memory & Language 58.541–73.

Ito, K., N. Jincho, R. Mazuka, N. Yamane, and U. Minai. 2007. Effect of contrastive intonation

in discourse comprehension in Japanese: An eye tracking study with adults and 6-yr olds. San Diego, CA: Poster presented to the 20th Annual CUNY Conference.

Ito, K., N. Jincho, U. Minai, N. Yamane, and R. Mazuka. 2008. Use of emphatic intonation for contrast resolution in Japanese: adults vs. 6-year olds. Chapel Hill, NC: Poster presented to the 21th Annual CUNY Conference.

——. forthcoming. Intonation facilitates contrast resolution: evidence from Japanese adults and 6-year olds.

Johnson, E., and P. W. Jusczyk. 2001. Word segmentation by 8 month olds: when speech cues count more than statistics. Journal of Memory & Language 44.548–67.

Jun, Sun-Ah (Editor). 2005. Prosodic typology: the phonology of intonation and phrasing. Oxford, UK: Oxford University Press.

Jusczyk, P. W., A. Frederici, J. M. Wessels, V. Y. Svenkerud, and A. M. Jusczyk. 1993. Infants' sensitivity to the sound patterns of native language words. Journal of Memory & Language 32.402–20.

Jusczyk, P. W., K. Hirsh-Pasek, D. Kemler-Nelson, L. Kennedy, A. Woodward, and J. Piwoz. 1992. Perception of acoustic correlates of major phrasal units by young infants. Cognitive Psychology 24.252–93.

Jusczyk, P. W., E. Hohne, and D. Mandel. 1995. Picking up regularities in the sound structure of the native language. Speech perception and linguistice experience: theoretical and methodological issues in cross-language speech research, ed. by W. Strange, pp. 91–119. Timonium, MD: York Press.

Katz, W. F., C. M. Beach, K. Jenouri, and S. Verma. 1996. Duration and fundamental frequency correlates of phrase boundaries in productions by children and adults. Journal of the Acoustical Society of America 99.3179–3191.

Kaye, K., and R. Charney. 1981. Conversational asymmetry between mothers and children. Journal of Child Language 8.35–49.

Kjelgaard, M., and S. R. Speer. 1999. Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. Journal of Memory and Langage 40.153–94.

Lieven, E. 1994. Crosslinguistic and crosscultural aspects of language addressed to children. Input and interaction in language acquisition, ed. by C. Gallaway and B. J. Richards, 56–73. Cambridge, UK: Cambridge University Press.

MacWhinney, B., and E. Bates. 1978. Sentential devices for conveying givenness and newness: a cross-cultural development study. Journal of Verbal Learning and Verbal Behavior 17.539–58.

Mandel, D. R., P. Jusczyk, and D. Nelson. 1994. Does sentential prosody help infants organize and remember speech information? Cognition 53.155–80.

Mandel, D. R., P. Jusczyk, and D. Pisoni. 1995. Infants' recognition of the sound patterns of their own names. Psychological Science 6.315–318.

Mazuka, R., Y. Igarashi, and K. Nishikawa. 2006. Input for learning Japanese: RIKEN Japanese mother–infant conversation corpus. The Institute of Electronics, Information and Communication Engineers, Technical Report/, TL2006-16 (2006-07).11–5.

Mazuka, R., N. Jincho, and H. Oishi. forthcoming. Development of executive control and language processing. Language and Linguistics Compass.

Mazuka, R. and Y. Tanaka. 2006. Children can be better than adults at using prosody to resolve syntactic ambiguity. Poster presented at the 19th Annual CUNY Conference on Human Sentence Processing, NY, NY.

Mehler, J., P. W. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini, and C. Amiel-Tison. 1988. A precursor of language acquisition in young infants. Cognition 29.143–78.

Meroni, L. and S. Crain. 2003. How children avoid kindergarten paths. Proceedings of 4th Tokyo Conference on Psycholinguistics. Hitsuji Shobo. Tokyo, Japan.

Morgan, J. L. 1996. A rhythmic bias in preverbal speech segmentation. Journal of Memory & Language 35.666–688.

Morgan, J. L., and J. R. Saffran. 1995. Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. Child Development 66.911–36.

Naigles, L. 1998. Developmental changes in the use of structure in verb learning. Advances in infancy research, Vol 12, ed. by C. Rovee-Collier, L. Lipsitt and H. Haynes, 298–318. London, UK: Ablex.

Nakassis, C., and Snedeker, J. 2002. Beyond sarcasm: Intonation and context as relational cues in children's recognition of irony. Proceedings of the Twenty-Sixth Boston University Conference on Language Development, ed. by A. Greenhill, M. Hughs, H. Littlefield and H. Walsh. Somerville, MA: Cascadilla Press.

Nazzi, T., D. G. Kemler-Nelson, P. W. Juczyk, and A. M. Jusczyk. 2000. Six-month olds' detection of clauses embedded in continuous speech: effects of prosodic well-formedness. Infancy 1.123–147.

Needham, W. P. 1990. Semantic structure, information structure, and intonation in discourse production. Journal of Memory and Language 29.455–68.

O'Brien, I., N. Segalowitz, J. Collientine, and B. Freed. 2006. Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learner. Applied Psycholinguistics 27.377–402.

Peppé, S. and J. McCann. 2003. Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. Clinical Linguistics & Phonetics 17.345–354.

Peppé, S., J. McCann, F. Gibbon, A. O'Hare, and M. Rutherford. 2006. Assessing prosodic and pragmatic ability in children with high-functioning autism. Journal of Pragmatics 38.1776–1791.

Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. PhD dissertation. Massachusetts Institute of Technology.

Pierrehumbert, J., and J. Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. Intentions in communication, ed. by P. Cohen, J. Morgan and M. E. Pollock, 342–365. Cambridge, MA: MIT Press.

Salerni, N., C. Suttora, and L. D'Odorico. 2007. A comparison of characteristics of early communication exchanges in mother–preterm and mother–full-term infant dyads. First Language 27.329–346.

Schafer, A. J., S. R. Speer, and P. Warren. 2003. Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. World situated language use: psycholinguistic, linguistic and computational perspectives on bridging the product and action tradition, ed. by M. Tanenhaus and J. Trueswell, 209–25. Cambridge, MA: MIT Press.

Schmitz, M., B. Höhle, A. Müller, and J. Weissenborn. 2006. The recognition of the prosodic focus position in german–learning infants from 4 to 14 months. Interdisciplinary studies on information structure 5, ed. by Schmitz Ishihara, 187–208. Potsdam, Germany: Potsdam University.

Selkirk, E. O. 1984. Phonology and syntax: the relation between sound and structure. Cambridge, MA: MIT Press.

Snedeker, J., and J. Trueswell. 2001. Unheeded cues: prosody and syntactic ambiguity in mother-child communication. Paper presented at the 26th Boston University Conference on Language Development.

——. 2004. The developing constraints on parsing decisions: the role of lexical biases and referential scenes in child and adult sentence processing. Cognitive Psychology 49(3).238–99.

Snedeker, J., and S. Yuan. 2008. Effects of prosodic and lexical constraints on parsing in young children (and adults). Journal of Memory and Language, 58.574–608.

Snow, D. 1994. Phrase-final syllable lengthening and intonation in early child speech. Journal of Speech and Hearing Research 37.831–40.

Soderstrom, M., A. Seidl, D. Kemler-Nelson, and P. Juscxzyk. 2003. The prosodic bootstrapping of phrases: evidence from prelinguistic infants. Journal of Memory and Language 49.249–67.

Solan, L. 1980. Contrastive stress and children's interpretation of pronouns. Journal of Speech and Hearing Research 23.688–98.

Terken, J., and S. G. Nooteboom. 1987. Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. Language and Cognitive Processes 2.145–63.

Weber, A., B. Braun, and M. W. Crocker. 2006. Finding referents in time: eye-tracking evidence for the role of contrastive accents. Language and Speech 49(3).367–92.

Wells, B., S. Peppé, and N. Goulandris. 2004. Intonation development from five to thirteen. Journal of Child Language 31.749–78.

Wieman, L. A. 1976. Stress patterns in early child language. Journal of Child Language 3.283–6.