

## Prosodic Bootstrapping of Clauses: Is it Language-Specific?

Kara Hawthorne (*University of Arizona*), Reiko Mazuka (*RIKEN Brain Science Institute & Duke University*), LouAnn Gerken (*University of Arizona*)

### Introduction

It has long been observed that prosody and syntax interact in our spoken language. For example, prosodic modulations such as pausing, final-syllable lengthening, and pitch resets occur reliably at clause and major phrase boundaries, particularly in infant-directed speech (Broen, 1972; Bernstein Ratner, 1986; Fernald et al., 1989; Soderstrom et al., 2008). Proponents of the prosodic bootstrapping hypothesis argue that infants may be able to take advantage of the correlations that exist between prosodic cues and syntactic boundaries to get a head-start in syntax acquisition (Gleitman & Wanner, 1982; Morgan, 1986). Since prosodic information is readily available in the acoustic signal and infants as young as 4.5 months have demonstrated sensitivity to correlations between prosodic information and syntactic boundaries (Jusczyk, 1989), prosody may facilitate the earliest stages of syntax acquisition.

Recent work in this area has asked whether prosodic structure influences infants' memory for strings of words. Several studies (e.g., Mandel et al., 1994; Soderstrom et al., 2005) have found that infants more easily remember strings of words when they are initially presented with well-formed prosody; for example, as a sentence rather than in a list. These studies suggest that the prosodic grouping of words influences how we remember those words, but they leave open the question of whether prosodic grouping influences syntactic learning.

In this paper, we explore two aspects of the prosodic bootstrapping hypothesis. First, we build on our previous work (Hawthorne & Gerken, 2011) and ask whether infants treat words within prosodic units as more cohesive and constituent-like than words that cross a prosodic boundary. More specifically, we ask whether infants recognize words with clause-like prosody (i.e., intonational phrase prosody) when they have moved to a new position in the sentence. One property of certain syntactic constituents in natural languages is that they can move, as seen in Examples (1) and (2) (brackets represent clausal constituents).

- (1) [we'll go to the zoo] [when Grandma gets here]  
(2) [when Grandma gets here] [we'll go to the zoo]

Note that the prosodic features of a clause are different depending on where that clause occurs in the sentence. "We'll go to the zoo," for example, has notably different prosody in Example (1) versus Example (2). Therefore, if a learner uses prosody to find a clause and then recognizes that clause when it has moved and its acoustics are quite different, she would appear to be treating the prosodically marked group of words as a constituent-like unit. This is tested in Experiment 1.

Experiment 2 investigates two potential mechanisms underlying the prosodic bootstrapping of syntax. One possibility is that *general acoustic* manifestations of prosody are sufficient to induce learning, since prosodic cues are perceptually salient. Even rats, for example, can use prosody to discriminate between languages (Toro et al., 2003). On the other hand, infants might first have to learn about the *language-specific* nature of the correlations between prosody and syntax in the target language before they can use that knowledge in syntax acquisition. We test these hypotheses by looking at whether Japanese-acquiring infants can use non-native English prosody for syntactic learning.

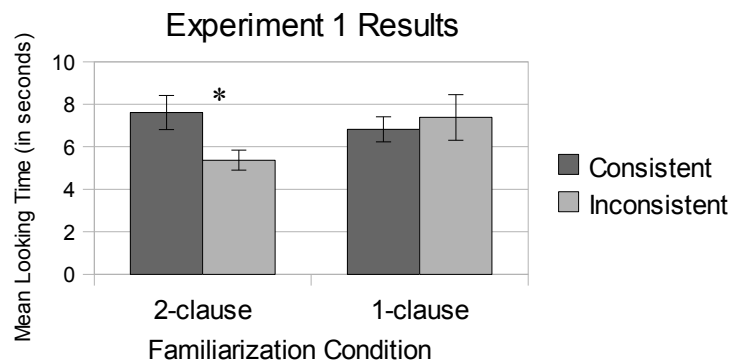
### Experiment 1

In Experiment 1, American English-acquiring 19-month-olds ( $n = 24$ ) were familiarized with six-word sentences from an artificial grammar. The infants were divided into two groups based on

the prosody of the familiarization stimuli. Half of the children heard utterances with 2-clause prosody of the form ABC, DEF, and half heard 1-clause utterances of the form ABCDEF. Each letter represents a class of nonsense words (class A = *bup* and *nim*). The familiarization period lasted for a total of 12 minutes.

Both groups were tested using the Head-Turn Preference Procedure on items with 2-clause prosody of the forms DEF, ABC and EFA, BCD. The DEF, ABC items are *consistent* with the clauses that were presented to the children from the 2-clause familiarization group, because the clauses (ABC and DEF) are simply switched (as in Examples (1) and (2), above). In contrast, the EFA, BCD items are *inconsistent* with the 2-clause familiarization items, since the clauses represent the ungrammatical movement of a non-constituent. If the infants in the 2-clause group treat words within prosodically marked clauses as more cohesive and constituent-like than words that cross a prosodic boundary, they should discriminate between these two types of test items. The 1-clause group, on the other hand, serves as a control condition: the 1-clause familiarization items provide no cues that parts of the utterance might be moveable constituents, so all of the test items should be equally unfamiliar.

As predicted, there was a significant interaction of familiarization condition by test condition ( $F(1, 110) = 4.72, p < .05$ ). The infants in the 2-clause group listened significantly longer to the consistent test items ( $F(1, 55) = 4.93, p < .05$ ), while the infants in the 1-clause group did not prefer one type of test item over the other ( $F < 1$ ). This pattern of results provides support for the prosodic bootstrapping hypothesis. The infants who were familiarized with utterances containing a prosodic break treated the two chunks like syntactic constituents, recognizing the prosodically grouped words when they had moved to a new position in the sentence and when they had a new prosodic contour.



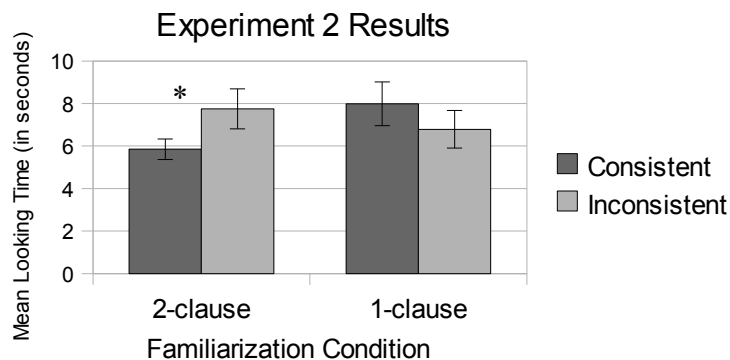
## Experiment 2

In Experiment 2, we asked whether the learning shown by the American infants in the 2-clause condition in Experiment 1 is due to their acquired knowledge of English prosody or to the acoustic salience of prosodic information. To explore these competing hypotheses, we replicated Experiment 1 using a different population: 19-month-old Japanese-acquiring infants in Japan ( $n = 24$ ). Importantly, the materials were the same English-based stimuli used in Experiment 1.

Japanese infants were chosen because Japanese and English have several important prosodic and syntactic differences (c.f. Venditti et al., 1996; Fisher & Tokura, 1996). English, for example, allows for complex syllable shapes and is typically considered a stress-timed language. Japanese allows only simple syllable shapes and is a mora-timed language. English marks prominent syllables with stress, while Japanese marks prominent syllables with pitch accent. Perhaps due in part to its use of pitch accents, Japanese infant-directed speech (IDS) has less pitch variation than American English IDS (Fernald et al., 1989). English and Japanese also differ at the level of basic word order: verbs precede objects in English (OV) and follow objects in Japanese (VO). Nespor et al. (2008) have argued that OV and VO languages differ in the manifestation of phrasal-level prosodic cues.

Japanese and American infants, by the time they are 19 months old, will have acquired quite different prosodic expectations, so if the language-specific hypothesis is correct, neither group of Japanese children should discriminate the two types of test items because the prosodic features of the English-based stimuli violate typical Japanese prosody. On the other hand, under the general acoustic hypothesis, Japanese infants should perform like American infants and discriminate consistent from inconsistent test items only in the 2-clause condition.

Results support the general acoustic version of the prosodic bootstrapping hypothesis. There was a significant interaction of familiarization condition by test condition ( $F(1, 110) = 7.54, p < .05$ ). The infants in the 2-clause group listened significantly longer to the inconsistent test items ( $F(1, 55) = 4.62, p < .05$ ), while the infants in the 1-clause group did not discriminate ( $F(1, 55) = 3.23, p = .08$ ). Note that the infants in the 1-clause group trended towards a preference for the consistent test items; however, this effect does not reach significance and it disappears when the first test item is discarded from the analysis ( $F < 1$ ). Therefore, Japanese-acquiring infants can make use of non-native prosodic information to locate clause-like units and recognize those clauses when they behave like syntactic-rule-abiding constituents. This suggests that prosodic information is sufficiently robust that it can be used in syntax acquisition before language-specific prosodic knowledge has been acquired.



## Discussion

The main finding from this research is that both American and Japanese infants discriminated between consistent and inconsistent test items only in the 2-clause condition. Therefore, both groups appear to have used English prosody to find groups of words which they treated as moveable, syntax-like constituents. One puzzle is that American infants in the 2-clause group showed a familiarity preference (listening longer to the more familiar consistent test items), while the Japanese infants in the 2-clause group showed a novelty preference (listening longer to the ungrammatical inconsistent test items). This is surprising, since a novelty preference is typically associated with a longer familiarization period or an easier task (Houston-Price & Nakai, 2004, e.g.). This result seems to suggest that the Japanese infants were *better* at using English prosody for syntax learning than American infants are. One likely explanation is that the difference in listening preference is a cultural effect. It has been noted that Japanese mothers spend more time rocking and carrying their infants compared to American mothers (Caudill & Weinstein, 1969), and this was observed by the experimenters during the experiment sessions. The familiarization phase began with a ten minute period during which the infant was allowed to play with toys while familiarization items played in the background. During this period, the American infants invariably played with the toys the entire time, while the majority of the Japanese infants spent a portion – or all – of that time on their mothers' laps, watching the experimenter play with the toys. Since the Japanese infants were, on average, less actively engaged with the toys during this period, they may have spent more time listening to the familiarization stimuli, such that by the testing period, they preferred to listen to the

novel, inconsistent test sentences.

These experiments provide the strongest evidence to date (1) that infants can use prosodic information for segmenting the speech stream into chunks that behave like syntactic constituents and (2) that universally detectable aspects of prosody can be used to support syntax acquisition.

## References

- Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics*, 12, 245-254.
- Broen, P.A. (1972). The verbal environment of the language-learning child. *Monograph of American Speech and Hearing Association*, 17.
- Caudill, W. & Weinstein, H. (1969). Maternal care and infant behavior in Japan and America. *Psychiatry: Journal for the Study of Interpersonal Processes*, 32, 12-43.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-linguistic study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16, 477-501.
- Fisher, C. & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development*, 67, 3192-3218.
- Gleitman, L.R. & Wanner, E. (1982). Language acquisition: The state of the art. In E Wanner & L.R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3-48). Cambridge: Cambridge University Press.
- Hawthorne, K., & Gerken, L.A. (2011). Infants use prosody to learn about clauses. *LSA Meeting Extended Abstracts 2011*. Retrieved March 12th, 2012 from <http://www.elanguage.net/journals/index.php/lsameeting/issue/view/269>.
- Houston-Price, C. & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, 13, 341-348.
- Jusczyk, P.W. (1989). *Perception of cues to clausal units in native and non-native languages*. Paper presented at the biennial meeting of the Society for Research in Child Development, Kansas City, Missouri.
- Morgan, J.L. (1986). *From simple input to complex grammar*. Cambridge, MA: MIT Press.
- Nespor, M., Shukla, M., Van de Vijver, R., Avesani, C., Schraudolf, H. & Donati, C. (2008). Different phrasal prominence realizations in VO and OV languages. *Lingue E Linguaggio*, VII, 1-29.
- Soderstrom, M., Kemler Nelson, D.G., & Jusczyk, P.W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior & Development*, 28, 87-94.
- Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J.L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, 35, 869-902.
- Toro, J.M., Trobalon, J.B., & Sebastian-Galles, N. (2003). The use of prosodic cues in language discrimination tasks by rats. *Animal Cognition*, 6, 131-13.
- Venditti, J.J., Jun, S.A., & Beckman, M.E. (1996). Prosodic cues to syntactic and other linguistic structures in Japanese, Korean, and English. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 287-311). Mahway, NJ : Erlbaum.