

The Effect of Prosody on Distributional Learning in 12- to 13-Month-Old Infants

Malathi Thothathiri^{*1}, Jesse Snedeker and Erin Hannon²
Department of Psychology, Harvard University, Cambridge, MA, USA

Distributional information is a potential cue for learning syntactic categories. Recent studies demonstrate a developmental trajectory in the level of abstraction of distributional learning in young infants. Here we investigate the effect of prosody on infants' learning of adjacent relations between words. Twelve- to thirteen-month-old infants were exposed to an artificial language comprised of 3-word-sentences of the form aXb and cYd, where X and Y words differed in the number of syllables. Training sentences contained a prosodic boundary between either the first and the second word or the second and the third word. Subsequently, infants were tested on novel test sentences that contained new X and Y words and also contained a flat prosody with no grouping cues. Infants successfully discriminated between novel grammatical and ungrammatical sentences, suggesting that the learned adjacent relations can be abstracted across words and prosodic conditions. Under the conditions tested, prosody may be only a weak constraint on syntactic categorization. Copyright © 2011 John Wiley & Sons, Ltd.

Key words: artificial grammar; distributional learning; prosody; syntax; infants

INTRODUCTION

Syntactic categorization—the grouping of words into grammatical categories—is an important prerequisite of productivity in language. Knowing that the words *dog* and *cat* belong to the same syntactic category allows us to generate 'I have a dog' after having heard utterances like 'I have a cat'. Young language learners clearly master this task before adulthood, but precisely when and how they do so

*Correspondence to: Malathi Thothathiri, Department of Psychology, Swarthmore College, Swarthmore, PA 19081-1397, USA. E-mail: mthotha1@swarthmore.edu

is a still an open question. Distributional information, that is, the statistics of co-occurrence between different words is one possible way to categorize words. Research in the past few decades has uncovered rich distributional learning abilities in young infants but also important constraints on such learning in infants of different ages. Here, we ask whether 12- to 13-month-olds can learn abstract distributional categories that generalize across surface properties. In particular, we ask whether the distributional relations learned by infants of this age are conditioned by prosody.

Our study was motivated by two observations. First, distributional learning shows a developmental trajectory wherein younger infants show less precise and/or abstract learning compared with older infants. Second, prosody plays an important role in early language acquisition. It is important to know then whether the earliest stages of distributional learning of syntactic categories are constrained by prosody. Below, we summarize developmental changes in distributional learning and consider the role of prosody in language acquisition.

Function words such as 'the' and 'a' may serve as anchor points in the distributional learning of categories (Valian and Coulson, 1988). Function words may be easily detectable in the language input by virtue of their higher frequency of occurrence and distinct phonological properties relative to content words (Gervain, Nespor, Mazuka, Horie, & Mehler, 2008; Morgan, Shi, & Allopenna, 1996). The distributional relation between specific function and content words may enable the categorization of content words into nouns and verbs (e.g. *the* <noun>, *are* <verb>). The ability to represent and use function words for syntactic categorization develops gradually during the first 2 years after birth. Newborn infants can use a constellation of perceptual cues (e.g. vowel duration, amplitude) to discriminate between open and closed class words (Shi, Werker, & Morgan, 1999). By 7 months, infants are able to detect function words in continuous speech (Höhle & Weissenborn, 2003). By 13 months, they can represent function words with sufficient accuracy so as to distinguish between valid English function words (e.g. *his*) and minimally different nonsense words (e.g. *ris*. Shi, Werker, & Cutler, 2006). Fourteen- to sixteen-month-old infants go further; they can use function words to syntactically categorize co-occurring novel words in continuous speech (Höhle Weissenborn, Kiefer, Schulz, & Schmitz, 2004). In the study of Höhle *et al.*, infants familiarized to novel words in a noun context (e.g. German *ein glamm*, a glamm) subsequently distinguished between passages containing those words in novel verb contexts from passages containing those words in novel noun contexts. The results summarized in this paragraph show an early sensitivity to the perceptual properties that identify function words (at birth) followed by increasing accuracy in the representation of such words (7 to 13 months) and finally the use of function words for syntactic categorization (14 to 16 months).

Within the domain of adjacent dependency learning, there are suggestions of increasing abstraction over age. Several studies of distributional learning have suggested that even adults more readily learn adjacent relations when there are multiple correlated cues as opposed to a single distributional cue (see e.g. Braine, 1987; Mintz, 2002; Smith, 1966). Gerken, Wilson, and Lewis (2005) exposed 17-month-old infants to a Russian gender paradigm with some members of the paradigm withheld. The familiarization stimuli contained partial correlations between phonology and distributional context (in this example, case endings). Specifically, half of the feminine nouns ended in phonological markers—k and co-occurred with the case endings—oj and —u. Half of the masculine nouns ended in phonological markers—tɛl and co-occurred with case endings—ya and—yem. During test, 17-month-old infants were able to generalize to novel combinations

not heard during familiarization. Importantly, they could do so even for test items that did not contain the phonological markers—*k* or—*tɛl*. This level of abstraction—extending distributional relations learned from phrases containing correlated phonological and distributional cues to phrases that do not contain the phonological cues—has not been reported in infants younger than 17 months. In fact, Gómez and Lakusta (2004) have suggested that younger infants, specifically 12-month-olds, might be at an earlier stage in their distributional learning, succeeding only in generalizing to novel cases where the same correlated cues exist as during familiarization. These authors familiarized 12-month-old infants with an artificial language consisting of *aX* and *bY* type phrases where *X* words were disyllabic and *Y* words were monosyllabic. Subsequent to familiarization, infants were tested on sentences containing novel *X* and *Y* words. Infants successfully discriminated between novel grammatical (*aX*, *bY*) and ungrammatical (*aY*, *bX*) sentences suggesting that they had learned the distributional relations between 'a' words and monosyllabic words and 'b' words and disyllabic words.

How abstract is the learning of adjacent dependencies at the age of 12 months? Given the growing abstraction of learned distributional relations between 12 and 17 months of age, we asked whether infants can generalize distributional relations to word pairings never seen before and whether they can do so even when test utterances differ prosodically from the utterances from which they learned the relations. This is relevant because infants show an early sensitivity to prosody. By 10 months of age, they are sensitive to correlations amongst the acoustic cues that signal clausal or phrasal boundaries (Hirsh-Pasek *et al.*, 1987; Jusczyk *et al.*, 1992). They prefer passages containing pauses at clausal or phrasal boundaries to those containing pauses in the middle of a clause or phrase. They also use these prosodic cues to group words; infants as young as 6 months of age are more likely to recognize a sequence of words presented within a single prosodic phrase than a sequence of words crossing a prosodic boundary (Nazzi, Kemler Nelson, Jusczyk, & Jusczyk, 2000; Soderstrom, Seidl, Kemler Nelson, & Jusczyk, 2003). Thus, the effect of prosody on infants' recognition of units has been well documented. Prosody also seems to influence word segmentation and lexical access. Johnson (2008) familiarized 12-month-olds to passages containing intended words that did not cross a word boundary (e.g. *toga* from *toga#lore*) and unintended words whose syllables crossed a word boundary (e.g. *dogma* from *dog#maligns*). Subsequently, infants showed a preference for familiar intended words over unfamiliar words, but not for familiar unintended words over unfamiliar words. This suggests that the presence of prosodic word boundaries influences infants' segmentation of words from continuous speech. Gout, Christophe and Morgan (2004) familiarized 13-month-olds with isolated words (e.g. *paper*). Subsequently, infants showed a preference for passages containing the familiarized word within one prosodic unit (e.g. *The scandalous paper...*) over passages where the syllables of the familiarized word were separated by a prosodic boundary (e.g. [*The outstanding pay*] [*persuades him...*]). This suggests that the presence of prosodic boundaries constrains lexical access. Other researchers have hypothesized that prosody may similarly constrain distributional analysis for syntactic categorization (prosodic bootstrapping: Gleitman & Wanner, 1982; Morgan & Newport, 1981).

In naturalistic language production, prosody has been hypothesized to constrain the production of grammatical determiner and noun combinations. For example, corpus analyses show that French-speaking 1- and 2-year-old children produce obligatory determiners earlier with monosyllabic than with multisyllabic nouns (Bassano, Maillochon & Mottet, 2008; Demuth & Tremblay, 2008). This has been

interpreted as evidence for grammatical markers first appearing in prosodically licensed contexts (as part of a binary foot; Demuth & Tremblay, 2008).

Are learned distributional relations prosodically conditioned in 12- to 13-month-old infants? As described above, infants at this age show some distributional learning (Gómez & Lakusta, 2004) but such learning may be more constrained compared with older children. To investigate the possible effect of prosody on distributional learning, we familiarized infants to utterances of the form aXb or cYd . These training sentences contained either a prosodic break after the first word ($a\#Xb$) or after the second word ($aX\#b$). During the test phase, infants heard sentences containing new X and Y words (so all distributional relations were novel). The test sentences did not contain any prosodic breaks between words, so these sentences were prosodically different from the training sentences. This design allowed us to explore two questions as follows: (i) whether infants would generalize the learned patterns to novel utterances containing new X and Y words; and (ii) whether this generalization was 'supra-prosodic', that is, whether infants would generalize the learned distributional relations to novel prosodic structures.

METHODS

Participants

Sixteen infants from the Boston area participated (four in each experimental condition). The mean age was 12 months and 22 days (range: 11; 28 to 14; 8). Eleven of the infants were female. Eight other infants were tested but excluded because of programme error (5) and fussiness (3).

Stimuli

We adapted the stimuli used by Gómez and Lakusta (2004). Infants were randomly assigned to one of two training languages. In L1, strings were of the form aXb or cYd . In the counterbalanced language L2, they were of the form aYb or cXd . The categories a , b , c and d consisted of two words each; X and Y categories consisted of six words each (Table 1). This resulted in a total of 48 sentences, which were randomly split into two sets of 24 sentences. Each infant heard both sets. X words were disyllabic; Y words were monosyllabic. Thus, discovering the structure of the training language amounted to discovering the contingency between a and/or b and disyllabic words and c and/or d and monosyllabic words (and the reverse for L2).

Table 1. Words used during training

a	b	c	d	X	Y
alt	omp	ong	ast	coomo	deech
ush	enk	erd	ulf	fengle	ghop
				kicey	jic
				loga	skeej
				paylig	vabe
				wazil	tam

Grammatical sentences were of the form aXb/cYd in Language 1 and aYb/cXd in Language 2.

X words were disyllabic; Y words were monosyllabic.

Infants were randomly assigned to one of two prosodic training conditions: in Prosody 1, there was a prosodic break after the first word; in Prosody 2, there was a break after the second. Thus, in Prosody 1, *alt coomo omp* was produced as *alt # coomo omp* (where # denotes a prosodic boundary). In Prosody 2, the same sentence was produced as *alt coomo # omp*. Sentences were recorded by a female English speaker who was asked to produce the sentences naturally using infant-directed intonation. Figures 1a and 1b show the pitch contours and waveforms for a training sentence produced using each of the two prosodies. As expected, the two versions differed in a number of ways. Because we were particularly interested in the intended prosodic boundary, we evaluated three well-known correlates of boundaries, namely preboundary lengthening, pitch change and pause duration (Beckman & Pierrehumbert, 1986; Cooper & Paccia-Cooper, 1980; Wightman, Shattuck Hufnagel, Ostendorf, & Price, 1992). All measurements were made using Praat (Boersma & Weenink, 2005). We measured the three characteristics for the first five training sentences in each of the four experimental conditions (two languages and two prosodies). As shown by the averaged results in Table 2, the position at which a prosodic boundary was intended (after the first and second words in Prosody 1 and Prosody 2, respectively) is, in fact, associated with longer pre-boundary rhymes, larger pitch changes and longer pauses compared with the position at which the speaker did not intend to produce a boundary. To further confirm that our prosodic manipulation was effective, we undertook two steps. First, we played the entire set of sentences to a naïve adult. This adult perceived all sentences as having three words, with the first two or the last two words grouped together as we had intended. Second, we played a subset of the training sentences⁷ to a trained ToBI transcriber. For all sentences, the transcriptions indicated a strong prosodic boundary (break index 4) at the intended position and no such boundary in the other position.

Unlike training sentences, test sentences were intended to contain no grouping cues. Figure 1c shows the pitch contour and waveform for an example test

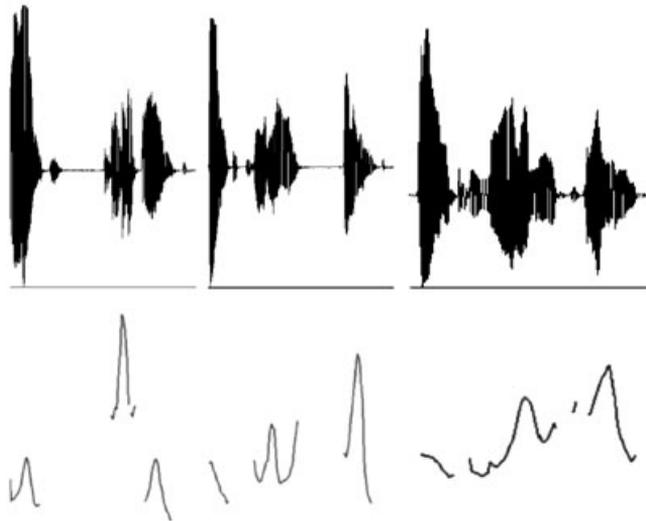


Figure 1. Waveforms and pitch contours of example sentences. a. Training (*alt coomo omp*) in Prosody 1; b. Training (*alt coomo omp*) in Prosody 2; c. Test (*alt nawlup omp*).

Table 2. Acoustic correlates of prosody in training and test sentences

Acoustic correlate	Sentence position	Prosody 1	Prosody 2	Test
Final syllable rhyme duration (s)	1st word	0.687 (0.004)	0.379 (0.005)	0.356 (0.033)
	2nd word	0.286 (0.087)	0.447 (0.173)	0.360 (0.075)
Pitch change (Hz)	Between 1st and 2nd words	99.920 (6.296)	47.024 (37.655)	35.870 (25.556)
	Between 2nd and 3rd words	56.188 (17.152)	106.068 (14.595)	42.195 (36.946)
Pause duration (s)	Between 1st and 2nd words	0.510 (0.006)	0.085 (0.001)	0.047 (0.032)
	Between 2nd and 3rd words	0.089 (0.021)	0.511 (0.017)	0.079 (0.035)

The figures in parentheses are standard deviations.

sentence. The rightmost column of Table 2 shows the averaged acoustic correlates of the boundaries in 24 test sentences.⁴ As expected, the magnitudes of preboundary rhyme length, pitch change and pause duration for the two relevant sentence positions (after the 1st and 2nd words) appear comparable.

During test, infants heard three-word sentences of two types. Grammatical sentences were of the form aXb/cYd; ungrammatical sentences were of the form aYb/cXd (and the reverse for L2). We used the same a, b, c and d words as before, but the X and Y words were new (Table 3), resulting in novel sentences that were not heard during training. Again, X words were disyllabic and Y words were monosyllabic. Twenty four of the 48 possible novel grammatical sentences were chosen and split into two lists, resulting in 12 test sentences per list. Within each list, the X and Y words appeared once each. The others (a, b, c and d) appeared three times each. The two grammatical lists were played twice each for a total of four grammatical trials played in two blocks such that there was no repetition within a block. The same procedure was used to create two ungrammatical lists, again repeated twice across blocks, for a total of four ungrammatical trials. All four test lists are shown in Table 4.

In both grammatical and ungrammatical sentences, all words appeared in their correct absolute positions. Success at this task therefore required an understanding of the relationship between a/b/c/d words on the one hand and the middle words (X/Y) on the other.

Procedure

We used a modification of the head-turn preference procedure (Kemler Nelson, Jusczyk, Mandel, Myers, Turk, & Gerken, 1995). Infants were seated on a caregiver's lap in a sound-attenuated booth. Caregivers wore headphones throughout the session. An experimenter controlled presentation of the stimuli

Table 3. Novel X and Y words used during test

X	Y
roosa	pel
bevit	foge
gackle	tood
meeper	vot
binow	rud
nawlup	biff

The other words (a, b, c and d) were the same as in training.

Table 4. Test stimuli

Type 1 (grammatical in L1, ungrammatical in L2)		Type 2 (ungrammatical in L1, grammatical in L2)	
List 1	List 2	List 1	List 2
erd vot ulf	erd foge ulf	alt tood omp	alt vot enk
ush meeper omp	alt nawlup omp	ush vot omp	ush pel enk
ong tood ulf	ush roosa enk	erd nawlup ast	ush tood enk
alt roosa omp	ong biff ulf	ong roosa ulf	ong meeper ast
erd rud ast	ush gackle enk	ush rud enk	ush foge omp
alt gackle omp	erd pel ast	ong bevit ast	ong binow ulf
ush binow enk	ong vot ast	ush biff enk	alt rud omp
ong foge ast	alt meeper enk	ong gackle ulf	erd roosa ast
ong pel ulf	ong rud ulf	erd meeper ulf	erd gackle ast
alt bevit enk	ush bevit omp	alt foge enk	ong nawlup ulf
ush nawlup enk	erd tood ast	erd binow ast	alt biff omp
erd biff ast	alt binow omp	alt pel omp	erd bevit ulf

These sentences either preserved or violated the dependencies found in training sentences (e.g. compare with L1 training: *alt coomo omp/erd deech ulf* and L2 training: *alt deech omp/erd coomo ulf*).

from outside the booth. All auditory stimuli were accompanied by a short, animated movie that remained constant across conditions. Each movie played on one of two computer monitors located to the left and right of the infant. Each session consisted of two training and eight test trials. All trials were preceded by a red blinking screen to orient the infants' attention. Each training trial consisted of a visual animation of swirling galaxies accompanied by 24 of the 48 training sentences played for a fixed duration until the end (~75 s) irrespective of infant looking. The two training trials alternated between the left and right monitors with order determined randomly.

The first test trial was presented immediately after training with order of presentation sides and test sentence types determined randomly. Movies consisted of a novel visual stimulus (swirling fractal patterns) accompanied by 12 test sentences. The experimenter, who was unable to hear the auditory stimuli, monitored infants looks to and away from the target monitor. Each test list was ~30 s long. It continued playing (looping if necessary) until the infant looked away for more than 2 s. Grammatical and ungrammatical lists alternated for a total of eight trials. All test trials used the same visual stimulus regardless of test sentence type ensuring that successful discrimination would depend on attention to the auditory stimuli.

RESULTS

The dependent measure was the infants' listening times to grammatical and ungrammatical lists. We excluded those trials during which the infant listened for less than 2 s.⁵ This eliminated a total of six trials across all participants (4.7%). We found no significant differences between the two training languages, so this variable was excluded from the analysis. A 2x2 mixed design ANOVA with the within-subjects variable trial type (grammatical/ungrammatical) and the between-subjects variable prosody (Prosody 1/2) yielded a significant main effect of trial type [$F(1,14) = 7.425$, $p < 0.02$]. Infants listened longer to grammatical ($M = 14.14$ s; $SE = 2.3$) compared with

ungrammatical sentences ($M = 8.01$ s; $SE = 1$) in both prosodic conditions (Figure 2). Twelve out of 16 infants showed this preference [sign test, $p = 0.077$]. There were no other main effects or interactions [$F_s < 1$, $p_s > 0.4$].

DISCUSSION

The results from our experiment show that 12- to 13-month-old infants can track adjacent relationships between frequently occurring context words (a/b/c/d) and less frequently occurring category words (X/Y) and generalize those relationships to novel X and Y words (similar to Gómez and Lakusta, 2004). Further, our experiment shows that the learned adjacent relations are not prosodically conditioned. Infants generalized the adjacent relations to utterances that differed prosodically from those heard during training. Previous research has suggested that the ability to detect words (i.e. *fixed* phonological sequences) across perceptual variation develops between 7.5 and 10.5 months of age (Singh, 2008). Our results go beyond such results in showing that 12- and 13-month-old infants can generalize, across prosodic variation, any learned relations between phonological units to sequences that they have never heard before. That is, infants in our study, having been trained on *alt coomo omp*, successfully discriminated between *alt nawlup omp* and *alt deech omp* despite never having heard the latter two sequences before and despite the fact that neither was prosodically similar to the first.

This study adds to a growing body of evidence, which shows that young language learners have impressively abstract distributional learning abilities. In some paradigms, 17-month-olds show similar levels of abstraction as adults. They can use partial correlations between distribution and phonology to learn adjacent relations and then extend them to novel instances that do not contain those correlations. As described in the Introduction, 12-month-olds' learning may be less abstract, but it is abstract nonetheless. The current study and a previous study by Gómez and Lakusta (2004) used stimuli that always contained correlated cues (distribution and number of syllables). Nevertheless, the generalization to *novel* instances suggests that the learning of adjacent relations was abstract (from one set of words to another set of words) and not simply limited to previously seen word pairings. The learned relations are also abstract across prosodic differences;

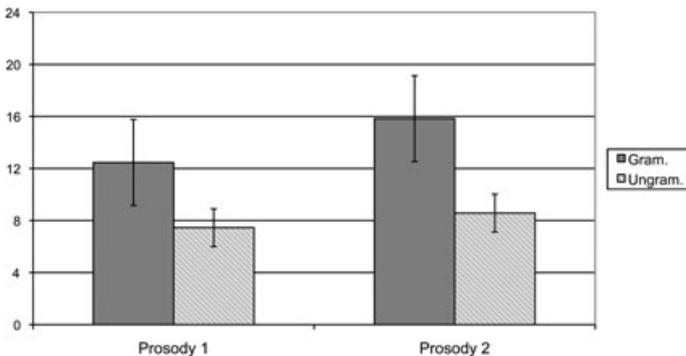


Figure 2. Listening times for grammatical and ungrammatical sentences (in seconds). Infants preferred grammatical sentences in both prosodic conditions.

in the current study, infants generalized the learned adjacent dependencies to novel prosodies.

Our stimuli were three-word utterances (e.g. aXb) each of which contained two adjacent relations (aX and Xb) as well as a relation between a frame and a middle word (a__b and X). Because our test stimuli contrasted novel grammatical sentences in which all three relations were valid (e.g. aX, Xb, a__b with X) with ungrammatical sentences in which all three relations were invalid (e.g. aY, Yb, a__b with Y), we cannot determine the exact relation(s) that infants were tracking and using in this task. Previous evidence suggests, however, that is unlikely that the 12- and 13-month-olds tested here were using frame representations like a__b for categorization. Several studies have suggested that learning non-adjacent dependencies may be harder (and may emerge later) than learning adjacent dependencies. Gómez and colleagues varied the middle element in artificial three-word utterances (aXc, bXd) and found that infants over 15 months of age were able to track non-adjacent relations, but only when the adjacent relations were highly variable and thus unpredictable (Gómez, 2002; Gómez & Maye, 2005). Twelve-month-old infants did not succeed even under these conditions (Gómez & Maye, 2005). The evidence from these artificial language studies dovetails with studies using natural language utterances. Some sensitivity to non-adjacent relations in infants' native languages has been shown in 18-, 19- and 24-month-old infants, but there have been no such demonstrations in infants younger than 15 months of age (van Heugten & Johnson, 2010; Höhle *et al.*, 2006; Santelmann & Jusczyk, 1998). Thus, the learning of non-adjacent relations may only emerge at a later stage of acquisition or under special input conditions (e.g. prior exposure to a relation in *adjacent* form; see Lany and Gómez (2008)).

Although the use of frames per se for categorization seems unlikely at this age, it is possible that infants benefitted from the presence of such frames. For example, infants could have tracked both bigrams in an integrated fashion (aX + Xb) as opposed to tracking each bigram separately (aX, Xb. See St. Clair, Monaghan, & Christiansen (2010) for a computational proposal regarding flexible versus fixed frames). The finding from our study cannot distinguish between these possibilities, or indeed whether infants were succeeding based on a *single* bigram (aX or Xb). In fact, it might be particularly interesting to evaluate in the future whether prosody influences, which dependencies are tracked or weighted higher. If prosody helps learners to restrict distributional analysis to linguistically relevant contexts—as suggested by the prosodic bootstrapping hypothesis—we might expect relations within a prosodic unit to be tracked better than those that straddle a prosodic boundary. A recent study by van Heugten and Shi reported that 17-month-old but not 14-month-old French infants successfully distinguished between grammatical and ungrammatical determiner-auxiliary non-adjacent dependencies in French. The 17-month-olds succeeded in this discrimination even though the non-adjacent dependency straddled a phonological phrase boundary (van Heugten & Shi, 2010). It remains to be determined whether younger infants like the 12- and 13-month-olds tested here would similarly ignore prosodic boundaries when tracking syntactic dependencies. In a pilot experiment not reported here, we failed to find a difference in infants' preference for within-prosodic-unit dependencies over across-prosodic-boundary ones. Future research should explore the conditions under which prosody may possibly constrain distributional analysis (e.g. for longer utterances or for tracking non-adjacent dependencies at younger ages). For the present, our results suggest that prosody does not exert an overly strong influence on the tracking of adjacent distributional relations between words.

Notes

1. Now at Swarthmore College.
2. Now at University of Nevada, Las Vegas.
3. We picked the first three sentences from each of eight training lists for a total of 24 sentences, split equally between the two languages and the two prosodies.
4. We picked the first three sentences from each of the four test lists.
5. Test sentences ranged in length between 1.5 and 2 s, so we estimated that this was the minimum time required to decide whether a given test list was grammatical.

ACKNOWLEDGMENTS

We wish to thank Stefanie Shattuck-Hufnagel for feedback on the prosody manipulation, Beth Johnson for ToBI transcriptions, Rachel Levine for help in setting up the experiments and Niloufar Rahi and Emily D'Amato for assistance in scheduling and testing participants. Part of this work was funded by a Harvard University Mind/Brain/Behavior (MBB) award to Malathi Thothathiri. We thank the MBB Initiative for their generous support.

REFERENCES

- Bassano, D., Maillochon, I., & Mottet, S. (2008). Noun grammaticalization and determiner use in French children's speech: A gradual development with prosodic and lexical influences. *Journal of Child Language*, 35(02), 403–438.
- Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255–309.
- Boersma, P., & Weenink, D. (2005). Praat: doing phonetics by computer (version 4.3.28) [Computer program]. Available from <http://www.praat.org/>
- Braine, M. D. S. (1987). What is learned in acquiring words classes: A step toward acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 65–87). Hillsdale, NJ: Erlbaum.
- Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.
- Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, 35(01), 99–127.
- Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32(02), 249–268.
- Gervain, J., Nespor, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognitive Psychology*, 57, 56–74.
- Gleitman, L., & Wanner, E. (1982). Language acquisition: The state of the art. In E. Wanner, & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3–48). Cambridge, UK: Cambridge University Press.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431–436.
- Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science*, 7(5), 567–580.
- Gómez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7(2), 183–206.
- Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, 51(4), 548–567.

- Hirsh-Pasek, K., Kemler Nelson, D., Jusczyk, P. W., Wright, K., Druss, B., & Kennedy, L. J. (1987). Clauses are perceptual units for young infants. *Cognitive Psychology*, *24*, 252–293.
- Höhle, B., Schmitz, M., Santelmann, L. M., & Weissenborn, J. (2006). The recognition of discontinuous verbal dependencies by German 19-month-olds: Evidence for lexical and structural influences on children's early processing capacities. *Language Learning and Development*, *2*(4), 277–300.
- Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, *6*(2), 122–127.
- Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, *5*(3), 341–353.
- Johnson, E. K. (2008). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech.
- Jusczyk, P. W., 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*, 252–293.
- Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. A. (1995). The headturn preference procedure for testing auditory perception. *Infant Behavior & Development*, *18*, 111–116.
- Lany, J., & Gómez, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychological Science*, *19*(12), 1247.
- Mintz, T. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition*, *30*, 678–686.
- Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior*, *20*, 67–85.
- Morgan, J., Shi, R., & Allopenna, P. (1996). Perceptual bases of rudimentary grammatical categories. In J. Morgan & K. Demuth (Eds.), *Signal to syntax*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Nazzi, T., Kemler Nelson, D. G., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six month olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy*, *1*, 123–147.
- Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, *69*(2), 105–134.
- Shi, R., Werker, J. F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10*(2), 187–198.
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*, B11–B21.
- Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition*, *106*, 833–870.
- Smith, K. H. (1966). Grammatical intrusions in the recall of structured letter pairs: Mediated transfer or position learning? *Journal of Experimental Psychology*, *72*, 580–588.
- Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, *49*, 249–267.
- St. Clair, M. C., Monaghan, P., & Christiansen, M. H. (2010). Learning grammatical categories from distributional cues: Flexible frames for language acquisition. *Cognition*, *116*, 341–360.
- Valian, V., & Coulson, S. (1988). Anchor points in language learning: The role of marker frequency. *Journal of Memory and Language*, *27*, 71–86.
- van Heugten, M., & Johnson, E. K. (2010). Linking infants' distributional learning abilities to natural language acquisition. *Journal of Memory and Language*, *63*(2), 197–209.
- van Heugten, M., & Shi, R. (2010). Infants' sensitivity to non-adjacent dependencies across phonological phrase boundaries. *The Journal of the Acoustical Society of America*, *128*, EL223–EL228.
- Wightman, C. W., Shattuck Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, *91*, 1707.