Distributional category learning by 12-month-old infants: an investigation into the role of prosody and distributional frames

Malathi Thothathiri

Jesse Snedeker

Erin Hannon

Department of Psychology

Harvard University

Address correspondence to:

Malathi Thothathiri

Department of Psychology, Harvard University

William James Hall, 33 Kirkland Street

Cambridge, MA 02138

Ph: (617) 384-8357

Email: malathi@wjh.harvard.edu

Abstract

Distributional information is a potential cue for learning syntactic categories. Recent artificial grammar studies demonstrate sophisticated distributional learning by young infants. Here we investigate the possible mechanisms and representations underlying this ability. Does prosody constrain distributional analysis? What specific distributional relations do learners track? Twelve-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. Subsequently they were tested on novel utterances that were consistent or inconsistent with the training sentences. In Experiment I, infants showed evidence for having learned the relevant relations by successfully discriminating between novel grammatical and ungrammatical sentences. In Experiment II, we asked whether prosody influences infants' distributional analysis. Contrary to our expectations, infants did not show a preference for relations between words that fell within a prosodic unit over those that straddled a prosodic boundary. In Experiment III, we explored whether infants' success in the first experiment arose from their representation of nonadjacent relations or distributional frames. Our results did not support a frames hypothesis. We discuss these results and offer hypotheses regarding the nature of infants' distributional learning abilities.

Keywords: Artificial Grammar, Distributional Learning, Prosody, Frames, Infants

Introduction

Humans use language productively, combining known words in diverse but constrained ways to create novel sentences. An important prerequisite of this productivity is the grouping of words into grammatical categories. 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"*. Researchers have proposed a variety of information sources that can help a child acquire syntactic categories, including phonology, semantics and distributional statistics (Kelly, 1992; Maratsos & Chalkley, 1980; Pinker, 1984). Somewhat frustratingly for language researchers (and perhaps the child), none of these sources are completely predictive of the category membership of words (the reliability problem). For example, although words that refer to actions are typically verbs, some words (e.g., *action*) violate this generalization. The same problem arises with distributional information. An algorithm that depends solely on matching sentential contexts may lead the child to incorrectly assume that *happy* and *here* belong to the same category upon hearing "*I am happy*" and "*I am here*". Pinker (1984) highlighted an additional problem for distributional accounts of categorization. In a multi-word utterance, there are many possible relations between words, only some of which are syntactically informative. How does the child know *which* relations to track (the computability problem)?

Research over the past few decades has suggested two ways that young learners may overcome these issues. First, infants may constrain their use of distributional evidence by focusing on a limited set of distributional relations for the purpose of categorization. This could alleviate the problem of computability (and if these cues were particularly valid, it could also alleviate the problem of reliability). Second, infants may look for redundancies between sources of information whereby multiple types of cues delineate the same categories. This might improve the reliability of categorization.

One way to constrain distributional analysis is to use function words as anchor points (Valian & Coulson, 1988). Function words often delineate phrase boundaries and may aid learners by constraining the contexts in which the analyses have to be performed (Gleitman & Wanner, 1982; Morgan, Meier & Newport, 1987). They are easily detectable in the language input by virtue of their higher frequency and different phonological properties relative to content words (Gervain, Nespor, Mazuka, Horie & Mehler, 2008; Morgan, Shi & Allopenna, 1996). Recent results with infants make these proposals viable. Infants are able to detect function words in continuous speech as well as use them to syntactically categorize co-occurring novel words (Hohle & Weissenborn, 2003; Hohle, Weissenborn, Kiefer, Schulz & Schmitz, 2004).

Another solution is to use redundant or correlated cues. Categories that are delineated by multiple cues (distributional, phonological and semantic) are more likely to be correct than those suggested by a single cue. Starting with Smith (1966), several artificial grammar studies have demonstrated that while adults can learn the absolute positions of words rather easily, they do not learn the relations between words unless there are additional semantic or phonological cues (Braine, 1987; Frigo & McDonald, 1998). For example, Frigo & McDonald (1998) showed that adults were able to learn relations such as aX and bY in an artificial language when a subset of the X and Y words were phonologically similar, but they failed to learn the dependencies in the absence of such similarities.

In this paper, we investigate the role of two other constraints, namely prosody and distributional frames, in making distributional analysis more tractable and reliable. Below we consider each of these factors in turn.

With respect to prosody, proponents of the prosodic bootstrapping hypothesis have suggested that prosody may help learners in restricting distributional analysis to linguistically relevant contexts (Gleitman & Wanner, 1982; Morgan & Newport, 1981). Utterance boundaries are reliably marked by acoustic cues such as vowel-lengthening and pauses (Fisher & Tokura, 1996). Phrasal boundaries are less reliably correlated with prosody (e.g., see Gerken, Jusczyk & Mandel, 1994) but infants may nevertheless benefit by splitting an utterance into smaller chunks for analysis (Jusczyk, 1998). Consistent with this proposal, a recent computational analysis of CHILDES corpora showed that categorization improves when distributional analysis is restricted to local phrasal contexts, even beyond categorization using large 8-word contexts (Mintz, Newport & Bever, 2002). Behaviorally, Morgan, et al. (1987) used a miniature language to show that adult learners are capable of using prosodic information to learn syntactic structure. Other studies show that infants less than 10 months of age are sensitive to correlations amongst the acoustic cues that signal clausal or phrasal boundaries (Hirsh-Pasek, Kemler Nelson, Jusczyk, Wright, Druss & Kennedy, 1987; Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward & Piwoz, 1992). Infants prefer passages containing pauses at clausal or phrasal boundaries to those containing pauses in the middle of a clause or phrase. Most tellingly, infants use these prosodic cues to group words; infants as young as six 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. What is unclear however is whether infants use such prosodic information to constrain syntactic categorization.

With respect to distributional frames, Mintz (2003) suggested that infants could use frequently occurring sentence frames with empty slots in the middle (e.g., you ____ it) for categorizing the words in the slots. Using computational analyses of child-directed speech, he found highly accurate classification of words that occur in the middle of such frames. In a similar vein, Childers and Tomasello (2001) have argued that repeated templates (e.g., He's *verb*-ing it) may aid in the acquisition of abstract constructions and categories. However, it is not clear if young language learners represent frames and use them for categorization, or even if they should. Contrary to Mintz (2003), Monaghan & Christiansen (2004) found superior categorization when their connectionist model used multiple bigrams[1] rather than frames. In artificial language experiments, 18-month-old infants and adults track nonadjacent relations (i.e., frames) only when the adjacent relations are least predictable (Gomez, 2002). Twelve-month-old infants do not succeed even under these conditions (Gomez & Maye, 2005). Thus it appears that learning nonadjacent relations may be harder than learning adjacent ones. For this reason it would be surprising but significant if nonadjacent dependencies played a critical role in categorization.

---

[1] Throughout this paper, we use the word "bigram" to refer to the representation of an adjacent relation between two words.

In the present paper, we attempt to answer two important questions about distributional learning: a) Do infants use prosody as a constraint on distributional analysis? and b) Do they preferentially encode frames rather than bigrams for the purpose of syntactic categorization? We based our experiments on a previous study of distributional learning in young infants. Gomez and Lakusta (2004) 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 relevant distributional relations. Our studies used an artificial language similar to the one used by Gomez and Lakusta, with one important change: the sentences in our language were three-word instead of two-word utterances. As explained below, this allowed us to investigate the role of prosody and frames in constraining distributional learning.

First, we manipulated the prosody of the 3-word sentences in the artificial language by inserting a prosodic break either between the first and the second words, or between the second and the third words. Subsequently we evaluated infants' preferences between novel sentences containing the grammatical relation that had previously occurred within a single prosodic unit and those containing the relation that had previously straddled a prosodic boundary. Because infants preferentially encode prosodically grouped units (Nazzi et al., 2000; Soderstrom et al., 2003) we might expect them to preferentially extract the relation that was highlighted by prosody and therefore prefer the former sentence type.

Second, we manipulated the type of test sentences to ask whether infants were representing the 3-word utterances (e.g., aXb) as a single nonadjacent dependency with a middle word (e.g., a___b) or as two adjacent bigrams (aX, Xb). In a previous artificial grammar learning study (Thothathiri & Snedeker, 2005), adult participants heard 3-word sentences where both adjacent and nonadjacent relations predicted the category of the middle word (e.g., aXb, pXq, cYd, rYs). During test, participants successfully discriminated between novel grammatical (aXb, cYd) and ungrammatical sentences (aYb, cXd). However, they found novel sentences that violated nonadjacent relations but respected adjacent ones (e.g., aXq, cYs) to be just as acceptable as the grammatical sentences, suggesting that they were not representing frames. Here we probe what infants will do in a similar situation.

We report three experiments with 12-month-old infants below. To begin, we wanted to determine whether infants can track the relevant relations in 3-word sentences and extend them to test sentences that were prosodically different from the sentences used during training (see Experiment I: Stimuli for details). Thus Experiment I served both as an extended replication of Gomez and Lakusta (2004) and as a precondition for our subsequent experiments. The grammatical test sentences obeyed all possible dependencies and the ungrammatical test sentences violated all of them. Infants therefore could succeed by attending to any one of the several possible dependencies. Having satisfied this necessary precondition (to preview our results), Experiments II and III further explored the role of prosody and different relations (adjacent or non-adjacent) in the learning of distributional categories.

Experiment I

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 due to: program error (5) and fussiness (3).

Stimuli

We adapted the stimuli used by Gomez 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 2 words each; X and Y categories consisted of 6 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).

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. Since 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 & Weenik, 2005). We measured the three characteristics for the first 5 training sentences in each of the four experimental conditions (2 languages and 2 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 to 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 sentences[2] 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.

< Table 1 about here >

< Figure 1 about here >

< Table 2 about here >

[2] We picked the first 3 sentences from each of 8 training lists for a total of 24 sentences, split equally between the two languages and the two prosodies.

Unlike training sentences, test sentences were intended to contain no grouping cues. Figure 1c shows the pitch contour and waveform for an example test sentence. The rightmost column of Table 2 shows the averaged acoustic correlates of the boundaries in 24 test sentences.[3] 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 3 times each. The two grammatical lists were played twice each for a total of 4 grammatical trials, played in 2 blocks such that there was no repetition within a block. The same procedure was used to create 2 ungrammatical lists, again repeated twice across blocks, for a total of 4 ungrammatical trials. All 4 test lists are shown in Table 4.

< Table 3 about here >

< Table 4 about here >

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

[3] We picked the first 3 sentences from each of the 8 test lists used across the three experiments.

other. In this experiment, such a discrimination could be made based on either the relationship between the first word and the middle word (e.g., aX compared to aY), or between the third word and the middle word (e.g., Xb compared to Xd), or both. Alternatively, discrimination could be based on the relationship between frames and middle words (e.g., a___b associated with X or Y).

Aside from our use of 3-word sentences, our stimuli differed from those used by Gomez and Lakusta (2004) in another critical way. The training sentences contained strong prosodic breaks between two of the three words and were thus prosodically different from the test sentences, which did not contain such strong grouping cues. Nevertheless, we expected infants to succeed at this task and listen longer to the novel grammatical sentences just as in the previous study.

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 soundproof booth. Caregivers wore headphones throughout the session. An experimenter controlled presentation of the stimuli 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 2 training and 8 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 seconds. Grammatical and ungrammatical lists alternated for a total of 8 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 and Discussion

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 seconds.[4] This eliminated a total of 6 trials across all participants (4.7%). We found no significant differences between the two training languages, so this variable was excluded from all further analyses. 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 compared to ungrammatical sentences in both prosodic conditions (Figure 2). Twelve out of sixteen infants showed this preference [sign test, $p=0.077$]. There were no other main effects or interactions [F's < 1, p's > 0.4].

[4] Test sentences ranged in length between 1.5 to 2 seconds, so we estimated that this was the minimum time required to decide whether a given test list was grammatical.

These results are consistent with a previous study that demonstrated 12-month-olds' preference for grammatical over ungrammatical artificial language stimuli (Gomez & Lakusta, 2004). As in that study, the results may be interpreted as demonstrating infants' ability to track relationships between frequently occurring context words (a/b/c/d) and less frequently occurring category words (X/Y), as well as the generalization of those relationships to novel X and Y words. In Experiment II, we explored the (possible) influence of prosody on such distributional learning. We contrasted two types of test sentences, neither of which was completely grammatical. One type of test sentence preserved only the relation highlighted by the prosody during the training phase; the other preserved only the cross-boundary relation. Will infants preferentially encode or extract the within-prosodic-unit dependency and therefore prefer the former sentence type?

<Figure 2 about here>

Experiment II

Participants

Twenty new infants participated (five in each experimental condition). The mean age was 12 months and 16 days (range: 11; 18 to 13; 18). Thirteen of the infants were female. Five other infants were excluded due to fussiness.

Stimuli

Training stimuli were the same as in Experiment I (aXb/cYd in language 1; aYb/cXd in language 2; produced in one of two prosodies). Test stimuli were one of two types, neither of which was completely grammatical. The first type (aXd/cYb) contained a correct 1st bigram and an incorrect 2nd bigram. The second (aYd/cXb) contained an incorrect 1st bigram and a correct 2nd bigram (as always, the converse was true for language 2). See Table 5 for a complete list of test stimuli.

< Table 5 about here >

Procedure

We used the same procedure as in Experiment I. The two types of test lists (1st bigram correct and 2nd bigram correct) alternated for a total of 8 trials.

Results and Discussion

We analyzed infants' listening times to the two types of test trials. A total of three trials were excluded because infants listened for less than 2 seconds (1.9%). As before, the results were collapsed across the two training languages. A 2x2 ANOVA (Type x Prosody) revealed no main effects or interactions (F's < 3, p's > 0.1). Infants did not listen significantly longer to either test type and there was no difference between the two prosodic conditions (Figure 3).

The rationale behind Experiment II was as follows: if prosody constrains distributional analysis, we would expect that relations within a prosodic unit would be better learned than those that straddle a prosodic boundary. This should translate into a preference for a correct 1st bigram over a correct 2nd bigram in the Prosody 2 conditions and a preference for a correct 2nd bigram over a correct 1st bigram in the Prosody 1 conditions. In the analysis above this preference would appear as an interaction between prosody and test type. We failed to find such an interaction. This suggests that prosody does not constrain the learning of abstract distributional relations which can be generalized to novel utterances (at least under these conditions). The null finding forms an interesting contrast with previous positive findings of the effect of prosody in tasks that do not test generalization. We discuss the implications under General Discussion.

The pattern of findings in these two experiments (a clear preference in Experiment I and a lack of preference in Experiment II) is consistent with two alternate accounts of distributional learning. First, this pattern of performance could reflect distributional learning on the basis of bigrams. In Experiment I infants were given a choice between a sentence that contains two correct bigrams and another that contains no correct bigrams. In contrast, the comparison in Experiment II involves two sentences both of which contain one correct and one incorrect bigram. Thus a bigram learner (with no recourse to prosody) would have no basis for preferring one to the other.

However, this pattern of performance is equally compatible with the proposal that children use frequent frames to guide distributional analysis (Childers & Tomasello, 2001; Mintz, 2003). In Experiment I infants were given a choice between a sentence that contains the appropriate frame and another that contains an inappropriate frame. In contrast, the comparison in Experiment II involves two sentences neither of which contains a frame that the infant has heard before. Thus a frame based learner would also have no basis on which to judge the test sentences.

In Experiment III, we tested the frames hypothesis by contrasting sentences with a correct frame and 2 correct bigrams with those that contain an incorrect frame and 1 correct bigram. If infants are encoding frames and their relations to the middle words, we would expect a strong preference for the former sentence type. In fact, the effect should be just as strong as that found in Experiment I (which also contrasted correct and incorrect frame sentences). If infants are encoding multiple bigrams however, the predictions are less clear cut since they depend on the specifics of how well children discriminate between utterances with two, one or no familiar bigrams. Critically, only the latter hypothesis is consistent with finding a smaller preference for the grammatical sentences in Experiment III compared to Experiment I.

<Figure 3 about here>

Experiment III

Participants

Sixteen infants participated (four in each experimental condition). The mean age was 12 months and 14 days (range: 11; 17 to 13; 27). Nine of the infants were female. One other infant was excluded due to fussiness.

Stimuli

The training stimuli were the same as in Experiments I and II. During test, we contrasted sentences that were not fully grammatical (incorrect frame and 1 correct bigram: aXd/cYb for both languages) with those that were completely grammatical (correct frame and 2 correct bigrams: aXb/cYd for language 1; aYb/cXd for language 2). For infants exposed to language 1, the sentences of the first type had a correct 1st bigram and an incorrect 2nd bigram. For infants exposed to language 2, they had an incorrect 1st bigram and a correct 2nd bigram. Test stimuli for language 1 are shown in Table 6.

< Table 6 about here >

Procedure

We used the same procedure as in previous experiments. The two sentence types alternated for a total of 8 test trials.

Results and Discussion

A total of six trials were excluded due to listening times less than 2 seconds (4.7%). We performed a 2x2 ANOVA (Type x Prosody) over listening times for the remainder of the trials. There were no significant main effects or interactions [F's < 2, p's > 0.2]. Infants preferred neither sentence type; prosody did not seem to influence their preference (Figure 4). Because the grammar violation was different for the two languages in Experiment III (incorrect 2^{nd} bigram for L1 and incorrect 1^{st} bigram for L2), we analyzed the two languages separately. We found no effects for either language [F's < 2, p's > 0.2]. This suggests that infants did not discriminate between fully grammatical sentences and those that contain either an incorrect 1^{st} or 2^{nd} bigram. This contrasts with infants' successful discrimination between fully grammatical and fully ungrammatical sentences in Experiment I. Combining data from Experiments I and III, a 2x2x2 ANOVA (Experiment x Type x Prosody) revealed a significant interaction between Experiment and Type [$F(1, 28) = 8.851$, $p<0.01$].

The lack of preference for the fully grammatical sentence type in Experiment III (but not in Experiment I) is informative in two ways. First, it suggests that infants' preferences are not based on nonadjacent dependencies (i.e., frames). In Experiment III, the fully grammatical test sentences contained correct frames while the other sentences did not. Nevertheless, infants showed no preference for the former type. Second, the results suggest that infants were tracking both adjacent dependencies in the 3-word sentences because we obtained a robust preference when we presented a contrast of 0 versus 2 correct bigrams (Experiment I) but no preference when we contrasted 1 versus 2 correct bigrams (Experiment III).[5]

<Figure 4 about here>

General Discussion

We now return to the two questions that motivated this research, namely the role that frames and prosody each play in constraining infants' distributional learning. Below we tackle each issue in turn.

Frames vs. Bigrams

In Experiment I, we found that infants learned the distributional relations in 3-word utterances. They successfully discriminated between novel grammatical and ungrammatical sentences. Experiment III ruled out the possibility that infants' success in the first experiment (and their failure in the second) was due solely to their representation of frames and the relations between frames and middle words. Infants did not discriminate between fully grammatical and partially grammatical sentences even though the latter contained invalid frames. Together, our results suggest that infants were evaluating the adjacent rather than the nonadjacent relations in 3-word sentences to judge

[5] The results are also consistent with each infant tracking a single adjacent dependency but with variation among the infants in which dependency is tracked.

the grammaticality of novel utterances.[6] However, these results do not imply that infants never use frames for categorization. Successful categorization may well involve the representation of frames when the numbers and frequencies of context and middle words are vastly different (Gomez, 2002). This may be the case in some parts of natural language e.g., when different verbs are surrounded by high-frequency function words, as in "*is* ***danc***-*ing*", "*is* ***sing***-*ing*", etc. What our results do suggest however is that infants can succeed using an alternate route, namely, by tracking multiple adjacent relations. Two questions regarding the use of adjacent and non-adjacent dependencies remain to be answered. First, will infants switch to using frames when the variability of middle words is increased? Gomez (2002) demonstrated that infants can *track* nonadjacent dependencies under these conditions. But that study used the same middle words in all the frames. Thus, it remains to be seen whether infants can *use* frame representations to categorize middle words. A variant of the paradigm used here, with a greater number of X and Y words, could be used to address this question. A second remaining issue is what children will do in the absence of predictive phonological cues (like the one-syllable, two-syllable distinction used in the present study). Several prior studies with adult learners have demonstrated the value of having correlated distributional and semantic/phonological cues (e.g., Braine, 1987). In addition, Gomez & Lakusta (2004) review evidence for 12-month-olds' failure to categorize when there were no distinguishing phonological cues. But all of these studies investigated gender categories that involved relations between stems and case endings. In other words, the training

---

[6] Alternatively, it is possible that infants were using *both* adjacent and non-adjacent dependencies. However, the pattern of results (success in discriminating between two sentence types with correct frames and 0 or 2 correct bigrams, and failure to discriminate between two sentence types with different frames and 1 or 2 correct bigrams) suggests, at the very least, that adjacent dependencies play a greater role in infants' discrimination.

stimuli were *two*-morpheme utterances. It remains to be seen whether categories in *three* -word utterances are more learnable based on distributional information alone (see Mintz, 2002). This is plausible because learners potentially have two additional sources of information in three-word compared to two-word utterances. First, they may represent nonadjacent dependencies or frames (Mintz, 2003). Second, they may use correlations between multiple distributional relations (Maratsos & Chalkley, 1980; Mintz, 2002). This issue can be explored by extending the current paradigm to test generalization to cases where there are no phonological cues to category membership.

The contrast between the first and third experiments warrants further discussion. Infants successfully discriminated fully grammatical sentences from fully ungrammatical sentences but not from partially ungrammatical sentences. We can think of three possible reasons for the latter null result. One reason is lack of statistical power. This seems unlikely because there is no numerical difference between grammatical and partially grammatical sentences in Experiment III. In fact, the preference goes in the opposite direction (mean listening time = 12.33 s for partially grammatical and 9.97 s for fully grammatical sentences). A second possibility is non-linearity in the mapping between the dependent measure and grammaticality. Listening times may not increase or decrease monotonically with increasing or decreasing grammatical status. A third possibility is that infants use their knowledge of previously encountered relations (or bigrams) to parse utterances and discover new relations between words. This could explain infants' (relative) interest in partially ungrammatical sentences that contained 1 previously encountered relation and their (relative) non-interest in fully ungrammatical sentences that contained 0 familiar relations. This suggestion is akin to previous proposals that infants may use the sentence fragments found in infant-directed speech (e.g., ***A ball)*** to parse longer utterances containing that fragment (e.g., *This is* ***a ball***) (Fisher & Tokura, 1996). Further research is required to evaluate this speculative hypothesis.

The role of prosody

Given the extensive literature on infants' sensitivity to prosodic grouping, our failure to find any effect of prosody is surprising. In Experiment II, infants did not distinguish between test sentences that contained a relation that was highlighted by the prosody used during training and those that contained a relation that was not highlighted. There are three possible explanations for this null result. First, our prosody manipulation may have been ineffective, i.e., infants may not have perceived the training sentences as expected. This seems unlikely given the magnitude of the acoustic differences between the utterances (see Table 2) and the fact that two adult listeners (1 trained and 1 untrained) perceived the training sentences as we intended. There is evidence for continuity in the processing of prosodic cues throughout ontogeny (Pannekamp, Weber & Friederici, 2006). A second possibility is that infants perceived the training sentences as expected, but simply did not need to constrain their distributional analysis. The hypothesis tested here was that prosody may aid grammar acquisition by chopping utterances into structurally relevant chunks. This may benefit young language learners who have memory or processing limitations compared to adults (Newport, 1988). We began testing the hypothesis with simple 3-word utterances because they were minimally different from the stimuli used in previous studies and we wanted to be sure that infants could succeed at the task. But our choice of short stimuli may have had a disadvantageous effect. The memory or processing load may have been simply inadequate for prosody to be relevant.

Future studies that use longer utterances may have a better chance at finding an effect of prosody. The third possible reason for the null effect is related to the second. A main hypothesized role for prosody is to ease computational burden. Thus, it is possible that infants use this cue only to constrain the analysis of dependencies that are potentially numerous and not obvious. Prosodic cues may not be useful for constraining the analysis of adjacent dependencies because these are already constrained (e.g., consider one word before and after each word). In contrast, the analysis of nonadjacent dependencies may benefit more from a prosodic constraint because it is unclear which nonadjacent dependencies one should be tracking (the relevant dependencies may be separated by varying lengths of intervening material). Extending the current paradigm to test the acquisition of nonadjacent dependencies can help evaluate this possibility.

Conclusions

The results of this study suggest that 12-month-old infants can track multiple adjacent distributional relations and generalize them to novel utterances. Under the conditions tested here, prosody does not appear to constrain infants' distributional analysis. Further research is required to elucidate the precise contribution of prosody and frames to the distributional learning of categories.

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

Boersma, P., & Weenik, D. (2005). Praat: doing phonetics by computer (version 4.3.28) [Computer program]. Available from //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.

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology, 37,* 739–748.

Fisher, C., & Tokura, H. (1996). Prosody in speech to infants: Direct and indirect cues to syntactic structure. In J. Morgan, & C. Demuth (Eds.), Signal to syntax (pp. 343–364). Mahwah, NJ: Erlbaum.

Frigo, L., & McDonald, J.L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language*, *39*, 218–245.

Gerken, L.-A., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: Nine month olds' sensitivity to phonological versus syntactic phrases. *Cognition*, *51*, 237–265.

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., & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science, 7*(5), 567-580.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13,* 431–436.

Gómez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, *7*(2), 183-206.

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., & 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.

Jusczyk, P. W. (1998). Dividing and conquering the linguistic input. In M. C. Gruber, D. Higgins, K. Olson, & T. Wysocki (Eds.), CLS 34, Vol.2: The panels (pp. 293–310). Chicago: University of Chicago.

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.

Kelly, M. H. (1992) Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*, 349-364.

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 and Development, 18, 111–116.

Maratsos, M. & Chalkley, M. A. (1980) The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (Ed.), *Children's language* (Vol. 2, pp. 127-214). Gardner.

Monaghan, P. & Christiansen, M.H. (2004). What distributional information is useful and usable for language acquisition? In *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (pp. 963-968). Mahwah, NJ: Lawrence Erlbaum.

Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, *19*, 498–550.

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.

Mintz, T. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition, 90,* 91–117.

Mintz, T. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition, 30,* 678–686.

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.

Newport, E. L. (1988). Constraints on learning and their role in language acquisition: Studies of the acquisition of American Sign Language. *Language Sciences, 10,* 147–172.

Pannekamp, A., Weber, C., & Friederici A. D. (2006). Prosodic processing at the sentence level in infants. *Neuroreport, 17*(6), pp. 675-678.

Pinker, S. (1984) *Language learnability and Language Development.* Harvard University Press.

Shi, R., Werker, J.F., & Morgan, J.L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*, B11–B21.

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.

Thothathiri, M. and Snedeker, J. (2005). Distributional learning of syntactic categories. *Poster presented at the Society for Research in Child Development Biennial Meeting*, Atlanta, GA.

Valian, V., & Coulson, S. (1988). Anchor points in language learning: the role of marker frequency. *Journal of Memory and Language*, *27*, 71–86.

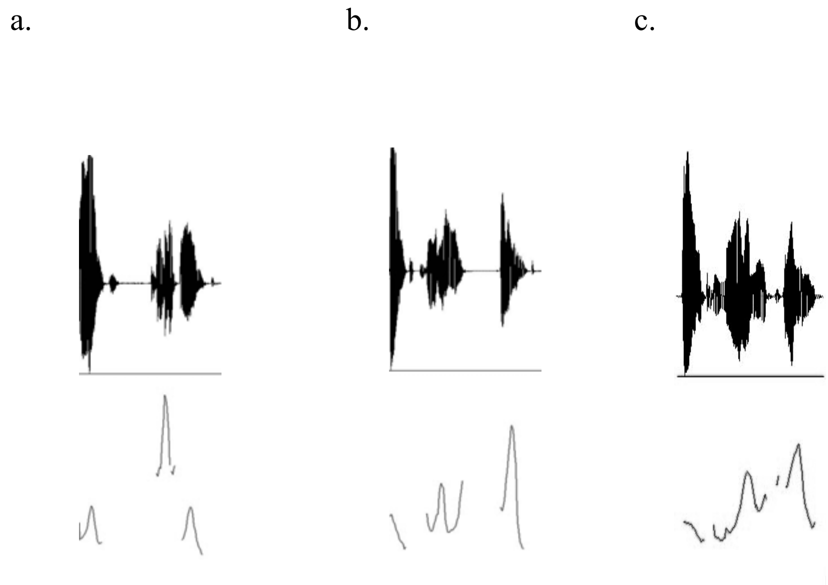

Figure 1. Waveforms and pitch contours of example sentences a. Training (*alt coomo omp*) in Prosody 1; b. Training (*alt coomo omp*) in Prosody 2; c. Test (*alt nawlup omp*)

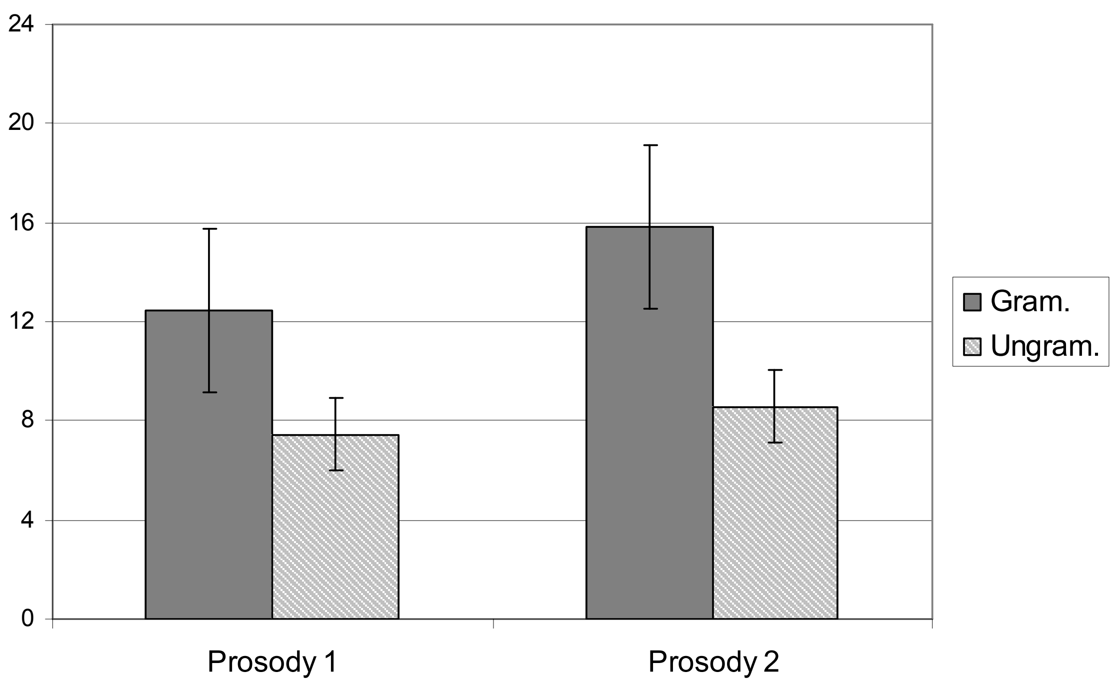

Figure 2. Experiment I: Listening times for grammatical and ungrammatical sentences (in seconds). Infants preferred grammatical sentences in both prosodic conditions.

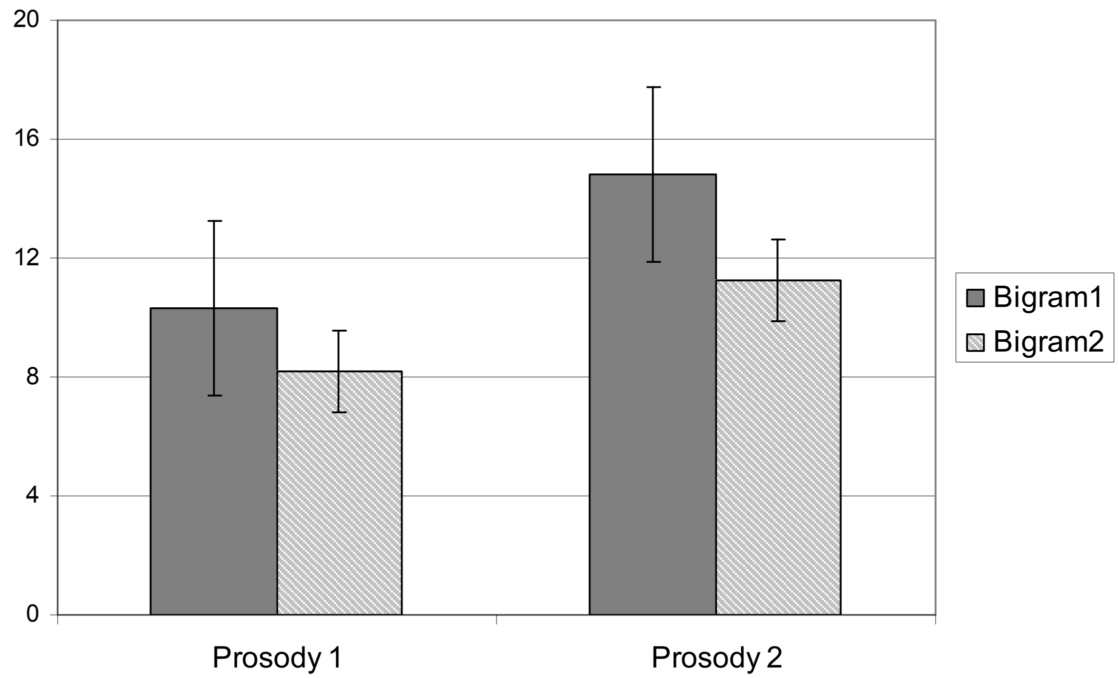

Figure 3. Experiment II: Listening times for two types of test sentences (1st bigram correct or 2nd bigram correct) (in seconds). Infants showed no preference. There was no difference between prosodic conditions.

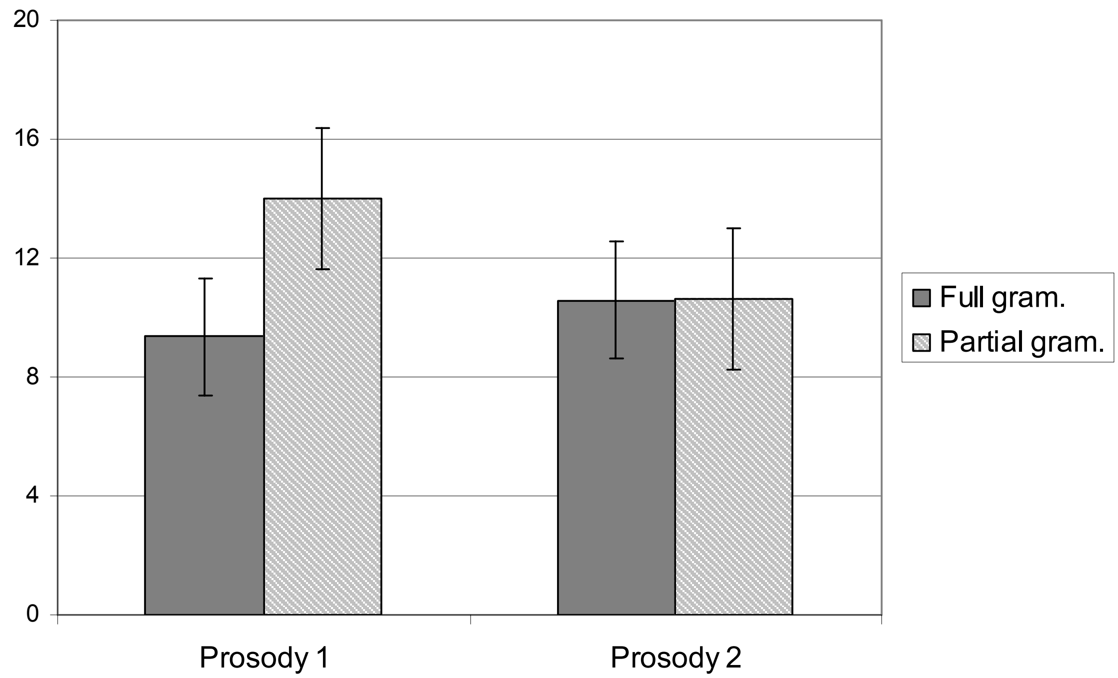

Figure 4. Experiment III: Listening times for two sentence types (fully grammatical and not fully grammatical) (in seconds). Infants showed no preference.

Table 1: Words used during training. 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.

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

Table 2. Acoustic correlates of prosody in training and test sentences. The figures in parentheses are standard deviations.

Acoustic correlate	Sentence position	Prosody 1	Prosody 2	Test
Final syllable rhyme duration (sec)	1st word	0.687 (0.004)	0.379 (0.005)	0.362 (0.031)
	2nd word	0.286 (0.087)	0.447 (0.173)	0.351 (0.083)
Pitch change (Hz)	Between 1st and 2nd words	99.92 (6.3)	56.19 (17.15)	35.91 (28.61)
	Between 2nd and 3rd words	47.02 (37.65)	106.07 (14.59)	48.8 (43.35)
Pause duration (sec)	Between 1st and 2nd words	0.51 (0.006)	0.085 (0.021)	0.084 (0.067)
	Between 2nd and 3rd words	0.089 (0.001)	0.511 (0.017)	0.116 (0.051)

Table 3: Novel X and Y words used during test. The other words (a, b, c and d) were the same as in training.

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

Table 4: Test stimuli for Experiment I. These sentences either preserved or violated the dependencies found in training sentences (e.g., compare to L1 training: *alt coomo omp* / *erd deech ulf* and L2 training: *alt deech omp / erd coomo ulf*).

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

Table 5. Test stimuli for Experiment II. These sentences contained either a correct 1st bigram or a correct 2nd bigram (e.g., compare to L1 training: *alt coomo omp* / *erd deech ulf* and L2 training: *alt deech omp / erd coomo ulf*).

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

Table 6. Test stimuli for Experiment III (language 1 only). These sentences contained either one correct bigram (and an incorrect frame) or two correct bigrams (and thus a correct frame) (e.g., compare to L1 training: *alt coomo omp* / *erd deech ulf*).

Type 1 (fully grammatical: 2 correct bigrams, correct frame)		Type 2 (not fully grammatical: 1 correct bigram, incorrect frame)	
List 1	List 2	List 1	List 2
erd vot ulf	erd foge ulf	erd vot enk	erd foge enk
ush meeper omp	alt nawlup omp	ush meeper ast	alt nawlup ast
ong tood ulf	ush roosa enk	ong tood enk	ush roosa ulf
alt roosa omp	ong biff ulf	alt roosa ast	ong biff enk
erd rud ast	ush gackle enk	erd rud omp	ush gackle ulf
alt gackle omp	erd pel ast	alt gackle ast	erd pel omp
ush binow enk	ong vot ast	ush binow ulf	ong vot omp
ong foge ast	alt meeper enk	ong foge omp	alt meeper ulf
ong pel ulf	ong rud ulf	ong pel enk	ong rud enk
alt bevit enk	ush bevit omp	alt bevit ulf	ush bevit ast
ush nawlup enk	erd tood ast	ush nawlup ulf	erd tood omp
erd biff ast	alt binow omp	erd biff omp	alt binow ast