Available online at www.sciencedirect.com
SCIENCE DIRECT
COGNITION
Cognition 93 (2004) B89–B96
www.elsevier.com/locate/COGNIT

Brief article

Nine-month-olds extract structural principles required for natural language

LouAnn Gerken

Department of Psychology, University of Arizona, Tucson, AZ 85721, USA

Received 21 July 2003; revised 5 November 2003; accepted 25 November 2003

Abstract

Infants' ability to rapidly extract properties of language-like systems during brief laboratory exposures has been taken as evidence about the innate linguistic state of humans. However, previous studies have focused on structural properties that are not central to descriptions of natural language. In the current study, infants were exposed to 3- and 5-syllable words from one of the two artificial languages that employed the same stress assignment constraints found in natural languages. Infants were able to generalize beyond the stress patterns encountered during familiarization to new patterns reflecting the same constraints. The results suggest that infants are able to rapidly extract the types of structural information required for human language.

Keywords: Learner; Infants; Metrical phonology

Recent studies reveal that infants are able to rapidly extract many properties of language-like input (Gerken, 2002; Gómez, 2002; Naigles, 2002; Saffran, Aslin, & Newport, 1996). However, none of these studies has examined structural properties of the type central to discussions of linguistic grammars (Gerken, 2002). For example, two studies have examined artificial grammars that are distinguishable based on patterns of repeating or alternating syllables (Gómez & Gerken, 1999; Marcus, Vijayan, Rao, & Vishton, 1999). Although detecting repetition is important in languages with reduplication, it is also important in other domains, such as music.[1,2] Further, reduplicative processes in language

E-mail address: gerken@u.arizona.edu (L.A. Gerken).

[1] E.g. Try to imagine the opening measure of Beethoven's Fifth with the first note played just twice.

[2] Marcus et al. (1999) suggest that the ability to recognize identity between elements in a string like "wo fe wo" is akin to recognizing that a sentence can comprise a noun–verb–noun sequence. However, identity is defined over tokens in the former example and over types in the latter, making learning in the two situations computationally distinct (Gómez & Gerken, 2000).

doi:10.1016/j.cognition.2003.11.005

often show complex interactions with segmental and prosodic structure (e.g. McCarthy & Prince, 1996). Thus, existing studies have shown sensitivity to the building blocks of human language but have not investigated learning of formal systems with comparable complexity.

The goal of the current research was to examine infants' ability to learn a formal system central to discussions of human language. The domain chosen was metrical phonology, which concerns structural principles for assigning stress to syllables in multi-syllabic words (Demuth, 1996; Dresher, 1999; Dresher & Kaye, 1990; Gupta & Touretzky, 1994; Hayes, 1994; Prince & Smolensky, 1997; Tesar & Smolensky, 1998). The metrical input to the learner is words with different patterns of strong (stressed) and weak (unstressed) syllables. In many languages, words with the same number of syllables can exhibit different stress patterns, based on a set of linguistic principles (Dresher, 1999). The current study adopts an Optimality Theory approach to stress assignment, within which ranked stress assignment constraints can conflict in their application to a particular word (Prince & Smolensky, 1997). When two constraints conflict, only the more highly ranked applies.

Within this framework, linguistic generalization requires learners to determine the constraints and rankings that account for the surface stress patterns in the input. The main question addressed in the current investigation is whether 9-month-olds show evidence of extracting such information after a very brief exposure.

1. Experiment 1

Infants were familiarized with a word list generated by one of the two languages (L1 & L2), differing in two of the four ranked stress assignment constraints (see Table 1). Both languages employed constraints A and B. L1 employed C^1 and D^1, while L2 employed C^2 and D^2. Familiarization items in each language provided evidence for three rankings of these constraints, shown in the third column of Table 1. For example, constraint A states that two stressed syllables cannot occur in sequence, while B states that heavy syllables should be stressed. The L1 familiarization word "TON ton do RE mi" provides evidence for the ranking $A \gg B$, because the second syllable should be stressed under constraint B, which if obeyed would lead to two stressed syllables in sequence, thereby violating constraint A. Since, in this example, constraint B is violated and A is not, a learner should infer that $A \gg B$.

Familiarization stimuli never provided evidence for a fourth ranking, $A \gg D$ ($A \gg D^1$ in L1 and $A \gg D^2$ in L2), however, this ranking is logically inferable from the three for which evidence was provided during familiarization. The $A \gg D$ ranking was used to create the test stimuli. Two types of test stimuli from each language were employed. They were L1 and L2 *abstraction test items*, which were consistent with the $A \gg D$ ranking from each language. These test items exhibited two important properties: first, they had entirely different patterns of stressed and unstressed syllables than did the familiarization items from either L1 or L2. Second, L1 and L2 abstraction test items had the same stress pattern, differing only in the location of the heavy syllable. Therefore, a significant familiarization language by test string interaction would provide evidence that infants are able to generalize beyond stress patterns encountered during familiarization to the abstract system underlying these patterns. Although it is difficult *a priori* to predict

Table 1
Sample familiarization and test stimuli used in Language 1 and Language 2 and stress constraint rankings attested

L1 familiarization	L2 familiarization	Stress constraint ranking attested[a]
TON ton do RE mi	do RE mi ton TON	A ≫ B
TON do re	do re TON	B ≫ C^1(L1), B ≫ C^2 (L2)
DO re TON	TON do RE	B ≫ C^1(L1), B ≫ C^2 (L2)
DO re TON mi fa	do re TON mi FA	B ≫ C^1(L1), B ≫ C^2 (L2)
DO re mi FA so	do RE mi fa SO	C^1 ≫ D^1 (L1), C^2 ≫ D^2 (L2)
L1 abstraction test	L2 abstraction test	Stress constraint ranking attested
do TON re MI fa	do RE mi TON fa	A ≫ D^1 (L1), A ≫ D^2 (L2)
L1 stress pattern test[b]	L2 stress pattern test	
DO re mi TON fa (Experiment 1)	do TON re mi FA (Experiment 1)	A ≫ D^1 (L1), A ≫ D^2 (L2)
MI re TON (Experiment 2)	TON mi RE (Experiment 2)	A ≫ ^{1} (L1), A ≫ D^2 (L2)

Syllables in upper case are stressed, and those in lower case are unstressed. Seven variants of each familiarization and test word were created using the 7 solfège syllable (do, re, mi, fa, so, la, ti) and substituting “re” for “do”, “mi” for “re”, etc. in the sample stimuli. No substitutions were made for the syllable “ton.”

[a] The following four stress constraints were used in the study (“ ≫ ” means “outranks”). (A) Two stressed syllables cannot occur in sequence. (B) Heavy syllables (those ending in a consonant) should be stressed, (C^1/C^2) syllables should be stressed if they are second to last (L1)/second (L2), and(D^1/D^2) alternating syllables should be stressed, starting from the left (L1)/right (L2).

[b] The stress pattern test items have the same stress patterns as the 5th and 3rd familiarization word in Experiments. 1 and 2, respectively.

the direction of preference in familiarization studies, several such studies have revealed a novelty preference—longer listening times to test strings from the unfamiliar language (e.g. Marcus et al., 1999; Saffran et al., 1996).

There were also L1 and L2 *stress pattern test items*, which were not only consistent with the A >> D ranking in the relevant familiarization language, but also exhibited the same stress pattern found in 20% of the familiarization words in that language. These test items were included to compare infants’ performance with that found in a previous study with adults using similar materials to those used in Experiment 1 (Guest, Dell, & Cole, 2000). When adults produced abstraction vs. stress pattern test words, they showed evidence of learning only for the latter items. In another experiment, they also showed weak evidence of extracting the A >> D ranking, but only when presented with words containing differing numbers of syllables in training vs. test. That is, adults appeared to extract both stress pattern information and the unattested ranking, but were most influenced by the former. Therefore, a secondary question in Experiment 1 was whether infants would show sensitivity to stress patterns instead of ranked stress assignment constraints, in which case, they should discriminate only L1 and L2 stress pattern test items.

1.1. Methods

1.1.1. Materials

L1 and L2 each generated seven tokens of each of the five familiarization word types (see Table 1). The 35 words in each language were recorded by the experimenter using

SoundEdit 16 on a Power Macintosh. In order to create lists of 1.5 min in duration, 24 words were selected roughly equally from the five word types and copied to the end of the familiarization list. The final familiarization lists comprised 59 randomly ordered words with 500 ms pauses between. Test items lists comprised seven tokens of each of the four test word types (L1 Abstraction, L1 Stress Pattern, L2 Abstraction, L2 Stress Pattern), each in two random orders, yielding eight lists. Each test item list was approximately 20 s in duration and was presented twice, once in each half of the experiment, for a total of 16 test trials.

1.1.2. Participants

Data were included from 18 infants (mean age 9 months 0 days) from English-speaking homes with no history of hearing or speech/language disorder. Half were familiarized with L1 and half with L2. An additional 15 infants were tested but failed to provide useable data for at least 12 test trials (14), or had overall looking times greater than 2 SDs from the group mean (1).

1.1.3. Procedure

Each infant sat on a caregiver's lap in a sound proof booth with an amber light in front of the infant and two red lights over speakers to each side. During familiarization, the entire word list (L1 or L2) was played from both speakers, and the light under one speaker flashed. After familiarization, the infant participated in 16 test trials. A trial began when the infant oriented to the flashing center light. One of the sidelights would then begin to flash, and when the infant turned toward the flashing light, the test trial would be played from the corresponding speaker. The trial lasted until the infant looked away from the light for 2 s. Looking times shorter than 2 s were excluded from the analyses.

1.2. Results and discussion

Looking times were subjected to a two familiarization language (L1 vs. L2) × 2 test language (L1 vs. L2) × 2 test word type (abstraction vs. stress pattern) ANOVA (Fig. 1), which revealed significant familiarization language × test language ($F(1,16) = 7.78$, $P < 0.02$) and three-way interactions ($F(1,16) = 4.90$, $P < 0.05$.) The three-way interaction indicates that infants responded differently on the two types of test items (abstraction vs. stress pattern). Therefore, their responses for each type of test item were examined separately. A 2 familiarization language × 2 test language ANOVA on the abstraction test items revealed only a significant interaction ($F(1,16) = 12.86$, $P < 0.01$). An examination of Fig. 1 reveals that, infants listened longer to test strings from the unfamiliar language ($t(17) = 3.15$, $P < 0.01$). These data suggest that infants can engage in the type of abstraction required for natural language.

In contrast, an ANOVA on stress pattern trials revealed no learning, but only an effect of test language ($F(1,16) = 4.85$, $P < 0.05$), with infants from both familiarization languages listening longer to L1 test items. The L1 stress pattern test words began in a stressed syllable, a pattern that is consistent with a majority of English words and shown in previous studies to be preferred by English-learning infants of the age studied here

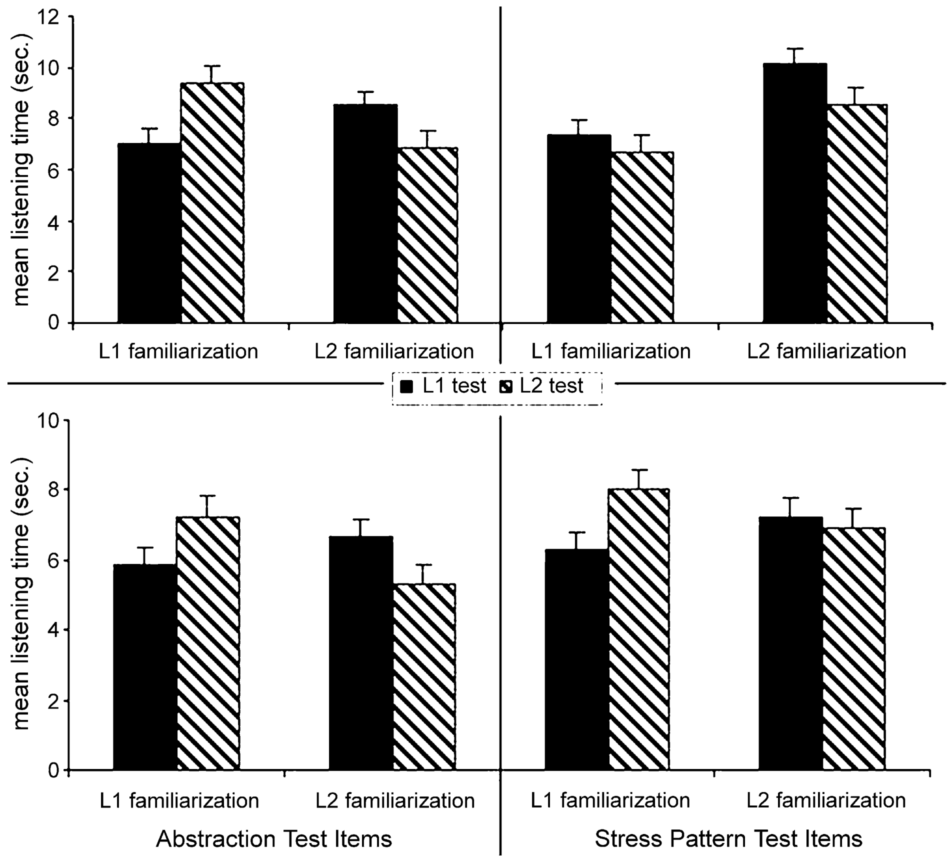

Fig. 1. Mean listening times and standard error bars in each condition of Experiments 1 (top panels) and 2 (bottom panels).

(Jusczyk, Cutler, & Redanz, 1993). Because infants' response to the stress pattern test trials in Experiment 1 was based on extra-experimental factors, we cannot determine whether they encoded stress pattern information in addition to inferring the A $\gg$ D ranking.

2. Experiment 2

The main finding of Experiment 1 was that infants, unlike adults in a previous study, were able to rapidly generalize to new words with new stress patterns that reflected an abstract language system. Given the potential importance of this effect for theories of human language, one goal of Experiment 2 was to confirm it with a replication. A secondary goal was to attempt again to determine if, in addition to generalization based on highly abstract properties of the input (the unattested A $\gg$ D ranking), infants also generalize based on stress pattern.

2.1. Methods

2.1.1. Materials

To replicate the main finding of Experiment 1, abstraction test items remained the same. To explore the secondary hypothesis that infants extract both stress constraints and rankings and stress pattern information, new stress pattern test items were created. Stress pattern test items had the same stress pattern as 20% of the familiarization stimuli in each familiarization language. Importantly, L1 and L2 stress pattern test items had identical stress patterns but differed in the location of the heavy syllable. If infants attend to stress pattern in addition to more abstract information, both L1 and L2 stress pattern test trials should have been familiar regardless of familiarization language, thereby leading to decreased discriminability of stress pattern test items compared with abstraction test items.

2.1.2. Participants

Data were included from 18 infants (mean age 9 months 3 days) from English-speaking homes with no history of hearing or speech/language disorder. Half were familiarized with L1 and half with L2. An additional 13 infants were tested but failed to provide useable data for at least 12 test trials (11), or had overall looking times greater than 2 SDs from the group mean (2).

2.1.3. Procedure

Familiarization and test procedures were the same as those in Experiment 1.

2.2. Results and discussion

A three-way analysis of variance again revealed a significant familiarization language × test language interaction ($F(1,16) = 9.97$, $P < 0.01$) such that infants responded differently to test stimuli, depending on their familiarization language. This interaction indicates learning occurred during familiarization, and, as in Experiment 1, infants listened longer to test strings from the unfamiliar language ($t(17) = 3.10$, $P < 0.01$). Unlike in Experiment 1, the three-way interaction involving test type (abstraction vs. stress pattern) was not significant ($F < 1$), suggesting that learning occurred for both abstraction and stress pattern items. However, an examination of Fig. 1 suggests that infants' preference for test strings from the language on which they were not familiarized is more robust for abstraction test items. A two-way ANOVA performed on infants' listening times to the abstraction test items revealed the predicted interaction of familiarization language and test language ($F(1,16) = 9.09$, $P < 0.01$). Thus, Experiment 2 confirms the finding from Experiment 1 that infants can very rapidly abstract linguistic structure and generalize to new items with the same structure. The comparable ANOVA on looking times to stress pattern test items revealed only a marginal interaction ($F(1,16) = 2.61$, $P < 0.13$). The marginal effect of learning for stress pattern test trials contrasts with the strong effect for abstraction test trials and is consistent with the notion that higher and lower level generalizations competed in the stress pattern test items. However, the lack of a significant three-way interaction means that such a conclusion must remain tentative.

3. Summary and conclusion

The data from the abstraction test items in both experiments clearly demonstrate that 9-month-old human infants are able to generalize beyond particular word stress patterns to the abstract system that generates these patterns. This is the first study to demonstrate such rapid generalization of the type required for human language.

One potential barrier to this conclusion is that the stressed syllable TON appeared on average in serial position 2 (of 5) in L1 familiarization items and position 3 in L2 items. TON appeared in positions 2 and 4 in L1 and L2 abstraction test items, respectively, raising the possibility that infants responded based on the serial position of TON, rather than constraint rankings.[3] However, this interpretation of the data is made less likely when we consider that TON appeared in positions 3 and 1 in L1 and L2 stress pattern test items, respectively, in Experiment 2. That is, TON appeared earlier in L1 than L2 for abstraction test items but later in L1 than L2 in stress pattern items. Nevertheless, infants responded similarly to both types of test items in Experiment 2, suggesting that an effect outside of serial position governed their responses.

The data from the current studies raise at least two questions. First, are infants more likely than adults to abstract beyond stress patterns to more abstract principles? A comparison of the current data to adult data reported by Guest et al. (2000) are suggestive. However, differences in testing conditions (passive listening vs. language production) may cause learners to emphasize different aspects of their input. Additional studies are needed in which adults are tested under conditions more similar to the infants in Experiments 1 and 2.

Second, are the generalizations they make consistent with characterizations of Universal Grammar (UG), or are they better thought of in other terms? For example, the only heavy syllable used in the current studies was TON. Would infants generalize to other heavy syllables, thereby giving evidence for a UG parameter setting "heavy syllables should be stressed" (Dresher, 1999)? Or, is the principle extracted by infants in the current experiment more local: "TON should be stressed"? This question is currently being addressed by new experiments. The results will help us to understand how far beyond their language input infants are able to generalize.

Acknowledgements

I thank T. Bever, A. Carter, G. Dell, R. Gómez, M. Hammond, P. Neher and M. Piattelli-Palmerini for comments and suggestions on earlier drafts, and D. Guest for help with

[3] I thank an anonymous reviewer for suggesting this analysis.

the experimental materials. This research was funded by NIH grant R01HD42170 to R. Gómez, L.A. Gerken and E. Plante.

References

Demuth, K. (1996). The prosodic structure of early words. In J. L. Morgan, & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 171–184). Hillsdale, NJ: Lawrence Erlbaum.

Dresher, B. E. (1999). Child phonology, learnability, and phonological theory. In W. C. Ritchie, & T. K. Bhatia (Eds.), *Handbook of child language acquisition* (pp. 299–346). San Diego: Academic Press.

Dresher, B. E., & Kaye, J. D. (1990). A computational learning model for metrical phonology. *Cognition*, *34*(2), 137–195.

Gerken, L. A. (2002). Early sensitivity to linguistic form. *Annual Review of Language Acquisition*, *2*, 1–36.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436.

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135.

Gómez, R. L., & Gerken, L. A. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*, 178–186.

Guest, D. J., Dell, G. S., & Cole, J. S. (2000). Violable constraints in language production: Testing the transitivity assumption of optimal theory. *Journal of Memory and Language*, *42*(2), 272–299.

Gupta, P., & Touretzky, D. S. (1994). Connectionist models and linguistic theory: Investigations of stress systems in language. *Cognitive Science*, *18*(1), 1–50.

Hayes, B. (1994). *Metrical stress theory*. Chicago: University of Chicago Press.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*(3), 675–687.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*(5398), 77–80.

McCarthy, J., & Prince, A. (1996). *Prosodic morphology 1986*. Rutgers, NJ: Rutgers University Center for Cognitive Science.

Naigles, L. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition*, *86*, 157–199.

Prince, A., & Smolensky, P. (1997). Optimality: From neural networks to universal grammar. *Science*, *275*(5306), 1604–1610.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Tesar, B., & Smolensky, P. (1998). Learning in optimality-theoretic grammars. *Lingua*, *106*, 161–196.