



Brief article

Decisions, decisions: infant language learning when multiple generalizations are possible

LouAnn Gerken

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

Received 4 January 2005; accepted 9 March 2005

Abstract

Two experiments presented infants with artificial language input in which at least two generalizations were logically possible. The results demonstrate that infants made one of the two generalizations tested, the one that was most statistically consistent with the particular subset of the data they received. The experiments shed light on how learners might go about solving the induction problem for human language.

© 2005 Elsevier B.V. All rights reserved.

An ever-growing collection of published studies demonstrates that, when infants are familiarized with a language-like system in the laboratory, they are able to extract patterns from that system (Chambers, Onishi, & Fisher, 2003; Gómez, 2002; Gómez & Gerken, 1999; Maye, Werker, & Gerken, 2002; Saffran & Thiessen, 2003; Saffran, Aslin, & Newport, 1996) and generalize beyond the specific stimuli encountered during familiarization to new stimuli embodying the same pattern (Gerken, 2004; Gerken, Wilson, & Lewis, in press; Gómez & Gerken, 1999; Gómez & LaKusta, 2004; Marcus, Vijayan, Rao, & Vishton, 1999; Maye & Weiss, 2003). A handful of these studies has also begun to explore what generalizations infants make when presented with different samples generated by the same formal system. For example, Gómez (2002) familiarized 18-month-olds with an artificial grammar of the form AXB and CXD, in which there is a dependency between the A and B elements and between the C and D elements. She found that it was only when the middle element was selected from a large pool (24) that infants could detect the relation between the first and third elements in the grammar. In a similar vein,

E-mail address: gerken@u.arizona.edu.

Gerken, Wilson and Lewis (2005) demonstrated that 17-month-olds can discern a Russian noun gender paradigm, but only if a subset of the nouns in the paradigm have both an derivational and inflectional marking for gender.

These studies demonstrate that when researchers create a set of stimuli with a particular formal structure, infants can discern that structure if they are given sufficient evidence supporting it. Such studies arguably take us a long way in understanding what sorts of information human language learners can compute, given that the information is available in the input. However, they do not directly address one puzzle that has plagued cognitive scientists for decades: Any set of experiences potentially supports an infinite number of possible generalizations (e.g., Goodman, 1954). Consider the classic example of Chomsky (1980), who notes that a child hearing the subset of her input consisting of the statement-question pair *The man is tall, Is the man tall?* is potentially in trouble. She has no basis on which to determine whether the first instance of *is* is moved to the front of the sentence to make a question, or whether the more abstract unit *main verb* is moved. If a child made the former generalization, she would believe that *Is the man who tall is sad?* is a possible English sentence. In contrast, a child who believed the basis of generalization to concern the main verb would reject that sentence in favor of the grammatical *Is the man who is tall sad?* This example reflects the essence of the induction problem.

I will not discuss here the additional ‘poverty of the stimulus’ claim made by Chomsky (1980) and others, in which input that would ultimately allow the learner to decide which hypothesis is correct is not available (see Pullum & Scholz, 2002, and associated replies). Rather, I will focus on the induction problem - the situation in which a subset of input clearly has at least two formal descriptions. What does an infant learner exposed to such input do? There are at least three possibilities: One is that the infant discerns both patterns embodied in the input and can generalize based on either one. A second possibility is that being faced with evidence of two possible generalizations prevents the learner from generalizing at all. Finally, and perhaps most interestingly, the infant might show evidence of having discerned different formal descriptions for different subsets of the input, depending on which description better accounts for that particular input.

I will return in the discussion to what might constitute a ‘better account’ of the input. However, for the time being, consider the stimuli employed by (Marcus et al., 1999), shown in Table 1. In that study, infants were briefly exposed to strings exhibiting an AAB or ABA pattern and subsequently discriminated new AAB vs. ABA strings. Note that while the AAB grammar illustrates an AAB pattern, it also supports more stimulus-bound generalizations; for example, all strings end in the syllables *di, je, li* or *we*. The fact that

Table 1

AAB familiarization stimuli used by Marcus et al. (1999). The first column and diagonal were used as familiarization stimuli in the current experiments

B	di	je	li	we
A				
le	leledi	leleje	leleli	lelewe
wi	wiwidi	wiwije	wiwili	wiwiwe
ji	jijidi	jijije	jijili	jijiwe
de	dededi	dedeje	dedeli	dedewe

infants were able to generalize to entirely new test strings suggests that they made the more abstract, AAB or ABA, generalization.

Now consider two different subsets of four stimuli in [Table 1](#), those on the diagonal and those in the first column. The stimuli on the diagonal exhibit the same property as the entire set of 16. They exhibit an AAB pattern, and a more stimulus-bound generalization would entail memorizing all four stimuli. Only noting the AAB pattern would allow learners to generalize to an entirely new set of syllables. In contrast, the stimuli in the first column exhibit an AAB pattern, but they also end in the syllable *di*. It is this subset of the stimuli that is of most interest, because both descriptions (AAB and ends in *di*) would allow learners to generalize beyond the particular syllables to which they had been exposed to a new set. The question addressed in the two experiment presented here is, what information do infants discern in the stimuli in the first column of [Table 1](#)? In particular, do they make only the AAB generalization, only the ends in *di* generalization, both, or neither?

1. Experiment 1

Experiment 1 asked whether infants demonstrate an abstract AAB (or ABA) generalization to the four stimuli in the first column of [Table 1](#) as well as to the stimuli on the diagonal.

1.1. Methods

1.1.1. Materials

Stimuli for both experiments were generated with the speech function of a Power Macintosh, system 8.6, using the Victoria voice at the default rate. One sec. pauses were inserted between the 3-syllable words, using speech analysis software. Infants in the column condition of Exp. 1 were familiarized with a 2 min., randomly ordered sample of the 3-syllable nonsense items *leledi*, *wiwidi*, *jijidi*, and *dededi* (AAB group) or *ledile*, *widiwi*, *jidiji*, and *dedide* (ABA group).¹ Infants in the diagonal condition were familiarized with *leledi*, *wiwije*, *jijjili*, and *dedewe* (AAB condition) or *ledile*, *wijewi*, *jiliji*, and *dewede* (ABA condition). During test, all infants heard *kokoba* and *popoga* in random order on AAB trials and *bakoba* and *gapoga* on ABA test trials.²

1.1.2. Participants

Thirty-two 9-month-olds contributed data for Exp. 1, 16 in the column condition (8 AAB and 8 ABA, mean age 9 mos. 3 days) and 16 in the diagonal condition

¹ Pilot testing on 7-month-olds of the AAB vs. ABB generalization reported by Marcus et al. using the full set of 16 stimuli failed to replicate the original result, but revealed a marginal overall preference for ABB strings. This replication failure prompted the use of the AAB vs. ABA generalization used in the experiments reported here. Additional pilot testing on 7- and 9-month-old revealed that the latter group yielded more robust data.

² Based on a suggestion by Eimas (1999), the ABA test trials were generated to be maximally similar to the AAB trials, differing only in the first syllable.

Table 2

Mean listening times and standard errors on trials with patterns consistent and inconsistent with familiarization stimuli in experiments 1 and 2, and significance tests comparing listening times

	Exp. 1	Exp. 1	Exp. 2
	Diagonal condition	Column condition	
Consistent	13.51 (1.82)	10.74 (1.56)	9.33 (1.18)
Inconsistent	10.14 (1.39)	10.18 (1.60)	6.25 (0.64)
Difference in sec. and significance	3.37 ($p < 0.05$)	0.56 (n.s.)	3.08 ($p < 0.05$)

(8 AAB and 8 ABA, mean age of 9 mos., 9 days). All infants were from English-speaking homes with no history of hearing or speech/language disorder. An additional 17 infants were tested but failed to listen for at least 2 sec. to all four test trials.

1.1.3. Procedure

In both experiments, infants participated in a familiarization phase followed by test trials. Each infant sat on a caregiver's lap in a sound-proof booth. The caregiver listened to masking music over headphones to avoid influencing the infant. An experimenter viewed the session on a video monitor outside the booth and controlled stimulus presentation via a Power Macintosh computer. In front of the infant was an amber light, and to each side were red lights under speakers. Once the infant appeared calm and attentive, the experimenter began the familiarization phase, during which the relevant word list was played from both speakers, and the light under one speaker flashed. During the 2 min. familiarization, infants heard 4 3-syllable strings in random order, with 1 sec. pauses between. After familiarization, the infant participated in 4 test trials (2 AAB and 2 ABA). Pilot testing revealed that no significant discrimination emerged if a greater number of trials was used. A test trial began when the infant oriented to the flashing center light. One of the side lights would then begin to flash, and when the infant turned toward the flashing light, an auditory word list would be played from the corresponding speaker. The trial lasted until the infant looked away from the light for 2 sec. In keeping with standard procedures using the Headturn Preference Procedure, looking times shorter than 2 sec. were excluded from the analyses (Kemler Nelson, Jusczyk, Mandel, Myers, Turk and Gerken, 1995).

2. Results and discussion

Overall, infants listened longer to test stimuli that were consistent with their familiarization condition (Table 2).³ An analysis of variance, with familiarization condition (column vs. diagonal) as the between subjects variable and test condition

³ The direction of preference is the opposite of that found by by Marcus et al. (1999). The reason for this difference may be the number of test trials (4 in the current experiment vs. 12 in Marcus et al.), the smaller number of familiarization items (4 vs. 16), or the fact that the AAB and ABA test trials differed only by the first syllable.

(consistent vs. inconsistent with training) as the within subjects variable revealed a main effect of consistency ($F(1, 30) = 5.20, p = .03$) and a marginal interaction between condition and consistency ($F(1, 30) = 2.65, p = 0.11$). Planned t-tests were done to examine the performance of the column vs. diagonal groups separately. Infants in the column group showed no evidence of discriminating the test items ($t(15) = 0.57, p > 0.50$, 2-tailed), while infants in the diagonal group showed significant discrimination ($t(15) = 2.39, p = 0.03$, 2-tailed).

Infants in the diagonal condition were familiarized with a subset of the stimuli in which the only common feature was an abstract AAB or ABA pattern. Like the infants studied by Marcus et al. (1999), infants in this condition were able to generalize to new test stimuli, suggesting that they had made the intended generalization, having only been exposed to four stimulus types. Infants in the column condition, who were exposed to a different subset of the same larger data set, failed to make the generalization. This pattern of results is consistent with two possible interpretations: Infants exposed to input consistent with two different formal systems make no generalization at all. Or, infants generalize based on the formal description that is more likely to have generated the input.⁴ The latter alternative suggests that infants in the column condition of Exp. 1 noted that the middle or last syllable was *di* in ABA and AAB stimuli, respectively. Exp. 2 pursued this alternative.

3. Experiment 2

Exp. 2 was identical to the column condition of Exp. 1 in the familiarization phase. However, in the test phase, the syllable *di* replaced the B element in both AAB and ABA test strings. If infants in Exp. 1 made a generalization based on the location of *di*, they should discriminate the new test strings.

3.1. Methods

3.1.1. Materials

As noted above, the familiarization stimuli were identical to those employed in the column condition of Exp. 1. The test stimuli were generated in the same manner as those in Exp. 1, but were the strings *kokodi* and *popodi* on AAB test trials and *podipo* and *kodiko* on ABA trials.

3.1.2. Participants

Sixteen infants contributed data in the Exp. 2 (8 AAB and 8 ABA, mean age 9 mos. 6 days). All infants were from English-speaking homes with no history of hearing or speech/language disorder. An additional 6 infants were test but failed to listen for at least 2 sec. to all four test trials.

⁴ A third interpretation is that infants in the column condition learned better, however, the data from Exp. 2 argues against this possibility.

3.1.3. Procedure

The procedure was identical to that used in Exp. 1.

4. Results and discussion

As in Exp. 1, infants showed a preference for stimuli consistent with their familiarization condition (Table 2). A t-test on mean listening times for test trials consistent vs. inconsistent with familiarization was significant ($t(15) = 2.26$, $p = .04$, 2 tailed). These data, coupled with infants' failure to discriminate under the same familiarization conditions in Exp. 1, suggest that infants in the column condition made only the generalization involving the position of the syllable *di*.

4.1. General discussion

Although the two experiments reported here employed artificial grammars that are quite different from those found in natural language, and can be learned by non-humans (Cumming & Berryman, 1961; Giurfa et al., 2001; Hauser, Weiss, & Marcus, 2002), they capture a critical aspect of the induction problem in language: Learners were faced with a choice between at least two types of generalization. A question raised by the experiments is what caused infants in the column condition to generalize based on the location of *di* rather than making the more abstract generalization?

One possibility is that the data are consistent with the Subset Principle (Manzini & Wexler, 1987), in which learners select among possible parameter values based on which value generates the smallest language compatible with the input data. Note that, if we interpret 'language' to mean 'set of sentences,' a learner would need to generate all of the sentences for each parameter value and determine which value generated fewer sentences (but see Wexler, 1993). Wexler and Manzini reduce the computational task for the learner by placing relevant parameters in a markedness hierarchy, in which the learner begins with the least marked value, which generates the smallest language.

Another possibility is consistent with Bayesian approaches to generalization (e.g., Tenenbaum & Griffiths, 2001), in which learners compare the subset of the input they have received to the range of input generated by different formal descriptions. For example, an infant might tacitly compute that it is extremely unlikely, given an AAB grammar, the only input ends in *di*. Depending on its implementation, this approach might also be computationally challenging. However, it has the advantage of applying to a more general (e.g., non-parameterized) learning problems, and it allows increasing confidence in hypothesis selection with increasing input set size. Importantly, this solution to the induction problem entails learners choosing among formal descriptions that they have already generated from the data using general purpose mechanisms (Saffran, Reeck, Niebuhr, & Wilson, 2005) or that are part of their innate endowment for language (e.g. Valian, 1990). For example, Saffran et al. (2005) demonstrated that the structure of the input determines the primitives (in this case absolute vs. relative pitch) over which generalizations are made. This type of research, in which learners 'choose' among

different generalizations allowed by input data, may ultimately allow us to distinguish between theories of language development.

Acknowledgements

This research was supported by NIH grant #R01 HD042170. Thanks to Frances Balcomb, Andrew Barss, Rebecca Gómez, Jill Lany, Patrick Neher and Mary Peterson for comments on earlier drafts.

References

- Chambers, K. E., Onishi, K. H., & Fisher, C. L. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, *87*, B69–B77.
- Chomsky, N. (1980). The linguistic approach. In M. Piattelli-Palmarini (Ed.), *Language and Learning* (pp. 109–116). Cambridge, MA: Harvard University Press, 109–116.
- Cumming, W. W., & Beryman, R. (1961). Some data on matching behavior in the pigeon. *J. Exp. Anal. Behav.*, *4*, 281–284.
- Gerken, L. A. (2004). Nine-month-olds extract structural principles required for natural language. *Cognition*, *93*, B89–B96.
- Gerken, L. A., Wilson, R., & Lewis, W. (2005). 17-month-olds can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*, 249–268.
- Giurfa, M., Zhang, S. W., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature*, *410*, 930–933.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436.
- Gómez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135.
- 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.
- Goodman, N. (1954). *Fact, fiction, and forecast*. University of London: Athlone Press.
- Hauser, M. D., Weiss, D., & Marcus, G. F. (2002). Rule learning by cotton-top tamarins. *Cognition*, *86*(1), B15–B22.
- Kemler Nelson, D., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A. E., & Gerken, L. A. (1995). The headturn preference procedure for testing auditory perception. *Infant Behavior and Development*, *18*, 111–116.
- Manzini, R., & Wexler, K. (1987). Parameters, binding theory and learnability. *Linguistic Inquiry*, *18*, 413–444.
- Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80.
- Maye, J., & Weiss, D. J. (2003). Statistical cues facilitate infants' discrimination of difficult phonetic contrasts. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th annual Boston University conference on language development* (pp. 508–518). Somerville, MA: Cascadilla Press, 508–518.
- Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101–B111.
- Pullum, G. K., & Scholz, B. C. (2002). Empirical assessment of stimulus poverty arguments. *Linguistic Review*, *19*(1), 9–50.
- Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, *39*, 484–494.
- Saffran, J. R., Aslin, R. N., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Saffran, J. R., Reeck, K., Niebuhr, A., & Wilson, D. (2005). Changing the tune: the structure of the input affects infants' use of absolute and relative pitch. *Developmental Science*, *8*, 1–7.

- Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24, 629–640.
- Valian, V. V. (1990). Null subjects: A problem for parameter-setting models of language acquisition. *Cognition*, 35, 105–122.
- Wexler, K. (1993). The subset principle is an intensional principle. In E. Reuland, & W. Abraham (Eds.), *Knowledge and Language: Orwell's Problem and Plato's Problem*. Dordrecht: Kluwer.