LANGUAGE LEARNING AND DEVELOPMENT, *4*(3), 228–248, 2008

ISSN: 1547-5441 print / 1547-3341 online
DOI: 10.1080/15475440802143117

Psychology Press
Taylor & Francis Group

Three Exemplars Allow at Least Some Linguistic Generalizations: Implications for Generalization Mechanisms and Constraints

LouAnn Gerken and Alex Bollt
University of Arizona

The mechanism that allows learners to generalize over linguistic input and its relation to constraints on possible generalizations was explored in three experiments. Infants were familiarized briefly with words exhibiting stress patterns generated by a set of ordered principles, and then tested on new words that were either consistent or inconsistent with those principles. In Experiment 1, 9-month-olds generalized from three different heavy syllables (those ending in a consonant) heard during familiarization, to a new heavy syllable at test, but did not generalize from multiple tokens of one heavy syllable. Experiment 2 demonstrated that, although 9-month-olds were able to use three syllables to generalize the linguistically natural principle assigning stress to heavy syllables, they failed to use them to generalize a principle that does not occur in natural languages, suggesting that the mechanism requiring three input examples does not apply to all input equally. Experiment 3 demonstrated that although 9-month-olds failed to generalize the unnatural principle, 7.5-month-olds succeeded, suggesting that constraints on generalization can be acquired over development.

INTRODUCTION

Human language is a marvel of creativity; from a finite set of words we can produce and comprehend an infinite number of utterances. Human language learners are no less marvelous; from a finite number of input utterances they are able to

This research was supported by NIH HD042170 to LouAnn Gerken. The authors warmly thank Rebecca Gómez, Tom Shultz, Andy Wedel, Josh Tenenbaum, and two anonymous reviewers for helpful comments and suggestions on earlier drafts.

Correspondence should be addressed to LouAnn Gerken, Cognitive Science, Communications 302, University of Arizona, Tucson, AZ 85721. E-mail: gerken@email.arizona.edu

produce novel utterances of their own. The fact that language learners are able to produce utterances they have never heard before suggests that they do more than simply store and reproduce their input. Rather, learners appear to have converged, at least to some degree of approximation, on the abstract linguistic system that generated their input. Determining how learners achieve this feat requires us to address at least three related questions.

First, what is the mechanism that infants use to generalize from the input that they have encountered to new instances? Of course, establishing a possible mechanism requires data from many converging sorts of investigations. However, we can narrow the scope of the question by asking how much evidence learners need in order to make a generalization (e.g., Xu & Tenenbaum, 2007b). All other things being equal, learners who generalize based on a single piece of data run the risk of wildly overgeneralizing or generalizing on the wrong basis altogether. But learners who require thousands of consistent exemplars for each generalization may be making poor use of their time, especially if they are using unconstrained induction, which cannot guarantee convergence on the correct generalization in the absence of negative evidence (e.g., Gold, 1967).

The “all other things equal” phrase in the preceding paragraph brings us to the second question: can the learning mechanism, whatever it might be, be applied to all data equally, or are learners constrained to consider some generalizations more likely than others? Clearly an individual infant cannot consider an infinite number of possible bases of generalization. Nevertheless, we must ask whether human infants as a group tend to make certain generalizations more than others. If we are using the amount of input required for generalization to provide a clue about the learning mechanism, we might ask whether the same amount of input is sufficient for making any equally complex generalization.

If the answer to the second question is *no*, and learners can generalize only from certain types of input, we must ask a third question: is the set of possible generalizations innate (e.g., Chomsky, 1981; Gibson & Wexler, 1994; Roeper & Williams, 1987), or might certain generalizations be acquired as learners note statistical patterns in their environment (e.g., Chambers, Onishi, & Fisher, 2003; Maye, Werker, & Gerken, 2002; Saffran, Aslin, & Newport, 1996)? One way to address the third question is to ask whether younger learners are *less* constrained in their generalizations than older learners.

These three questions—how much input is required by the linguistic generalization mechanism, whether the same amount of input is required for any generalization, and whether the focus of generalization narrows over development—can each be linked to a small handful of quite different developmental studies. These will be reviewed, briefly, below. The three experiments reported here are consistent with existing findings. The intended contribution of these experiments is to address the three questions about infant linguistic generalization that we have outlined using uniform stimuli and methods, and focused on a domain that is central to linguistic theorizing, thereby providing a more complete sketch than has previously been offered of infants converging on an appropriate linguistic generalization.

A review of the literature on how much input is required for infants and children to make generalizations from the input they hear shows remarkable convergence. Gerken (2006) demonstrated that 9-month-olds familiarized with just four different AAB or ABA patterns of syllables (*leledi, wiwije, jijili, dedewe* or *ledile, wijewi, jiliwi, dewede*) were able at test to discriminate AAB vs. ABA patterns instantiated in new syllables (*popoga, kokoba* vs. *pogapo, kobako*). This study indicates that infants can infer a pattern and generalize it from four different examples or *types*. (The word "types" in this work is used to refer to different examples yielded by a particular category or abstract structure.) Research on younger infants' visual generalization suggests that generalization can occur when only three input types are presented. Needham and colleagues exposed 4-month-olds to between one and three different types of a visual category and then tested them on new items that were either consistent or inconsistent with the category (Needham, Dueker, & Lockhead, 2005). They found that infants did *not* generalize based on one or two types generated by the category, but they did generalize from three types. Quinn and Bhatt (2005) confirmed the lack of generalization with two types of visual input stimuli and significant generalization with three types.

Moving away from infant pattern learning to child word learning, Xu and Tenenbaum (2007b) have shown that when 3- to 4-year-olds hear a label applied to a single dog (e.g., a dalmation), they extend that label both to other dalmations and to other dogs. When the label is applied to three similar dalmations, children extend it only to other dalmations, but when it is applied to three different dogs (different types), they treat the label as applying to the more general category dogs and not only dalmations. The authors explain this finding within a Bayesian hypothesis selection framework, arguing that a single labeled dalmation is equally consistent with a number of hypotheses, including that the label refers to just dalmations or to dogs. In contrast, hearing the same label applied to three different dalmations and to nothing else increases the probability that the appropriate extension is just dalmations (see also Gerken, 2006). Finally, hearing the label applied to three different types of the more general category of dogs provides positive evidence that the label does not apply solely to dalmations. In keeping with our first question, Xu and Tenenbaum (2007b) compare the Bayesian model that predicted their findings with one associative (Hebbian) learning model and find that the latter does not converge on the subordinate (dalmation) category with three input examples (also see Xu & Tenenbaum, 2007a). They also compare the Bayesian model with an instantiation of the subset principle (e.g., Berwick, 1986), in which the most narrow hypothesis is always preferred. The subset principle also failed to account for the data, but its error was different from the error in the associative model; it selected the subordinate (narrower) category regardless of whether one or three examples were presented.

Summarizing the data on the amount of input required for generalization, the existing studies suggest that one input type from a category is not sufficient to lead learners to appropriate generalization. Two studies examined the effectiveness of two input types and found that number to be insufficient for generalization of visual stimuli. Therefore, a tentative conclusion from this research is that at least three input types, but not any/many more are required to account for the data on generalizations made by infants and young children. This hypothesis will be explored in Experiment 1 specifically with reference to phonological structure, a domain in which the amount of input required for generalization has not been explored. Moreover, although we will not draw firm conclusions here about specific generalization mechanisms, the model comparisons reported by Xu and Tenenbaum (2007b) support our suggestion that determining the amount of input required for generalization has the potential to inform us about possible mechanisms.

We next turn to the literature on constraints on generalization and where they might come from. Few studies have been conducted to explore this issue with infants, probably because the studies would entail demonstrating that infants are *unable* to make a particular generalization, and null results are difficult to interpret in infant studies. Nevertheless, several existing studies are instructive. Gómez (2002) found that infants were unable to learn an aXb cXd pattern, in which there were dependencies between the first and third string elements when the number of X elements was small, but not when the number was large. She interprets this finding to mean that infants are biased against generalizations that involve nonadjacent dependencies, but are able to overcome this bias when the evidence is sufficiently strong (also see Newport & Aslin, 2004). This constraint on generalization over nonadjacent elements appears to apply to adults, children, and infants (Gómez, 2002; Newport & Aslin, 2004) and may, therefore, be one that is innate in humans and other animals. Another examination of constraints on generalization was carried out by Marcus and colleagues (Marcus, Fernandes, & Johnson, 2007) who demonstrated that 7.5-month-old infants were able to learn AAB or ABB patterns instantiated in syllables, but not in tones. They suggest that these results demonstrate a possibly innate bias in favor of extracting relations among elements in a string, but only in the domain of language. However, preliminary results from Dawson and Gerken (2006) indicate that 4-month-olds are able to generalize over AAB vs. ABA musical strings, while 7.5-month-olds fail to generalize over the same strings—the latter finding replicating Marcus et al. (2007). Recent research also demonstrates that rats are able to learn similar patterns with both syllables and tones (Murphy, Mondragon, & Murphy, 2008). The positive results from younger infants and rats, coupled with the null results from older infants, suggest that at least some constraints on generalizations are acquired as learners determine the relevant properties of a particular domain. Relations among specific notes account for very little of the variance in music (Saffran & Griepentrog, 2001), but relations among specific words are highly relevant to language. The notion that increased experience with music causes infants to demonstrate less sensitivity to some patterns instantiated in musical stimuli is at least superficially similar to the way in which infants show reduced sensitivity to nonnative linguistic contrasts based on statistical properties in the input (Maye et al., 2002; Werker & Tees, 1984).

Taken together, the data on constraints on generalization and their potential source(s) support two tentative conclusions. First, as we saw in the work of Marcus et al. (2007), the same learning mechanism (in that case, one that looks for relations among identical elements) might apply differentially to different input. We will investigate this notion in Experiment 2. Second, while some constraints on generalization, such as the span over which learners look for dependencies, may be inherent to human learners, others might arise as learners become familiar with the statistical distribution of their input. We will investigate this notion in Experiment 3.

The three experiments presented here explore the number of different input types needed for generalization, possible constraints on generalizations, and whether constraints change over development. The domain of linguistic generalization under consideration for all three explorations is metrical phonology. Metrical phonology concerns structural principles for assigning stress to syllables in multisyllabic words, and it is central to a discussion of linguistic generalization (Demuth, 1996; Dresher, 1999; Dresher & Kaye, 1990; Gupta & Touretzky, 1994; Hayes, 1994; Prince & Smolensky, 1997; Tesar & Smolensky, 1998). Recent research from our laboratory provided evidence that 9-month-olds are able to rapidly extract principles of metrical phonology in a brief laboratory exposure (Gerken, 2004). Infants in two experiments were exposed to 3- and 5-syllable words from two artificial language systems (L1 and L2), which were originally developed for a study with adults (Guest, Dell, & Cole, 2000, see Table 1), and which employed stress principles akin to those proposed for actual human languages.

Because modified versions of the same two languages will be used in all of the experiments reported here, let us briefly describe them. Table 1 shows that there are five types of familiarization words (familiarization word types 1–5) and two possible types of test words (Stress Pattern test words; Abstraction test words). In the actual materials there were seven different versions of each of the familiarization and test words, with the versions created by replacing each of the seven solfège syllables (*do, re, mi,* etc.) with the next syllable in the solfège sequence (e.g., *do* in the words shown in Table 1 was replaced by *re*, *re* by *mi*, etc.). The syllable *TON* in Table 1 always stayed in the same location for each of the familiarization and test word types. The familiarization and test words for each

TABLE 1
Sample Familiarization and Test Words Used in L1 and L2[a]

Word Type	*L1 sample words*	*L2 sample words*
Familiarization word type 1	TON ton do RE mi	do RE mi ton TON
Familiarization word type 2	TON do re	do re TON
Familiarization word type 3	DO re TON	TON do RE
Familiarization word type 4	DO re TON mi fa	do re TON mi FA
Familiarization word type 5	DO re mi FA so	do RE mi fa SO
Stress Pattern test word[b]	MI re TON	TON mi RE
Abstraction test word	do TON re MI fa	do RE mi TON fa

[a]Syllables in upper case are stressed, and those in lower case are unstressed.
[b]The stress pattern test items have the same stress patterns as familiarization word type 3 with the solfège syllables in different orders.

language were created from the application of four principles. The principles were ranked with respect to each other, so that when two principles could apply to a word, only the most highly ranked applied (Prince & Smolensky, 1997). The principles employed in L1, in order from high to low ranking, were: (A) two stressed syllables cannot occur in sequence, (B) heavy syllables (in the current context, those ending in a consonant) should be stressed, (C) syllables should be stressed if they are second to last, and (D) alternating syllables should be stressed, starting from the left. L2 employed principles A and B, but replaced principles C and D with: (C') syllables should be stressed if they are second, and (D') alternating syllables should be stressed, starting from the right.

To understand how the principles apply to yield the familiarization and test words in Table 1, first consider the L1 familiarization word *TON ton do RE mi*. Both principles B and D dictate that the first (heavy and leftmost) syllable should be stressed, and no other principle contradicts. Principle B dictates that the second (heavy) syllable should be stressed, but it is outranked by principle A, which disallows two stressed syllables in sequence. Principle D dictates that alternating syllables, and therefore the third syllable, should be stressed, but it is outranked by principle C, which dictates that the second to last (fourth) syllable should be stressed, and principle A prevents both the third and fourth from being stressed. Now consider the Abstraction test word *do TON re MI fa* in L1. The pattern of stressed and unstressed syllables found in this word never occurs in the familiarization stimuli; therefore these Abstraction test words are a true test of generalization beyond the input. In these words, Principle D dictates that the first syllable should be stressed, but the more highly ranked principle B says that the second (heavy) syllable should be stressed, and principle A prevents both first and second syllables from being stressed. Principles C and D both dictate that the fourth syllable be stressed, with no more highly ranked principle contradicting.

Infants in the experiments reported in Gerken (2004) showed significantly different listening times in the Headturn Preference Procedure (Kemler Nelson et al., 1995) for test items consistent vs. inconsistent with their familiarization language. Important for the present study is that the clearest indication of learning was exhibited for the Abstraction test words, where stress patterns were different from those encountered during familiarization. This finding makes clear that infants were able to go beyond learning stress patterns (Jusczyk, Cutler, & Redanz, 1993) to the principles that are hypothesized to underlie those patterns. However, the degree to which infants extracted the intended principles and rankings was left partially unexplored by the earlier study. That is because one of the key principles employed was that stress should be placed on heavy syllables (in the current context, those ending in a consonant); however, only a single example of a heavy syllable (a single input type) was presented across the set of words used in familiarization and test (*TON*). Therefore, we cannot determine from the previous results whether infants extracted the intended principle "stress heavy syllables," or a more stimulus-bound principle (e.g., "stress TON," Gerken, 2004, 2006). Experiment 1 avoids this problem by testing infants on new words that obey the stress principles, but contain a heavy syllable never heard during familiarization.

EXPERIMENT 1

Experiment 1 tested the prediction, based on the literature reviewed above, that infants would fail to generalize from multiple tokens of one heavy syllable to another, but succeed in generalizing from three heavy syllables to a fourth. The logic of that prediction is as follows: the paradigm used by Gerken (2004; also Guest et al., 2000) familiarized infants with five word types to provide evidence for a set of stress principles and rankings that learners can then apply to test items exhibiting new stress patterns (the Abstraction test items in Table 1). The data from Gerken (2004) suggest that infants are able to generalize over the different word types. All three of the current experiments employ the same five-word-type design; thus, infants in Experiment 1 who are familiarized with words containing *one* heavy syllable (*BOM)* and tested on words containing a second heavy syllable (*TON*) are familiarized with the same heavy syllable in multiple word contexts.[1] We reasoned that this situation should prompt learners to conclude that there is only one such syllable (*BOM*) and therefore not to generalize to a new heavy syllable at test (see Gerken, 2006). In contrast, being familiarized with

[1]Presenting a single heavy syllable in multiple word contexts is similar to presenting a single label to several dalmations, as was done by Xu and Tenenbaum (2007b).

three different heavy syllables occurring in the same word contexts should allow learners to abstract beyond the specific three syllables. This abstraction should allow them to generalize to a fourth heavy syllable at test.

Methods

Materials

Materials for the 1 Heavy Syllable familiarization condition were seven tokens of each of the five familiarization word types shown in Table 1 (L1 and L2). The only heavy syllable used during familiarization was *BOM*. That is, all instances of *TON* in the familiarization stimuli from Gerken (2004) were replaced with *BOM*. Materials for the 3 Heavy Syllable familiarization condition were like those for the 1 Heavy Syllable condition, except that two additional heavy syllables were added to the familiarization stimuli, making the set of heavy syllables used in familiarization *BOM*, *KEER*, and *SHUL*. These syllables were rotated through the seven familiarization word versions of each word type, with five versions of each word type appearing with two of the heavy syllables. For example, the first familiarization word type for L1 in Table 1 had the following familiarization words in the 3 Heavy Syllable familiarization condition: *do re BOM, do re KEER, re mi KEER, re mi SHUL, mi fa SHUL, mi fa BOM, fa so BOM, fa so KEER, so la KEER, so la SHUL, la ti SHUL, ti do BOM*.

The 35 words in each language were recorded by the first author 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 msec. pauses between.[2]

Test item lists for all four conditions (1 Heavy Syllable L1, 1 Heavy Syllable L2, 3 Heavy Syllables L1, 3 Heavy Syllables L2) comprised seven tokens of each of the four test word types shown in Table 1 (L1 Stress Pattern, L1 Abstraction, L2 Stress Pattern, L2 Abstraction). Stress Pattern test items shared a stress pattern and absolute syllable location with familiarization items. For example, there are three-syllable words in the L1 familiarization set that are stressed on first and third syllables, and in which the first syllable is one of the seven solfège syllables (e.g., *DO re TON*). Stress Pattern test items for L1 shared this property, although

[2]To ensure that the words containing heavy syllables were roughly equivalent in the two familiarization conditions, acoustic duration measurements were made of all familiarization words containing heavy syllables. The mean duration in msec. (and SE) of words containing heavy syllables in the 1 Heavy Syllable condition was 924 (24) and 911 (24) in L1 and L2, respectively. The comparable values in the 3 Heavy Syllables condition were 930 (32) and 936 (31) in L1 and L2, respectively. Based on the duration measure, the words in the two conditions appear to be acoustically very similar, with a mean duration difference of only 16 msec., less than one standard error.

the specific test words were never heard during familiarization (e.g., *MI re TON*). Abstraction test items had entirely different stress patterns from familiarization items. For example, L1 had no 5-syllable familiarization words with stress on the second and fourth syllables or with a heavy second syllable. However, such words were generated by the stress principles and rankings attested by the familiarization stimuli and therefore were used as the Abstraction test items for L1. Parallel Stress Pattern and Abstraction test items were created for L2 (see Table 1). Each test item list was approximately 20 sec. in duration and was presented twice, once in each half of the experiment for a total of 16 test trials. Note that our main interest in Experiment 1 was in the Abstraction test items, because these required infants to generalize beyond the particular stress patterns encountered during familiarization. We included the Stress Pattern test items only so that the test items in Experiment 1 would be identical to those in the previous study by Gerken (2004).

Participants

Participants were 36 infants (mean age 9 mos. 2 days; range 8 mos. 15 days to 9 mos. 15 days) from English-speaking homes with no history of hearing or speech/language disorder. Eighteen infants were in the 1 Heavy Syllable familiarization condition and the other 18 in the 3 Heavy Syllables familiarization condition. Half of the infants in each familiarization condition were familiarized with the L1 version and half with the L2 version of the language. An additional 11 infants were tested but failed to provide useable data for at least 12 test trials.

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. The caregiver listened to masking music over headphones and was instructed to look forward so as not to influence the infant's behavior. During familiarization, the entire word list (L1 or L2) was played from both speakers, and the light under one speaker flashed. An experimenter outside the test booth, who could not hear the stimuli, watched each infant's looking behavior over a closed circuit television and recorded direction of looking (center, left, right, elsewhere) on a button box connected with the computer running the testing software. After familiarization, the infant participated in 16 test trials, which were organized into two blocks of eight trials (2 different L1 pattern, 2 different L1 abstraction, 2 different L2 pattern, 2 different L2 abstraction), with the order randomized within each trial. 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 two seconds. Following the conventions in our laboratory and in the field, looking times shorter than two seconds were excluded from the analyses.

Results and Discussion

Listening times for each infant were averaged over the four tokens of each of the four word types (L1 Abstraction, L2 Abstraction, L1 Stress Pattern, L2 Stress Pattern). Following other studies using the two-grammar design, the L1 and L2 data labels were changed to “consistent” and “inconsistent” based on the infant’s familiarization language. For example, L1 Abstraction test items were relabeled “consistent abstraction” for infants familiarized with L1 and “inconsistent abstraction” for infants familiarized with L2. A t-test was performed to determine if the difference in listening times on consistent vs. inconsistent trials varied as a function of familiarization language (L1 vs. L2; t (34) = 0.08, $p > 0.90$, 2-tailed). Because the two languages yielded similar patterns of listening time differences, subsequent analyses collapse across familiarization language.

Because the primary comparison of interest was whether infants showed differences in learning for the Abstraction test items for 1 vs. 3 heavy syllable conditions, a 2 familiarization condition (1 heavy syllable vs. 3) X 2 consistency (consistent vs. inconsistent with familiarization language) ANOVA was first performed on the average amount of time infants attended to the Abstraction test stimuli of each type (see Fig. 1). Neither the main effect of familiarization

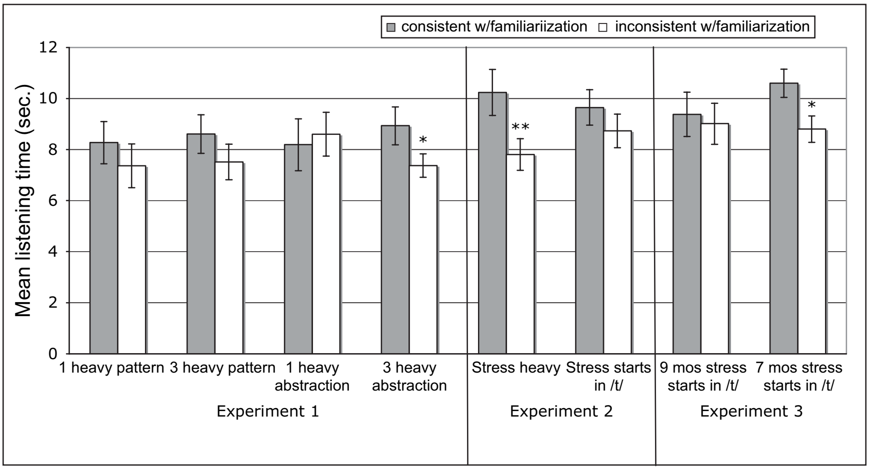

FIGURE 1 Mean listening times with standard error bars of the test conditions in Exps. 1–3. Note that Exps. 2–3 only employed test items comparable to the Abstraction test items of Exp. 1. * = $p < .05$ on a 2-tailed t-test ** = $p < .001$ on a 2-tailed t-test.

condition (F (1, 34) = 0.35, p = 0.56), nor the main effect of consistency (F (1, 34) = 1.15, p = 0.29) was significant. However, as predicted, there was a significant familiarization condition X consistency interaction (F (1, 34) = 4.08, p = 0.05). Follow-up t-tests on the listening time for consistent vs. inconsistent test items revealed that infants listened longer to stimuli consistent with the familiarization language in the 3 heavy syllable condition (t (17) = 2.64, p < 0.02 2-tailed). In contrast, infants in the 1 heavy syllable condition failed to demonstrate a significant effect of familiarization (t (17) = 0.38, p = 0.71). A parallel ANOVA was performed on listening times to the Stress Pattern test items. There was a trend toward a main effect of consistency (F (1, 34) = 3.20, p = 0.08), such that infants listened longer to items with stress patterns consistent with patterns that they had heard during familiarization. Neither the main effect of the familiarization condition nor the interaction approached significance (F's < 1).[3] The relatively weaker effect of the Stress Pattern condition compared with the Abstraction condition is consistent with Gerken (2004). Perhaps because the stress patterns in L1 and L2 Stress Pattern test items were familiar, infants were influenced by familiarity at this level, even though the placement of the heavy syllable differed between L1 and L2 test items.

The predicted interaction between familiarization condition and consistency for the Abstraction test items suggests that generalization over multiple tokens from a single input type is more difficult than generalization over three different input types. This finding is consistent with the existing developmental literature (Gerken, 2006; Needham et al., 2005; Quinn & Bhatt, 2005; Xu & Tenenbaum, 2007b). The data are also consistent with a suggestion by Albright & Hayes (2003) that induction across phonologically diverse input forms yields more general linguistic rules. The data from the current study support the view that infants in the previously published study on stress principle learning (Gerken, 2004) were not generalizing based on the abstract linguistic principle "stress heavy syllables," because infants in that experiment were familiarized with only a single heavy syllable. Rather, they may have generalized based on a more narrow principle of "stress *TON*." A similar finding in a different domain was reported by Gerken (2006), in which infants generalized to new AAB or ABA patterns, only when all of the A and B syllables varied. When there was a single B syllable that

[3]Interestingly, adults exposed to the same familiarization stimuli as infants in the 3 Heavy Syllables condition and then asked at test whether each test word was consistent or inconsistent with what they had heard during familiarization showed a strong effect of learning for the stress pattern test items, but no effect of learning at all for the Abstraction test items (also see Guest et al., 2000). One interpretation of these findings is that, as native English-speakers, adults have realized that syllable position (e.g., first syllable of a word) is the most reliable associate of lexical stress and therefore have come under most circumstances to ignore more subtle correlations (e.g., stress and ending in a consonant).

occurred in four different strings, infants generalized only to new strings containing that particular syllable.

Infants in the 3 Heavy Syllables condition demonstrated a familiarity preference, which stands in contrast to the novelty preference shown by infants in previously published studies of metrical stress learning (Gerken, 2004). One likely reason for the reversal in preference is that generalization in previous studies appeared to rest on relatively stimulus-bound properties of the familiarization—a principle akin to ”stress TON.” In the current study infants had a much more challenging task of generalizing over three different heavy syllables to a fourth one. Consistent with accounts of novelty vs. familiarity preference in the literature, the more challenging generalization may not have been fully mastered at the time of the test, resulting in infants' continued interest in stimuli that were like the familiarization stimuli (Hunter & Ames, 1988).

A question raised by infants' successful generalization in the 3 Heavy Syllable condition of Experiment 1 is whether any generalization of comparable complexity to “stress heavy syllables” can be inferred with just three different input types. In particular, can a generalization that is not typical of human language be learned with similar ease?

EXPERIMENT 2

In Experiment 2 we compared infants' ability to generalize from three heavy syllables to a fourth (a replication of Experiment 1), to a generalization that might not be equally learnable. This research follows the lead of Seidl and Buckley (2005), who in two experiments found that 9-month-olds were equally able to learn a phonetically natural generalization, which occurs in human languages, and an unnatural generalization, which does not occur in human languages. Experiment 2 asked if infants could learn a generalization that is atypical in characterizations of the metrical phonology of human languages.

Human languages appear to assign stress primarily based on the syllable rimes (vowel and/or final consonant), but not on syllable onsets (e.g., Dresher, 1996; Halle & Vergnaud, 1987; Prince, 1990; but see Gordon, 2005). Therefore, we reasoned that learners might be more likely to consider a generalization about syllable stress involving heavy syllables than syllables starting with a particular consonant. So that we could use the same test stimuli employed in Experiment 1 (in which *TON* was the heavy syllable), thereby leading to the greatest possibility for cross-experiment comparison, we generated stimuli in which nonheavy (CV in the current context) syllables beginning with /t/ were stressed. If, by age 9 months, learners are biased to consider linguistic principles that are attested in human languages, they should show evidence of generalizing “stress heavy syllables,” but not “stress syllables starting with /t/.” In contrast, if 9-month-olds' generalization in Experiment 2 reflects entirely unconstrained application of a particular generalization mechanism, they should generalize based on either principle, given that they have three different input types providing evidence for each generalization.

Methods

Materials

Materials for the "stress heavy syllables" familiarization condition were identical to L1 and L2 used in Experiment 2, with two exceptions. First, the solfège syllable /ti/ was replaced by the syllable /si/ in all familiarization and test stimuli. Second, Stress Pattern test items were excluded, because they are not of interest for the question of generalization. This change left a total of eight test trials. Materials for the "stress syllables starting with /t/" familiarization condition replaced the heavy syllables *BOM*, *KEER*, and *SHUL* in the stress heavy syllable familiarization condition with the CV syllables *TO*, *TI*, and *TU*. The test stimuli for the two pairs of languages were identical, each containing the syllable *TON*, which is both heavy (consistent with the stress heavy syllables familiarization condition) and starts in /t/ (consistent with the stress syllables staring with /t/ familiarization condition). Therefore, if infants in the two familiarization conditions respond differently to the identical test items, we can infer that what they learned during familiarization is responsible.

Participants

Participants were 36 infants (mean age 9 mos. 2 days; range 8 mos. 15 days to 9 mos. 15 days) from English-speaking homes with no history of hearing or speech/language disorder. Eighteen infants were in the stress heavy syllables familiarization condition and the other 18 in the stress syllables starting with /t/ familiarization condition. Half of the infants in each familiarization condition were familiarized with the L1 version and half with the L2 version of the language. An additional 8 infants were tested but failed to provide useable data for at least six test trials.

Procedures

Procedures were identical to those used in Experiment 1.

Results and Discussion

As in Experiment 1, a t-test was performed to determine if the difference in listening times on consistent vs. inconsistent trials varied as a function of familiarization language (L1 vs. L2; t (34) = 0.50, $p > 0.60$, 2-tailed). Because the two languages yielded similar patterns of listening time differences, subsequent analyses collapse across familiarization language. A 2 familiarization condition (stress heavy vs. stress starts with /t/) X 2 consistency (consistent vs. inconsistent with familiarization language) ANOVA was performed on the amount of time infants attended to test stimuli of each type (see Figure 1). Familiarization condition was a between-subjects variable. There was a significant main effect of consistency, such that infants listened longer on test trials that were consistent with their familiarization language (F (1, 34) = 7.77, $p < 0.01$). This finding mirrors the familiarity preference observed for the abstraction test items in the 3 heavy syllable condition of Experiment 1. Neither the main effect of familiarization condition nor the interaction was significant (p's > .25). Planned t-tests were performed to determine if infants showed a reliable familiarity preference for both the stress heavy syllables and the stress syllables starting with /t/ principles. The t-test for infants familiarized with the stress-heavy syllables language was highly significant (t (17) = 3.86, $p < 0.002$, 2-tailed). Thus, the central finding of Experiment 1 was replicated—infants are able to generalize from three heavy syllables to a fourth. In contrast, the t-test for infants familiarized with the stress t-onset language did not approach statistical significance (t (17) = 1.04, $p < 0.4$, 2-tailed).

The data from Experiment 2 hint at the possibility that infants were better able to make a generalization that characterizes a number of languages of the world (stress heavy syllables) than one that does not (stress syllables starting with /t/). However, the lack of an interaction between familiarization condition and consistency means that we must seek additional support for the view that the same learning mechanism (one that requires a small number of input examples for generalization) does not apply equally to all input. The small difference in generalization for linguistically typical vs. atypical principles observed in Experiment 2 has two possible sources. One possibility raised by the literature on parameter setting in language acquisition is that learners are born to expect languages in which syllable content either has no effect on stress or that heavy syllables are stressed (Dresher, 1999; Dresher & Kaye, 1990). On this view, encountering words with the same number of syllables, but with different stress patterns, should be enough to support the latter (quantity sensitive) parameter setting. However, since there is no proposed linguistic parameter associating stress with syllable onsets, no evidence allows such a generalization. This view suggests that increasing the number of participants in Experiment 2 should reveal the expected interaction between familiarization condition and consistency.

However, another reason why we might expect differences in generalization of typical vs. atypical linguistic principles concerns the statistics of the input. Our examination of English words from the MRC Psycholinguistic Database

(Coltheart, 1981) revealed that, among words with an estimated age of acquisition of five years or younger, syllables ending in a vowel (nonheavy syllables in the current context) constituted only 8% and 12% of stressed syllables in monosyllabic and disyllabic iambic (unstressed-stressed) words, respectively. However, syllables ending in a vowel constituted over one-third (38%) of the unstressed syllables in disyllabic trochaic words (stressed-unstressed) words. This rough measure suggests a relation between stress and syllable shape in English, such that stressed syllables are more likely to end in a consonant than are unstressed syllables. In contrast, an examination of the same database revealed that syllables beginning with /t/ constituted only 6% of stressed syllables in both monosyllabic and disyllabic trochaic words, and importantly, only 4% of the unstressed syllables in iambic words. That is, unlike syllables ending in a vowel, syllables beginning with /t/ did not pattern differently in stressed and unstressed syllables.[4] We have already noted instances in which infants appear to narrow the generalizations they make based on input statistics (Dawson & Gerken, 2006; Maye et al., 2002, Werker & Tees, 1984). If the statistics of English stressed syllables serve a similar developmental function, we might expect that while 9-month-olds do *not* generalize based on syllable onset, younger infants *would*. Experiment 3 explored that possibility.

Experiment 3 also addressed a difference in the relation between familiarization and test words for the stress syllables starting with /t/ stimuli in Experiment 2. Recall that we employed the same test materials in Experiment 2 as Experiment 1, in order to provide the greatest potential for comparison across experiments. However, this approach meant that the three stressed syllables in the “starting in /t/” familiarization condition were open syllables (CV), while the stressed syllable starting in /t/ at test was closed (CVC, *TON*). In line with the foregoing discussion on hypothesis selection, infants may well have viewed the familiarization stimuli as providing evidence for the narrower generalization that CV syllables starting in /t/ are stressed. This generalization would not apply to the test items, which were all CVC syllables. Therefore, Experiment 3 employed open syllables during familiarization and test.

[4]Two anonymous reviewers also did counts on two other databases, CHILDES and CELEX. The data from the CHILDES database supported our finding of an asymmetry in patterning of stressed syllables ending in consonants vs. starting in /t/. The data from the CELEX database did not reveal the same asymmetry. Perhaps one reason for the difference is that our analysis and the CHILDES-based analysis used words that occur in young children’s utterances. A related fact is that our analysis was restricted to mono- and disyllabic words. The reviewer using the CHILDES database also noted the large number of monosyllabic words. In contrast, the analysis using CELEX was unrestricted by number of syllables. The specific input that young learners receive and what they notice in that input clearly warrants further study.

EXPERIMENT 3

As noted above, Experiment 3 addressed both the possibility that younger infants make a generalization that older infants do not make, and a possible artifact of the stress syllables starting in /t/ stimuli of Experiment 2. If younger infants show evidence of learning the starts in /t/ generalization, while older learners do not, we would have evidence that constraints on generalization can be acquired over development, perhaps based on input statistics.

Methods

Materials

Materials for familiarization were identical to the stress syllables starting with /t/ familiarization stimuli of Experiment 2, in which the stressed syllables were *TU, TO,* and *TI*. As in Experiment 2, there were L1 and L2 versions. The test stimuli were changed from Experiment 2 to address the potential criticism that infants in the stress syllables starting with /t/ familiarization condition of Experiment 2 would have to generalize from three stressed open syllables (*TU, TO, TI*) to a stressed closed syllable (*TON*). Therefore, all instances of *TON* in the test stimuli from Experiment 2 were replaced by the open syllable *TA*.

Participants

Participants were 32 infants, 16 in a younger group and 16 in an older group. Half of the infants in each age group were familiarized with L1 and half with L2.

All infants were from English-speaking homes and had no reported history of hearing or speech/language disorder. The younger group ranged in age from 7 mos. 3 days to 8 mos. 2 days, with a mean of 7 mos. 19 days. The older group ranged from 8 mos. 16 days to 9 mos. 13 days, with a mean of 9 mos. 4 days. Four additional infants from each group were tested, but did not contribute usable data from at least eight test trials.

Procedures

Procedures were identical to those used in Experiments 1 and 2.

Results and Discussion

As in the previous experiments, a t-test was performed to determine if the difference in listening times on consistent vs. inconsistent trials varied as a function of familiarization language (L1 vs. L2; t (30) = 0.82, $p > 0.40$ 2-tailed). Because the two languages yielded similar patterns of listening time differences, subsequent analyses collapse across familiarization language. A 2 age (younger vs. older) X 2 consistency (consistent with familiarization language vs. inconsistent) ANOVA (see Figure 1) was performed on the mean listening times for each test condition. Neither main effect was significant (p's > .25), but the interaction was (F (1, 30) = 6.00, p < 0.02). Planned t-tests demonstrated that 7.5-month-olds discriminated the consistent vs. inconsistent items, listening longer to the former (t (15) = 2.58, p < 0.03 2-tailed). In contrast, 9-month-olds showed no evidence of learning (t (15) = 0.32, p = 0.38 2-tailed).

The data from the 9-month-olds in Experiment 3 replicate those in the stress syllables starting in /t/ condition of Exp. 2. Therefore, it appears that the learning mechanism requiring a small number of input types does not apply equally to all input. The 9-month-olds' data suggest that there are constraints on generalization, and in this case, that the constraints are consistent with properties typical of human languages. The fact that 7.5-month-olds were able to generalize under the same conditions that appeared to thwart generalization in the 9-month-olds demonstrates that the materials for the stress syllables starting with /t/ are not somehow unlearnable. Importantly, 7.5-month-olds' ability to make a generalization that is atypical of human languages suggests that all constraints may not be present from birth, but some may emerge over a relatively short developmental window. Given the statistics of English stressed syllables, a possible explanation for the difference between 7.5-month-olds and 9-month-olds is that the older infants have noted the strong (but imperfect) correlation between stressed syllables and syllables ending in a consonant. Thus, just as infants appear to use input statistics to focus on the phoneme inventory of the target language (Maye et al., 2002; Werker & Tees, 1984), and distinguish relevant from irrelevant information in language vs. music (Dawson & Gerken, 2006), they may also use input statistics to determine likely vs. unlikely bases of generalization about metrical phonology.

The difference between 7.5- and 9-month-olds is consistent with existing developmental data. For example, Jusczyk, Cutler, and Redanz (1993) found that 9-month-olds, but not 6-month-olds, preferred the strong-weak over weak-strong word lists, the former stress pattern being statistically more frequent in English (also see Echols, Crowhurst, & Childers, 1997). Such data suggest that 9-month-olds have noted something about lexical stress that younger infants have not. Turk, Jusczyk, and Gerken (1995) found that 9-month-olds prefer words that begin with a strong syllable that is also heavy, further indicating that 9-month-olds have noted statistical regularities about the segmental content of stressed syllables in English. More recently, Curtin, Mintz, and Christiansen (2005) found that even English-exposed 7-month-olds demonstrated a lexical segmentation strategy in which stressed syllables were treated as word-initial. What is added by the current data is the possibility that, while 7-month-olds have noted regularities about stress location on the beginning of English words, it is not until around

9 months that infants note the statistical bias for these stressed syllables to end with a consonant.

GENERAL DISCUSSION

We began with three questions concerning how infants generalize from limited linguistic input: What is the generalization mechanism? Does the mechanism apply to all input equally, or is it constrained? And if it is constrained, where do the constraints come from?

With respect to the generalization mechanism, we asked more specifically how many different input types are required for infants to make a generalization. This way of framing the mechanism question focuses on how rapidly infants are able to converge on a generalization without undue risk of overgeneralization. Consistent with research in several quite different learning domains, we found in Experiment 1 that three different input types are sufficient for generalization, but that multiple tokens of one input type are not. What sort of mechanism requires a small handful of examples, but will not generalize with a single example? To our knowledge, the only currently proposed generalization mechanisms for which the number of input examples is key are Bayesian hypothesis selection mechanisms (e.g., Tenenbaum & Griffiths, 2001; Xu & Tenenbaum, 2007a, 2007b). Although Parameter Theory holds that a parameter could logically be set with a single relevant example (e.g., Chomsky, 1981; Gibson & Wexler, 1994; Roeper & Williams, 1987), work within this framework does not typically specify the actual number of input examples required for real learners in real time. Associative approaches to generalization can sometimes generalize based on the same number of overall input as human infants (Shultz & Bale, 2001; Shultz & Gerken, 2005), and such approaches may well be able to capture the findings of Exp. 1. However, these approaches typically do not treat the specific number of input types needed for generalization as relevant to the comparison of model vs. human performance (Xu & Tenenbaum, 2007b). Given the growing consensus on the number of input types required for infant generalization across domains, proposals about generalization mechanisms should begin to take these data into account.

With respect to the question of whether the generalization mechanism applies to all input equally, the answer appears to be *no*. Nine-month-olds in Experiments 2 and 3 failed to generalize the linguistically unnatural principle stress syllables starting with /t/, despite the fact that the same amount of input was sufficient to generalize a principle that is typical of natural language.

With respect to the question of where such constraints on generalization might come from, the data from the 7.5-month-olds in Experiment 3 suggest that at least some constraints on generalization develop over time, perhaps as learners become increasingly familiar with the statistical regularities of their input.

Although English does not have an absolute requirement for heavy syllables to be stressed or for stressed syllables to be heavy, there is a strong statistical relation between a syllable's being stressed and its ending in a consonant. The data from Experiment 3, as well as a number of other developmental studies of English stress, suggest that 7-month-olds have not fully noted the asymmetrical relation between stress and syllable beginnings vs. endings. Therefore, they appear to be open to a greater range of generalizations about stress than are 9-month-olds.

Although all of the issues addressed in the current experiments have been to some extent addressed previously, the contribution of the current studies is that they examine questions about the generalization mechanism and constraints using minimally varying materials and in a linguistic domain that is potentially relevant to broad questions about linguistic generalization. The sketch of infant generalization that emerges from these data is one in which infants use their well-documented abilities to track input statistics (e.g., Gómez & LaKusta, 2004; Maye et al., 2002; Saffran et al., 1996) in service of determining which dimensions of their input account for the most variance. They then use some form of evidence assessment to determine the reliability of generalizations concerning these dimensions. This sketch can be put into Bayesian terms by saying that infants employ frequency-based priors to weight some generalizations more heavily in Bayesian hypothesis selection (Kemp, Perfors, & Tenenbaum, 2007; Tenenbaum, Griffiths, & Kemp, 2006).

One challenge inherent in filling in such a sketch entails determining the input dimensions over which infants perform statistics. Are these dimensions just the ones yielded by sensory information, or are more abstract dimensions also potential grist for the infant's statistical machinery? For example, is the generalization over heavy syllables in the current experiments based on acoustic properties of the stimuli (Gordon, 2005) or a more abstract property like "closed syllable"? The question of the dimensions over which statistical analyses are performed is a challenge not only to a hypothesis selection approach, but to all forms of inductive learning, including purely associative approaches. A second challenge that is particular to a hypothesis selection approach is determining the origin of the hypotheses (e.g., Shultz, 2007). Unlike finding statistical regularities in the input, hypothesis selection entails making some guess about the source(s) of those statistics. Although such an approach allows for rapid convergence on plausible bases of generalization, it also entails imputing considerable sophistication to learners and perhaps, particularly, to infant learners.

Regardless of whether the characterization of infant generalization suggested above is supported by future research, the current studies contribute to a growing literature on infants' abilities to generalize rapidly, without overgeneralizing, and with attention to the statistical properties of their input. Such studies provide boundary conditions for models of linguistic generalization and begin to elucidate why human language learners are such marvels of creativity.

REFERENCES

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational-experimental study. *Cognition, 90*, 119–161.

Berwick, R. C. (1986). Learning from positive-only examples: The subset principle and three case studies. In J. G. Carbonell, R. S. Michalski, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (Vol. 2, pp. 625–645). Los Altos, CA: Morgan Kauffman.

Chambers, K. E., Onishi, K. H., & Fisher, C. L. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition, 87*, B69–B77.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Netherlands: Foris.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology, 33*(A), 497–505.

Curtin, S., Mintz, T., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition, 96*, 233–262.

Dawson, C., & Gerken, L. A. (2006). 4-month-olds discover algebraic patterns in music that 7.5-month-olds do not. In *Proceedings of the Twenty-ninth Annual Conference of the Cognitive Science Society* (pp. 1198–1203). Mahwah, NJ: Erlbaum.

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 Associates.

Dresher, B. E. (1996). Introduction to metrical and prosodic phonology. In J. Morgan & K. Demuth (Eds.), *Signal to syntax*. Mahwah, NJ: 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, CA: Academic Press.

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

Echols, C., Crowhurst, M., & Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language, 36*, 202–225.

Gerken, L. A. (2004). Nine-month-olds extract structural principles required for natural language. *Cognition, 93*, B89–B96.

Gerken, L. A. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition, 98*, B67–B74.

Gibson, E., & Wexler, K. (1994). Triggers. *Linguistic Inquiry, 25*, 355–407.

Gold, E. M. (1967). Language identification in the limit. *Information and Control, 10*, 447–474.

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

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

Gordon, M. (2005). A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic Theory, 23*, 595–653.

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 & 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.

Halle, M., & Vergnaud, J. R. (1987). *An essay on stress*. Cambridge, MA: MIT Press.

Hayes, B. (1994). *Metrical Stress Theory*. Chicago: University of Chicago Press.

Hunter, M., & Ames, E. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research, 5*, 69–95.

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.

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.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science, 10*(3), 307–321.

Marcus, G. F., Fernandes, K., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science, 18*(5), 387–391.

Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101–B111.

Murphy, R. A., Mondragon, E., & Murphy, V. A. (2008). Rule learning by rats. *Science, 319(*5871), 1849–1851.

Needham, A., Dueker, G., & Lockhead, G. (2005). Infants' formation and use of categories to segregate objects. . *Cognition, 94*, 215–240.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*(127–162).

Prince, A. (1990). Quantitative consequences of rhythmic organization. In K. Deaton, M. Noske & M. Ziolkowsk (Eds.), *Papers from the Parasession on the Syllable in Phonetics and Phonology* (Vol. 26-II, pp. 355–398). Chicago: Chicago Linguistics Society.

Prince, A., & Smolensky, P. (1997). Optimality: From Neural Networks to Universal Grammar. *Science, 275*(5306), 1604–1610.

Quinn, P. C., & Bhatt, R. S. (2005). Learning perceptual organization in infancy. *Psychological Science, 16*(7), 511–515.

Roeper, T., & Williams, E. (Eds.). (1987). *Parameter Setting*. Dordrecht, Netherlands: Reidel.

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

Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology, 37*(1), 74–85.

Seidl, A., & Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development, 3–4*, 289–316.

Shultz, T. R. (2007). The Bayesian revolution approaches psychological development. *Developmental Science, 10*(3), 357–364.

Shultz, T. R., & Bale, A. C. (2001). Neural network simulation of infant familiarization to artificial sentences: Rule-like behavior without explicit rules and variables. *Infancy, 2*(4), 501–536.

Shultz, T. R., & Gerken, L. A. (2005). A model of infant learning of word stress. In *Proceedings of the Twenty-seventh Annual Conference of the Cognitive Science Society* (pp. 2015–2020). Mahwah, NJ: Erlbaum.

Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences, 24*, 629–640.

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences, 10*(7), 309–318.

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

Turk, A. E., Jusczyk, P. W., & Gerken, L. A. (1995). Do English-learning infants use syllable weight to determine stress? *Language & Speech, 38*(2), 143–158.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7, 49–63*.

Xu, F., & Tenenbaum, J. B. (2007a). Sensitivity to sampling in Bayesian word learning. *Developmental Science, 10*, 288–297.

Xu, F., & Tenenbaum, J. B. (2007b). Word learning as Bayesian inference. *Psychological Review, 114*, 245–272.