



# Musical Learning and Language Development

JENNY R. SAFFRAN

*Waisman Center and Department of Psychology, University of Wisconsin – Madison, Madison, Wisconsin 53706, USA*

AU:  
Do you wish  
to insert  
several  
headings  
through the  
text?

**ABSTRACT:** How do infant learners acquire structure, given complex environments? In this chapter, we consider the role played by statistical learning—tracking patterns in the environment—in the acquisition of language and music. The results from a series of experiments suggest that similar learning mechanisms may operate in both domains, but that these mechanisms are also influenced by domain-specific perceptual biases and input structure.

**KEYWORDS:** infant learning; statistical learning; music; language

How do infants begin to make sense of their world? Filled with patterned stimuli and replete with sensory information of all kinds, infants must somehow begin to derive structure amidst all of the noise—both figuratively and literally—that characterizes their environments. Innate predispositions are one way to sort through the noise to find the signal. Learning is another. In this chapter we consider one type of learning—statistical learning, or the detections of patterns in the environment—and the possible role played by statistical learning mechanisms in the acquisition of complex structure.

In particular, we focus on the acquisition of language and music. These two domains share much in common structurally: both are auditory (with the exception of signed languages), highly patterned, and internally consistent. Music and language also share the distinction of being two of the stimuli that are most interesting to developing humans. Along with faces, young infants are most consistently engaged by speech and by music (making singing a particularly welcome combination of face, speech, and music). Linguistic and musical knowledge are arguably the most complex systems universally acquired by humans early in life. Finally, both language and music consist of some structures that are present cross-culturally, along with other structures that vary across cultures. This means that young learners must be capable of learning in these domains, or else they could not acquire the features of their native language that are not universal (e.g., the particular sounds, words, and grammatical devices of English versus Thai) or the aspects of their native musical system that are not universal (e.g., the scale tones of Western tonal music versus Javanese music).

Address for correspondence: Jenny R. Saffran, Department of Psychology, University of Wisconsin – Madison, Madison, WI 53706. Voice: 608-262-9942; fax: 608-262-4029. jsaffran@wisc.edu

Ann. N.Y. Acad. Sci. 999: 1–5 (2003). © 2003 New York Academy of Sciences.  
doi: 10.1196/annals.1284.001



How might this learning occur? To answer this question, we have begun to explore the potential contributions of statistical learning: the ability to track consistent patterns in the input to discover units and structures. In particular, we have explored the role of statistical learning in the discovery of consistent patterns that lie within continuous streams of sound. For example, consider the problems faced by learners exposed to a novel language. In order to acquire the language, learners must find the sound sequences corresponding to words in speech; learners can then map those sound sequences to meanings and discover their syntactic relationships. However, the initial task of finding words is nontrivial, because fluent speech does not contain pauses or other consistent acoustic cues marking word boundaries.<sup>1</sup> Nevertheless, infants can rapidly extract words from continuous speech by 7.5 months of age.<sup>2</sup>

There are multiple cues that infants might exploit to discover word boundaries in fluent speech.<sup>3</sup> The statistical properties of language might be particularly useful to infants engaged in segmenting words.<sup>4,5</sup> For example, consider the two-word sequence *pretty baby*. The syllable *pre* precedes a small set of other syllables, including *ty*, *tend*, and *cede*; the probability that *pre* is followed by *ty* is thus quite high (roughly 80% in speech to young infants). However, because the syllable *ty* occurs word-finally, it can be followed by any syllable that can begin a word in the English language. Thus, the probability that *ty* is followed by *ba*, as in *pretty baby*, is extremely low (roughly 0.03% in speech to young infants). Given the statistical properties of the input language, the ability to track sequential probabilities could be an extremely useful tool for infant language learners.

To ask whether infants possess statistical learning mechanisms usable for word segmentation, we exposed 8-month-olds to a “nonsense” language in which the only cues to word boundaries were the statistical properties of the syllable sequences.<sup>5,6</sup> The infants first listened to a 2-minute continuous sequence of syllables, in which “words” were present but word boundaries were not marked, for example, *golabup-abikututibubabupugolabu etc.* We then tested the infants to determine if they could discriminate the words in the language from sequences spanning word boundaries (part-words); these were sequences that the infants had heard, but that did not contain the statistical properties of words. To succeed in this task, the infants had to track the statistical properties of the input. We tested the infants using the Headturn Preference Procedure,<sup>7</sup> in which discrimination of the two types of test items (words versus part-words) was assessed by calculating how long infants listened to each test item. Our results confirmed that infants can indeed use sequential statistics to find word boundaries, despite the brevity of their exposure to the novel language and the lack of any other cues to word boundaries.

Our language experiments as well as those from other research groups demonstrate that infants have access to powerful statistical learning mechanisms that are deployed in the absence of any training or explicit reinforcement and that may be useful for many aspects of language acquisition.<sup>8-10</sup> One immediate question raised by these results is whether the learning mechanism in question is dedicated solely to the acquisition of language or whether it might be used by infants for learning in other domains. Although the issue of domain-specificity versus domain-generality (modularity) plays a prominent role in discussions of adult psychological and neural structure, studies of learning have typically not directly addressed this issue directly (for discussion, see Elman *et al.*, Ref. 11).

The domain of music provided a natural arena in which to extend our work on statistical language learning. We thus developed a pseudo-musical analog of our language task by translating each syllable of our nonsense language into a sine-wave tone; for example, *golabu* became CFE. We then exposed both infants and adults to a statistical learning task in which “tone-words” were discoverable solely by virtue of their statistical properties.<sup>12</sup> As in our linguistic tasks, we discovered that learners were adept at tracking the probabilities with which particular tones co-occurred to locate the boundaries between tone-words. It is thus likely that learners can use the same statistical learning mechanisms in both linguistic and musical tasks; subsequent studies have since demonstrated that the same mechanism can also be used for visuospatial and visuomotor learning.<sup>13–15</sup>

The musical nature of the tone sequence task we used raised an interesting issue concerning the nature of musical learning. Consider the tone sequences used in our experiment, for example, CFEGAD#F#A#B etc. What are the units over which learners computed statistics? There are multiple perceptual primitives that might have entered into learners’ computations. Given a sequence like CFE, learners might have tracked the probabilities with which the absolute pitches C, F, and E co-occurred, an analogy with tracking the probabilities with which the syllables *go*, *la*, and *bu* co-occurred. Alternatively, infants might have tracked the probabilities with which the relative pitches, or the intervals between those pitches, co-occurred: what was the likelihood that an ascending perfect fourth (the interval between C and F) was followed by a descending minor second (the interval between F and E)?

Given the design of our original tone-learning experiment, it was impossible to determine which of these perceptual primitives was tracked during learning; did learners detect the statistical properties of absolute pitch sequences or relative pitch sequences? Learners at various ages appear to have access to both types of cues, but they may use one type of pitch information preferentially in some tasks relative to others.<sup>16,17</sup> To address this question, we manipulated our tone sequence statistical learning task so that in some conditions, successful test discrimination required the use of absolute pitch cues during learning, whereas in other conditions, successful test discrimination required the use of relative pitch cues during learning.<sup>18</sup> The results suggested that learners of different ages capitalized on different perceptual primitives to perform our task: 8-month-old infants tracked absolute pitch cues, while adults tracked relative pitch cues, given exactly the same stimulus materials.

Because infants in other types of tasks are skilled at detecting relative pitch information,<sup>17</sup> we hypothesized that the lack of musicality in our tone sequence stimuli might have influence infants’ performance. The materials used in the Saffran and Griepentrog<sup>18</sup> experiment were atonal and did not conform to any of the conventions of Western tonal music. It is therefore possible that infants’ reliance on absolute pitch cues in these experiments was a function of how the infants processed our stimuli; the lack of melodic structure may have led infants to focus on the nongeneralizable absolute pitch information rather than on the melodic information carried by relative pitch. To test this hypothesis, we exposed infants and adults to a tone sequence segmentation task in which the materials conformed to the key of C major.<sup>19</sup> We reasoned that the increment in musicality afforded by the inclusion of some tonal structure might lead infants to begin to track relative pitch in our materials. Interestingly, however, the infants continued to depend on absolute pitch cues for the test discrimination rather than relative pitch cues.



Why might some tasks lead infants to show one type of perceptual processing whereas other tasks suggest the use of a different set of perceptual cues? One possibility concerns an interaction between learners' existing perceptual sensitivities and the structure of the tasks themselves. This hypothesis was previously explored in the domain of pitch processing by birds.<sup>20</sup> When European starlings were presented with a pitch discrimination task that could be performed using either absolute or relative pitch cues, the birds initially solved the task using absolute pitches. However, when the task was changed to require transfer of the pitch sequences, the birds began to use relative pitch cues. These results suggest that the birds had access to both types of pitch cues, but that the demands of the task determined which pitch dimensions they used to perform the test discriminations.

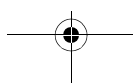
It is possible that a similar process occurs during pitch learning in human infants. That is, infants may have access to both types of pitch cues, and the structure of the learning task and/or test discrimination affects which aspects of pitch are tracked. If this hypothesis is correct, then when absolute pitch cues are rendered unreliable, infants should begin to track relative pitch cues instead. We thus created tone sequences that no longer contained consistent absolute pitch sequences; this was accomplished by continually transposing the "tone words."<sup>21</sup> Relative pitch sequences, however, remained highly predictable. Given these materials, infants were able to capitalize on relative pitch cues, which they failed to do when absolute pitch information was also available. By essentially removing absolute pitch cues as a signal of structure in the input, infants began to reliably discriminate the test materials using relative pitch cues.

The results of this line of research, taken together, support a nuanced view of infant learning capabilities. On the one hand, infant learners are powerful: they can detect structure using statistical cues, this learning mechanism operates rapidly and in the absence of reinforcement, and similar learning occurs across domains. On the other hand, learning is constrained, both by the infants' perceptual capabilities and by the structure of the material to be learned. Infants may avoid William James' "blooming buzzing confusion," at least in part, by virtue of a perceptual system that weights some cues more highly than others and by a learning system that can flexibly adapt to the task at hand. Little is yet known about the role played by such systems in acquiring the native musical system and how the infants' inherent preferences and biases interact with information to be learned. We are hopeful, however, that comparisons of learning across domains, and careful consideration of the tasks facing learners within a given domain, will lead us to understand how the developing brain so masterfully acquires the intricacies of the native environment.

#### ACKNOWLEDGMENTS

The preparation of this chapter was funded by a grants to J.R.S. from the International Foundation for Music Research, National Institute of Child Health and Human Development (RO1HD37466), and National Science Foundation (BCS-9983630). Thanks to Erik Thiessen for helpful comments on the manuscript.

AU: OK as written out?



## REFERENCES

1. COLE, R. & J. JAKIMIK. 1980. A Model of Speech Perception. Lawrence Earlbaum, Hillsdale, NJ.
  2. JUSCZYK, P.W. & R.N. ASLIN. 1995. Infants' detection of the sound patterns of words in fluent speech. *Cognit. Psychol.* **29**: 1–23.
  3. JUSCZYK, P.W. 1999. How infants begin to extract words from fluent speech. *Trends Cognit. Sci.* **3**: 323–327.
  4. SAFFRAN, J.R., E.L. NEWPORT & R.N. ASLIN. 1996. Word segmentation: the role of distributional cues. *J. Mem. & Lang.* **35**: 606–621.
  5. SAFFRAN, J.R., R.N. ASLIN & E.L. NEWPORT. 1996. Statistical learning by 8-month-old infants. *Science* **274**: 1926–1928.
  6. ASLIN, R.N., J.R. SAFFRAN & E.L. NEWPORT. 1998. Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* **9**: 321–324.
  7. KEMLER NELSON, D.G., P.W. JUSCZYK, D.R. MANDEL, *et al.* 1995. The Headturn Preference Procedure for testing auditory perception. *Infant Behav. Dev.* **18**: 111–116.
  8. SAFFRAN, J.R. & D.P. WILSON. 2003. From syllables to syntax: multi-level statistical learning by 12-month-old infants. *Infancy*. In press.
- AU: Update Ref. 8?
9. THIESSEN, E.D. & J.R. SAFFRAN. 2003. When cues conflict: statistical and stress cues in infant word segmentation. *Devel. Psychol.* In press.
- AU: Update Ref. 9?
10. MAYE, J., J.F. WERKER, & L. GERKEN. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* **82**: B101–111.
  11. ELMAN, J.L., E.A. BATES, *et al.* 1996. Rethinking Innateness: A Connectionist Perspective on Development. MIT Press. Cambridge, MA.
  12. SAFFRAN, J.R., E.K. JOHNSON, R.N. ASLIN & E.L. NEWPORT. 1999. Statistical learning of tone sequences by human infants and adults. *Cognition* **70**: 27–52.
  13. HUNT, R.H. & R.N. ASLIN. 2001. Statistical learning in a serial reaction time task: simultaneous extraction of multiple statistics. *J. Exp. Psychol. Gen.* **130**: 658–680.
  14. FISER, J. & R.N. ASLIN. 2002. Statistical learning of new visual features by infants. *Proc. Natl. Acad. Sci. USA* **99**: 15822–15826.
  15. KIRKHAM, N.Z., J.A. SLEMMER & S.P. JOHNSON. 2002. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* **83**: B35–B42.
  16. TAKEUCHI, A.H. & S.H. HULSE. 1993. Absolute pitch. *Psychol. Bull.* **113**: 345–361.
  17. TREHUB, S.E., G. SCHELLENBERG & D. HILL. 1997. The origins of music perception and cognition: a developmental perspective. *In Perception and Cognition of Music*. I. Deliège & J. Sloboda, Eds. :103–128. Psychology Press. East Sussex, UK.
  18. SAFFRAN, J.R. & G.J. GRIEPENTROG. 2001. Absolute pitch in infant auditory learning: evidence for developmental reorganization. *Dev. Psychol.* **37**: 74–85.
  19. SAFFRAN, J.R. 2003. Absolute pitch in infancy and adulthood: the role of tonal structure. *Dev. Sci.* **6**: 37–45.
  20. MACDOUGALL-SHACKLETON, S.A. & S.H. HULSE. 1996. Concurrent absolute and relative pitch processing by European starlings. *J. Comp. Psychol.* **110**: 139–146.
  21. SAFFRAN, J.R., K. REECK, A. NIEBUHR & D. WILSON. 2003. Changing the tune: the structure of the input affects infants' use of absolute and relative pitch. Manuscript submitted.