The acquisition of language by children

Jenny R. Saffran*, Ann Senghas†, and John C. Trueswell‡§

*Department of Psychology, University of Wisconsin, Madison, WI 53706-1696; †Department of Psychology, Barnard College of Columbia University, New York, NY 10027-6598; and ‡Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104-6196

Imagine that you are faced with the following challenge. You must discover the internal structure of a system that contains tens of thousands of units, all generated from a small set of materials. These units, in turn, can be assembled into an infinite number of combinations. Although only a subset of those combinations is correct, the subset itself is for all practical purposes infinite. Somehow you must converge on the structure of this system to use it to communicate. And you are a very young child.

This system is human language. The units are words, the materials are the small set of sounds from which they are constructed, and the combinations are the sentences into which they can be assembled. Given the complexity of this system, it seems improbable that mere children could discover its underlying structure and use it to communicate. Yet most do so with eagerness and ease, all within the first few years of life.

Below we describe three recent lines of research that examine language learning, comprehension, and genesis by children. We begin by asking how infants break into the system, finding the words within the acoustic stream that serves as input to language learning. We then consider how children acquire the ability to rapidly combine linguistic elements to determine the relationships between these elements. Finally, we examine how children impose grammatical structure onto their perceived input, even to the extent of creating a new language when none is available. These investigations provide insight into the ways in which children extract, manipulate, and create the complex structures that exist within natural languages.

Discovering the Units of Language

Before infants can begin to map words onto objects in the world, they must determine which sound sequences are words. To do so, infants must uncover at least some of the units that belong to their native language from a largely continuous stream of sounds in which words are seldom surrounded by pauses. Despite the difficulty of this reverse-engineering problem, infants successfully segment words from fluent speech from ≈7 months of age.

How do infants learn the units of their native language so rapidly? One fruitful approach to answering this question has been to present infants with miniature artificial languages that embody specific aspects of natural language structure. Once an infant has been familiarized with a sample of this language, a new sample, or a sample from a different language, is presented to the infant. Subtle measures of surprise (e.g., duration of looking toward the new sounds) are then used to assess whether the infant perceives the new sample as more of the same, or something different. In this fashion, we can ask what the infant extracted from the artificial language, which can lead to insights regarding the learning mechanisms underlying the earliest stages of language acquisition (1).

One important discovery using this technique has come from the work of Saffran and colleagues (2–5), who have examined the powerful role that statistical learning—the detection of consistent patterns of sounds—plays in infant word segmentation. Syllables that are part of the same word tend to follow one another predictably, whereas syllables that span word boundaries do not. In a series of experiments, they found that infants can detect and use the statistical properties of syllable co-occurrence to segment novel words (2). More specifically, infants do not detect merely how frequently syllable pairs occur, but rather the probabilities with which one syllable predicts another (3). Thus, infants may find word boundaries by detecting syllable pairs with low transitional probabilities. What makes this finding astonishing is that infants as young as 8 months begin to perform these computations with as little as 2 min of exposure. By soaking up the statistical regularities of seemingly meaningless acoustic events, infants are able to rapidly structure linguistic input into relevant and ultimately meaningful units.

To what extent do infants' capacities to detect the statistics of linguistic sounds extend to learning in nonlinguistic domains? Interestingly, infants are also able to detect the probabilities with which musical tones predict one another, suggesting that the statistical learning abilities used for word segmentation may also be used for learning materials such as music (4). In particular, infants, but not adults, can track the statistical structure of sequences of absolute pitches in a tone sequence learning task (5). These findings suggest that at least some of the statistical learning mechanisms described above are not applied solely to language learning.

The Child Parser: Packaging Words into Meaningful Units

Discovering the words of a language, and what they mean in the world, is only the first step for the language learner. Children must also discover how the distribution of these elements, including grammatical endings (*-s, -ed, -ing*) and function words (*of, to, the*) convey the further combinatorial meaning of an utterance. That is, children must implicitly discover and use the grammar of their language to determine who-did-what-to-whom in each sentence. This applies even for simple sentences like *Mommy gave Daddy the milk* as opposed to *Daddy gave Mommy the milk.* The parsing process is therefore an essential component of the language comprehension device, because it allows children to assemble strings of elements in such a way as to compute crucial, and even novel, relational conceptions of the world.

Adults are quite adept at parsing sentences to determine relational meaning. In fact, studies of adult language comprehension indicate that readers and listeners are so skilled at this process that they typically achieve it in real time, as each word is perceived. By measuring eye fixation and reaction time midsentence, these studies confirm that adults rapidly package incoming words into likely phrases using a variety of probabilistic cues gleaned from the sentence and its referential context (e.g., refs. 6 and 7).

Recently, Trueswell and colleagues (8–10) have examined how this rapid parsing system develops. In a series of studies, eye movements of children age 4 and older were recorded as they heard

---

This paper is a summary of a session presented at the 12th annual symposium on Frontiers of Science, held November 2–4, 2000, at the Arnold and Mabel Beckman Center of the National Academies of Science and Engineering in Irvine, CA.

§To whom reprint requests should be addressed. E-mail: trueswel@psych.upenn.edu.

www.pnas.org/cgi/doi/10.1073/pnas.231498898

instructions to move objects about on a table. Children's visual interrogation of the scene during the speech provided a window into the ongoing interpretation process. Of particular interest was their reaction to ambiguous instructions that required an implicit grammatical choice, e.g., *Tap the doll with the stick.* Here the phrase *with the stick* can be linked to the verb *Tap*, indicating how to do the tapping, or it can be linked to the noun *doll*, indicating which doll to tap. Adults tend to rely on the referential context when making choices like these, picking the analysis that is most plausible given the current scene. Which analysis did children choose? It depended heavily on the kind of linguistic cues found in the utterance itself. For instance, regardless of how likely the analysis was given the scene, children would interpret *with the stick* as how to carry out the action when the verb was of the sort like *Tap,* which tends to mention an instrument as part of its event. In contrast, they would interpret this same phrase as picking out a particular doll when the verb was of the sort that tends not to mention an instrument, e.g., *Feel* (8, 10).

Thus, like the Saffran *et al.* infants who used probabilistic cues to package syllables into likely words, older children package words into likely phrases using similar distributional evidence regarding these larger elements. Further experience is apparently necessary to detect the contingencies of when phrases are likely in given referential settings. Indeed, Trueswell *et al.* found that by age 8, children begin parsing ambiguous phrases in a context-contingent manner (8).

Language Acquisition as Creation

Although distributional analyses enable children to break into the words and phrases of a language, many higher linguistic functions cannot be acquired with statistics alone. Children must discover the rules that generate an infinite set, with only a finite sample. They evidently possess additional language-learning abilities that enable them to organize their language without explicit guidance (11). These abilities diminish with age (12) and may be biologically based (13). However, scientific efforts to isolate them experimentally encounter a methodological complication: given that today's languages were acquired by children in the past, language input to children already includes products of innate biases. It is therefore difficult to determine whether any particular linguistic element observed in a child's language is inborn or derived.

We can break this logical circle by examining those rare situations in which the language environment is incomplete or impoverished. Can children who are deprived of exposure to a rich, complete language nevertheless build a structured native language? The recent situation of deaf children in Nicaragua presents such a case.

Nicaraguan Sign Language first appeared only two decades ago among deaf children attending new schools for special education in Managua, Nicaragua. Their language environment provided incomplete linguistic input: they could not hear the Spanish spoken around them, and there was no previously developed sign language available. The children responded by producing gestures that contained grammatical regularities not found in their input, and in the process created a new, natural sign language. The language continues to develop and change as new generations of children enter school and learn to sign from older peers. Thus, there is a measurable discrepancy between the input to which each wave of arrivals was exposed and the language they acquired, evident in comparisons between the first wave of children (now adults in their 20s) and the second wave of children (now adolescents) (14).

One such development is in their expression of semantic roles, that is, in their use of language structure to indicate who-did-what-to-whom in an event (as in the difference in English between *the girl pushes the boy* and *the boy pushes the girl*). The first group of children invented signs for the things they needed to talk about (*girl, boy, push, give, fall,* etc.) and immediately began developing ways to string them together into sentences. For example, to describe events, they would name each participant followed by its role, such as *girl push boy fall*, or *boy give girl receive.*

The second wave of children to acquire the language added even more structure. Within a few years, not only was the order of the signs important, it also mattered where signs were produced. Once the boy and girl had been mentioned, *push* produced to one side would mean the girl was pushed, and to the other side would mean the boy was pushed. The children had developed spatial devices to indicate semantic roles, a feature typical of sign languages (e.g., ref. 15). The use of such constructions is evident today among Nicaraguan adolescents, but not adults (16). In fact, without contextual cues, adolescent signers will give a more narrow interpretation than that intended by adult signers, despite the fact that such signing represents their initial input.

These findings indicate that children can apply their own organizational biases to input that is not richly structured. Even when cues are absent from their environment, children can turn to inborn learning abilities to converge on a common language as a community.

The Acquisition of Language by Children

These examples of language learning, processing, and creation represent just a few of the many developments between birth and linguistic maturity. During this period, children discover the raw materials in the sounds (or gestures) of their language, learn how they are assembled into longer strings, and map these combinations onto meaning. These processes unfold simultaneously, requiring children to integrate their capacities as they learn, to crack the code of communication that surrounds them. Despite layers of complexity, each currently beyond the reach of modern computers, young children readily solve the linguistic puzzles facing them, even surpassing their input when it lacks the expected structure.

No less determined, researchers are assembling a variety of methodologies to uncover the mechanisms underlying language acquisition. Months before infants utter their first word, their early language-learning mechanisms can be examined by recording subtle responses to new combinations of sounds. Once children begin to link words together, experiments using real-time measures of language processing can reveal the ways linguistic and nonlinguistic information are integrated during listening. Natural experiments in which children are faced with minimal language exposure can reveal the extent of inborn language-learning capacities and their effect on language creation and change. As these techniques and others probing the child's mind are developed and their findings integrated, they will reveal the child's solution to the puzzle of learning a language.

1. Jusczyk, P. W. (1997) *The Discovery of Spoken Language* (MIT Press, Cambridge, MA).
2. Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996) *Science* **274,** 1926–1928.
3. Aslin, R. N., Saffran, J. R. & Newport, E. L. (1998) *Psychol. Sci.* **9,** 321–324.
4. Saffran, J. R., Johnson, E. K., Aslin, R. N. & Newport, E. L. (1999) *Cognition* **70,** 27–52.
5. Saffran, J. R. & Griepentrog, G. J. (2001) *Dev. Psychol.* **37,** 74–85.
6. Altmann, G. & Steedman, M. (1988) *Cognition* **30,** 191–238.
7. Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (1995) *Science* **268,** 1632–1634.
8. Trueswell, J., Sekerina, I., Hill, N. & Logrip, M. (1999) *Cognition* **73,** 89–134.
9. Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R. & Trueswell, J. C. (2000) *J. Psycholinguistic Res.* **29,** 597–626.
10. Snedeker, J., Thorpe, K. & Trueswell, J. C. (2001) in *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, eds. Moore, J. & Stenning, K. (Earlbaum, Hillsdale, NJ), pp. 964–969.
11. Chomsky, N. (1965) *Aspects of the Theory of Syntax* (MIT Press, Cambridge, MA).
12. Newport, E. L. (1990) *Cognit. Sci.* **14,** 11–28.
13. Pinker, S. (1994) *The Language Instinct* (Morrow, New York).
14. Senghas, A. & Coppola, M. E. V. (2001) *Psychol. Sci.* **12,** 323–328.
15. Supalla, T. (1982) Ph.D. thesis (University of California, San Diego).
16. Senghas, A., Coppola, M., Newport, E. L. & Supalla, T. (1997) in *Proceedings of the 21st Annual Boston University Conference on Language Development*, eds. Hughes, E., Hughes, M. & Greenhill, A. (Cascadilla, Boston), pp. 550–561.