

# Signal-Driven Computations in Speech Processing

Marcela Peña,<sup>1</sup> Luca L. Bonatti,<sup>1,2</sup> Marina Nespor,<sup>3</sup>  
Jacques Mehler<sup>1,4\*</sup>

Learning a language requires both statistical computations to identify words in speech and algebraic-like computations to discover higher level (grammatical) structure. Here we show that these computations can be influenced by subtle cues in the speech signal. After a short familiarization to a continuous speech stream, adult listeners are able to segment it using powerful statistics, but they fail to extract the structural regularities included in the stream even when the familiarization is greatly extended. With the introduction of subliminal segmentation cues, however, these regularities can be rapidly captured.

To learn an unknown language, listeners must segment connected speech into constituents and discover how words are organized. When adults try to cope with an unknown language or when infants learn their native language, they do so by listening to speech before they know either the words or the grammatical system of that language, and without receiving explicit instruction. To extract words as well as their organization from the speech stream, infants and adults must possess efficient computational procedures.

Several solutions have been proposed to

account for speech segmentation (1, 2). In particular, some investigators (3–5) have shown that adults and 8-month-old infants confronted with unfamiliar concatenated artificial speech tend to infer word boundaries at loci where the transitional probability between two adjacent syllables drops. That is, word boundaries are inferred between two syllables that rarely appear in sequence and not between two syllables that always appear together (6). Saffran *et al.* (5) demonstrated that participants exposed for several minutes to continuous speech judge trisyllables delimited by dips in transitional probability as being more familiar than trisyllables enclosing a transitional probability dip. Other studies have helped establish the importance of statistics in parsing speech as well as nonspeech sequences: adults can take advantage of statistics to segment speech streams, sequences of tones (7), and sequences of visual stimuli

(8–10), among other types of sequences.

As to the mechanisms responsible for the extraction of structural information, little is known. In one study (11), 7-month-old infants behaved as if they had inferred a rule after having been familiarized with a large number of trisyllabic items consistent with it. After familiarization, infants were presented with previously unheard items, and they behaved differently according to whether or not the items conformed to the rule. This result was observed using segmented strings of items composed of three separate consonant-vowel syllables (12). This suggests that infants tend to extract rule-like regularities, at least when they process a corpus of clearly delimited items. This study emphasizes the specific computational abilities that favor the discovery of the structural properties of a corpus. Conceivably, in the absence of such abilities, language would be impossible to acquire.

Assessing the scope and limits of statistical and structural computations for learning words and grammar in language remains an elusive problem. One reason is that the methodologies and stimuli used in the above-cited studies are sufficiently different that the relative importance of the two underlying mechanisms cannot be directly compared. The aim of our study is to explore, by means of easily comparable experimental situations, what such mechanisms accomplish and when precisely they operate in language processing. To this purpose, building on a suggestion by Newport and Aslin (13), we explore whether participants can segment a stream of speech by means of nonadjacent transition probabil-

<sup>1</sup>International School for Advanced Studies, Trieste, Italy. <sup>2</sup>University of Paris VIII at Saint Denis, France. <sup>3</sup>University of Ferrara, Italy. <sup>4</sup>Laboratoire de Science Cognitive et Psycholinguistique, CNRS and EHESS, Paris, France.

\*To whom correspondence should be addressed. E-mail: mehler@sissa.it

REPORTS

ities, and we also ask whether the same computations are used to promote the discovery of its underlying grammatical structure.

In experiment 1, we used essentially the same procedure described in Newport and Aslin (13) but tested speakers of French (14). We presented a 10-min-long stream of synthetic speech syllables composed of trisyllabic items mainly characterized by their non-adjacent transitional probabilities. We chose to call this the “AXC language” to denote that for every item, A predicts exactly C. We familiarized adults with a continuous stream of AXC “words.” An  $A_iXC_i$  item appears with three different X’s creating a family of words, for example [puliki], [puRaki], [pu-foki]. Three such families were pseudo-randomly arranged into the stream (15); we used the same three X syllables for all three families. Hence, the transitional probabilities between any  $A_i$  and the adjacent X, or between any X and any adjacent  $C_i$ , is 0.33; the transitional probabilities between the last syllable of any item and the first syllable of the following one is 0.5; and the transitional probabilities between any  $A_i$  and its  $C_i$  is always 1. If French speakers can segment on the basis of nonadjacent syllable transitional probabilities, they will organize the stream into meaningless trisyllabic words. To evaluate this, after familiarization we presented participants ( $n = 14$ ) with couples of words ( $A_iXC_i$ ) and “part words” ( $C_kA_iX$  or an  $XC_kA_i$ ), asking them to judge, for each couple, which item seemed to them more like a word of the imaginary language they heard in the stream. The results, presented in Fig. 1A, show that words were selected significantly more often than part words (16). Thus, participants appear to take advantage of nonadjacent statistical dependencies between consonant-vowel syllables to automatically segment a continuous stream (17). This outcome shows that humans can perform more powerful statistical computations than previously reported.

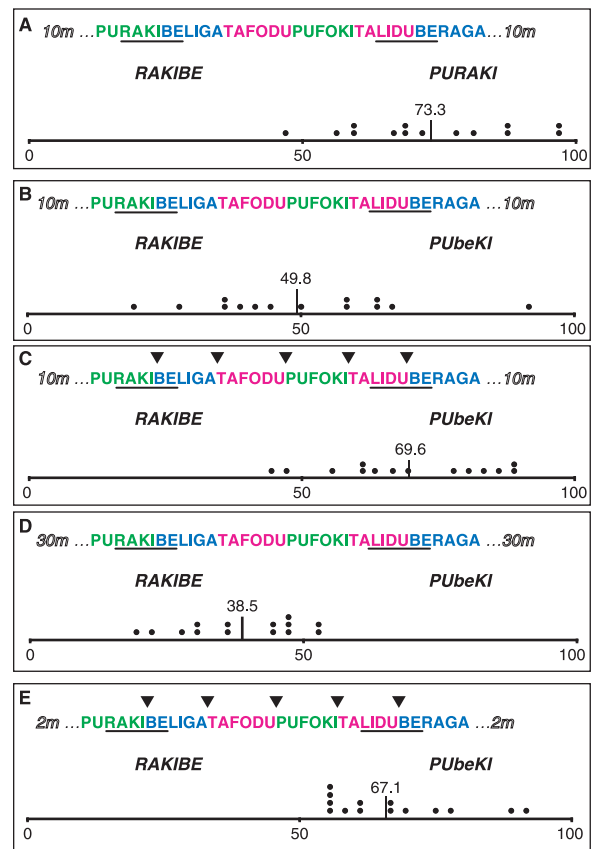
Experiment 1, which showed that distant transitional probabilities are used to identify items in the stream, could also be interpreted as evidence for the learning of a structural regularity. That is, our AXC language also respects the generalization “If  $A_i$  occurs then  $C_i$  will follow after an intervening X.” When participants in experiment 1 select words over part words, do they only identify the words in the stream, or do they also identify the structural generalization? We addressed this question in experiment 2. We familiarized another group of adults ( $n = 14$ ) with the same stream used in experiment 1. However, during the test phase, one of the items of each test pair had not appeared in the stream but was congruent with the generalization (we call it the “rule word”), whereas the other was a part word as defined above. The rule words

have an intervening syllable that appears in the stream but never between  $A_i$  and  $C_i$ ; thus, rule words have a novel surface form. The part words are the same as the ones used in experiment 1 (18) and, although relatively infrequent, have a familiar surface form. The results of this experiment are presented in Fig. 1B. Participants failed to choose the rule words over the part words. This shows that they failed to discover the underlying regularity; had they done so, they would have selected rule words over part words. We can therefore conclude that a computational mechanism sufficiently powerful to support segmentation on the basis of nonadjacent transitional probabilities is insufficient to support the discovery of the underlying grammatical-like regularity embedded in a continuous speech stream (19).

In experiment 1, we showed that participants are able to compute nonadjacent transitional probabilities, whereas in experiment 2 we showed that they fail to exploit the outcome of the same computation to extract

the underlying regularities. Why are statistical computations efficient for identifying components of a stream but not for achieving generalizations? We conjecture that this reflects the fact that the discovery of components of a stream and the discovery of structural regularities require different sorts of computations (20, 21), each requiring a specific input. When given a continuous speech stream, the listener must first “chunk” it into discrete word candidates. The role of statistical computations is precisely that of attaining this segmentation into components. In contrast, to discover grammatical-like regularities, the listener must be able to inspect memory traces of such discrete representations and project generalizations that encompass but go beyond the surface form of these items in memory. This process of projecting generalizations, we submit, may not be statistical in nature. This conjecture leads us to make the following prediction: it is the type of signal that is being processed rather than the amount of familiarization that determines

**Fig. 1.** In experiments 1 to 5, the familiarization stream consists of meaningless, monotonous synthesized speech composed by trisyllabic items in which the transitional probability between the first and third syllables is 1.0 (we call these the “words”). The first line of each frame contains a sample of the familiarization stream, and the numbers indicate its duration. Different colors highlight words; examples of part words are underlined. The second line contains an example of a test pair. Test pairs always compare a part word to either a word or a rule word, which is obtained from a word by substituting its middle syllable with a syllable that never occurred in that position during familiarization (lowercase in the examples). After being instructed to listen carefully to the familiarization stream, participants were asked to decide, for each test pair, which item looked more like a word of the imaginary language. The dots over the line at the bottom of each frame represent individual scores; the number above the vertical mark indicates the general mean. Each dot represents the percentage of choices for either words (A) or rule words (B to E) of individual subjects averaged across items. (A) After 10-min familiarization participants preferred words to part words ( $P < 0.0005$ ), indicating that they can segment the stream on the basis of distant syllable transitional probabilities. (B) After 10-min familiarization participants did not show a preference for rule words over part words (n.s.). (C) After 10-min familiarization with a stream that contains, at the edge of each word, 25-ms subliminal gaps (indicated by triangles above the first line), participants showed a preference for rule words over part words ( $P < 0.0005$ ). (D) Increasing familiarization to a continuous stream to 30 min induced participants to prefer part words over rule words ( $P < 0.002$ ). (E) A familiarization reduced to 2 min with a stream containing 25-ms gaps led participants to prefer rule words over part words ( $P < 0.0005$ ).



the type of computation in which participants will engage.

This prediction has two major consequences. First, changing a signal even slightly may induce a change in computation. Second, if segmentation and generalizations arise from different modes of computation, then the critical factor determining the selection of one or the other computation is not the amount of evidence that the listeners have received during familiarization, but rather the manner in which the input stream is packaged.

To test the first consequence, we reasoned as follows. If listeners are exposed to the stream used in experiment 2, with subtle segmentation cues added to it, then they will be relieved of the task of computing probabilities and will be able to capture the generalizations that otherwise eluded them. In experiment 3, our aim was to introduce cues to segmentation in the signal without making the participants aware of them. To this end, we introduced subliminal gaps of 25-ms duration (22) after each word in the familiarization stream, leaving the stream otherwise identical to the one used in the previous experiments. We predict that although the stream used in experiments 1 and 2 triggers statistical computations, the stream in experiment 3 will prompt participants to respond to its structure. After participants ( $n = 14$ ) were familiarized with the new stream, they were tested with the same pairs of items used in experiment 2. Figure 1C illustrates that participants judged that the rule word was more likely to be a word of the imaginary language than the part word, even though the rule word had a novel surface form compared with the part word. The present result, obtained under conditions subjectively very similar to those of experiment 2, entails that the insertion of minor silent gaps radically alters behavior (23). Even though participants were neither overtly told nor aware that the familiarization stream was segmented, they spontaneously formulated an implicit grammatical-like generalization that corresponds to the structure of the represented items. Indeed, even though participants had never heard items like [pubeki] or [pugaki] they were persuaded that these were in the familiarization stream, whereas part words like [likita] or [Radube], which they did hear in the stream, were not. This seems to be due to the fact that the selected items are compatible with a generalization of the kind "If there is a [pu] now, then there will be a [ki] after an intervening X." Our interpretation meshes well with previous research on artificial language learning, showing that adults can acquire certain syntactic structures if the input includes explicit bracketing cues (24–26).

Because transitional probability computations do not account for the participants' choices in experiment 3 (27), we propose that different computations, possibly of an algebraic or rule-governed nature (20, 21,

28), are responsible for the observed behavior.

To test the second consequence of our prediction, we reasoned that if the crucial factor determining whether participants perform a statistical or a grammatical-like computation is the type of signal they are processing, then we should expect to observe two facts. Even substantially prolonging the familiarization used in experiment 1 should not give rise to generalizations, because this kind of signal does not generate that type of computation. On the contrary, even reducing dramatically the familiarization used in experiment 3 should leave the listener's ability to establish the underlying generalization intact, because that kind of signal triggers a computation which makes hypotheses about the structure in a nonstatistical fashion. To assess whether both phenomena can be observed, we ran experiments 4 and 5.

In experiment 4, we posited that if participants in experiment 2 had failed to extract the relevant generalization because of the lack of time to consolidate the statistical computations, then a significant increase of exposure to the familiarization stream ought to help them reach the generalization. However, if their failure in experiment 2 is related to the type of computation performed rather than to the amount of exposure, then this modification should have the opposite effect: It would consolidate memory traces for the items in the stream and would inhibit the projection of the generalization. To test which of the two hypotheses is correct, we familiarized a new group of participants ( $n = 14$ ) with the same stream used in experiments 1 and 2 but tripled their exposure to 30 min. After familiarization, participants were tested with the same rule words and part words used in experiments 2 and 3. This time participants selected the part words over the rule words significantly more often (Fig. 1D), suggesting not only that they failed to notice that rule words can be described with an appropriate generalization but also that their memory representations for the items that actually appeared in the stream had a tendency to improve. That is, a greater exposure to the stream appears to solidify memory traces rather than yield information about its structure. Thus, participants are sensitive to the statistical contingencies contained in the stream. However, these statistical computations do not give rise to grammatical-like generalizations, despite the big increase in exposure. This shows that making the underlying structure of a stream emerge is not just a matter of strengthening the representation of its items. The result also shows that the mere existence of a represented corpus may be necessary but not sufficient to trigger grammatical-like computations.

What, then, triggers computations that lead to the projection of structural regularities? If this process is not statistical but more like an unconscious projection of conjectures

from examples, then amount of exposure should not be the most critical factor. In experiment 5, we presented a new group of participants ( $n = 14$ ) with the same stream used in experiment 3, but we reduced exposure by a factor of five, thus allowing only 2 min of familiarization. Because the gaps contained in the stream may help participants (who remained unaware of them) to segment without computing transitional probabilities, a minimal familiarization with such a stream might induce generalizations almost immediately. To assess this, at the end of the 2 min of familiarization we tested participants with the same pairs of rule words and part words as in experiment 3. The results are presented in Fig. 1E. They indicate that two min of exposure suffice for grammatical-like generalizations to be computed, suggesting that generalizations arise very rapidly when subliminal signals to segmentation are available. Indeed, participants' performance is comparable to that obtained with exposure to a longer familiarization in experiment 3 (29).

It is important to note that experiments 4 and 5 are symmetrical. Though in experiment 4 we showed that no generalization arises even after a very long familiarization period with a continuous stream, in experiment 5 we showed that a very short exposure to a stream containing subtle cues to segmentation suffices to capture the underlying regularity. Thus, we propose that two different behaviors arise from entirely different computational processes that may be triggered by subtle differences in the signal: one is biased toward the discovery of its statistical patterns, and the other is oriented toward the discovery of its structure. Silent gaps in the stream appear to cause the listener to switch from one computational mode to the other; yet, we do not claim that only these specific cues can bring about this change. Rather, we predict that the role of silent gaps is to make the stream slightly more similar to natural language. Speech is by nature discontinuous. A system looking for structure in speech is naturally attuned to a signal modulated by rhythm and intonation (26, 30); our silent gaps may be the last resort that this system exploits to make a stream more "natural" (31).

The discovery that adults and infants can perform powerful statistical computations over a continuous corpus has stirred an intense debate. Some have suggested that, considering the mind's statistical dexterity, learning based on frequency and distributions may be rich enough to explain the emergence of linguistic abilities (32). Our results suggest that even though learners can compute powerful statistical relations, they do not appear to use this ability to extract simple structural generalizations. The ability to use statistical information for processing an unknown language stream seems to be confined to the individuation of segments. The discovery of



the grammatical system underlying linguistic competence appears to require a different type of computation.

References and Notes

1. A. Christophe, E. Dupoux, J. Bertoncini, J. Mehler, *J. Acoust. Soc. Am.* **95**, 1570 (1994).
2. R. N. Aslin, J. Z. Woodward, N. P. LaMendola, T. G. Bever, in *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition*, J. Morgan, K. Demuth, Eds, (Erlbaum, Hillsdale, NJ, 1996), pp. 117–134.
3. J. R. Hayes, H. H. Clark, in *Cognition and the Development of Language*, J. R. Hayes, Ed., (Wiley, New York, 1970), pp. 221–234.
4. J. R. Saffran, E. L. Newport, R. N. Aslin, *J. Mem. Lang.* **35**, 606 (1996).
5. J. R. Saffran, R. N. Aslin, E. L. Newport, *Science* **274**, 1926 (1996).
6. Saffran, Aslin, and Newport (5) use continuous streams of artificial consonant-vowel syllables. All syllables have the same duration, loudness, and pitch. Range of transitional probabilities used with adults is 0.25 to 1.00 within trisyllabic items and 0.05 to 0.60 between. Range of transitional probabilities in infant studies is 1.00 within and 0.33 between trisyllabic items.
7. J. R. Saffran, E. K. Johnson, R. N. Aslin, E. L. Newport, *Cognition* **70**, 27 (1999).
8. J. Fiser, R. N. Aslin, *Psychol. Sci.* **12**, 499 (2001).
9. ———, *J. Exp. Psychol.* **28**, 458 (2002).
10. N. Z. Kirkham, J. A. Slemmer, S. P. Johnson, *Cognition* **83**, B35 (2002).
11. G. F. Marcus, S. Vijayan, S. B. Rao, P. M. Vishton, *Science* **283**, 77 (1999).
12. Marcus *et al.* (11) used lists of trisyllabic items (all consonant-vowel syllables) with either an ABA structure (e.g., *GATIGA*, *LINALI*, etc.) or an ABB structure (e.g., *GATITI*, *LINANA*, etc.). Each syllable was separated by 250-ms gaps, and each trisyllabic item was presented after a 1-s pause. Infants prefer new items consistent with the familiarization structure to new items inconsistent with it.
13. E. L. Newport, R. N. Aslin, in *Proceedings of the 24th Annual Boston University Conference on Language Development*, S. C. Howell, S. A. Fish, T. Keith-Lucas, Eds., November 1999, Boston, MA (Cascadia Press, Somerville, MA, 2000), pp. 1–21.
14. In a first study using a very similar procedure as Saffran *et al.* (5), we demonstrated that when adjacent transitional probabilities are manipulated, French participants ( $n = 12$ ) behave like the English participants they tested. We showed that after 21 min of familiarization, 10 out of 12 participants significantly chose words over part words. The within-word transitional probability range was 0.25 to 1.0 and the between-word transitional probability range was 0.05 to 0.6.
15. The  $A_1XC_1$  family is (in International Phonetic Alphabet): [puliki], [puraki], [pufoki]; the  $A_2XC_2$  is: [beliga], [beraga], [befoga]; and the  $A_3XC_3$  is: [talidu], [taradu], [tafodu]. The part words are either formed by the last syllable of one word and the first two syllables of the following word, i.e., structure CAX ([kitaRa], [kitafo], [gapufo], and [dubera]), or by the last two syllables of one word and the first syllable of the following word, i.e., structure XCA ([likita], [lidube], [radube], [ragapu], [fogapu]). All words and part words are meaningless in French. Materials and methods, as well as a sample of the familiarization sound file, are available as supporting material on *Science Online*.
16. Because, in the stream, part words occur less frequently than words, the outcome of experiment 1 might be due to absolute frequency of syllable co-occurrence, rather than to distant transitional probabilities. To disentangle the two explanations, we ran a control experiment in which a 10-min stream was constructed in such a way that each family of words contained one high-frequency item and two items occurring half as frequently. We then tested the two low-frequency words of each family against part words having exactly the same frequency of occurrence in the stream. Adult participants ( $n = 14$ ) were

- exposed to the 10-min stream and were tested with 24 word-part word couples equalized in frequency. Even in this case, 70.5% of participants' choices were for the words ( $P < 0.03$ ). Thus it is not the absolute frequency of occurrence of the trisyllabic items that directs participants' choices, but the distant transitional probability relations among their syllables.
17. Although the stream was synthesized to be monotonous, continuous, and composed of words and part words of equal duration, it is possible that listeners do not perceive the signal as such. The perceptual centers of (monosyllabic and bisyllabic) words in a noncontinuous stream do not correspond to their physical centers [J. Morton, S. Marcus, C. Frankish, *Psychol. Rev.* **83**, 405 (1976)]. A sense of rhythm might arise from an uneven distribution of the perceptual centers of the syllables composing words and part words in the stream used in experiment 1. This might have induced participants to favor words over part words in the test phase irrespective of the transitional probabilities between the syllables composing them. To rule out this hypothesis, we ran a control experiment. We constructed a new familiarization stream transforming most of the part words of experiment 1 into the words of the control, and most of the words of experiment 1 into the part words of the control (by manipulating the nonadjacent syllable transitional probabilities in the familiarization stream). Thus, we chose part words and words so that their average perceptual center distributions, as well as other phonological properties, would become the mirror image of those in experiment 1. We reasoned that if the result of experiment 1 were due to a difference in average perceptual center distributions of words and part words, in this control participants should now prefer the part words over the words. The outcome of the control experiment showed that participants ( $n = 14$ ) still preferred words (as defined by their nonadjacent transitional probabilities) over part words [mean preference for words = 58%,  $t(13) = 2.7$ ,  $P < 0.02$ ]. Thus, perceptual grouping based on syllable perceptual centers or other phonological properties specific to the words and part words we selected in experiment 1 accounts neither for the result of experiment 1 nor for the results of the following experiments reported in this paper.
18. Thus, the rule words have a structure  $A_iX^*C_j$  and the part words have either structure  $C_kA_iX$  or  $XC_jA_i$ . The asterisk indicates that  $X^*$  appears in the stream, but not as a middle syllable of any family  $[A_iXC_j]$ . For instance, if the stream contained [puliki], [puraki], and [pufoki] as one family and [beliga] in another family, we tested whether participants preferred [pubeiki], which had not appeared in the stream, to the part word [likibe] that did appear in it. The syllable [be] appears in the stream, but not as a middle syllable in the family [puXki].
19. The difference between performance in experiments 1 and 2 is significant [ $t(26) = 3.625$ ,  $P < 0.001$ ].
20. N. Chomsky, *Syntactic Structures* (The Hague, Mouton, 1957).
21. S. Pinker, *Science* **253**, 530 (1991).
22. To make sure that the 25-ms gaps remained at subthreshold level, we ran a control experiment. Participants ( $n = 14$ ) were asked to judge two sequences of synthetic speech drawn from the familiarization used in experiments 2 and 3: one of them had silent gaps at the end of words while the other one was continuous. Both sequences lasted 1 min. Participants were told that they would have to listen to two sequences of an imaginary language and were informed that they would have to answer questions about these speech streams at the end of the presentation. The first question was whether they had heard any silent gaps. Thirteen out of 14 participants responded negatively. Next, all participants were informed that only one of the sequences had gaps and were asked to identify which of the two sequences contained them. Participants responded at chance. Therefore the 25-ms gaps are processed at some level, but participants do not become aware of their presence. A sample of the familiarization sound file is available as supporting material on *Science Online*.
23. The difference between performance in experiments 2 and 3 is significant [ $t(26) = -3.087$ ,  $P < 0.005$ ].
24. J. L. Morgan, E. L. Newport, *J. Verb. Learn. Behav.* **20**, 67 (1981).

25. J. L. Morgan, R. P. Meier, E. L. Newport, *Cognit. Psychol.* **19**, 498 (1987).
26. ———, *J. Mem. Lang.* **28**, 360 (1989).
27. Participants might have included the gaps as separate elements for computing transitional probabilities. As a result, they may have preferred rule words, not because they extracted the structure of the stream, but because they computed probabilities over syllables, pauses, and absence of pauses in the stream and the test items. Thus participants may have analyzed the rule words in the test as having the structure #A1X\*C1# and the part words as having the structure #XC2@A3# (where # indicates a pause and @ the absence of a pause). In this case, the transitional probabilities between adjacent elements would favor rule words over part words and no sensitivity to the structure of the rule words would be needed to prefer rule words. This hypothesis makes a prediction that has not been confirmed in a control experiment. Though in experiment 3 the test items do not contain pauses, in this control experiment we tested participants ( $n = 14$ ) with items including the pauses. Thus, participants compared rule words with structure #A1X\*C1# to part words with structure #XC2#A3#. In this case, the presence of the pause in the part words makes the transitional probability of the part word higher than that of the rule word. Therefore, if pauses counted as separate events in the computation, participants should favor the part words over the rule words. Nevertheless, contrary to this prediction, participants still preferred rule words to part words [61.5%,  $t(13) = 2.8$ ,  $P < 0.015$ ].
28. G. F. Marcus, *The Algebraic Mind: Integrating connectionism and cognitive science* (MIT Press, Cambridge, MA, 2001).
29. It could be thought that in experiment 5, as well as in experiment 3, participants do not extract a structural regularity of the kind "If  $A_i$  occurs then  $C_j$  will follow after an intervening  $X$ ", but rely on lower-level contingencies between parts of the stimuli. For example, participants could have been learning that A would be followed by C at a certain latency, extracting only a temporal association: given A, wait for the occurrence of C after a given temporal interval. At a more abstract level, it could be thought that participants have induced a sequential rule of the kind "If  $A_i$  occurs at position 1, then  $C_j$  will follow at position 3." To evaluate these possibilities, we ran a control experiment. In it, participants were familiarized with the 2-min stream of experiment 5, but in the test phase they listened to pairs of quadrisyllabic items obtained from the words and the part words of experiment 5 by substituting the middle syllable with two syllables that never occurred in the stream. Participants ( $n = 14$ ) still preferred the item conforming to the structural generalization [mean preference for rule words = 63%,  $t(13) = 5.012$ ,  $P = 0.0002$ ]. This shows that participants did not respond on the basis of strict temporal or sequential templates derived from the familiarization stream.
30. L. R. Gleitman, E. Wanner, in *Language Acquisition: The State of the Art*, L. R. Gleitman, E. Wanner, Eds. (Cambridge Univ. Press, Cambridge, 1982), pp. 3–48.
31. We cannot exclude that longer familiarization with continuous streams might also yield generalizations, but the underlying computations would be different from the rapid projection of generalization that we found in experiments 3 and 5.
32. M. S. Seidenberg, *Science* **275**, 1599 (1997).
33. This research has been supported by a grant from the Regione di Friuli-Venezia-Giulia. We are grateful to D. Aslin, A. Caramazza, S. Dehaene, J. Fodor, J. Morton, G. Marcus, L. Newport, B. Scholl, M. Shukla, and T. Shallice for discussions and suggestions.

**Supporting Online Material**  
[www.sciencemag.org/cgi/content/full/1072901/DC1](http://www.sciencemag.org/cgi/content/full/1072901/DC1)  
 Materials and Methods  
 Audio Files S1 and S2  
 15 April 2002; accepted 19 August 2002  
 Published online 29 August 2002;  
 10.1126/science.1072901  
 Include this information when citing this paper.