

## Statistical computations over a speech stream in a rodent

JUAN M. TORO and JOSEP B. TROBALÓN  
*Universitat de Barcelona, Barcelona, Spain*

Statistical learning is one of the key mechanisms available to human infants and adults when they face the problems of segmenting a speech stream (Saffran, Aslin, & Newport, 1996) and extracting long-distance regularities (Gómez, 2002; Peña, Bonatti, Nespor, & Mehler, 2002). In the present study, we explore statistical learning abilities in rats in the context of speech segmentation experiments. In a series of five experiments, we address whether rats can compute the necessary statistics to be able to segment synthesized speech streams and detect regularities associated with grammatical structures. Our results demonstrate that rats can segment the streams using the frequency of co-occurrence (not transitional probabilities, as human infants do) among items, showing that some basic statistical learning mechanism generalizes over nonprimate species. Nevertheless, rats did not differentiate among test items when the stream was organized over more complex regularities that involved nonadjacent elements and abstract grammar-like rules.

Recent work on statistical learning has suggested that this computational process may be partly responsible for two important aspects of language development: the detection of word boundaries (Aslin, Saffran, & Newport, 1998; Saffran, 2001b; Saffran, Aslin, & Newport, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997) and the extraction of grammar-like structures (Newport & Aslin, 2004; Peña, Bonatti, Nespor, & Mehler, 2002; Saffran, 2001a, 2002). In a series of studies, Saffran, Aslin, and Newport (1996) and Saffran, Newport, and Aslin (1996) showed that human adults and infants could segment a speech stream composed of trisyllabic nonsense words on the basis of statistical regularities alone. In their adult study, the authors presented the stream to the participants for 21 min. After this exposure, the participants could differentiate, at above-chance levels, words that composed the stream from foils. In their study of these abilities in infants, the authors familiarized 8-month-olds with 2 min of the stream. During the test, the infants showed differen-

tial responding to words and to foils, reflecting the extraction of sequential regularities. In a later work, Aslin et al. (1998) explored the complexity of the computations performed by the infants. Given the structure of the streams used in the authors' previous studies, the infants could be detecting either frequency of co-occurrence among syllables or their transitional probabilities. Whereas the former implies learning that the two elements tend to appear together in sequences, the latter provides information about how predictable one element is with respect to another<sup>1</sup> and better reflects the types of dependencies infants have to learn in natural languages (Aslin et al., 1998). In fact, results showed that the sort of information the 8-month-olds were detecting during the speech segmentation experiments consisted of transitional probabilities, and not just frequency of co-occurrence.

However, it has been demonstrated that the ability to compute the sort of statistics used in this task is not restricted to the detection of word boundaries. It also applies to nonlinguistic stimuli such as tones (Saffran, Johnson, Aslin, & Newport, 1999) and visual shapes (Fiser & Aslin, 2001, 2002a, 2002b; Kirkham, Slemmer, & Johnson, 2002). In addition, the cotton-top tamarin (a New World monkey) can segment a speech stream using statistical cues, just as human adults and infants do (Hauser, Newport, & Aslin, 2001). In that study, following the general procedure used with human infants, the authors familiarized tamarins to the stream for 20 min and tested them on the following day using their orienting responses toward a concealed speaker as the dependent measure. The tamarins were more likely to orient to foils than to words, a result that shows the learning of sequential regularities from exposure to the stream. With respect to the extraction of nonadjacent dependencies, it has been found that tamarins can also detect them in a speech stream, although their

---

This research was supported by Grant JSMF-20002079 from the James S. McDonnell Foundation, HFSP Grant 301870, Catalan Government Research Grant SGR00034, and Spanish MECED Fellowship AP2000-4164. The procedure was approved by the Comité de Ética en Experimentación Animal from the Universitat de Barcelona and complied with the guidelines of the Catalan and Spanish governments for the treatment of laboratory animals. We are indebted to Richard Aslin and Marcela Peña for kindly supplying some of the stimuli used in this study. We thank Ferran Pons and members of the GRNC Group for their suggestions during all stages of the study; Tere Rodrigo and others at the Laboratori de Psicologia Animal for their help with the animals used in the experiments; Gary Marcus for suggesting Experiment 3C; and Albert Costa, Núria Sebastian-Gallés, Salvador Soto-Faraco, Scott Sinnett, and two anonymous reviewers for their constructive and insightful comments on an earlier version of the manuscript. Correspondence concerning this article should be addressed to J. M. Toro, Cognitive Neuroscience Sector, SISSA/ISAS, 2-4 via Beirut, 34014 Trieste, Italy (e-mail: jmtoro@sissa.it).

selectivities are somewhat different than those of human adults (Newport, Hauser, Spaepen, & Aslin, 2004).

These results suggest that the computational abilities used to detect word boundaries and long-distance regularities may have evolved for purposes other than language and that they may be found in other nonhuman animals (Hauser, Chomsky, & Fitch, 2002). Furthermore, available data suggest that the kind of statistical learning that would suffice for the segmentation of speech would also work on the extraction of more complex long-distance regularities (see Seidenberg, MacDonald, & Saffran, 2002). That is, an organism capable of computing the statistical information necessary for the detection of word boundaries should also be capable of extracting grammar-like rules. In this article, we report a test of this hypothesis through exploration of a rodent's statistical computations in a speech stream. If the computations required for the development of grammar-like structures depend not on specialized structures but on computational abilities used in speech segmentation, an animal capable of statistical learning should display both.

We have hints that rats are able to perform some statistical learning, since they can compute conditional statistics to calculate the predictive value of one stimulus with respect to another (Rescorla, 1968; Rescorla & Wagner, 1972). Also, rats provide an interesting model because, in addition to being a nonprimate species, they do not produce elaborate sequential calls as a means of interspecific communication,<sup>2</sup> as cotton-top tamarins do (Ghazanfar & Hauser, 2001; Weiss, Garibaldi, & Hauser, 2001; Weiss & Hauser, 2002). Therefore, any finding about statistical learning in rats can be detached from an explanation in terms of their own communication requirements.

In order to establish whether or not response sensibility to adjacent sequential regularities is separable in an animal from sensibility to nonadjacent grammar-like regularities, we ran a series of experiments using synthesized speech streams identical to those used in previous studies on statistical learning with human adults, infants, and tamarin monkeys (e.g., Aslin et al., 1998; Hauser et al., 2001; Saffran, Aslin, & Newport, 1996).

## EXPERIMENT 1 Segmentation of a Speech Stream

Although there are several useful regularities in speech, including phonotactic or prosodic regularities (see, e.g., Cairns, Shillcock, Chater, & Levy, 1997; Cutler, Mehler, Norris, & Segui, 1986; Johnson & Jusczyk, 2001; Thissen & Saffran, 2003), the ability to track the statistical distribution among different elements (such as syllables or words) is one of the key cues that humans can use when confronting the task of segmenting speech (Saffran, 2003). The goal of Experiment 1 was to study whether rats can segment a speech stream using these regularities. In order to test this, we synthesized two artificial languages with the same words organized in the same manner as those used by Hauser et al. (2001). In an attempt to keep the procedure as similar as possible to the previous ones, in

the first session we exposed the rats to the speech stream. In the second session, we measured responses to words present in the stream and responses to different foils. Instead of measuring orienting responses, we measured leverpressing by the rats.

### Method

**Subjects.** The subjects were 32 male Long-Evans rats. Their mean age was 3 months. The rats were caged in pairs and kept at 80% of their free-feeding weights during the experiment. The rats had no previous experience in experiments involving leverpressing or acoustic stimuli.

**Stimuli.** All of the stimuli were synthesized with MBROLA software (Dutoit, Pagel, Pierret, Bataille, & van der Vrecken, 1996), using a Spanish male diphone database (es2).<sup>3</sup> They were digitized at a sampling rate of 16 kHz and played through a Pioneer Stereo Amplifier A-445, using a Mediavision Proaudio Spectrum 16 sound card and two E.V. (S-40) speakers. The loudness levels for the stimulus presentation were set at 68 dB SPL using an Onsoku SM-6 sound level meter.

We used two artificial languages made out of four trisyllabic "words." The words used were the same as those in the Saffran, Aslin, and Newport (1996) and Hauser et al. (2001) experiments. Language A consisted of the words *tupiro*, *golabu*, *bidaku*, and *padoti*, whereas Language B consisted of *tudaro*, *pigola*, *bikuti*, and *budopa*. Words were concatenated in a semirandom order, with the restriction that there were no immediate repetitions of words, following the structure of the above-mentioned studies. The transitional probability between syllables forming a word was always 1.0, meaning that these syllable pairs always appeared together. Each word appeared around 431 times during the 20 min of exposure. Twelve test items were also created, including two words, two part-words, and two nonwords for each language. Part-words were formed by joining the last syllable of one word and the first two syllables of another. Transitional probabilities between syllables forming a part-word were .33 between the first and second syllables and 1.0 between the second and third syllables. Nonwords had not appeared in the stream, but they were formed by syllables that had appeared, though not in the same order; thus, the transitional probability among their component syllables was 0. For Language A, the test items were the words *tupiro* and *golabu*, the part-words were *tibida* and *kupado*, and the nonwords were *dapiku* and *tilado*. For Language B, the test items were the words *tudaro* and *pigola*, the part-words were *pabiku* and *tibudo*, and the nonwords were *tigobu* and *kudabi*. Each test item was presented for 695 msec.

**Procedure.** The experiments were run using Leticia Skinner boxes connected to a PC. The computer was used to control stimuli, to deliver food during sessions, and to register leverpress responses. The rats were taken from their home cages to the testing room, where they were placed individually in the Skinner boxes. They were exposed to magazine training until they reached an average leverpressing ratio of 1,000 presses over 20 min for at least 3 consecutive days on a variable ratio  $\pm 10$  (VR-10) program.

When this ratio was reached for all rats, the experiment began. Half of the rats ( $n = 16$ ) were assigned to Language A and the other half to Language B. The experiment consisted of two sessions. In Session 1, all the rats were exposed to the corresponding language for 20 min. The rats were placed in the Skinner boxes and received food on the VR-10 program, just as on any training day, except that while they were in the box the speech stream was presented. Therefore, feeding was not contingent on the stream. During this exposure session, food delivery was not directly paired with any segment of the stream, since there were no pauses between words. Thus, the first session was not a discrimination training session but simply involved exposure to the speech stream while, independently, the rats received food as a reward for their leverpressing responses. When the stream was over, the rats were returned to their home cages.

Session 2 was run the following day. The rats were placed in the Skinner boxes, and test items were presented in a balanced manner, with the restriction that no more than two items of the same type followed each other. There were 1-min intervals between items, during which leverpressing responses were registered. The food dispensers were disconnected, so the rats did not obtain food during the whole test session. The rationale for this paradigm is that if the rats segmented the stream, words should elicit a stronger leverpressing response than other test items should because of their higher frequency of appearance during the first session, which was run while the rats were receiving food.

## Results

For all the experiments, in order to account for individual differences in response ratio, leverpressing frequency for each type of test item during the entire test session was converted to  $Z$  values and averaged across subjects. An analysis of variance was performed with language (A vs. B) as a between-subjects variable and test (word vs. part-word vs. nonword) as a within-subjects variable. There were no significant differences between languages [ $F(1,30) = 0.38, p = .542$ ], but the variable test was significant [ $F(1,30) = 6.48, p < .05$ ; see Figure 1]. Since no differences between languages were found, data were collapsed across languages; the interaction between languages and test was not significant [ $F(1,30) = 0.621, p = .437$ ]. Further analysis revealed that the differences in response ratios to words (.491) versus nonwords (.315) were significant [ $t(31) = -4.94, p < .001$ ], as were the differences between words and part-words [.374;  $t(31) = -2.56, p < .05$ ]. That is, results reflected differential responding to test items in the direction of stronger leverpressing responses to items that frequently appeared in the stream (words) over items that appeared less often (part-words) or not at all (nonwords). Differences in the response ratios to nonwords and part-words were not significant [ $t(31) = -1.89, p = .068$ ]. These results strongly suggest that rats are able to track on line the distributional properties of elements composing the stream, resulting in the differentiation of test items and representing a first step in statistical learning in rodents. Note, however, that

the present results do not indicate whether rats use a computation for the segmentation of the streams similar to that used by human infants. As was said in the introduction, infants can compute transitional probabilities within a speech stream, which helps them to extract some regularities present in natural languages. It might be the case that rats also perform this computation during the segmentation of a speech stream, or that they rely on a different strategy to solve the task. To further explore this issue, we conducted Experiment 2.

## EXPERIMENT 2 Transitional Probabilities Versus Frequency of Co-Occurrence

Regarding the specific computations necessary to solve a segmentation task using streams such as those of the former experiment, Aslin et al. (1998) demonstrated that the strategy used by human infants was to calculate the transitional probabilities among items. Because of the structure of the speech streams used in Experiment 1, words appeared twice as often as did part-words, whereas nonwords did not appear at all. Thus, words, part-words, and nonwords were differentiated not only by transitional probabilities but also by the overall frequency of co-occurrence of their syllables (Aslin et al., 1998). That is, the rats could be attending to either of these two cues (transitional probabilities or frequency of co-occurrence) to segment the streams.

In Experiment 2, we wanted to identify which computation rats were performing during the segmentation task. We used streams with the same structure as in the Aslin et al. (1998) study, in which two of the component words were matched for transitional probabilities whereas the other two were matched for frequency of co-occurrence. If rats compute transitional probabilities in the segmentation tasks, differences should arise in a test with frequency-matched items. If rats rely on frequency of co-occurrence, there should not be differences in this test. However, there should be differences in a test with items matched on tran-

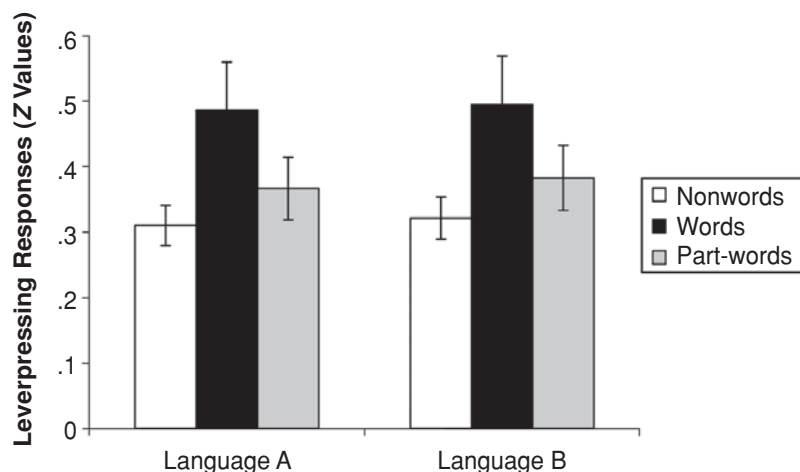


Figure 1. Experiment 1: mean  $Z$  values for responses in the test session for nonwords, words, and part-words in Language A and Language B.

sitional probabilities but having different overall frequencies of co-occurrence.

### Method

**Subjects.** We used two new groups of 32 male Long-Evans rats. Their mean age and housing conditions were the same as those of Experiment 1. The rats had no previous experience in experiments involving leverpressing or acoustic stimuli.

**Stimuli.** The structure of the streams, the four words that composed them, and the test items were the same as those in Aslin et al.'s (1998) experiment. In these new streams, two of the words appeared twice as often as the other two. Part-words formed by the union of syllables from the two high-frequency words had the same absolute frequency in the stream as did the two low-frequency words. Although part-words and low-frequency words were matched for overall frequency, their transitional probabilities were different. For each pair of syllables composing the words, transitional probability was 1.0, whereas for pairs of syllables composing the part-words it was .5 for one syllable and 1.0 for the other. Language A was composed of the words *pabiku*, *tibudo*, *golatu*, and *daropi*, with *golatu* and *daropi* as high-frequency words and *pabiku* and *tibudo* as low-frequency words. Part-words used during the test were *tudaro* and *pigola*. Language B had the words *tudaro*, *pigola*, *bikuti*, and *budopa*, with *bikuti* and *budopa* as high-frequency words and *tudaro* and *pigola* as low-frequency words. Part-words used during the test were *pabiku* and *tibudo*.

**Procedure.** The procedure was identical to that of Experiment 1. Thirty-two rats received the low-frequency words versus part-words test, whereas the remaining 32 rats received the low-frequency versus high-frequency words test.

### Results

As in the previous experiment, no significant interaction between language and test was found for the group of rats presented with the low-frequency words versus part-words test [ $F(1,30) = 0.001, p = .98$ ] or for the group presented with the low-frequency versus high-frequency words test [ $F(1,30) = 0.647, p = .427$ ], so the data were collapsed across languages. Unlike in Experiment 1, the rats responded similarly to both the low-frequency word (.426) and part-word (.408) test items [ $t(31) = -0.36, p = .716$ ]—that is, when items differed in their transitional probabilities, the rats did not detect the difference and responded similarly to them. Nevertheless, response ratios to high-frequency words (.514) and those to low-frequency words (.406) were significantly different [ $t(31) = 2.73, p < .05$ ], so the rats responded in a different manner to items that could be distinguished by their frequencies of co-occurrence (see Figure 2). Together, these results indicate that the rats did not use the more complex type of computation (transitional probabilities) but the simpler one (frequency of co-occurrence) during the speech segmentation task.

Experiment 2 confirmed that, as was demonstrated in Experiment 1, rats can segment a speech stream. However, they do so by using the overall frequency with which syllables occur together, a different kind of computation than that used by human infants (who rely on transitional probabilities). Although it is still unknown what type of computation tamarin monkeys perform over speech streams, these results suggest a difference in the sensitivity to sequential regularities that some mammal species have. It is interesting now to explore whether or not these differ-

ences expand from the type of computations used to the complexity of the regularities that can be extracted from the streams. Even though rats do not show the same type of statistical learning in word segmentation as do humans, perhaps they show the same type of learning in grammatical tasks. Peña et al. (2002) argued that the statistical computations required for speech segmentation are qualitatively different from the computations involved in the extraction of rules for developing grammatical structures. Furthermore, Fitch and Hauser (2004) and Newport et al. (2004) have shown different constraints across species for more elaborate computations over artificial languages. If such computations are not just a continuation of those needed for speech segmentation, it is possible that further constraints should be observed in rats.

## EXPERIMENT 3

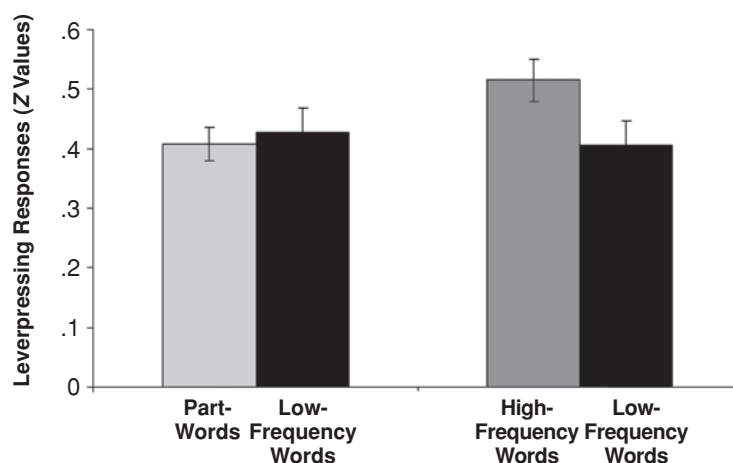
### Extraction of Grammar-Like Structures

#### Experiment 3A

#### Nonadjacent Regularities

Natural languages are structured in hierarchies, exhibiting dependencies and relationships between adjacent and nonadjacent elements at multiple levels (Chomsky, 1957; Jackendoff, 2002). These dependencies are especially salient in grammatical rules that are so important in some languages, such as noun-verb agreement. Peña et al. (2002) tested human adults' abilities to segment and extract the structure-forming rule of a speech stream composed of items with nonadjacent regularities. The authors used a synthesized speech stream composed of nine trisyllabic words in which the first and third syllables always appeared together with a variable syllable between them. Results showed that after 10 min of exposure, adults were able to segment the stream.<sup>4</sup> But even when exposure was increased to 30 min, the participants were unable to detect the structure of the segments. However, if pauses lasting 25 msec were introduced between words, the participants finally extracted the structure. The interpretation offered by Peña et al. is that the computational processes required for detecting regularities between items are qualitatively different from those used for extracting the underlying structure of the stream; the first is sufficient for the segmentation of speech, whereas the second is necessary for developing grammatical structures.

In Experiment 3A, we studied whether rats could track nonadjacent regularities of the type previously described in humans. With the use of Peña et al.'s (2002) stimuli, two tests could be performed. The first is a segmentation test, in which response ratios to words and response ratios to part-words are compared. If subjects can segment the stream, they must be able to differentiate between these two classes of items. The second test is a rule-finding test in which response ratios to rule words (i.e., items that have not appeared in the stream but comply with the rule underlying the words) and part-words are compared. If subjects not only segment the stream but also can extract the rule that forms its words, they must be able to differentiate between rule words and part-words.



**Figure 2. Experiment 2: Mean Z values for responses in the test session for part-words and low-frequency words, and for high-frequency words and low-frequency words collapsed for Languages A and B.**

## Method

**Subjects.** The subjects were 32 male Long-Evans rats. Their age and housing conditions were identical to those of the previous experiments. They had no previous experience in experiments involving leverpressing or acoustic stimuli.

**Stimuli.** The words used in the Peña et al. (2002) study were used here and were organized in the same manner. We used two 20-min streams: one with no pauses between words and another with 25-msec pauses between words. The streams were composed of nine words, three of which had the form [puxki], 3 [bexga], and 3 [taxdu], with syllables [li], [ra], and [fo] serving as  $x$ . The resulting words were *puliki*, *puraki*, *pufoki*, *beliga*, *beraga*, *befoga*, *talidu*, *taradu*, and *tafodu*. Test items consisted of the nine words, nine part-words, and nine rule words. Part-words resulted from the union of the last syllable of one word with the first two syllables of another, or the union of the last two syllables of one word with the first syllable of another (*likita*, *radube*, *fogapu*, *lidube*, *gapufo*, *kitara*, *dubera*, *ragapu*, and *ligabe*). Rule words were words that complied with the underlying rule but had not appeared in the stream (*pubeki*, *pugaki*, *putaki*, *bepuga*, *bekiga*, *beduga*, *tabedu*, *takidu*, and *tagadu*). Each test item was presented for 695 msec. The stimuli were synthesized and presented in the same manner as in Experiment 1.

**Procedure.** The procedure was identical to that of the previous two experiments. Sixteen rats were presented with the stream with no pauses, whereas the other 16 were presented with the stream containing 25-msec pauses. Half of the rats in each group were presented with a test of words and part-words, and the other half were presented with a test of rule words and part-words.

## Results

The rats presented with the speech stream without pauses showed no significant differences in response ratios to words (.381) and part-words [.429;  $t(7) = -1.47$ ,  $p = .184$ ] or to rule words (.445) and part-words [.442;  $t(7) = -0.54$ ,  $p = .605$ ], which demonstrates that they did not segment the words out of the stream and differentiate them from the part-words. This finding contrasts with that of Experiment 1 and can be explained only by the increased complexity of the stream, which included long-distance regularities. The rats also did not extract the rule underlying the formation of words and discriminate rule words from part-words. In a different manner, the rats presented

with the speech stream that had 25-msec pauses between words showed differences between words (.477) and part-words [.321;  $t(7) = 2.94$ ,  $p < .05$ ], whereas the difference between rule words (.404) and part-words (.466) was not significant [ $t(7) = 1.25$ ,  $p = .250$ ]. This pattern of results suggests that the rats could recognize the words that appeared in the stream when gaps were introduced between words, but they could not extract the underlying rule.

Contrary to what has been found with humans (Peña et al., 2002), the introduction of 25-msec pauses between words does not help rats to find the grammatical rule. When rats are exposed to a stream containing these pauses, they are able to distinguish the words from the part-words. However, it is very likely that this result is due to the fact that there were acoustic cues marking the beginnings and the ends of words which may have been detectable by the rats. Psychophysical data show that mammals, including humans and rats, have a minimum gap-detection threshold that ranges from 2.5 to 3.5 msec (Fay, 1992). If this cue is used, differences between words and part-words would not reflect a segmentation process with the use of statistical cues, but only the ability of rats to recognize sound items that in fact they had heard separately in the stream.

## Experiment 3B Increased Variability

Gómez (2002) showed that the performance of human adults and 18-month-old infants learning nonadjacent dependencies improved when the variability of the middle element increased. Gómez created four sets of words in which the value of the middle syllable was drawn from a group of 2, 6, 12, or 24 items. After an exposure phase, improvement in the identification of the items that followed the rule was observed only for subjects that heard the sets in which the syllable in the middle was most variable—that is, those in which the middle syllable took 24 values. Gómez concluded that high variability in the second syllable caused subjects to focus on the invariant structure of the

stimulus—that is, the constant first and third syllables—and helped them detect the regularity. To test whether or not rats could also benefit from higher variability in extracting the nonadjacent dependencies, we modified the stream we used in the previous experiment and ran Experiment 3B.

### Method

**Subjects.** The subjects were 32 male Long-Evans rats. Their mean age and housing conditions were the same as those of the previous experiments. They had no previous experience in experiments involving leverpressing or acoustic stimuli.

**Stimuli.** A new speech stream was created that was identical in structure to the one used in Experiment 3A. Words composing the stream had a trisyllabic structure, with the first and third syllables having the same values as in the previous experiment—that is, [puxki], [bexga], and [taxdu]. Nevertheless, *x* could take any of 24 values (le, di, bo, ma, to, ne, ka, fi, ro, do, ke, ba, gu, mo, ti, fe, pi, lo, pa, mi, nu, li, ra, fo), for a total of 72 words. The resulting stream, containing 25-msec pauses between words, was organized and synthesized in the same manner as the previous ones. Test items, including words, part-words, and rule words, were the same as those in Experiment 3A. Stimulus synthesis and presentation were identical to those used in the previous experiments.

**Procedure.** The procedure was the same as in Experiment 3A.

### Results

After exposure to the stream with increased variability in the middle syllable, the rats showed no significant differences in response ratios to words (.331) and part-words [.310;  $t(15) = 0.464, p = .649$ ]. There was a significant difference between rule words (.308) and part-words [.442;  $t(15) = -2.295, p < .05$ ], but in the opposite direction of what was expected; the rats preferred part-words over rule words. In contrast to what was found in Experiment 3A, the rats could not differentiate between words and part-words. This may be due to the fact that increased variability made individual words appear less often than they did in the previous experiment (around 190 times for 20 min in Experiment 3A, and around only 24 times for 20 min in Experiment 3B), so they were not readily recognized. As for the difference between part-words and rule words, part-words actually appeared in the stream whereas rule words did not; this result suggests that the rats did not extract the structure of the words and were simply responding to sounds they heard during the exposure session. In any case, it is surprising that this pattern was not observed in Experiment 3A, in which the rats could not differentiate between part-words and rule words even with the 25-msec pauses. The difference is that in Experiment 3A words were segmented from the stream by the rats, and this could have made the distinction between rule words and part-words irrelevant.

Experiment 3B shows that the rats did not benefit from the higher variability that made almost irrelevant the middle *x* position in the structure of the trisyllabic words forming the stream. That is, regarding the extraction of nonadjacent regularities, we could not find any result with rats that would parallel those found with human adults and infants, who improved performance when up to 24 syllables could take the middle position in such a structure (Gómez, 2002).

## Experiment 3C Abstract Rule Learning

Setting apart the problem of segmentation and focusing on the extraction of regularities from speech, researchers have studied the problem of rule learning. Marcus, Vijayan, Bandi Rao, and Vishton (1999) habituated 7-month-old infants to words conforming to either an AAB or an ABB rule, in which A and B could each take four different values but the structure was held constant. Then, the infants were tested on AAB and ABB words formed by new syllables. The authors found that by this age, infants can learn the constant rule and distinguish new tokens that conform to it from others that do not. In a replication of this study, Hauser, Weiss, and Marcus (2002) found that cotton-top tamarins also could extract this rule, suggesting generalities across species regarding the processes involved in these abstract relations.<sup>5</sup> The study of the abilities of rodents to extract grammar-like rules could be complete only if we explore their ability to find abstract rules on a speech stream, independent of the segmentation problem. We conducted Experiment 3C using the same stimuli as Marcus et al. (1999) did in their third experiment, which controlled for phonetic cues and reduplication on the rule.

### Method

**Subjects.** The subjects were 32 male Long-Evans rats. Their age, housing, and experience conditions were the same as those of the previous experiments.

**Stimuli.** Two sets of words were created. They were identical to the ones used in the third experiment of Marcus et al. (1999). One set followed the AAB structure and included the 16 words *leleje*, *lelewe*, *leledi*, *leleli*, *wiwije*, *wiwiwe*, *wiwidi*, *wiwili*, *jijije*, *jijiwe*, *jijidi*, *jijili*, *dedeje*, *dedewe*, *dededi*, and *dedeli*. The second set followed the ABB structure (A and B each having the same syllables as the previous structure) and included the 16 words *lejeje*, *lewewe*, *ledidi*, *lelili*, *wijeje*, *wiwewe*, *wididi*, *wilili*, *jijeje*, *jiwewe*, *jididi*, *jilili*, *dejeje*, *dewewe*, *dedidi*, and *delili*. Four test items were also created, two with the AAB structure (*babapo* and *kogoga*) and two with the ABB structure (*bapopo* and *kogaga*). Words were synthesized and presented in the same way as all other stimuli in this study.

**Procedure.** In the first session, half of the animals were presented with the AAB words and the other half with the ABB words for 20 min, with 500 msec between words. As in previous experiments, the rats received food for leverpressing responses on a VR-10 schedule that was not contingent on the words. In the second session, all the rats were presented with three random repetitions of the four test items. There was a 1-min interval between presentations, during which leverpressing responses were registered. Food dispensers were disconnected throughout the entire test session.

### Results

For the group of rats exposed to the words with the AAB structure, the mean response ratios were .469 for AAB items and .514 for ABB items [ $t(15) = -0.823, p = .423$ ]. For the group of rats exposed to the words with the ABB structure, the values were .431 for AAB items and .542 for ABB items [ $t(15) = -1.134, p = .275$ ]. There was no indication that the rats could extract the rule that formed the words in order to differentiate between test items. Through the use of a different methodology, it has been found that human newborns (Marcus et al., 1999)

and cotton-top tamarins (Hauser, Weiss, & Marcus, 2002) are able to identify as equivalent words conforming to an abstract structure. The rats did not succeed in this task even though we used a procedure that yielded positive results for speech segmentation (as in Experiments 1 and 2) and tried to keep it as similar as possible to those used with other species. So, even though it is difficult to support conclusions based on null effects, the data across the last three experiments show a general difficulty for rats in differentiating between test items after being exposed to artificial languages organized around long-distance regularities and abstract rules. A difficulty that was not appreciated when the artificial languages were based on adjacent regularities and their test items could be differentiated on the basis of relative frequency.

### GENERAL DISCUSSION

The present study shows that the ability to compute the necessary statistics to find word boundaries in a speech stream does not guarantee the ability to extract long-distance regularities of the type that have been associated with grammar-like structures. Our results demonstrate that rats are able to segment words out of a speech stream (Experiment 1) by relying on the frequency of co-occurrence between syllables and not on transitional probabilities (Experiment 2). Nevertheless, the rats did not differentiate between test items when they were exposed to a speech stream structured around the distributional regularities of nonadjacent items (Experiment 3A), even when the task was made easier by increased variability of the middle syllable (Experiment 3B) or when they had to extract an abstract rule for word formation (Experiment 3C). Two aspects of our results are linked to this conclusion.

First, unlike the responses of human infants, with the present methodology rats' responses did not vary as a function of transitional probabilities over a speech stream, but only as a function of frequency of co-occurrence—a computation that involves fewer parameters than the former one. This does not mean that rats are not able to compute transitional probabilities at all. It has been demonstrated that rats, as well as dogs and pigeons, can readily detect contingent relations among events, which implies the computation of several conditional probabilities, including transitional ones (Rescorla & Wagner, 1972). For example, animals succeeded in a task of learning the contingencies between paired tones and electric shocks in which transitional probabilities between stimuli were manipulated (Rescorla, 1968). The difference is that this task does not involve the on-line computation of several statistics among elaborate acoustic stimuli, as do the tasks in the present series of experiments. But our results certainly begin to draw clear differences in the analysis of complex sound streams across species—differences that may be responsible for the failure of rats in the detection of long-distance regularities.

Second, statistical learning in the rat may be constrained by the structure of the stimuli over which these computa-

tions are performed. This is the case of nonadjacent dependencies. Cross-species differences in the detection of these dependencies are reflected by the distinctive pattern of results obtained with human adults and cotton-top tamarins (Newport & Aslin, 2004; Newport et al., 2004) and may even reflect different processing levels (e.g., statistical vs. grammatical; Peña et al., 2002; but see Seidenberg et al., 2002, and a reply by Marcus & Berent, 2003). The computation of statistics among adjacent items would play a special role in the segmentation of speech (Peña et al., 2002) and the finding of regularities with the use of lower order transitions between items (Gómez, 2002), whereas the computation of statistics among nonadjacent items would be necessary for finding grammar-like structures and higher order conditionals on which natural languages are based. Accordingly, our results show that rats have the ability to track distributional regularities among different elements in a sound stream, a result that has also been found in human infants (Saffran, Aslin, & Newport, 1996) and cotton-top tamarins (Hauser et al., 2001). But, unlike human adults (Newport & Aslin, 2004; Peña et al., 2002) and tamarins (Newport et al., 2004), rats do not display discriminative responses after being exposed to streams based on long-distance regularities or grammar-like rules, a task that implies focusing on abstract sequences that apply to all the words forming the stream. Note, however, that the rats also did not use transitional probabilities in the segmentation task. Thus, with the present results, we cannot definitively establish whether the rats failed to detect more complex regularities because they did not track transitional probabilities to begin with, or because a different process, specific to the detection of such probabilities, was not triggered.

In contrast to human adults, rats do not profit from the 25-msec pauses to extract the underlying structure of the streams. Without any acoustic marker, rats are incapable of segmenting these kinds of words. Even with the help of pauses, they cannot generalize them in the frame of abstract rules. This shows a limit in the regularities they can extract from an auditory stream, which limit may be linked to other abilities in the domain of interspecific communication, such as sequential vocalizations. Cotton-top tamarins have shown remarkable abilities in detecting regularities by using statistical cues (Hauser et al., 2001; Newport et al., 2004), extracting algebraic-like rules (Hauser, Weiss, & Marcus, 2002), and mastering finite-state grammars (Fitch & Hauser, 2004). However, tamarin monkeys produce natural calls comprised of combinations of single units in fixed sequences (Ghazanfar & Hauser, 2001; Weiss et al., 2001; Weiss & Hauser, 2002). Thus, there is a possibility that the production and perception of these sequences of sounds predicts the ability to detect long-distance regularities in an acoustic stream.

It is important, though, to acknowledge the limitations of the methodology used in the present study. Most experiments on speech segmentation and the extraction of rules from speech streams familiarize either infants or tamarins to a set of sounds, and then measure looking times

or head-turning responses to the test items. In contrast, we had to rely on an indirect measure (leverpressing) of the rats' sensitivities. Even though this procedure allowed us to observe significant differences in some conditions, it is also true that a direct comparison of the present results with those from other species is made more difficult. Likewise, the nature of the task is likely to be responsible for some disparities in the results reported here as well as for those that suggest the learning of statistical regularities in pigeons (Keen & Machado, 1999) and rats (Fountain, Rowan, & Benson, 1999). But central to this work was the assessment of rats' abilities to detect regularities using the same kind of stimuli used with humans and tamarins. More experimental work is needed to establish whether the results reported here generalize to other kinds of stimuli and settings.

Finally, the application of widely shared statistical learning mechanisms to the problem of communication has an important implication. It may begin to highlight the basis of the peculiarities of the human communicative system and how it contrasts with that developed by other species. As Hauser, Chomsky, and Fitch (2002) suggested, an animal's constraints on the computation and discovery of statistical regularities among sound items may be one of its limits for developing a combinatorial communication system. These differences in the processes and constraints involving statistical learning over a speech stream between rats and other species (such as cotton-top tamarins and humans) may be an example of the kinds of differences in mechanisms involved in the perception and development of human language that make it seem so special.

#### REFERENCES

- ASLIN, R. N., SAFFRAN, J. R., & NEWPORT, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321-324.
- BURGDORF, J., KNUITSON, B., PANKSEPP, J., & SHIPPENBERG, T. (2001). Evaluation of rat ultrasonic vocalizations as predictors of the conditioned aversive effects of drugs. *Psychopharmacology*, *155*, 35-42.
- CAIRNS, P., SHILLOCK, R., CHATER, N., & LEVY, J. (1997). Bootstrapping word boundaries: A bottom-up corpus-based approach to speech segmentation. *Cognitive Psychology*, *33*, 111-153.
- CHOMSKY, N. (1957). *Syntactic structures*. The Hague: Mouton.
- CUTLER, A., MEHLER, J., NORRIS, D., & SEGUI, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory & Language*, *25*, 385-400.
- DUTOIT, T., PAGEL, V., PIERRET, N., BATAILLE, F., & VAN DER VRECKEN, O. (1996). The MBROLA project: Toward a set of high-quality speech synthesizers free of use for noncommercial purposes. *Proceedings of the Fourth International Conference of Spoken Language Processing* (pp. 1393-1396). Philadelphia: ICSLP.
- FAY, R. (1992). Structure and function in sound discrimination among vertebrates. In D. Webster, R. Fay, & A. Popper (Eds.), *The evolutionary biology of hearing*. New York: Springer-Verlag.
- FISER, J., & ASLIN, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Sciences*, *12*, 499-504.
- FISER, J., & ASLIN, R. N. (2002a). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 458-467.
- FISER, J., & ASLIN, R. N. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Science*, *99*, 15822-15826.
- FITCH, T., & HAUSER, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, *303*, 377-380.
- FOUNTAIN, S., ROWAN, J., & BENSON, D., JR. (1999). Rule learning in rats: Serial tracking in interleaved patterns. *Animal Cognition*, *2*, 41-54.
- GHAZANFAR, A., & HAUSER, M. D. (2001). The auditory behaviour of primates: A neuroethological perspective. *Current Opinion in Neurobiology*, *11*, 712-720.
- GÓMEZ, R. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431-436.
- HAUSER, M. D., CHOMSKY, N., & FITCH, T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569-1579.
- HAUSER, M. D., NEWPORT, E. L., & ASLIN, R. N. (2001). Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*, B53-B64.
- HAUSER, M. D., WEISS, D., & MARCUS, G. (2002). Rule learning by cotton-top tamarins. *Cognition*, *86*, B15-B22.
- JACKENDOFF, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- JOHNSON, E., & JUSCZYK, P. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory & Language*, *44*, 548-567.
- KEEN, R., & MACHADO, A. (1999). How pigeons discriminate the relative frequency of events. *Journal of the Experimental Analysis of Behavior*, *72*, 151-175.
- KIRKHAM, N. Z., SLEMMER, J. A., & JOHNSON, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*, B35-B42.
- MARCUS, G. F. (1999). Do infants learn grammar with algebra or statistics? Response to Seidenberg, Elman, Negishi and Eimas. *Science*, *284*, 436-437.
- MARCUS, G. F. (2000). *Pabiku and Ga Ti Ga*: Two mechanisms infants use to learn about the world. *Current Directions in Psychological Science*, *9*, 145-147.
- MARCUS, G. F., & BERENT, I. (2003). Are there limits to statistical learning? *Science*, *300*, 53-54.
- MARCUS, G. F., VIJAYAN, S., BANDI RAO, S., & VISHTON, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77-80.
- NEWPORT, E. L., & ASLIN, R. N. (2004). Learning at a distance: I. Statistical learning of nonadjacent dependencies. *Cognitive Psychology*, *48*, 127-162.
- NEWPORT, E. L., HAUSER, M. D., SPAEPEN, G., & ASLIN, R. N. (2004). Learning at a distance: II. Statistical learning of nonadjacent dependencies in a nonhuman primate. *Cognitive Psychology*, *49*, 85-117.
- PEÑA, M., BONATTI, L., NESPOR, M., & MEHLER, J. (2002). Signal-driven computations in speech processing. *Science*, *298*, 604-607.
- RESCORLA, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative & Physiological Psychology*, *66*, 1-5.
- RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York: Appleton-Century-Crofts.
- SAFFRAN, J. R. (2001a). The use of predictive dependencies in language learning. *Journal of Memory & Language*, *44*, 493-515.
- SAFFRAN, J. R. (2001b). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, *81*, 149-169.
- SAFFRAN, J. R. (2002). Constraints on statistical learning. *Journal of Memory & Language*, *47*, 172-196.
- SAFFRAN, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, *12*, 110-114.
- SAFFRAN, J. R., ASLIN, R. N., & NEWPORT, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.
- SAFFRAN, J. R., JOHNSON, E., ASLIN, R. N., & NEWPORT, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27-52.
- SAFFRAN, J. R., NEWPORT, E. L., & ASLIN, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory & Language*, *35*, 606-621.
- SAFFRAN, J. R., NEWPORT, E. L., ASLIN, R. N., TUNICK, R., & BAR-



- RUECO, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, **8**, 101-105.
- SEIDENBERG, M. [S.], ELMAN, J., NEGISHI, M., & EIMAS, P. (1999). Do infants learn grammar with algebra or statistics? *Science*, **284**, 433-435.
- SEIDENBERG, M. S., MACDONALD, M. C., & SAFFRAN, J. R. (2002). Neuroscience: Does grammar start where statistics stop? *Science*, **298**, 553-554.
- SHAIR, H., MASMELA, J., BRUNELLI, S., & HOFER, M. (1997). Potentiation and inhibition of ultrasonic vocalization of rat pups: Regulation by social cues. *Developmental Psychobiology*, **30**, 195-200.
- THIESSEN, E. D., & SAFFRAN, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, **39**, 706-716.
- WEISS, D. J., GARIBALDI, B. T., & HAUSER, M. D. (2001). The production and perception of long calls by cotton-top tamarins (*Saguinus oedipus*): Acoustic analyses and playback experiments. *Journal of Comparative Psychology*, **115**, 258-271.
- WEISS, D. J., & HAUSER, M. D. (2002). Perception of harmonics in the combination long call of cottontop tamarins, *Saguinus oedipus*. *Animal Behaviour*, **64**, 415-426.

## NOTES

1. Transitional probabilities are calculated by  $P(XY) = \frac{(\text{Frequency}XY)}{(\text{Frequency}X)}$ .
2. Rats' calls are limited to squealing and punctual ultrasonic vocalizations related to appetitive or aversive states (Burgdorf, Knutson, Panksepp, & Shippenberg, 2001; Shair, Masmela, Brunelli, & Hofer, 1997).
3. Available at <http://tcts.fpms.ac.be/synthesis/mbrola.html>.
4. Nevertheless, Newport and Aslin (2004) report different results when participants are presented with a slightly different stream.
5. There has been some debate regarding whether the mechanisms responsible for the extraction of these regularities are statistical or algebraic (Marcus, 1999, 2000; Seidenberg, Elman, Negishi, & Eimas, 1999). Nevertheless, these types of regularities are interesting for this study as long as they involve the generalization of a structure, independent of the tokens that instantiate it, and may represent a first step toward the formation of grammatical constructions.

(Manuscript received November 24, 2003;  
revision accepted for publication October 26, 2004.)