



Brief article

Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins

Marc D. Hauser^{a,*}, Elissa L. Newport^b, Richard N. Aslin^b

^a*Department of Psychology and Program in Neurosciences, Harvard University, Cambridge, MA 02138, USA*

^b*Department of Brain & Cognitive Sciences, Meliora Hall – River Campus, University of Rochester, Rochester, NY 14627, USA*

Received 13 May 2000; accepted 4 October 2000

Abstract

Previous work has shown that human adults, children, and infants can rapidly compute sequential statistics from a stream of speech and then use these statistics to determine which syllable sequences form potential words. In the present paper we ask whether this ability reflects a mechanism unique to humans, or might be used by other species as well, to acquire serially organized patterns. In a series of four experimental conditions, we exposed a New World monkey, the cotton-top tamarin (*Saguinus oedipus*), to the same speech streams used by Saffran, Aslin, and Newport (Science 274 (1996) 1926) with human infants, and then tested their learning using similar methods to those used with infants. Like humans, tamarins showed clear evidence of discriminating between sequences of syllables that differed only in the frequency or probability with which they occurred in the input streams. These results suggest that both humans and non-human primates possess mechanisms capable of computing these particular aspects of serial order. Future work must now show where humans' (adults and infants) and non-human primates' abilities in these tasks diverge. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Learning; Speech perception; Serial order; Non-human primates

* Corresponding author. Fax: +1-617-496-7077.

E-mail addresses: hauser@wjh.harvard.edu (M.D. Hauser), newport@bcs.rochester.edu (E.L. Newport), aslin@cvs.rochester.edu (R.N. Aslin).

1. Introduction

Humans have remarkable serial order abilities, demonstrated most clearly by their aptitude for acquiring natural languages. In a series of recent studies we have begun to investigate these abilities experimentally, by designing miniature languages which contain particular serial order properties exemplified in natural languages, and then asking whether, and how, human adults, children and infants acquire these properties from brief exposure to the languages. In the present studies we ask whether a non-human primate, the cotton-top tamarin monkey, can perform the most basic of these learning tasks on the same sequences of human speech.

1.1. *Word segmentation*

A fundamental ability that human language learners must possess is the ability to keep track of the order in which elements of the sound stream occur. One learning problem in which this ability might be used is *word segmentation* – the process of determining which sequences of sounds in the stream of speech form words. While native speakers of a language perceive breaks between the words within an utterance, acoustic analyses reveal that there are no physical cues, such as silence, which reliably signal the end of one word and the beginning of another (Cole & Jakimik, 1980). However, one way of identifying words in an unknown language is to find the recurring sequences of sounds in a corpus of speech (Chomsky, 1955/1975; Harris, 1951). Those sound sequences that occur repeatedly (e.g. *happy* or *children*) are candidates for words, while those that occur only occasionally (e.g. the sequence *pychil*, when these words occur together) are likely to be the accidental juxtapositions at word boundaries. To determine which sound sequences in a stream of speech form words, learners might compute the conditional probabilities between sequential syllables, called *transitional probabilities*¹ (Christophe, Dupoux, Bertocini, & Mehler, 1994; Goodsitt, Morgan, & Kuhl, 1993; Miller & Selfridge, 1950; Saffran, Newport, & Aslin, 1996). Over a speech corpus, syllable sequences with relatively high transitional probabilities are likely to be inside words, while those with relatively low probabilities are likely to be at the word boundaries.

Until recently, it has seemed implausible to imagine that human language learners could conduct such an analysis on-line. Although several studies have shown that infants can learn the order of events (e.g. Bauer & Mandler, 1989; Mandel, Kemler Nelson, & Jusczyk, 1996; Trehub, Thorpe, & Morrongiello, 1987), none have asked whether they can compute more complex order statistics. But, following a study by Hayes and Clark (1970), recent studies have shown that human learners (as young as

¹ Technically, our studies have asked whether subjects compute either the frequency of co-occurrence of syllables or some conditionalized probability of such syllable co-occurrences. In one study (Aslin et al., 1998), we demonstrated that infants do not merely compute co-occurrence frequency, but rather compute conditionalized probability. But in all cases our results are compatible with a number of conditionalized statistics, such as conditional entropy or mutual information, as well as with transitional probability. We cite the notion ‘transitional probability’ in our papers because of its venerable history in psycholinguistics, but we use this term as a reference to a class of conditionalized statistics.

8-month-old infants) are capable of keeping track of such sequential statistics in rapidly presented auditory materials (Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). In these studies, subjects were presented with a stream of speech, composed of a random ordering of trisyllabic nonsense ‘words’ from an artificial language. There were no acoustic cues to word boundaries; words were specified only by the statistics of the syllable sequences. Adults, children, and infants exposed to such streams discriminated the words (sequences with relatively high transitional probabilities) from sequences that had occurred in the familiarization stream with low transitional probabilities. Human learners are also capable of using statistical cues to segment non-speech input, such as musical tones (Saffran, Johnson, Aslin, & Newport, 1999), visual patterns (Asaad, 1998; Fiser & Aslin, 2000), and visuomotor sequences (Hunt & Aslin, 1998). In short, human learners are readily able to acquire the statistics of sequential patterns across a number of modalities and domains.

1.2. Sequential abilities in non-humans

Since humans can compute such statistics across a wide range of patterned materials, it is of interest to ask whether non-human animals can perform such basic computations on at least some of the same materials. Studies of natural vocal communication systems, as well as operant studies of serial learning, suggest that both birds and primates may have some clear, though limited, sequential abilities. For example, chickadees, zebra finches, and a number of other oscine species produce the notes and syllables of their songs in a fixed or stereotyped sequence, with order playing a crucial role in species, dialect, and individual identity (Hailman & Ficken, 1987; Hailman, Ficken, & Ficken, 1985, 1987; Immelman, 1969; Marler, 1977, 2000; Price, 1979). For some oscines (e.g. zebra finches), the particular syllables produced and the order in which they are sung is acquired from hearing the father’s song.² Tamarins, titi monkeys, and capuchins (Cleveland & Snowdon, 1981; Robinson, 1979, 1984) sometimes combine discrete call types into a sequence that results in a new call type. However, the acoustic structure of most non-human primate vocal repertoires appears to be relatively fixed at birth, with auditory experience playing only a minimal role in shaping call morphology (cf. Seyfarth & Cheney, 1999, for a review). Operant conditioning studies have shown that pigeons, capuchin monkeys, and rhesus macaques can learn arbitrary serial lists of up to six or so visually or auditorily presented items through training (D’Amato & Colombo, 1989; Schwartz, Chen, & Terrace, 1990; Terrace, 1987, 1991; Wright & Rivera, 1997). It has also been suggested that human-reared bonobos can perceive spoken words and can respond correctly to word order (Savage-Rumbaugh et al., 1993).

While these studies provide evidence of some serial order abilities in non-human

² For the present purposes we focus only on the issue of sequence and leave aside the question of whether varying sequences occur and have new combinatorial meanings. See Hauser (2000) for discussion.

species, many patterned domains (such as language and music) require serial order processing that is more complex than learning fixed sequences. In such domains, learners must keep track of *relative* serial order among pattern elements, and also the *frequency* or *consistency* with which various combinations and orders appear. Human learners are apparently capable of performing such computations on a large number of syllable sequences simultaneously, extremely rapidly as speech occurs in real time, and without training or reinforcement.

In the present studies we explore the species generality of these more complex serial order abilities to determine whether rapid statistical serial order learning is spontaneously present in tamarin monkeys – a species whose own vocalizations have some rudimentary sequential structure (Cleveland & Snowdon, 1981; Weiss & Hauser, unpublished data). To ensure that the abilities in question are comparable to those evaluated in humans, we ask whether tamarins can perform the same task on the same sequences of speech syllables that we presented to human infants, and tested with the same general methods.

2. Method

2.1. Subjects

Subjects were 13 adult cotton-top tamarins (*Saguinus oedipus*), seven females and six males. This species is native to the rainforests of Colombia. All subjects were born in captivity at the New England Regional Primate Research Center, Southborough, MA and have been housed at the Primate Cognitive Neuroscience Lab, Harvard University for approximately 8 years. Animals are housed in social groups consisting of a mating pair, and in some cases, their offspring. Tamarins' frequency sensitivity includes the range of human speech sounds (Cleveland & Snowdon, 1981; Stebbins, 1983).

All subjects have experience in playback experiments (Garibaldi, 1999; Ramus, Hauser, Miller, Morris, & Mehler, 2000; Weiss, Garibaldi, & Hauser, in press) and in experiments involving other cognitive abilities (Hauser, 1997, 1998; Santos & Hauser, 1999). Because of this, subjects voluntarily move in and out of their home cage and into a test area. When they arrive in the testing area, they are calm and will typically remain so for approximately 30 min. We can therefore present stimuli over a relatively long period of time without distressing them. All 13 subjects were tested in each of four conditions, but not all 13 provided useable data (see below).

2.2. Stimuli

Stimuli consisted of two of the same streams of synthetic speech syllables used in the human infant studies of Saffran, Aslin, and Newport (1996). The first speech stream (Language A) consisted of a 20 min random ordering of four three-syllable nonsense words (*tupiro*, *golabu*, *bidaku*, and *padoti*). The second speech stream (Language B) consisted of a similarly structured stream of four different words (*tudaro*, *pigola*, *bikuti*, and *budopa*). For both streams, the rate of presentation

was 4.5 syllables/s, with no acoustic cues at word boundaries. The only available information for extracting words was the greater statistical regularity of syllable sequences within words than of syllable sequences that spanned a word boundary. The coherence of syllables within words was tested against the reduced coherence of syllables within one of two types of foils: non-words, which contained syllables in an order that never occurred in the speech stream, or part-words, which contained syllables spanning a word boundary.

Two types of statistical cues differentiated words from part-words or non-words: the transitional probabilities between adjacent syllables, and the frequency of co-occurrence of adjacent syllables. Within the 20 min stream of speech, each of the four trisyllabic nonsense words occurred equally often and in random order, with the constraint that no word was immediately repeated. The transitional probabilities³ for syllables inside words were thus 1.0 (between syllables 1 and 2) and 1.0 (between syllables 2 and 3), the transitional probabilities for syllables within part-words were 0.33 (between syllable 3 of one word and syllable 1 of the next) and 1.0 (between syllable 1 and 2 of a word), and the transitional probabilities for syllables within non-words were 0.0 and 0.0. The words, part-words, and non-words also differed in the frequency with which these syllable sequences occurred in the exposure stream.⁴ Each of the words occurred three times more frequently than each of the part-words; each of the non-words never occurred in the stream. Thus, both transitional probabilities and trisyllabic co-occurrence frequencies differed between the words and the part- or non-words. In either case, in order to discriminate between the test items, subjects would have to compute complex statistics involving the frequency or consistency with which syllables occur in a particular relative order. Test items did not differ in the frequency of individual syllables.

2.3. Design

Because of the small number of available tamarins, a within-subject design was employed in which all of the tamarins were presented with two different speech streams, each tested on two different sets of test item contrasts, over four separate 2 day familiarization–test sessions. The first session consisted of Language A and the word (*tupiro*, *golabu*) versus non-word (*dapiku*, *tilado*) test. The second session, conducted 1 week later, consisted of Language A and the word (*tupiro*, *golabu*) versus part-word (*tibida*, *kupado*) test. The third session, conducted 7 weeks later, consisted of Language B and the word (*tudaro*, *pigola*) versus part-word (*pabiku*, *tibudo*) test. The final session, run 1 week later, consisted of Language B and the word (*tudaro*, *pigola*) versus non-word (*tigobu*, *kudabi*) test. This design (see Table 1) ensured that differential responding to the words versus non- or part-words could not result from intrinsic preferences for certain test items, because the words, non-words, and part-words in Languages A and B were different syllable sequences.

³ $[P(y|x) = P(xy)/P(x)]$.

⁴ See Aslin et al. (1998) for a paradigm used with infants in which only transitional probabilities differed between the test items, and co-occurrence frequency was matched.

Table 1
Design of Languages A and B and test items comparing words versus non-words or words versus part-words

	Language A	Language B
Words	tupiro, golabu, bidaku, padoti	tudaro, pigola, bikuti, budopa
Test words	tupiro, golabu	tudaro, pigola
Test non-words	dapiku, tilado	tigobu, kudabi
Test part-words	tibida, kupado	pabiku, tibudo

Moreover, the order of the conditions across the four sessions ensured that performance on session 1 (Language A, non-words) and session 3 (Language B, part-words) was uncontaminated by previous exposure to the test items.⁵

2.4. Experimental procedure

Each of the four test sessions involved the same 2 day familiarization–test procedure, and differed only in the stimuli presented for familiarization and test. On day 1, the entire colony of tamarins was familiarized to a 20 min continuous speech stream, while housed in their home room. On day 2, each tamarin was placed individually in the sound proof chamber (Fig. 1A) and re-familiarized with a 1 min corpus of the speech stream, followed immediately by four test trials, in random order: single instances of two words and two non-words, or two words and two part-words. The mean inter-trial interval for test trials was 23.1 s (SD 10.5). We did not run subjects who failed to leave their home room cage on the day of testing, or those who jumped around the test cage and failed to sit quietly (three in session 3, four in session 4). The dependent measure was an orienting response to a test stimulus presented from a concealed loudspeaker. For clarity in scoring a response, stimuli were presented when the subject was still and facing 180° away from the concealed speaker (that is, looking down and away from the speaker; Ramus et al., 2000). Subjects were scored as responding if they turned and looked in the direction of the speaker either during the presentation of the test stimulus, or within 2 s afterwards. Fig. 1B presents an example of a test trial and an orienting response following stimulus presentation. Orienting responses were scored independently by two observers from digitized video recordings; the observers were blind to the test condition (the inter-observer reliability was 82–92%). Trials on which subjects were not facing away from the speaker at stimulus onset, or on which observers could not code a response unambiguously, were eliminated (6%). This procedure has been used reliably in other experiments (Garibaldi, 1999; Ramus et al., 2000; Weiss et al., in press).

⁵ Since the same words are used as test items for the part-word and the non-word test sessions, it is possible that a novelty effect for the part-word/non-word on the second test session for a given language could result from having already been tested with the words in the previous session. To test for this effect, we ran the word/non-word contrast first in Language A and the word/part-word contrast first in Language B. No difference appeared between the first and second use of the same words in this design, and both first sessions showed a significant discrimination.

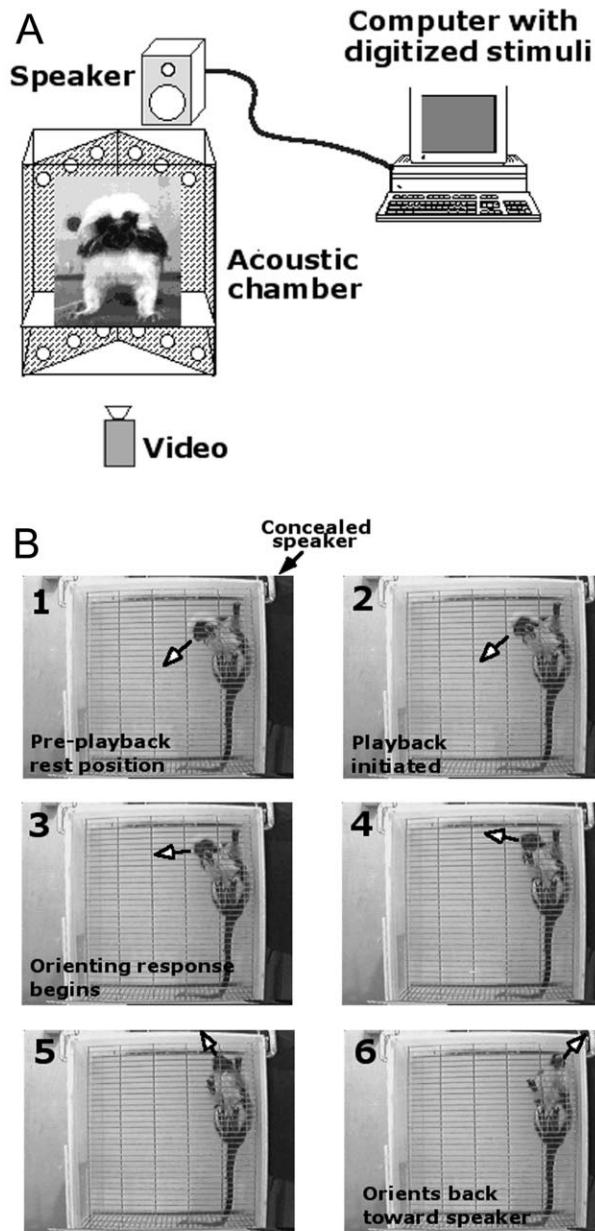


Fig. 1. (A) Experimental set-up for running cotton-top tamarins. (B) Frame-by-frame illustration of a tamarin's behavior during a playback trial.

3. Results

Each animal's responses in each test condition (words versus non- or part-words) were converted to a proportion, and these proportions were averaged across animals. Fig. 2 presents the mean proportion of trials showing an orienting response to words versus non-words. Non-words consisted of three syllables from the language in an order that had never occurred in the familiarization corpus. This contrast thus assesses subjects' learning of syllable order. As can be seen, the tamarins were significantly more likely to orient to non-words than to words (Wilcoxon Signed Ranks: $z(13) = 2.45$, $P = 0.01$ for Language A; $z(9) = 1.54$, $P = 0.12$ for Language B; $z(13) = 2.67$, $P < 0.01$ for A and B combined across each animal's useable test trials). Words and non-words were matched in the frequency of each of the individual syllables, and differed only in the order in which these syllables occurred. This discrimination therefore shows that tamarins acquired syllable order from their exposure to the speech streams, a result that parallels those obtained by Saffran, Aslin, and Newport (1996) with human infants.

Fig. 3 presents the mean proportion of trials showing an orienting response to words versus part-words. Part-words consist of the final syllable of one word and the first two syllables of another word; part-words are therefore sequences which did occur in the corpus. Comparing subjects' responses to words versus part-words assesses their learning of the frequency or probability with which various syllable orders occurred. Tamarins were significantly more likely to orient to part-words than to words (Wilcoxon Signed Ranks: $z(13) = 2.93$, $P = 0.003$ for Language A;

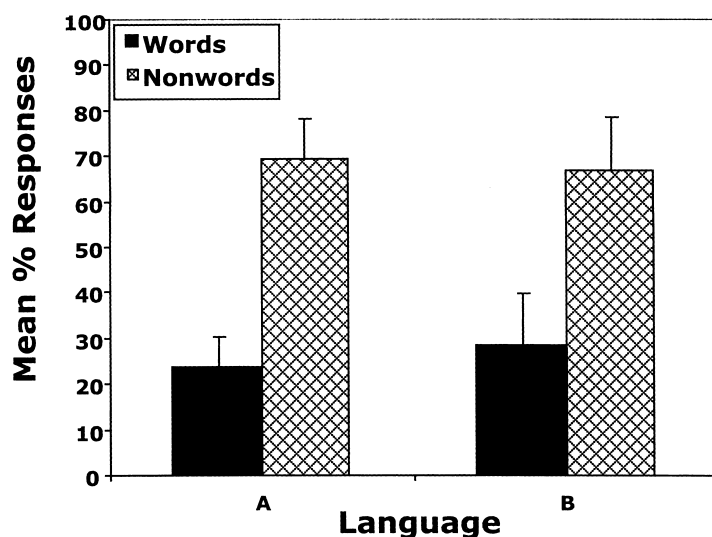


Fig. 2. Mean (standard error) percent of *word* versus *nonword* test trials on which subjects responded, for Language A (left) and Language B (right). Black bars indicate responses to *word* trials, cross-hatched bars indicate responses to *nonword* trials.

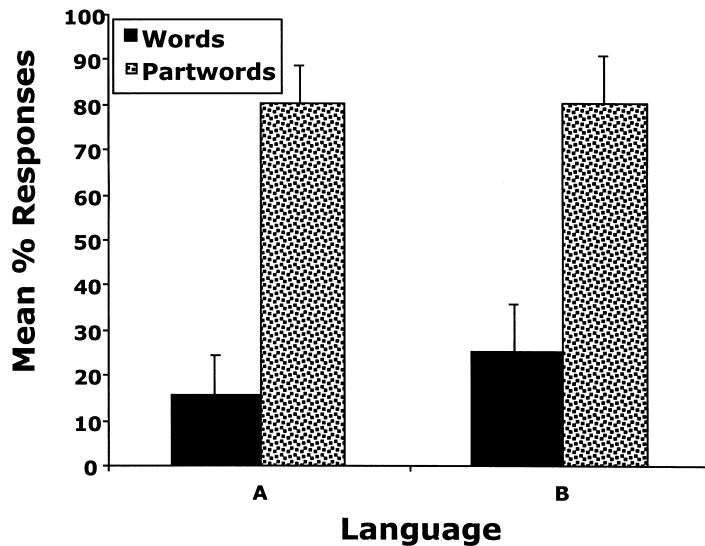


Fig. 3. Mean (standard error) percent of *word* versus *partword* test trials on which subjects responded, for Language A (left) and Language B (right). Black bars indicate responses to *word* trials, stippled bars indicate responses to *partword* trials.

$z(10) = 2.24$, $P = 0.025$ for Language B; $z(13) = 2.93$, $P = 0.003$ for A and B combined across each animal's useable test trials).

These results show much more complex serial order learning than the word/non-word contrast, and again parallel the findings of Saffran, Aslin, and Newport (1996) for human infants. Since words and part-words differ only in the transitional probability (or co-occurrence frequency) of one syllable pair within the trisyllabic unit, successful discrimination shows that tamarins, like human infants, are computing fairly complex serial order statistics during presentation of the familiarization corpus, and that they are performing these computations rapidly, reasonably accurately over a large number of syllable pairs, and without training or reinforcement for doing so.

Although the overall results for non-words and for part-words across the two languages were significant, the key data come from Language A (non-words) and Language B (part-words). These two conditions, which both showed significant evidence of discrimination between the two types of test items, could not have resulted from the tamarins' previous exposure to the specific test items (which was possible for Language A, part-words and Language B, non-words).

4. Discussion

Our results show that tamarin monkeys – a species known to produce their natural calls with some fixed order properties and to perceive certain contrasts in natural human speech (Garibaldi, 1999; Ramus et al., 2000; Weiss et al., in press) – are

capable of computing the same rather complex statistical regularities on fluent and rapid streams of synthetic human speech that have been demonstrated, by similar methods, in human infants (Saffran, Aslin, & Newport, 1996). These discriminations entail a number of separate abilities worthy of note.

First, to perform this task successfully, tamarins must be capable of discriminating speech syllables from one another, and treating these syllables (or the phonemes of which they are composed) as elementary units. Second, they must be capable of keeping track of the serial order in which these syllables occur, as well as the frequency and/or transitional probability of each syllable pair or syllable triplet computed over a sizeable sample of the rapid speech stream.⁶ Moreover, tamarins, like human infants, must be capable of performing such computations rapidly and without training or reinforcement.

These results suggest that humans and tamarins may share a common mechanism for rapidly processing and retaining these aspects of serial order information. However, even if humans and some non-human primates share the ability to compute certain elementary patterns of serial order, these abilities in humans (at least for the domain of language) have been extended and elaborated well beyond what we have demonstrated in the present tasks, into complexities of serial order processing that we do not expect non-human primates to share. The type of statistical computation examined here (computing co-occurrence relations between immediately adjacent elements) is a fundamental aspect of serial learning, potentially important to providing learners with an initial segmentation of speech or many other types of patterned domains (Newport & Aslin, 2000; Saffran et al., 1999). However, such an analysis would not be adequate for acquiring higher level patterns in language; many structures in natural languages involve relationships between classes of elements, or relations between elements not immediately adjacent to each other (cf. Chomsky, 1957). This suggests that some computational mechanisms beyond those studied here may be specific to humans and absent or deficient in monkeys, since only humans show the full range of linguistic skills that characterize natural language systems. In ongoing studies comparing infants and tamarins on these more extended aspects of serial order learning, we hope to clarify precisely where these species' abilities diverge. In the present work, however, we believe we have shown a surprising commonality in the ability to compute certain fundamental statistics of serial order, in the absence of explicit training.

Acknowledgements

We are grateful to Ruskin Hunt and Kelly Kinde for their help in preparing stimuli, and to Cory Miller and Geertrui Spaepen for their help in running the experiments and scoring responses. This work was supported in part by NSF grant SBR-9873477 to all three authors.

⁶ For each language, our task presents 12 distinct syllables and 20 different transitional probabilities (or 20 distinct syllable pairs); eight of these transitional probabilities are tested in the test stimulus contrasts.

References

- Asaad, P. (1998). *Statistical learning of sequential visual patterns*. Honors thesis, Department of Brain & Cognitive Sciences, University of Rochester, Rochester, NY.
- Aslin, R., Saffran, J., & Newport, E. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *27*, 321–324.
- Bauer, P. J., & Mandler, J. M. (1989). One thing follows another: effects of temporal structure on 1- to 2-year-olds' recall of events. *Developmental Psychology*, *25*, 197–206.
- Chomsky, N. (1955/1975). *The logical structure of linguistic theory*. New York: Plenum Press.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical approach to the bootstrapping problem for lexical acquisition. *Journal of the Acoustical Society of America*, *95*, 1570–1580.
- Cleveland, J., & Snowdon, C. T. (1981). The complex vocal repertoire of the adult cotton-top tamarin, *Saguinus oedipus oedipus*. *Zeitschrift für Tierpsychologie*, *58*, 231–270.
- Cole, R. A., & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- D'Amato, M. R., & Colombo, M. (1989). Representation of serial order in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 131–139.
- Fiser, J., & Aslin, R. N. (2000). *Position-dependent and position-independent statistical learning of shape conjunctions*. European Conference on Visual Perception, Rotterdam, The Netherlands.
- Garibaldi, B. (1999). *Sources of acoustic variation and perception in cotton-top tamarins*. Doctoral dissertation, Harvard University, Cambridge, MA.
- Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language*, *20*, 229–252.
- Hailman, J. P., & Ficken, M. S. (1987). Combinatorial animal communication with computable syntax: chick-a-dee calling qualifies as 'language' by structural linguistics. *Animal Behaviour*, *34*, 1899–1901.
- Hailman, J. P., Ficken, M. S., & Ficken, R. W. (1985). The "chick-a-dee" calls of *Parus atricapillus*: a recombinant system of animal communication compared with written English. *Semiotica*, *56*, 191–224.
- Hailman, J. P., Ficken, M. S., & Ficken, R. W. (1987). Constraints on the structure of combinatorial "chick-a-dee" calls. *Ethology*, *75*, 62–80.
- Harris, Z. (1951). *Methods in structural linguistics*. Chicago, IL: University of Chicago Press.
- Hauser, M. D. (1997). Artfactual kinds and functional design features: what a primate understands without language. *Cognition*, *64*, 285–308.
- Hauser, M. D. (1998). Expectations about object motion and destination: experiments with a nonhuman primate. *Developmental Science*, *1*, 31–38.
- Hauser, M. D. (2000). *Wild minds: what animals really think*. New York: Henry Holt.
- Hayes, J. R., & Clark, H. H. (1970). Experiments in the segmentation of an artificial speech analog. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.
- Hunt, R. H., & Aslin, R. N. (1998). Statistical learning of visuomotor sequences: implicit acquisition of sub-patterns. *Proceedings of the twentieth annual conference of the Cognitive Science Society*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Immelman, K. (1969). Song development in the zebra finch and other estrildid finches. In R. A. Hinde (Ed.), *Bird vocalizations* (pp. 61–74). Cambridge: Cambridge University Press.
- Mandel, D. R., Kemler Nelson, D. G., & Jusczyk, P. W. (1996). Infants remember the order of words in a spoken sentence. *Cognitive Development*, *11*, 181–196.
- Marler, P. (1977). The structure of animal communication sounds. In T. H. Bullock (Ed.), *Recognition of complex acoustic signals* (pp. 17–35). Berlin: Springer-Verlag.
- Marler, P. (2000). Origins of music and speech: insights from animals. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 31–48). Cambridge, MA: MIT Press.

- Miller, G. A., & Selfridge, J. A. (1950). Verbal context and the recall of meaningful material. *American Journal of Psychology*, *63*, 176–185.
- Newport, E. L., & Aslin, R. N. (2000). Innately constrained learning: blending old and new approaches to language acquisition. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University conference on language development*. Somerville, MA: Cascadilla Press.
- Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *Journal of Comparative and Physiological Psychology*, *93*, 260–277.
- Ramus, F., Hauser, M. D., Miller, C. T., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*, 349–351.
- Robinson, J. G. (1979). An analysis of the organization of vocal communication in the titi monkey *Callicebus moloch*. *Zeitschrift für Tierpsychologie*, *49*, 381–405.
- Robinson, J. G. (1984). Syntactic structures in the vocalizations of wedge-capped capuchin monkeys, *Cebus nigrivittatus*. *Behaviour*, *90*, 46–79.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Saffran, J., Johnson, E., Aslin, R., & Newport, E. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.
- Saffran, J., Newport, E., & Aslin, R. (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.
- Saffran, J., Newport, E., Aslin, R., Tunick, R., & Barrueco, S. (1997). Incidental language learning: listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101–105.
- Santos, L. R., & Hauser, M. D. (1999). How monkeys see the eyes: cotton-top tamarins' reaction to changes in visual attention and action. *Animal Cognition*, *2*, 131–139.
- Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., & Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, *58*, 1–221.
- Schwartz, K., Chen, S., & Terrace, H. S. (1990). Multiple list learning by rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 396–410.
- Seyfarth, R. M., & Cheney, D. L. (1999). Production, usage, and response in nonhuman primate vocal development. In M. D. Hauser, & M. Konishi (Eds.), *The design of animal communication*. Cambridge, MA: MIT Press.
- Stebbins, W. C. (1983). *The acoustic sense of animals*. Cambridge, MA: Harvard University Press.
- Terrace, H. S. (1987). Chunking by a pigeon in a serial learning task. *Nature*, *325*, 149–151.
- Terrace, H. S. (1991). Chunking during serial learning by a pigeon: I. Basic evidence. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 81–93.
- Trehub, S. E., Thorpe, L. A., & Morrongiello, B. (1987). Organizational processes in infants' perception of auditory patterns. *Child Development*, *58*, 741–749.
- Weiss, D., Garibaldi, B., & Hauser, M. D. (in press). The production and perception of long calls in cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology*.
- Wright, A. A., & Rivera, J. J. (1997). Memory of auditory lists by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, *23*, 441–449.