2 (1997)]. Care was taken to equate crystal size and placement between littermate sets of wild type, heterozygote, and homozygote mutants. Placement and sizes of injection sites were verified in sections counterstained with bisbenzimide. Quantification was done blinded to genotype.

18. R. Tuttle, J. E. Braisted, Y. Nakagawa, D. D. M. O'Leary, unpublished observations on *Sey*/*Sey* mice; see also work by H. Kawano *et al*. [*J. Comp. Neurol*. **408**, 147 (1999)] for similar findings on *Pax6* mutant rats.
19. K. M. Bishop, G. Goudreau, D. D. M. O'Leary, unpublished observations.
20. A. Simeone *et al.*, *EMBO J.* **11**, 2541 (1992).
21. J. Altman and S. A. Bayer, *J. Comp. Neurol*. **275**, 346 (1988).
22. E18.5 *Emx2* embryos were collected as in (*13*). AChE histochemistry was as in work by B. L. Schlaggar, J. A. De Carlos and D. D. M. O'Leary [*Dev. Brain Res.* **75**, 19 (1993)]. In situ hybridization and counterstaining were as in work by R. Tuttle, Y. Nakagawa, J. E. Johnson and D. D. M. O'Leary [*Development* **126**, 1903 (1999)].
23. A delay in overall development of the cortex cannot account for our findings. The mutant patterns of cadherin expression and thalamocortical projections do not resemble patterns observed at earlier stages in wild types. For example, the domains of cadherin expression do not exhibit significant proportional expansions or contractions during normal embryonic cortical development (*6*); axons from VP neurons do not transiently invade the visual area at early stages and the somatosensory area later, and vice versa for dLG axons; and the distribution of corticothalamic projection neurons is area specific, even at the earliest times that they can be labeled (*16*). *Emx2* and *Pax6* mutants also show opposite changes in cadherin expression patterns, which cannot be accounted for by a delay in cortical development in both mutants.
24. P.-P. Gao *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5329 (1998); D. D. M. O'Leary and D. G. Wilkinson, *Curr. Opin. Neurobiol*. **9**, 55 (1999).
25. J. Rusch and M. Levine, *Curr. Opin. Genet. Dev*. **6**, 416 (1996).
26. D. Stanojevic, S. Small, M. Levine, *Science* **254**, 1385 (1991); S. Small, A. Blair, M. Levine, *Dev. Biol*. **175**, 314 (1996).
27. J. L. R. Rubenstein *et al.*, *Cereb. Cortex* **9**, 524 (1999).
28. Independent regulation of *Emx2* and *Pax6* is suggested by their temporal coexpression and the findings that *Emx2* is expressed in the *Pax6* mutant and that *Pax6* is expressed in the *Emx2* mutant in graded patterns in the developing neocortex similar to those in wild-type mice [A. Stoykova, R. Fritsch, C. Walther, P. Gruss, *Development* **122**, 3453 (1996); K. M. Bishop and D. D. M. O'Leary, unpublished observations; see Web fig. 2, available at www.sciencemag.org/feature/data/1045964.shl]. In addition, analysis of the *extra toes* mouse mutant (a null mutation of *Gli3*) shows that the loss of *Gli3* expression correlates with a loss of *Emx2* expression in the rostral forebrain but does not affect *Pax6* expression in the developing neocortex (*30*).
29. P. H. Crossley and G. R. Martin, *Development* **121**, 439 (1995); E. A. Grove, S. Tole, J. Limon, L. Yip, C. W. Ragsdale, *Development* **125**, 2315 (1998); J. L. R. Rubenstein and P. A. Beachy, *Curr. Opin. Neurobiol.* **8**, 18 (1998).
30. T. Theil, G. Alvarez-Bolado, A. Walter, U. Rüther, *Development* **126**, 3561 (1999).

31. We are grateful to P. Gruss for the gift of *Emx2* mice, M. Goulding for *Sey* mice and *Pax6* plasmid, and Y. Nakagawa for *Cad6*, *Cad8*, and *Gbx2* plasmids and helpful discussions. This work was supported by NIH grant NS31558 (D.D.M.O'L.), the Natural Sciences and Engineering Research Council of Canada (K.M.B.), and the Max-Planck Society and European Commission grant BI04-CT96-0378 (G.G.).

4 October 1999; accepted 22 February 2000

Language Discrimination by Human Newborns and by Cotton-Top Tamarin Monkeys

Franck Ramus,[1]*† Marc D. Hauser,[2] Cory Miller,[2] Dylan Morris,[2] Jacques Mehler[1]

Humans, but no other animal, make meaningful use of spoken language. What is unclear, however, is whether this capacity depends on a unique constellation of perceptual and neurobiological mechanisms or whether a subset of such mechanisms is shared with other organisms. To explore this problem, parallel experiments were conducted on human newborns and cotton-top tamarin monkeys to assess their ability to discriminate unfamiliar languages. A habituation-dishabituation procedure was used to show that human newborns and tamarins can discriminate sentences from Dutch and Japanese but not if the sentences are played backward. Moreover, the cues for discrimination are not present in backward speech. This suggests that the human newborns' tuning to certain properties of speech relies on general processes of the primate auditory system.

A fundamental question in the study of language evolution and acquisition is the extent to which humans are innately endowed with specialized capacities to comprehend and produce speech. Theoretical arguments have been used to argue that language acquisition must be based on an innately specified language faculty (*1*, *2*), but the precise nature and extent of this "language organ" is mainly an empirical matter, which notably requires studies of human newborns as well as nonhuman animals (*3*–*5*). With respect to studies of humans, we already know that newborns as young as 4 days old have the capacity to discriminate phonemes categorically (*6*) and perceive well-formed syllables as units (*7*–*9*); they are sensitive to the rhythm of speech, as shown in experiments in which newborns distinguish sentences from languages that have different rhythmic properties but not from languages that share the same rhythmic structure (*10*, *11*); however, newborns do not discriminate languages when speech is played backward (*10*), and neurophysiological studies suggest that both infants and adults process natural speech differently from backward speech (*12*, *13*). All these studies indicate that humans are born with capacities that facilitate language acquisition and that seem well attuned to the properties of speech. Studies of nonhuman animals, however, show that some of these capacities may predate our hominid origins. For example, insects, birds, nonprimate mammals, and primates process their own, species-typical sounds in a categorical manner, and some of these species perceive speech categorically (*14*–*18*).

Our aim here is to extend the comparative study of speech perception in three directions. First, using the same design and the same material, we have conducted joint experiments on human newborns and on monkeys. Second, whereas most studies of nonhuman animal speech perception involve extensive training before testing on a generalization task, our experimental approach—the habituation-dishabituation paradigm—involves no training and parallels the method used in studies of infant speech perception. Thus, conditions are met to appropriately compare the two populations. Third, most studies of speech processing in animals involve tests of phonemic perception. Here, we extend the analysis to sentence perception, thereby setting up a much broader range of perceptual problems.

Our experiments were run on human newborns and cotton-top tamarin monkeys (*Saguinus oedipus oedipus*). The stimuli consisted of 20 sentences in Japanese and 20 sentences in Dutch uttered by four female native speakers of each language. Conditions in which the two languages are pitted against one another were compared with conditions in which speakers of the same language are contrasted. In addition, sentences within a session were played either forward or backward. To more readily control for prosodic features of the signal, we reran all conditions with synthesized exemplars of the original sentences. Synthesized sentences were created with the MBROLA diphone synthesizer (*19*). Phoneme duration and fundamental frequency were preserved, whereas the phonetic inventory was narrowed to only one phoneme per manner of articulation: all fricatives were synthesized as /s/, vowels as /a/, liquids as /l/, plosives as /t/, nasals as /n/, and glides as /j/.

---

[1] Laboratoire de Sciences Cognitives et Psycholinguistique, l'Ecole des Hautes Etudes en Sciences Sociales/Centre National de la Recherche Scientifique, 54 Boulevard Raspail, 75006 Paris, France. [2] Department of Psychology, Harvard University, Cambridge, MA, 02138, USA.

*To whom correspondence should be addressed. E-mail: f.ramus@ucl.ac.uk

†Present address: Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR, UK.

Thus, each synthesized sentence preserved only the prosodic characteristics of its natural counterpart while eliminating lexical and phonetic information (*20*).

We tested newborns with the high-amplitude sucking procedure and a habituation/dishabituation design. Sentences were elicited by the newborns' sucking on a pacifier. In the language change condition, newborns were habituated to 10 sentences uttered by two speakers in one language and then switched to 10 sentences uttered by two different speakers in the other language. In the speaker change condition, newborns were habituated to 10 sentences uttered by two speakers from one language and then switched to two different speakers in the same language. A significant increase in sucking after the language change, compared with the speaker change, is taken as evidence that newborns perceive a significant difference between the two languages (*21*).

We tested 32 newborns (*22*) on the natural language-forward experiment: 16 in the language change condition and 16 in the speaker change condition. Figure 1A shows that the two groups did not differ significantly and thus that newborns failed to discriminate the two languages ($F_{(1,29)} < 1$) (*23*). This result appears to conflict with previous experimental work showing that newborns discriminate English and Japanese. However, our experiment exposes newborns to great speaker variability (four voices) (*24*), and this factor has previously been shown to impair the discrimination abilities of infants (*25*). If speaker variability is responsible for the absence of discrimination, then we would predict successful discrimination with fewer speakers. To test for this possibility, we ran a second experiment using synthesized speech, thereby reducing the number of voices to one, that of the speech synthesizer (*26*).

We tested 32 additional newborns (*27*) on the forward language and speaker discrimination using the synthesized versions of the original sentences. Figure 1B shows that newborns in the language change condition increased their sucking significantly more during the 2 min after the switch than newborns in the speaker change condition ($F_{(1,29)} = 6.3$, $P = 0.018$). This indicates that, relying exclusively on prosodic cues, newborns discriminate sentences of Dutch from sentences of Japanese. Moreover, this result shows that the failure of newborns to discriminate in experiment 1A was probably due to speaker variability.

To determine the specificity of the newborns' capacity to discriminate languages, we tested 32 more newborns with the same synthesized sentences played backward (*28*). Figure 1C shows that newborns fail to discriminate languages played backward ($F_{(1,29)} < 1$) (*29*). Moreover, the interaction between experiments 1B and 1C (forward vs. backwards) is marginally significant ($F_{(1,59)} = 3.6$, $P = 0.06$). The finding that newborns discriminate two nonnative languages played forward but not backward suggests that the newborns' language discrimination capacity may depend on specific properties of speech that are eliminated when the signal is played backward. However, before drawing such a conclusion, it is important to directly assess the speech specificity of this capacity by testing it on another species.

We tested cotton-top tamarins ($n = 13$) with the same stimulus set as the newborns. Instead of sucking rate, however, we used a head orientation response toward the loudspeaker. During the habituation phase, a tamarin was presented with sentences uttered by two speakers in one language and then tested with a sentence uttered by a different speaker, either in the same language (speaker change condition) or in the other language (language change condition). Recovery of orientation toward the loudspeaker was interpreted as an indication that the tamarin perceived a difference between the habituation and test stimuli (*30*).

Experiment 2A involved natural sentences of Dutch and Japanese played either forward or backward (*31*). Figure 2A shows that 10 of 13 tamarins ($P < 0.05$; binomial test) dishabituated in the language change condition, whereas only 5 of 13 dishabituated to the speaker change ($P = 0.87$). The difference between language and speaker change is significant ($P < 0.05$; χ^2 test). This result suggests that the tamarins discriminated Dutch from Japanese regardless of speaker variation. Surprisingly, such a pattern was not observed when the sentences were played backward: only 5 of 13 tamarins dishabituated to the backward language change ($P = 0.87$); this pattern is not significantly different from the speaker change condition ($P > 0.2$). These results parallel those obtained with newborns on the synthetic stimuli.

In experiment 2B, we tested the same tamarins on both the speaker and the language conditions but with synthesized sentences. Figure 2B shows that 10 of 13 tamarins dishabituated to the forward language change ($P < 0.05$). Although the number of subjects dishabituating to the speaker change failed to reach statistical significance ($P = 0.29$), the increased numbers in this condition led to a nonsignificant difference between language and speaker change for the synthesized sentences ($P > 0.3$). For backward sentences, subjects failed to show a statistically significant level of dishabituation to either the language or the speaker change ($P = 0.29$ and $P = 0.13$). Experiment 2B suggests that the ability of tamarins to discriminate Dutch and Japanese is diminished when only prosodic cues are available.

When the data from experiments 2A and 2B are pooled (Fig. 2C), the overall result is clear:

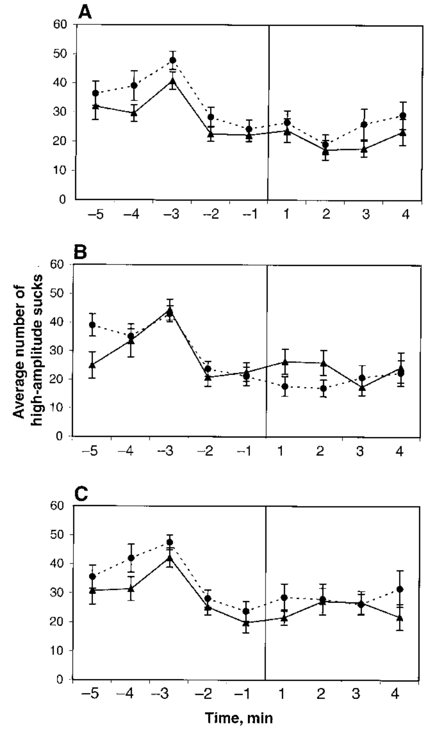

Fig. 1. Average number of high amplitude sucks per minute for babies in the control (speaker change, dotted lines) and experimental (speaker and language change, solid lines) groups. Minutes are numbered from the time of change. Error bars represent ±1 SEM. (**A**) Natural sentences played forward. (**B**) Same sentences synthesized. (**C**) Same sentences synthesized and played backward.

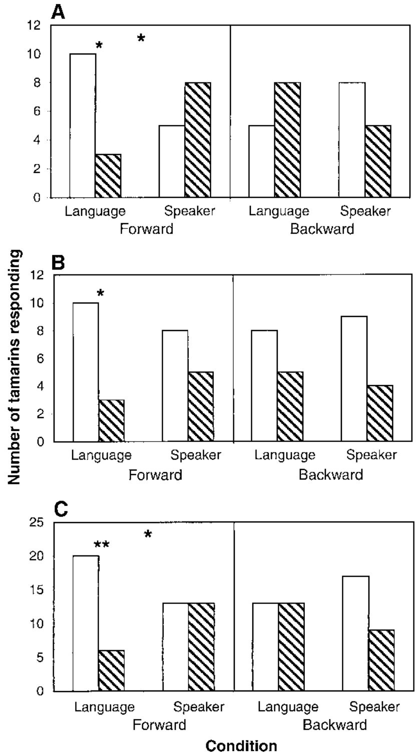

Fig. 2. Number of tamarins responding positively (white bars) and negatively (hatched bars) to test sentence depending on condition: language or speaker change, sentences played forward or backward. (**A**) Natural sentences. (**B**) Synthesized sentences. (**C**) Data from experiments 2A and 2B pooled together. $^*P < 0.05$. $^{**}P < 0.01$.

when sentences are played forward, tamarins significantly dishabituate to the language change ($P = 0.005$) but not to the speaker change ($P = 0.58$), and the difference between language and speaker change is significant ($P < 0.05$). When sentences are played backward, no such effect is observed. This overall result parallels that obtained with human newborns: both species discriminate sentences of Dutch and Japanese played forward but not backward.

The pattern of our results suggests striking similarities as well as differences between the monkey and the human auditory systems. First, we have shown that tamarins, like human newborns, are able to process not just isolated syllables but also whole strings of continuous speech and to extract enough information to discriminate between Dutch and Japanese. Second, their ability to do so above and beyond speaker variability suggests that they are able to extract auditory equivalence classes—that is, to extract abstract linguistic invariants despite highly variable acoustic shapes (*17*, *32*). Third, the fact that, like newborns, tamarins fail to discriminate when speech is played backward suggests that their language discrimination capacity relies not on trivial low-level cues but rather on quite specific properties of speech. Because tamarins have not evolved to process speech, we infer in turn that at least some aspects of human speech perception may have built upon preexisting sensitivities of the primate auditory system. Finally, unlike newborns, tamarins fail to discriminate the language change more than the speaker change when speech is resynthesized. This leaves open the possibility that human newborns and tamarins may not be responding to exactly the same cues in the sentences: tamarins might be more sensitive to phonetic than to prosodic contrasts.

References and Notes

1. N. Chomsky, *Language and Problems of Knowledge* (MIT Press, Cambridge, MA, 1988).
2. S. Pinker, *The Language Instinct* (William, Morrow, and Co., New York, 1994).
3. A. Doupe and P. Kuhl, *Annu. Rev. Neurosci.* **22**, 567 (1999).
4. A. Ghazanfar and M. D. Hauser, *Trends Cogn. Sci.* **3**, 377 (1999).
5. E. S. Spelke and E. L. Newport, in *Handbook of Child Psychology, Volume 1: Theoretical Models of Human Development*, R. M. Lerner, Ed. (Wiley, New York, 1998), pp. 275–340.
6. P. D. Eimas, E. R. Siqueland, P. W. Jusczyk, J. Vigorito, *Science* **171**, 303 (1971).
7. J. Bertoncini and J. Mehler, *Infant Behav. Dev.* **4**, 247 (1981).
8. J. Bertoncini, C. Floccia, T. Nazzi, J. Mehler, *Lang. Speech* **38**, 311 (1995).
9. R. Bijeljac-Babic, J. Bertoncini, J. Mehler, *Dev. Psychol.* **29**, 711 (1993).
10. J. Mehler *et al., Cognition* **29**, 143 (1988).
11. T. Nazzi, J. Bertoncini, J. Mehler, *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 756 (1998).
12. H. J. Neville and D. L. Mills, *Mental Retard. Dev. Disabil. Res. Rev.* **3**, 282 (1997).
13. D. Perani *et al.*, *Neuroreport* **7**, 2439 (1996).
14. G. Ehret and B. Haack, *Naturwissenshaften* **68**, 208 (1981).
15. R. A. Wyttenbach, M. L. May, R. R. Hoy, *Science* **273**, 1542 (1996).
16. P. K. Kuhl and J. D. Miller, *Science* **190**, 69 (1975).
17. P. Kuhl, in *Categorical Perception: The Groundwork of Cognition*, S. Harnad, Ed. (Cambridge University Press, Cambridge, 1987), pp. 355–386.
18. M. D. Hauser, *The Evolution of Communication* (Bradford Books/MIT Press, Cambridge, MA, 1996).
19. T. Dutoit, V. Pagel, N. Pierret, F. Bataille, O. van der Vrecken, *The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-commercial Purposes* (ICSLP'96, Philadelphia, 1996). This software is available free from tcts.fpms.ac.be/synthesis/mbrola.html.
20. Production of this kind of stimuli is described in detail in [F. Ramus and J. Mehler, *J. Acoust. Soc. Am.* **105**, 512 (1999)]. Supplementary material is available at www.sciencemag.org/feature/data/1047866.shl.
21. The experiment takes place in a sound-attenuated booth with only the baby and the experimenter inside. The experimenter is blind to the experimental condition and listens to masking noise during the test. Newborns are randomly assigned to the control or experimental group. Order of presentation of languages is counterbalanced across subjects. During a given phase, sentences corresponding to the condition are played in random order. The habituation phase lasts at least 5 min. The habituation criterion is a 25% decrease in sucking during two consecutive minutes compared with the maximum number of sucks previously produced in 1 min. Because increases in sucking rate can occur in the absence of stimulation (C. Floccia, thesis, l'Ecole des Hautes Etudes en Sciences Sociales, Paris, 1996), we compared sucking rate increase of the experimental and the control groups. This is done with a covariance analysis, comparing the average number of sucks during the 2 min after the switch (dependent variable) between the two groups and taking account of the average number of sucks during the 2 min before the switch (covariate). This analysis is detailed in [A. Christophe, E. Dupoux, J. Bertoncini, J. Mehler, *J. Acoust. Soc. Am.* **95**, 1570 (1994)] and has been used since.
22. Subjects are full-term healthy newborns, between 2 and 5 days old, recruited at the Port-Royal maternity hospital in Paris. Forty-two additional babies were tested, and their results were discarded for the following reasons: rejection of the pacifier (1), sleeping or insufficient sucking before the switch (12), crying or agitation (9), failure to meet the habituation criterion (9), sleeping or insufficient sucking after the switch (6), loss of the pacifier after the switch (4), and computer failure (1).
23. The slight tendency for babies in the control group to suck more overall than those in the experimental group, although visible in Fig. 1A, is not significant ($F_{(1,31)} = 2.6$, $P = 0.12$).
24. In the English-Japanese discrimination experiments by Nazzi *et al.* (*11*), the variability due to the four voices was much reduced by low-pass filtering the stimuli. In other experiments (*10*), a single bilingual speaker was used.
25. P. W. Jusczyk, D. B. Pisoni, J. Mullenix, *Cognition* **43**, 253 (1992).
26. Although the resynthesis process reduces all voices to one, other (prosodic) characteristics of the different speakers are preserved. We did not test discrimination of the natural sentences played backward because of failure of the newborns to discriminate these sentences even when played forward.
27. Twenty additional babies were tested, and their results were discarded for the following reasons: sleeping or insufficient sucking before the switch (6), crying or agitation (4), failure to meet the habituation criterion (1), sleeping or insufficient sucking after the switch (3), and loss of the pacifier after the switch (6).
28. Seven additional babies were tested, and their results were discarded for the following reasons: crying or agitation (1), sleeping or insufficient sucking after the switch (4), and loss of the pacifier after the switch (2).
29. The tendency for babies in the control group to suck more than those in the experimental group during the habituation phase is not significant ($F_{(1,31)} = 2.7$, $P = 0.11$).
30. Before these experiments were run, all tamarins had participated in a habituation-dishabituation experiment involving their own species-typical vocalizations. Thus, all subjects were familiar with the general test set-up. Subjects sit in the test cage within the acoustic chamber and do so without stress. They also sit or hang on the front panel during testing, thereby allowing relatively unambiguous observations of head-turning behavior. Experiments were run by transporting a tamarin from the home room to a test room, which was acoustically and visually isolated from all other tamarins. Two observers watched the session from a monitor outside the test room. Stimuli were played back from a concealed speaker only when the subject's head and body were oriented about 180° away from the speaker. A positive response was scored if the subject turned and oriented toward the speaker within the playback period. If the response was ambiguous, we ran the trial again but with a different exemplar from the habituation series. A subject was considered habituated if he or she failed to respond on two consecutive trials. After habituation, the test stimulus was played. The final trial of the session was a post-test playback, presenting the long call of a tamarin. Given the salience of the long call, we expected the tamarins to respond. If they failed to do so, we excluded the entire session, under the assumption that failure to respond to all stimuli represents general habituation to the test setup. The intertrial interval within a session was set at a minimum of 15 and a maximum of 60 s. All trials were videotaped. After a session was run, trials were digitized onto a computer and then scored blind with respect to test condition by stepping through the experiment frame-by-frame. Two observers scored each test trial; interobserver reliability was 0.92.
31. Each tamarin was tested in the four conditions, with the order of presentation of languages counterbalanced across subjects. The intersession interval was no less than 4 days, with a median of 7 days. When a session's data were excluded because of an ambiguous response to the test sentence or failure to respond to the post-test trial, the tamarin was tested again in the same condition after 1 to 3 weeks in order to complete all the conditions. The significance of the proportion of monkeys reacting to a given change was assessed with a binomial test, and the difference between two conditions was assessed with a 2×2 χ^2 test.
32. The fact that newborns fail under this very condition in experiment 1A is likely due to their immature auditory system, because susceptibility to speaker variability appears to resolve a few months after birth (*25*).
33. Supported by the Délégation Générale pour l'Armement and a NSF Young Investigator Award to M.D.H. We thank the Port-Royal maternity hospital for providing access to the newborns and we thank the parents for their participation. All parents gave informed consent. We thank the New England Regional Primate Research Center (P51RR00168-37) for providing the tamarins; the research was approved by the Animal Care and Use Committee at Harvard University (92-16) and the CCPPRB Paris-Cochin.

13 December 1999; accepted 24 February 2000