

EARLY LANGUAGE ACQUISITION: CRACKING THE SPEECH CODE

Patricia K. Kuhl

Abstract | Infants learn language with remarkable speed, but how they do it remains a mystery. New data show that infants use computational strategies to detect the statistical and prosodic patterns in language input, and that this leads to the discovery of phonemes and words. Social interaction with another human being affects speech learning in a way that resembles communicative learning in songbirds. The brain's commitment to the statistical and prosodic patterns that are experienced early in life might help to explain the long-standing puzzle of why infants are better language learners than adults. Successful learning by infants, as well as constraints on that learning, are changing theories of language acquisition.

STATISTICAL LEARNING
Acquisition of knowledge through the computation of information about the distributional frequency with which certain items occur in relation to others, or probabilistic information in sequences of stimuli, such as the odds (transitional probabilities) that one unit will follow another in a given language.

The acquisition of language and speech seems deceptively simple. Young children learn their mother tongue rapidly and effortlessly, from babbling at 6 months of age to full sentences by the age of 3 years, and follow the same developmental path regardless of culture (FIG. 1). Linguists, psychologists and neuroscientists have struggled to explain how children do this, and why it is so regular if the mechanism of acquisition depends on learning and environmental input. This puzzle, coupled with the failure of artificial intelligence approaches to build a computer that learns language, has led to the idea that speech is a deeply encrypted 'code'. Cracking the speech code is child's play for human infants but an unsolved problem for adult theorists and our machines. Why?

During the last decade there has been an explosion of information about how infants tackle this task. The new data help us to understand why computers have not cracked the human linguistic code and shed light on a long-standing debate about the origins of language in the child. Infants' strategies are surprising and are also unpredicted by the main historical theorists. Infants approach language with a set of initial perceptual abilities that are necessary for language acquisition, although not unique to humans. They then learn rapidly from exposure to language, in ways that are unique to humans, combining pattern detection and computational abilities (often called **STATISTICAL LEARNING**) with special social skills. An absence of early exposure to the

patterns that are inherent in natural language — whether spoken or signed — produces life-long changes in the ability to learn language.

Infants' perceptual and learning abilities are also highly constrained. Infants cannot perceive all physical differences in speech sounds, and are not computational slaves to learning all possible stochastic patterns in language input. Moreover, and of equal importance from a neurobiological perspective, social constraints limit the settings in which learning occurs. The fact that infants are 'primed' to learn the regularities of linguistic input when engaged in social exchanges puts language in a neurobiological framework that resembles communicative learning in other species, such as songbirds, and helps us to address why non-human animals do not advance further towards language. The constraints on infants' abilities to perceive and learn are as important to theory development as are their successes.

Recent neuropsychological and brain imaging work indicates that language acquisition involves **NEURAL COMMITMENT**. Early in development, learners commit the brain's neural networks to patterns that reflect natural language input. This idea makes empirically testable predictions about how early learning supports and constrains future learning, and holds that the basic elements of language, learned initially, are pivotal. The concept of neural commitment is linked to the issue of a 'critical' or 'sensitive' period for language acquisition.

*Institute for Learning and
Brain Sciences and the
Department of Speech and
Hearing Sciences,
University of Washington,
Seattle, Washington 98195,
USA.
e-mail:
pkkuhl@u.washington.edu
doi:10.1038/nrn1533*

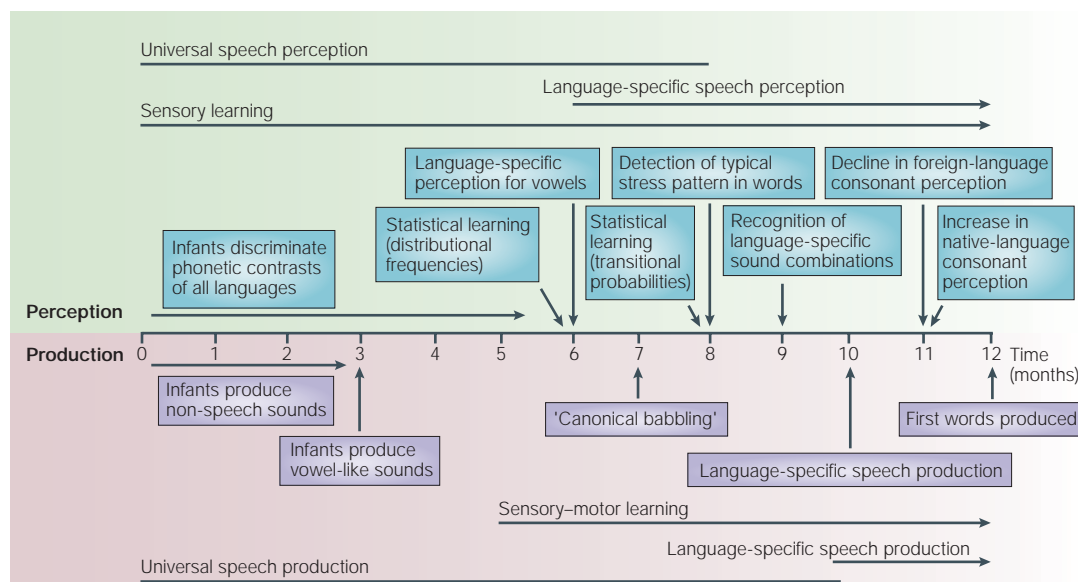


Figure 1 | **The universal language timeline of speech-perception and speech-production development.** This figure shows the changes that occur in speech perception and production in typically developing human infants during their first year of life.

NEURAL COMMITMENT

Learning results in a commitment of the brain's neural networks to the patterns of variation that describe a particular language. This learning promotes further learning of patterns that conform to those initially learned, while interfering with the learning of patterns that do not conform to those initially learned.

PHONEMES

Elements of a language that distinguish words by forming the contrasting element in pairs of words in a given language (for example, 'rake'–'lake'; 'far'–'fall'). Languages combine different phonetic units into phonemic categories; for example, Japanese combines the 'r' and 'l' units into one phonemic category.

PHONETIC UNITS

The set of specific articulatory gestures that constitute vowels and consonants in a particular language. Phonetic units are grouped into phonemic categories. For example, 'r' and 'l' are phonetic units that, in English, belong to separate phonemic categories.

CATEGORIZATION

In speech perception, the ability to group perceptually distinct sounds into the same category. Unlike computers, infants can classify as similar phonetic units spoken by different talkers, at different rates of speech and in different contexts.

The idea is that the initial coding of native-language patterns eventually interferes with the learning of new patterns (such as those of a foreign language), because they do not conform to the established 'mental filter'. So, early learning promotes future learning that conforms to and builds on the patterns already learned, but limits future learning of patterns that do not conform to those already learned.

The encryption problem

Sorting out the sounds. The world's languages contain many basic elements — around 600 consonants and 200 vowels¹. However, each language uses a unique set of only about 40 distinct elements, called **PHONEMES**, which change the meaning of a word (for example, from 'bat' to 'pat'). These phonemes are actually groups of non-identical sounds, called **PHONETIC UNITS**, that are functionally equivalent in the language. The infant's task is to make some progress in figuring out the composition of the 40 or so phonemic categories before trying to acquire words on which these elementary units depend.

Three early discoveries inform us about the nature of the innate skills that infants bring to the task of phonetic learning and about the timeline of early learning. The first, called categorical perception, focused on discrimination of the acoustic events that distinguish phonetic units (BOX 1)². Eimas and colleagues showed that young infants are especially sensitive to acoustic changes at the phonetic boundaries between categories, including those of languages they have never heard^{3–6}. Infants can discriminate among virtually all the phonetic units used in languages, whereas adults cannot⁷. The acoustic differences on which this depends are tiny. A change of 10 ms in the time domain changes /b/ to /p/, and equivalently small differences in the frequency domain change /p/ to /k/ (REF 8). Infants can discriminate these subtle differences from birth, and this ability is essential for the

acquisition of language. However, categorical perception also shows that infant perception is constrained. Infants do not discriminate all physically equal acoustic differences; they show heightened sensitivity to those that are important for language.

Although categorical perception is a building block for language, it is not unique to humans. Non-human mammals — such as chinchillas and monkeys — also partition sounds where languages place phonetic boundaries^{9–11}. In humans, non-speech sounds that mimic the acoustic properties of speech are also partitioned in this way^{12,13}. I have previously argued that the match between basic auditory perception and the acoustic boundaries that separate phonetic categories in human languages is not fortuitous: general auditory perceptual abilities provided 'basic cuts' that influenced the choice of sounds for the phonetic repertoire of the world's languages^{14,15}. The development of these languages capitalized on natural auditory discontinuities. However, the basic cuts provided by audition are primitive, and only roughly partition sounds. The exact locations of phonetic boundaries differ across languages, and exposure to a specific language sharpens infants' perception of stimuli near phonetic boundaries in that language^{16,17}. According to this argument, auditory perception, a domain-general skill, initially constrained choices at the phonetic level of language during its evolution. This ensured that, at birth, infants are prepared to discern differences between phonetic contrasts in any natural language^{14,15}.

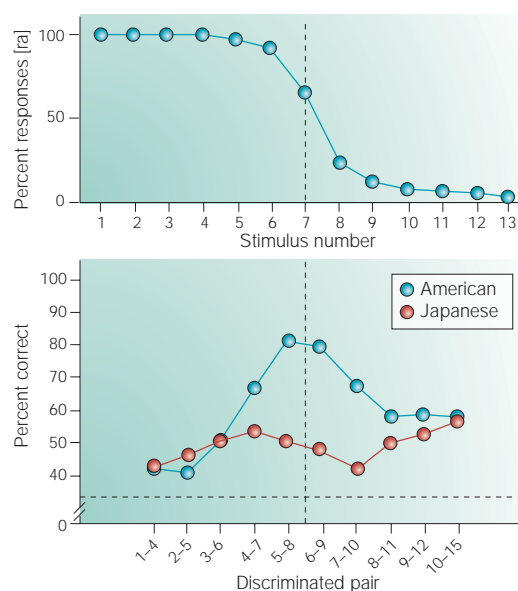
As well as discriminating the elementary sounds that are used in language, infants must learn to perceptually group different sounds that they clearly hear as distinct (BOX 2). This is the problem of **CATEGORIZATION**¹⁸. In a natural environment, infants hear sounds that vary on many dimensions (for example, talker, rate and phonetic context). At an early age, infants can categorize

Box 1 | What is categorical perception?

Categorical perception is the tendency for adult listeners of a particular language to classify the sounds used in their languages as one phoneme or another, showing no sensitivity to intermediate sounds. Laboratory demonstrations of this phenomenon involve two tasks, identification and discrimination. Listeners are asked to identify each sound from a series generated by a computer. Sounds in the series contain acoustic cues that vary in small, physically equal steps from one phonetic unit to another, for example in 13 steps from /ra/ to /la/.

In this example, both American and Japanese listeners are tested⁷. Americans distinguish the two sounds and identify them as a sequence of /ra/ syllables that changes to a sequence of /la/ syllables. Even though the acoustic step size in the series is physically equal, American listeners do not hear a change until stimulus 7 on the continuum. When Japanese listeners are tested, they do not hear any change in the stimuli. All the sounds are identified as the same — the Japanese 'r'.

When pairs of stimuli from the series are presented to listeners, and they are asked to identify the sound pairs as 'same' or 'different', the results show that Americans are most sensitive to acoustic differences at the boundary between /r/ and /l/ (dashed line). Japanese adults' discrimination values hover near chance all along the continuum. Figure modified, with permission, from REF. 7 © (1975) The Psychonomic Society.



speech sounds despite such changes^{19–23}. By contrast, computers are, so far, unable to recognize phonetic similarity in this way²⁴. This is a necessary skill if infants are to imitate speech and learn their 'mother tongue'²⁵.

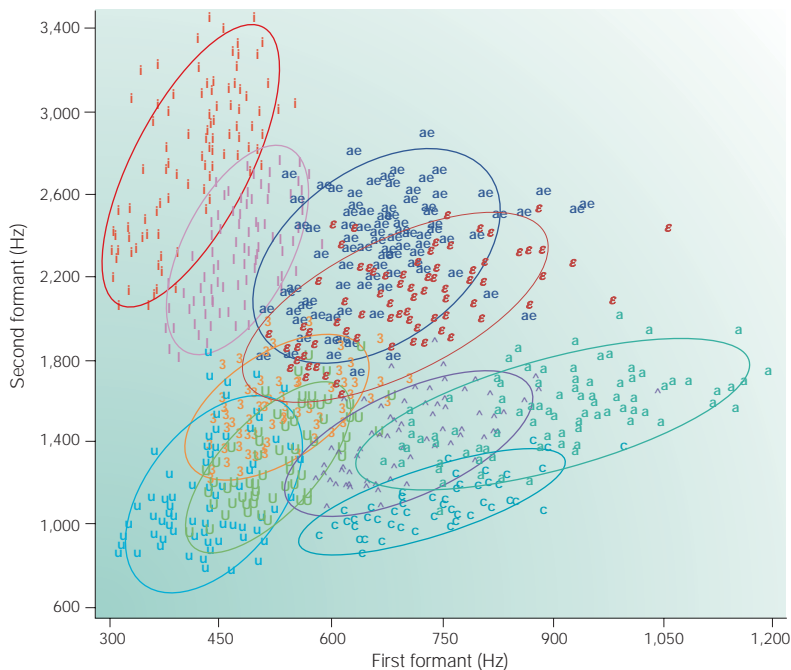
Infants' initial universal ability to distinguish between phonetic units must eventually give way to a language-specific pattern of listening. In Japanese, the phonetic units 'r' and 'l' are combined into a single phonemic category (Japanese 'r'), whereas in English, the difference is preserved ('rake' and 'lake'); similarly, in English, two Spanish phonetic units (distinguishing 'bala' from 'pala') are united in a single phonemic category. Infants can initially distinguish these sounds^{4–6}, and Werker and colleagues investigated when the infant 'citizens of the world' become 'culture-bound' listeners²⁶. They showed that English-learning infants could easily discriminate Hindi and Salish sounds at 6 months of age, but that this discrimination declined substantially by 12 months of age. English-learning infants at 12 months have difficulty in distinguishing between sounds that are not used in English^{26,27}. Japanese infants find the English r–l distinction more difficult^{28,29}, and American infants' discrimination declines for both a Spanish³⁰ and a Mandarin distinction³¹, neither of which is used in English. At the same time, the ability of infants to discriminate native-language phonetic units improves^{30,32,33}.

Computational strategies. What mechanism is responsible for the developmental change in phonetic perception between the ages of 6 and 12 months? One hypothesis is that infants analyse the statistical distributions of sounds that they hear in ambient language. Although adult listeners hear 'r' and 'l' as either distinct (English speakers) or

identical (Japanese), speakers of both languages produce highly variable sounds. Japanese adults produce both English r- and l-like sounds, so Japanese infants are exposed to both. Similarly, in Swedish there are 16 vowels, whereas English uses 10 and Japanese uses only 5 (REFS 34,35), but speakers of these languages produce a wide range of sounds³⁶. It is the distributional patterns of such sounds that differ across languages^{37,38}. When the acoustic features of speech are analysed, modal values occur where languages place phonemic categories, whereas distributional frequencies are low at the borders between categories. So, distributional patterns of sounds provide clues about the phonemic structure of a language^{39,40}. If infants are sensitive to the relative distributional frequencies of phonetic segments in the language that they hear, and respond to all instances near a modal value by grouping them, this would assist 'category learning'.

Experiments on 6-month-old infants indicate that this is the case (FIG. 2). Kuhl and colleagues⁴¹ tested 6-month-old American and Swedish infants with prototype vowel sounds from both languages (FIG. 2a). Both the American-English prototype and the Swedish prototype were synthesized by computer and, by varying the critical acoustic components in small steps, 32 variants of each prototype were created. The infants listened to the prototype vowel (either English or Swedish) presented as the background stimulus, and were trained to respond with a head-turn when they heard the prototype vowel change to one of its variants (FIG. 2b). The hypothesis was that infants would show a 'perceptual magnet effect' for native-language sounds, because prototypical sounds function like magnets for surrounding sounds⁴². The perceptual magnet effect is hypothesized to reflect prototype learning in cognitive psychology⁴³.

Box 2 | Why is speech categorization difficult?



Phonemic categories are composed of finite sets of phonetic units. Phonetic units are difficult to define physically because every utterance, even of the same phonetic unit, is acoustically distinct. Different talkers, rates of speech and contexts all contribute to the variability observed in speech.

Talker variability

When different talkers produce the same phonetic unit, such as a simple vowel, the acoustic results (FORMANT FREQUENCIES) vary widely. This is because of the variability in vocal tract size and shape, and is especially different when men, women and children produce the same phonetic unit. In the drawing, each ellipse represents an English vowel, and each symbol within the circle represents one person's production³⁵.

Rate variability

Slow speech results in different acoustic properties from faster speech, making physical descriptions of phonetic units difficult²².

Context variability

The acoustic values of a phonetic unit change depending on the preceding and following phonemes²³.

These variations make it difficult to rely on absolute acoustic values to determine the phonetic category of a particular speech sound. Despite all of these sources of variability, infants perceive phonetic similarity across talkers, rates and contexts^{19–23}. By contrast, current computer speech-recognition systems cannot recognize phonetic similarity when the talker, rate and context change²⁴. Figure reproduced, with permission, from REF. 35 © (1995) Acoustical Society of America.

The results confirmed this prediction —the infants did show a perceptual magnet effect for their native vowel category (FIG. 2c). American infants perceptually grouped the American vowel variants together, but treated the Swedish vowels as less unified. Swedish infants reversed the pattern, perceptually grouping the Swedish variants more than the American vowel stimuli. The results were assumed to reflect infants' sensitivities to the distributional properties of sounds in their language³⁹. Interestingly, monkeys did not show a prototype magnet effect for vowels⁴², indicating that the effect in humans was unique, and required linguistic experience.

FORMANT FREQUENCIES

Frequency bands in which energy is highly concentrated in speech. Formant locations for each phonetic unit are distinct and depend on vocal tract shape and tongue position. Formants are numbered from lowest frequencies to highest: F1, F2 and so on.

Infants can also learn from distributional patterns in language input after short-term exposure to phonetic stimuli (FIG. 2). Maye and colleagues⁴⁰ exposed 6- and 8-month-old infants for about 2 min to 8 sounds that formed a series (FIG. 2d). Infants were familiarized with stimuli on the entire continuum, but experienced different distributional frequencies. A 'bimodal' group heard more frequent presentations of stimuli at the ends of the continuum; a 'unimodal' group heard more frequent presentations of stimuli from the middle of the continuum. After familiarization, infants were tested using a listening preference technique (FIG. 2e). The results supported the hypothesis that infants at this age are sensitive to distributional patterns (FIG. 2f); infants in the bimodal group discriminated the two sounds, whereas those in the unimodal group did not. Further work on distributional cues shows that infants learn the PHONOTACTIC PATTERNS of language, rules that govern the sequences of phonemes that can be used to compose words. By 9 months of age, infants discriminate between phonetic sequences that occur frequently and those that occur less frequently in ambient language⁴⁴. These findings show that statistical learning involving distributional patterns in language input assists language learning at the phonetic level in infants.

Discovering words. The phonemes of English are used to create about half a million words. Reading written words that lack spaces between them gives some sense of the task that infants face in identifying spoken words (BOX 3). Without the spaces, printed words merge and reading becomes difficult. Similarly, although conversational speech provides some acoustic breaks, these do not reliably signal word boundaries. When we listen to another language, we perceive the words as run together and spoken too quickly. Without any obvious boundaries, how can an infant discover where one word ends and another begins? Field linguists have spent decades attempting to identify the words used by speakers of a specific language. Children learn implicitly. By 18 months of age, 75% of typically developing children understand about 150 words and can successfully produce 50 words⁴⁵.

Computational approaches to words. Word segmentation is also advanced by infants' computational skills. Infants are sensitive to the sequential probabilities between adjacent syllables, which differ within and across word boundaries. Consider the phrase 'pretty baby'; among English words, the probability that 'ty' will follow 'pre' is higher than the probability that 'bay' will follow 'ty'. If infants are sensitive to adjacent transitional probabilities in continuous speech, they might be able to parse speech and discover that pretty is a potential word, even before they understand its meaning.

Saffran and colleagues have shown how readily infants use sequential probabilities to detect words⁴⁶, greatly advancing an initial study that indicated that infants are sensitive to this kind of information. In the initial study⁴⁷, 8-month-old infants were presented with three-syllable strings made up of the syllables 'ko', 'ga',

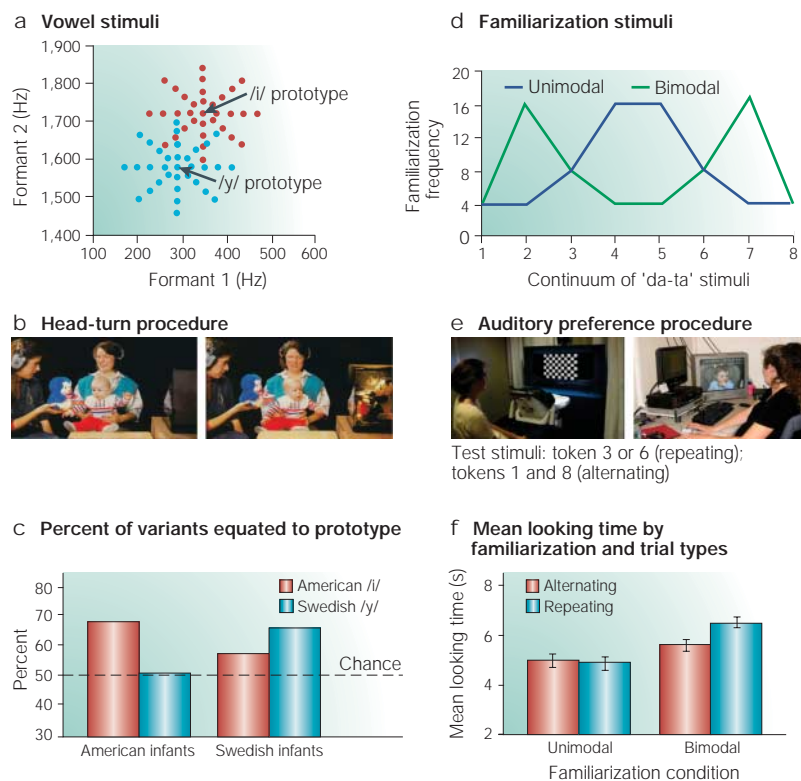


Figure 2 | Two experiments showing infant learning from exposure to the distributional patterns in language input. **a** | The graph shows differences in formant frequencies between vowel sounds representing variants of the English /i/ and the Swedish /y/ vowels used in tests on 6-month-old American and Swedish infants. **b** | Head-turn testing procedure: infants hear a repeating prototype vowel (English or Swedish) while being entertained with toys; they are trained to turn their heads away from the assistant when they hear the prototype vowel change, and are rewarded for doing so with the sight of an animated toy animal. Head-turn responses to variants indicate discrimination from the prototype. **c** | Infants perceive more variants as identical to the prototype for native-language vowel categories, indicating that linguistic experience increases the perception of similarity among members of a phonetic category⁴¹. **d** | In another study, infants are familiarized for 2 min with a series of 'da-ta' stimuli, with higher frequencies of either stimuli 2 and 7 (bimodal group) or stimuli 4 and 5 (unimodal group). **e** | Auditory preference procedure: two types of auditory stimulus, alternating 1 and 8, or repeating 3 or 6, are presented sequentially, along with visual stimuli to elicit attention. Looking time to each type of stimulus is measured; significantly different looking times indicate discrimination. **f** | Infants in the bimodal group looked for significantly longer during the repeating trials than during alternating trials, whereas infants in the unimodal condition showed no preference, indicating that only infants in the bimodal condition discriminated the 'da-ta' end-point stimuli⁴⁰. Panels **a** and **c** modified, with permission, from REF. 41 © (1992) American Association for the Advancement of Science. Panels **d** and **f** modified, with permission, from REF. 40 (2002) © Elsevier Science.

PHONOTACTIC PATTERNS

Sequential constraints, or rules, governing permissible strings of phonemes in a given language. Each language allows different sequences. For example, the combination 'zb' is not permissible in English, but is a legal combination in Polish.

and 'de' (FIG. 3a). Three groups of infants were tested, and the arrangement of the syllables 'ko' and 'ga' was manipulated across groups. For one group, they occurred in an invariant order, 'koga', with transitional probabilities of 1.0 between two syllables, as would be the case if the unit formed a word. For the second group, the order of the two syllables was variable, 'koga' and 'gako', with a transitional probability of 0.50. For the control group, one syllable was repeated twice, 'koko', consistent with a word, but not one that allowed a transitional probability strategy. The third syllable, 'de', occurred before or after the two-syllable combination. The three syllables were matched on all other acoustic cues (duration, loudness and pitch) so that parsing could not be based on some other aspect of the syllables. The infants' task was to

respond when the third syllable in the string, 'de', changed to 'ti' (FIG. 3b). Perceiving a phonetic change in a trisyllabic string is difficult for infants at this age, even though they readily discriminate the two syllables in isolation⁴⁸. The experiment tested whether infants can use transitional probabilities to 'chunk' the input, and whether doing so reduced the perceptual demands of phonetic processing in adjacent syllables. Infants in the invariant group performed significantly better than infants in the other two groups, whose performance did not differ from one another (FIG. 3c), indicating that only infants in the invariant group perceived 'koga' as a word-like unit, which made discrimination of 'de' and 'ti' significantly easier.

Saffran and colleagues⁴⁹ firmly established that 8-month-old infants can learn word-like units on the basis of transitional probabilities. They played to infants 2-minute strings of computer synthesized speech (for example, 'tibusopabikugolatudaropi') that contained no breaks, pauses, stress differences or intonation contours (FIG. 3d). The transitional probabilities were 1.0 among the syllables contained in four pseudo-words that made up the string 'tibuso', 'pabiku', 'golatu' and 'daropi', and 0.33 between other adjacent syllables. To detect the words embedded in the strings, infants had to track the statistical relations among adjacent syllables. After exposure, infants were tested for listening preferences with two of the original words and two part-words formed by combining syllables that crossed word boundaries (for example, 'tudaropi' — the last syllable of 'golatu' and the first two of 'daropi') (FIG. 3e). The infants showed an expected novelty preference, indicating that they detected the statistical regularities in the original stimuli by preferring stimuli that violated that structure (FIG. 3f). Further studies showed that this was due not to infants' calculation of the frequencies of occurrence, but rather to the probabilities specified by the sequences of sounds⁵⁰. So, 2 min of exposure to continuous syllable strings is sufficient for infants to detect word candidates, indicating a potential mechanism for word learning.

These specific statistical learning skills are not restricted to language or to humans. Infants can track adjacent transitional probabilities in tone sequences⁵¹ and in visual patterns^{52,53}, and monkeys can track adjacent dependencies in speech when Saffran's stimuli are used⁵⁴.

PROSODIC CUES also help infants to identify potential word candidates and are prominent in natural speech. About 90% of English multisyllabic words in conversational speech begin with linguistic stress on the first syllable, as in the words 'pencil' and 'stapler' (REF. 55). This strong-weak (trochaic) pattern is the opposite of that used in languages such as Polish, in which a weak-strong (iambic) pattern predominates. All languages contain words of both kinds, but one pattern typically predominates. At 7.5 months of age, English-learning infants can segment words from speech that reflect the strong-weak pattern, but not the weak-strong pattern — when such infants hear 'guitar is' they perceive 'taris' as the word-like unit, because it begins with a stressed syllable⁵⁶.

Box 3 | How do infants find words?

Unlike written language, spoken language has no reliable markers to indicate word boundaries. Acoustic analysis of speech shows that there are no obvious breaks between syllables or words in the phrase: 'There are no silences between words' (a).

Word segmentation in printed text would be equally difficult if the spaces between words were removed. The continuous string of letters below could be broken up in two different ways, as shown in b.

a Spoken: with no markers

"There are no silences between words"



ThereAre	NoS	ilen	ces	Bet	weenWord	s
----------	-----	------	-----	-----	----------	---

b Printed text: with no markers

THEREDONATEAKETTLEOFTENCHIPS

THE RED ON A TEA KETTLE OFTEN CHIPS or THERE, DON ATE A KETTLE OF TEN CHIPS

Natural speech also contains statistical cues. Johnson and Jusczyk⁵⁷ pitted prosodic cues against statistical ones by familiarizing infants with strings of syllables that provided conflicting cues. Syllables that ended words by statistical rules received word-initial stress cues (they were louder and longer, and had a higher pitch). They found that infants' strategies change with age; at 8 months, infants recover words from the strings on the basis of initial-word stress rather than statistical cues^{57,58}. By contrast, at 7 months, they use statistical rather than prosodic cues⁵⁹. How infants combine the two probabilistic cues, neither of which provides deterministic information in natural speech, will be a fruitful topic for future investigation.

How far can statistical cues take infants? Do these initial statistical strategies account for the acquisition of linguistic rules? At present, studies are focused on infants' computational limitations; if infants detect statistical regularities only in adjacent units, they would be severely limited in acquiring linguistic rules by statistical means. Non-adjacent dependencies are essential for detecting more complex relations, such as noun-verb agreement, and these specific relations are acquired only later in development^{60,61}.

Newport and colleagues⁶² have shown that adults can detect non-adjacent dependencies in the kinds of syllable strings used by Saffran when they involve segments (consonants or vowels) but not syllables. By contrast, Old World monkeys can detect non-adjacent syllable dependencies and segmental ones that involve vowels, but not consonants⁶³. Infants apparently cannot detect non-adjacent dependencies in the specific kinds of continuous strings used by Saffran⁶³.

However, there is some evidence that young children can detect non-adjacencies such as those required to learn grammar. Gomez and colleagues played artificial word strings (for example, 'vot-pel-jic-rud-tam') to 12-month-olds for 50–127 s to investigate whether they

could learn the rules that specified word order^{64,65}. Two grammars were used to generate the word strings. The grammars used the same word units and produced sequences that began and ended with the same words, but word order within the strings varied. After exposure to one of the artificial languages, infants preferred to listen to new words specifying the unfamiliar grammar, indicating that they had learned word-order rules from the grammar that they had previously experienced. The word items used during familiarization were not those used to test the infants, showing that infants can generalize their learning to a new set of words — they can learn abstract patterns that do not rely on memory for specific instances.

Similarly, Marcus showed that 7-month-olds can learn sequences of either an ABB ('de-li-li') or ABA ('we-di-we') form, and that they can extend this pattern learning to new sequences, a skill that was argued to require learning of algebraic rules^{66,67}. It has been proposed that infants compute two kinds of statistics, one arithmetic and the other algebraic^{66,67}; however, experimentally differentiating the two is difficult^{68,69}. Further tests are required to determine whether infants are learning rules or statistical regularities in these studies.

Social influences on language learning

Computational learning indicates that infants learn simply by being exposed to the right kind of auditory information — even in a few minutes of auditory exposure in the laboratory^{40,47,49}. However, natural language learning might require more, and different, kinds of information. The results of two studies — one involving speech-perception learning and the other speech-production learning — indicate that social interaction assists language learning in complex settings. In both speech production and speech perception, the presence of a human being interacting with a child has a strong influence on learning. These findings are reminiscent of the constraints observed in communication learning in songbirds^{70,71}.

The impact of social interaction on human language learning has been dramatically illustrated by the few instances in which children have been raised in social isolation; these cases have shown that social deprivation has a severe and negative impact on language development, to the extent that normal language skills are never acquired⁷². In children with autism, language and social deficits are tightly coupled — aberrant neural responses to speech are strongly correlated with an interest in listening to non-speech signals as opposed to speech signals⁷³. Speech is strongly preferred in typically developing children⁷⁴. Social deprivation, whether imposed by humans or caused by atypical brain function, has a devastating effect on language acquisition. Theories of social learning in typically developing children have traditionally emphasized the importance of social interaction on language learning^{75,76}. Recent data and theory posit that language learning is grounded in children's appreciation of others' communicative intentions, their sensitivity to joint visual attention and their desire to imitate⁷⁷.

PROSODIC CUES

Pitch, tempo, stress and intonation, qualities that are superimposed on phonemes, syllables, words and phrases. These cues convey differences in meaning (statements versus questions), word stress (trochaic versus iambic), speaking styles (infant- versus adult-directed speech) and the emotional state of a speaker (happy versus sad).

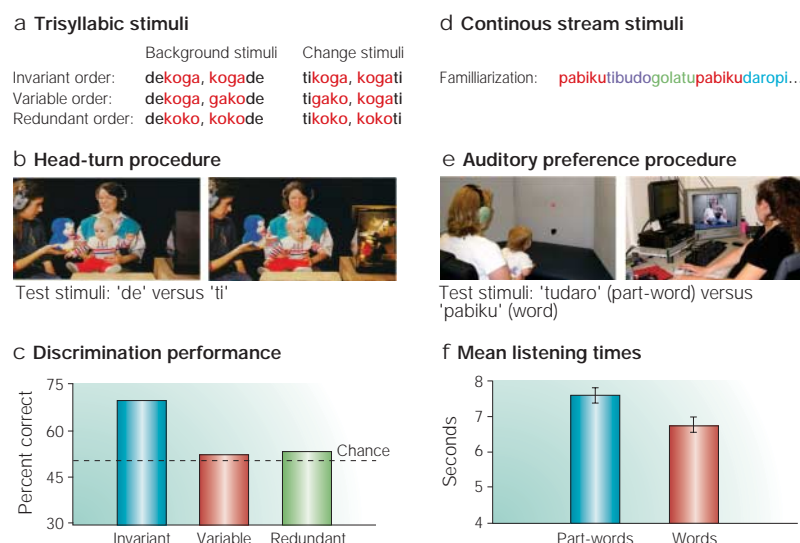


Figure 3 | Two experiments showing infant learning of word-like stimuli on the basis of transitional probabilities between adjacent syllables. **a** | Trisyllabic stimuli used to test infant learning of word-like units using transitional probabilities between the syllables 'ko' and 'ga'. In one group they occurred in an invariant order, with transitional probabilities of 1.0; in a second group they were heard in a variable order, with transitional probabilities of 0.50. A redundant order group served as a control. In all groups, the third syllable making up each word-like unit was 'de'. **b** | The head-turn testing procedure was used to test infants' detection of a change from the syllable 'de' to the syllable 'ti' in all groups. **c** | Only the invariant group performed above chance on the task, indicating that the infants in this group recognized 'koga' as a word-like unit⁴⁷. **d** | A continuous stream of syllables used to test the detection of word-like stimuli that were created from four words (different colours), the syllable transitional probabilities of which were 1.0. All other adjacent transitional probabilities were 0.33. **e** | After a 2-min familiarization period, blinking lights above the side speakers were used to attract the infant's attention. Once the infant's head turned towards the light, either a word or a part-word was played and repeated until the infant looked away, and the total amount of looking time was measured. Discrimination was indicated by significantly different looking times for words and part-words. **f** | Infants preferred new part-words, indicating that they had learned the original words⁴⁹.

A study that compared live social interaction with televised foreign-language material showed the impact of social interaction on language learning in infants³¹. The study was designed to test whether infants can learn from short-term exposure to a natural foreign language.

Nine-month-old American infants listened to four native speakers of Mandarin during 12 sessions in which the speakers read books to the infants and talked about toys that they showed to the infants (FIG. 4a). After the sessions, infants were tested with a Mandarin phonetic contrast that does not occur in English to see whether exposure to the foreign language had reversed the usual decline in infants' foreign-language speech perception (FIG. 4b). The results showed that infants learned during the live sessions, compared with a control group that heard only English (FIG. 4c)³¹.

To test whether such learning depends on live human interaction, a new group of infants saw the same Mandarin speakers on a television screen or heard them over loudspeakers (FIG. 4a). The auditory statistical cues available to the infants were identical in the televised and live settings, as was the use of 'motherese'^{78,79} (BOX 4). If simple auditory exposure to language prompts learning,

the presence of a live human being would not be essential. However, the infants' Mandarin discrimination scores after exposure to televised or audiotaped speakers were no greater than those of the control infants; both groups differed significantly from the live-exposure group (FIG. 4c). Infants are apparently not computational automatons — rather, they might need a social tutor when learning natural language.

Social influences on language learning are also seen in studies of speech production^{80–82}. Goldstein *et al.* showed that social feedback modulates the quantity and quality of utterances of young infants. In the study, mothers' responsiveness to their infants' vocalizations was manipulated (FIG. 4d). After a baseline period of normal interaction, half of the mothers were instructed to respond immediately to their infants' vocalizations by smiling, moving closer to and touching their infants: these were the 'contingent condition' (CC) mothers. The other half of the mothers were 'yoked controls' (YC) — their reactions were identical, but timed (by the experimenter's instructions) to coincide with vocalizations of infants in the CC group. Infants in the CC group produced more vocalizations than infants in the YC group, and their vocalizations were more mature and adult-like (FIG. 4e)⁸⁰.

In other species, such as songbirds, communicative learning is also enhanced by social contact. Young zebra finches need visual interaction with a tutor bird to learn song in the laboratory⁸³, and their innate preference for conspecific song can be overridden by a Bengalese finch foster father who feeds them, even when adult zebra finch males can be heard nearby⁸⁴. White crown sparrows, which reject the audiotaped songs of alien species, learn the same alien songs when they are sung by a live tutor⁸⁵. In barn owls⁸⁶ and white-crowned sparrows⁸⁵, a richer social environment extends the duration of the sensitive period for learning. Social contexts also advance song production in birds; male cowbirds respond to the social gestures and displays of females, which affect the rate, quality and retention of song elements in their repertoires⁸⁷, and white-crowned sparrow tutors provide acoustic feedback that affects the repertoires of young birds⁸⁸.

In birds, interactions can take various forms. Blindfolded zebra finches that cannot see the tutor, but can interact through pecking and grooming, learn their songs. Moreover, young birds that have been operantly conditioned to present conspecific song to themselves by pressing a key learn the songs they hear^{89,90}, indicating that active participation, attention and motivation are important⁷⁰.

In the human infant foreign-language-learning situation described earlier, a live person also provides referential cues. Speakers often focused on pictures in the books or on the toys that they were talking about, and the infant's gaze followed the speaker's gaze, which is typical for infants at this age^{91,92}. Gaze-following to an object is an important predictor of receptive vocabulary^{92,93}; perhaps joint visual attention to an object that is being named also helps infants to segment words from ongoing speech.

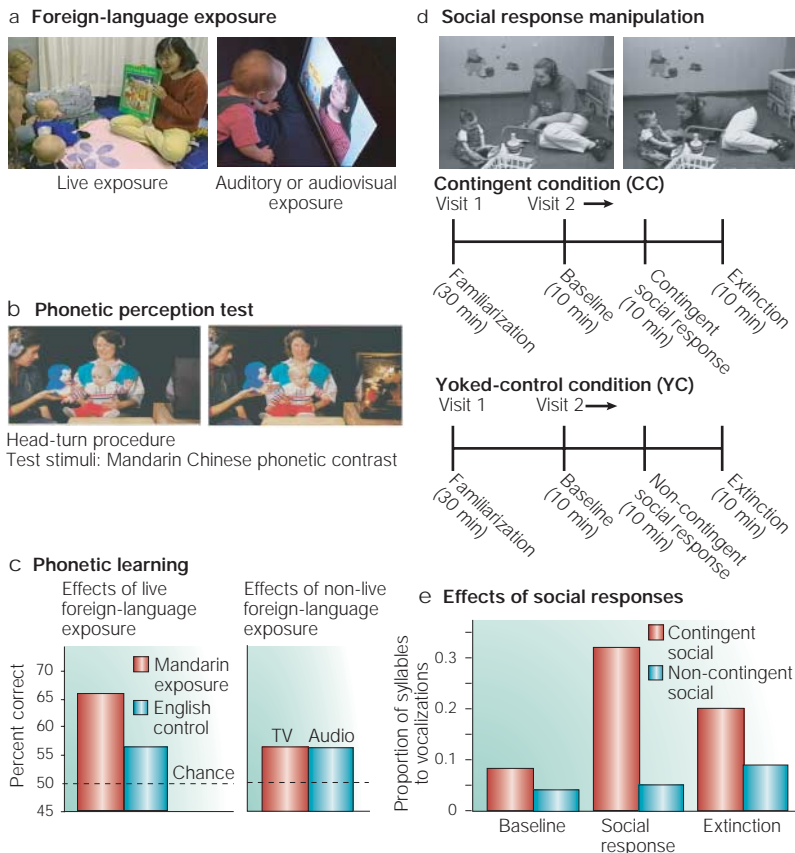


Figure 4 | Two speech experiments on social learning. **a** | Nine-month-old American infants being exposed to Mandarin Chinese in twelve 25-min live or televised sessions. **b** | After exposure, infants in the Mandarin exposure groups and those in the English control groups were tested on a Mandarin phonetic contrast using the head-turn technique. **c** | The results show phonetic learning in the live-exposure group, but no learning in the TV- or audio-only groups³¹. **d** | Eight-month-old infants received either contingent or non-contingent social feedback from their mothers in response to their vocalizations. **e** | Contingent social feedback increased the quantity and complexity of infants' vocalizations⁸⁰. Panel **c** modified, with permission, from REF. 31 © (2003) National Academy of Sciences USA. Panels **d** and **e** modified, with permission, from REF. 80 © (2003) National Academy of Sciences USA.

For both infants and birds, it is unclear whether social interaction itself, or the attention and contingency that typically accompany social interaction, are crucial for learning. However, contingency has been shown to be an important component in human vocalization learning^{81,82}, and reciprocity in adult–infant language can be seen in infants' tendency to alternate their vocalizations with those of an adult^{94,95}. The pervasive use of motherese (BOX 4) by adults is a social response that adjusts to the needs of infant listeners^{96,97}. For infants, early social awareness is a predictor of later language skills⁹².

Social interaction can be conceived of as gating computational learning, and thereby protecting infants from meaningless calculations⁷¹. The need for social interaction would ensure that learning focuses on speech that derives from humans in the child's environment, rather than on signals from other sources^{70,98,99}. Social interaction might also be important for learning sign language; both deaf and hearing babies who experience a natural sign language babble using their hands on the same schedule that hearing babies babble using their

mouths¹⁰⁰. Constraints are evident when infants hear or see non-human actions: infants imitate vocalizations rather than sine-wave analogues of speech¹⁰¹, and infer and reproduce intended actions displayed by humans but not by machines¹⁰².

Social factors might affect language acquisition because language evolved to address a need for social communication. There are connections between social awareness and other higher cognitive functions^{103,104}, and evolution might have forged connections between language and the social brain.

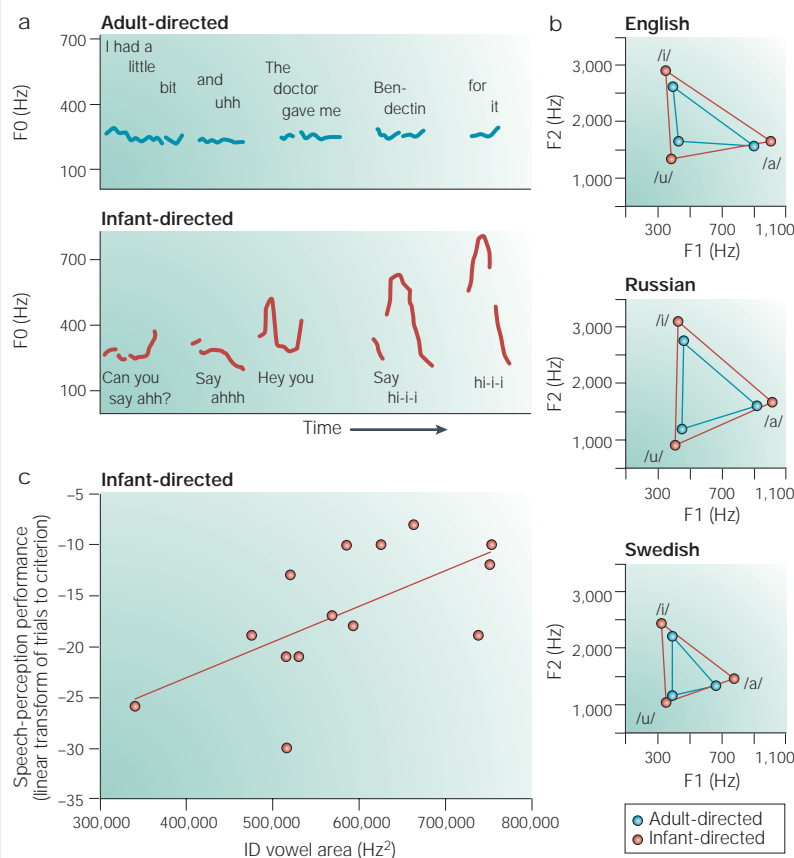
The mechanism that controls the interface between language and social cognition remains a mystery. The effects of social environments might be broad, general and 'top-down', and might engage special memory systems^{105,106}. People engaged in social interaction are highly aroused and attentive — general arousal mechanisms might enhance our ability to learn and remember, as well as prompting our most sophisticated language output. These effects could be mediated by hormones, which have been implicated in learning and song production in birds^{107,108}. On the other hand, learning might also involve more specific, 'bottom-up' mechanisms attuned to the particular form and information content of social cues (such as eye gaze). Further studies are needed to understand how the social brain supports language learning.

Native language neural commitment

A growing number of studies have confirmed the effects of language experience on the brain^{109–117}. The techniques used in these studies have recently been applied to infants and young children^{30,32,61,118–121}. For example, Dehaene-Lambertz and colleagues used functional MRI to measure the brain activity evoked by normal speech and speech played backwards in 3-month-old infants, and found that similar brain regions are active in adults and infants when listening to normal speech but that there are differences between adults' and infants' responses to backwards speech¹¹⁹. Pena and colleagues studied newborn infants' reactions to normal and backwards speech using optical topography, and showed greater left-hemisphere reaction when processing normal speech¹²⁰.

At present, studies tell us less about why our ability to acquire languages changes over time. One hypothesis, native language neural commitment (NLNC), makes specific predictions that relate early linguistic experience to future language learning¹²². According to NLNC, language learning produces dedicated neural networks that code the patterns of native-language speech. The hypothesis focuses on the aspects of language learned early — the statistical and prosodic regularities in language input that lead to phonetic and word learning — and how they influence the brain's future ability to learn language. According to the theory, neural commitment to the statistical and prosodic regularities of one's native language promotes the future use of these learned patterns in higher-order native-language computations. At the same time, NLNC interferes with the processing of foreign-language patterns that do not conform to those already learned.

Box 4 | What is 'motherese'?



When we talk to infants and children, we use a special speech 'register' that has a unique acoustic signature, called 'motherese'. Caretakers in most cultures use it when addressing infants and children. When compared to adult-directed speech, infant-directed speech is slower, has a higher average pitch and contains exaggerated pitch contours, as shown in the comparison between the pitch contours contained in adult-directed (AD) versus infant-directed (ID) speech (a)⁷⁸.

Infant-directed speech might assist infants in learning speech sounds. Women speaking English, Russian or Swedish were recorded while they spoke to another adult or to their young infants⁷⁹. Acoustic analyses showed that the vowel sounds (the /i/ in 'see', the /a/ in 'saw' and the /u/ in 'Sue') in infant-directed speech were more clearly articulated (b). Women from all three countries exaggerated the acoustic components of vowels (see the 'stretching' of the formant frequencies, creating a larger triangle for infant-directed, as opposed to adult-directed, speech). This acoustic stretching makes the vowels contained in motherese more distinct.

Infants might benefit from the exaggeration of the sounds in motherese (c). The sizes of a mother's vowel triangles, which reflect how clearly she speaks, are related to her infant's skill in distinguishing the phonetic units of speech⁹⁶. Mothers who stretch the vowels to a greater degree have infants who are better able to hear the subtle distinctions in speech. Panel a modified, with permission, from REF. 78 © (1987) Elsevier Science; panel b modified, with permission, from REF. 79 © (1997) American Association for the Advancement of Science; panel c modified, with permission, from REF. 96 © (2003) Blackwell Scientific Publishing.

Evidence for the effects of NLNC in adults comes from magnetoencephalography (MEG): when processing foreign-language speech sounds, a larger area of the adult brain is activated for a longer time period than when processing native-language sounds, indicating neural inefficiency¹¹¹. This neural inefficiency for foreign-language information extends beyond

speech — processing mathematical knowledge in a second language is also difficult¹²³. In both cases, native-language strategies can interfere with information processing in a foreign language¹²⁴.

Regarding infants, the NLNC hypothesis predicts that an infant's early skill in native-language phonetic perception should predict that child's later success at language acquisition. This is because phonetic perception promotes the detection of phonotactic patterns, which advance word segmentation^{44,125,126}, and, once infants begin to associate words with objects — a task that challenges phonetic perception^{127,128} — those infants who have better phonetic perception would be expected to advance faster. In other words, advanced phonetic abilities in infancy should 'bootstrap'¹²⁹ language learning, propelling infants to more sophisticated levels earlier in development. Behavioural studies support this hypothesis. Speech-discrimination skill in 6-month-old infants predicted their language scores (words understood, words produced and phrases understood) at 13, 16 and 24 months¹³⁰.

Neural measures provide a sensitive measure of individual differences in speech perception. Event-related potentials (ERPs) have been used in infants and toddlers to measure neural responses to phonemes, words and sentences^{30,61,121,131}. Rivera-Gaxiola and colleagues recorded ERPs in typically developing 7- and 11-month-old infants in response to native and non-native speech sounds, and found two types of neural responder³⁰. One group responded to both contrasts with positive-going brainwave changes ('P' responders), whereas the second group responded to both contrasts with negative-going brainwave changes ('N' responders) (BOX 5). Both groups could neurally discriminate the foreign-language sound at 11 months of age, whereas total group analyses had obscured this result³⁰.

In my laboratory, we use behavioural and ERP measures to take NLNC one step further. If early learning in infants causes neural commitment to native-language patterns, then foreign-language phonetic perception in infants who have never experienced a foreign language should reflect the degree to which the brain remains 'open' or uncommitted to native-language speech patterns. The degree to which an infant remains open to foreign-language speech (in the absence of exposure to a foreign language) should therefore signal slower language learning. As an open system reflects uncommitted circuitry, skill at discriminating foreign-language phonetic units should provide an indirect measure of the brain's degree of commitment to native-language patterns.

Ongoing laboratory studies support this hypothesis. In one study, 7-month-old infants from monolingual homes were tested on both native and foreign-language contrasts using behavioral and ERP brain measures¹³². As predicted, excellent native-language speech perception, measured with behavioural or brain measures, correlated positively with later language skills, whereas better foreign-language speech perception skills correlated negatively with later language skills.

Box 5 | What can brain measures reveal about speech discrimination in infants?

Continuous brain activity during speech processing can be monitored in infants by recording the electrical activity of groups of neurons using electrodes placed on the scalp. Event-related potentials (ERPs) are small voltage fluctuations that result from evoked neural activity. ERPs reflect, with high temporal resolution, the patterns of neuronal activity evoked by a stimulus. It is a non-invasive procedure that can be applied to infants with no risks. During the procedure, infants listen to a series of sounds: one is repeated many times (the standard) and a second one (the deviant) is presented on 15% of the trials. Responses are recorded to each stimulus.

Using a longitudinal design, Rivera-Gaxiola and colleagues³⁰ recorded the electrophysiological responses of 7- and 11-month-old American infants to native and non-native consonant contrasts. As a group, infants' discriminatory ERP responses to the non-native contrast are present at 7 months of age, but disappear by 11 months of age, consistent with behavioural data.

However, when the same infants were divided into subgroups on the basis of individual ERP components, there was evidence that the infant brain remains sensitive to the non-native contrast at 11 months of age, showing discriminatory positivities at 150–250 ms (P responders) or discriminatory negativities at 250–550 ms (N responders). Infants in both sub-groups increased their responsiveness to the native-language consonant contrast by 11 months of age.



Foreign phonetic test:

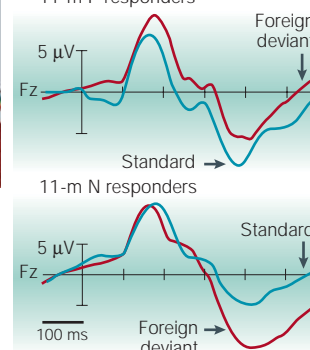
'ta-ta-ta-DA' (Spanish)

English listeners hear the Spanish syllable 'ta' as 'da'

Native contrast:

'da-da-da-T^HA' (English)

Responses to foreign contrast at 11 months of age



These results indicate that infants who remain open to all linguistic possibilities — retaining the innate state in which all phonetic differences are partitioned — do not progress as quickly towards language. To learn language, the innate state must be altered by input and give way to NLNC.

Neural commitment could be important in a 'critical period' or 'sensitive period'¹³³ for language acquisition¹³⁴. Maturation factors are a powerful predictor of the acquisition of first and second languages^{135–140}. For example, deaf children born to hearing parents whose first exposure to sign language occurs after the age of 6 show a life-long attenuation in ability to learn language¹⁴¹. Why is age so crucial? According to NLNC, exposure to spoken or signed language instigates a mapping process for which infants are neurally prepared¹⁴², and during which the brain's networks commit themselves to the basic statistical and prosodic features of the native language. These patterns allow phonetic and word learning. Infants who excel at detecting the patterns in natural language move more quickly towards complex language structures. Simply experiencing a language early in development, without producing it themselves, can have lasting effects on infants' ability to learn that language as an adult^{105,143,144} (but see REF. 145). By contrast, when language input is substantially delayed, native-like skills are never achieved¹⁴¹.

If experience is an important driver of the sensitive period, as NLNC indicates, why do we not learn new languages as easily at 50 as at 5? What mechanism or process governs the decline in sensitivity with age? A statistical process could govern the eventual closing of the sensitive period. If infants represent the distribution of a particular vowel in language input, and are sensitive to the degree of variability in that distribution, the closing

of the sensitive period would be cued by the stability of infants' phonetic distributions. In early childhood, caretakers' pronunciations would be overly represented in a child's distribution of vowels. As experience with more speakers occurred, the distribution would change to reflect further variability. With continued experience, the distribution would begin to stabilize. Given the variability in speech (BOX 2), this might require substantial listening experience. According to the hypothesis, when the 'ah' vowels of new speakers no longer cause a change in the underlying distribution, the sensitive period for phonetic learning would begin to close, and learning would decline. There are probably several sensitive periods for various aspects of language, but similar principles could apply.

In bilingual children, who hear two languages with distinct statistical and prosodic properties, NLNC predicts that the stabilization process would take longer, and studies are under way to test this hypothesis. Bilingual children are mapping two distinct systems, with some portion of the input they hear devoted to each language. At an early age neither language is statistically stable, and neither is therefore likely to interfere with the other, so young children can acquire two languages easily.

NLNC provides a mechanism that contributes to our understanding of the sensitive period. It does not deny the existence of a sensitive period; rather, it explains the fact that second language learning abilities decline precipitously as language acquisition proceeds. NLNC might also explain why the degree of difficulty in learning a second language varies depending on the relationship between the first and second language¹⁴⁶; according to NLNC, it should depend on the overlap between the statistical and prosodic features of the two languages.

Computation with constraints

In the first year of life, infants show avid learning of the information in language input. Infants learn implicitly when they hear complex, everyday language spoken in their homes, as well as in laboratories. By the age of 6 months, infants' experiences with the distributional patterns in ambient language have altered their perception of the phonetic units of speech. At 8 months, the sensitivity of infants to the statistical cues in speech allows them to segment words. At 9 months, infants can learn from exposure to a foreign language in natural infant-directed conversations, but not if the information is presented through a television or audiotape. Infants' successes in these experiments support specific hypotheses about mechanisms that explain how infant learning occurs. They represent substantial advances in our understanding of language acquisition.

At the same time, the data indicate that there are perceptual, computational, social and neural constraints. Infants do not distinguish all possible physical differences in speech sounds — only those that underlie phonetic distinctions^{3–6}. In word learning, infants compute transitional probabilities that assist them in identifying potential words, but computational constraints are also shown^{46,60,63,147,148}. Moreover, when learning natural language, constraints are seen in the potential need for a social, interactive human being^{31,80}. Finally, learning produces a neural commitment to the patterns of an individual's native language, and this constrains future success at acquiring a new language.

The origins of these constraints on infants' acquisition of language are of interest to theorists. Tests on non-human species and in domains outside the field of language have led to the view that aspects of language

might have evolved to match a set of domain-general perceptual and learning abilities^{14,15,62,63,122,148,149}. Further research will continue to explore which aspects of infants' language-processing skills are unique to humans and which reflect domain-general as opposed to language-specific skills. Current research highlights the possibility that language evolved to meet the needs of young human beings, and in meeting their perceptual, computational, social and neural abilities, produced a species-specific communication system that can be acquired by all typically developing humans.

Concluding remarks

Substantial progress has been made in understanding the initial phases of language acquisition. At all levels, language learning is constrained — perceptual, computational, social and neural constraints affect what can be learned, and when. Learning results in native language neural commitment (NLNC). According to this model, computers and animals, while capable of some of the skills demanded by language, do not duplicate the set of human perceptual, computational, social and neural constraints that languages exploit. Language was designed to enable people to communicate using a code that they would learn once and hold onto for a lifetime. The rules by which infants perceive information, the ways in which they learn words, the social contexts in which language is communicated and the need to remember the learned entities for a long time probably influenced the evolution of language. Identifying constraints on infant learning, from all sources, and determining whether those constraints reflect innate knowledge that is specific to language, or are more domain-general, will be a continuing focus in the next decade.

- Ladefoged, P. *Vowels and Consonants: An Introduction to the Sounds of Language* 2nd edn (Blackwell, Oxford, UK, 2004).
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. Perception of the speech code. *Psychol. Rev.* **74**, 431–461 (1967).
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. & Vigorito, J. Speech perception in infants. *Science* **171**, 303–306 (1971).
- Lasky, R. E., Syrdal-Lasky, A. & Klein, R. E. VOT discrimination by four to six and a half month old infants from Spanish environments. *J. Exp. Child Psychol.* **20**, 215–225 (1975).
- Eimas, P. D. Auditory and phonetic coding of the cues for speech: discrimination of the /r-l/ distinction by young infants. *Percept. Psychophys.* **18**, 341–347 (1975).
- Werker, J. F. & Lalonde, C. Cross-language speech perception: initial capabilities and developmental change. *Dev. Psychol.* **24**, 672–683 (1988).
- Miyawaki, K. et al. An effect of linguistic experience: the discrimination of /r/ and /l/ by native speakers of Japanese and English. *Percept. Psychophys.* **18**, 331–340 (1975).
- Stevens, K. N. *Acoustic Phonetics* (MIT Press, Cambridge, Massachusetts, 2000).
- Kuhl, P. K. & Miller, J. D. Speech perception by the chinchilla: voice-voiceless distinction in alveolar plosive consonants. *Science* **90**, 69–72 (1975).
- Kuhl, P. K. & Miller, J. D. Speech perception by the chinchilla: identification functions for synthetic VOT stimuli. *J. Acoust. Soc. Am.* **63**, 905–917 (1978).
- Kuhl, P. K. & Padden, D. M. Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *J. Acoust. Soc. Am.* **73**, 1003–1010 (1983).
- Pisoni, D. B. Identification and discrimination of the relative onset time of two component tones: implications for voicing perception in stops. *J. Acoust. Soc. Am.* **61**, 1352–1361 (1977).
- Jusczyk, A. M., Pisoni, D. B., Walley, A. & Murray, J. Discrimination of relative onset time of two-component tones by infants. *J. Acoust. Soc. Am.* **67**, 262–270 (1980).
- Kuhl, P. K. Theoretical contributions of tests on animals to the special-mechanisms debate in speech. *Exp. Biol.* **45**, 233–265 (1986).
- Kuhl, P. K. in *Plasticity of Development* (eds Brauth, S. E., Hall, W. S. & Doelling, R. J.) 73–106 (MIT Press, Cambridge, Massachusetts, 1991).
- Aslin, R. N. & Pisoni, D. B. in *Child Phonology: Perception and Production* (eds Yeni-Komshian, G., Kavanagh, J. & Ferguson, C.) 67–96 (Academic, New York, 1980).
- Burnham, D. Developmental loss of speech perception: exposure to and experience with a first language. *Appl. Psycholinguist.* **7**, 207–240 (1986).
- Kuhl, P. K. in *Neonate Cognition: Beyond the Blooming Buzzing Confusion* (eds Mehler, J. & Fox, R.) 231–262 (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1985).
- Hillenbrand, J. Speech perception by infants: categorization based on nasal consonant place of articulation. *J. Acoust. Soc. Am.* **75**, 1613–1622 (1984).
- Kuhl, P. K. Speech perception in early infancy: perceptual constancy for spectrally dissimilar vowel categories. *J. Acoust. Soc. Am.* **66**, 1668–1679 (1979).
- Kuhl, P. K. Perception of auditory equivalence classes for speech in early infancy. *Infant Behav. Dev.* **6**, 263–285 (1983).
- Miller, J. L. & Liberman, A. M. Some effects of later-occurring information on the perception of stop consonant and semivowel. *Percept. Psychophys.* **25**, 457–465 (1979).
- Eimas, P. D. & Miller, J. L. Contextual effects in infant speech perception. *Science* **209**, 1140–1141 (1980).
- Zue, V. & Glass, J. Conversational interfaces: advances and challenges. *Proc. IEEE* **88**, 1166–1180 (2000).
- Kuhl, P. K. & Meltzoff, A. Infant vocalizations in response to speech: vocal imitation and developmental change. *J. Acoust. Soc. Am.* **100**, 2425–2438 (1996).
- Vocalizations of infants watching a video of a female talker were recorded at 12, 16 and 20 weeks of age. The results show developmental change between 12 and 20 weeks of age and also provide evidence of vocal imitation in infants by 20 weeks of age.**
- Werker, J. F. & Tees, R. C. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* **7**, 49–63 (1984).
- Best, C. & McRoberts, G. W. Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Lang. Speech* **46**, 183–216 (2003).
- Tsushima, T. et al. *Proceedings of the International Conference on Spoken Language Processing* Vol. S28F-1, 1695–1698 (Yokohama, Japan, 1994).
- Kuhl, P. K., Tsao, F. M., Liu, H. M., Zhang, Y. & de Boer, B. in *The Convergence of Natural and Human Science* (eds Domasio, A. et al.) 136–174 (The New York Academy of Science, New York, 2001).
- Rivera-Gaxiola, M., Silva-Pereyra, J. & Kuhl, P. K. Brain potentials to native- and non-native speech contrasts in seven- and eleven-month-old American infants. *Dev. Sci.* (in the press).
- An ERP study showing that at 11 months, the infant brain remains sensitive to non-native-language contrasts. Infants' responsiveness to native-language consonant contrasts also increased over time.**
- Kuhl, P. K., Tsao, F.-M. & Liu, H.-M. Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl Acad. Sci. USA* **100**, 9096–9101 (2003).
- Two studies showing that learning can occur with only short-term exposure to a language in infants, and that it is enhanced by social interaction.**

32. Cheour, M. *et al.* Development of language-specific phoneme representations in the infant brain. *Nature Neurosci.* **1**, 351–353 (1998).
33. Kuhl, P. K., Tsao, F. M., Liu, H. M., Zhang, Y. & De Boer, B. Language/culture/mind/brain. Progress at the margins between disciplines. *Ann. NY Acad. Sci.* **935**, 136–174 (2001).
34. Fant, G. *Speech Sounds and Features* (MIT Press, Cambridge, Massachusetts, 1973).
35. Hillenbrand, J., Getty, L., Clark, M. & Wheeler, K. Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.* **97**, 3099–3111 (1995).
36. Perkell, J. & Klatt, D. *Invariance and Variability in Speech Processes* 1–604 (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1986).
37. Lacerda, F. The perceptual magnet effect: an emergent consequence of exemplar-based phonetic memory. *Proc. Int. Congr. Phonetic Sci.* **2**, 140–147 (1995).
38. Lisker, L. & Abramson, A. S. A cross-language study of voicing in initial stops: acoustical measurements. *Word* **20**, 384–422 (1964).
39. Kuhl, P. K. Early linguistic experience and phonetic perception: implications for theories of developmental speech perception. *J. Phonetics* **21**, 125–139 (1993).
40. Maye, J., Werker, J. F. & Gerken, L. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* **82**, B101–B111 (2002).
41. Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* **255**, 606–608 (1992).
42. Kuhl, P. K. Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Percept. Psychophys.* **50**, 93–107 (1991).
43. Rosch, E. Cognitive reference points. *Cognit. Psychol.* **7**, 532–547 (1975).
44. Jusczyk, P., Luce, P. & Charles-Luce, J. Infants' sensitivity to phonotactic patterns in the native language. *J. Mem. Lang.* **33**, 630–645 (1994).
- This study found that 9-month-old infants, but not 6-month-old infants, prefer frequently occurring phonetic patterns in monosyllables.**
45. Fenson, L. *et al.* *MacArthur Communicative Development Inventories: User's Guide and Technical Manual* (Singular Publishing Group, San Diego, California, 1993).
46. Saffran, J. R. Constraints on statistical language learning. *J. Mem. Lang.* **47**, 172–196 (2002).
- This study shows that learners can use the predictive relationships that link elements within phrases to acquire phrase structure. Predictive relationships improved learning for sequentially presented auditory stimuli, and for simultaneously presented visual stimuli, but not for sequentially presented visual stimuli.**
47. Goodsitt, J. V., Morgan, J. L. & Kuhl, P. K. Perceptual strategies in prelingual speech segmentation. *J. Child Lang.* **20**, 229–252 (1993).
48. Karzon, R. Discrimination of polysyllabic sequences by one- to four-month-old infants. *J. Exp. Child Psychol.* **39**, 326–342 (1985).
49. Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month old infants. *Science* **274**, 1926–1928 (1996).
50. Aslin, R. N., Saffran, J. R. & Newport, E. L. Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* **9**, 321–324 (1998).
51. Saffran, J. R., Johnson, E. K., Aslin, R. N. & Newport, E. L. Statistical learning of tone sequences by human infants and adults. *Cognition* **70**, 27–52 (1999).
52. Fiser, J. & Aslin, R. N. Statistical learning of new visual feature combinations by infants. *Proc. Natl Acad. Sci. USA* **99**, 15822–15826 (2002).
53. Kirkham, N. Z., Slemmer, J. A. & Johnson, S. P. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* **83**, B35–B42 (2002).
- This study provides evidence that infants' statistical learning from auditory input can be generalized to the visual domain.**
54. Hauser, M. D., Newport, E. L. & Aslin, R. N. Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* **78**, B53–B64 (2001).
55. Cutler, A. & Carter, D. The predominance of strong initial syllables in the English vocabulary. *Comput. Speech Lang.* **2**, 133–142 (1987).
56. Jusczyk, P. W., Houston, D. M. & Newsome, M. The beginnings of word segmentation in English-learning infants. *Cognit. Psychol.* **39**, 159–207 (1999).
57. Johnson, E. K. & Jusczyk, P. W. Word segmentation by 8-month-olds: when speech cues count more than statistics. *J. Mem. Lang.* **44**, 548–567 (2001).
- The authors showed that, when multiple cues are available, 8-month-olds weighed prosodic cues more heavily than statistical cues.**
58. Saffran, J. R. & Thiessen, E. D. Pattern induction by infant language learners. *Dev. Psychol.* **39**, 484–494 (2003).
- A study of how infants segment words according to stress patterns. Nine-month-old infants learned to segment speech using the iambic pattern whether the exposure consisted of 100% or 80% iambic words. Seven-month-olds could alter their segmentation strategies when the distribution of stress cues in words was altered.**
59. Thiessen, E. D. & Saffran, J. R. Learning to learn: infants' acquisition of stress-based strategies for word segmentation. *J. Mem. Lang.* (under revision).
60. Santelmann, L. M. & Jusczyk, P. W. Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition* **69**, 105–134 (1998).
61. Silva-Pereyra, J., Rivera-Gaxiola, M. & Kuhl, P. K. An event-related brain potential study of sentence comprehension in preschoolers: semantic and morphosyntactic processing. *Cognit. Brain Res.* (in the press).
62. Newport, E. L. & Aslin, R. N. Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognit. Psychol.* **48**, 127–162 (2004).
63. Newport, E. L., Hauser, M. D., Spaepen, G. & Aslin, R. N. Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognit. Psychol.* **49**, 85–117 (2004).
64. Gomez, R. L. & Gerken, L. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition* **70**, 109–135 (1999).
65. Gomez, R. L. & Gerken, L. Infant artificial language learning and language acquisition. *Trends Cogn. Sci.* **4**, 178–186 (2000).
66. Pena, M., Bonatti, L. L., Nespor, M. & Mehler, J. Signal-driven computations in speech processing. *Science* **298**, 604–607 (2002).
67. Marcus, G. F., Vijayan, S., Bandi Rao, S. & Vishton, P. M. Rule learning by seven-month-old infants. *Science* **283**, 77–80 (1999).
68. Seidenberg, M. S., MacDonald, M. C. & Saffran, J. R. Does grammar start where statistics stop? *Science* **298**, 553–554 (2002).
69. Seidenberg, M. S. & Elman, J. Do infants learn grammar with algebra or statistics? *Science* **284**, 433 (1999).
70. Doupe, A. J. & Kuhl, P. K. Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631 (1999).
71. Kuhl, P. K. Human speech and birdsong: communication and the social brain. *Proc. Natl Acad. Sci. USA* **100**, 9645–9646 (2003).
72. Fromkin, V., Krashen, S., Curtiss, S., Rigler, D. & Rigler, M. The development of language in Genie: a case of language acquisition beyond the 'critical period'. *Brain Lang.* **1**, 81–107 (1974).
73. Kuhl, P. K., Coffey-Corina, S., Padden, D. M. & Dawson, G. Links between social and linguistic processing of speech in preschool children with autism: behavioral and physiological measures. *Dev. Sci.* **7**, 19–30 (2004).
74. Vouloumanos, A. & Werker, J. F. Tuned to the signal: the privileged status of speech for young infants. *Dev. Sci.* **7**, 270–276 (2004).
- The authors investigated differences in 2- to 7-month old infants' perception of nonsense speech sounds and structurally similar non-speech analogues. They found a bias for speech sounds in infants as young as 2-months old.**
75. Bruner, J. *Child's Talk: Learning to Use Language* (W. W. Norton, New York, 1983).
76. Vigotsky, L. S. *Thought and Language: A Usage-Based Theory of Language Acquisition* (MIT Press, Cambridge, Massachusetts, 1962).
77. Tomasello, M. *Constructing a Language* (Harvard Univ. Press, Cambridge, Massachusetts, 2003).
78. Fernald, A. & Kuhl, P. Acoustic determinants of infant preference for motherese speech. *Infant Behav. Dev.* **10**, 279–293 (1987).
79. Kuhl, P. K. *et al.* Cross-language analysis of phonetic units in language addressed to infants. *Science* **277**, 684–686 (1997).
80. Goldstein, M., King, A. & West, M. Social interaction shapes babbling: testing parallels between birdsong and speech. *Proc. Natl Acad. Sci. USA* **100**, 8030–8035 (2003).
81. Bloom, K. Social elicitation of infant vocal behavior. *J. Exp. Child Psychol.* **20**, 51–58 (1975).
82. Bloom, K. & Esposito, A. Social conditioning and its proper control procedures. *J. Exp. Child Psychol.* **19**, 209–222 (1975).
83. Eales, L. The influences of visual and vocal interaction on song learning in zebra finches. *Anim. Behav.* **37**, 507–508 (1989).
84. Immelmann, K. in *Bird Vocalizations* (ed. Hinde, R.) 61–74 (Cambridge Univ. Press, London, 1969).
85. Baptista, L. F. & Petrinovich, L. Song development in the white-crowned sparrow: social factors and sex differences. *Anim. Behav.* **34**, 1359–1371 (1986).
86. Brainard, M. S. & Knudsen, E. I. Sensitive periods for visual calibration of the auditory space map in the barn owl optic tectum. *J. Neurosci.* **18**, 3929–3942 (1998).
87. West, M. & King, A. Female visual displays affect the development of male song in the cowbird. *Nature* **334**, 244–246 (1988).
88. Nelson, D. & Marler, P. Selection-based learning in bird song development. *Proc. Natl Acad. Sci. USA* **91**, 10498–10501 (1994).
89. Adret, P. Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Anim. Behav.* **46**, 149–159 (1993).
90. Tchernichovski, O., Mitra, P., Lints, T. & Nottebohm, F. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* **291**, 2564–2569 (2001).
91. Brooks, R. & Meltzoff, A. N. The importance of eyes: how infants interpret adult looking behavior. *Dev. Psychol.* **38**, 958–966 (2002).
92. Baldwin, D. A. in *Joint Attention: Its Origins and Role in Development* (eds Moore, C. & Dunham, P. J.) 131–158 (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1995).
93. Mundy, P. & Gomes, A. Individual differences in joint attention skill development in the second year. *Infant Behav. Dev.* **21**, 469–482 (1998).
94. Kuhl, P. K. & Meltzoff, A. N. The bimodal perception of speech in infancy. *Science* **218**, 1138–1141 (1982).
95. Bloom, K., Russell, A. & Wassenberg, K. Turn taking affects the quality of infant vocalizations. *J. Child Lang.* **14**, 211–227 (1987).
96. Liu, H.-M., Kuhl, P. K. & Tsao, F.-M. An association between mothers' speech clarity and infants' speech discrimination skills. *Dev. Sci.* **6**, F1–F10 (2003).
97. Thiessen, E. D., Hill, E. & Saffran, J. R. Infant-directed speech facilitates word segmentation. *Infancy* (in the press).
98. Marler, P. in *The Epigenesis of Mind: Essays on Biology and Cognition* (eds Carey, S. & Gelman, R.) 37–66 (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1991).
99. Evans, C. S. & Marler, P. in *Comparative Approaches to Cognitive Science: Complex Adaptive Systems* (eds Roitblat, H. L. & Meyer, J.-A.) 341–382 (MIT Press, Cambridge, Massachusetts, 1995).
100. Petitto, L. A., Holowka, S., Sergio, L. E., Levy, B. & Ostry, D. J. Baby hands that move to the rhythm of language: hearing babies acquiring sign language babble silently on the hands. *Cognition* **93**, 43–73 (2004).
- This study showed that hearing babies acquire sign language 'babble' with their hands in a way that differs from hearing babies acquiring spoken language.**
101. Kuhl, P. K., Williams, K. A. & Meltzoff, A. N. Cross-modal speech perception in adults and infants using nonspeech auditory stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* **17**, 829–840 (1991).
102. Meltzoff, A. N. Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Dev. Psychol.* **31**, 838–850 (1995).
103. Adolphs, R. Cognitive neuroscience of human social behaviour. *Nature Rev. Neurosci.* **4**, 165–178 (2003).
104. Dunbar, R. I. M. The social brain hypothesis. *Evol. Anthropol.* **6**, 178–190 (1998).
105. Knightly, L. M., Jun, S.-A., Oh, J. S. & Au, T. K.-F. Production benefits of childhood overhearing. *J. Acoust. Soc. Am.* **114**, 465–474 (2003).
106. Funabiki, Y. & Konishi, M. Long memory in song learning by zebra finches. *J. Neurosci.* **23**, 6928–6935 (2003).
107. Wilbrecht, L. & Nottebohm, F. Vocal learning in birds and humans. *Ment. Retard. Dev. Disabil. Res. Rev.* **9**, 135–148 (2003).
108. Nottebohm, F. The road we travelled: discovery, choreography, and significance of brain replaceable neurons. *Ann. NY Acad. Sci.* **1016**, 628–658 (2004).
109. Dehaene-Lambertz, G., Dupoux, E. & Gout, A. Electrophysiological correlates of phonological processing: a cross-linguistic study. *J. Cogn. Neurosci.* **12**, 635–647 (2000).
110. Callan, D. E., Jones, J. A., Callan, A. M. & Akahane-Yamada, R. Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *Neuroimage* **22**, 1182–1194 (2004).
111. Zhang, Y., Kuhl, P. K., Imada, T. & Kotani, M. Effects of language experience: where, when & how. *Cognitive Neuroscience Society Annual General Meeting program* 2003, 81–82.

112. Sanders, L. D., Newport, E. L. & Neville, H. J. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nature Neurosci.* **5**, 700–703 (2002).
113. Golestani, N. & Zatorre, R. J. Learning new sounds of speech: reallocation of neural substrates. *Neuroimage* **21**, 494–506 (2004).
114. Wang, Y., Sereno, J. A., Jongman, A. & Hirsch, J. fMRI evidence for cortical modification during learning of Mandarin lexical tone. *J. Cogn. Neurosci.* **15**, 1019–1027 (2003).
115. Winkler, I. *et al.* Brain responses reveal the learning of foreign language phonemes. *Psychophysiology* **36**, 638–642 (1999).
116. Koyama, S. *et al.* Cortical evidence of the perceptual backward masking effect on /l/ and /r/ sounds from a following vowel in Japanese speakers. *Neuroimage* **18**, 962–974 (2003).
117. Temple, E. *et al.* Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. *Proc. Natl Acad. Sci. USA* **100**, 2860–2865 (2003).
This fMRI study of children with dyslexia showed that an auditory processing and oral language remediation programme produced increased brain activity in areas that are usually activated in children who have no difficulty in reading.
118. Cheour, M. *et al.* Magnetoencephalography (MEG) is feasible for infant assessment of auditory discrimination. *Exp. Neurol.* (in the press).
119. Dehaene-Lambertz, G., Dehaene, S. & Hertz-Pannier, L. Functional neuroimaging of speech perception in infants. *Science* **298**, 2013–2015 (2002).
The authors used fMRI to show that, like adults, language activates areas in the left hemisphere, with additional activation in the prefrontal cortex of awake infants.
120. Pena, M. *et al.* Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl Acad. Sci. USA* **100**, 11702–11705 (2003).
121. Mills, D. L., Coffey-Corina, S. & Neville, H. J. Language comprehension and cerebral specialization from 13–20 months. *Dev. Neuropsychol.* **13**, 397–445 (1997).
122. Kuhl, P. K. A new view of language acquisition. *Proc. Natl Acad. Sci. USA* **97**, 11850–11857 (2000).
123. Dehaene, S., Spelke, E., Pinel, P., Stanescu, R. & Tsivkin, S. Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science* **284**, 970–974 (1999).
124. Iverson, P. *et al.* A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* **87**, B47–B57 (2003).
125. Friederici, A. D. & Wessels, J. M. I. Phonotactic knowledge of word boundaries and its use in infant speech perception. *Percept. Psychophys.* **54**, 287–295 (1993).
126. Mattys, S., Jusczyk, P., Luce, P. & Morgan, J. L. Phonotactic and prosodic effects on word segmentation in infants. *Cognit. Psychol.* **38**, 465–494 (1999).
127. Werker, J. F., Fennell, C., Corcoran, K. & Stager, C. Infants' ability to learn phonetically similar words: effects of age and vocabulary size. *Infancy* **3**, 1–30 (2002).
This study showed that 14-month-old infants could not learn to pair phonetically similar words with different objects, whereas 20-month-old infants could. Vocabulary size was a predictive factor in the younger infants.
128. Stager, C. & Werker, J. F. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* **388**, 381–382 (1997).
129. Morgan, J. L. & Demuth, K. *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1996).
130. Tsao, F. M., Liu, H. M. & Kuhl, P. K. Speech perception in infancy predicts language development in the second year of life: a longitudinal study. *Child Dev.* **75**, 1067–1084 (2004).
131. Pang, E. *et al.* Mismatch negativity to speech stimuli in 8-month-old infants and adults. *Int. J. Psychophysiol.* **29**, 227–236 (1998).
132. Kuhl, P. K., Nelson, T., Coffey-Corina, S., Padden, D. M. & Conboy, B. Early brain and behavioral measures of native and non-native speech perception differentially predict later language development: the neural commitment hypothesis. *Soc. Neurosci. Abstr.* 15935 (2004).
133. Knudsen, E. I. in *Fundamental Neuroscience* (ed. Zigmund, M. J.) 637–654 (Academic, San Diego, 1999).
134. Lenneberg, E. H. *Biological Foundations of Language* (Wiley, New York, 1967).
135. Newport, E. Maturation constraints on language learning. *Cognit. Sci.* **14**, 11–28 (1990).
136. Johnson, J. & Newport, E. Critical period effects in sound language learning: the influence of maturation state on the acquisition of English as a second language. *Cognit. Psychol.* **21**, 60–99 (1989).
137. Piske, T., MacKay, I. & Flege, J. Factors affecting degree of foreign accent in an L2: a review. *J. Phonetics* **29**, 191–215 (2001).
138. Long, M. Maturation constraints on language development. *Stud. Second Lang. Acquis.* **12**, 251–285 (1990).
139. Birdsong, D. & Molis, M. On the evidence for maturational constraints in second-language acquisition. *J. Mem. Lang.* **44**, 235–249 (2001).
140. Flege, J. E., Yeni-Komshian, G. H. & Liu, S. Age constraints on second-language acquisition. *J. Mem. Lang.* **41**, 78–104 (1999).
A study of second language learning in Korean speakers who arrived in the United States at different ages. Age of arrival in the United States predicted the strength of perceived foreign accent, but grammaticality scores were more related to education and use of English.
141. Mayberry, R. I. & Lock, E. Age constraints on first versus second language acquisition: evidence for linguistic plasticity and epigenesis. *Brain Lang.* **87**, 369–84 (2003).
142. Greenough, W. T. & Black, J. E. in *The Minnesota Symposia on Child Psychology, Vol. 24: Developmental Behavioral Neuroscience* (eds Gunnar, M. & Nelson, C.) 155–200 (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1992).
143. Oh, J. S., Jun, S.-A., Knightly, L. M. & Au, T. K.-F. Holding on to childhood language memory. *Cognition* **86**, B53–B64 (2003).
144. Au, T. K.-F., Knightly, L. M., Jun, S.-A. & Oh, J. S. Overhearing a language during childhood. *Psychol. Sci.* **13**, 238–243 (2002).
This study showed that adults speak a second language with a more native-like accent if they overheard the language regularly during childhood.
145. Pallier, C. *et al.* Brain imaging of language plasticity in adopted adults: can a second language replace the first? *Cereb. Cortex* **13**, 155–161 (2003).
146. Flege, J., Bohn, O. & Jang, S. Effects of experience on non-native speakers' production and perception of English vowels. *J. Phonetics* **25**, 437–470 (1997).
147. Morgan, J. L., Meier, R. & Newport, E. L. Structural packaging in the input to language learning: contributions of intonational and morphological marking of phrases to the acquisition of language. *Cognit. Psychol.* **19**, 498–550 (1987).
148. Saffran, J. R. Statistical language learning: mechanisms and constraints. *Curr. Dir. Psychol. Sci.* **12**, 110–114 (2003).
149. Hauser, M. D., Chomsky, N. & Fitch, W. T. The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1569–1579 (2002).

Acknowledgements

The author is supported by grants from the National Institutes of Health, the Santa Fe Institute, the National Science Foundation (Science of Learning Center), and the William P. and Ruth Gerberding University Professorship Fund. The author thanks D. Padden, J. Pruitt, L. Yamamoto and T. Knight for assistance in preparing the manuscript, and A. Meltzoff and G. Cardillo for helpful comments on earlier drafts.

Competing interests statement

The author declares no competing financial interests.

Online links

FURTHER INFORMATION

Encyclopedia of Life Sciences: <http://www.els.net/Language>

Kuhl's homepage: <http://labs.washington.edu/kuhl/>

Access to this interactive links box is free online.