Theories of early language acquisition

Kim Plunkett

What features of brain processing and neural development support linguistic development in young children? To what extent is the profile and timing of linguistic development in young children determined by a pre-ordained genetic programme? Does the environment play a crucial role in determining the patterns of change observed in children growing up? Recent experimental, neuroimaging and computational studies of developmental change in children promise to contribute to a deeper understanding of how the brain becomes wired up for language. In this review, the multidisciplinary perspectives of cognitive neuroscience, experimental psycholinguistics and neural network modelling are brought to bear on four distinct areas in the study of language acquisition: early speech perception, word recognition, word learning and the acquisition of grammatical inflections. It is suggested that each area demonstrates how linguistic development can be driven by the interaction of general learning mechanisms, highly sensitive to particular statistical regularities in the input, with a richly structured environment which provides the necessary ingredients for the emergence of linguistic representations that support mature language processing. Similar epigenetic principles, guiding the emergence of linguistic structure, apply to all these domains, offering insights into phenomena ranging from the precocity of young infant’s sensitivity to speech contrasts to the complexities of the problem facing the young child learning the Arabic plural.

Very few of the major landmarks in language development have been tagged to specific aspects of brain development. Although many critical events in neural development occur during the language learning years and their potential role in language development has been discussed\(^1\), in most cases the causal link between language development and behaviour remains unclear. Language development may depend on neuroanatomical and neurophysiological changes in the brain. Conversely, language development may reflect a fine tuning of pre-fabbed, genetically-determined structures and a consolidation of cortical representations. Recent neuroimaging and neuropsychological research on infants and children shows that the developing language system undergoes striking changes in neural organization, displaying extensive plasticity (see Box 1), and that its final layout depends critically on experience. At the same time, experimentalists have uncovered an increasingly detailed picture of the sophisticated linguistic propensities of the young infant. In this article, I propose that these findings indicate the need for learning mechanisms that are finely tuned to extracting statistical properties of the speech input. Computational implementations of these learning mechanisms provide the language acquisition researcher with important clues as to how the brain becomes wired up for language.

Early speech perception

Newly born infants can discriminate the speech contrasts of all human languages\(^3\). Furthermore, their way of categorizing speech sounds is universal, so that a child born to Japanese-speaking parents has the same phonemic category boundaries as a child born to Spanish-speaking parents. For some non-native speech contrasts, the ability to discriminate remains intact until fairly late in the first year. For example, infants of 6–8 months of age from an English-speaking background have been shown to be able to distinguish the glottalized velar/uvular stop contrast [k\textcircled{i}]-[q\textcircled{i}] in Nthlakapmx, a language spoken by people in a tribe indigenous to the Northwest Pacific, and the Hindi voiceless aspirated versus breathy voiced contrast [t\textcircled{a}]-[da\textcircled{a}]), neither of which are exploited in English. For English-learning infants, this ability declines after the age of 10–12 months as their phonological processing develops\(^6\). For other contrasts, discrimination sensitivity can decline even earlier. For example, Polka and Werker\(^8\) have shown that English infants have already lost the ability to discriminate the German vowel contrasts [Y\textcircled{a}]-[U\textcircled{a}] by 6–8 months of age. It would appear that sensitivity to non-native contrasts declines earlier for vowels than for consonants\(^9\). Interestingly, however, not all sensitivities to non-native contrasts
Decline in this fashion. Best et al. have shown that even English-learning infants aged 12–14 months are able to discriminate Zulu click contrasts, as are English-speaking adults.

The mechanisms by which the newly born infant can discriminate all human speech sounds, and the process by which the child becomes attuned to the parental language are not well understood. Dehaene-Lambertz and Dehaene have shown that auditory evoked related potentials (ERPs) can be used to unravel the temporal and spatial organization of the neuronal processes underlying phoneme discrimination. They played two-month-old infants synthesized speech stimuli as groups of five syllables (e.g., /ba/, /ba/, /ba/, /ga/) where the first four syllables were identical (the standard) and the fifth was either identical or phonetically different (deviant). A significant difference in auditory ERPs between standard and deviant stimuli showed that the infant could discriminate the deviant stimuli. Christophe and Morton suggest that this technique might be used to study the developmental profile of responses to native and non-native contrasts, thereby shedding light on whether the brain is still sensitive to non-native speech contrasts, but ignores the information, or whether the ability to discriminate non-native speech contrasts is truly lost.

Auditory ERPs offer an important new tool for studying which aspects of the acoustic stimulus young infants are sensitive to, when they are sensitive to it and even, in some cases, the parts of the brain that are the most revealing of these discriminative capacities. However, ERP measurements are unlikely to tell us how the brain actually accomplishes the task. Nakisa and Plunkett have developed a neural network model of early phonological development. The model is based roughly on Jusczyk and Bertoncini’s proposal that the development of speech perception should be viewed as an innately guided learning process: learning the speech contrasts of the native language takes place rapidly because the system is innately structured to be sensitive to correlations of certain distributional properties of the speech stimulus and not others.

The model is based on vertebrate neuronal development. Neurons are allowed to migrate, grow axons and synapses under the control of genes for various trophic factors. Other genes then control the means by which synapses are modified by experience. A population of these neural networks is generated and allowed to breed, with a selective pressure for networks that respond in the desired way to speech sounds. Network fitness is calculated using the stored output unit activities after the network has been exposed to a test set of spoken English sentences. The fitness function favours networks that represented occurrences of the same phoneme as similarly possible and different phonemes as differently as possible. When, after many generations of this evolutionary process, one of these neural networks is exposed to speech sounds from any of 14 human languages (including English, Cantonese, Swahili, Farsi, Czech, Hindi, Hungarian, Korean, Polish, Russian, Slovak, Spanish, Ukrainian and Urdu), it rapidly modifies its connections and creates a representation of speech sounds that is the same regardless of the language to which it has been exposed. Furthermore, the internal representations of speech in the network show the same categorical boundaries that are observed in adult and infant perception (see Box 2). Once a network architecture has been selected by the evolutionary process, only two minutes of speech are required to train the network.

The innately guided learning exhibited by this network enables it to learn very quickly and makes it less dependent on the ‘correct’ environmental statistics. The model offers an account of how infants from different linguistic environments can learn the same featural representation so soon after birth. In this sense, innately guided learning as implemented in this model is half-way between nativism and constructivism. It shows how genes and the environment can interact to ensure rapid development of a featural representation of speech on which further linguistic development depends.
Box 2. Categorical perception in the Nakisa and Plunkett* model

Categorical perception of phonemes is a robust phenomenon observed in both infants and adults. The network was trained on a series of 11 spectra which formed a linear continuum from the pure /sh/ to a pure /s/. Individually, each of the 11 spectra in the continuum were fed into a network that had been trained on 30 sentences of continuous speech in English. The output feature responses were stored for each spectrum in the continuum. The distances of these feature vectors from the pure /sh/ and pure /s/ indicated the categorical nature of the network's internal representations of the speech spectra, as shown in the figure below.

All of the human languages tested seemed to be equally effective for training the network to represent English speech sounds. To see whether any sounds could be used for training, the network was trained on white noise. This resulted in slower learning and a lower final fitness. The fitness for a network trained on white noise never reached that of the same network trained on human speech. An even worse impediment to learning was to train on low-pass filtered human speech.

There is also evidence that infants are sensitive to the prosodic organization of their native language. Jusczyk et al.16 report that prelinguistic infants have identified a regularity of English wherein disyllabic words tend to adhere to a trochaic (strong-weak) stress pattern17. Newcombe and Jusczyk18 show that infants aged seven and a half months can use this knowledge to segment disyllabic words from the main speech stream. Young children are also more likely to imitate syllables that are stressed or word-final than syllables that are both unstressed and nonfinal19. Fernald et al.20, using a preferential looking task, have shown that infants are more likely to recognize familiar words in utterance-initial or final position than when the word occurs in the middle of the utterance.

Saffran et al.21 have focused on the ability of young children to acquire linguistic structure via statistical cues. They point out that the statistical properties of multisyllabic words are potentially useful for infant word segmentation. Over a corpus of speech sounds, there are measurable regularities that distinguish those recurring sound sequences that comprise words from the more accidental sound sequences which occur across word boundaries. Using the familiarization-preferential looking procedure (see Box 3), Saffran et al.22 showed that eight-month-old infants are able to perform the necessary statistical computations. Following a two minute exposure to a synthetic speech stream containing only statistical cues to word boundaries, the infants' listening preferences demonstrated that they had extracted and remembered serial order information about the familiarization items, distinguishing 'words' (recurring syllable sequences) from syllable strings spanning word boundaries. This preferential behaviour indicates that the infants computed the co-occurrence frequencies for pairs of sounds across the familiarization corpus. Jusczyk et al.23 report that nine-month-old infants (but not six-month-olds) are attentive to the frequency with which phonotactic sequences occur within English. These results, together with the findings of Saffran et al.24, suggest that infants have access to a powerful mechanism for the computation of statistical properties of the language input even from very brief exposures and that this develops sometime between the age of six and nine months. They indicate that infants may be far better at deriving structure from statistical information than has often been assumed in the acquisition literature. In particular, certain aspects of language that are argued to be unlearnable and thus innately specified may be discoverable by appropriately constrained statistical learning mechanisms.

Word recognition

During the first year of life, infants become attuned to more than the phonemic contrasts of their native language. They pick up knowledge that enables them to identify words and other linguistic units in speech. Jusczyk and Aslin24 have used the familiarization-preferential looking procedure (see Box 3) to demonstrate that even infants aged seven and a half months have some ability to detect words in fluent speech. Jusczyk et al.25 showed that nine-month-old American and Dutch infants prefer to listen to word lists that conform to the phonetic and phonotactic structure of their own language. In contrast, six-month-old American infants showed no preference for lists from either language.

Reference


*Source: Trends in Cognitive Sciences - Vol. 1, No. 4, July 1997
Box 3. Familiarization-preference head turn procedure

The familiarization-preference procedure was developed by Jusczyk and Adlin. In this procedure, infants are exposed to auditory material that serves as a potential learning experience. Subsequently, they are presented with two types of test stimuli: (a) items that were contained within the familiarization material, and (b) items that were highly similar but are not contained within the familiarization material. During a series of test trials that immediately follows familiarization, infants control the duration of each test trial by their sustained visual fixation of a blinking light. If infants have extracted the crucial information about the familiarization items, they may show differential durations of fixation (listening) during the two types of test trials.

Reference

Word learning
One of the most dramatic manifestations of language development during the first two years of life is the rapid increase in the rate of vocabulary development observed at around 18 months of age. This spurt in development usually occurs in both comprehension and production. There are three main families of theories about the mechanism underlying this vocabulary spurt. These are linguistic development, conceptual development, and the development of constraints on word learning. All of these theories postulate the triggering of a new principle of organization into the child's understanding of the object-label relationship. Woodward et al. argue that these explanations imply that learning a new word prior to the vocabulary spurt is likely to be a time-consuming process, requiring considerable exposure to a new word. There is growing evidence, however, that the young (precocious vocabulary spurt) child may not be as hampered in learning new words as was thought previously. Woodward et al. report that, under favourable circumstances, 13-month-old infants can learn novel words from as few as nine presentations of a novel word token. Baldwin argues that joint attention between the infant and its instructor is necessary for word learning. However, Schafer and Plunkett have succeeded in replicating the findings of Woodward et al. under conditions that do not require the presence of an instructor, suggesting that the precocious vocabulary spurt child is already equipped with a powerful learning mechanism for forming object-label associations.

Imaging early language development
Event-related potentials have been used to examine developmental changes in neural processing in normal children. Mills et al. examined the changes in the organization of brain activity linked to comprehension of single words in infants aged from 13 to 20 months. Auditory evoked related potentials were recorded as children listened to both a series of words whose meanings they did or did not understand, and to words pronounced backwards. The ERPs differed as a function of word comprehension within 200 ms after word onset. At 13–17 months of age comprehension-related differences were bilateral and broadly distributed over the anterior and posterior cortex. In contrast, at 20 months of age these effects were limited to temporal and parietal regions of the left scalp. These results indicate that the neural organization for word comprehension shifts, and that this shift occurs precisely during the period of development when language acquisition is most pronounced. The implication is that plasticity and reorganization may be natural properties of the developing system and are not restricted to compensatory changes in damaged brains. Mills et al. suggest that aspects of their
ERP findings are linked to changes in early lexical development that occurs typically between 13 and 20 months of age. However, it is still unclear whether the changes in ERP reflect qualitative changes in the underlying language processing of lexical items or a consolidation of existing lexical representations. Schafer and Plunkett\cite{1996} have developed a technique based on the preferential-looking task for training infants on novel words. Combining novel word learning with ERP measurements offers an opportunity to evaluate whether the observed hemispheric specialization for lexical processing arises from prolonged experience with words or from the development of new cognitive processing strategies.

**Connectionist modelling of non-linear word learning**

Recent work in connectionist modelling has demonstrated that small and gradual changes in a network, not involving the maturation of new systems, can lead to dramatic non-linearities in its behaviour. For example, Plunkett et al.\cite{1993} have developed a connectionist model of children’s vocabulary development that involves an auto-associative process of relating labels to images. Although the training of the model involved small continuous changes in the connection strengths within and across the different processing modalities in the network (see Box 4), the linguistic behaviour of the network exhibited dramatic non-linearities in vocabulary development, mimicking the well-known vocabulary spurt that occurs in young children towards the end of their second year. Furthermore, the model showed clear-cut dissociations in receptive and expressive vocabulary development, suggesting that the asymmetries between comprehension and production that are often observed in young children may be a natural outcome of the child’s attempt to integrate multiple representations. This modelling work suggests that the results observed in children with focal lesions (see Box 1) and in ERP studies of word comprehension need not necessarily imply prewired, dedicated modules. The results are entirely consistent with the non-linear onset of overt behaviours linked to gradual experience-driven learning processes and coordination of multiple representations in the underlying neural system.

**Inflectional morphology**

Symbolic accounts of the acquisition of inflectional morphology\cite{1994,1997} assume a dual-route mechanism for the processing of regular and exceptional words: a rule-governed process attempts to inflect all words, while an associative memory attempts to identify the exceptions to the rule and block its application. For example, in this view, the plural formation of 'sheep's is blocked by the identification of the exceptional plural form 'sheep' in associative memory whereas plural formation of 'boys' is achieved by application of the rule (add /s/) to the word 'boy'. The rule-governed process acts as a default that applies to any word,
offering the language user economy in representation (no need to store information about inflected forms that conform to the default) and creativity (the capacity to inflect forms previously not encountered).

In contrast, connectionist accounts of the acquisition of inflectional morphology\(^\text{42-46}\) assume a single-route mechanism for the processing of both regular and exceptional forms. There is no distinction in the manner in which regular and exceptional forms are handled in this account. They are processed by the same network of connections that maps an uninflected form of the word to its inflected form. The network's capacity to inflect novel forms is shaped by its experience with the forms on which it has already been trained.

In English, the inflectional systems of the past tense and the plural are highly regular. Irregular past tense forms and irregular noun plurals constitute only 14% and 2% of their respective systems\(^\text{47}\). The dual-route account of inflectional morphology is very efficient at representing these systems as only a minority of forms need to be stored in associative memory and the default rule can deal with the majority of forms. A connectionist network stores information about all the words. Nevertheless, the dominance of the regular words in the system results in the network producing regular responses to novel words. Consequently, both dual-route and connectionist approaches can explain the preponderance of regular responses to novel words by English speakers but for different reasons: the dual-route account exploits a default rule that attempts to regularize any word available to the language user; the connectionist account exploits the skewed distribution in favour of regular words in the language.

Minority defaults

There is evidence from speakers of other languages that their ability to produce a default response to novel words or overgeneralize the default to exceptional words does not rely upon a numerical superiority of the words that epitomize the default in the language. For example, Clahsen et al.\(^\text{48}\) and Marcus et al.\(^\text{49}\) claim that the 's' plural in German is the default process even though it constitutes a minority of the plural forms in the language. A similar claim is made for the default status of the 'sound' plural in Arabic. These authors claim that languages whose speakers conform to a minority default pattern, appear to present a major challenge to connectionist accounts of inflectional morphology as networks operating on the principle of 'similar inputs produce similar outputs' are unlikely to produce a default response to novel forms.

Hare et al.\(^\text{50}\) have demonstrated that connectionist models of inflectional morphology can learn a default response even in the absence of superior numbers for the default class. Two factors contribute to a network's capacity to respond in a default-like fashion: First, words which look similar at the input need not have similar internal representations. Second, the distribution of the words in the language influences the ability of the network to act in a default-like fashion (see Box 5). Under appropriate conditions (see Box 5), it is possible for a network to learn a distributional default.

Plunkett and Nakisa\(^\text{51}\) trained a neural network on the Arabic plural and evaluated its performance on words not encountered in the training set. They showed that the network was superior to the dual-route model at predicting the plural class of Arabic words on which it had never been trained. In particular, prediction of membership in the sound plural class was more accurate in the neural network model. In a similar fashion, Nakisa et al.\(^\text{52}\) have shown that a connectionist network trained on a subset of German plurals predicts accurately the class membership of German plurals that it has never seen before. The network is in much the same position as the Arabic or German child who may have to guess how to form the plural of a word. These results indicate that the distribution of nouns in Arabic and German may provide subtle clues to plural class membership which are not obvious even to sophisticated professional linguists.

Conclusions

As yet, none of the domains of language acquisition described above are understood properly. However, the picture of the language learning child is becoming increasingly refined as we uncover the details of what is developing and when development occurs, where the neural systems in the brain for controlling linguistic behaviours are located and how these systems actually function. Behavioural, neuropsychological and computational studies reveal that the young infant is richly endowed with neural systems, well-adapted to the business of linguistic information processing. At the same time, I believe that a multidisciplinary approach to the study of language acquisition points to the utility of viewing linguistic development as driven by the interaction of powerful general learning mechanisms with a richly structured environment that provides the necessary ingredients for the emergence of mature linguistic representations.

**Outstanding questions**

- What are the mechanisms that enable infants to tune in to the speech contrasts that are specific to their native language? Why do some non-native speech contrasts remain discriminable by adults and children while others are no longer perceived as distinct? Does the brain remain sensitive to non-native speech contrasts even though discrimination experiments demonstrate a lack of such sensitivity?
- Are the word-like linguistic chunks that infants extract from continuous speech stored in a prelexical mental repository awaiting adequate conceptual development to achieve semantic grounding? How does conceptual development influence the process of lexical segmentation? What developments underpin the dramatic changes in the statistical processing of speech that seems to occur between six and nine months of age?
- What is the nature of the label-object associations shown in recent demonstrations of rapid word learning in prevocabulary spurt infants? Is the lateralization of early word representations to the left hemisphere in postvocabulary spurt infants a consequence of prolonged experience with specific words or is it due to the emergence of new lexical processing strategies?
- What are the facts of acquisition in complex inflectional systems like the Arabic and German plural? Are there default processes operating in these and other languages or are over-regularizations and generalizations better understood as operating through processes governed by analogy and frequency?
Box 5. How to obtain a minority default from a neural network

Forrester trained a neural network to categorize stimuli belonging to one of three classes. Each input point corresponds to a point on a two-dimensional plane. The distribution of points is shown below. The majority of the points are contained in two squares. All the points within a square are assigned to belong to the same class. These can be thought of as exceptions. The minority of the points are distributed outside these square regions. All the points show the square regions belong to the same class. These can be thought of as patterns representing the default category (minority class). The question of interest is how does a neural network trained on this distribution of points respond to novel patterns, that is, the points in the two-dimensional plane on which it has never been trained? The classifier network used by Forrester and Plunkett contained two input units to specify the x, y coordinates in the two-dimensional plane, 20 hidden units that formed internal representations of the input patterns permitting a non-linear classification of the input space and three output units to classify the input patterns. The network was trained with the points shown below and then tested on every point in the plane. The surface plots show the activation of the three classifier units at different stages in training. Darker regions indicate higher activation. The final column (late training) shows that all the classifier units do a good job at partitioning the space by the end of training. In particular, most of the points in the two-dimensional plane are treated as though they belong to the third class, the so-called defaults, even though the training set contained a minority of forms in this class. This example demonstrates how a neural network can be trained to produce a default-like response provided it has the resources to construct internal representations that permit a non-linear partitioning of the input space and provided the forms in the 'language' are appropriately distributed.

---

**References**


37 Schaffer, G. and Plunkett, K. Rapid word learning by 10-month-olds under tightly controlled conditions. Cogn. Brain Res. 13, 597-611
49 Marcus, G. et al. (1995) German inflection: the exception that proves the rule. Cognitive Psychol. 30, 186-235
52 Nakisa, R.C., Plunkett, K. and Hahn, U. A cross-linguistic comparison of single and dual-route models of inflectional morphology, in Models of Language Acquisition: Inductive and Deductive Approaches (Broader, P. and Murre, J., eds), MIT Press (in press)

In the September issue....

The Trends Guide to the Internet (1997)

Feeling left behind? Worried about time-wasting? This free guide will lead you through those all-important first steps towards using the Internet efficiently and effectively. Written by an international panel of experts, the guide will tell you what to expect and where to find it. Articles include:

- The origins of the Internet
- Key terms: the jargon explained
- Basic Internet facilities and how to connect
- The World Wide Web
- USENET: setting up and joining newsgroups
- Creating your own home page
- FTP: how to retrieve files from around the world
- What's new in online journals
- Where to go next for help

Also including a poster listing useful sites to visit.

For further information concerning bulk sales, contact Thelma Reid c. mail, t.reid@elscvier.co.uk; Tel, +44 1223 311114; Fax, +44 1223 321410.