

The periodicity bias

Anne Cutler

MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, U.K.

Jacques Mehler

Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS & EHESS, 54 Boulevard Raspail, 75006 Paris, France

The evidence presented in the target articles suggests that infants' phonetic development is driven by the acquisition of words as holistic units; and that segmental representations, previously considered to become language-specific only late in the first year of life, actually arise throughout this first year, with language-specific representations of vowels arising earlier than of consonants. We propose that these circumstances are unified by a bias towards attention to periodic sounds, present at birth. This equips the child to exploit linguistic rhythm to achieve initial segmentation of continuous speech signals to extract word units; the use of linguistic rhythm for segmentation is still present in adult processing. It also has the effect of making vowels achieve effective prototypical representations earlier than consonants.

Children are born without knowledge of a language. Nothing in their genetic endowment predisposes them to acquire one human language rather than another. Yet a child who is exposed to a human language will, in the absence of exceptional circumstances, acquire that language. In fact, children display clear evidence of knowledge about the specific language of their environment before the end of the first year of life, well before they can effectively communicate language. In addition to all the other development work which children undertake during their first 12 months, they make formidable strides towards acquisition of the highly complex skill of language. What is the precise course of this process? And in particular, which part of the child's innate endowment plays the main role in how the process unfolds?

Below we will suggest an answer to this latter question, and show how our suggestion relates to the theme of this special issue: phonetic development. This theme is somewhat controversial in its own right. It is probably fair to say that a majority of work in language acquisition has dealt with lexical/syntactic/semantic development, with phonetic development trailing somewhat behind—despite the impressive body of work so comprehensively reviewed, from varying perspectives, by the four target articles in this issue.

On the one hand, this imbalance is inappropriate, since the extent to which language acquisition as a whole depends on the development of phonetic competence is all too frequently glossed over. For instance, theories which propose that the child

uses semantic abilities as a basis for the development of syntactic competence (Pinker, 1984), or syntactic abilities as a basis for the development of semantic competence (Gleitman, 1990) depend crucially on the child's capacity to manipulate individual words, and this in turn depends on the child's awareness that speech signals are composed of individual words. Since speech signals are in fact continuous, with in many languages few robust and reliable cues to where one word ends and the next begins, the highly important question of how the child learns to extract word units from speech is in large part a phonetic one.

On the other hand, there is a sense in which a preference on the part of language acquisition researchers for the investigation of lexical, syntactic and semantic development is entirely reasonable. One could even ask reasonably whether there is such a process as phonetic development at all. Do children (in contrast to adult learners of new languages) focus at all on acquiring sounds? Or is the acquisition of phonetic abilities something which arises as a necessary effect of the successful acquisition of linguistic skills in general, wherein the child's focus is firmly on the most general aspect of these skills, namely the ability to communicate? As adult learners of new languages know well, imperfect development of phonetic skills is not necessarily a bar to successful communication; imperfect lexical, syntactic and semantic skills often can be.

Something like this point of view seems to underlie the approach of several of the target articles in this issue. Thus the articles by WERKER & POLKA and by JUSCZYK explicitly point out that the process of phonetic (or, by implication, any other) development is most profitably considered in terms of the needs of the child at each point in the developmental course; the primary need, according to JUSCZYK, is the establishment of communication. With respect to the development of segmental skills, JUSCZYK, SUOMI and VIHMAN all see these as arising out of the production of words as holistic units. SUOMI goes furthest, in arguing that such an approach can obviate the need for segmental representations at any level of processing by either the child or the adult language user. SUOMI's DAPHO model is, explicitly, a model of adult phonetic processing which is driven by evidence from the developmental situation; in this it again resembles JUSCZYK's approach, since JUSCZYK's WRAPSA model is likewise developmental but designed to provide an account as well of adult processing.

Our own approach, too, is based on an integration of the child and the adult case. The results of our experimental programme investigating speech segmentation by adults suggest to us a way in which adult processing reflects the child's earliest linguistic experience. Our work has focused on the problem posed for adult listeners by the fact that there are no robust and reliable cues to word boundaries. In practice this problem for the adult is driven by memory constraints: in order to understand an utterance, listeners have to match it against what is in memory, but human memories could not possibly store as a whole every utterance which might ever be spoken. Therefore we have to identify the individual units of which an utterance is made up, and match these against discrete entries in the mental lexicon. To do this we must segment the speech stream into portions which correspond to lexically stored units.

The most significant outcome of our research over a number of years is the finding that the way in which the speech segmentation problem is solved differs across languages. For French listeners, our evidence suggests that the syllable can act as a

segmentation unit (Mehler, Dommergues, Frauenfelder & Segui, 1981; Cutler, Mehler, Norris & Segui, 1986). In English, on the other hand, syllabic segmentation is not used (Cutler *et al.*, 1986); instead, listeners segment speech according to stress units (Cutler & Norris, 1988; Cutler & Butterfield, 1992). The situation is different again in Spanish and in Catalan (Sebastián-Gallés, Dupoux, Segui & Mehler, 1992). Most recently, our series of studies of speech segmentation has been extended to Japanese (Otake, Hatano, Cutler & Mehler, 1993); the results suggest that Japanese listeners can use a subsyllabic unit called the mora in segmenting speech. English and French listeners presented with the Japanese speech materials responded quite differently, suggesting that mora-based processing is specific to Japanese listeners. In fact, we found that French listeners segmented Japanese speech by syllables, just as they segment both French (Mehler *et al.*, 1981) and English (Cutler *et al.*, 1986).

The Japanese result is particularly interesting because it suggests a way of unifying these language-specific speech segmentation findings. The mora is the unit of rhythm in Japanese (for instance, in Japanese poetry), just as English rhythm is stress-based, while French has syllabic rhythm. In other words, across a phonologically quite diverse set of languages listeners appear to rely on linguistic rhythm to help them solve the speech segmentation problem.

Our interpretation of this pattern of results invokes, as already mentioned, the earliest steps in language acquisition. The infant learning to distinguish meaning in the speech signals which occur in its environment faces a segmentation problem which is compounded by the absence of a store of meaningful units such as the adult possesses. The infant's task is, indeed, to *build* a vocabulary, from scratch. On what basis can this process be started? We have suggested that it may be the case that the characteristic rhythmic pattern of a language is sufficiently salient to assist the newborn child in segmenting the continuous speech stream into discrete units.

This translates to a proposal about the capacities with which the child is already equipped when embarking upon the language acquisition task. In its most neutral formulation, our proposal is that the child starts with the expectation that meaning will map to form, and, moreover, to a particular kind of form: input which is periodically structured. Speech signals have periodic structure, and for the majority of children speech will be among the most salient forms of input available. Of course, there is evidence that in the first few months and even days of life infants prefer to listen to speech rather than to other auditory input (Colombo & Bundy, 1981; Glenn, Cunningham & Joyce, 1981), to speech in their own language rather than in another language (Mehler, Jusczyk, Lambertz, Halsted, Bertoni & Amiel-Tison, 1988), and to child-directed speech rather than adult-directed speech (Fernald, 1985). Moreover the salience in particular of the rhythmic structure of speech is demonstrated by a finding by Condon & Sander (1974) that the neonate is able to synchronize its movement with speech structure. Of particular interest is that Condon & Sander demonstrated clear synchronies between infants' movements and speech input whether the speech was spoken directly to the child or played from a tape recorder, and whether the speech was in the parental language or a foreign language. (Tapping sounds, on the other hand, did not invoke synchrony in the infant's movement.)

The underlying motivation for our proposal comes from the speech segmentation problem. This is not an issue which looms large in the arguments made in the target articles; in fact, the issue is in effect avoided by SUOMI and VIHMAN, both of whose

proposals concern children already in possession of a few words at the very least. How the very first words are acquired is an issue they do not address in detail. JUSCZYK does consider this issue, and his suggestions, though couched in general terms, are very close to the spirit of our own: he suggests that prosodic structure is the dimension which infant listeners exploit to accomplish speech segmentation. (Note that in fact VIHMAN cites Jusczyk's work in this area in agreeing with the suggestion that prosody could provide "an entry point" to the segmentation of continuous speech, and WERKER & POLKA also acknowledge the importance of his demonstration of infants' sensitivity to prosodic structure.) We believe that our approach displays further interesting consistencies with some aspects of the claims made in the target articles.

Consider, for instance, the proposals of WERKER & POLKA. Their paper is the one which is most strongly focused both on perceptual development and on the development of segmental skills. In a clear introductory review they describe what has been until recently a consensus view in this area: that the first year of life is a time of gradually refining the ability to perceive phonologically relevant contrasts. Moreover, it was generally agreed that the process of refinement towards a language-specific target set of phonemes does not begin before six months of age. As WERKER & POLKA go on to recount, however, that consensus view has undergone considerable upheaval as a result of recent work.

The way in which this new work on segmental development differs from the work upon which the consensus view was based is in the nature of the phonetic segments under investigation. For good historical reasons, most work on segmental perception abilities concerned consonants. (Specifically this was because categorical perception of phonemes was first demonstrated in adult consonant perception, and the extension of this work to infant perception was both an obvious and, since categorical perception was indeed replicated in infants, a rewarding move). The new work simply transfers the focus to vowel perception. The result of this simple shift in focus, however, is the comprehensive collapse of the model which seemed to account so well for phonetic development as long as only consonants were considered.

As WERKER & POLKA describe, however, there is evidence that infants are sensitive to vowel contrasts well before they are sensitive to consonant contrasts. Kuhl, Williams, Lacerda, Stevens & Lindblom (1992) have demonstrated that language-specific vowel prototypes are in place by six months of age. The vowel-consonant distinction is, properly speaking, a phonological one: vowels form syllabic nuclei, while consonants occur in the margins of syllables—in onsets and codas. Nevertheless, there are also acoustic-articulatory correlates of the distinction. Vowels are relatively steady-state sounds, produced with vibration of the vocal cords and without obstruction of the airflow from the lungs. Consonants, in contrast, are relatively transient, produced with full or partial obstruction of the airflow from the lungs, and with or without vocal cord vibration. In effect this means: vowels exhibit greater periodicity than consonants.

It is also typically the case that vowels have longer duration than consonants (Crystal & House, 1988). This too could be an important factor in infant perception. Cowan & Morse (1986), on the basis of a finding that (adults') mental representations of vowels gradually become more diffuse over time, argued that memory processes are more critical in vowel than in consonant perception. Cowan, Suomi &

Morse (1982) found evidence consistent with longer-lasting echoic traces in infant than in adult memory; they argued that this could be a valuable compensation for a slower processing rate in infants. If we assume that the speed of speech processing gradually increases through the first year of life, and that at slower rates only those stimuli can be stored which reach a certain minimum steady-state duration, then it makes sense that language-specific representations of prototype vowels will be attained before language-specific representations of prototypes of the more transient speech sounds, the consonants. (Note that not all consonants are equally transient. This view of the order of acquisition of segmental prototypes must predict that prototypes of steady-state consonants such as nasals would be attained earlier than prototypes of, say, stop consonants. As WERKER & POLKA point out, the extension of the infant perceptual studies to all phoneme categories promises interesting insights.)

As JUSCZYK suggests, his evidence on infants' sensitivity to prosodic marking of clausal structure, which precedes their sensitivity to marking of phrasal structure, can be interpreted as a gradual refinement of the ability to perceive patterns: at first relatively gross structure, then ever finer structure (see also Jusczyk & Bertoncini, 1988; Mehler & Dupoux, 1990). Our proposal posits the same process of refinement operating within as well as between such structural levels. That is, the infant could pay attention to larger prosodic chunks, then to metrical units such as the syllable or foot, and last to phonetic segments; in the same way, once attention is focussed on the phonetic segment, attention to the "larger", more steady-state segments would precede attention to the "smaller", more transient units.

In conclusion, we feel that our proposal is very much in agreement with the approaches taken in the target articles, even where these are in apparent disagreement. Children do pay attention to holistic words; and the way they do this in the first instance, to extract their first word units, is by focusing on the characteristic rhythm of the input language. Their bias towards periodicity in the search for meaningful input leads them, in turn, to apply effective processing procedures to vowels before they can do so with consonants. The more periodic a sound (and, as a correlate, the longer its effective duration), the earlier it will be usefully acquired in language-specific form. The rather unsatisfying discontinuity with which WERKER & POLKA conclude their paper—that an abrupt reorganization in the child's phonological system occurs towards the end of the first year of life, effectively requiring unlearning of some phonetic knowledge—may in fact simply fall out of the attainment of the capacity to deal with shorter-duration segments, with the consequent expansion of the phonemic repertoire.

Of course, as VIHMAN so persuasively argues, the effective experience may be highly individual and may lead to individual differences in the precise course of development of phonetic or any other skills. This is one reason why we have framed our proposal in very general terms. We believe that a bias towards attending to periodically structured input is present in all infants, and that its exploitation in language acquisition is not absolutely dependent on the availability of spoken input. Of course it is true that for most children the most salient form of stimulus will be auditory. But not all children are lucky enough to be able to attend to auditory input—and among those that cannot, language acquisition still follows a developmental path with recognizable similarities to that of the acquisition of speech (Pettito & Marentette, 1991). To account for the strong similarities which have been

demonstrated between phonetic development and the development of manual communication skills, it is helpful to consider phonetic organization as merely one example of a particular type of structure to which the newborn infant is inherently biased, which we have called periodic structure. Finally, then, our proposal has the added advantage of greater universality in comparison with a model of phonetic development alone.

Financial support is acknowledged from the Human Frontier Science Program.

References

- Colombo, J. & Bundy, R. (1981) A method for the measurement of infant auditory selectivity, *Infant Behavior and Development*, **4**, 219–233.
- Condon, W. S. & Sander, L. W. (1974) Synchrony demonstrated between movement of the neonate and adult speech, *Child Development*, **45**, 456–462.
- Cowan, N. & Morse, P. A. (1986) The use of auditory and phonetic memory in vowel discrimination, *Journal of the Acoustical Society of America*, **79**, 500–507.
- Cowan, N., Suomi, K. & Morse, P. A. (1982) Echoic storage in infant perception, *Child Development*, **53**, 984–990.
- Crystal, T. H. & House, A. S. (1988) Segmental durations in connected-speech signals: Current results, *Journal of the Acoustical Society of America*, **83**, 1553–1573.
- Cutler, A. & Butterfield, S. (1992) Rhythmic cues to speech segmentation: Evidence from juncture misperception, *Journal of Memory and Language*, **31**, 218–236.
- Cutler, A., Mehler, J., Norris, D. G. & Segui, J. (1986) The syllable's differing role in the segmentation of French and English, *Journal of Memory and Language*, **25**, 385–400.
- Cutler, A. & Norris, D. G. (1988) The role of strong syllables in segmentation for lexical access, *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 113–121.
- Fernald, A. (1985) Four-month-old infants prefer to listen to motherese, *Infant Behavior and Development*, **8**, 181–195.
- Gleitman, L. R. (1990) The structural sources of verb meanings, *Language Acquisition*, **1**, 1–55.
- Glenn, S. M., Cunningham, C. C. & Joyce, P. F. (1981) A study of auditory preferences in nonhandicapped infants and infants with Down's Syndrome, *Child Development*, **52**, 1303–1307.
- Jusczyk, P. W. & Bertoncini, J. (1988) Viewing the development of speech perception as an innately guided process, *Language and Speech*, **31**, 217–236.
- Kuhl, P., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. (1992) Linguistic experience alters phonetic perception in infants by six months of age, *Science*, **255**, 606–608.
- Mehler, J., Dommergues, J.-Y., Frauenfelder, U. & Segui, J. (1981) The syllable's role in speech segmentation, *Journal of Verbal Learning and Verbal Behaviour*, **20**, 298–305.
- Mehler, J. & Dupoux, E. (1990) *Naitre humain*. Paris: Editions Odile Jacob.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988) A precursor of language acquisition in young infants, *Cognition*, **29**, 144–178.
- Otake, T., Hatano, G., Cutler, A. & Mehler, J. (1993) Mora or syllable? Speech segmentation in Japanese, *Journal of Memory and Language*, **32**, in press.
- Pettito, L. A. & Marentette, P. F. (1991) Babbling in the manual mode: Evidence for the ontogeny of language, *Science*, **251**, 1493–1496.
- Pinker, S. (1984) *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Sebastián-Gallés, N., Dupoux, E., Segui, J. & Mehler, J. (1992) Contrasting syllabic effects in Catalan and Spanish, *Journal of Memory and Language*, **31**, 18–32.