3 8 2

Discovering the Sound Pattern of a Language

Grace H. Yeni-Komshian, James F. Kavanagh, and Charles A. Ferguson (Eds.)
Child Phonology, Vol. 1: Production
New York: Academic Press, 1980.
318 pp. $27.00

Grace H. Yeni-Komshian, James F. Kavanagh, and Charles A. Ferguson (Eds.)
Child Phonology, Vol. 2: Perception
New York: Academic Press, 1980.
268 pp. $24.00

Review by
Michael Studdert-Kennedy

Grace H. Yeni-Komshian is associate professor in the Department of Hearing and Speech Sciences at the University of Maryland (College Park). ■ *James F. Kavanagh is associate director of the Center for Research for Mothers and Children at the National Institute of Child Health and Development (Bethesda, Maryland). He is coeditor of* Orthography, Reading, and Dyslexia *with R. L. Venezky.* ■ *Charles A. Ferguson, professor of linguistics at Stanford University, is coeditor of* Language in the U.S.A. ■ *Michael Studdert-Kennedy, reviewer, is professor of communications at Queens College of the City University of New York, as well as research associate at Haskins Laboratories in New Haven, Connecticut. He is coeditor (with U. Bellugi) of* Signed and Spoken Language: Biological Constraints on Linguistic Form.

A typical middle-class American child of six years can recognize nearly 8,000 root words (according to Mildred Templin, 1957). The child has learned these words over roughly four years, at an average rate of 5 or 6 a day. Each word is formed, according to a set of language-specific rules for constructing syllables, by combining a few of the several dozen articulatory patterns that generate the consonants and vowels of an English dialect. How, we may well ask, does the child learn the perceptual and motor patterns that will permit it to build so large a lexicon in so short a time?

That is the question to which these two volumes are addressed. They comprise the proceedings of a conference of 34 linguists and psychologists, convened by the National Institute of Child Health and Human Development in Bethesda, Maryland, during May 1978. They form a compendium of theory and research done over the previous decade in the young field of child phonology. According to a rough count by Jenkins (given in a chapter of shrewd comments, criticism, and advice at the end of Volume 2), over 90% of the references in these volumes are to works published since 1968, and over 60% to works published since 1975.

Child phonology begins (as Ferguson and Yeni-Komshian [Vol. 1, chap. 1] remind us in their useful introductory survey of its history) with the publication of Jakobson's *Kindersprache, Aphasie and Allgemeine Lautgesetze* in 1941. Jakobson's proposals quickly became standard dogma because they offered an elegant integration of phonological development into the then-dominant structuralist account of phonology. Central to Jakobson's position was the view that babbling during the child's first year was mere random articulatory exercise and that learning to speak was a linguistic matter, abrupt in onset and entailing the development of particular phonemic oppositions before other particular oppositions in a fixed, universal order.

However, the discontinuity between babbling and speech is more apparent than real, the consequence, Lieberman (Vol. 1, chap. 7) suggests, of the phonetician's lack of a descriptive framework for pre-speech. MacNeilage (Vol. 1, chap. 2) points out that this lack is now being rectified. He concludes a succinct account of what we know and do not know about adult control of speech production with the enticing suggestion that studies of pre-speech may be amenable to treatment in terms of the coordinative structures of action theory. A *coordinative structure*, or *synergism*, is a set of muscles constrained to act as a unit. For example, Stark (Vol. 1, chap. 5) provides a framework for classifying vocal behavior during the first 15 weeks of life and finds that many features of adult speech are present but uncoordinated. Thus, variations in pitch and vocalic structure are observed during infant cry, whereas consonantal sounds such as clicks, friction noises and trills occur during vegetative processes. Stark proposes that the development of speech involves the harnessing and coordinating of these features into the precisely timed patterns of babble.

Stark's approach meshes neatly with that of Oller (Vol. 1, chap. 6), who reports validating studies of a framework for describing the development of phonetic control during the first year of life, from what he terms the *quasi-resonant nuclei* of nonreflexive vocalizations in the first month to the *variegated babbling* of the 11th and 12th months. His system promises to break a bottleneck in the study of pre-speech vocalization, taking the first step toward norms that may permit early diagnosis of deafness or other pathologies. However, Oller's chief concern is with the theoretical issue of explaining the regularities of infant development. Do they simply reflect general anatomical and physiological maturation? Is there evidence of conscious, speech-related vocal activity during the first year of life? When do the first signs of shaping by the language community appear?

The last question is also raised by Lieberman (Vol. 1, chap. 7) in a preliminary report on a longitudinal acoustic study of the speech of a small group of normal, middle-class children from birth through pre-school. Particularly valuable here, both for normative purposes and as evidence of changes in phonetic scope of the vocal tract, are a dozen formant frequency plots on which one can observe the steadily increasing extent of each child's vowel quadrilateral. Interestingly, the children do not mimic adult formant frequencies, even though for many vowels they could do so, by appropriate vocal tract maneuvers. Instead, already by the fourth month, vowels are falling into their "proper" acoustic relations, a fact consistent with the hypothesis of an innate normalization mechanism. The data also discount Jakobson's claim of discontinuity by illustrating the smooth emergence of the vowels of words from the vowels of babble.

More on Jakobson

Lest it seem that I am flogging Jakobson's horse past death, let me note that his theories are cited (and disputed) in 10 of the 13 chapters in Volume 1. Indeed, Menn in a lucid and thought-provoking chapter (Vol. 1, chap. 3) on the historical development of phonological theory (with the witty epigraph "Beware Procrustes bearing Occam's razor") suggests that "the entire cautious and meticulous modern tradition of child phonology field-work was forged by . . . [the] necessity" of establishing counter-evidence to Jakobson's arguments (p. 28).

This is, in fact, precisely the focus of Macken's chapter on the acquisition of syllable-initial stop systems (Vol. 1, chap. 8). There are two possible tests of Jakobson's claim of a fixed, universal order of development—across children within a language and across languages. Macken does both. She tests Jakobson's prediction of an invariant sequence for stops (/p/ before /p:t/ before /p:k/) on case studies (by others) of five English-learning children and concludes that the best she can do is to reformulate the prediction as "front before back" and then assign it no more than a high probability of being right. Testing Jakobson's prediction that the first stops will be voiceless unaspirated, on her own English and Spanish data, she finds strong support, but also evidence of language-specific patterns in the timing, ordering, and phonetic structure of the first stop contrasts, /bdg/, that seem to reflect relative frequencies of these stops in the language being learned.

The role of language-specific frequencies is, of course, very much to the point and still far from clear. Locke (Vol. 1, chap. 10), reporting a novel and ingenious study on the prediction of child speech errors, presents arguments and evidence that there is no such effect. Nonetheless, there does seem to be much more cross-language and within-language variation than Jakobson would predict. Thus in a careful study of the production of word-initial, English fricatives and affricates by 73 children between 2 and 6 years, Ingram and his colleagues (Vol. 1, chap. 9) found much the same order that previous studies have reported, but with considerable variation from child to child, from word to word, and even from time to time within a word.

Contextual variability has, incidentally, no less clinical than theoretical interest. Menyuk (Vol. 1, chap. 11) reports studies of both perception and production, demonstrating that children with suspected central nervous system abnormalities may present quite different patterns of error according to whether they are assessed with nonsense syllables or familiar words, in a test situation or while playing with other children. Taken with the numerous studies reported in these volumes in which normal children display their diversity, Menyuk's report should encourage caution in the assessment of a child's phonological capacity.

Continuity and discrimination abilities

In Volume 2, *Perception*, we again confront the continuity issue—though not explicitly formulated, perhaps because Jakobson himself did not consider the infant's perceptual capacities. However, Blumstein, once Jakobson's student, fills the gap in a chapter (Vol. 2, chap. 2) reporting her work with Stevens on the spectral structure of stop consonant release bursts. Crossing the psychology of Hume with the linguistics of Jakobson, Blumstein posits "innate biological mechanisms . . . selectively tuned to *primary*, [linguistically] unmarked, invariant acoustic cues" for place of articulation, in conjunction with "marked . . . *secondary* context-dependent cues" whose linguistic function the infant learns "as a direct consequence of the co-occurrence of these cues with the invariant acoustic properties" (p. 19).

The hypothesis of "innate biological mechanisms" stems, of course, from the many studies precipitated by Eimas and his colleagues (1971) when they successfully transposed the high amplitude sucking procedure for assessing an infant's discriminative capacity during the first 3 to 4 months of life, from visual to auditory research. Eilers (Vol. 2, chap. 3) describes the paradigm and others suited to later age ranges—heart rate variation as an index of attention (1–8 months) and visual reinforcement (by an animated toy) of head turning toward the locus of a stimulus change (6–18 months). Eilers also reviews many studies using these techniques to demonstrate that infants can discriminate virtually every major acoustic property that underlies a phonemic contrast in English. Few negative findings have been reported, and such inconsistencies as there are between studies seem to be due to inadequate acoustic specification. For example, researchers generally (Eilers is no exception) seem unaware that voice onset time (VOT) was originally defined as a special case of a general *articulatory* variable (the timing of laryngeal action) that would generate any and all of more than a dozen acoustic cues to voicing distinctions. Unwary synthesis can thus produce different responses to the same value of VOT due to differences in, say, release burst energy or the onset frequency of the first formant.

In any event, what we now have is a rough taxonomy of infant psychoacoustic capacity for discriminating (not categorizing) certain dimensions of speech sounds—in all likelihood, a general mammalian capacity that tunes, rather than is tuned to, speech. Kuhl (Vol. 2, chap. 4) has higher goals. Her current research makes direct tests, by the head-turning technique, of an infant's capacity to form categories of speech sounds. Her data show that 6-month-old infants can learn to categorize: (1) tokens of /a/ versus /i/ and of /a/ versus /ɔ/, spoken by a male, a female and a (synthesized) child on two different pitches; (2) tokens of syllable-initial or syllable-final /s/ versus /ʃ/, and /b/ versus /θ/, spoken by several talkers with /i,a,u/; (3) (according to preliminary data on a single infant) tokens of initial, medial or final /d/ versus /g/, spoken with /i,a,u/. This research directly confronts crucial issues of segmentation and invariance, across speakers and phonetic contexts, and is, in my view, the most interesting current work in the area.

Nonetheless, if 6-month-old infants are indeed able to segment syllables and form categories of their component consonantal and vocalic portions, what are we to make of the apparent perceptual difficulties of older children? Barton (Vol. 2, chap. 6) provides a critical analysis of the methods used to assess a child's capacity to discriminate (that is, distinguish between two stimuli) and identify (that is, refer a stimulus to an internal representation, perceive phonemically). Whatever the task, performance varies with many factors, such as word status (real vs. nonsense), word familiarity, feature composition, and of course, age. In general, 2- to 3-year-old children seem to identify at least familiar words quite accurately. But why should familiarity be a factor at all?

Of course, some sounds are more difficult than others. Barton shows that there is no evidence for any general order of perceptual acquisition in either Russian or English (the only languages on which there have been studies, it seems). But certain distinctions are notoriously difficult—for example, /f/ versus /θ/ (on which Kuhl's infants were successful), or /r/ versus /l/. For the latter contrast, Strange and Broen (Vol. 2, chap. 7) report a careful study of 21 3-year-olds in which they found evidence of a perception–production link: If a child had difficulty with the identification task, he was more likely to have difficulty producing /r/ or /l/ than if he did not.

Perhaps the solution to the puzzle lies in paying more attention to just how a child's perceptual capacity is measured. Strange and Broen (Vol. 2, chap. 7) provide an excellent discussion of this matter as it bears on the relation between perception and production. They suggest that measures of the two processes should be in some sense coordinate: "It would seem . . . more reasonable to compare . . . [the] . . . kind of perceptual capacity [assessed in infants] with an empirical assessment of the physiological capacity to produce these sounds (i.e., with motoric capabilities independent of linguistic volition) . . . [as in] . . . prebabbling vocalizations" (p. 149). They point to our lack of "a concept of 'intentional, coordinated perception' . . . comparable to our understanding of speech production as the articulation of lexical items with the intent to communicate linguistically" (p. 150).

Implicit in this argument is the assumption that perception and production somehow march together. Straight (Vol. 1, chap. 4) in a somewhat naughty and polemic chapter argues, to the contrary, for two separate and distinct components in auditory and articulatory processing. Much of his argument stems from what he himself acknowledges to be an "egregious . . . lack of knowledge of the literature on child and adult speech perception and production" (p. 67). But he has also been overly impressed by those well-known cases in which a child knows that she is saying, for example, [fIs], when she should be saying [fIʃ]. This, of course, is what we would expect if learning to speak entailed the gradual marshaling of subtly interleaved motoric structures so as to capture the delicacies of dialect.

Perception and action

In fact, perhaps the most striking achievement of the child in learning to speak is that he learns to reconstruct the language of his community with such precision. One is not surprised that mothers begin to exaggerate their articulation, clarifying their phonetic execution, just when the child begins to utter its first words (Malsheen, Vol. 2, chap. 9), nor that a Spanish child learning English as a second language will display an appropriate shift of a few milliseconds, away from the Spanish and toward the English boundary, in judgments of a VOT continuum (Williams, Vol. 2, chap. 10). Perception has evolved to control action (and action to control perception). There is no sound reason to believe that the evolution of language has led to their divorce.

In conclusion, what do these volumes lack? Nothing, I think, except perhaps a chapter on the pre-speech development and communicative use of prosody. Allen and Hawkins (Vol. 1, chap. 12) do, in fact, provide a thorough review of a sizable literature on the development of syllable stress and rhythm, as well as a report of their own research on syllabic weight and accentuation in 3- and 5-year-olds. And Clumeck (vol. 1, chap. 13) reviews the acquisition of tone in Thai and Mandarin Chinese, showing that pitch begins to be used for lexical contrast only when the child begins to use words modeled on the adult language. What we miss among the chapters on pre-speech is some account of the infant's first attempts to communicate, and of the gradual differentiation of segmental from suprasegmental utterance.

Nonetheless, these volumes provide a solid review of an increasingly complex field with deep implications for our understanding of the biological bases of speech and language. The editors are to be congratulated on collecting a group of essays that will certainly influence the direction of research in the field during the coming decade.

References

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. Speech perception in infants. *Science*, 1971, *171*, 303–306.

Jakobson, R. [Child language, aphasia and phonological universals] (A. Keiler, Trans.) The Hague: Mouton, 1968.

Templin, M. *Certain language skills in children*. Minneapolis, Minn. University of Minnesota Press, 1957.