Signed and Spoken Language: Biological Constraints on Linguistic Form, eds. U. Bellugi and M. Studdert-Kennedy, pp. 3-12. Dahlem Konferenzen 1980. Weinheim: Verlag Chemie GmbH.

Introduction

M. Studdert-Kennedy* and U. Bellugi**
*Queens College, City University of New York, Flushing, NY 11367 and Haskins Laboratories, New Haven, CT 06510
**The Salk Institute for Biological Studies
La Jolla, CA 92037, USA

INTRODUCTION

Speech is the natural medium of language. Specialized structures and functions have evolved for spoken communication: vocal tract morphology, lip, jaw and tongue innervation, mechanisms of breath control, and perhaps even matching perceptual mechanisms. Moreover, language processes are controlled by the left cerebral hemisphere in over 95% of the population, and damage to certain areas of the left cortex, adjacent to the primary auditory area or to motor areas that control muscles important for speech, may induce aphasic syndromes that do not follow from damage to corresponding areas of the right. Such facts demonstrate that humans have evolved structures and physiological mechanisms adapted for communication by speech and hearing.

However, new questions are forced on us by the recent discovery that there are primary visual-manual languages that have taken their own course of development as autonomous languages, yet nonetheless share many key grammatical properties with spoken languages (4). What properties of language are mere consequences of the modality in which it has evolved, and what properties are essential to linguistic expression in any form? Surely, the linearity of spoken language, for example - that is, the sequencing of segments over time - follows naturally from its use

of sound. But can the same be said of the segments themselves? Of the dual pattern of lexical form and syntax, commonly cited as a distinctive property of language? And what of the hierarchical structure of language? Of its varied array of structure-dependent rules in phonology and syntax?

Are such properties even purely linguistic? This is the question asked by Fromkin and Klima in their introductory chapter, on which, they suggest, comparisons between signed and spoken languages may throw light. Presumably, whatever properties survive the instantiation of language in other modalities must be either general cognitive properties or special supramodal properties, unique to language. One goal of the following chapters is to isolate candidates for such properties and to see how their surface form is shaped by their modality of expression.

For example, Studdert-Kennedy and Lane ask how the different modalities of sign and speech have forced the formational structure of language into different patterns. They propose that the sublexical structures of both systems serve to match a limited set of signaling devices to the cognitive demand for an essentially unlimited lexicon. They then compare the parallel and serial structures of the two kinds of languages, arguing that, while the oral-auditory mode favors temporal sequence for purposes of linguistic contrast, but simultaneity for smooth and rapid execution, the manual-visual system favors just the reverse. They attribute the different lower-level structures of the two languages to their different motor and perceptual systems.

Similarly, Bellugi (whose chapter surveys the main structural properties of American Sign Language (ASL) at the lexical and phrasal levels) examines the effect of gestural mode not only on the structure of basic lexical signs, but also on the grammatical devices for linking them: where speech prefers to concatenate, sign prefers to conflate, the one serial, the other parallel. In an ASL utterance, a sign stem, composed of simultaneous hand

configuration, place of articulation, and movement is displayed at the same time as its modulating inflection (itself composed of simultaneous orthogonal dimensions of movement in space) and at the same time as other grammatical devices, such as facial gesture and spatial indexing. Thus, the structural components of ASL are displayed concurrently as multidimensional, layered configurations in space, while those of speech, at a comparable level of analysis, form a sequence of independent segments distributed over time. Noting that ASL signs on average take twice as long to form as English words to say, Bellugi suggests that the contrasting serial and parallel structures of speech and sign may serve to match peripheral performances to a central proposition rate common to both languages.

For Bellugi the serial-parallel difference is a pivotal difference between the surface structures of spoken and signed language. But Levelt suggests that this difference does not carry through into processing. He points out that what is sequential in the processing may be parallel in the structure, and vice versa. He argues that as processing moves from the sign or word to the phrase, clause, and discourse, the issue becomes increasingly blurred. On each serially added item, in both sign and speech, information from several sources is brought to bear in parallel: both languages evidently make simultaneous use of different sources of information in the interpretation of a given constituent. Of course, this fact cannot dissolve the differences in surface form between signed and spoken language. What Levelt emphasizes, rather, is that differences fade (as indeed they do across spoken languages) when our attention shifts from periphery to center.

Nowhere is this more evident than in acquisition. At the surface, signed languages differ radically from spoken languages. Since their gestures are formed in the very space to which they typically refer, their representational scope is far richer than that of speech: the signer is always free to break into "mimetic depiction," and he often

does. One might have expected the young deaf child to take advantage of the potential for analog representation offered by the visual-gestural mode. However, Newport and Supalla show that this is precisely what he does not do. Rather than imitate a complex analog movement, the child breaks it into stylized, discrete components, making the gesture learnable by a simplifying analysis. Acts of learning within and across the generations, these authors suggest, imprint the analytic character of the learning process on the adult language. Slobin sketches a similar process for spoken languages, describing the diachronic oscillation between simple one-to-one mappings over meaning and surface forms, preferred by language learners, and the relatively opaque mappings, forced by the demand for rapid, automatic processing among proficient speakers. What both these papers suggest is that systematic comparison of the acquisition of signed and spoken languages may contribute not only to the development of learnability theory, discussed by Wexler, but also to our understanding of the origins of language structure.

Yet another area to which comparative studies of acquisition may contribute is cerebral localization of function. As Zaidel demonstrates in his extensive review, we are still unsure what the left hemisphere is specialized for. Is it specialized for motor functions? For abstract phonological and syntactic processes? For some general cognitive style of analysis and feature extraction? Moreover, although the regularity of language onset and stages of language acquisition suggest that the developmental trajectory of the underlying neural processes is highly constrained, we know very little about its actual epigenetic course. Congenitally deaf adults who have learned a primary sign language, but no language through the auditory modality, offer a privileged testing ground for hypotheses concerning the nature of cerebral specialization and the role of experience in its development. How do characteristics of the primary language, including its transmission modality, affect its neural representation? Should we expect that representation to be similar for signed and spoken languages because they draw on some

common "pre-wired" function? Or should we expect it to be different because of their differences in modality? The relevant studies will be fraught with difficulties of measurement and interpretation, as Zaidel makes plain, but they also hold out a unique promise of progress.

Up to this point we have been discussing contributors who see the discovery of languages instantiated in different modes as an opportunity to separate the surface effects of modality from the underlying structure of language itself. We turn now to a more diverse group. On the one hand, there are those who welcome the discovery as an invitation to redo our whole approach to language in a thoroughly earth-bound way, focusing on principles of motor control and perception that must shape all the behaviors of terrestrial organisms. On the other hand, there are those who see the discovery as mere confirmation of what some linguists have already assumed, that the underlying form of language is essentially independent of its medium.

The most explicit statement of this view comes from Morton. Morton does not seriously consider the possibility that the vocal-auditory mode has influenced the evolution of underlying language capacity, since he sees no evidence that the manual-visual system is disadvantaged. Starting therefore from the premise that the two forms of peripheral apparatus are essentially equivalent, Morton proceeds to press natural selection to its limits. He argues that if we take language to be a fully adapted function, no longer in the process of evolution, then we must suppose that selective pressures have shaped it to a form fitted to its function. Whatever we take that function to be, then the necessary linguistic means to fulfill it must have evolved. Apparent limitations - as, for example, on the degree of acceptable center-embeddings or on the length of an utterance - are limitations on the non-linguistic cognitive, motoric, or perceptual systems with which language cooperates. In short, language is neither more nor less powerful than it needs to be: it is a perfectly adapted function like, say, human bipedal

locomotion or avian flight. How could it be otherwise?, Morton asks. How could natural selection have evolved a software program more powerful than the hardware that must execute it?

A position close to Morton's is reached by Marshall on quite other grounds, from a review of the neuropsychological evidence. Marshall takes the occasion to test Hughlings Jackson's view (a view founded on the associationist and evolutionary psychology of Herbert Spencer (8)) that the mind can be nothing more than "...processes representing movements and impressions." He concludes from a rich variety of evidence - most notably from double-dissociations of function across clinically-defined disorders - that "...there is more structure to language, both normal and pathological, than can be captured by sensory or motoric constraints." Marshall caps his argument with the case of a child who neither recognized nor produced speech sounds, but who learned as his first language the British equivalent of American Signed English (Paget-Gorman).

Of course, the fact that the linguistic system, as we now know it, can, in such a case, be directly engaged by a derivative writing or signing system, parasitic on spoken language, does not mean that its present form has not been determined, in greater or less degree, by the mode in which it originally evolved. Writing, fingerspelling, and Signed English do, after all, share properties of discreteness and linearity with a primary natural language. Perhaps, as Levelt suggests, these are precisely the properties that language acquired from its evolution in a vocal-auditory mode.

Alternatively, as Pattee argues in a broadly speculative paper, "discrete, rate-independent, linear symbol strings" may be the ultimate, impenetrable property of all biologically realized symbol systems. Pattee explores an analogy, first drawn by Jakobson, between the structural principles of the genetic code and of language. The analogy invites (perhaps, rather, warns of) a central distinction between natural and artificial (or

machine) systems: while artifical systems are necessarily explicit in stating the relations between a symbol and its referent, natural systems are not. The gene carries no explicit symbols for the details of the structure that it yields. Rather, it operates by "...an implicit harnessing of natural laws and structures which need no instructions." Pattee proposes this metaphor as a model of how the symbol strings, the words or signs, of an utterance may constrain the deep structures of the brain to yield meaning.

The metaphor is beguiling, but perhaps not therefore empty, when brought to bear on the surface structures of language. How this might be done is illustrated in the papers by Turvey and by Shaw and Cutting and in the ambitious discussions of Group I, focused by Summerfield. Turvey believes, with Morton and Marshall, that language is modality free, but he has very different reasons. For Turvey (as for Herbert Spencer and Hughlings Jackson) language is activity: its structure must inevitably "...reflect the organizational style that characterizes the control and coordination of acts" - a style common, presumably, to every modality of expression. We will advance our understanding of linguistic form only when we abandon the attempt to characterize it as unique, autonomous, formal, and adopt a more physicalistic "style of inquiry," as exemplified in recent approaches to the theory of action and the organization of motor control. Turvey's paper (like Shaw and Cutting's and like the ensuing discussion) is not easy. Some readers - particularly those teethed on the subtleties of syntax or phonology - may find these papers at best irrelevant, at worst irksome and overly protreptic. However, they make no pretence to address linguistic issues directly. Rather, their goal is to reframe our approach to the immediate problems of motor control and perception, under which, it must be conceded, many problems of speaking and signing are subsumed.

For example, a central notion of action theory (developed and applied to speech in (3)) is that we can resolve many of the paradoxes of standard approaches to motor control, if we abandon the dualistic notion of a central "higher level" control system and

a "lower level" executive. We do not have to suppose that the detailed "features" of a motor act, or gesture, are specified at the center. The supraspinal role is, rather, that of an organizer, marshaling groups of muscles (coordinative structures) for an action, but leaving the details of that action free to emerge from a combination of non-muscular forces (e.g., gravity, friction) and the current state of the muscles. In other words, features are not the hypostatized entities of linguistic theory, but, as Turvey remarks, the "...necessary consequences of distributed physical principles" (cf. (2)).

A complementary notion from the theory of event perception is developed by Shaw and Cutting. Many of the apparent paradoxes of perceptual theory (for example, the invariant speech segments that we perceive in the face of a variable and continuous acoustic signal) may be resolved, if we abandon the dualistic notion of a proximal array and an incommensurate perceptual experience. We are inevitably drawn into paradox as soon as we "explain" the percept by attributing function (e.g., feature extraction) to its descriptive predicates. Perception cannot create structure. Rather, we, as perceivers, are tuned to a structure that exists and the scientist's task is to specify that structure correctly. The reader will find in Shaw and Cutting's difficult, but richly promising, account of event theory and its goals many hints of possible novel approaches to both signed and spoken language. We will not presume to guide the reader further.

Indeed, we have only gone so far in illustrating the complementary approaches of action and event theory, because the feature example enables us to draw attention to a point that may easily be missed, and that may disarm the captious linguist or, at least, encourage him to suspend his disbelief while reading these chapters. This is the fact that the posture of Turvey and Shaw and Cutting is fully compatible with that of several current writers on the role of phonetics in linguistic theory (e.g., 5-7). These writers, frustrated by the seeming incommensurability of phonological theory and phonetic fact - particularly, by the lack of isomorphism among the units of linguistic analysis, of articulation

and of the acoustic signal - have proposed that it is time for students of production and perception to develop their own descriptive units. Of course, as Fowler and her colleagues (3) point out, this could only exacerbate the problem, if the outcome were not ultimately a rapprochement between theory and observation.

However, it became clear during the discussions of Group I that Lindblom and his colleagues (e.g., (1)) have for some years been pursuing a project close in spirit to action-event theory. The project is summarized in the Summerfield et al. report and we will not go into detail here. But its goal, simply put, is to show how the featural structure of vowel (and, in due course, other phonetic) systems can be derived from perceptual and articulatory constraints. More generally, Lindblom (6) has emphasized that explanatory theory must refer "...to principles that are independent of the domain of the observation themselves" and has urged that phonetic theory "... move [its] search for basic explanatory principles into the physics and physiology of the brain, nervous system and speech organs..."

Such an approach starts from the assumption that the form of language must reflect general physical, motoric, and perceptual constraints, and, more narrowly, the constraints of the particular apparatus that language has found a way to use. That these last constraints are likely to have played a role in shaping linguistic form seems obvious as soon as we try to imagine a language based on, say, smell. And once we concede that not every sensorimotor system seems equally effective as a medium of linguistic expression, it becomes appropriate to ask what similarities of oral-auditory and manual-visual modes have fitted them to language and how their differences may force language into different surface forms. The instantiation of language in a new modality may then permit us to strip off whatever is contingent and modality-bound, to discover language starkly set at the intersection of sign and speech.

REFERENCES

(1) Bladon, R.A.W., and Lindblom, B.E.F. Modeling the judgment of vowel quality differences. J. Acous. Soc. Am., in press.

(2) Cooper, F.S. 1972. How is language conveyed by speech? In Language by Ear and by Eye, eds. J.F. Kavanagh and I.G. Mattingly, pp. 25-45. Cambridge, MA: The MIT Press.

(3) Fowler, C.A.; Rubin, P.; Remez, R.E.; and Turvey, M.T. 1980. Implications for speech production of a general theory of action. In Language Production, ed. B. Butterworth, pp. 373-420. New York: Academic Press.

(4) Klima, E.S., and Bellugi, U. 1979. The Signs of Language. Cambridge, MA: Harvard University Press.

(5) Lindblom, B. 1972. Phonetics and the description of language. In Proceedings of the 7th International Congress of Phonetic Sciences, eds. A. Rigault and R. Charbonneau, pp. 63-92. The Hague: Mouton.

(6) Lindblom, B. 1980. The goal of phonetics, its unification and application. Language and Speech 23.

(7) MacNeilage, P., and Ladefoged, P. 1976. The production of speech and language. In Handbook of Perception, eds. E.C. Carterette and M.P. Friedman, vol. 7, pp. 75-120. New York: Academic Press.

(8) Young, R.M. 1970. Mind, Brain and Adaptation in the 19th Century. Oxford: Clarendon Press.