In: Knight, C., Studdert-Kennedy, M. & Hurford, J.R. (eds.) The Evolutionary Emergence of Language. Cambridge: Cambridge University Press.

1188

10

Evolutionary Implications of the Particulate Principle: Imitation and the Dissociation of Phonetic Form from Semantic Function

MICHAEL STUDDERT-KENNEDY

Introduction

At least three unique properties distinguish language from other systems of animal communication: unlimited semantic scope, freedom from control by identifiable external stimuli (displaced reference), and transduction into alternative perceptuomotor modalities (writing, fingerspelling). All three properties, it will be argued, depend on dissociating phonetic form from semantic function. Such a dissociation arose with the emergence of vocal imitation, a necessary condition of the protolanguage that evolved when our hominid ancestors chanced on 'the particulate principle of self-diversifying systems' (Abler 1989). This is a physical principle to which all natural systems that, in Humboldt's (1836/1972: 70) famous phrase, 'make infinite use of finite means' (physics, chemistry, genetics, language) necessarily conform. In such systems, discrete units from a finite set of meaningless elements (e.g. atoms, chemical bases, phonetic segments) are repeatedly sampled, permuted and combined to yield larger units (e.g. molecules, genes, words) that are higher in a hierarchy and both different and more diverse in structure and function than their constituents.

The particulate principle rationalises both the hierarchical structure of language and the discrete combinatorial mechanisms on which the hierarchy is raised. The principle has many implications for the evolution of language. For example, it casts doubt on the likely communicative scope of any prelinguistic symbolic system, limited to a purely analog representation of the world (e.g. Donald 1998). And the necessity for hiorarchical organisation discourages the notion that syntax might have emerged before the combinatorial mechanisms of phonology were well established (Bickerton 1998: 344). My concern in what follows, however, is with implications of the particulate principle for the separation of phonetic form from semantic function, and for the emergence of an independent level of phonetic representation, preadaptive not only for displaced reference and syntax, but also for the very much later cultural development of writing.

We begin by briefly considering the language hierarchy as it has generally been conceived in recent decades.

The Nature of the Language Hierarchy

Duality of Patterning

Fifty years ago André Martinet (1949) coined the phrase 'la double articulation', to describe the two-level hierarchy of phonology and syntax that characterises all human languages. Since then, 'duality of patterning', as Hockett (1958) termed the concept, has come to be generally recognised as critical to the unbounded semantic scope of language. At the lower level of the hierarchy, phonology evades the limits of a finite vocal apparatus by permuting and combining discrete articulatory actions to construct an unbounded lexicon of words. At the higher level, syntax permutes and combines words to represent a potential infinity of relations among objects, events and concepts. Despite general agreement on these functions, the questions of whether or why duality is necessary for language, why constituents are discrete and why they form a hierarchy of independent levels, were first addressed only a decade ago by William Abler (1989). The starting point for Abler's work was the relation between genetics and language.

Genes and Language

The first to see a possible parallel between the gene and language was the Austrian physicist Erwin Schrödinger (Jacob 1977; Pollack 1994). A decade before the structure of the DNA molecule was described, Schrödinger (1944) correctly predicted that the gene would prove to be a molecule with the structure of an aperiodic crystal. The molecule would offer, through rearrangements of its component atoms, a range of isomeric variants analogous to those offered by rearrangements of dots and dashes in Morse code. The vast combinatorial resources of the gene's 'code-script' (Schrödinger 1944: 62) would then suffice to specify all the diverse lines of development of every structure in the living world.

Since Schrödinger wrote, the analogy between gene and language has become a textbook commonplace (e.g. Pollack 1994), and a topic of papers by both linguists (e.g. Jakobson 1970) and biologists (e.g. Jacob 1977), although satisfactorily explained by none. Jakobson, for example, remarked: 'among all the information-carrying systems, the genetic code and the verbal code are the only ones based upon the use of discrete components which, by themselves, are devoid of inherent meaning but serve to constitute the minimal senseful unit, i.e., entities endowed with their own intrinsic meaning in the given code' (1970: 438). But how and why the parallel arose are questions he left unanswered.

Jacob went further, recognising that the principle of combining discrete units is 'not limited to language and heredity . . . [but] . . . appears to operate in nature each time there is a question of generating a large diversity of structures using a restricted number of building blocks' (1977: 188). Writing of language and heredity, Jacob also emphasised, that 'for such a system to function implies that the basic units, phonemes or chemical radicals, are by themselves devoid of meaning' (ibid.). As we shall see, this fact is crucial for understanding the early evolution of language.

Yet even Jacob stopped short of explaining why, as he wrote, 'Such a method of construction appears to be the only logical one' (1977: 188). For a particulate genetics Fisher (1930) had already supplied the logic. Fisher reasoned that, if characteristics of parents blended, they would be lost in the average of their offspring, and the characteristics of the offspring would lie between, not outside, those of their parents; variation, critical to the process of natural selection, would then diminish from generation to generation. In fact, of course, variation is conserved, or even increased, across generations, and parental characteristics may reappear unmodified in their descendants. From such facts, Fisher (like Mendel before him) inferred that biological inheritance was effected by a particulate mechanism, an inference confirmed by the description of the DNA molecule some twenty years later. Fisher's logic was the impetus for Abler's (1989) independently developed account of the gene-language parallel, which subsumes both systems under a single physical principle.

The Particulate Principle of Self-Diversifying Systems

Abler's key insights were two. First, Fisher's arguments concerning the mechanism of inheritance could be extended to the hierarchical structures of physics, chemistry and language. All four domains have an unbounded range of properties, and so 'must be based on particles rather than on blending constituents, because blending constituents would form combinations whose properties lie between rather than outside, the properties of the original constituents' (Abler 1989: 1). Here, it is to sustain variation at each level of the hierarchy that the units combined must be discrete.

Abler's second insight was that von Humboldt's (1836/1972) characterisation of language could be extended to the hierarchical structures of physics, chemistry and genetics. All four domains exploit discrete combinatorial mechanisms to 'make infinite use of finite means' (p. 70) by 'a synthetic process … [that] … creates something … not present *per se* in any of the associated constituents' (p. 67). It is this 'creativity', this automatic emergence of novelty, that Abler terms 'self-diversifying'. Here, then, it is for the emergence of novel structures and functions at each level of the hierarchy that the units combined must be discrete. Only because units do not blend and disappear, but combine as integral units to form new integral units, can novel structures arise whose properties are not limited by, and cannot be predicted from, the properties of their constituents.

'We cannot derive the properties of common salt from those of sodium and chlorine, nor of a protein from the genes that control its formation; in language we cannot derive the meaning of a word from the phonetic elements that compose it, nor the meaning of a proposition from the lexical meanings of its words without regard to their syntactic grouping' (Studdert-Kennedy 1998: 204). Because units cannot be derived from their constituents, each level of structure is subject to new and characteristic rules of combination, giving rise to a separate independent level of function. Hence the independence of syntax and phonology.

For a fuller account of the particulate principle than is possible here readers may turn to Abler (1989) and Studdert-Kennedy (1998). For present purposes, perhaps enough has been said to make clear that duality of patterning is not a unique cognitive property peculiar to language, but reflects rather a general physical and mathematical property of the natural world to which all self-diversifying systems necessarily conform. Under the particulate principle the axioms of earlier formulations of duality – combinatorial mechanisms and independence of successive levels in the hierarchy – receive a rational explanation.

The Dissociation of Sound and Meaning

Nonetheless, despite striking correspondences across domains, language differs from its particulate congeners in several respects. Most important are differences in the relations between form and function. For example, although we may not be able to predict the properties of a chemical compound from its component elements, or the functions of a gene from the sequence of its base pairs, these properties and functions evidently emerge from, and are in some sense intrinsic to, the physical systems of which they are parts: structure and function are inseparable, if not identical. In language, by contrast, form and function arise from distinct and incommensurate sources. The meaning of a word, or of a sentence, does not emerge from phonetic or syntactic structure as, say, water

'emerges' from the structures of hydrogen and oxygen. Rather, meaning is assigned by some arbitrary, extralinguistic process within a language community. Thus, while the particulate principle affords the hierarchical *structure* necessary to support the unlimited semantic scope of language, semantic *function* itself enters the system from outside.

The dissociation of sound and meaning has no precedent in other animal vocalisations, whose signal inventories are limited and not subject to cultural modification. The dissociation is, in fact, the critical discontinuity that separates human language from other primate systems of vocal communication – critical because, as we have seen, meaningless units at the base of a hierarchy are essential to operation of the particulate principle in all its domains. In language, it is only if they are meaningless that the same units can be repeatedly permuted and combined to form different units of meaning. And only because the basic units are meaningless can the meanings assigned to their combinations be arbitrary – as required for a lexicon of unbounded semantic scope.

The key evolutionary questions then are: What are the basic units, and how did they come to be without meaning?

Units and the Vocal Mechanism

Gestures

Several authors have proposed that the basic units of language are neither phonemes (consonants and vowels) nor their descriptive features, but gestures, such as those adopted by the theory of articulatory phonology (Browman and Goldstein 1986, 1992; cf. Bell 1911; Fowler 1996; Liberman and Mattingly 1985, 1989; Lindblom 1992, 1998; Saltzman and Munhall 1989; Studdert-Kennedy 1987, 1991, 1998; Studdert-Kennedy and Goodell 1995). Gestures are the irreducible, objectively observable units of phonetic action and perception, from which segments and syllables are formed. In the framework of articulatory phonology, a gesture is a constriction formed by lips, tongue, velum or larynx, at a certain point in the vocal tract. The function of a gesture is to shape a vocal tract configuration, controlling the flow of air so as to produce a characteristic pattern of sound. How did such gestures come to take their place at the base of the language hierarchy?

From Syllable to Gesture to Segment

Elsewhere I have sketched a possible line of phonetic evolution, modelled on the ontogeny of speech development over the first years of life (Studdert-Kennedy

1987, 1991, 1998; see also Lindblom 1992, 1998; MacNeilage and Davis, this volume). Briefly, the model takes the protosyllable, a single cycle in the repeated closed-open cycles of the mandible in a primate call, as the original unit of both articulatory action and meaning (cf. MacNeilage 1998a, 1998b). Under pressure for lexical diversity (and, as will be argued, for vocal imitation), gestures emerged from the holistic protosyllable by differentiation of the vocal apparatus into a coordinated system of more or less independent articulators; gestures were selected for acoustic distinctiveness and articulatory economy, including ease of combination and sequencing (Carré and Studdert-Kennedy 1998; cf. Lindblom 1986).

As the lexicon grew, segments (consonants and vowels) emerged to facilitate rapid, successive activation of recurrent multigestural routines within the syllabic frame. Such routines arose because certain combinations of the limited set of physiologically possible laryngeal and supralaryngeal gestures repeatedly occurred either simultaneously (as in the labial/lingual and glottal gestures for voiceless fricatives) or very closely overlapping in time (as in the labial/lingual and glottal gestures for voiced and voiceless stops, or the labial and lingual gestures for rounded vowels). Separate instructions for their simultaneous or rapidly successive and overlapping activation were then readily superseded by an integral instruction for the entire gestural constellation (i.e. segment) of which they were a part. (For fuller discussion of segments and gestural coordination, see Fowler 1996: 530ff.).

Critical to this posited evolutionary sequence and to the emergence of the gesture-segment-syllable hierarchy at the meaningless base of language was the evolution of vocal learning, or imitation. But before we come to this we must consider the vocal mechanism itself.

The Vocal Mechanism

We should not underestimate the complexity of the vocal mechanism nor the importance of its idiosyncratic properties for the evolution of language. In our preoccupation with the challenge of syntax we may be tempted to take evolution of the vocal mechanism and of its extraordinary motor facility for granted. Carstairs-McCarthy (1998: 291), for example, assumes that descent of the larynx (whether a side effect of bipedal posture or, as Lieberman (1984) has argued, directly selected for increased phonetic scope) sufficed 'to produce the full range of modern articulatory possibilities'. But 'the full range' is not merely a matter of an enlarged pharynx and the novel configurations it affords.

If we look at fluent speech as a purely motor skill, we note several salient characteristics. First is the rhythmic cycle of syllables. Cyclicity characterises all motor activities, most obviously locomotion; in speech the metrical beat of the syllables rests on finely modulated control of the respiratory apparatus, a degree of control apparently absent in other primates (Lenneberg 1967: ch. 3; MacLarnon and Hewitt 1998). Other characteristics are speed and precision: a typical rate of fluent speech (between pauses) is of the order of 6–7 syllables, or 10–15 segments, per second (Liberman et al. 1967). There would be nothing remarkable here perhaps, if the movements were repetitive – a musician's tremolo finger movements can reach 16 Hz, and hummingbirds beat their wings at over 70 Hz. Speakers, however, execute different movements of different articulators within a syllable, and different patterns of movement from syllable to syllable, all precisely phased and coordinated within a tolerance of millimetres and milliseconds.

A final characteristic of speech as a motor activity is that the form and sequence of its movements are arbitrary and internally controlled. Again, if the movements were guided by the environment, as in the swift manoeuvres of a downhill skier, or of a bird through the branches of a tree, we might find nothing extraordinary in speech. But as it is, the rapidly changing patterns of articulation, carried on the brisk beat of the syllable and on the slower rhythm of the respiratory cycle, all under endogenous control, call for a specialised premotor planning mechanism perhaps unique among motor systems. Evolution of this system, including a capacity for short- and long-term phonetic memory, must have entailed sustained selective pressure over many generations, not only for a growing lexicon, but for increasingly lengthy word strings.

We turn now to a crucial, though largely unrecognised, factor in the evolution of the speech motor system, the capacity for vocal imitation.

Imitation

The role of imitation in language acquisition has often been discounted, because, in learning to construct sentences, children clearly do more than repeat what they have heard: they extract and apply rules. They can do so, however, only because they have previously built a repertoire of words by copying their companions. The preservation of local dialects from generation to generation, not only in the words chosen for use, but in their characteristic phonetic variants, attests to the accuracy of the copies. Here I shall argue that imitation was the critical factor not only in building a lexicon, but in precipitating the dissociation of phonetic form and semantic function.

To imitate, say, a facial expression or an utterance of a conspecific is to transduce an optic or acoustic pattern into muscular actions that yield an acceptable replica of the model. No doubt the process entails empathetic recognition of a conspecific's acts as potentially one's own by means of some supramodal representation of body organs, such as Meltzoff and Moore (1997) propose in their model of facial imitation. But how, in physiological fact, do light and sound get into the muscles (Michael Turvey, in conversation)?

Studies of macaques have isolated brain cells that respond to social events, including facial displays and locomotion by a conspecific or a human (e.g. Brothers, Ring and Kling 1990). Recently, Rizzolatti and his colleagues (Rizzolatti et al. 1996) have found so-called mirror neurons in macaque cortex that fire not only when a monkey grasps or manipulates food, but also when it sees a human experimenter do the same. Moreover, these neurons lie in an area of macaque cortex arguably homologous with human Broca's area. Rizzolatti and Arbib (1998) review data from transcranial magnetic stimulation and positron emission tomography studies that demonstrate a mirror system for manual grasping in humans. They postulate 'a fundamental mechanism for action recognition' (Rizzolatti and Arbib 1998: 190) in both monkey and human, and discuss the implications of such a system for the evolution of manual gesture and speech.

Here we consider what a specialised capacity for vocal imitation implies behaviourally and cognitively for the evolution of speech and language.

Vocal Imitation

The value of imitation to a member of a social species is not hard to imagine. Yet few species imitate. The apes, and some monkeys, perhaps copy general bodily actions, but vocal learning is peculiar to a few species of songbirds, certain marine mammals, and humans (Hauser 1996).

We may never know when, or how, human vocal imitation began to evolve. We should perhaps expect it to have been a factor in, and to have coevolved with, differentiation of the vocal apparatus. Certainly, we cannot doubt that imitation was critical to the early evolution of language. Without it there could have been no shared vocabulary of learned meanings, and so no basis for the evolution of propositional utterance and syntax. As we shall see, other properties preadaptive for syntax, for displaced reference and for the cultural development of writing are also implicit in the process of imitation itself.

The Process of Vocal Imitation and the Loss of Meaning

When a child learns to say a word, it finds in the acoustic patterns of its companions' speech detailed specifications for a corresponding pattern of motor organisation. Imitating a word entails (conceptually) at least three steps: (1) analysis of the sound pattern into its underlying articulatory components,

(2) storage of the analysed structure for a shorter or longer period, depending on the delay between model and copy and (3) reassembly of the components in their correct temporal sequence.

Notice that the meaning of the word plays no part. The desire to communicate meaning may motivate vocal imitation, but meaning contributes nothing to the act itself. Imitation is a purely formal process that temporarily strips the model of whatever semantic function it may have, reducing it to its bare perceptuomotor components. Here then in the act of imitation, we may reasonably surmise, is where structure and function first came to be dissociated in hominid communication. We will return to this point.

First we should note that breaking a word into its articulatory components does not automatically set those components free for independent use in another word. In fact, early words in the speech of modern children often seem to be holistic patterns, quite unrelated to one another. For example, a young child who says *no* correctly may substitute [m] for [n] in *night*, and [b] for [m] in *moo* (Ferguson and Farwell 1975); or a child who says [n∧t] correctly in *doughnut* may say [p∧p] for *nut* in *peanut* (Studdert-Kennedy and Goodell 1995). Thus, gestures firmly executed in one context are not necessarily ready for use in another.

A striking example of motoric analysis without independent control of the analysed components comes from a deaf boy learning American Sign Language (ASL) (Akamatsu 1985). His deaf mother, lacking an ASL sign for the name of a supermarket, 'Safeway', habitually fingerspelled it as a rapid string of seven hand configurations corresponding to the English letters of the name. The boy learned to recognise the string as an integral sign, and to reproduce it; only when he was nine years old and was himself learning to fingerspell did he realize that the sign he had learned was, in fact, a string of discrete, but coarticulated handshapes (cf. Armstrong, Stokoe and Wilcox 1995: 106ff.) Despite habitual use, the handshapes had not escaped from their context.

Early acts of hominid vocal imitation may then have been holistic, or formulaic, routines. Target utterances were perhaps analysed, stored and replicated as a sequence of gestures, but they were not yet represented as a sequence of independent phonetic elements that could be marshalled for use in other contexts (cf. Wray, this volume). At this stage of evolution (as in development) phonetic structure was merely implicit in the mechanism of imitation.

The Emergence of an Independent Level of Phonetic Representation

We do not know the stages by which the vocal tract gradually differentiated into the six active articulators of modern languages (larynx, tongue root, body and blade (or tip), velum and lips). But it is evident that emergence of the gesture as a commutable unit follows naturally from the limited potential of the vocal apparatus. Each articulator has relatively few motorically and perceptually discriminable states. Moreover, none of the articulators works alone; several are active and all are at least passively engaged in every utterance (Mattingly 1991). Speakers could therefore build a sizeable repertoire of words, or formulaic utterances, only by enlisting the same articulator for the same gesture again and again in different contexts. As the number of different contexts for the same gesture increased, speed and economy of motor organisation precipitated the recurrent element as a unit of motor control, independent of its context (Lindblom 1992, 1998; Studdert-Kennedy 1987, 1991). As crystallisation of recurrent gestures (and of recurrent constellations of gestures, or segments) pervaded speakers' lexicons, there emerged a new level of representation, intermediate between signal and message.

We have been describing the elements of the newly emerged level in motor terms, on the assumption that, as argued elsewhere (Studdert-Kennedy 1998), it was through particulation of the vocal machinery that gestures first arose; but it is through their acoustic and/or optic effects that a listener recognises a speaker's gestures. The intermediate representation must therefore have been isomorphic with both act and percept and to that extent abstract; the representation must also have been specific to speech, that is, phonetic. The consequences of a new independent and abstract phonetic level, induced by the evolution of vocal imitation, ramified through much of what eventually became language.

Phonetic Memory: A Preadaptation for Displaced Reference and Syntax

In the early stages of the evolution of vocal imitation (as in the early stages of its development in the modern child) the phonetic representation of an utterance was perhaps stored little longer than necessary to execute the copy, and so little longer than the event that elicited the utterance. Lengthening the store, so that the phonetic form of a word, once learned, became a permanent entry in a speaker's vocal repertoire, had the obvious advantage of enabling a speaker to reproduce a word without a model. With the separation of a word from its model, and so from the original occasion of its use, its phonetic form became freestanding, as it were, and open to arbitrary shifts in meaning. Arguably, then, the separation of form from meaning, an incidental consequence of imitation, was the first step away from vocalisations expressing current emotional state towards vocal units ('words') of arbitrary and displaced reference. Certainly, an independent phonetic memory is necessary, though obviously not sufficient, for a sizeable vocabulary of displaced reference.

Both short- and long-term phonetic memory were also essential preadaptations for syntax. Long-term phonetic memory was necessary both to formulate and to understand the meaning of a syntactically organised utterance. Short-term, or working, phonetic memory was necessary, in speaking, to hold upcoming words in premotor store for rapid execution. Ample evidence for such a store has come from spoonerisms and other speech errors where single words, syllables, segments or gestures are exchanged or misplaced. In listening, working phonetic memory was necessary to hold words, without final commitment to meaning, while computing their syntactic relations. Without a preadapted system for storing phonetic structure, independently of its meaning, syntax could not have begun to evolve. But what exactly do we mean by phonetic structure? Where do the constraints that distinguish a word from a random string of gestures come from?

The Syllabic Origins of Phonetic Structure

Gestures and segments are not so abstract as to have entirely lost their moorings in the syllable from which they are derived: the major phonetic classes, consonants and vowels, are in fact defined by their role in controlling the flow of air over the course of a syllable. Consonants, formed by complete or narrow constrictions at the onset or offset of a syllable, impede the flow of air; vowels, the nucleus of the syllable, allow air to flow freely. Implicit therefore in the repeated opening and closing of the vocal tract over the course of an utterance is a pattern of gestural alternation that specifies the boundaries of possible words in the sequence. Thus, in evolution, the syllable or syllable string, the integrative articulatory 'frame' of which meaningless gestures and segments came to be the 'content' (MacNeilage 1998b), retained its initial status as the basic unit of meaning by virtue of the restrictions it imposed on segmental sequence.

To make this clear, consider a random number table where, by definition, we cannot predict a number, or any string of numbers, however many preceding numbers we know. Such a table is devoid of structure because all possible sequences of numbers can occur. Suppose, however, that we compose the table from an English discourse, substituting a pair of digits from 01 to roughly 40, for each phoneme. The table will now be structured because not all possible sequences of numbers can occur: the sequences will reflect restrictions on number combinations imposed by the syllable structure of English. Skilled cryptanalysts would now quickly see that the distribution of numbers was not random and, if they knew that the coded text was English, they could apply the rules of English syllable structure (and, of course, their knowledge of English words) to recover the discourse. (It was, in fact, by assuming that the language was

Greek, and the unit of transcription the syllable, that Michael Ventris deciphered Minoan Linear B (Chadwick 1958).)

My point here is that a word is defined in the first instance by its phonetic structure, secondarily by its semantic function. Definition by meaning is secondary, because meaning is unstable and varies with context. Phonetic structure, by contrast, is invariant (within the tolerance of diachronic and synchronic variation), because it arises from restrictions on the combination of gestures imposed by the biophysical structure of the syllable (MacNeilage and Davis, this volume).

Let me emphasise that I am not suggesting that all phonological constraints arise as phonetic constraints from the syllable. My point is merely that in evolution the vocal machinery itself was the only possible initial source of constraint on combinations of prephonological gestures and segments. Subsequently, of course, a diverse array of universal phonetic and language-specific phonological constraints arose, from other aspects of linguistic function, to supplement the universal rubric of the syllable.

Whatever the source of phonetic and phonological constraints, it is these constraints, not meaning, that determine whether a string of segments is, or is not, a possible word in a particular language. Recognition of this fact, many thousands of years later, was a critical step in the cultural development of writing and reading.

The Particulate Basis of Writing and Reading

Language is unique among systems of animal communication in its transduction into an alternative perceptuomotor modality. I use 'transduction' rather than 'translation' because my concern here is with writing, and writing is not an independent natural language, like a sign language, that happens to use hand and eye instead of mouth and ear. Rather, writing is parasitic on speech.

The earliest forms of writing were syllabic (Gelb 1952). Here an ambiguity may arise, in a morphosyllabic language such as Chinese, where every syllable is a unit of meaning (morpheme), as to whether a written symbol refers to the syllable as morpheme or as unit of phonetic structure. The ambiguity is reflected in the widespread, but mistaken belief that Chinese characters represent meaning directly without phonological mediation (DeFrancis 1989).

Early forms of pictographic writing did indeed represent meaning directly. But with the recognition of syllabic homonyms (syllables with the same sound, but different meanings) there became available the 'rebus principle' (DeFrancis 1989: 50), by which a syllabic symbol could be used to refer not to the syllable's meaning, but to its sound. For example, in rebus writing the symbol for the

word *date* (fruit) might be used to represent the word *date* (day). By this device form and function were dissociated in writing, as they had been dissociated in speaking by the emergence of imitation.

With the invention of the rebus, therefore, true writing, capable of representing the full semantic range of language, became possible. Writing could now exploit its own form of the particulate principle with a small set of meaningless symbols, whether for syllables or for phonemes. In fact, all full writing systems (DeFrancis 1989: ch. 2) represent the phonological forms of words, with varying degrees of explicitness, depending on the language. No system represents meaning alone, or meaning without phonological mediation. From what we have argued, a purely semantic representation of any language is indeed impossible, because a phonetic structure is intrinsic to every word. Every language has phonologically permissible words that happen not to have been assigned a meaning (so-called nonsense words). No language has words that have meaning, but have not yet been assigned phonological form. Writing systems therefore differ from one another only in the phonological level (syllable or phoneme) and in the precision of their phonological representation.

Apart from the syllabaries of Chinese and Chinese-influenced languages and of a few African and Amerindian languages, alphabets derived from Phoenician consonantal orthography are the only systems in general use today (Gelb 1952: 184). The letters of an alphabet represent neither sounds nor articulations, but phonological entities (phonemes) at a higher level of abstraction than the perceptuomotor entities that we must posit to account for imitation. Each phoneme encompasses a class of phonetic variants, or allophones, that vary across phonetic contexts, dialects and even individual speakers. What is important here is that alphabetic writing represents the phonological form of a word, not its meaning. That is why we can read a text without understanding it, but cannot understand a text without reading it.

Thus, the dissociation of meaning and phonetic structure, precipitated by the evolution of vocal imitation, planted a preadaptive seed no less critical for the relatively recent cultural development of writing than for the biological evolution of displaced reference and syntax, tens of thousands of years earlier.

Summary and Conclusion

Duality of patterning, the two-level hierarchy of phonology and syntax on which the unbounded semantic scope of language rests, is a special case of the particulate principle to which all systems that 'make infinite use of finite means' (physics, chemistry, genetics, language) necessarily conform. At the base of all such systems are discrete, meaningless (intrinsically functionless) units on which, by successive combinatorial mechanisms, a hierarchy is raised.

This chapter proposes articulatory gestures as the basic units of spoken language from which phonetic segments and syllables are formed. Dissociation of phonetic form from semantic function, the critical discontinuity that separates language from all other systems of animal communication, arose as a side-effect of the evolution of vocal imitation, a capacity unique among primates to humans. To imitate a word, a speaker must analyse, store and reassemble its gestures, a purely formal process that temporarily strips the model of whatever semantic function it may have, reducing it to its perceptuomotor components.

Imitation thus gave rise to a new level of processing between signal and message, comprising a phonetic representation and a mechanism for phonetic storage. The short-term phonetic store necessary for immediate imitation proved to be a preadaptive step toward the short- and long-term phonetic memory systems necessary for displaced reference and syntax and ultimately for the cultural development of writing and reading many thousands of years later.

Acknowledgements

My thanks to William Abler, Carol Fowler, James Hurford, Alvin Liberman and Ignatius Mattingly for instructive comments on earlier versions. Preparation of the paper was supported in part by Haskins Laboratories.

References

Abler, W. 1989. On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures* 12: 1–13.

Akamatsu, C. T. 1985. Fingerspelling formulae: a word is more or less the sum of its letters. In W. Stokoe and V. Volterra (eds), *SLR'83: Sign language research*. Silver Spring, MD: Linstok, pp. 126–132.

Armstrong, D. F., W. C. Stokoe and S. E. Wilcox. 1995. *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.

Bell, A. G. 1911. *The Mechanism of Speech*. New York: Funk and Wagnalls.

Bickerton, D. 1998. Catastrophic evolution: the case for a single step from protolanguage to full human language. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches to the Evolution of Language: Social and cognitive bases*. Cambridge: Cambridge University Press, pp. 341–358.

Brothers, L., B. Ring and A. Kling. 1990. Response of neurons in the macaque amygdala to complex social stimuli. *Behavioural Brain Research* 41: 199–213.

Browman, C. P. and L. Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook* 2: 219–252.

Browman, C. P. and L. Goldstein. 1992. Articulatory phonology: an overview. *Phonetica* 49: 155–180.

Carré, R. and M. Studdert-Kennedy. 1998. The origin of speech gestures. Paper read at the Second International Conference on the Evolution of Language, University of East London.

Carstairs-McCarthy, A. 1998. Synonymy avoidance, phonology and the origin of syntax. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches to the Evolution of Language: Social and cognitive bases*. Cambridge: Cambridge University Press, pp. 279–296.

Chadwick, J. 1958. *The Decipherment of Linear B*. Cambridge: Cambridge University Press.

DeFrancis, J. 1989. *Visible Speech*. Honolulu, HI: University of Hawaii Press.

Donald, M. 1998. Mimesis and the Executive Suite: missing links in language evolution. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches to the Evolution of Language: Social and cognitive bases*. Cambridge: Cambridge University Press, pp. 44–67.

Ferguson, C. A. and C. B. Farwell. 1975. Words and sounds in early language acquisition. *Language* 15: 419–439.

Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. Oxford: Clarendon.

Fowler, C. A. 1996. Speaking. In H. Heuer and S. Keele (eds), *Handbook of Perception and Action, Vol. 2*. London: Academic, pp. 503–560.

Gelb, I. J. 1952. *A Study of Writing*. Chicago: Chicago University Press.

Hauser, M. 1996. *The Evolution of Communication*. Cambridge, MA: MIT Press.

Hockett, C. F. 1958. *A Course in Modern Linguistics*. New York: Macmillan.

Humboldt, W. von 1836/1972. *Linguistic Variability and Intellectual Development*. Translated by G. C. Buck and F. A. Raven. Philadelphia: University of Pennsylvania Press.

Jacob, F. 1977. The linguistic model in biology. In D. Armstrong and C. H. Van Schooneveld (eds), *Roman Jakobson: Echoes of his scholarship*. Lisse: de Ridder, pp. 185–192.

Jakobson, R. 1970. Linguistics. In *Main Trends of Research in the Social and Human Sciences* Vol. 1. Paris/The Hague: UNESCO/Mouton, pp. 437–440.

Lenneberg, E. H. 1967. *Biological Foundations of Language*. New York: Wiley.

Liberman, A. M., F. S. Cooper, D. P. Shankweiler and M. Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review* 74: 431–461.

Liberman, A. M. and I. G. Mattingly. 1985. The motor theory of speech perception revised. *Cognition* 21: 1–36.

Liberman, A. M. and I. G. Mattingly. 1989. A specialization for speech perception. *Science* 243: 489–494.

Lieberman, P. 1984. *The Biology and Evolution of Language*. Cambridge, MA: Harvard University Press.

Lindblom, B. 1986. Phonetic universals in vowel systems. In J. J. Ohala and J. J. Jaeger (eds), *Experimental Phonology*. London: Academic, pp. 13–44.

Lindblom, B. 1992. Phonological units as adaptive emergents of lexical development. In C. A. Ferguson, L. Menn and C. Stoel-Gammon (eds), *Phonological Development: Models, research, implications*. Timonium, MD: York Press, pp. 131–163.

Lindblom, B. 1998. Systemic constraints and adaptive change in the formation of sound structure. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches* *to the Evolution of Language: Social and cognitive bases*. Cambridge: Cambridge University Press, pp. 242–264.

MacLarnon, A. and G. Hewitt. 1998. Protolanguage to language: evidence from the evolution of breathing control. Paper read at the Second International Conference on the Evolution of Language, University of East London.

MacNeilage, P. F. 1998a. Evolution of the mechanisms of language output: comparative neurobiology of vocal and manual communication. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches to the Evolution of Language: Social and cognitive bases*. Cambridge: Cambridge University Press, pp. 222–241.

MacNeilage, P. F. 1998b. The Frame/Content theory of evolution of speech production. *Behavioral and Brain Sciences* 21: 499–511.

Martinet, A. 1949. La double articulation linguistique. *Travaux du Cercle Linguistique de Copenhague* 5: 30–38.

Mattingly, I. G. 1991. The global character of phonetic gestures. *Journal of Phonetics* 18: 445–452.

Meltzoff, A. N. and M. K. Moore. 1997. Explaining facial imitation: a theoretical model. *Early Development and Parenting* 6: 179–192.

Pollack, R. 1994. *Signs of Life*. Boston: Houghton Mifflin.

Rizzolatti, G. and M. A. Arbib. 1998. Language within our grasp. *Trends in Neuroscience* 21: 188–194.

Rizzolatti, G., L. Fadiga, V. Gallese and L. Fogassi. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3: 131–141.

Saltzman, E. and K. Munhall. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1: 333–382.

Schrödinger, E. 1944. *What is Life?* Cambridge: Cambridge University Press.

Studdert-Kennedy, M. 1987. The phoneme as a perceptuomotor structure. In A. Allport, D. G. MacKay, W. Prinz and E. Scheerer (eds), *Language Perception and Production*. London: Academic, pp. 67–84.

Studdert-Kennedy, M. 1991. Language development from an evolutionary perspective. In N. A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch and M. Studdert-Kennedy (eds), *Biological and Behavioral Determinants of Language Development*. Hillsdale, NJ: Erlbaum, pp. 5–28.

Studdert-Kennedy, M. 1998. The particulate origins of language generativity: from syllable to gesture. In J. R. Hurford, M. Studdert-Kennedy and C. Knight (eds), *Approaches to the Evolution of Language: Social and cognitive bases*. Cambridge: Cambridge University Press, pp. 202–221.

Studdert-Kennedy, M. and E. H. Goodell. 1995. Gestures, features and segments in early child speech. In B. de Gelder and J. Morais (eds), *Speech and Reading*. Hove: Erlbaum (UK), Taylor and Francis, pp. 65–88.