OXFORD HANDBOOKS IN LINGUISTICS

The Oxford Handbook of Applied Linguistics
Second edition
Edited by Robert B. Kaplan

The Oxford Handbook of Case
Edited by Andrej Malchukov and Andrew Spencer

The Oxford Handbook of Cognitive Linguistics
Edited by Dirk Geeraerts and Hubert Cuyckens

The Oxford Handbook of Comparative Syntax
Edited by Guglielmo Cinque and Richard S. Kayne

The Oxford Handbook of Compounding
Edited by Rochelle Lieber and Pavol Štekauer

The Oxford Handbook of Computational Linguistics
Edited by Ruslan Mitkov

The Oxford Handbook of Compositionality
Edited by Markus Werning, Edouard Machery, and Wolfram Hinzen

The Oxford Handbook of Field Linguistics
Edited by Nicholas Thieberger

The Oxford Handbook of Grammaticalization
Edited by Heiko Narrog and Bernd Heine

The Oxford Handbook of Japanese Linguistics
Edited by Shigeru Miyagawa and Mamoru Saito

The Oxford Handbook of Laboratory Phonology
Edited by Abigail C. Cohn, Cécile Fougeron, Marie Hoffman

The Oxford Handbook of Language Evolution
Edited by Magggie Tallerman and Kathleen R. Gibson

The Oxford Handbook of Language and Law
Edited byLawrence Solan and Peter Tiersma

The Oxford Handbook of Linguistic Analysis
Edited by Bernd Heine and Heiko Narrog

The Oxford Handbook of Linguistic Interfaces
Edited by Gillian Ramchand and Charles Reiss

The Oxford Handbook of Linguistic Minimalism
Edited by Cedric Boeckx

The Oxford Handbook of Linguistic Typology
Edited by Jae Jung Song

The Oxford Handbook of Translation Studies
Edited by Kirsten Malmkjaer and Kevin Windle

THE OXFORD HANDBOOK OF

LANGUAGE EVOLUTION

Edited by

MAGGIE TALLERMAN

and

KATHLEEN R. GIBSON

OXFORD
UNIVERSITY PRESS

of communicative domains. There is no reason to suppose that early human language was so confined. Here, one should recall that the interlingual function is at the core of what characterizes a pidgin. All pidgins and creoles are—or were, formerly—MICs. The pressures driving the stabilization and structural elaboration of pidgins can be related to expanding domains of communication and the need for predictability and automaticity.

The elaboration of modern human language out of protolanguage can only have coincided with the increasing complexity of early human society (Givón 1998: 96; Johansson 2005: 239). Since Saussure (1916[1976: 263]), proposals that languages ultimately share a common origin in one place and time and in a single group have faced at best a mixed reception. Indeed, there can have been no single moment at which the language faculty itself crystallized (see Hurford and Dediu 2009). More probable is that human language commenced with the appearance of more or less discrete communication systems within small bands of people; see Nichols, this volume. Characteristic of human culture is its cumulative nature and growth in complexity and diversity with time, indeed at such a rate that our ancestors 'could not communicate about their world without constantly evolving better ways of communicating' (Odling-Smee and Laland 2009: 120). Diversity in material culture creates the opportunity for trade; 'with trade comes negotiation, and further selection for effective communication' (ibid). With higher frequencies of contact between bands came an increased need for individuals to exchange information in group-external settings and the establishment of cross-group communication networks. Increasingly complex grammars took shape as communicative requirements become more demanding.

This brings us to the commensurability of the cognitive capacities of the agents of grammar construction. Givón believes that *Homo sapiens* built its communication system on a pre-existing 'neuro-cognitive platform' that included semantic memory (words), event representation (word concatenation), and event-episodic representation (clause concatenation) (2009: 336). The adaptive impetus to go beyond this platform was rooted in the demands of communication. The windows approach postulates that modern pidgin speakers return to the pregrammatical mode of speaking and can (under the right conditions) create new, language-independent rules in order to meet their expanding communicative needs. The developmental processes they utilize (e.g. the Merge operation, extrapolation of functional categories out of lexical meaning) are presumed to recapitulate those of our hominin ancestors. But at present, it is unclear to what extent the possession of full language is a factor in modern constructions of a grammatical mode of speaking. Kihm (2000) and Benazzo (2009) believe that comparison with other processes that illustrate the creation of language anew is indicated, specifically the spontaneous invention of signed languages. Nevertheless, a properly constructed pidgin window on language evolution holds great heuristic promise.

CHAPTER 57

WHAT MODERN-DAY GESTURE CAN TELL US ABOUT LANGUAGE EVOLUTION

SUSAN GOLDIN-MEADOW

57.1 INTRODUCTION

Humans are equipotential with respect to language-learning—if exposed to language in the manual modality, children will learn a sign language as quickly and effortlessly as they learn a spoken language. Why then has the oral modality become the modality of choice for languages around the globe? The oral modality might have triumphed over the manual modality simply because it is so good at encoding messages in the segmented and combinatorial form that human languages have come to assume. But this is not the case—the manual modality is just as good as the oral modality at segmented and combinatorial encoding, as evidenced by sign languages of the deaf. There is thus little to choose between sign and speech on these grounds. However, language serves another important function—it conveys information imagistically. The oral modality is not well suited to this function, but the manual modality excels at it. Indeed, the manual modality has

taken over this role (in the form of spontaneous gestures that accompany speech) in all cultures. It is possible, then, that the oral modality assumes the segmented and combinatorial format not because of its strengths, but to compensate for its weaknesses (Goldin-Meadow and McNeill 1999).

This argument rests on a crucial assumption—that imagistic information is an important aspect of human communication and that it is well served by representation in the manual modality. The present chapter examines the gestures that hearing speakers produce when they talk to provide evidence for this assumption and focuses on two roles of gesture: when the manual modality works along with speech to fulfil the functions of language, (1) its imagistic encoding fills in gaps left by speech and thus plays an important communicative role for the listener, and (2) its imagistic encoding helps speakers think and thus plays an important cognitive role for the speaker. The chapter then explores the changes that take place in the manual modality when it is called upon to fulfil the functions of language on its own, that is, when it works alone without speech in both established and newly-emerging sign languages. The chapter ends with a brief discussion of the advantages of a language system that contains both an imagistic and a segmented format, and the implications of the phenomenon for linguistic evolution.

57.2 When the manual modality shares the functions of language with the oral modality

57.2.1 The properties of gesture accompanying speech

McNeill (1992) has argued that the gestures that accompany speech form a single integrated system with that speech, with each modality best suited to expressing its own set of meanings. Speech reflects a linear-segmented, hierarchical linguistic structure, utilizing a grammatical pattern that embodies the language's standards of form and drawing on an agreed-upon lexicon of words. In contrast, gesture reflects a global-synthetic image. It is idiosyncratic and constructed at the moment of speaking—it does not belong to a conventional code. Consider, for example, a speaker who is describing the east coast of the United States and produces a gesture tracing the shape of the coastline. The gesture conveys nuances of the coastline that are difficult, if not impossible, to capture in speech. Gesture thus allows speakers to convey thoughts that may not easily fit into the categorical system that their conventional language offers (Goldin-Meadow 2003a).

McNeill (1992: 41) lists the fundamental properties of the gestures that accompany speech as follows:

(i) Gestures are global in meaning. The meanings of the parts of a gesture are determined by the whole (and not *vice versa*, as is the case in speech). Indeed, the parts of a gesture cannot really be considered isolable units, as they are dependent for their meaning on the whole. In contrast to the bottom-up structure of sentences, there is consequently a top-down structure *within a gesture.*

(ii) Gestures are non-combinatoric. Gestures do not combine to form larger, hierarchically structured gestures. Most gestures are one to a clause and, even when there are successive gestures within a clause, each corresponds to an idea unit in and of itself. There is, as a result, no hierarchical structure *across gestures* (though there may be other kinds of non-hierarchical structure; see McNeill (1992)).

(iii) Gestures are context-sensitive. They are free to incorporate only the salient and relevant aspects of the context. Each gesture is created at the moment of speaking and highlights what is relevant. Because of the sensitivity of gestures to the context of the moment, there is variability in the forms gesture takes *within a speaker.*

(iv) Gestures do not have standards of form. Different speakers display the same meanings in idiosyncratic ways. There is consequently variability in the forms gesture takes *across speakers.* Even when there is cross-speaker similarity, this is not because of standards but because of similarity of meaning—similar meanings engender similar gestures.

It is often easy to analyse a given gesture into parts, but these parts have a different status from the parts of sentences—they are individually constructed with meanings that are determined by the context and that percolate from the top down. For example, in describing an individual running, a speaker moved his hand forward while wiggling his index and middle fingers (McNeill 1992). The parts of this gesture gain meaning because of the meaning of the whole; the wiggling fingers mean 'running' only because we know that the gesture, as a whole, depicts someone running and not because this speaker uses wiggling fingers to mean running in any other context. Indeed, in other gestures produced by this same speaker, wiggling fingers may well have a very different meaning (such as indecision between two alternatives). To argue that the wiggling fingers gesture is composed of separately meaningful parts, one would have to show that each of the three components that comprise the gesture—the V handshape, the wiggling motion, and the forward motion—is used for a stable meaning across the speaker's gestural repertoire. There is no evidence for stability of this sort in co-speech gestures (McNeill 1992; Goldin-Meadow et al. 1995).

Thus, the gestures that accompany speech are not composed of parts but instead have parts that derive from wholes. Moreover, they are wholes that represent by way of imagery. Because the gesture as a whole must be a good (that is, relatively transparent) representation of its referent, the addition of semantic information to a spontaneous gesture never decreases its iconicity.

57.2.2 The imagistic information encoded in the manual modality plays a role in communication

The imagistic base of gesture allows it to capture and reveal information that speakers may have difficulty expressing in speech. As a result, gesture offers listeners insight into information that cannot be gotten by listening. Gesture thus has the potential to play a unique role in communication.

Take as an example a child describing why she thinks that the water in a tall, thin container is a different amount from the water in a short, fat container. She says, 'it's different because this one's tall and that one's short', while holding a flat palm first at the height of the water in the tall container and then at the height of the water in the short container. The child focuses on the height dimension of the containers in both speech and gesture. Now consider a child who also focuses on height in his speech, but does so while producing first a narrow C-shaped gesture indicating the width of the tall container and then a larger C-shaped gesture indicating the width of the short container. This child focuses on the height of the containers in speech, but their width in gesture.

In order to fully understand that the amount of water in the two containers is the same, the child needs to understand that height and width compensate for one another—that the taller height of the first container is offset by its narrower width, and that the shorter height of the second container is offset by its larger width. The second child has noticed both dimensions and, although he says that the amount of water in the two containers is different, he is well on his way toward grasping the concept of conservation of quantity. When both children are given instruction in the concept, the second child, the one whose gestures convey information *not* found in speech, is more likely to benefit from the instruction than the first, whose gestures convey the same information in gesture and speech (Church and Goldin-Meadow 1986).

If gesture can reveal unspoken thoughts, those thoughts are then 'out there' and can be part of the conversation—assuming, of course, that gesture can be read by ordinary listeners in ordinary circumstances. And it can. Everyone can read gesture, young or old, in an experiment or in real-life communication. In fact, the information we take from speech is affected by the gestures that accompany speech (Goldin-Meadow and Sandhofer 1999). For example, children and adults are *more* likely to glean the message conveyed in speech when it is accompanied by gesture conveying the same information than when it is accompanied by no gesture at all. Conversely, listeners are *less* likely to glean the message in speech when it is accompanied by gesture conveying a different message than when it is accompanied by no gesture.

These facts raise the possibility that, by playing a role in communication, gesture can lead to cognitive change in the gesturer. If children reveal their readiness for instruction simply by moving their hands, and if listeners are attentive to those movements and change their responses accordingly, gesture can provide an indirect

way for children (and all learners) to tell their teachers what they need next. A teacher would be able to recognize that the two children described earlier in the conservation of quantity example differ in how well they understand conservation only if the teacher paid attention to the children's gestures as well as their speech. Teachers do, in fact, attend to the gestures that their pupils produce. Moreover, they alter the instruction they give their pupils (both their speech and their gestures) as a function of the children's gestures (Goldin-Meadow and Singer 2003). Gesture can thus change the course of learning by influencing the kind of input the learner receives.

Interestingly, speakers are often not aware that they are moving their hands when they speak, and listeners rarely know whether the information they glean from a conversation comes from the speaker's hands or mouth. Nonetheless, gestures that seem to be invisible have a noticeable impact on communication.

57.2.3 The imagistic information encoded in the manual modality plays a role in thinking

Gesture thus has an impact on listeners. But it can also have an effect on the speakers themselves. We have all had the experience of finding ourselves gesturing when no one is watching. We may feel sheepish about it but that does not stop us. Why do we gesture when we speak? Perhaps because gesturing helps us think. Indeed, there is evidence that gesturing can make it easier to retrieve words (Krauss et al. 2000), to package ideas into words (Kita 2000), to tie words to the real world (Glenberg and Robertson 1999), and to remember an unrelated list of words (Goldin-Meadow et al. 2001). By freeing resources that can be used for other tasks, gesturing has the potential to contribute to cognitive growth.

Gesture can also play a role in cognitive growth by providing an imagistic route through which ideas can be made active or brought into the learner's repertoire. For example, telling children to gesture while they explain their solutions to a maths problem brings out new, and correct, ideas in gesture about how to solve the problem. Interestingly, at the same time that they are producing these correct ideas in gesture, the children continue to solve the problems incorrectly and to produce incorrect problem-solving strategies in speech. However, if the children are then given instruction in how to solve the problems, they are more likely to profit from the instruction than children who were told not to gesture (Broaders et al. 2007). Gesturing thus brings out implicit ideas, which, in turn, can lead to learning.

Even more striking, we can introduce new ideas into children's cognitive repertoires by telling them how to move their hands. For example, if we make children sweep their left hand under the left side of the mathematical equation 3+6+4=__+4 and their right hand under the right side of the equation during instruction, they learn how to solve problems of this type. Moreover, they are more likely to succeed on

the problems than children told to say, 'The way to solve the problem is to make one side of the problem equal to the other side' (Cook et al. 2007).

How does gesturing promote new ideas? The children may be extracting meaning from the hand movements they are told to produce. If so, they should be sensitive to the particular movements they produce and learn accordingly. Alternatively, all that may matter is that the children are moving their hands. If so, they should learn regardless of which movements they produce. In fact, children who were told to produce movements instantiating a correct rendition of the grouping strategy during instruction (such as a V-hand placed under the 3 and 6 in the 3+6+4=__+4 problem, followed by a point at the blank) solved more problems correctly after instruction than children told to produce movements instantiating a partially correct strategy (such as a V-hand placed under the 6 and 4, followed by a point at the blank), and the latter group, in turn, solved more problems correctly than children told not to gesture at all (Goldin-Meadow et al. 2009). Importantly, this effect was mediated by whether children added the grouping strategy to their post-instruction spoken repertoires. Because the grouping strategy was never expressed in speech during instruction by either child or teacher, nor was it expressed in gesture by the teacher, the information that children incorporated into their post-instruction speech must have come from their own gestures. We may be able to lay foundations for new knowledge simply by telling learners how to move their hands.

Moreover, the manual modality may be a particularly good venue for innovation because ideas expressed in this modality may be less likely to be challenged (or even noticed) than ideas expressed in the more explicit and recognized oral modality. Because gesture is less codified and less monitored than speech, it may be more welcoming of fresh ideas than speech.

57.3 When the manual modality takes over all of the functions of language

We have seen that the manual modality conveys information imagistically, and that this information has an important role to play in both communication and thinking. The manual modality assumes an imagistic form when it is used in conjunction with a segmented and combinatorial system (i.e. speech). But what happens when the manual modality must fulfil all of the functions of language on its own? It turns out that, under these circumstances, the manual modality changes its form and itself becomes segmented and combinatorial. We see this phenomenon in conventional sign languages passed down from one generation to the next, but it is also found, and is particularly striking, in emerging sign languages.

57.3.1 Conventional sign languages

Sign languages of the deaf are autonomous languages, independent of the spoken languages of hearing cultures. Despite the fact that they are processed by the hand and the eye and not the mouth and the ear, sign languages have the essential properties of segmentation and combination that characterize all spoken language systems (Klima and Bellugi 1979; Sandler and Lillo-Martin 2006). Sign languages are structured at the sentence level (syntactic structure), at the sign level (morphological structure), and at the level of sub-sign, and have meaningless elements akin to phonemes ('phonological' structure). Just like words in spoken languages (but unlike the gestures that accompany speech), signs combine to create larger wholes (sentences) that are typically characterized by a basic order, for example, SVO (Subject–Verb–Object) in American Sign Language (ASL); SOV in Sign Language of the Netherlands. Moreover, the signs that comprise the sentences are themselves composed of meaningful components (morphemes).

Although the signs in a language like ASL often look iconic, this iconicity does not appear to play an important role in the way signers process sign, nor in the way children acquire sign. For example, young children are just as likely to learn a sign whose form does not resemble its referent as a sign whose form is an iconic depiction of the referent. Moreover, many signs and grammatical devices do not have an iconic relation to the meanings they represent. For example, the sign for 'slow' in ASL is made by moving one hand across the back of the other hand. When the sign is modified to be 'very slow', it is made more rapidly since this is the particular modification of movement associated with an intensification meaning in ASL (Klima and Bellugi 1979). Thus, modifying the meaning of a sign can reduce its iconicity in a conventional sign language simply because the meaning of the sign as a whole is made up of the meanings of the components that comprise it.

In contrast, as described earlier, the gestures that accompany speech are not composed of parts but are instead non-compositional wholes. Since the gesture as a whole must be a good representation of its referent, the addition of semantic information to a spontaneous gesture always increases its iconicity—if something is thought of as very slow, the gesture for it is also very slow (McNeill 1992). The gesture *as a whole* represents 'very slow', and although one could, in principle, break up the gesture into two parts (such as 'slow', a movement across the back of the hand, and 'very', an exaggerated and slowed movement), there is no evidence that these particular forms have independent and consistent meaning across a range of gestures—as they would have to if they were part of a combinatorial system in a conventional sign language (we later consider whether signers gesture).

57.3.2 Emerging sign languages

Not only is segmentation and combination characteristic of communication in the manual modality when that communication has been conventionalized within a community over generations, but it is also a salient feature of emerging manual communication systems. We consider three different systems, all of which display the properties of segmentation and combination: (1) the gestures invented by a deaf child who has not been exposed to a conventional sign language to communicate with hearing individuals—a system of homesigns developed over a period of years by a deaf child; (2) the gestures that arise when homesigners are brought together for the first time and that change over time as new learners enter the community—newly developing sign languages developed over decades by a community; (3) the gestures that hearing individuals create in an experimental situation when asked to communicate using their hands and not their mouths—signs developed on-the-spot by hearing adults.

57.3.2.1 *Signs invented by a homesigner*

Deaf children exposed from birth to a conventional sign language such as ASL acquire that language in stages comparable to those followed by hearing children acquiring a spoken language. However, 90% of deaf children are not born to deaf parents who can provide early exposure to conventional sign language. Rather, they are born to hearing parents who, not surprisingly, speak to their children. Unfortunately, it is extremely uncommon for deaf children with severe to profound hearing losses to acquire spontaneously the spoken language of their hearing parents and, even with intensive instruction, their speech is very likely to be markedly delayed. In addition, unless hearing parents send their deaf children to a school in which sign language is used, the children are not likely to be exposed to a conventional sign system.

Despite their lack of a usable model of conventional language, deaf children of hearing parents manage to communicate and do so by means of a self-created system of homesigns. These systems are characterized by a variety of language-like properties, including segmentation and combination (Goldin-Meadow 2003b, 2005). Rather than communicate the way a mime artist would, enacting an event as veridically as possible, the child conveys messages using segmented gestures combined into a consistently structured string. For example, rather than going over to the cookie jar and pretending to remove the cookie and eat it, the child points toward the cookie and then jabs her hand several times toward her mouth, effectively conveying 'cookie-eat'. The gesture strings generated by each of the deaf children can be described in terms of very simple patterns. These patterns predict which semantic elements are likely to be gestured and where in the gesture string those elements are likely to be produced. For example, deaf children

inventing homesigns in different countries (China, United States) tend to leave gestures for the agent (S) out of their gesture sentences, as do many languages (even English has a form that permits agent omission, the truncated passive: *the ball was hit*, as opposed to *the ball was hit by the boy*). The deaf children typically produce a gesture for the object (O) and a gesture for the action (V) and, importantly, they produce these gestures in a consistent order, placing the O gesture before the V gesture; in other words, they follow an OV order (Goldin-Meadow and Mylander 1998). The gesture systems thus have sentence-like structure.

In addition to structure at the sentence level, each deaf child's homesign system also has structure at the word level (Goldin-Meadow et al. 1995, 2007). Each gesture is composed of a handshape and a motion component, and the meaning of the gesture as a whole is determined by the meanings of each of these parts. For example, a child moves his hand shaped like an O in a short motion arcing downward to request the experimenter to lay a penny down flat. The O-handshape represents a 'round object' (the penny) in this gesture and in the child's entire corpus of gestures, and the short-arc motion represents 'put down', again across the entire gesture corpus. When produced together within a single gesture, the component parts combine to create the meaning of the whole, 'put down a round object'.

Importantly, the structure found at the sentence and word levels in each of the deaf children's gesture systems cannot be traced back to the spontaneous gestures that their hearing parents produced while talking to them. The children see the global and unsegmented gestures that their parents produce. But when gesturing themselves, they generate gestures that are discrete, segmented forms joined together into structured strings. The children thus transform the unsegmented gestures they see into a segmented and combinatorial system of their own.

57.3.2.2 *Signs invented by a community of homesigners*

Nicaraguan Sign Language offers a unique opportunity to watch a sign language become increasingly complex over generations of creators. The initial step in the creation process took place when deaf children in Managua were brought together for the first time in an educational setting. The deaf children had been born to hearing parents and each was likely to have invented his or her own homesign system. When brought together, they needed to develop a common sign language. Not surprisingly given its homesign roots, the system generated by this first cohort of signers was characterized by segmentation and combination (Kegl et al. 1999). But Nicaraguan Sign Language did not stop there. Every year, new students entered the school and learned to sign among their peers. The second cohort of signers had as its input the sign system developed by the first cohort and, over the course of two decades, changed the system so that it became more language-like (Senghas and Coppola 2001).

The second cohort of signers, in a sense, stands on the shoulders of the first. It does not need to introduce segmentation and combination into the system—those properties are already present in their input. They can therefore take the transformation process one step further. But it may be the Nicaraguan homesigners (and homesigners all over the globe) who take the first, and perhaps the most transformative, step—they change hearing speakers' gestures, which are global and synthetic, into a segmented and combinatorial system. Subsequent learners are then able to build on these properties, creating a system that looks more and more like the natural languages of the world.

The situation in Nicaragua is not unique. As another example, a community, now in its seventh generation and containing 3500 members, was founded 200 years ago in Israel by the Al-Sayyid Bedouins. Within the last three generations, 150 deaf individuals were born into this community, all descended from two of the founders' five sons. Al-Sayyid Bedouin Sign Language (ABSL) was thus born (Sandler et al. 2005). With three generations of signers, there is an opportunity not only to glimpse a language in its infant stages but also to watch how it has grown. For example, highly regular sign order evolved to mark grammatical relations in ABSL within the first generation; the particular order used is SOV. However, the language appears to have developed very little, if any, complex morphology (Aronoff et al. 2004, although it is worth noting that not all spoken languages have rich morphological structure either).

ABSL is not yet a mature language and thus is still undergoing change. As a result, signers from each of the three generations are likely to differ, and to differ systematically, in the system of signs they use. By observing signers from each generation, we can therefore make good guesses as to when a particular linguistic property first entered the language. Moreover, because the individual families in the community are tightly knit, with strong bonds within families but not across them, we can chart changes in the language in relation to the social network of the community. We can determine when properties remained within a single family and when they did not, and thus follow the trajectory that particular linguistic properties took as they spread (or failed to spread) throughout the community. This small and self-contained community consequently offers a unique perspective on some classic questions in historical linguistics.

Like Nicaraguan Sign Language, ABSL has arisen with no influence from any established language, either signed or spoken. However, ABSL differs from Nicaraguan Sign Language in that it is developing in a socially stable community with children learning the system from their parents. The differences and similarities between the two systems can thus provide useful information about the trajectories that languages follow as they grow from a homesign system into a fully formed conventional sign language.

57.3.2.3 *Signs invented by a hearing adult*

The findings reviewed thus far suggest that segmentation and combination are fundamental to human language. Note that these properties are not forced upon language by the modality in which it is expressed. Segmentation and combination are found in human language whether it is produced in the oral or manual modality. Moreover, segmentation and combination are not inevitable in the manual modality—they are not found when the manual modality is used along with speech, that is, when hearing people produce co-speech gestures.

What then determines when segmentation and combination will arise in the manual modality? One possibility is that segmentation and combination crop up in the manual modality only when it takes on the primary burden of communication. To test this hypothesis, we can examine hearing adults' gestures when those gestures are produced with speech (sharing the communicative burden) and when they are produced instead of speech (shouldering the entire communicative burden). The gestures adults produce without speech ought to display segmentation and combination and thus be distinct from the gestures the adults produce with speech.

This prediction was confirmed (Goldin-Meadow et al. 1996). When they produced gesture without speech, the adults frequently combined those gestures into strings and the strings were consistently ordered, with gestures for certain semantic elements occurring in particular positions in the string; that is, there was structure across the gestures at the sentence level. In addition, the verb-like action gestures that the adults produced could be divided into handshape and motion parts, with the handshape of the action frequently conveying information about the objects in its semantic frame; that is, there was structure within the gesture at the word level (although the adults did not develop a system of contrasts within their gestures, a characteristic of deaf homesigners' systems; Goldin-Meadow et al. 1995, 2007). Thus, the adults produced gestures characterized by segmentation and combination and did so with essentially no time for reflection on what might be fundamental to language-like communication.

The appearance of segmentation and combination in adults' gestures produced without speech is particularly striking given that these properties were *not* found in the gestures that these same adults produced *with* speech—their co-speech gestures were rarely combined into strings, and handshape was rarely used to convey object information within a gesture (Goldin-Meadow et al. 1996). In other words, the adults did not use their gestures as building blocks for larger sentence or word units. Rather, they used their gestures to imagistically depict the scenes they described, as speakers typically do when they spontaneously gesture along with their talk.

Interestingly, when hearing speakers of a variety of languages (Chinese, Turkish, and Spanish, as well as English) are asked to describe a series of events using only

their hands, they not only produce strings of segmented gestures characterized by consistent order, but those strings all display the same gesture order—even though the speakers use the predominant orders of their respective languages (and thus use different orders) when describing the same scenes in speech (Goldin-Meadow et al. 2008). The gesture order that all speakers-turned-signers use is SOV—precisely the order that we see in the early stages of other emerging sign systems (ABSL and the homesigns developed by individual deaf children). This order may reflect a natural ordering that humans exploit when creating a communication system over short and long time spans.

57.4 The advantages of a communication system with both a segmented and an imagistic representational format

Modern-day human communication systems are based on a segmented and combinatorial mode of representation (typically conveyed in the oral modality) that gives the system its generative capacity. But they also have an imagistic mode of representation (in the manual modality) that exists alongside, and that gives the system the ability to be responsive to the communicative needs of the moment. The gestures that speakers produce in the manual modality can express information that they are often not able to express within the codified spoken system. This information is processed by the listener (not necessarily consciously) and becomes part of the conversation. Moreover, once information has been expressed in the manual modality, it can catalyse change in the speaker and eventually find its way into the oral modality. Thus, there is an imagistic side to human communication that plays an important role in both communication and thinking.

If we grant that there are advantages to a communication system with both a segmented/combinatorial format and an imagistic format, we can then understand why language is the province of the oral modality. Whereas the oral modality and the manual modality can assume the segmented and combinatorial format equally well, the manual modality is particularly well suited to the imagistic format. It therefore takes over this function, leaving segmentation and combination, the hallmarks of the linguistic code, to the oral modality.

This speculation raises an interesting question about sign language. In sign, the manual modality assumes the segmented and combinatorial format essential to human language. Can the manual modality at the same time be used for imagistic expression? In other words, do signers gesture? They may (Emmorey 1999), perhaps with their mouths (for instance, one Israeli signer puffed out her cheek when signing about carrying a valise; her mouth gesture, and only her mouth gesture, made it clear that the valise was full; Sandler 2003). But the oral modality can also be used in limited ways for imagistic expression (for instance, the speed at which an object moves can be captured in the speed of the speech describing it; Shintel et al. 2006). Although it is possible to have both functions served by the same modality, it may be more efficient to separate the imagistic and segmented/combinatorial forms of representation by modality.

Does the fact that there is both an imagistic and a segmented side to modern-day communication bear on the question of linguistic evolution? It is possible, as Donald argues (1991, this volume), that the mimetic function preceded and, over evolutionary time, led to the current day analytic structures that characterize language. But it is equally possible that the two functions were present in our communicative efforts from the beginning, and that the oral and manual modalities have always worked together to fulfil our communicative needs, evolving together to produce the single system that characterizes our modern-day language.