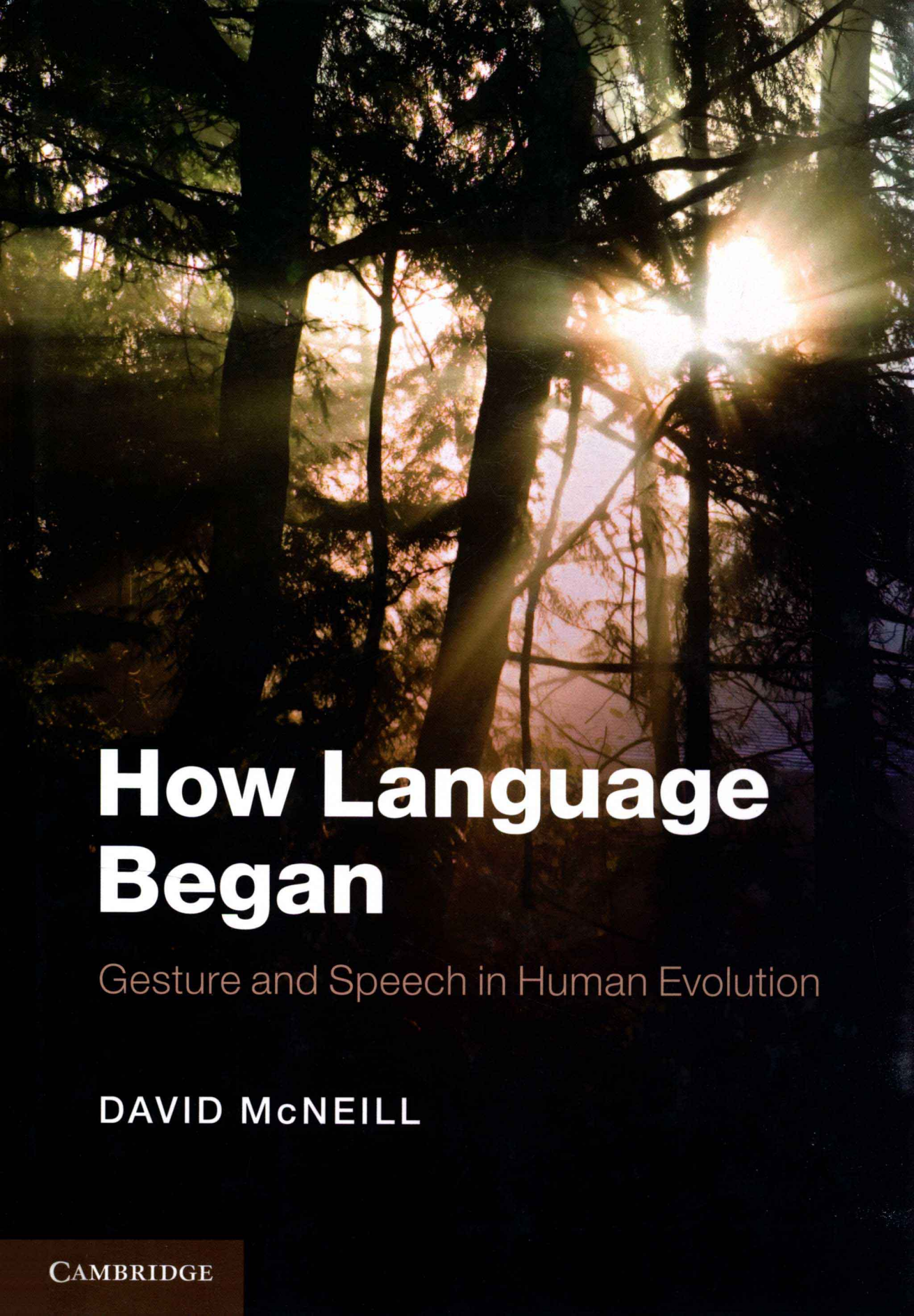

How Language Began

Gesture and Speech in Human Evolution

DAVID McNEILL

CAMBRIDGE

How Language Began

Gesture and Speech in Human Evolution

DAVID MCNEILL

常州大学图书馆
藏书章

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9781107605497

© David McNeill 2012

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2012

Printed and bound in the United Kingdom by the MPG Books Group

A catalog record for this publication is available from the British Library

Library of Congress Cataloging in Publication data

McNeill, David.
How language began : gesture and speech in human evolution / David McNeill.
pages ; cm
Includes bibliographical references and index.
ISBN 978-1-107-02121-1 (hardback : alkaline paper)
1. Language and languages – Origin. 2. Speech and gesture. I. Title.
P116.M455 2012
401 – dc23 2012021835

ISBN 978-1-107-02121-1 Hardback
ISBN 978-1-107-60549-7 Paperback

Cambridge University Press has no responsibility for the persistence or
accuracy of URLs for external or third-party internet websites referred to
in this publication, and does not guarantee that any content on such
websites is, or will remain, accurate or appropriate.

Preface – Out on a limb

This is my third book to focus on the nexus of gesture and language. Together, the three amount to a kind of unintended trilogy, what has turned out to be a sustained examination and ultimate explanation of a certain phenomenon. The first, *Hand and Mind* (1992), introduced what was then a newly discovered world of gesture, not the stand-alones (known as emblems) that have been acknowledged for millennia, but those overlooked but omnipresent gestures that wed themselves to speech itself. The second, *Gesture and Thought* (2005), developed an explanation of this wedding, the growth point. Now I tackle the origin of the growth point in evolution. By this third volume I am aware of having run far out on a limb. Out on a limb because in crafting the book I have followed a line of argument to its logical limit, or as close to a limit as I can get. The line is that language is more than the lexicosyntactic forms that one sees in written texts and the analyses of linguistics. It is also imagery. This imagery is in gesture, and is inseparable from language. The hypothesis of a growth point encompasses this idea. Taking seriously that language includes gesture as an integral component changes the look of everything. We see language in a new way, as a dynamic "language-as-action-and-being" phenomenon, not replacing but joining the traditional static (synchronic) "language-as-object" conception that has guided linguistics for more than a century.

One idea is more than one note, and in pursuing it I have discovered that it touches a wide range of other topics in language, children's development, brain, mind, and society. In this way, a great breadth of phenomena is linked. I cover, besides gesture and its binding power with speech, a specific mechanism for the origin of language, and the scenarios in which it could have arisen; an alternative, the "gesture-first" hypothesis, which fails both by predicting what did not evolve and not predicting what did evolve; the "equiprimordiality" of speech and gesture instead; a thought–language–hand brain link present in all humans but revealed directly in deafferentation cases where gestures occur normally but practical actions are impossible; phylogenetic echoes in ontogenesis of *two* language origins, one of which is extinct (which extinction is also echoed in ontogenesis); new forms of action of the hands and vocal tract orchestrated by significances other than the

actions themselves; the origin of syntax, while biological, to make these new actions shareable and portable in encounters with others in socio-culturally maintained templates (constructions); psychological sources of linguistic diversity; parallels and non-parallels to human language in chimpanzee and other primates; the remarkable, still-emerging discoveries in comparative genetics of the two or three kinds of humans known to have existed and how they may have differed in linguistic capacities; how consciousness and memory were reshaped by the origin of language; gestures during musical performances and the possibility of gestures hidden in written prose on the page; the loss of language at points of vulnerability left over from the origin; and the unlikelihood of language evolving in any species that lacks hands.

试读结束：需要全本请在线购买： www.ertongbook.com

Acknowledgments

For comments on the manuscript, I am grateful to Carolin Kirchhof, Liesbet Quaeghebeur, Kazuki Sekine, Randall L. B. McNeill, Frank Bechter, Jana Bressem, and above all Elena Levy, with whom years ago I started the serious study of speech and gesture and who has read and commented on the manuscript more than once. I have done my best to incorporate everything these responsive fellow gesture-world inhabitants have suggested.

My colleague, Susan Duncan, has played a huge role in the development of the growth point concept, and her many contributions are recognized throughout the text.

Many of the ideas developed here were first explored with my colleagues, Bennett Bertenthal, Jonathan Cole, Susan Duncan, and Shaun Gallagher (see McNeill *et al.* 2005, 2008).

I wish to acknowledge Michael Arbib for his spirited defense of "gesture-first" made in his editor's review of our McNeill *et al.* paper in *Interaction Studies* (2008), which has helped shape the discussion in Chapter 3.

Bencie Woll provided excellent feedback especially concerning Chapter 2, which is far more digestible as a result, and also for the precious Henry Sweet reference, quoted in Table 3.1.

Nobuhiro Furuyama made very helpful comments about the "supplantation" arguments of Chapter 3.

Sarah Thomason made saving comments about the "Psycho-Babel" section of Chapter 3.

Steve McCafferty's comments helped jell the discussion of metaphoricity in Chapter 4.

The title and subtitle emerged in interactions with Cambridge University Press.

I wish to thank both Susan and Elena for seeing that what started as notes had undergone a metamorphosis into this book.

For a second time Nobuko McNeill has inspired chapters with her amusing, pithy, clear-headed and invariably thought-provoking and (often) thought-shaking remarks.

I thank my family, Nobuko, Cheryl and Randall, for their patience, laughter, encouragement, and suggestions as they combated my wavering determination to see this project through, and provided ideas that figure in several chapters.

The work tapped in writing this book has been supported over the years by research grants from NSF, NIH, the Video Analysis and Content Extraction (VACE) project (with Francis Quek and Mary Harper), The Spencer Foundation, and The Samuel Beck Memorial Fund of the University of Chicago.

I have lectured on the topics in this book at two International Society for Gesture Studies Conferences, in Austin and in Evanston, at the Chicago Linguistics Society, at the German Semiotics Society in Dresden, at the two *ORAGEs* (in Besançon and Aix-en-Provence), and at universities and research institutes in Denmark, France, Germany, Italy, Japan, Sweden, and the US.

Contents

List of figures [*page* viii]
List of tables [x]
Preface [xi]
Acknowledgments [xiii]

1 Introduction – gesture and the origin of language [1]

2 What evolved (in part) – the growth point [19]

3 How it evolved (in part) – Mead's Loop [58]

4 Effects of Mead's Loop [114]

5 Ontogenesis in evolution – evolution in ontogenesis [165]

6 Alternatives, their limits, and the science base of the growth point [186]

Notes [223]
References [236]
Index [257]

Figures

1.1 The Gesture Continuum. [*page* 5]
1.2 Example of a speech-synchronized gesture. [8]
1.3 Jahai (Laos) lip point. [11]
1.4 Canonical "OK." [17]
2.1 "Rising hollowness" gesture. [27]
2.2 Illustrating gesture combinations. [30]
2.3 Isolated "inside" gestures do not highlight interiority. [36]
2.4 "Inside" after "outside" gestures highlight interiority. [37]
2.5 Exception that proves the rule. [39]
2.6 Timing highlights interiority. [40]
2.7 Gestures with a sign language. [41]
2.8 Catchments. [44]
2.9 Setting up the opposed forces paradigm. [50]
2.10 Setting up the unpacking. [52]
2.11 Meta-pragmagtics. [55]
3.1 Gesture and vocalization in chimpanzees. [72]
3.2 Bonobo gestures. [72]
3.3 Imitation but no instruction in chimpanzees. [74]
3.4 First occurrences of "King" and "Queen" morphs. [81]
3.5 A later abbreviated "Queen." [82]
3.6 "Queen" and "King" by the listener. [83]
3.7 Possible "Sylvester" proto-morphs. [84]
3.8 IW gesture with vision. [103]
3.9 IW two-handed iconic gesture without vision. [104]
3.10 IW cannot perform an instrumental action without vision. [105]
3.11 IW changes rate of speech and gesture without vision in perfect synchrony. [106]
3.12 Annotated brain. [107]
3.13 Two versions of gesture-first extinction. [113]
4.1 Ad hoc metaphor for "inaccessibility." [121]
4.2 Neapolitan gestures. [123]
4.3 English speaker's non-cultural equivalent. [125]
4.4 Communicative dynamism. [127]

4.5 Hierarchies of consciousness. [132]
4.6 "Microcsopic" consciousnss. [133]
4.7 Decline and return of energy in gesture. [137]
4.8 Proportion of items retained in working memory. [141]
4.9 Interpersonal mimicry. [149]
4.10 Scene from Canary Row. [150]
4.11 Gesture to be mimicked 1. [151]
4.12 Gesture to be mimicked 2. [151]
4.13 Gesture to be mimicked 3. [152]
4.14 Embodiment in two bodies. [153]
4.15 A dog surveys with its nose, a human with its eyes. [154]
4.16 Williams syndrome reduces overlap of thought and language. [157]
5.1 Gesture explosion. [175]
5.2 Motion "decomposition." [178]
5.3 2;6 year-old appears to combine path and manner. [179]
6.1 Hypothetical gestures during pauses and phonations. [189]
6.2 Questioned gesture grammatical construction. [193]
6.3 Gesture "pathfinder." [197]
6.4 Example illustrating GP hypothesis formulation and test. [211]

Tables

2.1 Semiotic oppositions within GPs. [*page* 20]
2.2 Catchment themes. [47]
3.1 Gesture-first advocates and their statements. [61]
3.2 "Constructions" in wordless motion event descriptions. [86]
4.1 Conversation snippet. [146]
4.2 Speech, gesture, and context in an autistic boy's narration. [159]
5.1 "Pivot" construction. [171]
5.2 Pivot-like slots. [172]
5.3 Gestural viewpoints at three ages. [176]
5.4 Gestures and speech by 2;4 year-old. [182]
6.1 Gesture rates (onsets per 1000 s) in fluent speech. [190]

试读结束：需要全本请在线购买： www.ertongbook.com

1 | Introduction – gesture and the origin of language

> It's like seeing someone's thought.
>
> – Mitsuko Iriye, historian, on learning to code gestures (in the 1980s)

1.1 THE ORIGIN OF LANGUAGE

The origin of language, a prodigal topic, has recently returned to respectability after a long exile.[1] Discoveries in linguistics, brain science, primate studies, children's development, and elsewhere have inspired new interest after the infamous nineteenth-century ban (actually, bans) on the topic – both the Société de Linguistique de Paris in 1865 and the Philological Society of London in 1873 prohibiting all contributions on the topic (London promising that any such would be tossed directly into the wastebasket; all of this described in Kendon 1991). The topic can be approached from many angles. Most common seems to be the comparative – differences and resemblances between humans and other primates. A related approach is to consider the brain mechanisms underlying communicative vocalizations and/or gesture. These have been recorded directly in some primate species and can be compared to humans on performance measures thought to depend on similar brain mechanisms. Or a linguistic angle – the key features of human language and whether anything can be said of how they came to be and whether other animal species show plausible counterparts. Approaches are combined in comparing human language to vocalizations, gestures, and/or the instructed sign language use of, say, orangutans or chimps. I will take a third approach, gestures, which also has its devotees, but I shall diverge from other approaches in crucial ways. I am not endorsing a popular current theory, appearing over and over in a veritable avalanche of recent books – what I dub "gesture-first." Despite the theory's name, the primatologist, neuroscientist, developmental psychologist, anthropologist, sign-language linguist, regular linguist, computer scientist, etc. proponents of gesture-first seemingly lack any serious acquaintance with gesture other than (it appears) its folk culture portrayals (so they do not recognize a key point of this book:

that language is misconstrued if it is not seen as a unity of language and gesture).

Gesture-first holds that the initial form of language lacked speech – it was a pantomimic or a sign language. I show that gesture-first (to put it delicately) is unlikely to be true because it is unable to capture the connections of speech and gesture that we, living counter-examples, display: it "predicts" what did not evolve (that gesture withered or was marginalized when speech arose) and does not predict what did evolve (that gesture is an integral part of speaking). A theory that says what didn't happen did, and what did happen didn't, can't generally be true, to say the least. That so many have adopted it I explain by the above-mentioned folk (and fabricated) beliefs about gestures.

The origins question homes in on what makes us human; how we diverge from other animal species, including our near neighbors, the Great Apes; it exposes in a fundamental way what comprises the gift of language. The approach here will ultimately synthesize various approaches to the question that modern authors have pioneered, not out of an urge to be all-encompassing but because these approaches will find a place in this approach's own inner logic.

The origin of language brought forth not only language but also new forms of action, new modes of thought, and new structures in the human brain. And these changes in action, thought, and brain are the sources of much else without which history, culture, and the human story could not have unfolded as they have. I hope by the end to clarify this story, how it began in unexpected ways, and on what it depends at a foundational level.

1.1.1 How this book differs

The approach here is to uncover the kind of mind that made the origin of language possible; and correspondingly, the kind of mind that language, once started, modified and extended. Other approaches emphasize the external aspects of the origin – communication, structure, parallels to other animal communication, all of which are valid but do not attempt to uncover the mindset of the creatures in which language came to be. My guiding idea and fundamental divergence is the following, proffered as an insight into the human mindset for language in general: Gestures are components of speech, not accompaniments but actually integral parts of it. Much evidence supports this idea, but its full implications have not always been recognized. The *growth point* (GP) hypothesis is designed to explicate this integral

linkage. It is presented briefly here, more fully with an example below, and explained in detail in Chapter 2.

Gestures offer one kind of symbol, language a different kind, and the two kinds of semiosis ("semiosis" and "semiotic" refer to the nature of symbols) are unified in GPs; in a GP symbols of these two different orders combine. A key insight is that speech on the one hand and gesture on the other, when combined in a GP, bring semiotically opposite modes of thinking together at the same time. A GP thus forms a single mental package or idea unit out of semiotically unlike components.

This "unity of opposites," as I will call it, creates a new form of human cognition that animates language and gives it a dynamic dimension. The semiotic opposition in a GP is intrinsically unstable; it seeks a resting place. The instability and the processes initiated to stabilize or resolve it, which I call "unpacking the GP," propel thought and speech forward, hence provide a dynamic dimension of language. All of these features of language were built in by how language began.

1.2 WHAT IS "GESTURE"?

1.2.1 Definitions of "gesture"

Gesture plays a central role in the arguments of this book. It is taken seriously and I need to explain what I mean when I refer to it. I cannot deny that the word is problematic. A journalist's cliché portrays a gesture as trivial, irrelevant, and slightly contemptible. It uses "gesture" to label something that a public figure, a politician or a magnate, has done as ungenuine and feckless; as sterile, futile, pointless, unfruitful, and untruthful, made for show and not effect. The cliché is worse than irrelevant. It positively obstructs understanding. Given the word's ragged appearance I would have preferred not to use it at all but there is no avoiding it; a suitable alternative simply does not exist in our language. I once concocted a term, "temaniosis," made from a Japanese root for "imitation in the hand"[2] and a Greek suffix for "of or relating to, of the nature of," which I thought would get close to the sense of "gesture" that I am using – but discovered that it is a combination so broad linguistically, exact though it is, that it offends some readers' sensibilities. And in any case it is vain to invent a word that will not gain general currency, and I judged there was little hope of that.

Adam Kendon (2004) placed gestures in the category of "actions that have the features of manifest deliberate expressiveness." I adopt this

definition but with one qualification and one proviso. The qualification is that gesture cannot be deliberate; as we regard them "gestures" are unwitting and automatic, and anything but deliberate. (Kendon may have meant by "deliberate" non-accidental, and with this I agree; but the word also conveys, "done for a purpose," and with that I do not agree.) The proviso concerns "action." In the sense that we intend (gesture as a special or what I later call a "new" kind of action) movements are orchestrated by significances created by the speaker him- or herself, not movements to attain external goals (goals lead to practical actions, not gestures). So our definition, based on Kendon's but excising "deliberate" and specifying the kind of action (and far from tripping off the tongue), is this:

A gesture is an unwitting, non-goal-directed action orchestrated by speaker-created significances, having features of manifest expressiveness.

Very often I use "gesture" still more restrictively to mean all of the above, plus:

A gesture is a manifestly expressive action that enacts imagery (not necessarily by the hands or hands alone) and is generated as part of the process of speaking.

1.3 THE GESTURE CONTINUUM

The remainder of this chapter is organized around Figure 1.1, The Gesture Continuum, a continuum of manifest expressiveness modes, all differing but all termed "gestures," annotated to show where these definitions and other important concepts apply. Later, I give examples of the Continuum and describe in detail certain features of it, especially the gesticulation pole – the focus of this book. The Gesture Continuum plays an important role as well in sorting out different explanations of the beginning of language, as Chapter 3 explains.

To begin, as the Continuum shows, the word "gesture" is problematic not only because of the ragged aspects but also because it is seriously ambiguous. It covers very different phenomena. The gestures of concern to us are integral components of speech, not substitutes, accompaniments or ornaments. Such gestures appear at one end of the Continuum, called by Kendon (1988b) "gesticulations."[3] These gestures are synchronous and co-expressive with speech, not redundant; and not signs, salutes, or what are called emblems (see below). They are by far the most frequent – in descriptive speech about 90% of utterances are accompanied by them

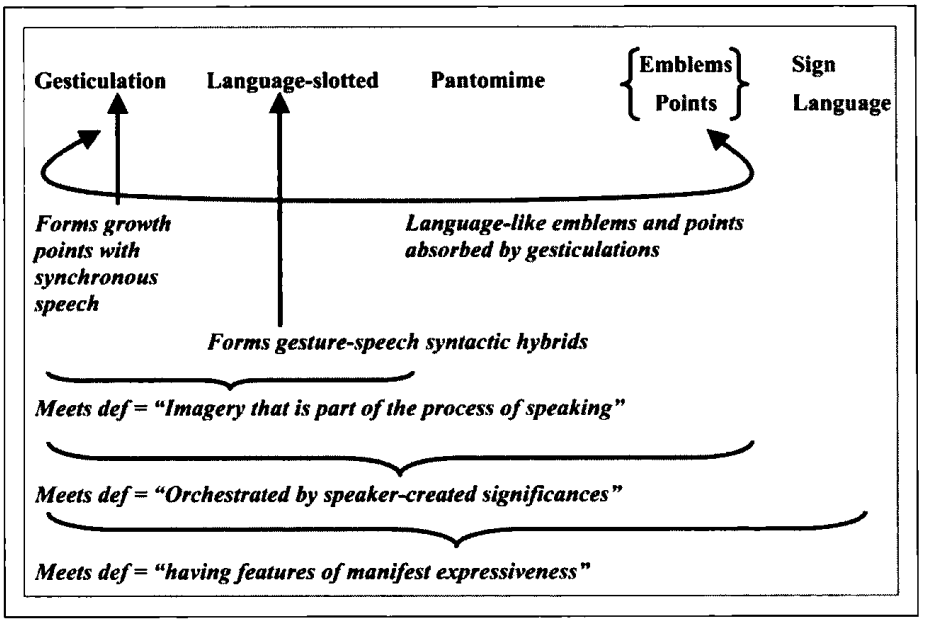

Figure 1.1 The Gesture Continuum annotated.

(Nobe 2000) and they occur in similar form and numbers across many languages.

1.3.1 Dimensions

Underlying the Continuum are three dimensions – how necessary speech is to the gesture; how language-like is the gesture; and how conventionalized is its form, so as one goes from gesticulation to sign the relationship of gesture to speech changes:

- The obligatory presence of speech decreases.
- Language-like properties increase.
- Socially regulated conventional signs replace self-generated form-meaning pairs.

1.3.2 Semiotic packages

We see the changes in how the positions along the Continuum form their own characteristic semiotic packages. At the gesticulation end (our concern) a *dual semiotic* prevails, imagery and linguistic encoding of the

same underlying idea in one package, the GP. At the language-slotted point a gesture is absorbed into its sentence, changing the relationship to language by losing co-expressivity and becoming *a constituent of the sentence itself.* At pantomime the semiosis is *reenactment*, and at the emblem/pointing point a gesture is *partly encoded* in itself. Finally, at the sign language pole gesture is *fully encoded* (cf. Klima and Bellugi 1979, Bechter 2009). The examples below illustrate these semiotic packages.

1.3.3 Timing

Also characteristic of each position is a different speech–gesture temporal arrangement (difficult to indicate in a linear layout). At the gesticulation end, the significant part of the gesture – the "stroke" – and its co-expressive speech are synchronous; at the language-slotted position, gesture slots into a vacancy in the sentence; at the pantomime and emblem/point position, gesture and speech have loose temporal relationships and speech may be completely absent. Sign language, finally, is freestanding and without speech.

1.3.4 Examples (from the most to the least language-like)

An example of an American Sign Language (ASL) *sign* is TREE – the dominant arm extended upward at the elbow, the fingers extended outward (a "5-hand") and the hand rotating back and forth at the wrist. The subordinate hand is extended flat and prone under the dominant hand elbow. The sign obviously depicts a schematic tree – trunk, leaves fluttering in the wind and the ground – but the iconicity is conventionalized and must include these specific features. A signer does not make up a new sign for each occasion. Arika Okrent (pers. comm.) calls it "non-specific," in that it is used equally for all kinds of trees and tree shapes, not just those with long bare trunks and fluttering leaves. This too is part of its conventionalization. Sign Languages such as American Sign Language of the deaf and others around the world are socioculturally maintained linguistic codes that have arisen where vocal/auditory communication is impossible. The most established are full language systems in their own right. While iconicity is present it too is conventionalized. The ASL sign is an iconic depiction, but it is a *standardized* selection of iconic features that other sign languages, also with signs that are iconic and regulated, may not use at all (Danish Sign Language traces an outline of a tree).

试读结束：需要全本请在线购买： www.ertongbook.com