The Study of Speech Processes

There has been a longstanding bias in the study of spoken language toward using writing to analyze speech. This approach is problematic in that it assumes language to be derived from an autonomous mental capacity to assemble words into sentences, while failing to acknowledge culture-specific ideas linked to writing. Words and sentences are writing constructs that hardly capture the sound-making actions involved in spoken language. This book brings to light research that has long revealed structures present in all languages but which do not match the writing-induced concepts of traditional linguistic analysis. It demonstrates that language processes are not physiologically autonomous, and that speech structures are structures of spoken language. It then illustrates how speech acts can be studied using instrumental records, and how multisensory experiences in semantic memory couple to these acts, offering a biologically grounded understanding of how spoken language conveys meaning and why it develops only in humans.

VICTOR J. BOUCHER is Senior Researcher and Professor of Speech Sciences at the Université de Montréal. His career work on the physiological processes of speech have led him to view human language as arising from constraints on motor-sensory systems and to a critical reappraisal of methods of language study.

The Study of Speech Processes

Addressing the Writing Bias in Language Science

Victor J. Boucher
Université de Montréal

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107185036
DOI: 10.1017/9781316882764

First published 2021

A catalogue record for this publication is available from the British Library.

ISBN 978-1-107-18503-6 Hardback

Contents

List of Figures *page* ix
List of Tables xii
List of Abbreviations xiii
Preface xv

Introducing a Fundamental Problem of Language Science 1

Part I Questions of Ontology: Writing and the Speech–Language Divide 11

1 How We Are Introduced to the Study of Spoken Language 13
1.1 Language as an "Autonomous" System, or the Effects of *Scriptism* 14
1.2 Defining "Speech" 18
1.3 Was the Speech–Language Division Ever Physiologically Grounded? 20
1.3.1 Saussure's Argument of a Separate Language Faculty in Broca's Area 21
1.3.2 Arguments of the Arbitrariness of Signs and Abstract Phonology 23
1.3.3 On the Primacy of Linguistic Criteria: The Historical Disconnect from Instrumental Observations 25
1.3.4 Explaining Systems of Distinctive Features: Lindblom's Demonstration (1986) 31

2 The Modality-Independence Argument and Storylines of the Origin of Symbolic Language 34
2.1 Cognitive Skills as Insufficient Factors in the Rise of Symbolic Communication 35
2.2 The Case against Modality-Independent Accounts of Symbolic Language 37
2.3 Modality-Dependent Accounts of the Rise of Symbolic Language 41
2.3.1 Mimesis, Procedural Learning, and the Case of Sign Languages 41
2.3.2 "Sound Symbolism": Questions of the Efficiency of Iconic Signs 44
2.3.3 Articulated Vocalization and the Rise of Symbolic Signs: A Laboratory Demonstration 45
2.4 The Phylogeny and Ontogeny of an Amodal Symbol Function as a Pseudo-Puzzle 49

3 The Recent History of Attempts to Ground Orthographic Concepts of Language Theory 52
3.1 From Orthographic Representations to "Substantive Universals" 53
3.2 Shoehorning Orthographic Concepts: Issues in Grounding the LAD 58
3.2.1 Biases and Limitations of Analyzing Language Development through Writing 59
3.2.2 The Search for Marks of Words and Phrases, versus "Chunks" 61
3.3 Neuroscience Falls upon Nonexistent Substantive Universals: Why This Invalidation Is Different 66
3.4 Abandoning the Competence–Performance Divide 69
Postscript – On the Use of the IPA and Terms of Latin Grammar in the Present Work 70

Part II Questions of Epistemology: The Role of Instrumental Observations 73

4 Recognizing the Bias 75
4.1 On the Tradition of Overlooking Instrumental Observations: The Case of the Phoneme 76
4.1.1 From Instrumental Records of Co-articulation to Transcribed Spoonerisms 77
4.1.2 On the Origin of Alphabet Signs: The Hypothesis of a Preliterate Awareness of Phonemes 85
4.1.3 Testing Phoneme Awareness: Issues in Defining *Reference Units* 89
4.2. The Looking-Glass Effect: Viewing Phoneme Awareness by Reference to IPA Transcripts 94
4.2.1 "Phonological" Evidence of Phonemes Versus Motor Processes 99
4.2.2 On Arguments of the "Logical Necessity" of Phonemes and the Success of Alphabet Systems 103
4.2.3 Effects of Writing on Speakers' Awareness of Words, Phrases, Sentences 105

5 (Re-)defining the Writing Bias, and the Essential Role of Instrumental Invalidation 108
5.1 On the Persistence of Scriptism in the Study of Spoken Language 108
5.2 The Need to Address Complaints of Cultural Centrism and Ethical Concerns 111

Part III The Structure of Speech Acts 115

6 Utterances as Communicative Acts 117
6.1 Describing Speech Acts and Their Meaning 117
6.2 The Parity Condition, Motor-Sensory Coupling, and the Issue of Utterance Structure 125
6.3 The Coding of Speech Acoustics in the Auditory Brain Stem and Effects of Motor-Sensory Coupling 129
6.4 Multimodal Sensory Integration: Introducing Neural Entrainment to Speech Structure 133

6.4.1 The Specificity of Neural Entrainment in the Speech Modality 133
6.4.2 Neural Entrainment to Structures of Motor Speech: Linking to Spiking Activity 134
6.4.3 On the Role of Subcortical Processes: Multisensory-to-Motor Integration and Chunking 137
6.5 Relating to Utterance Structure, or What the Brain Does Not Intrinsically Construct 141

7 Relating to Basic Units: Syllable-Like Cycles 143
7.1 Speech Production: On the Brain–Utterance Interface That Never Was 143
7.2 Basic Sequencing Units in Theories of Speech-Motor Control: Some Examples 147
7.2.1 The Equilibrium-Point (EP) Hypothesis 147
7.2.2 The Task Dynamics (TD) Model 149
7.2.3 Directions in Auditory Space Into Velocities of Articulators: The DIVA Model 152
7.3 Critical Evidence of Basic Sequencing Units and What Shapes Them 156
7.3.1 Intrinsic Muscle-Tissue Elasticity and Its Effect on Speech Motions 157
7.3.2 Other Intrinsic Effects of Muscle Tissues on Motion Sequencing within Syllable Cycles 159
7.3.3 Just How Many Units Are There in CV and VC, and Are These Represented in Memory? 162
7.3.4 Syllable Cycles within Chunks and Graded Motion Control without Phonemes 165

8 Relating Neural Oscillations to Syllable Cycles and Chunks 172
8.1 The Entrainment of Low-Frequency Oscillations and Speech Processing 172
8.1.1 On the Role of Theta- and Delta-Size Processing Windows 173
8.1.2 Reviewing Claims of a Non-sensory Entrainment of Delta to Content Units 174
8.2 Delta-Size Windows and the Sensory Chunking of Speech 176
8.2.1 Chunks and Their Signature Marks 176
8.2.2 Neural Entrainment in Speech Processing 179

9 Breath Units of Speech and Their Structural Effects 182
9.1 Utterances as Breath Units versus Sentences in Speaker–Listener Interaction 182
9.2 On Interpreting Measures of "Mean Length of Utterance" (MLU) 183
9.2.1 Utterance Complexity, Lexical Diversity, and MLU: Linking to Developing Motor Structures 187
9.2.2 Chunks in Breath Units of Speech and the Development of Vocabulary 190
9.2.3 On Explaining Developmental Milestones 192
9.3 The Structure of Spoken Language: An Interim Summary with a View on Addressing the Issue of Scriptism 194

Part IV The Processing of Speech Meaning 197

10 The Neural Coding of Semantics 199
10.1 Units of Writing, Structures of Utterances, and the Semantics of Speech 199

10.2 The Lexico-Semantic Approach: Context Information as "Nonessential" 200
10.2.1 Lexico-Semantics and Traditional Models of Language Processing 201
10.2.2 Embodied versus Disembodied Semantics 203
10.3 How Semantic Representations of Verbal Expressions Develop: On "Modes of Acquisition" 205
10.4 The Partitioning of Semantic Memory and Its Formatting in Spoken Languages 208
10.4.1 Words Are Not Biologically Grounded Units: Why Sensory Chunking Is Necessary 211
10.4.2 On Representations of Verbalized Forms in Memory: Activating Episodes of *Speech Acts* 214
10.5 The Nature of Semantic Representations: On the Neural Coding of Context Information in Action Blocks of Speech 216

11 Processes of Utterance Interpretation: For a Neuropragmatics 220
11. 1 The Issue of the Selective Activation of Semantic Representations in Speech Contexts 220
11.1.1 Context-Based Semantics: Clinical Observations Using Unconventional Test Batteries 222
11.2 On Context-Based Speech Comprehension: Selective Activation of Semantic Representations On-Line 225
11.2.1 Thalamocortical Interactions and the Integrating Role of the Motor Thalamus 227
11.2.2 The Semantics of Utterances: The Analogy of Action Selection in Spatial Navigation 230
11.2.3 Subcortical Mechanisms of Buffering and Context-Based Semantic Processing 234

Epilogue 239

References 244
Index 305

Figures

1.1 Typical schematic sagittal representation of the "speech" apparatus *page* 20
1.2 Observable acoustic structures of an utterance 25
2.1 Examples of the speech stimuli used by Boucher et al. (2018) in a sound–picture association task 48
2.2 Learners' sound–picture associations across trials and feedback conditions (Boucher et al., 2018) 48
3.1 Examples of formal syntactic analyses performed on orthographic units 54
4.1 Tracings of radiographic recordings for [ku] and [ki] 78
4.2 Tracings of radiographic recordings for a spontaneous speech error 83
4.3 Tracings of radiographic recordings for a corrected form following a spontaneous speech error 83
4.4 Faber's (1990) list of early Greek letters with their precursors in Old Aramaic script 86
4.5 Basic spectrographic values that are sufficient to synthesize /di/ and /du/, from Liberman et al. (1967) 90
4.6 Katakana signs representing how similar-sounding syllables can be variably interpreted depending on the choice of writing systems 96
4.7 Illustration of dynamic characteristics of speech production 102
6.1 A general scenario where a speaker offers food to a listener while producing utterances 119
6.2 The relationship between acoustic stimuli consisting of a synthesized "da" and ABRs (according to Skoe & Kraus, 2013) 130
6.3 Functional effects of the phase of neural oscillations (according to Schroeder et al., 2008) 136
7.1 Representations and transformations from input signal to lexical representation as viewed by Poeppel et al. (2007) 145
7.2 An illustration of the EP hypothesis of intrinsic force-length relationships in muscle control 148

7.3 Idealized unitary mass displacement in a critically damped and under-damped spring system at three times (Boucher, 2008) 150
7.4 Guenther's DIVA control scheme (2016) 153
7.5 Rectified and smoothed EMG activity of opener and closer muscles of lip and jaw motions, along with midline articulator displacements, bilabial compression, and intra-oral pressure (Boucher, 2008) 159
7.6 Velocity and mid-sagittal displacement of lip and jaw opening as a function of force-related measures of closing motions (Boucher, 2008) 160
7.7 Fiber composition of intrinsic tongue muscles in four parts of the tongue (Stål et al., 2003) 161
7.8 Glottal motion and EMG activity of laryngeal muscles preceding the first vowel in "*I say*" (Hirose et al., 1980) 163
7.9 Overall correct serial recall of auditorily presented CV and VC sequences of nonsense syllables 165
7.10 Audio signal, intra-oral pressure, and kinematics of midline labial motions during the production of *Bobby* ['bɑbi] and *poppy* ['pɑpi] 167
8.1 VOTs for [ta] and [da] produced at varying rates of speech (Boucher, 2002) 174
8.2 Pitch (F_0), and dB energy patterns of utterance stimuli along with corresponding ERPs of regions of interests (Gilbert et al., 2015) 178
8.3 Inter-trial phase coherence (ITPC) for three types of stimuli as a function of the frequency of oscillations (Boucher et al., 2019) 180
9.1 Age-related changes in speech breathing recorded via plethysmographic belts 186
9.2 Vital capacities in liters and MLU in morphemes and syllables (Boucher & Lalonde, 2015) 189
9.3 MLU in morphemes and syllables as a function of vital capacity (Boucher & Lalonde, 2015) 189
9.4 Overall percentages of nominal forms of 1, 2, and 3 syllables or more used by 50 speakers aged 5–27 years (Boucher & Lalonde, 2015) 191
10.1 An illustration of the embodied approach in which the semantics of the lexeme *banana* is seen to be grounded in perception and action systems (Kemmerer, 2015) 210
10.2 Effects of rehearsing lexical items in different production conditions on participants' recall of having produced the items (Lafleur & Boucher, 2015) 216

10.3 Phase-amplitude coupling between theta-band oscillations and gamma oscillations (Canolty et al., 2006) 218
11.1 Analyses of LFPs in the subthalamic nucleus by Watson and Montgommery (2006) 236
E.1 An illustration of the relationship between neural responses, assumed "interim" units of language analysis, and speech stimuli 242

Tables

1.1 Part of Lindblom's (1986) predictions of "vowel systems" using a maximal perceptual distance metric and two acoustic-to-auditory mapping approaches *page* 32

11.1 Test batteries suggested by Murdoch and Whelan (2009) to observe the effects of pallidotomy, thalamotomy, and pathologies of the cerebellum on "high level linguistics" 224

Abbreviations

ABR	auditory brainstem response
ABSL	Al-Sayyid Bedouin Sign Language
ASL	American Sign Language
CNS	central nervous system
EEG, MEG	electroencephalography, magnetoencephalography
EGG	electroglottography
EMG	electromyography
EMMA	electromagnetic articulography
ERP	event-related potential
F_0	fundamental frequency perceived as pitch
F_1, F_2	formants
FAF	frequency altered feedback
FFR	frequency following response
IPA	International Phonetic Alphabet
ITPC	inter-trial phase coherence
LAD	Language Acquisition Device
LCA	last common ancestor
LFP	local field potential
MEP	motor-evoked potential
MLU	mean length of utterance
MRI, fMRI	magnetic resonance imaging, functional magnetic resonance imaging
tDCS	transcranial direct current stimulation
TMS	transcranial magnetic stimulation
TP	transition probability between transcribed units
VOT	voice onset time
VC	vital capacity

Preface

It is customary in a preface to indicate the source of a book, why it was written and for whom, and to thank those who helped forwarding the work through to publication. The present monograph is primarily directed at students and researchers in sectors relating to spoken language. It is also aimed at any interested reader wishing to understand the historical course and recent developments of language science extending to techniques of neuroscience. The book addresses a long-standing problem that may be apparent to anyone with minimal training in methods of linguistic analysis. Such training is part of introductory courses that are often a prerequisite in subprograms of psychology, language neuroscience, communication disorders, and language teaching, among other disciplines. The tradition has been that all who engage in the study of language are trained in analyzing transcribed speech and thus come to conceptualize spoken language by reference to theories erected on these analyses. This has definite consequences across sectors. By such training, many view spoken language as containing letter- and word-like units, organized in terms of given categories that are reminiscent of those used in codes of alphabet writing. But perhaps because of my field of interest (speech science), it has been persistently clear to me, as a student and researcher, that there are hardly any links between instrumental observations of speech and formal analyses of transcripts. This discrepancy was the source of a career-long interrogation on how it was that empirical observations did not serve to correct assumptions shared by analysts and investigators of spoken-language processing. In the meantime, I became acquainted with a body of historical essays, including a publication by Linell (1982/2005) entitled *The Written Language Bias in Linguistics*. These works documented how spoken language came to be studied using text and essentially demonstrated that the formal analysis of language, as currently taught in universities, is conceptually based on orthographic code. The essays exposed an important bias with broad implications, although the implications were not spelled out except by reference to the sociology of literacy and language theory. There was a need, as I saw, for a work that documented the course of the writing bias, the arguments used to claim the existence of orthographic-like units and categories in the brain, and the

consequences for experimental research. More pressingly, there was a need to address the bias by detailing how speech can be studied, not through the prism of one's writing system, but through instrumental techniques that could identify motor-sensory elements and structures of speech processing. However, I fully recognized the risks of such an endeavor.

I realized that exposing a bias across sectors of language science ran the risk of appearing confrontational on all fronts and that readers might not be aware of the accumulating evidence of a basic problem. To avoid such judgments, the monograph had to discuss experimental findings. It had to be made clear that, throughout the history of language science, investigators have explicitly acknowledged a basic discrepancy between theories erected on orthographic concepts and observed processes of spoken language.

But it was also a concern that a book that documents a bias in research would constitute a wholly negative enterprise – unless it proposed a way of *addressing* the problem. The monograph had to show that there is a coherent set of findings that supports an approach to the study of spoken language that does not entail notions of letter- and word-like units as in text. The discussion of this evidence, however, presented yet another risk. In this case, there was the chance that some vital findings could be missed, or else that the evidence would refer to domains of research unfamiliar to some readers. On this problem, the challenge was to remain on topic while assuming a knowledge base by which readers could judge competing proposals. As a compromise, I provide, especially in the latter parts of the book that refer to neuroscience, multiple references to recent surveys and critical reviews from which readers can cull background information. In short, I fully recognized the risks of submitting a work dealing with the writing bias in language science. However, far outweighing these considerations was the prospect of a science that seeks to understand the biological underpinnings of spoken language based on culture-specific constructs of writing. In other words, weighing in the balance was the prospect of pursuing studies of speech processes in a way that made the scientific status of the field appear questionable at its base.

The present work aims to address the writing bias through an approach that rests on observable structures of speech. This offers a view of how research may move forward in elaborating biologically plausible accounts of spoken language where speech observations are commensurate with neural processes. The evidence that is marshaled in support of the approach draws mostly from published work, though some pivotal findings are the product of my collaboration with colleagues whom I wish to thank. In particular, I am indebted to Boutheina Jemel (Université de Montréal), who designed the image of the book cover, and Annie C. Gilbert (McGill University). Both have had a major influence on my view of the role of neural oscillations in speech processing. Both have convinced me of the value of small laboratories where

experimentalists can transgress the boundaries of academic disciplines and share expertise. I also gratefully acknowledge the support of members of our team, especially Julien Plante-Hébert and Antonin Rossier-Bisaillon without whom there would have been no time to write. The essential parts of this monograph were developed in answering an invitation from Philippe Martin (Université Paris-Diderot) to deliver a series of conferences in his department, and I am truly grateful for his ongoing encouragement and discussions on prosodic structure. The format of the subject matter that follows benefited from the commentaries of students who attended my courses at the Université de Montréal. Hopefully, the monograph can serve to foster critical thinking in future students and researchers. I also thank Douglas Rideout for revising the text under the pressure of impending deadlines. Finally, in the context where there is considerable controversy in language theory, I wish to express my gratitude to my editor Helen Barton and Cambridge University Press for their open-mindedness and support.