

Preface

Reading the papers in Rebuschat and Williams' volume, "*Statistical learning and language acquisition*," brings me both back in time and looking ahead to the future. I suppose that is appropriate, given that the kind of learning at issue is precisely the kind that avails itself of prior experience to predict future events.

The construct of statistical learning is both intuitively appealing and frustratingly vague. The appeal of statistical learning, I believe, derives from its apparent simplicity: it would be sensible for learners to exploit distributions of events in their environments to predict future events. Unfortunately, the flip side of this apparent simplicity is that the construct is so easily applied that it is difficult to decide where statistical learning rightly begins and ends. Appropriately, then, the chapters in the current volume bring out both the pleasures and the pitfalls of accounts that invoke statistical learning mechanisms.

When Dick Aslin, Elissa Newport, and I began to work on our collaborative studies on infant and adult statistical language learning in the early 1990s, we were keenly aware of the history surrounding these ideas. Their roots lie in the structural linguistics of Leonard Bloomfield (1933) & Zellig Harris (1955), and in prior experimental and theoretical work by Hayes & Clark (1970), Goodsitt, Morgan, & Kuhl (1993), Braine (1966), Reber (1967), Morgan & Newport (1981), Maratsos & Chalkley (1981), and many others.

Despite the long history of research and debate surrounding these ideas, I did not anticipate the field's reaction to our initial infant studies. There were two interesting and surprising dimensions to those reactions. The first dimension spanned responses ranging from "*Duh!*" to "*Impossible!*" Some colleagues, particularly those in the visual sciences, responded to our initial studies by saying: "*Of course learners track statistics in environmental input; how could they not?*" At the other extreme, some readers questioned the idea that statistical information could have any efficacy whatsoever given the complexities of natural language: "*How could a learning ability that allows you to remember wallpaper patterns possibly have anything to do with real linguistic input?*" While the current incarnations of these perspectives are markedly less extreme, they continue to provide necessary counterpoints as we work to expand and refine our theories.

The second dimension is also still quite current. After we published our first paper on infant statistical language learning, some readers responded

by saying: “*Wow! This is real evidence for a language learning device; look at the speed and ease with which infants in these studies learned a novel linguistic structure.*” Other readers responded by saying: “*Wow! This is real evidence for a general learning device; these results suggest that language learning must be subserved by the same machinery that we use to learn across varied domains.*” Of course, that initial paper was not intended to directly address questions of domain-specificity versus domain-generality. But, as is evident from the papers in this volume, the subsequent decade has seen a great deal of research focused on this issue.

Reading these chapters, I’m struck by the breadth of questions that have emerged and reemerged over the past 2 decades of research on statistical learning. Issues of nativism and empiricism continue to fascinate us. Because learning requires both innate machinery and experience as input to that machinery, it is a fertile domain within which to explore nature-nurture questions. We continually return to issues of ecological validity, even as we constantly attempt to refine our methods to move closer to studying learning “in the wild.” We continue to grapple with fundamental issues: What are the computations that learners perform? What are the units over which those computations are performed? Are there developmental differences that affect which units are tracked and which computations are prioritized? What is the locus of constraints on learning, and does this differ across domains? What are the most appropriate ways to model statistical learning processes? What is the relationship between statistical learning and other key cognitive constructs (implicit learning, associative learning, procedural learning, working memory, attention, conscious awareness)? How should statistical learning fit with current thinking about language evolution and neural plasticity?

These are not new questions. But the ways in which the authors in this volume address them are new and very exciting. It is notable that some consensus has emerged: Nobody takes statistical learning for granted (“*Duh!*”), and nobody is arguing that these learning mechanisms are entirely irrelevant. We are all working to determine the role that statistical learning should play within our broader theories.

Finally, the papers in this volume suggest that accounts that invoke statistical learning mechanisms have moved into the mainstream of sub-fields well beyond first language acquisition. Second language acquisition is of particular interest, given that immersion in an L2 is essentially an implicit learning experience. Application to language and cognitive disorders is a natural extension, and research on individual differences has huge potential. Music is an ideal companion domain for research on language; the role of expectation has a long and illustrious history in music

theory (Meyer, 1956), and I expect that we will see the continued emergence of rich statistical learning accounts in this domain.

What will be contained in the next edition of this volume, a few years down the line? It's hard to know. Statistical learning accounts may become a full-fledged alternative to more traditional perspectives in language acquisition, as well as in other domains where these models are beginning to be applied (e.g., social cognition, perception for action, etc.). Or aspects of ideas from this framework may be integrated into other types of accounts, playing a role where needed. To a large extent, the future of statistical learning hinges on the answers to the questions laid out by the authors in this volume. I can't wait to find out what the answers are!

Jenny Saffran

References

- Bloomfield, L.
1933 *Language*. New York: Holt.
- Braine, M. D. S.
1966 Learning the positions of words relative to a marker element. *Journal of Experimental Psychology*, 72, 532–540.
- Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K.
1993 Perceptual strategies in prelingual speech segmentation. *Journal of Child Language*, 20, 229–252.
- Harris, Z. S.
1955 From phoneme to morpheme. *Language*, 31, 190–222.
- Hayes, J. R., & Clark, H. H.
1970 Experiments in the segmentation of an artificial speech analog. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.
- Maratsos, M., & Chalkley, M. A.
1980 The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In Katherine Nelson (Ed.), *Children's language Vol. 2*. (pp. 127–213). New York: Gardner Press.
- Meyer, L.
1956 *Emotion and meaning in music*. Chicago: University of Chicago Press.
- Morgan, J. L., and Newport, E. L.
1981 The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior*, 20, 67–85.
- Reber, A. S.
1967 Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.

List of contributors

Althea Bauernschmidt

Department of Psychological Sciences
Purdue University

Morten H. Christiansen

Department of Psychology
Cornell University

Christopher M. Conway

Department of Psychology
Saint Louis University

Zoltan Dienes

School of Psychology and Sackler
Centre for Consciousness Science
University of Sussex

Nick C. Ellis

Department of Psychology and
English Language Institute
University of Michigan

Judit Gervain

Laboratoire Psychologie de la
Perception
Institut Neurosciences Cognition
Université Paris Descartes

Michael H. Goldstein

Department of Psychology
Cornell University

Kalim Gonzales

Department of Psychology
University of Arizona

Rebecca Gómez

Department of Psychology
University of Arizona

Michelle A. Grempp

Central Institute for the Deaf
Washington University School of
Medicine

Phillip Hamrick

Department of Linguistics
Georgetown University

Jessica F. Hay

Department of Psychology
University of Tennessee – Knoxville

Jill Lany

Department of Psychology
University of Notre Dame

Psyche Loui

Department of Neurology
Beth Israel Deaconess Medical Center
and Harvard Medical School

Jacques Mehler

Language, Cognition and
Development Lab
Scuola Internazionale Superiore di
Studi Avanzati (SISSA), Trieste

Jennifer B. Misyak

Department of Psychology
Cornell University

Daniel Navarro

School of Psychology
University of Adelaide

Marina Nespore

Facoltà di Psicologia
Università di Milano Bicocca

Matthew Brook O'Donnell

English Language Institute
University of Michigan

Amy Perfors

School of Psychology
University of Adelaide

Pierre Perruchet

Laboratoire d'Etude de l'Apprentissage
et du Développement
CNRS-UMR 5022 Université de
Bourgogne, Dijon

David B. Pisoni

Department of Psychological and
Brain Sciences
Indiana University

Bénédicte Poulin-Charronnat

Laboratoire d'Etude de l'Apprentissage
et du Développement
CNRS-UMR 5022 Université de
Bourgogne, Dijon

Michelle Sandoval

Department of Psychology
University of Arizona

Mohinish Shukla

Department of Brain and Cognitive
Sciences
University of Rochester

Luca Onnis

Department of Second Language
Studies
University of Hawaii, Manoa

Patrick Rebuschat

School of Linguistics and English
Language
Bangor University

Kenny Smith

Language Evolution and Computation
Research Unit
School of Philosophy, Psychology and
Language Sciences
University of Edinburgh

Anne D. Walk

Department of Psychology
Saint Louis University

Geraint A. Wiggins

Department of Computing and Centre
for Cognition, Computation and
Culture
Goldsmiths, University of London

John N. Williams

Department of Theoretical and
Applied Linguistics
University of Cambridge

Table of contents

Preface	v
<i>Jenny Saffran</i>	
List of contributors	ix
Introduction: Statistical learning and language acquisition	1
<i>Patrick Rebuschat and John Williams</i>	
Statistical-sequential learning in development	13
<i>Jennifer B. Misyak, Michael H. Goldstein and Morten H. Christiansen</i>	
Bootstrapping language: Are infant statisticians up to the job?	55
<i>Elizabeth K. Johnson</i>	
Sensitivity to statistical information begets learning in early language development	91
<i>Jessica F. Hay and Jill Lany</i>	
Word segmentation: Trading the (new, but poor) concept of statistical computation for the (old, but richer) associative approach.	119
<i>Pierre Perruchet and Bénédicte Poulin-Charronnat</i>	
The road to word class acquisition is paved with distributional and sound cues	145
<i>Michelle Sandoval, Kalim Gonzales and Rebecca Gómez</i>	
Linguistic constraints on statistical learning in early language acquisition	171
<i>Mohinish Shukla, Judit Gervain, Jacques Mehler and Marina Nespor</i>	
The potential contribution of statistical learning to second language acquisition	203
<i>Luca Onnis</i>	
Statistical learning and syntax: What can be learned, and what difference does meaning make?	237
<i>John N. Williams and Patrick Rebuschat</i>	

Statistical construction learning: Does a Zipfian problem space ensure robust language learning?	265
<i>Nick C. Ellis and Matthew Brook O'Donnell</i>	
Can we enhance domain-general learning abilities to improve language function?	305
<i>Christopher M. Conway, Michelle A. Greppe, Anne D. Walk, Althea Bauernschmidt and David B. Pisoni</i>	
Conscious versus unconscious learning of structure.	337
<i>Zoltan Dienes</i>	
How implicit is statistical learning?	365
<i>Phillip Hamrick and Patrick Rebuschat</i>	
What Bayesian modelling can tell us about statistical learning: what it requires and why it works.	383
<i>Amy Perfors and Daniel J. Navarro</i>	
Evolutionary perspectives on statistical learning	409
<i>Kenny Smith</i>	
Statistical learning: What can music tell us?	433
<i>Psyche Loui</i>	
“I let the music speak”: Cross-domain application of a cognitive model of musical learning	463
<i>Geraint A. Wiggins</i>	
Index	495

Introduction: Statistical learning and language acquisition

Patrick Rebuschat and John Williams

Recent years have witnessed an increasing interest in empiricist approaches to language acquisition (see Behrens, 2009; Ellis, 2006a, 2006b; Elman et al., 1996; Goldberg, 2006; MacWhinney, 1999; Redington & Chater, 1998; Tomasello, 2003). This development was driven, in part, by two observations, namely that (i) infants' environment is considerably richer in linguistic and non-linguistic cues than previously anticipated and that (ii) infants are able to make extensive use of these cues when acquiring language. Both findings suggest a greater role for learning than traditionally assumed by nativist approaches to language development (e.g. Anderson & Lightfoot, 2002; Chomsky, 1966, 1986, 1988; Crain & Pietroski, 2001; Roeper & Williams, 1987). Among empiricist approaches, research conducted on *statistical learning*, i.e. our ability to make use of statistical information in the environment to bootstrap language acquisition, has been particularly fruitful.

Statistical learning research was sparked by the work of Jenny Saffran, Elissa Newport, and Richard Aslin (Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996) and developed into a major research strand in developmental psychology (see Gómez, 2007; Saffran, 2003, for reviews). Statistical learning involves computations based on units or patterns, which can include linguistic elements such as speech sounds, syllables, syntactic categories and form-meaning mappings. The types of statistical computation range from simple frequency counts to the tracking of co-occurrence information and conditional probability. Research in statistical learning generally focuses on infant or child language acquisition, though studies with adult subjects are also common. In terms of methodology, the most distinctive features of statistical learning research are the careful manipulation of statistical information in the input and the use of artificial languages (see Gómez & Gerken, 2000, for a review).

In their seminal study, Saffran, Aslin, and Newport (1996) investigated whether 8-month-old infants could use statistical information to solve the problem of word segmentation, i.e. to discover word boundaries in running speech. Infants were exposed to two minutes of a continuous speech

stream that contained four three-syllable nonsense words (e.g., *tupiro*, *padoti*). The “words” were repeated in random order, and a speech synthesizer was used to generate a continuous auditory sequence (e.g., *bidakupadotigolabubidakupadotigolabubidakutupiro...*). The sequence contained no pauses, stress differences or any other acoustic cues between words, so that the only cue to word boundaries were the transitional probabilities between syllables. The transitional probability *within* words was 1.0, given that the first syllable of a word was always followed by the second, and the second syllable by the third (e.g., *tu-* was always followed by *-pi-*, and *-pi-* followed by *-ro*). The transitional probability *between* words was 0.33 because the final syllable of a given word could be followed by the initial syllable of three different words (e.g., *-ro* could be followed *go-*, *bi-*, or *pa-*). Infants were then tested by means of the head-turn preference procedure to determine whether they could recognize the difference between trained items (*tupiro*, *golabu*) and novel items (*dapiku*, *tilado*). Saffran, Aslin and Newport (1996) found that the 8-month-olds successfully discriminated between familiar and unfamiliar stimuli, which suggests that infants are highly sensitive to statistical information (here, transitional probabilities) and that they can use this information to succeed in a complex learning task (word segmentation).

This early research on statistical learning was important for demonstrating that infants are “intuitive statisticians” (Ellis, 2006b), who are able to make extensive use of environmental cues when acquiring language. Importantly, subsequent research has shown that the capacity for statistical learning is maintained throughout adulthood (e.g., Saffran, Newport, & Aslin, 1996) and that statistical learning is not restricted to the task of word segmentation. After more than a decade of experimental research, there is ample evidence that both infants and adults can exploit the statistical structure of their environment in order to succeed in a wide variety of linguistic tasks, including phonological learning (e.g., Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002), word learning (e.g., Estes, Evans, Alibali, & Saffran, 2007; Yu & Smith, 2007; Smith & Yu, 2008) and syntactic development (e.g. Gerken, Wilson, & Lewis, 2005; Saffran & Wilson, 2003; Thompson & Newport, 2007). There is also evidence that the cognitive mechanism involved in statistical learning is not specific to language acquisition but rather domain-general in nature, i.e. the learning mechanism applies to statistical information in the environment, irrespective of the nature of the stimulus (auditory, visual, tactile, etc.; see Saffran & Thiessen, 2007, for discussion). For example, several experiments have demonstrated that infants and adults can track sequential statistics in non-

linguistic auditory stimuli (e.g., Saffran, Johnson, Aslin, & Newport, 1999) and visual stimuli (e.g., Bulf, Johnson, & Valenz, 2011; Fiser & Aslin, 2002a, 2002b). Studies on cotton-top tamarin monkeys (Hauser, Newport, & Aslin, 2001) and rodents (e.g., Toro & Trobalón, 2004) further suggest that basic aspects of statistical learning are not unique to human learners. Finally, it is widely accepted that the process of statistical learning can occur incidentally, i.e. subjects can acquire the statistical structure of language without the conscious intention to learn, making the process of statistical learning analogous to that of implicit learning (see also Dienes, this volume; Hamrick & Rebuschat, this volume; Misyak, Goldstein, & Christiansen, this volume).

This Volume

The present volume brings together researchers from a variety of disciplines (cognitive psychology, computer science, corpus linguistics, developmental psychology, psycholinguistics) in order to assess the progress made in statistical learning research, to critically appraise the role of statistical learning in language acquisition, and to determine future directions to take in this interdisciplinary enterprise. The volume was inspired by an eponymous symposium which the editors organized for the 2009 edition of the Georgetown University Round Table on Languages and Linguistics (GURT). The feedback we received from the symposium presenters and conference delegates was very positive throughout, and when we were approached by Mouton de Gruyter regarding the possibility of producing an edited volume on the same topic we readily agreed to do so. Three presentations of our original symposium were converted into much expanded and updated chapters (Ellis & O'Donnell, Williams & Rebuschat, and Hay & Lany). The remaining contributors were recruited specifically for this volume.

Each chapter in this volume was peer-reviewed by 2–3 anonymous reviewers and by the two editors. In addition, many chapters were used as readings in a postgraduate course on the Implicit and Explicit Learning of Languages (Ling-494), offered by the first editor at Georgetown University. This enabled us to gain feedback on the readability of texts and on the clarity of the arguments expressed by the authors. The final product is a volume that is written in an accessible and engaging fashion and that gives readers a snapshot of the exciting research that has examined the role of statistical learning in language acquisition.

In Chapter 1, Jennifer Misyak, Michael Goldstein and Morten Christiansen focus on two distinct, but closely related research traditions, namely implicit learning (Reber, 1967) and statistical learning (Saffran, Aslin, & Newport, 1996). Both approaches focus on how we acquire information from the environment and both rely heavily on the use of artificial grammars. Perruchet & Pacton (2006) suggested that implicit and statistical learning represent two approaches to a single phenomenon. Conway & Christiansen (2006) go as far as combining the two in name: *implicit statistical learning*. Misyak, Goldstein and Christiansen's aim is to promote the synergistic fusion of the two approaches by highlighting theoretical and methodological similarities and by providing researchers with a thorough and much-needed synthesis of current research in both fields.

In Chapter 2, Elizabeth Johnson evaluates the contribution of statistical learning to solving the bootstrapping problem. The chapter focuses on infant learners and the task of word segmentation, but Johnson's observations apply to many levels of spoken language acquisition. She first provides a brief overview of the progress made in statistical learning research. This is followed by an engaging discussion of five questions and challenges faced by distributional models of language development. Does the ability to track patterns in an artificial language scale up to the challenge of natural language? What are the units that language learners keep track of, and what type of calculations do they perform? Can distributional models predict children's difficulties? How much knowledge is innate and how much is acquired? and How to interpret looking-time data in statistical learning research?

In Chapter 3, Jessica Hay and Jill Lany also concentrate on the role of statistical information in infant language development. Their chapter begins with three important observations. Firstly, many of the early statistical learning experiments employed artificial languages that lack the rich, multidimensional structure of natural language. Secondly, many studies presented subjects with stimuli that are devoid of semantic information. Both of these aspects arguably reduce the ecological validity of studies. Thirdly, early research has little to say about how statistical learning at one level (e.g. syllables) relates to statistical learning about other aspects of language (e.g., word classes). Hay and Lany then describe several recent studies that have begun to address these gaps. The work reviewed in their chapter shows that infants are highly adept at tracking statistical regularities in an artificial language even with tasks that closer approximate the problems faced over the course of learning a natural language. Importantly, this research also shows how sensitivity to statistical structure in one area

of language can bootstrap the learning of other, more complex dimensions of language structure.

In Chapter 4, Pierre Perruchet and Bénédicte Poulin-Charronat propose that statistical learning phenomena can be interpreted as end-products of associative learning processes and that the associative approach can provide a stronger and more appropriate framework within which to examine statistical learning. Their emphasis is on the widely-studied task of word segmentation. After describing their thesis in detail, Perruchet and Poulin-Charronat discuss different explanations for our sensitivity to statistical structure (associative, attention-based, and interference-based accounts). They then explore how statistical computation can be integrated with other factors that are known to play an important role in word segmentation (acoustical cues and contextual information) in a unified, dynamic perspective that is based on the associative learning tradition. They conclude by considering evidence from behavioural experiments and computational modelling.

In Chapter 5, Michelle Sandoval, Kalim Gonzales and Rebecca Gomez focus on the acquisition of word classes. They first consider three cues to word class – distributional, phonological and prosodic – and review studies that examined the role of these cues in the acquisition of lexical categories. Sandoval, Gonzales and Gomez then discuss how these multiple sources of information are integrated in word class acquisition. Their chapter concludes with a discussion of how learners might scale up from purely form-based categories to lexical classes.

In Chapter 6, Mohinish Shukla, Judit Gervain, Jacques Mehler and Marina Nespors suggest that a synthesis between rationalist and empiricist approaches might be necessary to account for a complex phenomenon like language acquisition and propose that three types of mechanisms – rule-based, distributional and perceptual – are required to explain how languages are acquired. The authors begin by defining statistical learning and by reviewing several key studies. In the following sections, they then investigate how a powerful, domain-general statistical learning mechanism interacts with other, language-specific and perceptual processes. Specifically, they consider how linguistic representations constrain the use of statistical information at the phonemic, morphological, syntactic, and prosodic levels.

In Chapter 7, Luca Onnis reflects on the potential contribution of statistical learning to second language (L2) acquisition. In the first part, Onnis discusses four principles based on statistical learning research that can be applied to L2 learning scenarios. These general learning principles are: (i) “Integrate probabilistic sources of information”, (ii) “Seek invariance in

the signal”, (iii) “Reuse learning mechanisms”, and (iv) “Learn to predict.” In the second part, Onnis then elaborates on how these principles can be put to use for specific problems arising in L2 acquisition and teaching. He considers evidence from both behavioural experiments and computational analyses of corpora.

In Chapter 8, John Williams and Patrick Rebuschat focus on the acquisition of second language (L2) syntax in adult learners. Their chapter discusses the contribution of statistical learning to L2 syntactic development and the role of prior linguistic knowledge. An obvious criticism of artificial language experiments is that learners are often exposed to meaningless stimuli. Williams and Rebuschat describe a series of experiments that employed semi-artificial languages, i.e. systems in which the complexity of natural language was maintained and semantic information present. Their findings support the view that syntactic structure can be induced from an analysis of the contingencies between words. However, they also suggest that there are limitations to what can be learned.

In Chapter 9, Nick Ellis and Matt O’Donnell present the results of a corpus analysis that was designed to test the generalizability of construction grammar theories of language learning. The linguistic focus is on Verb-Argument Constructions (VACs); the corpus in question is the British National Corpus (BNC). The chapter begins with a description of the main tenets of construction grammar and usage-based approaches to language acquisition. This is followed by a discussion of determinants of construction learning (frequency, function, and contingency of form-function mapping). The next section of the chapter is dedicated to a thorough description of the corpus analysis and its results. Ellis and O’Donnell find that constructions are Zipfian in their type-token distributions in usage, selective in their verb form occupancy, and coherent in their semantics. They suggest that these characteristics make linguistic constructions robustly learnable by a statistical learning mechanism.

In Chapter 10, Christopher Conway, Michelle Grep, Anne Walk, Althea Bauernschmidt, and David Pisoni discuss whether statistical learning abilities can be enhanced to improve language function. They begin by reviewing evidence highlighting the importance of statistical learning in language acquisition and processing. They then describe recent research that used computerized training techniques that were designed to improve working memory. This provides the background for a discussion of two studies that assessed the effectiveness of a new adaptive training task for improving domain-general learning abilities. The first study focuses on

adult subjects with normal hearing. The second study considers children who are deaf or hard of hearing. Conway, Grep, Walk, Bauernschmidt and Pisoni's findings confirm that the basic mechanisms of learning and memory can be trained, and that training tasks such as theirs might be employed as an intervention for treating disorders of language and learning.

One of the widely discussed questions in implicit learning research is whether the knowledge acquired in sequence learning and artificial grammar experiments is, in fact, implicit. In Chapter 11, Zoltan Dienes presents a methodology for determining the conscious (explicit) and unconscious (implicit) status of knowledge. Dienes first provides a definition of unconscious knowledge. He then discusses different measures of awareness, with a special emphasis on subjective measures. After introducing the distinction between structural and judgment knowledge, Dienes then presents extensive evidence in support of subjective measures of awareness.

In Chapter 12, Phillip Hamrick and Patrick Rebuschat describe an experiment that investigated whether a typical statistical learning experiment results in implicit knowledge, explicit knowledge, or both. The experiment combined the cross-situational word learning paradigm (Yu & Smith, 2007) and the subjective measures of awareness developed by Dienes (this volume; Dienes & Scott, 2005). Subjects were either exposed under incidental or intentional learning conditions. Hamrick and Rebuschat found clear learning effects under both conditions. However, subjects in the intentional group developed both implicit and explicit knowledge, while the subjects in the incidental group developed primarily implicit knowledge. The experiment illustrates the usefulness of including measures of awareness when researching statistical learning.

In Chapter 13, Amy Perfors and Daniel Navarro explore the why and what of statistical learning from a computational modelling perspective. Perfors and Navarro propose that Bayesian techniques can be particularly useful for understanding what kinds of learners and assumptions are necessary for successful statistical learning. Their chapter begins with a brief introduction to Bayesian modelling, contrasting it with the other widely-used computational approach to statistical learning (connectionism). The remaining chapter is structured around a series of key questions: What is statistical learning? What data does statistical learning operate on? What knowledge does learner acquire from the data? What assumptions do learners make about the data? What prior knowledge does the learner possess? Finally, why does statistical learning work?

In Chapter 14, Kenny Smith approaches the topic of statistical learning from an evolutionary perspective. Smith first describes generative and non-generative approaches to language universals and language evolution. He then discusses recent research on linguistic variation as a test-case for exploring debates on the link between learning biases and universals in language design. The chapter concludes with a discussion of the biological evolution of the language faculty.

In Chapter 15, Psyche Loui approaches the topic of statistical learning from a non-linguistic perspective, with a special focus on music. The central thesis of her chapter is that much of our musical knowledge can be acquired by means of experience with the statistical regularities in the input. Loui begins her chapter with a discussion of the modality-independence of statistical learning and then briefly reviews research on how we acquire implicit knowledge of music. This sets the stage for a description of several of Loui's experiments on the acquisition of an artificial musical system by adult learners. The artificial system is based on the Bohlen-Pierce scale, a novel scale that is entirely different from existing musical systems. Loui's paradigm allowed her to address several important questions, e.g. What aspects of musical structure can be learned? How quickly can we acquire pitch, timbre, etc.? How much does emotion in music depend on statistical regularities? The chapter concludes with an outline of possible future directions.

In Chapter 16, Geraint Wiggins presents the Information Dynamics of Music (IDyOM) model of musical melody processing. A special feature of this model is its multidimensionality, i.e. it is capable of modelling perceptual phenomena whose percepts are multidimensional constructs. Importantly, even though it was designed as a model of melody processing, IDyOM can be applied to other, non-musical domains. Wiggins takes a strong view of statistical learning, in which statistical estimation is paramount in cognition. IDyOM is not presented merely as a way of capturing regularities in the observed data, but as a theory of the processing mechanism itself. That is, IDyOM is viewed as a simulation of actual cognitive processing. The chapter begins with a discussion of the relationship between language and music and a survey of the relevant literature in statistical linguistics. Wiggins then presents a detailed overview of IDyOM. The chapter concludes with a study that explored whether IDyOM is able to model a basic linguistic task (syllable identification) by means of the same information theoretic principles that apply in melody segmentation.

Acknowledgements

The volume, and the symposium on which it was based, would not have been possible without the extensive help of many people. We would like to thank Ron Leow and Cristina Sanz, the organizers of the 2009 edition of GURT, for inviting us to organize a symposium, and we thank our presenters for making it such a successful event. With regards to the volume, we are most grateful to our authors for their excellent contributions and for agreeing to peer-review several chapters. Without their hard work, there would be no volume. At Mouton de Gruyter, we are very grateful to Cathleen Petree for proposing this project in the first place. Sadly, Cathleen passed away only a few months after we began working on the volume, and we are saddened that she is not around to see the finished book. We would like to thank Emily Farrell, our new editor at Mouton de Gruyter, for her continued support and Wolfgang Konwitschny, our production editor, for his assistance in producing the volume. At Georgetown, we are very grateful to Elizabeth Kissling, who worked as our editorial assistant and made our lives significantly easier. Finally, several PhD students provided their feedback on the chapters, for which we would like to thank them: Phillip Hamrick, Katie Kim, Elizabeth Kissling, Julie Lake, and Kaitlyn Tagarelli.

References

- Anderson, S. R., & Lightfoot, D.
2002 *The language organ: Linguistics as cognitive psychology*. Cambridge: Cambridge University Press.
- Behrens, H.
2009 Usage-based and emergentist approaches to language acquisition. *Linguistics*, 47(2), 383–411.
- Bulf, H., Johnson, S. P., & Valenza, E.
2011 Visual statistical learning in the newborn infant. *Cognition*, 121, 127–132.
- Chomsky, N.
1966 *Cartesian linguistics: a chapter in the history of rationalist thought*. New York: Harper & Row.
- Chomsky, N.
1986 *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Chomsky, N.
1988 *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: MIT Press.

- Conway, C. M., & Christiansen, M. H.
 2006 Statistical learning within and between modalities. *Psychological Science*, 17, 905–912.
- Crain, S., & Pietroski, P.
 2001 Nature, nurture and Universal Grammar. *Linguistics & Philosophy*, 24, 139–186.
- Dienes, Z.
 this volume Conscious versus unconscious learning of structure. In P. Rebuschat & J. N. Williams (Eds.) *Statistical learning and language acquisition*. Berlin: Mouton de Gruyter.
- Dienes, Z., & Scott, R.
 2005 Measuring unconscious knowledge: distinguishing structural knowledge and judgment knowledge. *Psychological Research*, 69, 338–351.
- Ellis, N. C.
 2006a Cognitive perspectives on SLA: The Associative-Cognitive CREED. *AILA Review*, 19, 100–121.
- Ellis, N. C.
 2006b Language acquisition as rational contingency learning *Applied Linguistics*, 27, 1–24.
- Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D. & Plunkett, K.
 1996 *Rethinking innateness: a connectionist perspective on development*. Cambridge, MA: MIT Press (1996).
- Fiser, J., & Aslin, R. N.
 2002a Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 458–467.
- Fiser, J., & Aslin, R. N.
 2002b Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*, 99, 15822–15826.
- Gerken, L., Wilson, R., & Lewis, W.
 2005 Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32, 249–268.
- Goldberg, A. E.
 2006 *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.
- Gómez, R.
 2007 Statistical learning in infant language development, In M. Gaskell (ed.), *The Oxford handbook of psycholinguistics* (pp. 601–616), Oxford: Oxford University Press.
- Gómez, R. L., & Gerken, L.
 2000 Infant artificial language learning and language acquisition, *Trends in Cognitive Sciences*, 4(5), 178–186.

- Graf Estes, K. M., Evans, J., Alibali, M. W., & Saffran, J. R.
2007 Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, 18, 254–260.
- Hamrick, P., & Rebuschat, P.
this volume How implicit is statistical learning? In P. Rebuschat & J. N. Williams (Eds.) *Statistical learning and language acquisition*. Berlin: Mouton de Gruyter.
- Hauser, M., Newport, E. L., & Aslin, R. N.
2001 Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition*, 78, B41–B52.
- Lieven, E., & Tomasello, M.
2008 Children's first language acquisition from a usage-based perspective: In P. Robinson & N. Ellis (eds). *Handbook of cognitive linguistics and second language acquisition*. (pp. 168–196) London: Routledge.
- MacWhinney, B. (Ed.)
1999 *Emergence of language*. Mahwah, NJ: Lawrence Erlbaum.
- Maye, J., Weiss, D. J., & Aslin, R. N.
2008 Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, 11, 122–134.
- Maye, J., Werker, J. F., & Gerken, L.
2002 Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.
- Misyak, J. B., Goldstein, M. H., & Christiansen, M. H.
this volume Statistical-sequential learning in development. In P. Rebuschat & J. N. Williams (Eds.) *Statistical learning and language acquisition*. Berlin: Mouton de Gruyter.
- Perruchet, P., & Pacton, S.
2006 Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, 10, 233–238.
- Reber, A. S.
1967 Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.
- Redington, M. & Chater, N.
1998 Connectionist and statistical approaches to language acquisition: A distributional perspective. *Language and Cognitive Processes*, 13, 129–191.
- Roeper, T., & Williams, E. (Eds.)
1987 *Parameter Setting*. Reidel: Amsterdam.
- Saffran, J. R.
2003 Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, 12, 110–114.

- Saffran, J. R., & Thiessen, E. D.
2007 Domain-general learning capacities, In E. Hoff & M. Schatz (eds.), *Blackwell handbook of language development* (pp. 68–86), Oxford: Blackwell.
- Saffran, J. R., & Wilson, D. P.
2003 From syllables to syntax: Multi-level statistical learning by 12-month-old infants. *Infancy*, 4, 273–284.
- Saffran, J. R., Aslin, R. N., & Newport, E. L.
1996 Statistical learning by 8-month-old infants, *Science*, 274, 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L.
1999 Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52.
- Saffran, J. R., Newport, E. L., & Aslin, R. N.
1996 Word segmentation: The role of distributional cues, *Journal of Memory and Language*, 35, 606–621.
- Smith, L., & Yu, C.
2008 Infants rapidly learn word–referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568.
- Thompson, S. P., & Newport, E. L.
2007 Statistical learning of syntax: The role of transitional probability, *Language Learning and Development*, 3, 1–42.
- Tomasello, M.
2003 *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Toro, J. M. & Trolalón, J. B.
2005 Statistical computations over a speech stream in a rodent, *Perception & Psychophysics*, 67, 867–875.
- Yu, C., & Smith, L. B.
2007 Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, 18, 414–420.