

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Keeping the hands in mind : what elicited pantomime reveals about language structure

Permalink

<https://escholarship.org/uc/item/72p8w0fd>

Author

Hall, Matthew L.

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Keeping the hands in mind: What elicited pantomime reveals about language structure

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor
of Philosophy

in

Psychology

by

Matthew L. Hall

Committee in charge:

Professor Victor Ferreira, Chair
Professor Rachel Mayberry, Co-Chair
Professor David Barner
Professor Karen Emmorey
Professor Tamar Gollan
Professor Carol Padden

2012

The dissertation of Matthew Lee Hall is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2012

DEDICATION

I dedicate this dissertation to all those who have inspired and supported me along the way. Although I cannot possibly list everyone who deserves recognition, I would like to thank the following people directly: my father, whose global travels first sparked my curiosity about language; Bill DeTillo, whose instruction in English grammar cemented my love of language; the late Ross Parham, who taught our 7th grade English class about Noam Chomsky and promptly instructed us to forget him; Bill Sayles, who helped me realize that I wanted scientific discovery to be my contribution to the world; the ASL and BCS faculty at the University of Rochester, who prepared me for graduate school; and the incredible community of language researchers at UC San Diego, whose input and feedback have been invaluable over the past 6 years- especially my advisors and the rest of my committee.

Finally, to my friends: Eugene, Rahul, Mike, Nicole, Fareed, Ami, Khalil, Jeanette, Ben, Jason, Jeremy, Mac, Erin, Lisa, Paula, Ashton, Rebecca, Stephanie, Dave, Emily, Andrew, Aaron, Jessie, Alicen, Ron, Matt, Dan, Jon, Cara, Naomi, Eric, Andre, Betsy, Ryan, Emily, Megan, and especially Ben: all of you, together with many others too numerous to name, have been a true godsend to me in more ways than I can count. Thank you for your love, companionship, acceptance, hospitality, guidance, patience, prayers, silliness, and friendship. This dissertation would not exist without you, and I can only hope to be half the friend to you that you have been to me.

Ecclesiastes 12:12b

Of the making of many books there is no end, and much study wearies the body.

EPIGRAPH

If we had no faculty of speech, how should we communicate with one another? Should we not use signs, like the deaf and dumb? The elevation of our hands would mean lightness; heaviness would be expressed by letting them drop. The running of any animal would be described by a similar movement of our own frames. The body can only express anything by imitation; and the tongue or mouth can imitate as well as the rest of the body.

-Plato (Cratylus)

TABLE OF CONTENTS

Signature Page.....	iii
Dedication.....	iv
Epigraph.....	v
Table of Contents.....	vi
List of Figures.....	viii
List of Tables.....	x
Acknowledgements.....	xi
Vita.....	xii
Abstract.....	xvii
Chapter 1: Introduction.....	1
Chapter 2: Cognitive constraints on constituent order: Evidence from elicited pantomime.....	17
Experiment 1	
Methods.....	27
Results.....	32
Discussion.....	34
Experiment 2	
Methods.....	39
Results.....	40
Discussion.....	43
Experiment 3	
Methods.....	45
Results.....	46
Discussion.....	50
Chapter 3: Investigating constituent order change with the hands: Consistency promotes the emergence of SVO in pantomime.....	68
Experiment 1	
Methods.....	83

Results.....	86
Discussion.....	90
Experiment 2	
Methods.....	92
Results.....	94
Discussion.....	98
Chapter 4: Comprehenders prefer SVO, but producers prefer SOV: Evidence from elicited pantomime.....	112
Experiment 1	
Methods.....	126
Results.....	130
Discussion.....	132
Experiment 2	
Methods.....	134
Results.....	136
Discussion.....	137
Chapter 5: General Discussion.....	156

LIST OF FIGURES

Chapter 2, Figure 1. Constituent orders across all participants, for non-reversible (upper) and reversible (lower) trials. For reversible events, Region A shows orders that are inconsistent with both the confusability hypothesis and the role conflict hypothesis. Region B shows orders that are uniquely consistent with confusability. Region C shows orders that are consistent with both confusability and role conflict. Region D shows orders that are uniquely consistent with role conflict. Region E shows orders that are uninformative with respect to these hypotheses. Rare orders are defined as those that account for less than 2% of trials for both non-reversible and reversible events.....	31
Chapter 2, Figure 2. Proportion of trials containing gestures that resembled case markers, for non-reversible and reversible trials.....	34
Chapter 2, Figure 3. Constituent orders across all participants, for non-reversible (upper) and reversible (lower) trials. For reversible events, Regions A-E are defined as in Figure 1.....	40
Chapter 2, Figure 4. Proportion of trials containing gestures that resembled case markers, for non-reversible and reversible trials.....	42
Chapter 2, Figure 5. Constituent orders across all participants in each condition, for non-reversible (upper) and reversible (lower) trials. For reversible events, Regions A-E are defined as in Figure 1.....	47
Chapter 2, Figure 6. Proportion of trials containing gestures that resembled case markers, for non-reversible and reversible trials, by condition.....	48
Chapter 3, Figure 1. Distribution of constituent orders in Experiment 1 (English speakers). Region A shows efficient orders with subject before object. Region B shows efficient orders with object before subject. Region C shows repetitious orders with subject before object. Region D shows repetitious orders with object before subject. Region E shows underinformative orders. Region F shows orders that were coded as ambiguous, simultaneous, or excluded due to lack of consensus.....	87
Chapter 3, Figure 2. Distribution of constituent orders in Experiment 2 (Turkish speakers). Region A shows efficient orders with subject before object. Region B shows efficient orders with object before subject. Region C shows repetitious orders with subject before object. Region D shows repetitious orders with object before subject. Region E shows underinformative orders. Region F shows orders that were coded as ambiguous, simultaneous, or excluded due to lack of consensus.....	94
Chapter 4, Figure 1. Schematic illustration of the experimental paradigm. Event clips and pantomime clips were not visible at the same time. Labels are included for illustration	

only; the only text displayed to participants was “less likely... more likely”, which was translated as “çok iliskisiz ... çok iliski” for Experiment 2.....	127
Chapter 4, Figure 2. Results of Experiment 1, which tested native English speakers. Error bars represent SEM.....	131
Chapter 4, Figure 3. Results of Experiment 2, which tested native Turkish speakers. Error bars represent SEM.....	136

LIST OF TABLES

Chapter 3, Table 1. Language screening questions for Experiment 2.....93

Chapter 4, Table 1. Language screening questions for Experiment 2.....135

ACKNOWLEDGEMENTS

Chapter 2, in part, has been submitted for publication of the material as it appears in 2012, Hall, M. L. Mayberry, R. I., & Ferreira, V. S. Cognitive constraints on constituent order: Evidence from elicited pantomime. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in part, has been submitted for publication of the material as it appears in 2012, Hall, M. L., Ferreira, V. S., & Mayberry, R. I. Investigating constituent order change with the hands: Consistency promotes the emergence of SVO in pantomime. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, has been submitted for publication of the material as it appears in 2012, Hall, M. L., Mayberry, R. I., & Ferreira, V. S. Comprehenders prefer SVO, but producers prefer SOV: Evidence from elicited pantomime. The dissertation author was the primary investigator and author of this paper.

CURRICULUM VITAE

EDUCATION

- 2012 Ph.D. in Cognitive Psychology
University of California, San Diego
Thesis: *Keeping the Hands in Mind: What Elicited Pantomime Reveals about Language Structure.*
Co-chairs: Victor Ferreira & Rachel Mayberry
- 2008 - 2012 Doctoral candidate in Cognitive Psychology
University of California, San Diego
- 2006 - 2008 M.A. in Cognitive Psychology
Thesis: *Is Syntactic Knowledge Modality-Independent? Evidence from syntactic priming.*
Co-chairs: Victor Ferreira & Rachel Mayberry
University of California, San Diego
- 1999 - 2003 B.A. in Brain & Cognitive Sciences with Highest Distinction and Honors in Research
B.A. in American Sign Language with High Distinction
Thesis: *Short-term memory in second language learners of American Sign Language.*
Chair: Elissa Newport
University of Rochester

PROFESSIONAL EXPERIENCE

- 2003 - 2006 Research Assistant, Brain & Vision Lab. PI: Daphne Bavelier
University of Rochester

ACADEMIC HONORS

- 2012 Outstanding Graduate Student Award Finalist, UC San Diego
- 2011 – 2012 Rita L. Atkinson Fellow, UC San Diego
- 2010 – 2011 Predoctoral Fellow; Center for Research in Language, UC San Diego
- 2010 Norman Anderson Travel Award
- 2009 TA Excellence Award; Linguistics Department, UC San Diego
- 2007 National Science Foundation Graduate Fellowship Honorable Mention
- 1999 – 2003 Magna Cum Laude, Phi Beta Kappa, Golden Key, Dean's List

RESEARCH INTERESTS

Sign languages, psycholinguistics, language evolution, bilingualism, language acquisition, cognitive development.

PUBLICATIONS

- Hall, M. L.,** Mayberry, R. I., & Ferreira, V. S. (submitted). Comprehenders prefer SVO, but producers prefer SOV: Evidence from elicited pantomime.
- Hall, M. L.,** Ferreira, V. S., & Mayberry, R. I. (submitted). Investigating constituent order change with the hands: Consistency promotes the emergence of SVO in pantomime.
- Hall, M. L.,** Mayberry, R. I., & Ferreira, V.S. (submitted). Cognitive constraints on constituent order: Evidence from elicited pantomime.
- Hall, M. L.,** Ferreira, V.S., & Mayberry, R. I. (2012). Phonological similarity judgments in ASL: Evidence for maturational constraints on phonetic perception in sign. *Sign Language & Linguistics*, 15, 104-127. doi: 10.1075/sll.15.1.05hal
- Hall, M. L.** (2011). Bilingual picture naming studies constrain theories of lexical access. *Frontiers in Psychology* 2, 381. doi: 10.3389/fpsyg.2011.00381
- Hall, M. L.,** & Bavelier, D. (2011). Short-term memory in sign and speech: The source of the serial span discrepancy. *Cognition*, 120, 54-66.
- Hall, M. L.,** & Bavelier, D. (2010). Working memory, deafness, and sign language. In M. Marschark and P. Spencer (Eds.) *Oxford Handbook of Deaf Studies, Language, and Education, Volume 2*.
- Bavelier, D., Newport, E. L., **Hall, M.,** Supalla, T., & Boutla, M. (2008). Ordered short-term memory differs in signers and speakers: Implications for models of short-term memory. *Cognition*, 107, 433-459.
- Bavelier, D., Newport, E. L., **Hall, M. L.,** Supalla, T., & Boutla, M. (2006). Persistent difference in short-term memory span between sign and speech: Implications for cross-linguistic comparisons. *Psychological Science*, 17(12), 1090-1092.
- Hall, M.,** Newport, E., Supalla, T., & Bavelier, D. (2003) Short Term Memory in Second Language Learners of ASL: the Competing Effects of Modality and Mother Tongue. *Journal of Undergraduate Research (University of Rochester)*, 2(2), 30-35.

PAPERS IN PREPARATION

- Hall, M. L.,** Ferreira, V.S., & Mayberry, R. I. (in preparation). Syntactic priming in American Sign Language.

Mayberry, R. I., **Hall, M. L.**, & Zvaigzne, M. T. (in preparation). Subjective frequency estimates for 432 signs in relation to age of ASL exposure.

REFEREED RESEARCH PRESENTATIONS

Hall, M. L., Mayberry, R. I., & Ferreira, V. F. (2012). Being human: experimental evidence that nominal gestures for humans shape constituent order preferences in pantomime. *To be presented at the 5th conference of the International Society for Gesture Studies, July 2012.*

Hall, M. L., Ferreira, V. S., & Mayberry, R. I. Producers SOV use. Perceivers like SVO. *Poster presented at the 52nd Annual Meeting of the Psychonomic Society, November 2011.*

Hall, M. L., Mayberry, R. I., & Ferreira, V. S. “WOMAN BOX LIFT”, but not “WOMAN BOY LIFT”: Why reversibility changes argument order in pantomime. *Poster presented at the 51st Annual Meeting of the Psychonomic Society, November 2010.*

Hall, M. L., Ferreira, V. S., & Mayberry, R. I. Syntactic priming in American Sign Language. *Talk presented at the 10th Theoretical Issues in Sign Language Research conference, October 2010.*

Hall, M. L., & Mayberry, R. I. Early language exposure effects the development of phonological representations in sign. *Poster presented at the 10th Theoretical Issues in Sign Language Research conference, October 2010.*

Hall, M. L., Mayberry, R. I., & Ferreira, V. S. Argument order in pantomime: Consistency fails when it's needed most. *Poster presented at the 16th Annual Conference on Architectures and Mechanisms for Language Processing, September 2010.*

Hall, M. L., Ferreira, V. S., & Mayberry, R. I. Syntactic Priming in American Sign Language. *Poster presented at the Sixth International Workshop on Language Production, September 2010.*

Hall, M.L., Mayberry, R. I., & Ferreira, V. S. Communication systems shape the natural order of events: Competing biases from grammar and pantomime. *Talk presented at the 4th conference of the International Society of Gesture Studies, July 2010.*

Hall, M. L., & Mayberry, R. I. Early language exposure effects the development of phonological representations in sign. *Poster presented at the 12th Laboratory Phonology conference, July 2010.*

Hall, M., & Ferreira, V. S. Is syntactic knowledge modality-independent? *Poster presented at the 21st CUNY Conference on Sentence Processing, March 2008.*

Hall, M., Supalla, T., Newport, E., & Bavelier, D. Short Term Memory in Signers: The Role of Memory Coding. *Poster presented at the 12th meeting of the Cognitive Neuroscience Society, April 2005.*

INVITED TALKS

Cognitive influences on language structure: Insights from experimentally-elicited pantomime. *Linguistics, UCLA, May 2012.*

Cognitive influences on language structure: Insights from experimentally-elicited pantomime. *Communication Sciences & Disorders, University of Iowa, April 2012.*

Bilingual picture naming studies constrain theories of lexical access. *Brain & Cognitive Sciences, University of Rochester, August 2011.*

Bilingual picture naming studies constrain theories of lexical access. *Psychology, University of Pittsburgh, August 2011.*

Keeping the hands in mind: What pantomime reveals about language structure. *Laboratory for Cognitive Neuroscience, San Diego State University, October 2010.*

Short-term memory in sign and speech: The source of the serial span discrepancy. *Deafness, Cognition, and Language Centre, University College London, September 2010.*

PROFESSIONAL SERVICE

Ad hoc reviewer: *Psychonomic Bulletin and Review, Memory & Cognition, Cognitive Science, Frontiers in Language Sciences, Journal of Deaf Studies and Deaf Education, European Journal of Cognitive Psychology*

Co-founder: Sign Language Reading Group of San Diego

Graduate Admissions Representative (2010-2011)

Career Transition Representative (2011-2012)

PROFESSIONAL MEMBERSHIPS

American Psychological Association

Sign Language Linguistics Society

International Society for Gesture Studies

TEACHING EXPERIENCE

Sign Language and Its Culture (TA: 2006-2009)

Developmental Psychology (TA: 2008, 2010)

Cognitive Neuroscience (TA: 2009)

TEACHING INTERESTS

Psycholinguistics

Language Acquisition

Working Memory, Short-Term Memory, and Language

Neurolinguistics

American Sign Language

Bilingualism

REFERENCES

Victor Ferreira

vferreira@ucsd.edu

UC San Diego, Psychology

Rachel Mayberry

rmayberry@ucsd.edu

UC San Diego, Linguistics

Daphne Bavelier

daphne@cvs.rochester.edu

University of Rochester, Brain & Cognitive Sciences

ABSTRACT OF THE DISSERTATION

Keeping the hands in mind: What elicited pantomime reveals about language structure

by

Matthew L. Hall

Doctor of Philosophy in Psychology

University of California, San Diego, 2012

Professor Victor Ferreira, Chair
Professor Rachel Mayberry, Co-Chair

This dissertation contains three studies that investigate whether attested patterns of constituent order distribution and change in the world's languages can be attributed, in part, to cognitive preferences for some constituent orders over others. To assess these preferences, seven experiments employed an *elicited pantomime* task. Communication in gesture without speech may reveal cognitive pressures that otherwise remain hidden.

Study 1 manipulates the whether the to-be-described events are semantically non-reversible (e.g. a woman lifting a box) or reversible (e.g. a woman lifting a boy). Across three experiments, participants reliably preferred SOV for describing non-reversible

events in pantomime, but avoided using SOV to describe reversible events. Although previous accounts have assumed similar behavior to be due to constraints on comprehension, the distribution of alternative orders (mainly SVO, OSV, & SOSV) suggests a greater role for constraints on production in explaining these patterns.

Whereas Study 1 focuses on explaining drifts away from SOV, Study 2 tests the hypothesis that languages drift toward SVO because natural languages are subject to additional constraints that pantomime may not engage, such as pressures on information structure and efficiency. These additional pressures might work against OSV and SOSV, leaving SVO as a preferred solution for reversible events. Experiment 1 (English speakers) created more language-like conditions by simulating the emergence of a lexicon and of shared lexical knowledge. Under these conditions, OSV and SOSV decreased, and SVO increased. Experiment 2 found similar patterns with native speakers of an SOV language (Turkish), suggesting that the increase in SVO among English speakers was not simply a result of influence from their native language.

Whereas Studies 1 and 2 focus on constituent order preferences in pantomime production, Study 3 focuses on pantomime comprehension. Speculation about constituent order preferences in comprehension has figured largely in linguistic and psycholinguistic theory, but little empirical evidence is currently available. Using a novel paradigm, Study 3 finds that constituent order preferences in comprehension are distinctly different from those in production, for both speakers of English (Experiment 1) and Turkish (Experiment 2). Results suggest that constraints on production may play a greater role than previously thought.

CHAPTER 1:

INTRODUCTION

Any attempt at defining what makes humanity different from other organisms will include language on the list of uniquely human abilities. Every human community has a language: it pervades our lives, structuring our social interactions, and even permitting some kinds of thoughts that might not be possible otherwise. I refer to this as the *ubiquity* of language. And yet even the most uninformed layman is well aware that languages vary enormously from one human community to the next. That is, despite their ubiquity, human languages nevertheless lack *uniformity*. This juxtaposition of ubiquity with the absence of uniformity has led generations of language researchers to search for core principles that, although they may be manifested differently across different languages, are nevertheless shared at a deeper level across all languages. I refer to this as the notion of *universality* in human language.

The notion of language universals means different things to different scholarly communities. To some, it means identifying a set of processes that can be used to generate all and only the grammatical sentences of any human language (e.g. Chomsky, 2011). To others, it means identifying observable patterns of linguistic structure that are common across all languages, or those that recur with a substantial degree of statistical regularity (Greenberg, 1963; Hawkins, 2007). To still others, the search for universals at those levels has proven fruitless, and should be supplanted by the observation that human language is the only known communication system that is fundamentally variable at all levels of structure (Evans & Levinson, 2009). Researchers with this view are likely to attribute apparent universals (or probabilistic tendencies) to humans' shared cognitive capabilities and social interaction (Chater & Christiansen, 2009). The present work is informed by all of these approaches, but does not neatly align with any single perspective.

My goal is to assess the extent to which we can account for patterns in the world's languages in terms of general cognitive processes, without reference to grammar-specific devices.

The work presented here focuses on patterns in constituent order, which refers to the order in which subjects (S), objects (O), and verbs (V) are mentioned in a transitive event (i.e., an event where an agent performs an action on a patient, such as a woman lifting a box). I chose this focus for three main reasons. First, constituent order is one of the dimensions where languages most obviously differ from one another; therefore, a principled account of this type of variation stands to make a substantial advance in our understanding. Second, constituent order has been very well characterized in the languages of the world, both in terms of its current distribution across the world's languages (also known as *synchronic variation*) and in terms of how the constituent order of a given language is likely to change across time (also known as *diachronic variation*). Third, despite the fact that theories of constituent order distribution and change have been central to linguistic theories of language universals, most claims in the linguistic literature are fundamentally retrospective in nature, rather than predictive. There has been little empirical research testing the psychological reality of various linguistic explanations. The research presented in this dissertation is intended to make a contribution toward our understanding of the cognitive processes that shape the nature of human language, both structurally and functionally. Below, I review several key observations about constituent order, and highlight the remaining questions that are addressed by the seven experiments in this dissertation.

Constituent order: observations and questions

The first observation that calls for explanation is that the synchronic distribution of constituent orders is highly uneven. There are six logical ways to order subjects, verbs, and objects: SOV, SVO, VSO, VOS, OVS, and OSV. A naïve assumption would be that all orders should be attested with approximately equal frequency across the world's languages, but this is not the case. Several large-scale surveys conducted in the latter half of the 20th century find that the vast majority (roughly 75%) of the world's languages use either SOV or SVO, with SOV being the single most common order (Dryer, 2008; Greenberg, 1963; Hawkins, 1983; Tomlin, 1986). VSO is in a distant third place, with the remaining three orders being even scarcer. (In fact, object-initial languages are so rare that they were claimed to be non-existent as recently as Vennemann, 1973.)

Thus, a key question concerns the extent to which these asymmetries are rooted in cognitive preferences. Is it the case that cognition is biased toward some orders and away from others? It could be that the synchronic distribution we observe today is simply the byproduct of historical accidents, such as language contact, military conquest, or geographical patterns that led some language families to proliferate more quickly or more broadly than others. If this were so, then we would not expect to see evidence of subtle cognitive biases toward particular constituent orders in the minds of people today. Alternatively, it could be that some constituent orders have proliferated while others have dwindled in part because some orders are more compatible with human cognition. If this were so, then we would expect to find evidence of consistent preferences for certain constituent orders across individuals, and across groups.

The second observation that calls for explanation is that the diachronic distribution of constituent order change is also highly uneven. Languages do change over

time, but some types of change are common while others are uncommon, or even unattested. In particular, languages commonly shift away from SOV, but do not shift toward it without some external influence such as language contact, conquest, and so forth. (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977; Vennemann 1974, 1975). Meanwhile, languages do shift toward SVO, which appears to be relatively more stable.

This presents an interesting paradox. If it is the case that SOV is so prevalent because it is especially compatible with human cognition, then why does it become less common over time? Likewise, why do languages gravitate toward SVO over time? If it is due to influences from cognition, why is SVO not more common in the first place? One possibility is that there are competing sets of cognitive biases. For example, Langus and Nespor (2010) suggested that there is a conceptual system (non-linguistic) that favors SOV and a computational system (linguistic) that favors SVO. They hypothesized that during the earliest stages of human language evolution, the non-linguistic system was still dominant, and so favored SOV. But over the course of centuries and millennia, the linguistic system has become dominant, and so now changes toward SVO are taking place. However, an alternative possibility is that there are some functions for which SOV is not well suited, and for which SVO is a useful substitute. Even if SVO is not as ideal as SOV in all cases, it may be that eventually the combined impact of these pressures is such that shifting to SVO affords greater utility overall. This latter account places less emphasis on a strict delineation between linguistic and non-linguistic processing, and focuses more on cognitive aspects of communication for which SOV may be problematic, such as describing events that are potentially confusable.

*Confusability, constituent order, case marking, and the
production/comprehension division of labor*

The primary purpose of human language is to convey a message from one mind to another, which necessarily entails both production and comprehension. These two processes are, in a sense, inverses of one another. The producer's challenge is to activate, select, and assemble the appropriate combination of words (or signs) to convey an intended meaning. The comprehender's challenge is to use the producer's output to reconstruct the intended message. It is conceivable, then, that what is easier for the producer might increase the difficulty of the comprehender's task. Likewise, the producer could endeavor to make the task as easy as possible for the comprehender, but that would come at a considerable cost on the producer's part. The grammars of natural languages represent various solutions to these challenges (Ferreira, 2008).

It is commonly assumed that both constituent order and case marking have evolved in response to the need to communicate about events where semantics alone are not sufficient to indicate who did what to whom (e.g. a boy, a girl, and a pushing event). In describing the function of case marking and constituent order, Vennemann (1975) states,

Much human communication is about interactions between human beings, and there is accordingly a large number of verbs which allow both human subjects and human objects; e.g. love, kiss, anger, invite, ask, call, scold, hit, kill, etc. It is essential that it is always clear what the relation of all noun phrases to their respective verbs is in any sentence, whether simple or complex; and the context does not always make this clear either. (p. 295)

Such accounts implicitly assume that these linguistic devices exist primarily for the benefit of comprehenders. Presumably, producers are aware of the message they intend

to convey, and so they are not at risk of confusing the agent and patient. Comprehenders, on the other hand, are assumed to need additional cues to accurately reconstruct the intended meaning, thus leading to the emergence of additional linguistic devices such as case marking and fixed word order.

Interestingly, there is a probabilistic relationship between case marking and constituent order: languages with case marking tend to have more flexible constituent order. However, in practice, many case-marked languages nevertheless have a dominant constituent order. This means that constituent order itself can serve as a cue to meaning, albeit a potentially weaker cue. This, in turn, means that producers can risk being somewhat sloppy in their use of case marking. Reducing (or omitting) case marking when it is not semantically necessary is one way to decrease producer effort without compromising communicative success. A natural prediction of this view, then, is that if a case system were to erode entirely, a language should simply deploy a fixed word order strategy, using whatever its most dominant order was. All else being equal, we should then expect to find that the distribution of constituent orders in languages without case systems mirrors the distribution of constituent orders in languages with case systems, but this is not the case. Instead, languages without case systems tend to be SVO, while SOV languages are more likely to have case systems (Greenberg, 1963; Sinnemäki, 2010). There are at least two plausible interpretations of this pattern, which are not mutually exclusive: SOV languages may be less likely to lose their case systems, and a loss of case may trigger an SOV language to become SVO. For present purposes, there is no need to distinguish between these possibilities; it suffices to note the implication that without the aid of case marking, SOV is dispreferred for *semantically reversible* events (e.g. a boy

pushing a girl). Vennemann (1975) makes this explicit: “As a substantive S-O marking system is eroded by phonological change, word order syntax must react to compensate for the ambiguities and perceptual complexities arising in a consistent verb-final language” (p. 293). Note that, in this framing, both SOV and OSV should be dispreferred for reversible events because of the potential difficulty they create for a comprehender.

This account seems fairly straightforward and entirely plausible, but it is important to note that it is fundamentally retrospective. Although it suggests a psychological mechanism, empirical evidence to that effect has not been brought forth. The present research aims to test both assumptions of this account: first, that SOV is dispreferred for describing reversible events, and second, that the source of this dispreference lies in constraints on comprehension.

Summary of research questions

The central questions addressed by the experiments below are these. First, is there evidence that the prevalence of SOV in the world’s language is related to a cognitive preference for SOV in the minds of individuals? Second, are there situations where participants disprefer SOV, and could these situations be responsible in part for the fact that languages shift away from SOV over time? Third, are there other cognitive preferences that give rise to SVO? Fourth, if so, to what extent are these preferences attributable to constraints on comprehension versus constraints on production?

Elicited pantomime as a window into constituent order preferences

Finding an appropriate method to test these questions is not a trivial task. Because only humans have language, animal models can be ruled out. And because humans have language, behavioral studies using language stimuli are likely to be subject

to massive confounds due to interference or transfer from a participant's native language. Computational simulations provide one way around the language confound, and they can be informative in demonstrating proof of concept. However, even when a candidate process is shown to be possible in theory, there is no guarantee that humans actually use it in practice. Thus, although computational studies offer a high degree of experimental control, it comes at the potential cost of external validity. Other useful insights come from observational studies of language creation, such as homesign, emerging sign languages, and creoles. Because people in these situations have no (or minimal) previous language input, extra-linguistic (e.g. cognitive, social) influences on language structure can be seen more clearly. In addition, there is at least the possibility of using existing theories to make predictions about how these new languages will evolve. However, any system as complex as human language will be subject to a multitude of influences, thus making it difficult to discern the impact of any one factor or process. The fact that these systems emerge spontaneously among human beings in real-world settings is a tremendous advantage in terms of external validity. However, that also means that they are not amenable to empirical manipulation, and therefore these insights come at the cost of experimental control.

An ideal method would combine the external validity of naturalistic cases of language emergence in human beings with the experimental control offered by empirical studies, while also finding a way to minimize interference from the participants' native language(s). These criteria are largely met by a technique that I call elicited pantomime, which was originally developed by Goldin-Meadow and her colleagues (Goldin-Meadow, So, Özyürek, & Mylander, 2008). In this paradigm, hearing adults with no knowledge of

any sign language are asked to describe transitive events in pantomime: that is, gesture without speech. Their utterances are then coded for the order in which they mention agents, patients, and actions. This means that they, like young homesigners or users of an emerging sign language, must construct representations of meaning by using their own bodies and the space around them. However, unlike those naturalistic settings, we can maintain experimental control over the situation, and systematically alter features of the stimuli or the communicative setting to determine what impact, if any, these manipulations have on the resulting constituent orders. The biggest drawback to this paradigm is that it relies on the behavior of human adults, who have already acquired a native language. However, there are at least three strategies for guarding against the possibility that the results we observe could be due to influence from the participants' native language, rather than to general cognitive principles. First, the original findings of Goldin-Meadow et al. (2008) demonstrated that although participants had the opportunity to use the structure of their native language to describe events, they did not appear to do so, even after having recently described the same events in speech. This demonstrates that pantomime is fairly robust to native language transfer. However, it is surely not impervious, which brings us to the second strategy: to the extent that participants' behavior in this task deviates from that of their native language, we can be confident that their behavior is not being driven by transfer from their native language. In other words, native language transfer is only a significant concern when the participants' behavior in the task is congruent with the structure of their native language. In those cases, we rely on the third strategy: to the extent that speakers of diverse languages nevertheless show similar behavior in the task, we can be reasonably confident that the behavior does not

originate from native language transfer, even when that pattern is congruent with one group's native language.

Study 1

In three experiments, Study 1 replicates previous work by showing that native speakers of an SVO language (English) do show a strong preference for using SOV order to describe non-reversible events in pantomime, but also extends previous work by showing that the same individuals avoid using SOV to describe reversible events in pantomime. Experiment 1 establishes this pattern of results when reversible and non-reversible events are mixed throughout the testing session. Experiment 2 finds that even when the testing session begins with an initial block of 45 non-reversible trials, participants still avoid SOV when the trials become reversible in the final block. Experiment 3 manipulates whether the reversible events are mixed throughout (as in Experiment 1), blocked at the end (as in Experiment 2), or blocked at the beginning. The question was whether participants who encountered reversible events earlier in the testing session would be less likely to use SOV to describe non-reversible events. This did not prove to be the case; however, the placement of reversible events did modulate the frequency of a behavior that resembles case marking in natural languages.

In each experiment, participants gravitated toward the same three alternatives to SOV for describing reversible events. One was to move the object to the end, yielding SVO. Another was to move the object to the beginning, yielding OSV. A third was to repeat a copy of the subject immediately before the verb, yielding SOSV. A common feature of all three of these alternatives is that V is immediately preceded by S, which allows participants to be in the agent (S) role at the time that they produce the action (V)

gesture. If they had used SOV to describe reversible events, they would likely have been in the role of patient (O) at the time they produced the action (V) gesture. This would have presented what I call a *role conflict* in production. In contrast to previous claims that SOV is dispreferred for reversible events because it creates potential confusability in comprehension, I argue that the potential for role conflict in production is what motivates participants to avoid using SOV to describe reversible events. I also consider the implications of the role conflict account for explaining patterns in spoken languages.

Whether role conflict, confusability, or an as-yet-unidentified hypothesis best accounts for the results of Study 1, the data clearly support the idea that the need to communicate about reversible events is part of what drives human languages away from SOV. This basic notion accords well with the prevailing views in functional linguistics. However, the current account suggests that constraints on production may play a greater role in driving these changes than has been previously considered. Therefore, even if role conflict is not the correct mechanism, the present work still opens new avenues of potential inquiry.

Study 2

Whereas Study 1 focuses on cognitive factors that drive languages away from SOV, Study 2 asks whether there are similarly motivated explanations for the rise in SVO. Although SVO was one of the prevalent alternatives to SOV in Study 1, it was not the only one. The experiments in Study 2 tested whether creating more language-like conditions (by simulating the emergence of a lexicon and shared knowledge) would implicitly engage additional cognitive pressures, such as a drive to be efficient and to mention subjects before objects. When combined with the pressure to avoid role conflict,

these three forces could conspire to promote the emergence of SVO. Experiment 1 found support for these ideas among English speakers, but that could easily have been due to influence from their native language. Therefore, Experiment 2 repeated the same study with native speakers of an SOV language (Turkish), and again found that SVO increased when participants had to describe reversible events under more language-like conditions.

These results suggest that SVO may begin to arise in the world's languages in part because it is an efficient way to describe both reversible and non-reversible events while avoiding the use of SOV to describe reversible events.

Study 3

Studies 1 and 2 focused exclusively on pantomime production, but it is possible that constituent order preferences in production and comprehension could be different. It is especially critical to test constituent order preferences in comprehension because previous research, both in linguistics and psychology, has made as-yet-untested assumptions about what would or would not be problematic in comprehension. We therefore developed a novel paradigm for collecting constituent order preferences in comprehension, and tested speakers of both English (Experiment 1) and Turkish (Experiment 2). In both cases, we found that constituent order preferences in comprehension were markedly different than those in production (Study 1 & Study 2). In particular, comprehenders favored SVO more than producers did for both reversible and non-reversible events. Comprehenders also did not penalize SOV for reversible events, as producers did.

These findings strengthen the claims that SOV-avoidance in pantomime production (and perhaps also in the world's languages) is driven to a considerable extent by constraints on production, rather than by comprehension.

Summary

Together, these three studies provide a new perspective on the ways that human language is shaped by human cognition. First, these studies provide empirical evidence that there are indeed cognitive preferences for certain constituent orders over others. Second, these studies validate the assumption that constituent order preferences are strongly influenced by whether events are semantically reversible. Third, these studies offer a cognitively motivated explanation for why SVO gradually supplants SOV over long time scales. Finally, the current studies suggest that constraints on production play a greater role in driving languages away from SOV than has been previously thought, and that comprehenders may not in fact have difficulty interpreting SOV descriptions of reversible events.

Although it is possible that some of the findings reported here are specific to the manual modality, a unified account of the parallels between our findings in gesture and patterns in spoken languages would be more parsimonious than independent explanations. Finally, even if some of the present findings are eventually shown to be unique to the manual modality, the present results can still inform our understanding of homesign and emerging sign languages, which are among our richest sources of insight into the relationship between language and cognition.

References

- Chater, N., & Christiansen, M. H. (2009). Language acquisition meets language evolution. *Cognitive Science*, 1-27.
- Chomsky, N. (2011). Language and other cognitive systems: What is special about language? *Language Learning and Development*, 7, 263-278.
- Dryer, M. S. (2008). Order of subject, object, and verb. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures online* (<http://wals.info>). Munich: Max Planck Digital Library.
- Evans, N. & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32, 429-492.
- Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation*, 49, 209-246.
- Gell-Mann, M. & Ruhlen, M. (2011). The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, 108, 17290-17295.
- Givón, T. (1979). *On understanding grammar*. New York: Academic Press.
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 105, 9163-9168.
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language: Report of a conference held at Dobbs's Ferry, NY, April 13-15, 1961* (pp. 73-113). Cambridge: MIT Press.
- Hawkins, J. (1983). *Word order universals*. New York: Academic Press.
- Hawkins, J. A. (2007). Processing typology and why psychologists need to know about it. *New Ideas in Psychology*, 25, 87-107.
- Langus, A., & Nespors, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, 291-318.
- Li, C. N. (1977). *Mechanisms of syntactic change*. Austin: University of Texas Press.
- Sinnemäki, K. (2010). Word order in zero-marking languages. *Studies in Language*, 34, 869-912.

- Tomlin, R. S. (1986). *Basic word order: functional principles*. London: Croom Helm.
- Vennemann, T. (1973). Explanation in syntax. In J. P. Kimball (Ed.), *Syntax and semantics, Vol 2* (pp. 1-50). New York: Seminar Press.
- Vennemann, T. (1974). Topics, subjects, and word order: from SXV to SVX via TVX. In J. M. Anderson & C. Jones (Eds.), *Historical linguistics, Vol 1* (pp. 339-376). Amsterdam: North Holland Publishing Company.
- Vennemann, T. (1975). An explanation of drift. In C. N. Li (Ed.), *Word order and word order change* (pp. 269-305). Austin, TX: University of Texas Press.

CHAPTER 2

COGNITIVE CONSTRAINTS ON CONSTITUENT ORDER: EVIDENCE FROM ELICITED PANTOMIME

Abstract

To what extent does human cognition influence the structure of human language? Recent experiments using elicited pantomime suggest that the prevalence of Subject-Object-Verb (SOV) order across the world's languages may arise in part because SOV order is most compatible with how we conceptually represent transitive events (Goldin-Meadow, So, Özyürek, & Mylander, 2008). However, this raises the question as to why non-SOV orders exist. Here we explore whether SOV might be suboptimal for describing certain kinds of events, thereby creating opportunities for other constituent orders to emerge. Specifically, we test whether SOV might be dispreferred for transitive events in which both the agent and patient are plausible agents: we call these "reversible" events. In three experiments, we use elicited pantomime to show that although participants have a strong tendency to produce SOV descriptions for non-reversible events (e.g., a woman lifting a box), they reliably avoid SOV descriptions for reversible events (e.g. a woman lifting a boy). These results are consistent with other recent findings using this paradigm (Gibson, Brink, Piantadosi, & Saxe, 2011; Meir, Lifshitz, Ilkbasaran, & Padden, 2010), but whereas those accounts emphasize the role of potential confusability in comprehension, we emphasize here a greater role of constraints on production.

In natural languages, words refer to entities (e.g., cats, girls), states (bliss), and actions (petting). But we use natural language to describe more than just disconnected entities, states, and actions; we also describe how these relate to one another. To do so, we string words together into sentences (e.g., “The girl blissfully pets the cat”), so as to convey which entities are doing which activities to which other entities, in which states, and so forth. Thus, an important property of natural languages is that we require devices that allow us to not only convey who, what, and whom, but also who *did* what *to* whom.

One way that different languages convey such information is by mentioning these major constituents in a specific order. For example, in an English active sentence like “The boy pushed the box,” the noun phrase before the verb is the subject of that verb and so denotes the entity that performed the action (sometimes called the *agent*), and the noun phrase after the verb is the object of the verb and so denotes the entity that had the action performed on it (sometimes called the *patient*). In fact, in most English sentences that have both a subject and an object, this particular ordering – subject-verb-object or *SVO* – is used, leading English to be termed an SVO language. Other languages use different orders of constituents than SVO. For example, in Turkish, sentences that have both a subject and object tend to order the subject first, followed by the object and then the verb, leading Turkish to be termed an SOV language. (It is important to note that this paper focuses mainly on the relative order of the semantic roles of agent, action, and patient, which can be dissociated from the syntactic roles of subject, verb, and object, as in passive sentences such as, “The box was pushed by the boy.” However, for ease of exposition, we adopt the nomenclature of S, V, and O.)

Curiously, the distribution of constituent orders across language families is far from even. SOV is dominant, with SVO a close second, and VSO a distant third. The remaining orders (OSV, VOS, OVS) are extremely rare (Dryer, 2008; Greenberg, 1966; Hawkins, 1983; Tomlin, 1986). Furthermore, languages are known to change from SOV toward SVO, but the reverse change is much rarer (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977). This suggests that in the past, SOV languages were once more dominant than they are now. Indeed, some research on the origins of language argues that human proto-language had SOV constituent order (Newmeyer, 2000). In short, SOV orders are not only predominant in the world's present languages; they may have been even more so in the past.

The predominance of SOV order is unlikely to be solely due to sociolinguistic factors, such as which language communities contacted, conquered, or emigrated from which. For instance, SOV has quickly emerged as the dominant order of a young sign language (Al-Sayyid Bedouin Sign Language) that has evolved without contact with other sign languages, and where the ambient spoken languages use SVO (Sandler, Meir, Padden, & Aronoff, 2005). This is especially noteworthy given the observation that similar convergence at the phonological level has not yet occurred in this language (Aronoff, Meir, Padden, & Sandler, 2008). SOV and the related orders SV and OV have also been reported in the first generation of signers of another young sign language evolving in Nicaragua, although OSV was also observed (Senghas, Coppola, Newport, & Supalla, 1997). Haviland (2011) also reports preliminary evidence of SV, OV, and SOV emerging in an even younger sign language currently evolving in Mexico. This latter case is particularly interesting in that the ambient spoken language (Tzotzil) is robustly

VOS. In all of these cases, the emergence of (S)OV cannot easily be attributed to any parent language, spoken or signed, and so is unlikely to be directly due to language contact.

Likewise, evidence suggests that persistent SOV preferences cannot be only due to learning biases during acquisition, or alignment among interlocutors. This is shown by cases where individuals create linguistic systems in the absence of input. Children are typically born into language-rich environments. However, children who are profoundly deaf from birth are sometimes not exposed to language until sometime after birth. While many of these children eventually receive linguistic input (either sign language or spoken language via hearing technology), others sometimes lack language input throughout childhood. In these latter cases, children typically create a manual communication system known as homesign (Goldin-Meadow & Feldman, 1977). Parental contributions to these systems are minimal (Goldin-Meadow & Mylander, 1983); thus, whatever structure they exhibit can be attributed to a large extent to the child's own innovation. Research has shown that the constituent order of homesign systems is remarkably consistent. Rather than being a grab bag of all possible orders, individuals' utterances are mainly composed of a limited subset of these orders: SV and OV, which are both consistent with SOV, and familiar from the descriptions of emerging sign languages reviewed above (for a review, see Goldin-Meadow, 2003). This tendency is robust cross-culturally, suggesting that individual homesigners all over the world have consistent preferences for constituent order (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, Özyürek, Sancar, & Mylander, 2008). These observations point toward a cognitive preference for SOV in the absence of language input.

More relevant to the experiments reported here is the finding that a preference for SOV orders is not unique to deaf children. Goldin-Meadow and colleagues (2008) asked hearing non-signers from diverse linguistic backgrounds to describe transitive and intransitive events in pantomime. They found that for transitive events, participants showed a strong tendency to produce SV, OV, and SOV descriptions. This tendency was equally robust among SVO speakers (English, Spanish, Chinese) and SOV speakers (Turkish). These findings have since been replicated by Langus and Nespors (2010).

In sum, evidence from typology, new sign languages, homesign systems, and pantomime suggests that the SOV order may have a ‘special’ status. The final observation – that adults spontaneously and systematically pantomime events with SOV orders – points to a possible reason for this special status: SOV may be more compatible with the way that human cognitive systems tend to organize event knowledge. That is, people may think about dynamic events in such a way that they are more efficiently linearized in an SOV order (Goldin-Meadow et al., 2008). If so, SOV ordering would naturally emerge in humans’ initial languages (both when language first emerged and today), which explains the appearance of SOV in homesign systems and in adults’ spontaneous pantomime behavior.

However, despite the above evidence favoring SOV, there are also systematic ways in which languages shift away from SOV. Chief among these are two phenomena related to language change: rapid convergence on SVO when a pidgin becomes a creole, and gradual but unidirectional drift away from SOV. We briefly review each situation in turn.

The term *pidgin* refers to contact language that results when speakers of different languages are forced to communicate. Pidgins generally retain only the bare bones of linguistic structure from the input languages (i.e. the various native languages represented in the community). Thus, these situations provide an opportunity to examine how linguistic structure re-emerges when the relatively disorganized pidgin is regularized into a *creole* via the language acquisition of young learners who are exposed to it from early childhood. A significant feature of creoles is that they are almost always SVO (Bakker, 2008; McWhorter, 2001). Indeed, SVO is found even when a creole's input languages were SOV (Kouwenberg, 1992).

Although this process is most easily observed in creoles, it is repeated on much slower time scales in natural languages. Languages are dynamic systems that change over centuries in response to various pressures. One systematic aspect of this diachronic change is that, as briefly noted above, languages commonly shift away from SOV, but do not shift toward it except through language contact (Venneman 1973). SVO-to-SOV change in Mandarin Chinese has been put forward as one counter-example (Li & Thompson, 1974), but even that claim has been strongly challenged (Sun & Givón, 1985). A recent analysis by Gell-Mann and Ruhlen (2011) showed that SOV languages can become SVO (or, less commonly, OVS and OSV). Meanwhile, SVO languages sometimes become VSO (and vice versa) or VOS (which can also become VSO), but not SOV.

If SOV has a special status because it corresponds to how human cognitive systems tend to order the elements of dynamic events, then what can explain the fact that languages systematically move away from SOV (whether through creolization or through

gradual diachronic change)? One clue comes from an asymmetry in the kinds of languages that use *case marking*. Case marking is another way that languages convey relational information (i.e. who did what to whom). Unlike content words, case markers - which are typically suffixes -- do not refer to entities, states, or actions; rather, they indicate the role that content words play in an utterance. For example, a case-marking language may attach one suffix to the subject of a sentence, and a different suffix to the object. The relevant asymmetry is that case marking is more common in languages whose preferred order is SOV than in languages whose preferred order is SVO. This pattern is nicely illustrated by data from the World Atlas of Language Structures (<http://wals.info>), which is a large database that catalogues various linguistic features of the world's languages. It includes data on the case-marking status of 101 SOV languages; of these, 83 (82%) use some form of case marking. Conversely, of the 71 SVO languages for which WALs includes case marking data, only 21 (29%) use some form of case marking. One plausible interpretation of these facts is that case marking performs a function that is more imperative for SOV languages than for SVO languages.

Another clue comes from a more specific pattern of case marking usage. In some languages, case marking is compulsory on all objects. However, many languages exhibit a pattern known as *differential object marking* (Bossong, 1991). In these languages, case marking is used selectively such that objects with more subject-like properties are more likely to be case marked. For example, subjects tend to be definite (e.g. "the woman", while objects tend to be indefinite ("a box"). More crucially for present purposes, subjects also tend to be human ("the woman") while objects tend to be non-human ("a box"). Aissen (2003) demonstrated that across languages, reliable hierarchies can be

inferred: if a language case-marks inanimate objects, it will also case-mark animate objects; likewise, if a language case-marks animate non-human objects, it will also case-mark human objects. The fact that object animacy is one of the specific triggers for differential object marking indicates that case marking may be especially important for *reversible* events: those where the patient is a plausible agent, such as “a woman lifted a boy”.

Further support for this interpretation comes from another way that languages change across time. Differential object marking systems are most familiar in languages whose case systems are in the process of decay. In languages in general, there is a pressure to reduce phonological content (and sometimes, morphology goes along with it; e.g., “iced tea” is referred to by many as “ice tea”). When the meaning of an utterance makes the relationships among referents obvious (e.g. women lift boxes; boxes don’t lift women), the pressure to reduce phonology may outweigh the incentive to retain case marking. However, when the meaning of an utterance leaves the relationships among referents ambiguous (e.g. women can lift boys, and boys can lift women), there may be greater benefits to retaining overt case marking. If these benefits outweigh the pressures to let the morphemes erode, a differential object marking system results. If, however, the case system erodes entirely, a language will tend to either re-evolve case marking or shift to another constituent order - usually SVO (Sinnemäki, 2010).

Taken together, these data suggest that, although SOV may be well-suited for describing non-reversible events, it might be less well-suited for describing reversible events. The fact that languages retain case longest for reversible events suggests that it is those events where case is most needed. Just as the gradual loss of case over time leads

many SOV languages to slowly shift to SVO, the sudden loss of case in pidgins may be what drives creoles to quickly shift to SVO: in the absence of case marking, SOV is not well-suited for communicating about reversible events.

Still missing, however, is an explanation for why SOV is not well suited for reversible events. The most intuitive possibility is based on the potential for confusability when two agent-like arguments are adjacent (Gibson et al., 2011; Vennemann, 1973). For example, in an utterance of the form “woman boy push,” the adjacency of “woman” and “boy” may make them more ambiguous with respect to which is the agent and which is the patient, so that an addressee who encounters this utterance might be at risk for misinterpreting the producer’s intended message. Such confusability-based accounts are inherently about the nature of the comprehension process, as of course, producers know the meanings they intend to convey (whether an event is reversible or non-reversible). One possible mechanism that could cause change away from SOV is that the relative confusability of S and O in unmarked SOV utterances leads to communicative inefficiency, which feeds back to producers, leading to linguistic innovations such changing constituent order or using case marking. Another possibility is that producers construct mental models of what would be difficult for or ambiguous to their addressees (Dell & Brown, 1991), and compensate for this anticipated confusability by changing constituent order or using case marking.

If indeed reversibility acts to drive languages away from SOV orders, then we may be able to observe this effect in the type of pantomime task used by Goldin-Meadow et al. (2008). In this paradigm, hearing adults who have no experience with sign language are shown simple events, such as a woman pushing a box, which they must then describe

in gesture without speaking. (We refer to this as pantomime to distinguish it from co-speech gesture, but participants do remain seated throughout, and generally use the same articulators as homesign and sign languages: hands, arms, torso, head, and the space in front of the torso). As mentioned above, the main finding of Goldin-Meadow et al. (2008), replicated by Langus and Nespors (2010), was that participants who spoke a variety of languages had a strong tendency to use SOV order in this task. However, all of the stimulus events in both previous studies were non-reversible; that is, the patient could not plausibly have been the agent. In the experiments reported here, we used a new set of items to replicate this non-reversible condition, which serves as a baseline measure of constituent order preferences in this task and with our materials. Crucially, we also included reversible events, where non-human patients were replaced with human patients, keeping the agent and action the same. If there is a cognitive pressure that biases against SOV for reversible events, we should observe SOV less on reversible trials than on non-reversible trials. In place of SOV, we might observe SVO, case marking, or possibly both. Furthermore, if the effect is driven by the goal of reducing confusability (Gibson et al. 2011), then we should observe fewer reversible events with constituent orders where the subject and object gestures both precede the verb.

Experiment 1

Method

Participants. We tested 24 undergraduate students at UC San Diego who reported being monolingual native English speakers. All participants gave consent to participate and be videotaped as part of the study.

Materials. Stimuli consisted of video clips of a human agent (man, woman, boy, or girl) performing a transitive action (kissing, lifting, petting, or pushing) on a patient that was either non-human (ball, box, bike, car, cat, dog) or human (man, woman, boy, or girl). Events where the patient was human were considered reversible; all others were considered non-reversible. (We did examine whether animate non-human objects [dog, cat] patterned differently from inanimate objects, but found very little evidence of such differences.) Among reversible events, the patient always differed from the agent in age or sex if not both.

The 61 stimulus events are provided in the Appendix. They were composed of 45 non-reversible events and 16 reversible events. Movies were filmed with a consumer-level digital camcorder in naturalistic settings and were displayed on a Macintosh laptop using PsyScope X software (Cohen et al., 1993; build 53, available from <http://psy.ck.sissa.it/>).

Design. The manipulation of reversibility was within-subjects; all participants saw all 61 items, which were arranged in a pseudo-random order, with the constraint that consecutive events differed in at least two of the three arguments involved, thus mitigating concerns about potential trial-to-trial carryover or contrast effects. To guard against possible order or learning effects, half of the participants encountered the items in reverse order. Reversible and non-reversible events were mixed throughout the block.

Participants performed three tasks: describing the events in English, describing the events in pantomime, and performing an unrelated filler task that involved describing shapes. Each task had its own dedicated block, and the filler task always separated the English and pantomime tasks, as in Goldin-Meadow et al. (2008). However, whereas

Goldin-Meadow et al. always began with English and ended with pantomime, we counterbalanced the order of these tasks. (Goldin-Meadow et al. (2008) aimed to make the strong demonstration that SOV order was used in pantomime even after those events have just been described in English; however, we counterbalanced task order because we expected that SVO might be more prevalent for reversible events, and we wanted to be able to demonstrate that effect even when participants had not yet described the events in English.)

Procedure. Participants were told that they would see short video clips, and they were instructed to describe the clip either in English, or in pantomime (i.e. gesture without speech). Practice trials consisted of four videos of events like those shown in the target trials: three non-reversible and one reversible. During the practice trials, participants were encouraged to produce gestures that conveyed all of the major information in each scene, but we very carefully avoided using any directives that mentioned subjects/agents, objects/patients, or verbs/actions. The aim was to ensure that on critical trials participants would use gestures for all three constituents; this was largely successful (97.2% contained agent gestures, 99.1% contained patient gestures, 99.9% contained action gestures, not counting excluded trials, or those classified as “ambiguous” or “simultaneous”).

Following the 4 practice trials, participants began the 61 critical trials. On each trial, a short video clip would play. Upon its completion, participants described the event in English to a sound recorder or in gesture to a digital camcorder. The experimenter remained in the room but did not interact with the participant, and was often hidden from

view. Participants were free to replay an event as often as desired. When the participant finished describing an event, they clicked a button to proceed to the next trial.

Coding. Subjects' gestures were transcribed by two trained coders, who identified the each gesture's referent as subject, verb, object, or other. Multiple *consecutive* gestures for the same referent were considered as belonging to the same constituent. For example, the string "TALL LONG-HAIR WHEELS PEDALS HANDLEBARS PUSH" would be parsed as SOV (woman, bicycle, push). However, if the string was non-consecutive, a repetition would be noted. For example, "TALL LONG-HAIR WHEELS PEDALS PUSH HANDLEBARS" would be coded as SOVO (woman bicycle push bicycle). In the absence of any principled way to determine phrase or utterance boundaries in pantomime, everything that the participant produced was counted as a single utterance unless there was a pause of more than 2 seconds or the string was interrupted by the participant replaying the video. In these cases, the most complete string containing an action gesture was coded. Gestures classified as "other" were ignored in determining constituent order; these mainly contained information about the environmental surroundings of the scene. If it was not possible to determine the referent of a gesture, the trial was coded as "ambiguous". If participants produced gestures that referred to more than one constituent at the same time, such that it was not possible to compute order of mention, it was considered "simultaneous" (for example, simultaneously producing two inverted "V" handshapes to represent two people standing near each other). Orders that accounted for less than 2% of both non-reversible and reversible events were considered "rare" and left unlabeled.

Coders were not blind to the experimental manipulations. However, they were blind to one another's ratings, and agreed on 1,104 of the 1,464 utterances (75.6%). In cases of disagreement, the first author coded the trial blind to the raters' responses. If this resulted in 2 of 3 coders agreeing, the trial was included. If all three coders disagreed, the trial was excluded from analysis (88 trials = 6%; labeled "excluded" in Figure 1).

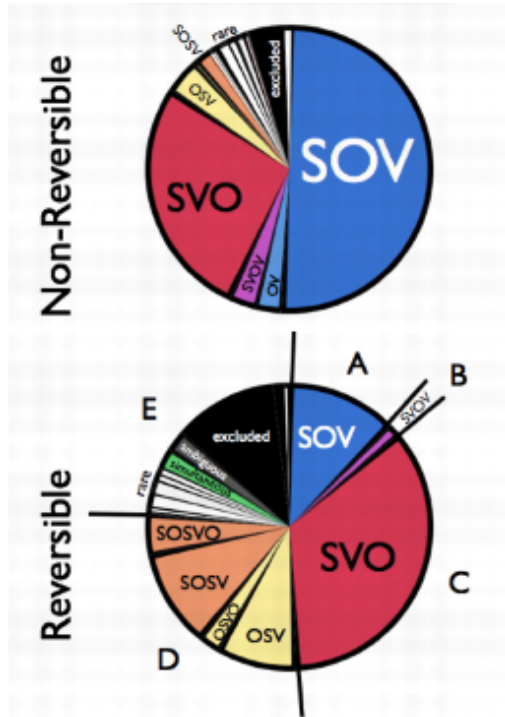


Figure 1. Constituent orders across all participants, for non-reversible (upper) and reversible (lower) trials. For reversible events, Region A shows orders that are inconsistent with both the confusability hypothesis and the role conflict hypothesis. Region B shows orders that are uniquely consistent with confusability. Region C shows orders that are consistent with both confusability and role conflict. Region D shows orders that are uniquely consistent with role conflict. Region E shows orders that are uninformative with respect to these hypotheses. Rare orders are defined as those that account for less than 2% of trials for both non-reversible and reversible events.

Results

Constituent order in English. As expected, English descriptions from all participants were over 98% SVO for both reversible and non-reversible events, and will not be discussed further.

Constituent order in pantomime.

Group results. Figure 1 presents the distribution of constituent orders in participants' pantomimes of non-reversible (upper panel) and reversible (lower panel) events. Among non-reversible events, SOV was the most common order, accounting for 50.9% of utterances; OV descriptions, which are consistent with SOV, constituted another 2.9% of utterances. SVO was roughly half as common, accounting for 27.5%. However, among reversible events, these proportions were reversed. SOV shrank to only 12.5% (Region A), while SVO became the most common order with 35.4% (Region C). OSV and OSVO grew from a combined 4.3% in non-reversibles to 11.5% in reversibles. SOSV and SOSVO also grew from a combined 1.9% in non-reversibles to 15.6% in reversibles. These four orders are shown together in Region D. The only other obvious change between non-reversible and reversible conditions was an increase in the percentage of excluded trials. This is due mainly to the fact that many participants produced highly complex utterances on reversible trials; the more components an utterance had, the less likely it was that at least 2 of 3 coders were in complete agreement. These trials are depicted in Region E, together with those classified as ambiguous, simultaneous, or rare.

To quantify the decrease in SOV, we followed Goldin-Meadow et al. (2008) by computing the proportion of each participant's utterances that were consistent with SOV

(i.e. SV, OV, & SOV). A within-subjects t-test showed that participants used significantly fewer SOV-consistent orders for reversible than for non-reversible events [$t(23) = 6.18, p < .001$].

Individual results. The group-level results above were echoed by the individual data. SOV became less common from non-reversible to reversible events in 20 participants, increased in only 1, and was absent or nearly so for the remaining 3. This decrease in SOV for reversible events was significant by a sign test ($p < .001$). We also identified the orders that each participant used most often for non-reversible events and for reversible events. For non-reversible events, SOV was dominant in 17 of 24 participants. For reversible events, only 3 participants remained SOV dominant.

“Case Marking” in Pantomime. In the current data, we observed gestures that functioned like case markers. They had no referential content, but indicated which role other gestures played in the utterance. They most commonly took the form of gestures where “1” indicated the agent and “2” indicated the patient. Some reversible utterances began with a “2” gesture, which we interpreted as signaling that the event to-be-described involved 2 people. In the analyses of constituent order discussed above, these gestures were treated as repetitions of the referent; that is, a string such as “1 MAN 2 WOMAN 1 LIFT 2” would be coded as SOSVO.

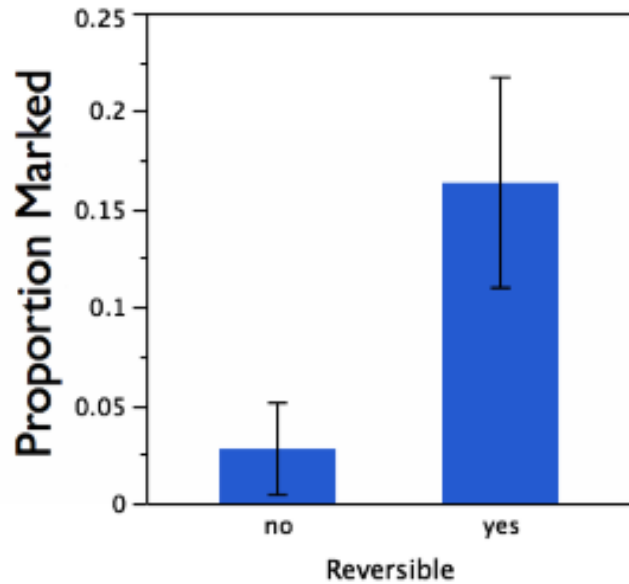


Figure 2. Proportion of trials containing gestures that resembled case markers, for non-reversible and reversible trials.

Group results. We coded the trials dichotomously for the presence versus absence of these devices (inter-rater reliability = 95.2% after 2 coders; 100% after 3) and found that gestures resembling case marking were significantly more common for reversible (16.4%) than non-reversible events (2.8%; $t(23) = 3.08$, $p < .001$).

Individual results. Gestures that functioned like case markers were used by 10 of the 24 participants, appearing anywhere from 1 to 39 times per individual. In each of these 10 participants, the proportion of reversible events with case marking gestures was greater than the proportion of non-reversible events with case marking gestures. Again, this pattern is significant by a sign test ($p < .01$). Thus, while not universal, this type of gesture is by no means an isolated phenomenon; we therefore examine it more systematically in Experiments 2 and 3.

Discussion

The main objective of the Experiment 1 was to test the hypothesis that there exists a cognitive pressure that biases against SOV for reversible events. Our results supported this hypothesis. We replicated the SOV bias found by Goldin-Meadow et al. (2008) and Langus and Nespors (2010) for non-reversible events, but we found a marked decrease in SOV for reversible events.

These results converge neatly with preliminary reports from similar studies by Meir, Lifshitz, Ilkbasaran, and Padden (2010), and Gibson, Brink, Piantadosi, and Saxe (2011). Thus, there is growing agreement about the existence of a cognitive bias against using SOV to describe reversible events. However, the exact nature of this bias is less clear. Both Meir et al. (2010) and Gibson et al. (2011) propose explanations based on the notion that reversible utterances contain potential confusability. That is, in a PERSON-PERSON-ACTION utterance, it may be more difficult to determine which person was the agent. Such utterances are not ambiguous for producers (who know what they intend to communicate), but if producers take their addressees' perspective into account, they might avoid SOV to minimize confusion or memory demands. Gibson et al. (2011) make this explicit, stating that participants should avoid "...putting two similar elements on the same side of the head [verb], where they may be more likely to be confused."

This hypothesis makes the specific prediction that utterances with S and O on the same side of V should be less common for reversible events than for non-reversible events. This prediction finds support insofar as SOV became less common for reversible events (Figure 1, Region A), and SVO became more common (Figure 1, Region C). However, it fails to account for our finding that a number of other orders with S and O on the same side of the verb also become more common for reversible events, including

OSV, OSVO, SOSV, and SOSVO (Figure 1, Region D). Similar findings were also reported by Meir et al. (2010), who observed a marked increase in OSV for reversible events. A confusability-based explanation would predict all of these orders to decrease, and yet they became more common. This invites us to consider an alternative hypothesis to account for the data.

To be more specific, note first that there is a consistent difference in how participants pantomime human versus non-human entities. For human entities, participants usually took on the role of the entity. For example, to gesture “man,” some male participants pointed to themselves, while other participants pantomimed the familiar flexed-biceps pose, and still others adopted a particular posture. These are all ways in which they described a human entity by embodying its role. In contrast, for non-human entities, participants typically did not take on the role of the entity. For example, to pantomime “box,” participants did not take on the role of the box; they usually just gestured out a box shape in space. Critically, when participants pantomimed actions, they invariably took on the role of the agent performing the action. To pantomime lifting, for example, participants acted as if they themselves were the agent doing the lifting. To describe a non-reversible event (e.g. a woman lifting a box) using SOV order, participants would generally adopt the role of the agent (long hair), then produce a gesture for the box without adopting any role. In this case, the participant doesn’t need to do anything special to re-inhabit the role of agent in time to produce the action gesture.

In contrast, using SOV for reversible events (e.g. a man lifting a woman) is likely to entail a *role conflict* between O and V. For example, if a participant described a reversible event using SOV order, she or he would first adopt the role of the agent

(flexing muscles), then the patient (long hair). The participant is now in the patient role but is ready to produce the action, which requires him or her to be in the agent role. If the participant were to produce an action gesture without first doing something to switch back into the agent role, it may “feel” to him or her as if it is the patient and not the agent that is carrying out the action. It is this that we refer to as role conflict.

We suggest that participants’ drive for avoiding role conflict in production is what best accounts for their behavior. Some participants avoided role conflict by repeating an agent gesture before the action (e.g. SOSV), others solved it by moving O to the front (e.g. OSV), and the most common solution was to move O to the end (SVO). Some also used gestures like “1” and “2” to overtly mark the difference between the two human entities.

Both hypotheses (minimizing confusability vs. avoiding role conflict) can account for the decrease in SOV and increase in SVO, but only the role conflict hypothesis can also explain increases in other orders such as those in Region D of Figure 1, all of which avoid having the patient immediately precede the action. In contrast, the only order that is uniquely consistent with the confusability hypothesis is SVOV (Figure 1, Region B), and it became slightly less frequent in reversible events than it had been in non-reversible events. This account is admittedly post hoc, and so it might simply describe the one sample of data we collected; our next goal is to replicate the effect. If the role conflict account is correct, we should see that participants in Experiment 2 also avoid SOV for reversible events, but that other orders like OSV, OSVO, SOSV, and SOSVO should not be similarly penalized.

In addition, Experiment 2 is designed to further restrict the possible influence of English on participants' responses. We observed that participants who began by describing the events in English used SVO 17% more in pantomime than participants who began with the pantomime task. Importantly, this tendency did not interact with reversibility [$F(1,22) = .33, p = .57$], so that our main observations stand unchanged. It does, however, suggest that asking participants to describe the events in English can influence their behavior on a subsequent pantomime task. Insofar as one of the key results of the reversibility manipulation was an increase in SVO, we want to guard against the possibility that it arose because participants had previously described the events in English. Therefore, all participants in Experiment 2 performed the pantomime task first, and then described the events in English.

A further concern is that the magnitude of the SOV bias in our non-reversible events was somewhat smaller than that reported by Goldin-Meadow et al. (2008); on those non-reversible trials, our participants used more SVO than theirs did. Because we used different events and more stimuli, it is possible that the increased SVO we observed is attributable to some property of our items. Alternatively, it is possible that the pressure to avoid SOV for reversible events was so strong that it spilled over and influenced the way that participants described non-reversible events, biasing them toward SVO. To distinguish these possibilities, Experiment 2 presented all of the non-reversible events first, with the reversible events all appearing at the end of the session. This manipulation makes it impossible for the reversible events to influence how participants described the non-reversible events. If participants' descriptions of reversible events attenuated SOV and increased SVO for the non-reversible events in Experiment 1, we should observe

relatively more SOV and relatively less SVO for the non-reversible events in Experiment 2. One consequence of this manipulation could be an attenuation of the SOV-avoidance for reversibles that we observed in Experiment 1; after 45 non-reversible trials, participants might become accustomed to their particular way of describing events. However, any effects of reversibility that we do see could thus be seen as even more robust in this design.

Experiment 2

Method

Participants. We tested 12 undergraduate students at UC San Diego who reported being monolingual native English speakers. All participants gave consent to participate and be videotaped as part of the study.

Materials. The stimuli were the same as those in Experiment 1.

Design. All participants described events in pantomime first and then described the events in English. There was no filler task.

To test whether our non-reversible items were capable of eliciting a stronger SOV bias, we designed a new pseudo-random order in which the 45 non-reversible events appeared first, followed by the 16 reversible events. As before, the only constraint was that consecutive trials always differed in at least two elements. This design allows a more straightforward comparison between our stimuli and those of Goldin-Meadow et al. (2008), which were all non-reversible.

Procedure. The only additional difference from Experiment 1 was changing the four practice trials to be all non-reversible events that were chosen with permission from the items used by Goldin-Meadow et al. (2008).

Coding. We used the same coding procedures as in Experiment 1.

Results

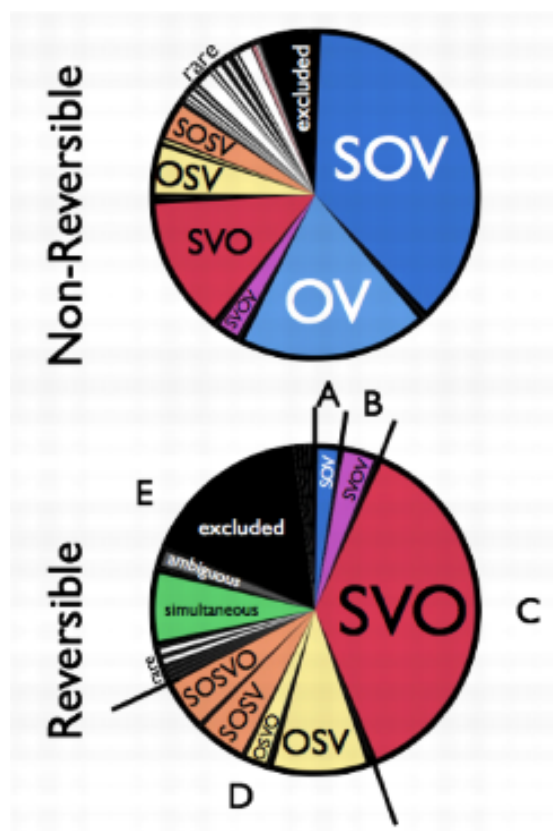


Figure 3: Constituent orders across all participants, for non-reversible (upper) and reversible (lower) trials. For reversible events, Regions A-E are defined as in Figure 1.

English. As before, the English results were 96.2% SVO and do not merit further discussion.

Constituent Order in Pantomime.

Group results. Figure 3 presents the distribution of constituent orders in participants' pantomimes of non-reversible (upper panel) and reversible (lower panel) events. We again found an SOV bias for non-reversible events, which – together with OV

– accounted for 57.8% of trials. SVO was again the second most common order, but accounted for only 14.3% of trials: roughly one fourth the size of the SOV bias. This SOV bias was numerically greater in this experiment (blocked) than in the previous experiment (mixed). We again found that participants strongly avoided SOV for reversible events, where they accounted for only 2.6% of reversible trials (Region A). Likewise, SVO increased from 14.3% in non-reversibles to 38% in reversibles (Region C). Orders that would be predicted to decrease by confusability-based accounts (Meir et al., 2010, Gibson et al., 2011) did not do so. Instead, OSV and OSVO together increased from 5.9% to 13%, with SOSV and SOSVO also increasing together from 4.1% to 10.4% (Region D).

As in Experiment 1, we quantified the decrease in SOV by using a paired t-test to compare the proportion of each participant's utterances that were consistent with SOV (i.e. SV, OV, & SOV) for non-reversible and reversible events. This test was significant [$t(11) = 6.18, p < .001$].

Individual results. Again, the individual data yield the same picture as the group data. SOV became less frequent from non-reversible to reversible events in every participant. This decrease is significant by a sign test ($p < .001$). While 8 of 12 participants were SOV-dominant for non-reversible events (i.e. SOV was their most common constituent order), 0 remained SOV-dominant for reversible events.

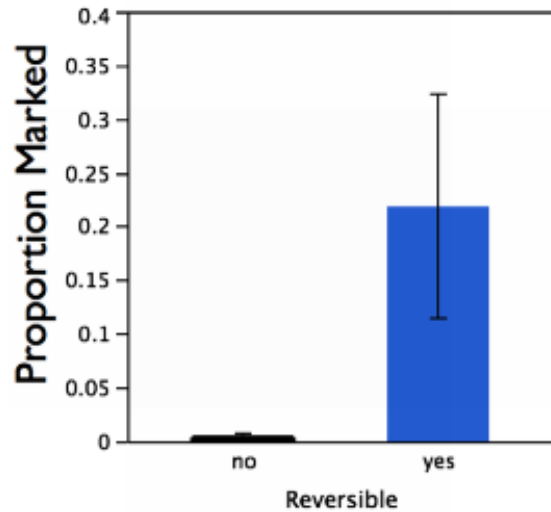


Figure 4. Proportion of trials containing gestures that resembled case markers, for non-reversible and reversible trials.

Case Marking.

Group results. As in Experiment 1, we observed gestures that functioned like case markers. We coded trials dichotomously for the presence versus absence of devices that carried no referential content except to indicate another gesture's semantic role (inter-rater reliability = 97.7% after 2 coders; 100% after 3). Again, we found that gestures resembling case marking were more common for reversible (22.9%) than non-reversible events (0.3%). Despite a smaller number of observations and high variability, this effect was significant by a 1-tailed t-test [$t(11) = 2.05$, 1-tailed $p = .033$].

Individual results. Gestures that functioned like case markers were found in 6 of 12 participants, appearing anywhere from 1 to 16 times per participant. They were more common in reversible events for 5 participants, and equal (0%) in 6; however, 1 participant produced 2 instances of case marking on non-reversible trials and 0 on

reversible trials. With this small number of observations, those two trials rendered the sign test non-significant.

Discussion

Experiment 2 had three main goals: (1) To replicate Experiment 1 while limiting the influence of English; (2) to test the role conflict account of SOV avoidance; and (3) to test whether the presence of reversible events in Experiment 1 influenced the ways in which participants described the non-reversible events.

The first goal was clearly achieved. Participants in both experiments were SOV-dominant for non-reversible events, but avoided SOV for reversible events, even after having seen a string of 45 consecutive non-reversible events. Gestures that functioned like case markers were present in both experiments; this phenomenon was numerically stronger in Experiment 2, but the smaller sample size weakened our power to detect a significant effect.

The second goal was to compare two accounts of the above effects: one based on confusability and one based on role conflict. Confusability-based accounts predict that constituent orders with S and O on the same side of V should become proportionally less common for reversible events. Our role conflict account posits that only those orders where a human object is followed by a verb gesture should be penalized. Both accounts correctly predict a decrease in SOV and an increase in SVO. Therefore, the critical test is to examine orders with S and O gestures on the same side of V, but where O is not followed by V (e.g. OSV, OSVO, SOSV, SOSVO, etc.). Here, the results once again favored the role conflict account: participants did not avoid those orders; if anything, they became more common, as shown in Figure 2, Region D. The only order that is uniquely

predicted to increase by confusability is SVOV, whose increase from 2.6% to 3.7% is modest (Figure 2, Region B). These results strongly resemble those of Experiment 1, suggesting that this pattern does not merely reflect random differences between individuals. Still, additional replication would further strengthen the case.

Experiment 2's third main goal was to explore an unexpected aspect of Experiment 1: namely, that the SOV bias was weaker than that reported by Goldin-Meadow et al. (2008), and SVO figured much more prominently in our non-reversible events than in theirs. It may have been that the presence of reversible events (which were mixed throughout the block in Experiment 1) influenced the way participants described the non-reversible events. If so, then moving all the reversible events to the end of the block should have resulted in more SOV and less SVO for the non-reversible events in Experiment 2 than in Experiment 1. This was in fact the case: among non-reversible events, the SOV bias was stronger in Experiment 2 (blocked, 59.6%) than in Experiment 1 (mixed, 54.4%). Similarly, SVO was less common in Experiment 2 (blocked, 14.3%) than in Experiment 1 (mixed, 27.5%). If these differences can be replicated, they could constitute the first (to our knowledge) empirical demonstration that the need to communicate about reversible events causes constituent order to shift away from SOV and toward SVO, even for non-reversible events.

However, there are several reasons to be cautious before reaching such a conclusion. First, these differences (especially the comparison of the SOV bias) are not large. Another limitation is that they are based on between-experiment comparisons, which admit the possibility of uncontrolled differences between studies. These differences would be more convincing if obtained via random assignment in a single

experiment. A third limitation is that we cannot distinguish effects of blocking versus mixing the reversible events from those of moving the English block to the end of the session: the differences we observed could be attributed to one, the other, or both changes. Experiment 3 was designed to distinguish these possibilities by manipulating where the reversible events appeared during the session, and randomly assigning participants to one of three conditions: reversibles-last (like Experiment 2), reversibles-mixed (like Experiment 1), or reversibles-first, where the presence reversible events might exert an even stronger influence on how participants describe non-reversible events.

Experiment 3

To determine whether the presence of reversible events influences the way participants describe reversible events, Experiment 3 manipulated (between subjects) where the reversible events were positioned over the course of the experiment. Based on the results of Experiment 1 and 2, we predicted that the prevalence of SOV for non-reversible events would be greatest when reversible events had not yet occurred (reversibles-last), less when reversible events are mixed throughout the block (reversibles-mixed), and least when the block begins with all of the reversible events (reversibles-first).

Method

Participants. We tested 36 undergraduate students at UC San Diego who reported being monolingual native English speakers. All participants gave consent to participate and be videotaped as part of the study. They were randomly assigned to one of the three conditions, yielding 12 participants in each group.

Materials. Stimuli were the same as those in Experiments 1 and 2.

Design. The only changes to the design were the between-subject manipulation of the position of the reversibles and removing English from the experiment altogether. The reversibles-last group ($n = 12$) saw the same order as the participants in Experiment 2, which began with the 45 non-reversible events and finished with the 16 reversible events. The reversibles-mixed group ($n = 12$) saw one of the orders that we used in Experiment 1, with reversible events mixed throughout the testing session. The reversibles-first group ($n = 12$) saw the same order as the reversibles-last group, except that the 16 reversible events came at the beginning of the experiment instead of the end.

Procedure. The only change to the procedure was in the practice trials. We wanted all participants to experience the same practice trials, but because some groups would be starting with reversibles and others with non-reversibles, we couldn't use any transitive events during practice. Instead, the practice consisted of 6 intransitive events (e.g. a woman waving), half of which also included a locative element (a girl walking toward a car). These materials were borrowed with permission from the original materials used by Goldin-Meadow et al. (2008). Locatives were included as a third element of some videos to encourage participants on critical trials to use gestures for all three arguments, which indeed occurred (82.3% contained agent gestures, 93.5% contained patient gestures, 99.8% contained action gestures).

Coding. We used the same coding procedures as in Experiments 1 & 2.

Results

Constituent Order

Between-subjects results. The presence of reversible events did not systematically influence the frequency with which participants used SOV to describe non-reversible

events. We followed Goldin-Meadow et al. (2008) in classifying each utterance as either consistent with SOV (SOV, OV, and SV) or not (anything else), and submitted the resulting proportions to a 2 x 3 mixed analysis of variance (ANOVA) with reversibility (reversible, non-reversible) as a within-subjects factor and group (reversibles-last, reversibles-mixed, reversibles-first) as a between-subjects factor. There was no main effect of group [$F(2,33) = .32, p = .73$] and no reversibility x group interaction [$F(2,33) = .69, p = .51$]. We used a planned contrast to test for linear trend (reversibles-last > reversibles-mixed > reversibles-first), but this too was far from significant [$F(1, 33) = .59, p = .45$].

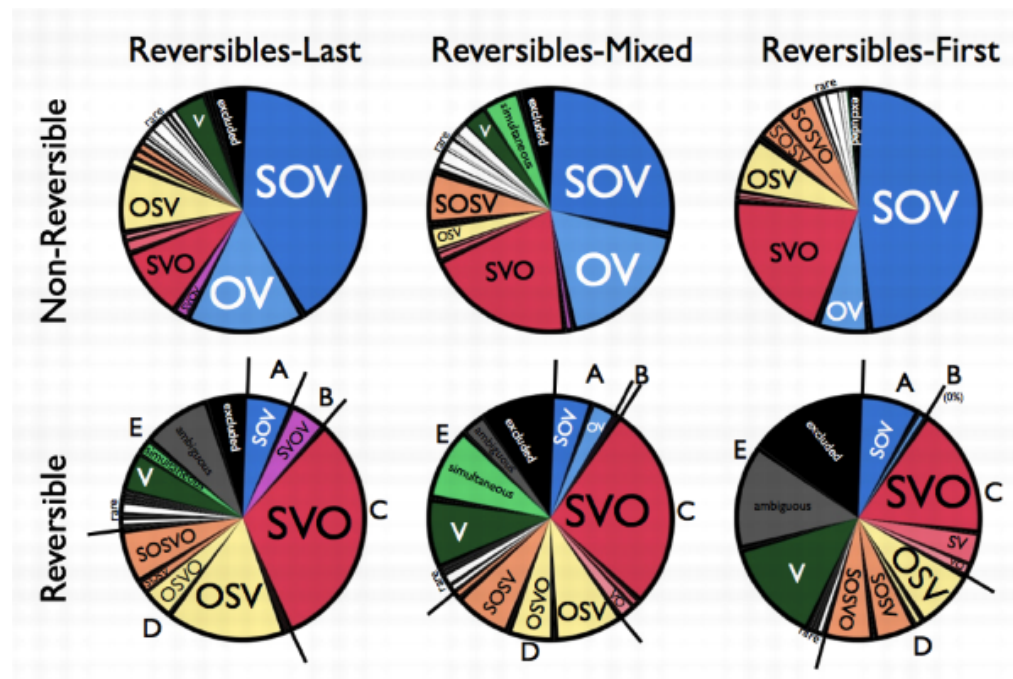


Figure 5. Constituent orders across all participants in each condition, for non-reversible (upper) and reversible (lower) trials. For reversible events, Regions A-E are defined as in Figure 1.

Within-subjects results. The effect of reversibility is evident in the contrast between the top panels of Figure 5 and their corresponding panels beneath. In each case, participants used SOV much less for reversible events than for non-reversible events

(Region A). This decrease was significant in the ANOVA above [$F(1,33) = 78.09, p < .001$] and by a sign test across individuals (34/36 decreased, $p < .001$).

As in the previous two experiments, orders with both S and O gestures on the same side of V did not become less common for reversible events (Figure 5, Region D). Collapsing across groups, OSV and OSVO increased from 7.4% in non-reversibles to 15.1% in reversibles. SOSV and SOSVO increased from 6.0% to 10.1%. The most common order for reversible events was SVO, which increased from 16.9% to 25.9% (Region C), collapsing across groups. SVOV, the only order uniquely consistent with the confusability hypothesis, increased from 1% to 1.7% (Region B), collapsing across groups.

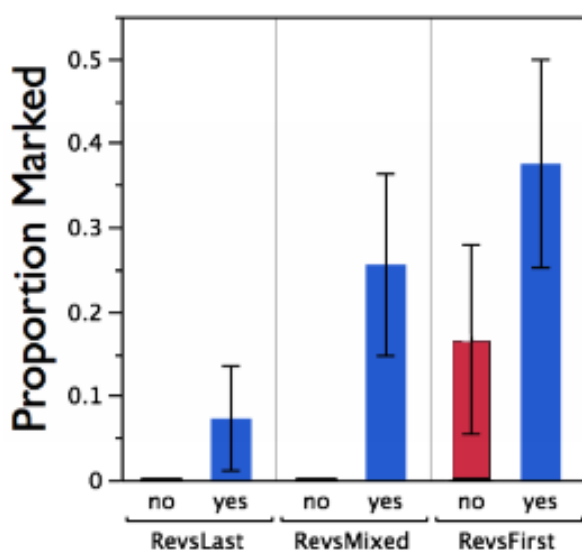


Figure 6. Proportion of trials containing gestures that resembled case markers, for non-reversible and reversible trials, by condition.

Individual Results. In the absence

of a main effect of group, we discuss the data from all 36 individuals together. SOV became less frequent from non-reversible to reversible events in 34 participants, which is significant by a sign test ($p < .001$). Of the 36 participants, 23 were SOV-dominant for

non-reversible events (i.e. SOV was their most common constituent order), with only 3 remaining SOV-dominant for reversible events.

Case Marking

Group results. As in Experiments 1 and 2, we observed gestures that functioned like case markers, and found that they were more common for reversible events than for non-reversible events. We computed the proportion of trials on which each participant used gestures that functioned like case markers. We submitted these to the same 2 x 3 ANOVA described above and once again found that case marking was more common for reversible events than for non-reversible events [$F(1,33) = 10.56, p < .01$]. Interestingly, there was also a trend toward a main effect of group [$F(2,33) = 2.82, p = .07$]. As shown in Figure 6, case marking was relatively rare in the reversibles-last group, and restricted to reversible events only. In the reversibles-mixed group, it became more common but was still restricted to reversible events. Case marking was most common in the reversibles-first group, where it began to be used even for non-reversible trials. A post-hoc test for linear trend was significant [$F(1,33) = 5.54, p < .03$].

Individual results. Gestures that functioned as case marking were found in 15 of 36 participants, appearing anywhere from 1 to 61 times per participant. They were proportionally more common in reversible events for 14 participants, less common for reversibles in 1, and tied (0% or 100%) in the remaining 21. This pattern is significant by a sign test ($p < .001$).

Discussion

The main purpose of Experiment 3 was to test whether the presence of reversible events would influence the way participants described non-reversible events. Based on a

comparison of Experiments 1 and 2, we expected to see such effects on constituent order, but did not find them. However, we did find such evidence in the frequency of case marking gestures. Placing reversible events earlier tended to lead to more case marking on both reversible and non-reversible events. These data are consistent with observations from the linguistics literature that in many languages, reversible events are more likely to be case-marked than non-reversible events. As noted in the introduction, these systems, known as Differential Object Marking systems, are most widely observed in languages whose case systems have begun to decay, such that reversible events are the ones that retain case marking the longest. However, systems that begin as differential object marking will sometimes expand to become “consistent object marking” systems, where all types of objects receive overt case marking (Aissen, 2003; Bosson, 1991). The present data are, to our knowledge, the first empirical demonstration that even speakers of languages without case systems feel a pressure to use additional devices to describe reversible events, and that the need to communicate about reversible events can drive the spread of case marking to non-reversible events. These data highlight the role of cognition in driving these changes. It appears that the same pressures that drive these long-term changes in a language over time are also at work in the minds of our participants, whose native language doesn’t use a case system.

Experiment 3 also replicated the main findings of Experiments 1 and 2. First, participants in all conditions avoided using SOV order for reversible events. Second, their alternative constituent orders are more consistent with our role conflict account than with previous confusability-based accounts. Whereas confusability-based accounts predict that orders like OSV, OSVO, SOSV, and SOSVO should also be avoided for

reversible events, our production-based account predicts that they should actually be acceptable solutions to the problem of having a human object gesture followed by a verb gesture. The data support the latter view.

General Discussion

In three experiments, we asked native English speakers to describe simple transitive events in gesture without speaking. When these events involved a human agent and a nonhuman patient (non-reversible), the most common constituent order was SOV. This replicates previous reports by Goldin-Meadow et al. (2008) and Langus and Nespor (2010), who take this phenomenon as evidence of a cognitive pressure that favors SOV as a default order for event representation. This is consistent with linguistic evidence that SOV was more common longer ago in linguistic history, perhaps including proto-language; our data help to establish that this was not merely an accident of history, because the same preference for SOV is clearly active in the minds of naïve participants today. Similarly, we reviewed evidence documenting the emergence of (S)OV in deaf homesigners and in the emerging sign languages of several deaf communities. However, the SOV preference is clearly not merely an artifact of deafness, because we see it at work in our participants, who all have normal hearing and no prior experience with sign language. This convergence across many different domains lends strong credence to the view that the early and widespread emergence of SOV is rooted in human cognition.

However, the central finding of these experiments is that when the events involved a human agent and a human patient (i.e., were reversible), participants strongly avoided SOV in favor of other constituent orders. Specifically, descriptions of reversible events appeared to be targeted toward breaking up O-V adjacency when the object was

human. Similar results have been reported by Meir et al. (2010) and Gibson et al. (2011), who attribute this SOV-avoidance to the potential confusability inherent in PERSON-PERSON-ACTION strings. In so doing, they place the locus of the drive to avoid SOV in the addressee, or else in the producer's mental model of what an addressee would find to be difficult.

There are three main reasons to reconsider this view. First, it rests on the assumption that comprehenders would disprefer SOV for reversible events. However, we recently asked speakers of English (SVO) and Turkish (SOV) to watch gestured events and to provide a subjective "goodness" rating. The stimuli varied in constituent order and reversibility; nonetheless, neither English nor Turkish comprehenders rated SOV as significantly worse in reversible than in non-reversible events (Hall, Ferreira, & Mayberry, 2011). The second reason is that the present task did not involve communicating to an addressee; thus, any effects of avoiding confusability cannot be due to feedback from the gradual accumulation of unsuccessful attempts at communication. Instead, the effects would need to be attributed to the participants' mental models of addressees (Dell & Brown, 1991). Insofar as those mental models tell producers to avoid SOV, they conflict with what we observed in actual comprehenders, and thus it is still appropriate to localize this effect to production. Furthermore, producers have been shown to be insensitive to ambiguities they create for the addressees when speaking their native language (Ferreira, 2008). Thus, it would be very surprising to see such large effects in this case.

The third and final reason to reconsider confusability-based accounts is that they miss a generalization in the data. Confusability-based accounts predict that participants

should avoid all constituent orders where S and O are on the same side of V. However, participants only avoided some of those orders: the ones where O and V are in sequence. Orders such as OSV, OSVO, SOSV, and SOSVO were either unchanged or became more frequent for reversible events. We have argued that this is because participants' behavior was mainly driven not by concerns about what would be ambiguous to an addressee, but by a constraint against role conflict in production. This conflict occurs because participants commonly take on the roles of both human agents and patients, but action gestures require the participant to be in the agent role. This, in turn, is because when participants denote the actions of kissing, lifting, petting, and pushing, they invariably do so by using their own bodies as if they were the agent.

One way to avoid role conflict is to repeat a subject gesture just before the verb, which yields orders like SOSV and SOSVO. Although solutions of this nature were fairly common in our data, the repetition of constituents makes them potentially inefficient, and so it is not surprising that we don't find much evidence of this strategy in natural language.

Another way to avoid role conflict is to produce the object first, and then the subject-verb unit. This yields OSV, which was also a persistent strategy adopted by our participants. In spoken languages, there is a strong crosslinguistic tendency for the subject to precede the object; OSV clearly violates that pattern. It remains unclear exactly how common OSV is (as a basic or alternative order) in the sign languages of the world, or even if the distribution of constituent orders in sign languages is the same as in spoken languages. To date there has been no comprehensive typological survey of

constituent order in the world's sign languages. This represents a significant knowledge gap that would be valuable to fill.

A third way to avoid role conflict is to move the object gesture to the end of the utterance, yielding SVO. This solves the problem without the need for repetition, and allows the subject to be produced before the object. Thus, it is not surprising that this was the most common solution in our current data. This may help to explain why languages change from SOV to SVO, but not vice versa, which also helps to explain why SVO is so prevalent in the world's languages, despite all of the forgoing evidence favoring SOV. One caveat here is that, unlike the previous two solutions, SVO is also consistent with English and could therefore reflect influence from the participants' native language, rather than underlying cognitive processes. To address this concern, we conducted a similar study with native speakers of Turkish (SOV), and found that they too avoid SOV, and that SVO emerges as one of the solutions (Hall, Ferreira, & Mayberry, 2011). However, SVO only emerged under conditions that more closely approximated natural language situations, which is consistent with the account we have offered here.

The last systematic pattern that we observed in the data was that some participants began to use gestures that functioned like case markers. That is, they did not refer to any real-world referent; instead, they signaled the role that another gesture played in the event. These gestures were more common for reversible events than for non-reversible events, which strongly resembles differential object marking in the languages of the world. One difference is that in our data, these gestures were not especially more common with SOV sequences, as they are in spoken languages. However, one significant parallel is that case marking gestures were especially uncommon in SVO sequences, just as case marking is

rare in SVO languages. This constitutes evidence that case marking and SVO order are independent solutions to the same problem (semantic reversibility), and once you have one solution, the other isn't necessary.

Finally, the above analyses implicitly assumed that all instances of SOV involved role conflict, because of the sequence O-V with a human O. However, it should be noted that some reversible trials that were coded as SOV did in fact satisfy the role conflict constraint, despite the fact that O and V were produced in sequence. One way that participants could avoid role conflict even in SOV descriptions was to avoid taking on the role of the human patient (e.g. referring to the patient by tracing the outline of a body in space, rather than referring to their own body). This accounted for 55 of the 93 reversible trials that were coded as SOV above (59%). But even when participants did take on the role of patient with their own bodies, they could use SOV and still avoid role conflict provided that they established the agent in one location, the patient in another, and then shifted back into the agent's space before producing the action gesture, even without producing any overt repetition of an agent gesture. This strategy was observed on 8/93 reversible SOV trials (9%). Interestingly, Fischer (1975) observed this type of SOV structure in old American Sign Language (ASL), but noted that this strategy is somewhat inefficient because it requires an extra transitional movement between O and V. She then used written transcriptions of signed narratives to make that case that in fact old ASL had SOV structure, but argued that modern ASL has transitioned to using OSV and SVO for these types of reversible events precisely because OSV and SVO avoid this extra transitional movement. The fact that we see naïve participants behaving similarly suggests that the factors involved are not specific to being deaf or having experience with

sign language, but are grounded in aspects of cognition that are broadly shared, and potentially universal.

Production versus Comprehension

The idea that the reversible events require additional linguistic devices is not new. It has long been recognized that a key function of morphosyntax is to guide the interpretation of utterances that cannot be interpreted based on semantics alone -- that is, reversible events. Traditional explanations generally assume that what makes SOV problematic is the potential for confusability when two human nominals both precede the verb. In so doing, such views implicitly place the locus of the constraint on addressees. According to these accounts, producers are not susceptible to confusability in this way, given that they already know who did what to whom. However, if they were to produce a sequence like MAN WOMAN LIFT, they would run the risk of their addressee sometimes picking the wrong interpretation. This communicative inefficiency would then feed back to producers over time, eventually leading to linguistic change.

Our results suggest that this traditional view is incomplete, if not simply incorrect. The explanation we have offered is operative during production. The problem is not that the sequence MAN WOMAN LIFT is produced but ambiguous to an addressee; rather, the problem is that the producer detects a potential mismatch between the intended agent of the action (man) and the most recently adopted role (woman), and so decides against producing the action gesture without adding any additional cues. This does not mean that confusability and comprehension play no role in how these processes unfold in natural languages, only that constraints on production may be much more responsible than previously thought.

A key question concerns whether this constraint is specific to the manual modality, or if it also applies to spoken language. For now, the answer is unclear. On the one hand, the account offered here depends on the unique ways in which physical bodily actions are literally embodied in gestural communication. Sequences of spoken words may not encounter quite the same conflict of roles when a word denoting a human object is followed immediately by an action word. On the other hand, if the constraint does not apply, then we need separate explanations to account for similar phenomena, which seems unparsimonious. Therefore it is worth considering whether similar processes might be at work in spoken languages.

Although our proposed role conflict constraint may not apply directly to speech, unmarked SOV sequences may still pose a similar problem in speech production as in pantomime production. For example, assume that the speaker's goal is to describe an event where a woman pushed a man. To the speaker, an unmarked SOV sequence such as "the woman the man pushed" might be dispreferred because there is a potential agent (man) intervening between the real agent (woman) and the action. Therefore, it may seem to the speaker as if he or she is momentarily producing an incorrect utterance ("the man pushed").

A stronger instance of this pattern might occur in well-known *local coherence* sentences like, "The coach smiled at the player tossed a Frisbee." Although difficult to understand, this sentence is grammatical; it is a reduced form of a relative clause ("The coach smiled at the player [who was] tossed a Frisbee"). These sentences have been used in psycholinguistic studies of sentence processing because they cause considerable difficulty in comprehension. According to Tabor, Galantucci, and Richardson (2004), the

main problem with these sentences is that they contain a *local coherence* -- the sequence, “the player tossed a Frisbee” is a locally coherent utterance, within which the player is the agent, tossed is the action, and the Frisbee is the patient. The problem is that this locally coherent unit is incompatible with the global meaning of the sentence, in which the player is not the agent of tossing, but the object of being smiled at. Psycholinguistic studies have used sentences such as this to gain insight about the mechanisms that support sentence comprehension (e.g. Levy, Bicknell, Slattery, & Rayner, 2009), but it is readily acknowledged that these sentences are “fairly obscure” (Tabor et al., 2004, p. 364). Our findings here raise the question of whether there are similar constraints on production, which would explain why these sentences are so rare. For example, one possibility is that instances of local coherence also seem like errors to the producer, who detects the local coherence during production (perhaps via monitoring) and avoids producing the utterance in that form. One way to test this hypothesis might be to elicit descriptions of such scenes but manipulate whether the critical verb uses the same form for its past tense and past participle (c.f. tossed/tossed vs. threw/thrown). If producers are sensitive to this local coherence, they should be more likely to produce the full relative clause for verbs like toss (“...at the player who was tossed the Frisbee”) than for verbs like throw (“...at the player who was thrown the Frisbee”). For now, we hypothesize that the pressure to avoid unmarked SOV for reversible events is universal, but that other factors – including communicative modality along with additional constraints on comprehension and acquisition – may influence the form of the solution.

If the role conflict constraint is ultimately shown to be modality-specific, at least two insights would result. First, we would learn that there are limits on the extent to

which results of elicited pantomime can explain attested patterns in spoken languages. Delineating the boundaries of any new tool is an important part of scientific progress. But second, we would gain additional insight into the cognitive foundations of emerging systems of manual communication (i.e. homesign and young sign languages). Recently, Meir, Padden, Aronoff, and Sandler (2007) reported a lexicalization pattern that was common across three sign languages of varying age, which they call “body as subject”. In this pattern, the signer’s own body (as distinct from the signer’s hands) represents the grammatical category of subject. This often corresponds to the semantic category of Agent, as was typically the case in the present studies, but other thematic roles are also possible, including Patient, Experiencer, Undergoer, and others. However, in each case the signer’s body represents the highest-ranking thematic role in its clause, which (among other factors) is argued to earn subject status. Meir et al. (2007) convincingly demonstrate that this lexicalization pattern is part of the grammar of these sign languages, and that it is one of the earliest syntactic patterns to emerge. They find sign languages that use body-as-subject while lacking other forms of verb agreement, but do not find sign languages that use other forms of verb agreement but lack body-as-subject.

Our results from elicited pantomime shed new light on the source of this pattern by demonstrating that the same phenomenon appears when naïve participants describe events in pantomime. This suggests that the roots of the body-as-subject pattern lie in aspects of cognition that are common across deaf and hearing individuals, both signers and non-signers. When a gestural system undergoes grammaticalization, the emerging syntax capitalizes on regularities that are already present in non-linguistic (or pre-

linguistic) gesture, such that body-as-subject becomes incorporated into a new sign language grammar before other types of verb agreement evolve.

Conclusions

We began by asking whether cognition tends to favor some constituent orders over others. The answer appears to be yes, and the semantic properties of the event are among the factors that determine which order is favored. By asking participants to describe reversible and non-reversible events in pantomime, we identified a cognitive constraint against the conjunction of a human object followed by an action gesture. The parallels between how this constraint is satisfied by our participants and by natural languages suggest that this cognitive pressure is active in shaping the structure of natural languages, and that it is operative during production. The extent to which these findings apply across both signed and spoken language modalities remains an open question for future research, but either outcome would be of scientific value.

Chapter 2, in part, has been submitted for publication of the material as it appears in 2012, Hall, M. L. Mayberry, R. I., & Ferreira, V. S. Cognitive constraints on constituent order: Evidence from elicited pantomime. The dissertation author was the primary investigator and author of this paper.

Appendix

Agent	Action	Patient	Reversibility
boy	kiss	ball	Non-reversible
woman	kiss	ball	Non-reversible
woman	kiss	bike	Non-reversible
woman	kiss	box	Non-reversible
boy	kiss	box	Non-reversible
boy	kiss	car	Non-reversible
man	kiss	car	Non-reversible
woman	kiss	cat	Non-reversible
boy	kiss	cat	Non-reversible
man	kiss	dog	Non-reversible
girl	kiss	trike	Non-reversible
man	lift	ball	Non-reversible
girl	lift	ball	Non-reversible
woman	lift	bike	Non-reversible
boy	lift	bike	Non-reversible
woman	lift	box	Non-reversible
girl	lift	box	Non-reversible
girl	lift	car	Non-reversible
boy	lift	cat	Non-reversible
woman	lift	cat	Non-reversible
woman	lift	dog	Non-reversible
woman	pet	ball	Non-reversible
boy	pet	ball	Non-reversible
woman	pet	bike	Non-reversible
boy	pet	bike	Non-reversible
woman	pet	box	Non-reversible
girl	pet	box	Non-reversible
boy	pet	car	Non-reversible
man	pet	car	Non-reversible
woman	pet	cat	Non-reversible
boy	pet	cat	Non-reversible
boy	pet	dog	Non-reversible
man	pet	dog	Non-reversible
boy	push	ball	Non-reversible
man	push	ball	Non-reversible
woman	push	bike	Non-reversible
man	push	box	Non-reversible
boy	push	box	Non-reversible
man	push	car	Non-reversible
boy	push	car	Non-reversible

boy	push	cat	Non-reversible
woman	push	cat	Non-reversible
man	push	dog	Non-reversible
boy	push	dog	Non-reversible
girl	push	trike	Non-reversible
woman	kiss	boy	Reversible
boy	kiss	girl	Reversible
woman	kiss	man	Reversible
boy	kiss	woman	Reversible
woman	lift	boy	Reversible
boy	lift	girl	Reversible
man	lift	woman	Reversible
girl	lift	woman	Reversible
woman	pet	boy	Reversible
boy	pet	girl	Reversible
man	pet	woman	Reversible
boy	pet	woman	Reversible
girl	push	boy	Reversible
woman	push	girl	Reversible
woman	push	man	Reversible
boy	push	woman	Reversible

References

- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, 21(3), 435-483.
- Aronoff, M., Meir, I., Padden, C., & Sandler, W. (2008). The roots of linguistic organization in a new language. *Interaction Studies*, 9(1), 131-150.
- Bakker, P. (2008). Pidgins versus creoles and pidgincreoles. In S. Kouwenberg & J. V. Singler (Eds.), *The Handbook of Pidgin and Creole Studies* (pp. 130-157). Walden, MA: Wiley-Blackwell.
- Bossong, G. (1991). Differential object marking in Romance and beyond. In D. Wanner & D. A. Kibbee (Eds.), *New Analyses in Romance Linguistics* (pp. 143-170). Philadelphia: John Benjamins.
- Comrie, B. (1981). *Language universals and linguistic typology: Syntax and morphology*. Basil Blackwell, Oxford.
- Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptation during language production: Limiting the role of the “model of the listener”. In D. Napoli & J. A. Kegl (Eds.), *Bridges Between Psychology and Linguistics: A Swarthmore Festschrift for Lila Gleitman* (pp. 105-129). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25, 257-271.
- Dryer, M. S. (2008). Order of subject, object, and verb. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The World Atlas of Language Structures Online* (<http://wals.info>). Munich: Max Planck Digital Library.
- Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation*, 49, 209-246.
- Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40(4), 296-340.
- Ferreira, V.S. (2007). How are speakers' linguistic choices affected by ambiguity? In A. S. Meyer, A. Krott, & L. R. Wheeldon (Eds.), *Automaticity and Control in Language Processing* (pp. 63-92), London: Psychology Press.
- Fischer, S. (1975). Influences on word order change in American Sign Language. In C. N. Li (Ed.), *Word order and word order change* (pp. 3-25). Austin, TX: University

of Texas Press.

- Gell-Mann, M. & Ruhlen, M. (2011). The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, 108, 17290-17295.
- Gibson, E., Brink, K., Piantadosi, S. & Saxe, R. (2011, April). *Cognitive pressures explain the dominant word orders in language*. Presented at the CUNY Conference on Human Sentence Processing, Palo Alto, CA.
- Givón, T. (1979). *On understanding grammar*. New York: Academic Press.
- Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.
- Goldin-Meadow, S. & Feldman, H. (1977). The development of language-like communication without a language model. *Science*, 197, 401-403.
- Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication in deaf children: Noneffect of parental input on language development. *Science*, 221, 372-374.
- Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, **391**, 279-281.
- Goldin-Meadow, S., Özyürek, A., Sancar, B., & Mylander, C. (2008). Making language around the globe: A crosslinguistic study of homesign in the United States, China, and Turkey. In J. Guo, E. Lieven, N. Budwig, & S. Ervin-Tripp (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 27-39). New York: Psychology Press.
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 105, 9163-9168.
- Greenberg, J. H. (1966). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language: Report of a conference held at Dobb's Ferry, NY, April 13-15, 1961* (pp. 73-113). Cambridge: MIT Press.
- Hall, M. L., Ferreira, V. S., & Mayberry, R. I. (2011, November). *Producers SOV use. Perceivers like SVO*. Poster presented at the annual meeting of the Psychonomic Society, Seattle, WA.
- Haviland, J. B. (2011). Nouns, verbs, and constituents in an emerging 'Tzotzil' sign language. In R. Gutiérrez-Bravo, L. Mikkelsen, & E. Potsdam (Eds.),

Representing Language: Essays in honor of Judith Aissen (pp. 157-171). Santa Cruz, CA: Linguistics Research Center.

Hawkins, J. (1983). *Word order universals*. New York: Academic Press.

Langus, A., & Nespors, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, 291-318.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, 106, 21086-21090.

Li, C. N. (1977). *Mechanisms of syntactic change*. Austin: University of Texas Press.

Li, C. N., & Thompson, S. A. (1974). An explanation of word order change SVO → SOV. *Foundations of Language*, 12, 201-214.

Kouwenberg, S. (1992). From OV to VO: Linguistic negotiation in the development of Berbice Dutch Creole. *Lingua*, 88, 263-299.

McWhorter, J. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology*, 5, 125-166.

Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. A. (2010, April). *The interaction of animacy and word order in human languages: A study of strategies in a novel communication task*. Paper presented at the 8th International Conference on the Evolution of Language, Utrecht, Germany.

Meir, I., Padden, C. A., Aronoff, M., & Sandler, W. (2007). Body as subject. *Journal of Linguistics*, 43, 531-563.

Newmeyer, F. J. (2000). On the reconstruction of "proto-world" word order. In C. Knight, M. Studdardt-Kennedy, & J. R. Hurford, (Eds.), *The Evolutionary Emergence of Language* (pp. 372-388). New York: Cambridge University Press.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, 102, 2661-2665.

Senghas, A., Coppola, M., Newport, E. L., & Supalla, T. (1997). Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the Boston University Conference on Language Development*, Vol 21 (pp. 550-561). Boston: Cascadilla Press.

- Sinnemäki, K. (2010). Word order in zero-marking languages. *Studies in Language*, 34, 869-912.
- Sun, C. & Givón, T. (1985). On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implications. *Language*, 61, 329-351.
- Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, 50, 355-370.
- Tomlin, R. S. (1986). *Basic word order: functional principles*. London: Croom Helm.
- Vennemann, T. (1973). Explanation in syntax. In J. P. Kimball (Ed.), *Syntax and Semantics, Vol. 2* (pp. 1-50). New York: Seminar Press.

Author note

The authors gratefully acknowledge members of the Laboratory for Multimodal Language Development, the Language Production Laboratory, the Sign Language Research Group, and the Center for Research in Language at UC San Diego for helpful feedback. We also thank Susan Goldin-Meadow for sharing stimulus materials and providing feedback, and Carolyn Mylander for guidance in using the coding system. This research would not have been possible without assistance with data collection and coding from Jessica Liu, Caitlin Scott, Shereen Cohen, Isabel Sharman, Kenny Oyama, Alex Kuo, Elena Churilov, Jayd Blankenship, Kristi Cheng, Lena Sun, and Youngju (Danbi) Ahn. Funding for this research was provided by NIH grant HD051030, NIH grant 5T32DC000041-19, and by the Rita L. Atkinson Graduate Fellowship.

CHAPTER 3

INVESTIGATING CONSTITUENT ORDER CHANGE WITH THE HANDS:
CONSISTENCY PROMOTES THE EMERGENCE OF SVO IN PANTOMIME

Abstract

One of the most basic functions of human language is to convey who did what to whom. In the world's languages, the order of these three constituents (subject (S), verb (V), and object (O), respectively) is uneven, with SOV and SVO being most common. Recent experiments using experimentally-elicited pantomime can account for the prevalence of SOV, but extant explanations for the prevalence of SVO could benefit from further empirical support. Here, we test whether SVO might emerge because (a) SOV is not well suited for describing reversible events (a woman pushing a boy), and (b) pressures to be efficient and mention subjects before objects conspire to rule out many other alternatives. We tested this by asking participants to describe reversible and non-reversible events in pantomime, and instructed some participants to be consistent in the form of their gestures and to teach them to the experimenter. These manipulations led to the emergence of SVO in speakers of both English (SVO) and Turkish (SOV).

One of the hallmarks of human language is that it is organized at multiple levels. For example, all languages allow only certain sounds and combinations of sounds; our knowledge of language is what tells us that “hing” is a possible word in English, but “ngih” is not. Likewise, languages allow only certain combinations of words in sentences. For example, English allows us to say, “The woman pushed the box”, but not typically “The woman the box pushed”. This level of organization is commonly called *constituent order*. The constituents in the preceding example can be described in both semantic and syntactic terms: “the woman” is the semantic agent (she does the action) and the syntactic subject; “the box” is the semantic patient (it receives the action) and the syntactic object; and “pushed” both identifies the semantic action and functions syntactically as the verb. Although our focus here is on the semantic categories of agent, patient, and action, we will describe these constituents with shorthand borrowed from the syntactic notions of subject (S), object (O), and verb (V).

English is an example of a language with fairly strict rules about constituent order. The grammatical example above uses SVO order; the ungrammatical example is SOV. And although it is possible to say, “The box was pushed by the woman” (OVS, in semantic terms), it is far less common in natural speech. However, other languages have different rules about which orders are preferred, unusual, or ungrammatical. Turkish is an example of a language where the default order is SOV, while Hixkaryana (spoken in northern Brazil) is a language that primarily uses OVS. In fact, linguists have established that all six logically possible orders of S, V, and O are attested as basic orders in the languages of the world. However, the distribution of these orders is extremely uneven, and these asymmetries have long been the subject of scrutiny in both linguistics and

psychology for what they tell us about way the human mind organizes the basic semantic elements of sentences.

One type of asymmetry concerns the distribution of constituent order across the languages of the world today, often called *synchronic variation*. Several large-scale surveys in recent decades (e.g. Dryer, 2008; Greenberg, 1963; Hawkins, 1983; Tomlin, 1986) have led to the widely accepted conclusion that the great majority of the world's languages belong to one of just two types: SOV or SVO, with SOV being slightly more common. VSO languages come in a distant third, with VOS, OVS, and OSV tapering off dramatically.

A second type of asymmetry concerns *diachronic variation*: predictable changes in constituent order over time. Just as not all constituent orders are equally common at any given point in time, it appears that languages do not change toward all orders equally often. Rather, there is a particularly intriguing pattern in which languages drift away from SOV, but do not toward it; languages tend to drift toward SVO (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977).

Identifying the underlying sources of these asymmetries has motivated a great deal of linguistic and psycholinguistic research in recent decades. This work has revealed multiple sources of influence. For example, geographic proximity (*areal effects*) increases the probability of direct contact between languages, which is one way that constituent order can change (Dryer, 1992). A language's own history is another powerful predictor. Indeed, recent work using analytical methods borrowed from evolutionary biology has revealed that at least some longstanding claims about putative universals of language may in fact be attributable to the languages in question having a

shared genealogy (e.g. Atkinson, 2011; Dunn, Greenhill, Levinson, & Gray, 2011; Levinson & Gray, 2012). In contrast, the primary focus of the present research is on a third potential source of language change: the nature of cognition itself.

Why this may be so can be illustrated by considering two thought experiments. First, consider whether and how language might first arise in a group of children who were raised under appropriate physical and social conditions, but who lacked language input. Would they develop language at all, and if so, what would its structure be? Any systematicity in their behavior could be attributable to their shared cognitive preferences as humans. A second thought experiment would be to imagine a closed community of people who share a single language, and who have no contact with speakers of any other language. If we could sample their language today and then again in 500 years time, would there be any differences? If so, any differences we observe could not be due to areal or genealogical factors, and would again be attributable to humans sharing certain cognitive preferences. Note that we use the term “cognitive” here in a very broad sense, encompassing communicative and social functions, since all complex human behavior is at some level a byproduct of the human mind. We intend it to stand in contrast to external influences on language change, such as contact with other languages. In all, it may be the case that over many successive iterations of production, comprehension, and acquisition, people’s language behavior gravitates toward certain predictable states.

Of course, these experiments are necessarily fictitious; conducting them would be both unethical and impractical. However, there are several approximations of experiments like these, and to the extent that diverse paradigms converge on similar answers, we can gain confidence in the accuracy of these approximations. One

substantial body of work has focused on the case of language emergence *de novo*, as in the first thought experiment above. This research considers how deaf children create inchoate communicative systems called *homesign*, and how the aggregation of multiple deaf homesigners (and sometimes their hearing interlocutors) can lead to the evolution of a full sign language. A robust finding in this literature is that when homesigners produce utterances with two arguments (e.g., subject and verb, or verb and object), they have a strong tendency to use SV and OV, rather than VS or VO (for a review, see Goldin-Meadow, 2003). This tendency is observed cross-culturally (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, Özyürek, Sancar, & Mylander, 2008), and cannot be attributed to the spoken or gestural input from hearing parents or the surrounding culture (Goldin-Meadow & Mylander, 1983). Furthermore, this pattern appears to carry over into the early stages of some young sign languages, where three-argument utterances (those with subject, verb, and object) are more common. Haviland (2011) reports preliminary evidence of SOV, SV, and OV order in a very young sign language used by a single family with three deaf siblings where the ambient spoken language, Tzotzil, uses VOS order. Sandler, Meir, Padden, and Aronoff (2005) have analyzed the constituent order of Al-Sayyid Bedouin Sign Language (ABSL), which has emerged over the past 75 years as a gene for recessive deafness has become relatively widespread within the village community. They have found that ABSL is a robust SOV language; that such consistency has emerged so quickly is all the more significant considering that this language has not yet converged on a unified phonological system (Aronoff, Meir, Padden, & Sandler, 2008). Nicaraguan Sign Language (ISN) emerged when large numbers of isolated homesigners were brought together in a school setting. Senghas, Coppola,

Newport, and Supalla (1997) report that SOV, SV, and OV were commonly used by the first cohort of ISN signers, although OSV was also present. Thus, these findings are consistent with the idea that in the absence of linguistic input something about human cognition privileges SOV order. However, no exhaustive typological work has yet established how stable this pattern is across all emerging sign languages, nor whether new sign languages vary in how long they retain SOV order as the language matures.

This latter question relates to the second thought experiment we proposed above: whether constituent order would change across time, even without outside influence from other languages. As mentioned previously, abundant research on spoken languages indicates that over long time scales, many languages do shift away from SOV and toward SVO, but not vice versa (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977). Although all of the large-scale studies to date have focused exclusively on spoken languages, Fischer (1975) argues that the same SOV-to-SVO shift has also occurred in American Sign Language.

Although the phenomenon of constituent order change has been well described, assessing the role of cognition in shaping constituent order over such time scales is difficult. More compelling evidence would come from empirical demonstrations that constituent order preferences can vary under different circumstances even within a single individual and within a single testing session. Three recent studies have aimed to do just that by using a novel empirical technique for measuring constituent order preferences: elicited pantomime. In all three, naïve hearing participants of varying language backgrounds were shown pictures or video clips of simple events and were asked to

describe those events in pantomime (i.e., gesture without speech). We discuss each study in turn.

Goldin-Meadow, So, Özyürek, and Mylander (2008) were the first to use elicited pantomime to assess constituent order preferences. They tested participants whose native language used SOV (Turkish), SVO (English, Spanish), or allowed both orders (Chinese), and found that when describing events in pantomime, speakers of all these language backgrounds showed a strong propensity to use orders that were consistent with SOV ordering (SOV, SV, and OV). In this respect, the data strongly resemble the constituent orders found in homesign and young sign languages, as discussed above. Goldin-Meadow et al. (2008) view these data as evidence that humans organize knowledge about events in a way that is most naturally linearized in SOV order, and conclude that one reason SOV order is so prevalent among the world's languages is that it is compatible with mental representations of these events.

Thus far, the elicited pantomime results nicely account for the prevalence of SOV, but do not explain why SVO is nearly as common, or why languages commonly change from SOV to SVO but not vice versa. The second study, by Langus and Nespors (2010), investigated this issue. To explain the prevalence of SVO, Langus and Nespors suggest that there are two separate cognitive systems at work. One, the *conceptual* system, is responsible for non-linguistic processing and prefers SOV. The other, the *computational* system, is responsible for linguistic processing and prefers SVO. They replicated the basic SOV preference from Goldin-Meadow et al. (2008) and conducted a second experiment showing that pantomimed utterances did not have the characteristic structure that would be expected if the pantomimes were being generated by a truly linguistic

(syntactic) system. In two additional experiments, they tested how quickly participants responded to various constituent orders when scenes were described with gestures (e.g., GIRL BALL THROW) versus written words (e.g. “girl ball throw”). They found that for gestured stimuli, participants responded more quickly to OV-type orders, but if stimuli were printed words, participants responded more quickly to VO-type orders. Importantly, this was true for speakers of both Italian (SVO) and Turkish (SOV). They concluded that gestures engage the conceptual system, which prefers SOV, whereas words engage the computational system, which prefers SVO. They then suggested that the prevalence of both SOV and SVO in the world’s languages results from the interaction of these two systems.

This account fares well in explaining the extant data, but assumes that these two systems are distinct. Langus and Nespors (2010, p. 291) explicitly state, “We rely on the proposal that the human faculty of language is modular and that it is possible to identify different cognitive systems responsible for specific linguistic tasks (Chomsky, 2000; Fodor, 1983)”. According to this account, the SVO preference is posited to be part of the human innate language faculty in need of no further motivation or justification.

The goal of the present study is to test whether a preference for SVO might be explicable in terms of cognitive-functional pressures that might bias certain constituent orders over others. That is, rather than positing a built-in preference for SVO, is it possible to identify specific factors that would cause a communication system to shift from SOV to SVO? The third study using elicited pantomime (Hall, Mayberry, and Ferreira, submitted) has already taken a step in this direction. This study drew on three longstanding observations from linguistics. First, most SOV languages identify agents

and patients overtly by using case marking (suffixes (or equivalent) that indicate a word's grammatical role), whereas this is less true of SVO languages (Greenberg, 1963). Second, overt object marking is especially common cross-linguistically when the object is a potential subject such as a human, known as *differential object marking* (Aissen, 2003; Bossong, 1991). Third, shifts from SOV to SVO are often triggered when a language loses its case marking system (Sapir, 1921; Sinnemäki, 2010; Vennemann, 1973). Taken together, these observations suggest that in the absence of case marking, SOV order may not be well-suited for describing events in which both the agent and the patient are plausible agents (e.g. a boy kissing a woman). We refer to these as *reversible events*. Neither Goldin-Meadow et al. (2008) nor Langus and Nespors (2010) used reversible events in their experiments; all of their stimuli involved *non-reversible events*, where semantic role assignment was clear from the meaning of the gesture alone, regardless of constituent order (e.g. a boy kissing a box). To address this gap, Hall, Mayberry, and Ferreira (submitted) asked native speakers of English (SVO) to describe both non-reversible and reversible events in pantomime. They found that although participants showed a robust preference for SOV for non-reversible events, they strongly avoided using SOV for reversible events. Instead, the most common solution that participants used was SVO, suggesting that perhaps SVO arises in the world's languages because it is well suited for describing both reversible and non-reversible events.

However, the participants in Hall, Mayberry, and Ferreira (submitted) were all native speakers of English (SVO). In order for these results to be fully generalizable, we need to know whether SOV speakers also avoid using SOV to describe reversible events in pantomime. Preliminary support for this hypothesis comes from Meir et al. (2010)

who report pilot data in which 9 Turkish (SOV) speakers tended to avoid SOV for reversible events. These findings are echoed by Gibson et al.'s (2011) finding that bilingual speakers of Japanese (SOV) and English (SVO) also avoided using SOV for reversible events in pantomime. Although the available data indicate that SOV speakers are likely to avoid SOV for reversible events, it is not yet known whether they, like English speakers, would also favor SVO as a solution in a pantomime task.

In the experiments reported by Hall, Mayberry, and Ferreira (submitted), participants produced a wide variety of non-SOV constituent orders to describe reversible events. Although SVO was the most common of these, it accounted for only 18-38% of reversible utterances across all experiments, leaving many other alternative orders besides SOV and SVO. Interestingly, many of these alternative orders are not commonly found in natural languages. In particular, these unusual orders tended to be inefficient, or to mention the object before the subject (or both). These unusual orders are highly relevant to the issues explored here, and so we describe each type of order in turn.

All of the stimuli in Hall et al. (submitted) contained one agent, one patient, and one action. Thus, a maximally efficient way to describe the events would be to use one of the six constituent orders consisting of three and only three elements (SVO, SOV, VSO, VOS, OVS, and OSV). Utterances that omit one or more elements (e.g., SV, VO, V, etc.) are underinformative when the to-be-described event is transitive. Those that repeat at least one element (e.g., SOSV, SOSOV, SOSVO, etc.) include redundant information. It is important to note that among the languages of the world, none are known to have a basic constituent order that is inefficient (either underinformative or repetitious). Although there is substantial variability in which constituent order a

language will use most often, the order used is always one that is maximally efficient, under this definition. Accordingly, Slobin (1977) argues that the pressure for language to be efficient is one of the major forces that shapes linguistic structure. In this respect, it appears that there are pressures that constrain natural languages that did not similarly constrain elicited pantomime in previous experiments. Perhaps the inefficient orders we observed in pantomime would decrease under conditions that were closer to natural language, leaving more efficient orders like SVO and OSV.

Another common strategy that participants used for describing reversible events in pantomime was to put the object before the subject (e.g. OSV, OSOV, OSVO, etc.). Once again, this tendency is uncommon as a dominant order in natural language; some early researchers even claimed that true O-before-S languages were unattested (Parker, 1980; Venneman, 1973), but we now know that such languages do exist (Derbyshire & Pullum, 1981; Pullum, 1981). Still, they are confined mainly to the Amazon basin and surrounding areas, and orders with O-before-S (VOS, OVS, OSV) are unquestionably the rarest of the six efficient constituent orders.

Various explanations have been proposed for why subjects generally precede objects, including those that emphasize grammar-internal motivations (Greenberg, 1963; Hockett, 1958; Venneman, 1974, 1975), those based on psychological processing (e.g. Dayley, 1985; Gibson, 1998; Hawkins, 2002; 2007; Newmeyer, 2002), and those based on evolution in terms of both phylogeny and ontogeny (Givón, 1979). Discriminating among these various accounts is not the goal of the present study; rather, for our purposes, it suffices to note that the vast majority of the world's languages use a basic constituent order in which the subject precedes the object. Therefore, the fact that O-before-S orders

commonly appear when participants describe reversible events in pantomime (Hall et al., submitted; Meir et al., 2010) once again suggests that some of the pressures that constrain natural languages did not constrain elicited pantomime under the conditions tested in these experiments.

In the present experiments, we evaluate the hypothesis that SVO emerges in natural languages in part because SVO allows participants to avoid SOV for reversible events while still being efficient and keeping the subject before the object. This hypothesis is not novel; however, no previous studies have demonstrated that these forces are latently present and at work in the minds of naïve participants, even over very short time scales (i.e., within a single testing session). In particular, we explore the possibility that being efficient and ordering the subject before the object might become increasingly relevant constraints under more language-like circumstances, compared to previous studies of elicited pantomime, which have been fairly unconstrained. Therefore, we aimed to create a controlled laboratory situation where participants would describe both non-reversible and reversible events in pantomime, but where some participants would be in more language-like situations than others. We hypothesized that under the demands of more language-like situations, pantomime might start to exhibit more of the characteristics of natural languages, such as a decrease in inefficient and O-before-S orders, and increases in SVO, especially for reversible events.

To manipulate the extent to which the pantomime situation was language-like, we considered the ways in which new natural human languages emerge in the visual-manual modality: that is, homesign and new sign languages. One of the first features to emerge in these systems is a lexicon, in which a consistent reciprocal mapping develops between

a gesture and a basic category of referents. For example, a long-hair gesture that originated as a way to describe a particular woman with long hair might be extended to refer to the general category “woman”, regardless of the length of a woman’s hair. We therefore chose to simulate this stage of language emergence by asking some participants to create and use a consistent gestural lexicon for describing transitive events in pantomime.

Another difference between previous experiments with pantomime and natural languages is that natural languages (including homesign) arise in the presence of other people. In theorizing about the various forces that shape language structure, Singleton, Goldin-Meadow, and McNeill (1995, p. 307) suggest that,

Having people to share your communication system with also may be a key to... reaching a truly arbitrary, efficient linguistic system... It is evident that by entering into and sharing a communication system with a willing partner, these kinds of evolutionary pressures would be introduced. Consequently, what might begin as a rudimentary communication system (e.g., iconic gestures created in the absence of speech) would evolve into a system that more closely approximates a full-fledged natural language.

Therefore, in the present study, the participants who received instruction to be consistent in the form of their gestures were further divided into two groups: shared and private.

In the shared condition, participants taught their gestures to the experimenter, who confirmed understanding of the form-meaning mapping of each gesture in isolation, but did not otherwise use the gestures for any communicative purpose. We contrast the shared condition with the *baseline condition*, in which (as in previous experiments) participants were simply instructed to describe events in pantomime with no requirement to be consistent in the form of their gestures.

In the private condition, participants were instructed to create and use a consistent gestural lexicon, but did not share it with the experimenter, who left the room during testing. We predict the private condition to be intermediate between the baseline and shared conditions in terms of its language-likeness.

To summarize, the present study builds on previous accounts of why SOV and SVO are both prevalent among the world's languages, with SOV common in a language's earlier stages while diachronic change favors SVO. To explain the prevalence of SOV, Goldin-Meadow et al. (2008) suggest that the way we organize our mental representations of event knowledge is especially compatible with SOV order. To explain the prevalence of SVO, Langus and Nespors (2010) suggest that, in addition to such conceptual representations, the human mind also houses a separate computational (linguistic) system, which prefers SVO. However, no explanation is offered for why that might be the case beyond an appeal to ideas about built-in preferences that are innately specified in the human language faculty (e.g. Chomsky, 2000). The present study aims to offer an explanation grounded in cognitive constraints on processing. Specifically, we note that there is already strong evidence that participants avoid using SOV to describe reversible events, but that in previous studies, many people opted for alternative orders that are dispreferred in natural language, including those that are inefficient, and those that put object before subject. We hypothesize that under circumstances that are a closer approximation of natural language situations, we will observe a decrease in constituent orders that are inefficient and that put object before subject. The interaction of these pressures would leave SVO as a preferred solution for describing reversible events, potentially explaining its prevalence both synchronically and diachronically. Therefore,

to make the elicited pantomime task slightly more language-like, we instructed some participants to create and use a consistent gestural lexicon, and we contrast the constituent orders produced by participants in the consistent-gesture conditions with those produced by participants in a baseline condition, where no instruction was given about being consistent in the form of their gestures.

Experiment 1

Method

Participants. We tested 36 undergraduates at UC San Diego who reported being native monolingual speakers of English. All participants gave consent to participate and be video recorded as part of the study, and received course credit for their participation.

Materials. Stimuli were 61 video clips of a human agent (man, woman, boy, girl) performing a transitive action (lift, kiss, pet, push) on a patient that was either non-human (ball, box, bike, car, cat, dog) or human (man, woman, boy, girl). Events with non-human patients are termed “non-reversible”. Events with human patients are termed “reversible”; in these cases, the agent and patient always differed in age, sex, or both. A full list of stimuli is given in Appendix A; these same stimuli were used in Hall, Mayberry, and Ferreira (submitted). Stimuli were presented in a Macintosh laptop using PsyScope X software (Cohen, MacWhinney, Flatt, & Provost, 1993).

Design and procedure. Participants were randomly assigned to one of three groups (baseline, private, or shared), yielding 12 participants per condition. In all cases, instructions were immediately followed by a brief exposure phase, where participants saw a preview of the kinds of things about which they would need to gesture. The exposure phase began with video sequences showing diverse instances of each action

(e.g., several clips of lifting, several kissing, etc.). The videos were constructed so as to emphasize the common action shared by each short segment. After briefly exposing the participants to each of the four action elements, the nominal elements were presented as still images, one category at a time. For example, the first screen displayed still images of the various different balls that would appear in the test clips. The next screen displayed the various boxes, followed by bikes, cars, cats, dogs, and finally men, women, boys, and girls, in that order. This order of exposure (essentially VOS) was chosen so that it would not bias either SOV or SVO.

The critical manipulation was the instructions that were given during the exposure phase. Participants in the baseline condition were told that they would see short video clips that they would later describe in gesture, and that the exposure phase was just a preview of the kinds of things they might see. For them, there was no task to perform during the exposure phase. Participants in the private group were given the same basic instructions, but were also asked to use the exposure phase to think of a gesture that they would then use consistently throughout the session to refer to each action and each nominal element. However, they were not asked to produce those gestures overtly until the beginning of practice. The goal was to help the participants create a gestural lexicon that they would use consistently, but that would not be shared with the experimenter, who left the room after the practice trials in this condition. Participants in the shared group received the same instructions as in the private group, except that they were asked to overtly teach their gestures to the experimenter, who both confirmed understanding, and remained visible in the room throughout the testing session.

Following the exposure phase, participants performed 8 practice trials during which the experimenter encouraged them to include gestures for all three components of each event (agent, patient, action), but carefully avoided using language that would bias participants toward any single order. After the practice trials, participants were given 61 test trials. To facilitate comparison with previous studies that have used only non-reversible events (Goldin-Meadow et al., 2008; Langus & Nespors, 2010), participants began with 45 non-reversible events, followed by 16 reversible events. Hall, Mayberry, and Ferreira (submitted) found no differences in constituent order when reversible trials came at the end of the testing session compared to when they came at the beginning or were mixed throughout. Within these blocks of non-reversible and reversible trials, we used a fixed random order with the constraint that consecutive trials shared no more than one element.

Coding and analysis. Subjects' gestures were transcribed by two trained coders, who identified the each gesture's referent as either Subject, Verb, Object, or other. Multiple *consecutive* gestures for the same referent were considered as belonging to the same constituent. For example, the string "TALL LONG-HAIR WHEELS PEDALS HANDLEBARS PUSH" would be parsed as SOV (woman, bicycle, push). However, if the string were non-consecutive, a repetition would be noted. For example, "TALL LONG-HAIR WHEELS PEDALS PUSH HANDLEBARS" would be coded as SOVO (woman bicycle push bicycle). In the absence of any principled way to determine phrase or utterance boundaries in pantomime, everything that the participant produced was counted as a single utterance unless there was a pause of more than 2 seconds or the string was interrupted by the participant replaying the video. In these cases, the most complete

string containing an action gesture was coded. Gestures classified as “other” were ignored in determining constituent order; these mainly contained information about the environmental surroundings of the scene. If it was not possible to determine the referent of a gesture, the trial was considered “ambiguous” (in gray below). If participants produced gestures that referred to more than one constituent at the same time, such that it was not possible to compute order of mention, it was considered “simultaneous” (in green below). Trials coded as ambiguous or simultaneous were excluded from statistical analysis.

Coders were not blind to the experimental manipulations. However, they were blind to one another’s ratings, and agreed on 2018 of the 2196 utterances (91.9%). In cases of disagreement, a third rater coded the trial blind to the other raters’ responses. If this resulted in 2 of 3 coders agreeing, the trial was included. If all three coders disagreed, the trial was excluded from statistical analysis (45 trials = 2%, in black below).

Results

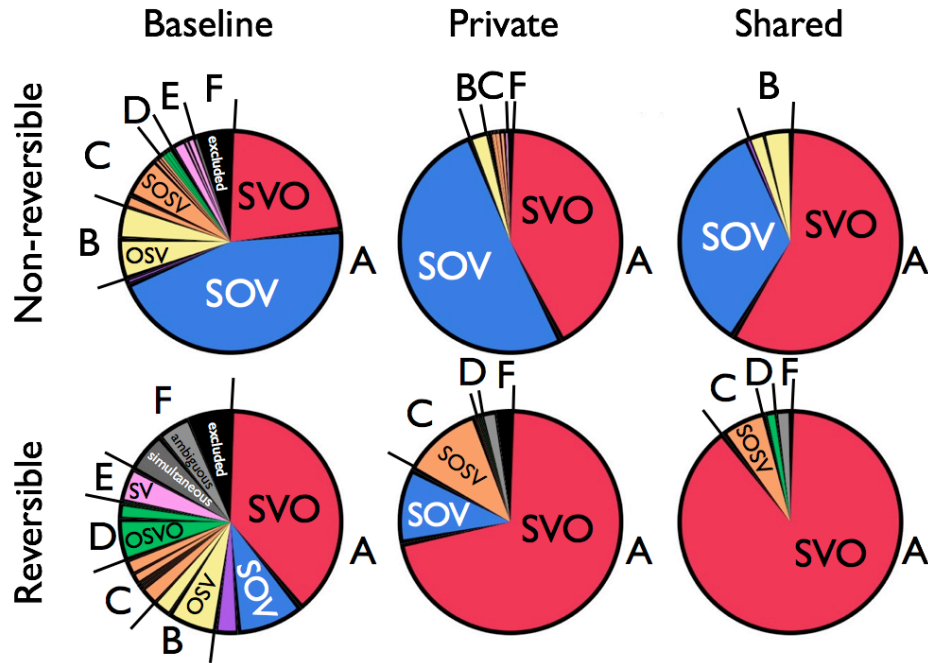


Figure 1. Distribution of constituent orders in Experiment 1 (English speakers). Region A shows efficient orders with subject before object. Region B shows efficient orders with object before subject. Region C shows repetitious orders with subject before object. Region D shows repetitious orders with object before subject. Region E shows underinformative orders. Region F shows orders that were coded as ambiguous, simultaneous, or excluded due to lack of consensus.

Prevalence of SOV. Only trials with SOV order were coded as consistent with SOV; all other trials were coded as inconsistent. We analyzed whether our manipulations of semantic reversibility (within subjects) or gestural consistency (between subjects) influenced the prevalence of SOV at both the group level and at the level of individual participants. Group-level data are displayed in Figure 1.

Group results. For each participant, we computed the proportion of SOV trials and submitted these proportions to a 2 x 3 mixed ANOVA with reversibility (non-reversible, reversible) as a within-subjects factor and group (baseline, private, shared) as a between subjects factor. We found that group did not influence the prevalence of SOV

[$F(2,33) = .79, p = .46$], but that participants used significantly less SOV for reversible trials [$F(1,33) = 29.66, p < .001$]. These effects did not interact [$F(2,33) = .07, p = .93$].

Individual results. It was not the case that every participant resembled the group average; rather, individual participants tended to have a particular order that they used more often than any other within each level of reversibility. We defined this as the participant's dominant order, and used Fisher's exact test to determine whether the reversibility manipulation influenced the probability of participants being SOV-dominant. In the baseline group, 7/12 participants were SOV-dominant for non-reversible events but only 1/12 was SOV-dominant for reversible events ($p < .02$). In the private group, 6/12 participants were SOV-dominant for non-reversible events but only 1/12 was SOV-dominant for reversible events ($p < .04$). In the shared group, 4/12 participants were SOV-dominant for non-reversible events, but 0/12 were SOV-dominant for reversible events ($p < .05$).

We also used Fisher's exact test to determine whether the between-subjects manipulation of gestural consistency influenced the probability of participants being SOV-dominant by comparing baseline to private and baseline to shared for each level of reversibility. For non-reversible events, the number of SOV-dominant participants in the baseline group (7/12) was not significantly different from the number of SOV-dominant participants in the private (6/12, $p = .5$) or shared (4/12, $p = .2$) groups. The same was true for the reversible events: the number of SOV-dominant participants in the baseline group (1/12) was not significantly different from the number of SOV-dominant participants in the private (1/12) or shared (0/12) groups.

Prevalence of SVO. Following the above logic, only trials with SVO order were coded as consistent with SVO; all other orders were coded as inconsistent. Unless otherwise noted, the statistical methods below are identical to those mentioned above. It is important to note that these analyses of SVO are not strictly independent of the preceding analyses of SOV, because a trial cannot be both SOV and SVO. However, because the hypotheses that we aim to test make different predictions about the circumstances under which SOV and SVO should increase and decrease, it is important to examine both dependent measures. We note (a) that when the same manipulation affects SOV rates and SVO rates in opposite directions, this should only be taken as a single observation in support of the effect of that manipulation (because the non-independence could drive the second effect), and (b) that we do not compare rates of SOV and SVO directly, mitigating some concerns regarding their non-independence.

Group results. The 2 x 3 ANOVA revealed that SVO was more prevalent in some groups than others [$F(2,33) = 3.90, p < .04$]. Planned contrasts found that SVO was more common in the shared group than in the baseline group [$F(1,33) = 7.68, p < .01$]. For the private group, the increase in SVO was not significantly different from baseline [$F(1,33) = 2.80, p = .10$]. SVO was also significantly more common for reversible trials than for non-reversible trials [$F(1,33) = 17.11, p < .001$], in complement to the results with SOV rates. There was no group x reversibility interaction [$F(2,33) = .62, p = .54$].

Individual results. We again used Fisher's exact test to determine whether reversibility influenced the probability that individual participants in each condition would be SVO-dominant. Although more participants in each group were SVO-dominant for reversible than for non-reversible events, these differences did not reach

significance [baseline: 3/12 vs. 5/12, $p = .24$; private: 6/12 vs. 9/12, $p = .16$; shared: 7/12 vs. 11/12, $p = .07$].

Finally, we used Fisher's exact test to determine whether the between-subjects manipulation of gestural consistency influenced the probability of participants being SVO-dominant by comparing baseline to private and baseline to shared, for each level of reversibility. For non-reversible events, the number of SVO-dominant participants in the baseline condition (3/12) was not significantly different from the number in private (6/12, $p = .16$) or shared (7/12, $p = .09$). For reversible events, the increase in the number of SVO-dominant participants did not reach statistical significance from baseline (5/12) to private (9/12, $p = .09$), but did reach significance from baseline (5/12) to shared (11/12, $p < .02$).

Discussion

Consistent with previous research (Gibson et al., 2011; Hall, Mayberry, & Ferreira, submitted; Meir et al., 2010), participants avoided using SOV to describe reversible events. However, the instruction to create and use a consistent gestural lexicon did not influence how often participants used SOV descriptions. This pattern was robust at both the group and individual levels. Such findings are to be expected if indeed these two manipulations interact with two different cognitive pressures. If the instruction to use a consistent gestural lexicon incentivizes efficiency (i.e., including S, O, and V exactly once each) and information structure (i.e., keeping S before O), then SOV is a perfectly adequate order, and so is not predicted to decrease across groups. SOV becomes problematic only when the to-be-described events are reversible, at which point participants should avoid it.

We did, however, find that the prevalence of SVO was influenced by both reversibility and by the differential instructions given to each group. As in previous studies (Gibson et al., 2010; Hall, Mayberry, & Ferreira, submitted), participants used more SVO for reversible events than for non-reversible events. This is consistent with the idea that using SVO is one way to avoid role conflict, minimize confusability, or both. A novel finding of the present study is that SVO became more prevalent when participants were instructed to create and use a consistent gestural lexicon, especially in the shared group (where that lexicon was also understood by a passive interlocutor). For reversible events, this effect is a straightforward consequence of the interaction of three cognitive pressures: if SOV is not a good option for describing reversible events (because of role conflict, confusability, or both), and if it is important to maximize efficiency and to keep the subject before the object, then SVO is the only order that satisfies those three constraints.

One unexpected finding, however, was that the instruction to create and use a consistent gestural lexicon increased SVO not only for reversible events, but also for non-reversible events. Because SVO is also an efficient order with S before O, it should be preferred to orders like SOSOV, OSV, and VOS, which all occurred more in the baseline group than in the private and shared groups. The unexpected aspect of this finding was that SOV should have been just as good a solution on those grounds, and so we might have expected to see both SOV and SVO increase, but only SVO became more frequent across groups. This, in turn, raises the possibility that some or potentially all of the increase in SVO across groups could come from another source: the participants' native language. It may be that the process of creating and using a gestural lexicon encourages

participants to silently recode their gestures into words in their native language. That, in turn, could then bias the order in which participants gesture to more closely reflect the order of their native language: in this case, SVO.

Therefore, the data from Experiment 1 cannot determine the extent to which the increase in SVO across groups reflects a potentially universal cognitive pressure, a language-specific preference for SVO, or a combination of both. To explore this question in further detail, we replicated Experiment 1 with native speakers of Turkish, whose language uses SOV structure. Our hypothesis predicts that SVO should still emerge in reversible events when participants are instructed to create and use a gestural lexicon. If so, it cannot be attributed to influence from participants' native language, which would instead work against this finding.

Experiment 2

Method

Participants. All testing was conducted in Turkey by a native Turkish speaker, mainly in Sariyer and Istanbul. An ideal participant would have no contact with or knowledge of any SVO language. Since that is highly unlikely, we asked all potential participants about their experience with other languages prior to selection. Potential participants were excluded if an SVO language was spoken in their home. All but one of the participants were raised in a home where only Turkish was spoken; the one exception had one parent who spoke Arabic (VSO) at home. (Two participants reported having one parent who was fluent in an SVO language (Albanian), but did not indicate that it was spoken in their home.) Roughly two thirds of potential participants reported having some contact with English or another SVO language in school. We therefore asked potential

participants to self-report their level of proficiency in these languages using the multiple-choice scale in Table 1. Potential participants were excluded if they reported “3” or above in any SVO language. This left 33 participants, of whom 9 reported “0”, 19 reported “1”, and 5 reported “2”. These questions were selected because they were easy to administer and score without participants or experimenters needing to travel, and because they were easy for potential participants to understand (even those with limited schooling). We did not verify these self reports with any empirical tests (a) because of limited time in experimental sessions, (b) because it would have been impossible to prepare test materials in all possible SOV languages, and (c) because we wanted to encourage participants to be in fully “monolingual mode” throughout their recruitment and testing (Grosjean, 1992). All participants gave consent to be videotaped as part of the study, and were paid for their participation.

Table 1. Language screening questions for Experiment 2.

Language screening questions
5 = I am comfortable using this language in all situations.
4 = I am mostly comfortable using this language, but I prefer another one in some situations.
3 = I can use this language when necessary, but prefer not to.
2 = I can usually make myself understood, but it’s difficult.
1 = I would only use this language if I had no other options.
0 = I don’t remember any of this language.

Materials. We used the same materials as in Experiment 1.

Design and procedure. The design and procedure were identical to Experiment 1, except that written and spoken instructions were delivered in Turkish.

Coding and analysis. Coding procedures were identical to Experiment 1. The first two coders agreed on 1915/2013 utterances (95.1%). After the third coder, only 27 trials (1.3% of the data) were excluded. Unless otherwise noted, the statistical methods were identical to those in Experiment 1.

Results

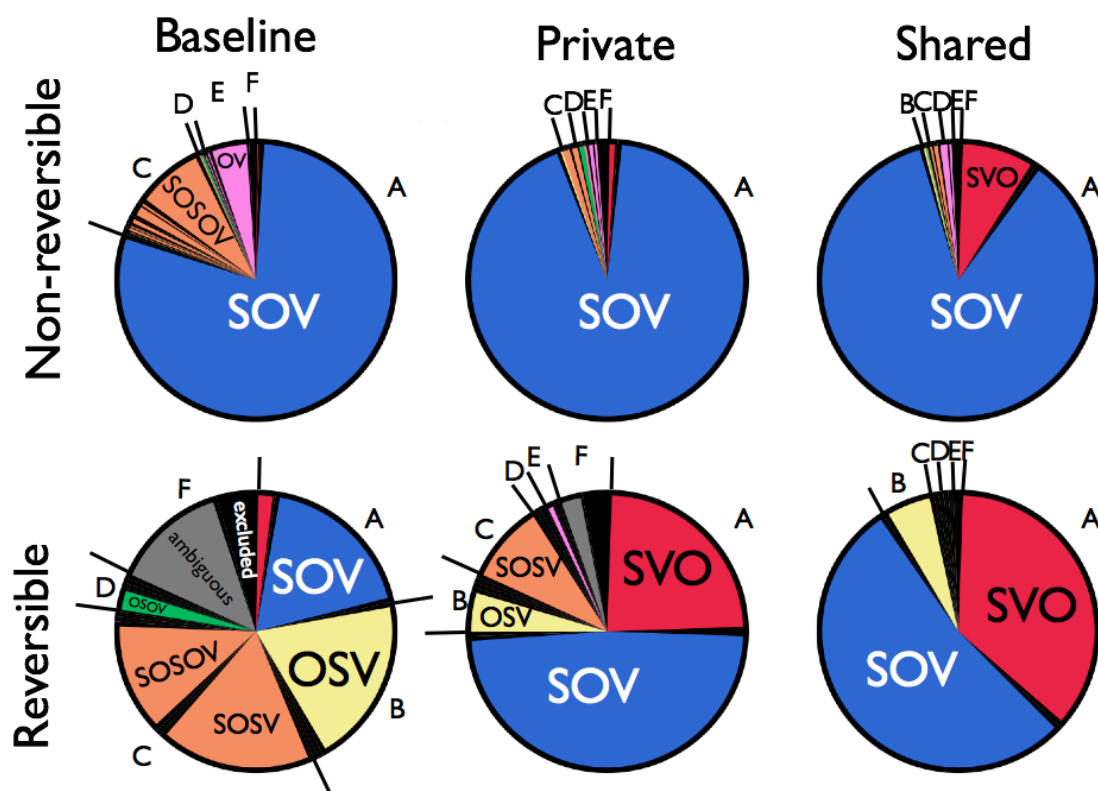


Figure 2. Distribution of constituent orders in Experiment 2 (Turkish speakers). Region A shows efficient orders with subject before object. Region B shows efficient orders with object before subject. Region C shows repetitious orders with subject before object. Region D shows repetitious orders with object before subject. Region E shows underinformative orders. Region F shows orders that were coded as ambiguous, simultaneous, or excluded due to lack of consensus.

Prevalence of SOV. As in Experiment 1, the proportion of trials that had SOV order was analyzed at both the group and individual level. Group-level data are displayed in Figure 2.

Group results. The 2 x 3 ANOVA revealed a trend for SOV to be more common in some groups than others [$F(2,30) = 2.84, p = .07$]. Planned comparisons found that SOV was slightly more common in the private group than in the baseline group [$F(1,30) = 4.49, p < .05$], and that SOV was marginally more common in the shared group than in the baseline group [$F(1,30) = 4.02, p = .05$]. SOV was significantly less common on reversible events than on non-reversible events [$F(1,30) = 47.02, p < .001$]. There was no interaction between group and reversibility [$F(2,30) = 1.53, p = .23$].

Individual results. At the individual level, we used Fisher's exact test to determine whether the reversibility manipulation influenced the probability of participants being SOV-dominant. In the baseline group, 10/11 participants were SOV-dominant for non-reversibles, whereas 0/10 were SOV-dominant for reversibles ($p < .001$). In the private group, 11/11 participants were SOV-dominant for non-reversibles, whereas 4/11 were SOV-dominant for reversibles ($p < .01$). In the shared group, 10/11 participants were SOV-dominant for non-reversibles, and 6/11 were SOV-dominant for reversibles ($p = .07$).

We also used Fisher's exact test to determine whether the between-subjects manipulation of gestural consistency influenced the probability of participants being SOV-dominant by comparing baseline to private and baseline to shared for each level of reversibility. For non-reversible events, the number of SOV-dominant participants in the baseline group (10/11) was not significantly different from the number of SOV-dominant

participants in the private (11/11) or shared (10/11) groups. For reversible events, the number of SOV-dominant participants in the baseline group (0/11) was significantly different from the number of SOV-dominant participants in the private group (4/11, $p < .05$) and the shared group (6/11, $p < .01$).

Prevalence of SVO. Following the above logic, the proportion of trials that had SVO order was analyzed at both the group and individual level. Group-level data are displayed in Figure 2.

Group results. The 2 x 3 ANOVA revealed a trend for SVO to be more common in some groups than others, as evidenced by a trend toward a main effect of group [$F(2,33) = 2.54, p = .10$]. Planned comparisons found that the prevalence of SVO was not significantly different between the baseline and private groups [$F(1,30) = 1.47, p = .24$], but that SVO was significantly more common in the shared group than in the baseline group [$F(1,30) = 5.08, p < .04$]. SVO was also significantly more common in reversible events than in non-reversible events [$F(1,30) = 7.64, p < .01$], complementing the results with SOV orders. There was no interaction between group and reversibility [$F(2,30) = 1.56, p = .23$].

Individual results. At the individual level, we used Fisher's exact test to determine whether the reversibility manipulation influenced the probability of participants being SVO-dominant. In the baseline group, 0/11 participants were SVO-dominant for non-reversible events, and 0/11 were SVO-dominant for reversible events ($p = 1$). In the private group, 0/11 were SVO-dominant for non-reversibles, and 3/11 were SVO-dominant for reversibles ($p = .11$). In the shared group, 1/11 was SVO-dominant for non-reversibles, and 4/11 were SVO-dominant for reversibles ($p = .14$).

For the crucial comparison, we also used Fisher's exact test to determine whether the manipulation of group influenced the probability of participants being SVO-dominant by comparing baseline to private and baseline to shared for each level of reversibility. For non-reversible events, there were no differences between the number of SVO-dominant participants in the baseline (0/11), private (0/11), and shared (1/11) groups (all p values $> .5$). For reversible events, the difference between baseline (0/11) and private (3/11) did not reach significance ($p = .11$), but there was a significant difference between the number of SVO-dominant participants in the baseline (0/11) and shared (4/11) groups ($p < .05$).

Combined analysis of Experiments 1 and 2 (group level). Finally, we combined the results from both experiments using a $2 \times 2 \times 3$ ANOVA with reversibility (no, yes) as a within-subjects factor, and native language (English, Turkish) and group (baseline, private, shared) as between-subjects factors.

With percent SOV as the dependent measure, there was a main effect of native language [$F(1,63) = 37.29, p < .001$], indicating that Turkish speakers used more SOV than English speakers overall. There was also a main effect of reversibility [$F(1,63) = 75.07, p < .001$], indicating that SOV was less common for reversible events overall. Importantly, native language did not interact with reversibility [$F(1,63) = .82, p = .37$] or group [$F(2,63) = 2.01, p = .14$]. No other significant effects approached significance.

With percent SVO as the dependent measure, there was a main effect of native language [$F(1,63) = 29.77, p < .001$], indicating that English speakers used more SVO than Turkish speakers overall. There was also a main effect of reversibility [$F(1,63) = 23.59, p < .001$], indicating that SVO was more common for reversible events overall.

Here, the main effect of group was significant [$F(2,63) = 6.07, p < .01$]. Planned comparisons revealed that SVO was more common in the shared group than in the baseline group [$F(1,63) = 12.04, p < .001$]. SVO was also more common in the private group than in the baseline group [$F(1,63) = 4.07, p < .05$]. Importantly, native language did not interact with group or reversibility (all F s < 1), and no other effects approached significance.

Discussion

The present data demonstrate that SVO begins to emerge for reversible events in the shared group, and to a lesser extent, for reversible events in the private group as well. Importantly, significantly more participants in the shared group used SVO to describe reversible events relative to the baseline group. Participants in all groups avoided SOV, but many of the alternative orders employed by participants in the baseline group tended to involve repetition (SOSV, SOSOV), or put O before S (OSV). However, those tendencies decreased in the private and shared groups, with both SOV and SVO increasing instead. However, whereas the increase in SOV is potentially attributable to influence from the participants' native language, the increase in SVO is not. Instead, we propose that it emerges because it uniquely satisfies the constraints against using SOV for reversible events while still being efficient and keeping S before O.

As in English speakers, the instruction to create and use a consistent gestural lexicon did not lead to reduced SOV. This is again consistent with the notion that SOV is an efficient order that keeps S before O. However, in contrast to English speakers, more participants in the private and shared groups were SOV-dominant for reversible events than in the baseline group. This pattern suggests that the instruction to create and use a

consistent gestural lexicon may indeed have encouraged some participants to use verbal recoding, which in turn led to increased use of native-language order in the private and shared groups than the baseline group. (The lack of a corresponding increase in SOV among non-reversible events may be due to a ceiling effect.) Therefore, it seems likely that at least some of the increase in SVO that we observed in Experiment 1 might be attributable to influence from the participants' native language, rather than reflecting potentially universal principles of cognition. However, the crucial question of Experiment 2 is whether we have any evidence that SVO emerges as a response to our manipulations when it cannot be attributed to influence from the participants' native language. As we have noted above, SVO does emerge when Turkish speakers describe reversible events with a self-generated gestural lexicon, an effect that cannot be attributed to the speakers' native language word order.

Although some effects were statistically marginal in omnibus analyses, the crucial tests are the planned comparisons, especially those between the baseline and shared conditions. In those cases, both the parametric and nonparametric tests found significantly more SVO in the language-like (shared) condition than in the baseline condition. Furthermore, Fisher's exact test is known to be conservative; therefore, a significant result is unlikely to reflect Type I error. Ideally, more participants would be tested; however, our sample size in each group is 50% larger than that of the original Goldin-Meadow et al. (2008) study. In addition, each hour of data collection requires roughly 4.5 to 12 hours of coding, which means that testing additional participants comes at substantial time costs.

One final aspect of the present data deserves comment. We found that native Turkish speakers avoided using SOV descriptions for reversible events, which replicates a pattern described by Hall, Mayberry, & Ferreira (submitted). The present observation is especially noteworthy because SOV is the characteristic order of Turkish participants' native language for both reversible and non-reversible events. Therefore, the pressure that drove these participants to avoid SOV must have been strong enough to outweigh the natural tendency to describe events by using the structure of one's native language. Similar findings in SOV speakers have also been observed by Gibson et al. (2011), who tested Japanese-English bilinguals, and by Meir et al. (2010), who reported preliminary data from 9 Turkish monolinguals.

General Discussion

The experiments presented here show two main points. First, we demonstrated that even native speakers of an SOV language (Turkish) avoid using SOV to describe reversible events in pantomime. This is consistent with earlier results from English speakers (Gibson et al., 2011; Hall, Mayberry, & Ferreira, submitted), as well as preliminary data from 9 Turkish monolinguals (Meir et al., 2010) and from Japanese-English bilinguals (Gibson et al., 2011). Despite giving contrasting explanations for why people avoid SOV for reversible events, these authors all agree that there is some functional motivation behind this behavior.

Second, the present experiments show that SVO may arise in part because it is an efficient way to describe reversible events while still keeping subjects before objects. In previous studies, participants often used constituent orders that were inefficient (either underinformative or repetitious) or placed objects before subjects; this happened

especially often for reversible events. We hypothesized that such orders were relatively common primarily due to the absence of other pressures that act on natural language. To test this hypothesis, we manipulated two aspects of the pantomime task. First, since a lexicon is one of the earliest language structures to emerge in new languages, we instructed some participants to create and use a gestural lexicon. Second, because natural languages arise in the context of human relationships, we instructed half of the participants to teach their gestures to the experimenter (the shared condition), while the other half performed the task alone (the private condition). We compared the constituent orders produced by the participants in each of these conditions against those produced by participants in the baseline condition, who received no special instructions (as in previous experiments). We found that both English and Turkish speakers were more likely to use SVO to describe reversible events in the shared condition than in the baseline condition, which was again largely characterized by constituent orders that were either inefficient (Regions C, D, & E of Figures 1 & 2) or mentioned objects before subjects (Regions B & D of Figures 1 & 2), or both (Region D). Although the results from the English speakers could reflect covert influence of L1, the same cannot be true of Turkish speakers. We therefore take these results as evidence that at least part of the reason that SVO emerges in the world's languages is because it allows language users to satisfy these three constraints.

We return, then, to Langus and Nespor's (2010) proposal that the constituent order distribution in the world's languages stems from two systems: a conceptual system that prefers SOV, and a computational system that prefers SVO. Under a strict interpretation of this account, our data suggest that it is incorrect, or at least incomplete.

First, there is now strong evidence that elicited pantomime does not always prefer SOV: for reversible events, participants clearly avoid SOV. Therefore, insofar as pantomime engages the conceptual system alone, this system is sensitive to more than simply the order of agents, patients, and actions. Second, the present study has also established that it is possible to account for the emergence of SVO even in the absence of the posited computational (syntactic) system. We agree with Langus and Nespors that participants' behavior in elicited pantomime tasks is not being governed by *bona fide* syntactic processing, and yet we have demonstrated that merely instructing participants to be consistent in the form of their gestures and to share them with the experimenter was sufficient to increase the frequency of participants using SVO to describe reversible events. Because participants were never exposed to other people's pantomimes, this result cannot be due to a process like creolization, which is known to give rise to SVO. Because we observed an increase in SVO even among Turkish (SOV) speakers, the effect likewise cannot be attributed merely to covert influence from the participants' native language. Thus, it would seem that at least some aspects of a strict interpretation of Langus and Nespors's model requires modification.

Relaxing some of the assumptions of Langus and Nespors (2010) allows the model to capture the data with minimal modification. In particular, the model is largely successful if it drops the assumption that the two systems are strictly segregated, and that a push for SVO is triggered by exposure to linguistic input during infancy. In that case, the model would essentially propose that, over time, the various constraints that languages face tend to be best satisfied by SVO order. Here we have considered only three potential constraints, but there are likely many others as well. For example, there

may also exist a cognitive preference for mentioning the subject before the verb, as suggested by Givón (1979). If so, that could explain why our data contain so few instances of VSO, which is the only other efficient non-SOV order that keeps subjects before objects. And indeed, perhaps it is not a coincidence that VSO is the third most common order across spoken languages.

Future research should also consider the potential role of modality (i.e., vocal or manual) in influencing constituent order preferences. This is especially critical since our two main sources of information about word order (studies of homesign/emerging sign languages and large-scale typological surveys of spoken languages) focus on different modalities. It will likely be important to characterize which pressures are invariant across modalities, and which appear to be stronger constraints on one modality than another. For example, it is possible that part of the reason that SOV is so prevalent in the manual modality is due to the spatial or embodied nature of the arguments involved. During debriefing, many participants reported that it felt most natural to establish the nominal gestures in space before acting on them. If so, that could explain why even in the shared condition, nearly half of the non-reversible responses from English-speaking participants were SOV. Future experiments could forbid the use of space to test whether SVO emerges as a preferred constituent order when physical space is not available. Likewise, one of the proposed explanations for why participants avoid SOV for reversible events relies on the observation that gesturers commonly take on the role of the agent (subject) when producing an action (verb) gesture, and are therefore motivated to avoid being in the role of the patient (object) immediately prior to producing the action gesture (Hall, Mayberry, & Ferreira, submitted). Future experiments could require participants to

describe reversible and non-reversible events in a modality that is neither speech nor gesture, such as creating a lexicon by using a slide whistle (Verhoef, Kirby, & Padden, 2011).

Conclusions

In sum, we suggest that the distribution of constituent orders across the world's languages, both synchronically and diachronically, likely reflects cognitive pressures that operate on language. Although attributing some of these pressures to an innate human language faculty may be descriptively successful, the current research sought more functional explanations. The present experiments illustrate how systematic patterns of language structure, such as the emergence of SVO, can be explained by general cognitive principles. We also suggest that elicited pantomime is a helpful empirical tool for studying cognitive influences on language structure, especially with regard to testing hypotheses derived from observations of homesign and emerging sign languages. By the same token, the generalizability of these results to spoken languages should also be considered by comparing results from this paradigm to those obtained from experiments with spoken language as well as computational simulations and other non-speech paradigms.

Chapter 3, in part, has been submitted for publication of the material as it appears in 2012, Hall, M. L., Ferreira, V. S., & Mayberry, R. I. Investigating constituent order change with the hands: Consistency promotes the emergence of SVO in pantomime. The dissertation author was the primary investigator and author of this paper.

Appendix

Agent	Action	Patient	Reversibility
boy	kiss	ball	Non-reversible
woman	kiss	ball	Non-reversible
woman	kiss	bike	Non-reversible
woman	kiss	box	Non-reversible
boy	kiss	box	Non-reversible
boy	kiss	car	Non-reversible
man	kiss	car	Non-reversible
woman	kiss	cat	Non-reversible
boy	kiss	cat	Non-reversible
man	kiss	dog	Non-reversible
girl	kiss	trike	Non-reversible
man	lift	ball	Non-reversible
girl	lift	ball	Non-reversible
woman	lift	bike	Non-reversible
boy	lift	bike	Non-reversible
woman	lift	box	Non-reversible
girl	lift	box	Non-reversible
girl	lift	car	Non-reversible
boy	lift	cat	Non-reversible
woman	lift	cat	Non-reversible
woman	lift	dog	Non-reversible
woman	pet	ball	Non-reversible
boy	pet	ball	Non-reversible
woman	pet	bike	Non-reversible
boy	pet	bike	Non-reversible
woman	pet	box	Non-reversible
girl	pet	box	Non-reversible
boy	pet	car	Non-reversible
man	pet	car	Non-reversible
woman	pet	cat	Non-reversible
boy	pet	cat	Non-reversible
boy	pet	dog	Non-reversible
man	pet	dog	Non-reversible
boy	push	ball	Non-reversible
man	push	ball	Non-reversible
woman	push	bike	Non-reversible
man	push	box	Non-reversible
boy	push	box	Non-reversible
man	push	car	Non-reversible
boy	push	car	Non-reversible
boy	push	cat	Non-reversible
woman	push	cat	Non-reversible

man	push	dog	Non-reversible
boy	push	dog	Non-reversible
girl	push	trike	Non-reversible
woman	kiss	boy	Reversible
boy	kiss	girl	Reversible
woman	kiss	man	Reversible
boy	kiss	woman	Reversible
woman	lift	boy	Reversible
boy	lift	girl	Reversible
man	lift	woman	Reversible
girl	lift	woman	Reversible
woman	pet	boy	Reversible
boy	pet	girl	Reversible
man	pet	woman	Reversible
boy	pet	woman	Reversible
girl	push	boy	Reversible
woman	push	girl	Reversible
woman	push	man	Reversible
boy	push	woman	Reversible

References

- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, 21(3), 435-483.
- Aronoff, M., Meir, I., Padden, C., & Sandler, W. (2008). The roots of linguistic organization in a new language. *Interaction Studies*, 9(1), 131-150.
- Atkinson, Q. D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, 332, 346-349.
- Bossong, G. (1991). Differential object marking in Romance and beyond. In D. Wanner & D. A. Kibbee (Eds.), *New Analyses in Romance Linguistics* (pp. 143-170). Philadelphia: John Benjamins.
- Chomsky, N. (2000). Linguistics and brain science. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 13-28). Cambridge, MA: MIT Press.
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25, 257-271.
- Dayley, J. (1985). Why all languages aren't SOV or VOS, or, how competing motivations lead to natural inconsistency. *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 52-64.
- Dryer, M. S. (1992). The Greenbergian word order correlations. *Language*, 68, 81-138.
- Dryer, M. S. (2008). Order of subject, object, and verb. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The World Atlas of Language Structures Online* (<http://wals.info>). Munich: Max Planck Digital Library.
- Derbyshire, D. S., & Pullum, G. K. (1981). Object-initial languages. *International Journal of American Linguistics*, 47, 192-214.
- Dunn, M., Greenhill, S. J., Levinson, S. C., & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473, 79-82.
- Fischer, S. (1975). Influences on word order change in American Sign Language. In C. N. Li (Ed.), *Word order and word order change* (pp. 3-25). Austin, TX: University of Texas Press.
- Gell-Mann, M. & Ruhlen, M. (2011). The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, 108, 17290-17295.

- Gibson, E., (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68, 1-76.
- Gibson, E., Brink, K., Piantadosi, S. & Saxe, R. (2011, April). *Cognitive pressures explain the dominant word orders in language*. Presented at the CUNY Conference on Human Sentence Processing, Palo Alto, CA.
- Givón, T. (1979). *On understanding grammar*. New York: Academic Press.
- Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.
- Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication in deaf children: Noneffect of parental input on language development. *Science*, 221, 372-374.
- Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, 391, 279-281.
- Goldin-Meadow, S., Özyürek, A., Sancar, B., & Mylander, C. (2008). Making language around the globe: A crosslinguistic study of homesign in the United States, China, and Turkey. In J. Guo, E. Lieven, N. Budwig, & S. Ervin-Tripp (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 27-39). New York: Psychology Press.
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 105, 9163-9168.
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language: Report of a conference held at Dobb's Ferry, NY, April 13-15, 1961* (pp. 73-113). Cambridge: MIT Press.
- Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One Mind, Two Languages* (pp. 1-22). Oxford: Blackwell Publishers.
- Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2012) *Cognitive constraints on constituent order: Evidence from elicited pantomime*. Manuscript submitted for publication.
- Haviland, J. B. (2011). Nouns, verbs, and constituents in an emerging 'Tzotzil' sign language. In R. Gutiérrez-Bravo, L. Mikkelsen, & E. Potsdam (Eds.), *Representing Language: Essays in honor of Judith Aissen* (pp. 157-171). Santa

Cruz, CA: Linguistics Research Center.

- Hawkins, J. (1983). *Word order universals*. New York: Academic Press.
- Hawkins, J. (2002). Symmetries and asymmetries: their grammar, typology, and parsing. *Theoretical Linguistics*, 28, 95-149.
- Hawkins, J. (2007). Processing typology and what psychologists need to know about it. *New Ideas in Psychology*, 25, 87-107.
- Hockett, C. F. (1958). *A course in modern linguistics*. New York: The MacMillan Co.
- Langus, A., & Nespors, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, 291-318.
- Levinson, S. C., & Gray, R. D. (2012). Tools from evolutionary biology shed new light of the diversification of languages. *Trends in Cognitive Sciences*, 16, 167-173.
- Li, C. N. (1977). *Mechanisms of syntactic change*. Austin: University of Texas Press.
- Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. A. (2010, April). *The interaction of animacy and word order in human languages: A study of strategies in a novel communication task*. Paper presented at the 8th International Conference on the Evolution of Language, Utrecht, Germany.
- Newmeyer, F. J. (2000). On the reconstruction of "proto-world" word order. In C. Knight, M. Studdardt-Kennedy, & J. R. Hurford, (Eds.), *The Evolutionary Emergence of Language* (pp. 372-388). New York: Cambridge University Press.
- Newmeyer, F. J. (2002). The SO/OS asymmetry and two interpretations of the MaOP: commentary on John A. Hawkins, 'symmetries and asymmetries: their grammar, typology, and parsing'. *Theoretical Linguistics*, 28, 171-176.
- Parker, F. (1980). Typology and word order change. *Linguistics*, 18, 269-288.
- Pullum, G. K. (1981). Languages with object before subject: a comment and catalogue. *Linguistics*, 19, 147-155.
- Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, 102, 2661-2665.
- Sapir, E. (1921). *Language: an introduction to the study of speech*. New York: Harcourt Brace and Company.

- Senghas, A., Coppola, M., Newport, E. L., & Supalla, T. (1997). Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the Boston University Conference on Language Development*, Vol 21 (pp. 550-561). Boston: Cascadia Press.
- Singleton, J. L., Goldin-Meadow, S., & McNeil, D. (1995). The cataclysmic break between gesticulation and sign: Evidence against a unified continuum of gestural communication. In K. Emmorey & J. Reilly (Eds.), *Language, gesture, and space* (pp. 287-311). Hillsdale, NJ: Erlbaum Associates.
- Sinnemäki, K. (2010). Word order in zero-marking languages. *Studies in Language*, 34, 869-912.
- Slobin, D. I. (1977). Language change in childhood and history. In J. Macnamara (Ed.), *Language learning and thought* (pp. 185-214). New York: Academic Press.
- Tomlin, R. S. (1986). *Basic word order: functional principles*. London: Croom Helm.
- Vennemann, T. (1973). Explanation in syntax. In J. P. Kimball (Ed.), *Syntax and Semantics, Vol. 2* (pp. 1-50). New York: Seminar Press.
- Vennemann, T. (1974). Topics, subjects, and word order: from SXV to SVX via TVX. In J. M. Anderson & C. Jones (Eds.), *Historical Linguistics, Vol 1* (pp. 339-376.)
- Vennemann, T. (1975). An explanation of drift. In C. N. Li (Ed.), *Word order and word order change* (pp. 269-305).
- Verhoef, T., Kirby, S., & Padden, C. (2011). Cultural emergence of combinatorial structure in an artificial whistled language. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 483-488). Austin, TX: Cognitive Science Society.

Author note

The authors gratefully acknowledge members of the Laboratory for Multimodal Language Development, the Language Production Laboratory, the Sign Language Research Group, and the Center for Research in Language at UC San Diego for helpful feedback. We are especially grateful for the partnership of Aylin Kuntay and Ecem Çoban at Koç Üniversitesi. This research would not have been possible without assistance with data collection and coding from Caitlin Scott, Shereen Cohen, Isabel Sharman, Kenny Oyama, Alex Kuo, Elena Churilov, Jayd Blankenship, Kristi Cheng, Lena Sun, and Youngju (Danbi) Ahn. Funding for this research was provided by NIH grant HD051030, NIH grant 5T32DC000041-19, and by the Rita L. Atkinson Graduate Fellowship.

CHAPTER 4

COMPREHENDERS PREFER SVO, BUT PRODUCERS PREFER SOV: EVIDENCE FROM ELICITED PANTOMIME

Abstract

To what extent can psycholinguistic processing preferences account for variation in constituent order (both across languages, and within a language over time)? Previous studies have used elicited pantomime production to gain insight into cognitive preferences for constituent order, but have largely overlooked the possibility that constituent order preferences in comprehension might be different. We therefore developed a novel paradigm to measure constituent order preferences in pantomime comprehension. Participants saw a video of a transitive event (e.g. a woman pushing a box), followed by a video of a confederate describing that event in pantomime, using either Subject-Object-Verb (SOV), Subject-Verb-Object (SVO), or Object-Subject-Verb (OSV) order. Participants rated the correspondence between the pantomime clip and the event clip. Relative to previous studies of pantomime production, participants showed an increased preference for SVO, and less sensitivity to whether events were reversible (e.g. a woman pushing a boy) or non-reversible (e.g. a woman pushing a box). These patterns were stable across speakers of English (an SVO language) and Turkish (an SOV language). These results suggest that the constituent orders of the world's languages are shaped in part by the dynamic interplay of production and comprehension.

The primary goals of the language sciences are to describe and explain patterns in natural human languages. One type of pattern that has received much attention is *constituent order* (sometimes called “word order”). Constituent order is most frequently discussed with respect to *transitive events*: those in which an agent performs some action on a patient. In most languages, agents are canonically expressed as subjects (S), actions as verbs (V), and patients as objects (O). For example, the canonical constituent order of English sentences is SVO, as in, “The woman lifts the box”. Other languages have different canonical orders; for example, Turkish uses SOV order. Indeed, all six logically possible orderings of these three elements are attested in the world’s languages; however, their synchronic distribution is highly uneven. SOV and SVO together account for roughly 75% of the world’s languages, with VSO in a distant third (roughly 10%), while VOS, OVS, and OSV together account for the remaining 15% (e.g. Dryer, 2008; Greenberg, 1963; Hawkins, 1983; Tomlin, 1986). The exact figures vary between typological surveys, but there is very little disagreement about the relative ordering.

The unevenness of this distribution is one of two major observations that have intrigued language researchers since the seminal work of Greenberg (1963). The second intrigue is that when constituent order changes over time, languages tend to change away from SOV but not toward it. Languages do change toward SVO, which tends to be more stable (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977; Vennemann 1974, 1975). The paths of constituent order change across time are the result of many interacting factors. At this point, it is important to distinguish between two general categories of influence: those that are language-external (e.g. migration, conquest, geography, contact) and those that are language-internal (e.g. would be likely to occur even in a stable, closed,

monolingual community, given enough time). In the present studies, our goal is to gain a deeper understanding of language-internal factors. We assume that when diverse languages change in similar ways that cannot be explained through language-external forces, they do so in response to functional pressures from within language: that is, from the interface of communication and cognition (Beckner et al., 2009). The present experiments aim to test this assumption and to provide more information about what those functional pressures may or may not be.

The concept of functional pressures shaping language structure is by no means new. Some researchers have focused on pressures that are claimed to be internal to grammar, such as drives toward “consistent” or “harmonic” orderings of phrasal heads and complements, which may explain observations such as OV languages tending to use postpositions while VO languages tend to use prepositions (Dryer 1992; Greenberg, 1963; Lehmann 1973; Vennemann, 1973, 1974). Other researchers have emphasized how grammar is shaped by discourse (DuBois 1985, 2003), pragmatics (Givón, 1979), memory (Gibson, 1998), child language acquisition (Kirby, 1999), processing efficiency (Hawkins 1994, 2002, 2004; Newmeyer 2002), and information theory (Maurits, Perfors, & Navarro, 2010). Still others point out that, for the most part, these pressures are not mutually exclusive, and that many different types of functional pressures are involved in shaping language structure (Beckner et al., 2009; Bybee, 2009; Dayley, 1985; Malkiel, 1968). Studies in this tradition typically analyze the frequency of particular language structures (e.g. constituent order), either across many different languages or within a given language across time, or they measure processing efficiency within a given language whose grammar permits several options. In all cases, researchers draw their

conclusions from examining the extant (and sometimes past) grammars of natural human languages.

The aim of the present studies is not to advocate for or against any of the above accounts, but rather to propose an additional dimension of language processing that might influence constituent order distributions and change: namely, the interplay between how production and comprehension prefer to represent and process elements in dynamic events. It has long been recognized that there are tradeoffs between what is easier to produce and what is easier to comprehend (Givón, 1979; Ladefoged, 1982; Jaeger & Tily, 2011; Slobin, 1980; Zipf, 1949), such that the grammars of natural languages come to represent some type of compromise between the two (Ferreira, 2008). In addition, even when the demands of production and comprehension are compatible, it is possible to identify situations where linguistic behavior is driven by one system or the other (e.g. production-centered behavior in Dell & Brown, 1991). However, as Haspelmath (2002) points out, conspicuously little has been said about how the potentially divergent processing demands of production and comprehension contribute to attested patterns of constituent order distribution and change. Could it be that some constituent orders are preferred in production while others are preferred in comprehension? The present experiments address this question.

Because natural human languages involve both production and comprehension, simply looking at the grammars of natural languages will not be sufficient to reveal influences that are specific to production or specific to comprehension. What we want to know is, if left to its own devices, what constituent order would the production system prefer? And likewise, if left to its own devices, what constituent order would the

comprehension system prefer? It may be that the two systems have concordant preferences. This possibility seems intuitively advantageous, from both a processing and evolutionary perspective. But it is also possible that production and comprehension could have discordant preferences for constituent order. That is, that the information that a comprehender wants to receive first might not be the same information that a producer wants to provide first.

Natural languages emerge and are acquired with people interchangeably playing the role of producer and comprehender, but there is at least one situation where a communicative system emerges in which its creator is almost always a producer and rarely a comprehender: homesign. A homesign system is the creation of a deaf child who lacks access to both spoken and sign language. In the absence of language input, children in this situation reliably invent their own systems of communication, and these systems contain a surprising number –though not all– of the features that characterize human language (for a review, see Goldin-Meadow et al., 2003). Goldin-Meadow and colleagues have shown that hearing parents' gestures cannot account for the salient features of homesign (Goldin-Meadow & Mylander, 1983). The structure of homesign, then, offers clues as to the preferred constituent order of the production system. A robust finding is that constituent order in homesign tends to be analogous to either SV or OV, and not VS or VO. This pattern is stable across individual homesigners in several countries, cultures, and spoken language environments (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, Özyürek, Sancar, & Mylander, 2008), Goldin-Meadow and colleagues have pointed out that SV and OV are consistent with SOV order, but not with SVO order. This observation serves as an initial window into the constituent order

preferences of the production system, but it is not an unobstructed view. Although homesigners themselves are rarely in the role of comprehender, they do have interlocutors who try to understand their utterances, and it is possible that they have learned through trial and error that SV and OV order are what best suit their comprehenders. To more directly test the preferences of the production system alone, it would be useful to examine a situation in which only production is involved, but which is also unconstrained by the rules of any particular language's grammar.

Goldin-Meadow and her colleagues have developed a paradigm that has these exact properties: elicited pantomime. In this paradigm, hearing people with no knowledge of a sign language are shown video clips of transitive events (e.g. a woman lifting a box), and are asked to describe those events in pantomime: that is, gesture without speech. Their descriptions are then coded for the relative order in which they mention agents, actions, and patients. (Because natural languages typically encode these as subjects, verbs, and objects, respectively, we will use the latter nomenclature when referring to participants' gestures, and will use the terms agent, patient, and action to describe the semantic content of events). Because the elicited pantomime task involves only production, and participants do not communicate with any addressee, this task offers a wider window into the constituent order preferences of the production system. Still, any participants in this task will already speak another language, which introduces the possibility that their pantomime behavior may be influenced by the structure of their native language, rather than being a pure measure of the production system's constituent order preferences. This concern is mitigated by the results first observed by Goldin-Meadow et al. (2008): namely, that participants show a consistent preference for using

SV, OV, and SOV order to describe canonical transitive events (such as a woman lifting a box) in pantomime, regardless of the constituent order of their native language. The initial study tested native speakers of English, Turkish, Spanish, and Chinese, and the SOV preference has subsequently been replicated in speakers of Italian and Turkish (Langus & Nespors, 2010), Hebrew and Turkish (Meir, Lifshitz, Padden, & Ilkbasaran, 2010), English and Japanese (Gibson, Piantadosi, Brink, & Saxe, 2011), English (Hall, Mayberry, & Ferreira, submitted), and English and Turkish (Hall, Ferreira, & Mayberry, submitted). Taken together, these results suggest that, left to its own devices, the production system prefers to describe canonical transitive events in SOV order.

The question of constituent order preferences in comprehension has received far less attention. In fact, only one prior study has attempted to address this question directly. Langus and Nespors (2010) presented participants with scene descriptions in all six logically possible constituent orders, and used a two-alternative forced-choice paradigm to assess how quickly participants could identify which picture corresponded to the description. They tested native speakers of Italian (SVO) and Turkish (SOV) on two types of stimuli: those constructed from gestures, and those constructed from spoken words (in the appropriate language, but without any grammatical marking or function words). Their primary motivation was to determine whether constituent order preferences differed for stimuli that were putatively linguistic (words) versus non-linguistic (gestures). However, their results are also informative about constituent order preferences in comprehension. In gesture comprehension, they found faster reaction times for OV orders (SOV, OSV, OVS) than for VO orders (SVO, VSO, VOS), with SOV eliciting the fastest responses. Importantly, this pattern held true for both Italian

and Turkish speakers. At first glance, this seems to suggest that the comprehension system might prefer SOV structure just like the production system. However, data from the word comprehension condition showed a different pattern. There they found faster reaction times for VO orders (SVO, VSO, VOS) than for OV orders (SOV, OSV, OVS). Although both Italian and Turkish speakers showed this pattern when averaging across the three orders in each category, Italian (SVO) speakers were fastest for SVO sequences and Turkish (SOV) speakers were fastest for SOV. This discrepancy between the word and gesture conditions could indicate fundamental differences in how people process words and gestures (as argued by Langus & Nespore), or it could indicate unforeseen confounds in the experimental design. At present, it remains unclear whether the gesture condition or the word condition provides a clearer window in constituent order preferences in production.

One drawback of using a two-alternative forced-choice paradigm is that the participants' task essentially reduces to identifying the single word or gesture that is incongruent with one of the response alternatives. For example, if a given trial displays a picture of a woman lifting a box and a picture of a woman pushing a box, the participant need only attend to whether the spoken or gestured stimulus includes lifting or pushing. Reaction times in these studies therefore indicate how quickly participants can discriminate correct from incorrect descriptions (e.g. WOMAN PUSH BOX vs. WOMAN LIFT BOX). By assessing participants' responses to differences in semantic content, the task may have introduced strategies or biases that confounds the comparison of constituent orders. Instead, the chief question of interest concerns processing for different constituent orders that have all of the same semantic content (e.g. WOMAN PUSH BOX vs.

WOMAN BOX PUSH). To address this question, we developed a novel experimental paradigm (described below) in which we contrast pantomime comprehension for stimulus sequences that varied only in constituent order, keeping semantic content the same. Because Langus and Nespors (2010) found that using words as stimuli led to more language-specific patterns of behavior, we used gestures as stimuli.

A second limitation of previous research –in both production and comprehension– is that both Goldin-Meadow et al. (2008) and Langus and Nespors (2010) focused only on events where semantics alone makes it possible to determine agent versus patient. For example, given any ordering of a woman, a box, and pushing, it is clear that the woman is the agent and the box is the patient. In contrast, natural languages need to communicate about not only these events, which we term *non-reversible* (because the agent and patient referents cannot plausibly be reversed), but also events with a human agent and a human patient (e.g. a woman pushing a boy). We refer to these as *reversible* events. Several previous studies of pantomime production have found that constituent order preferences for reversible events are markedly different from those for non-reversible events. However, those studies reached different conclusions as to the source of these differences, with some arguing that producers' behavior was designed for the benefit of a hypothetical comprehender (Gibson et al., 2011; Meir et al., 2010), and others arguing that their behavior was primarily due to constraints on production (Hall, Mayberry, & Ferreira, submitted). Because these observations played a key role in motivating the current study, we describe these studies in greater detail.

Common to all three studies was the general methodology of asking hearing participants with no knowledge of a sign language to describe transitive events in

pantomime: that is, gesture without speech. In addition, all three studies compared constituent order preferences in production for non-reversible events versus reversible events. For non-reversible events, all three studies observed that SOV sequences (including SV and OV) were the most common across speakers of all languages tested, thus replicating the central result of Goldin-Meadow et al. (2008). However, all three studies also found that participants described reversible events differently. Meir et al. (2010) reported that their Hebrew speakers replaced SOV with SVO, while their Turkish speakers primarily replaced SOV with OSV. Gibson et al. (2011) also reported that as SOV decreased, SVO increased, but did not report whether their English or Japanese speakers also used OSV. The English speakers tested by Hall, Mayberry, and Ferreira (submitted) generally replaced SOV with one of three solutions: SVO, OSV, or SOSV. Thus, the key observation that holds across all three studies is that participants avoided using SOV to describe reversible events in pantomime. Where the studies differ is in their interpretation of the source of this effect.

Both Meir et al. (2010) and Gibson et al. (2011) attributed their participants' SOV-avoidance to the notion that for reversible events, having both human nominal gestures precede (or follow) an action gesture would be potentially confusable to an addressee. Thus, these accounts rely on the idea that producers' mental models of comprehenders' preferences might account for the data. (Although there were no addressees in these experiments, participants may still have behaved as though they were communicating.) In contrast, Hall, Mayberry, and Ferreira (submitted) point out that if SOV events were confusable, then OSV events should also be confusable, because both involve sequences of PERSON-PERSON-ACTION. The same is true for SOSV, which

involves PERSON-PERSON-PERSON-ACTION. However, OSV and SOSV became more common for reversible events, not less, as confusability-based accounts would predict. These observations led Hall, Mayberry, and Ferreira (submitted) to propose an account that emphasized constraints on production. This account was based on a generalization they observed in the data: namely, that for reversible events, verb gestures were almost always preceded by subject gestures. According to this account, producers' constituent order preferences stem from their self-imposed constraint to embody the role of the agent at the time that they produced the action, so as to avoid a *role conflict* (for further discussion, see Hall, Mayberry, & Ferreira, submitted). SOV does not satisfy that constraint, but SVO, OSV, and SOSV all do.

A significant weakness of both the comprehension-centered accounts (minimize confusability) and the production-centered account (avoid role conflict) is that they make assumptions about what would or would not be beneficial to a comprehender without actually testing constituent order preferences in pantomime comprehension. The primary goal of the present study was to address this gap. In so doing, we aimed to test (a) whether Langus and Nespors's (2010) observation of an SOV preference in pantomime comprehension for non-reversible events would generalize to a new paradigm, and (b) how comprehenders' constituent order preferences for reversible events compare to those in production. In the service of this second goal, we focus on four key aspects of constituent order preferences in pantomime production, listed below. These patterns, which were the most stable across previous studies of pantomime production, will serve as a benchmark for assessing the extent to which constituent order preferences in pantomime comprehension are concordant or discordant with those in production.

1. For non-reversible events, producers used SOV far more than any other order (Gibson et al., 2011; Goldin-Meadow et al., 2008; Hall, Mayberry, & Ferreira, submitted; Langus & Nespors, 2010; Meir et al., 2010).

2. For reversible events, producers used SOV far less than for non-reversible events (Gibson et al., 2011; Hall, Ferreira, & Mayberry, submitted; Hall, Mayberry, & Ferreira, submitted; Meir et al., 2010).

3. Producers used SVO more for reversible than for non-reversible events. Importantly, this was true not only for SVO speakers (English: Gibson et al., (2011); Hall, Mayberry, & Ferreira (submitted); Hebrew: Meir et al. (2010)), but also for SOV speakers (Turkish: Hall, Ferreira, & Mayberry (submitted)).

4. OSV was not commonly used for non-reversible events, but became more frequent for reversible events (English: Hall, Ferreira, & Mayberry, submitted; English & Turkish: Hall, Mayberry, & Ferreira, submitted; Hebrew & Turkish: Meir et al., 2010).

To the extent that pantomime comprehension also reveals these four patterns, constituent order preferences in comprehension and production can be considered concordant.

To measure constituent order preferences in pantomime comprehension, we developed a new experimental paradigm that would allow us to contrast preferences for stimuli that varied only in constituent order, keeping semantic content the same. In this paradigm, described below, participants first saw a clip of a transitive event, and then saw a confederate describing that event in pantomime, using either SOV, SVO, or OSV order. The participants' job was to make likelihood judgments about the pantomime clips in relation to the event that had been shown.

Consider first a non-reversible event where, for example, a woman pushes a box. That event would, over the course of the experiment, be paired with the gesture sequences WOMAN BOX PUSH (SOV), WOMAN PUSH BOX (SVO), and BOX WOMAN PUSH (OSV). (On filler or “catch” trials, it would also be paired with a semantically incorrect gesture sequence, such as WOMAN BOX LIFT or BOY PUSH BOX.) We reasoned that if production and comprehension are sensitive to the same sets of cognitive biases, then comprehenders should give higher likelihood ratings to constituent orders that were more common in pantomime production. For example, they should give high ratings to SOV pantomime clips for non-reversible events, because that was the most common order in production.

The same logic applies to reversible events. For example, imagine now that the event clip shows a woman pushing a boy, and the pantomime clip contained the gesture sequence WOMAN BOY PUSH (SOV). We can then contrast participants’ responses to SOV descriptions of reversible events with their responses to SOV descriptions of non-reversible events. Again, if constituent order preferences in comprehension are concordant with those in production, we would expect participants to give lower ratings to descriptions of reversible events relative to non-reversible events.

Characterizing participants’ behavior in this task will therefore allow us to assess the extent to which constituent order preferences are stable across comprehension and production. If so, it would suggest that dynamic interplay between production and comprehension is unlikely to be responsible for the typological distribution of constituent orders across the world’s languages or for the direction of diachronic change from SOV to SVO. On the other hand, if constituent order preferences in comprehension differ from

those in production, the nature of those differences could provide insight into which aspects of typology and diachronic change are more influenced by production versus comprehension.

Experiment 1

Method

Participants. We tested 16 undergraduate students at UC San Diego who reported being monolingual native speakers of English with no knowledge of any sign language (familiarity with letters or numbers was acceptable). Informed consent was obtained from all participants, who received course credit for their participation.

Materials. Stimuli consisted of two types of movies: event clips and pantomime clips. Event clips showed a human agent (man, woman, boy, girl) performing a transitive action (kiss, lift, pet, push) on a patient that was either non-human (ball, bike, box, cat) or human (man, woman, boy, girl). Events with a non-human patient were considered non-reversible. Events with a human patient were considered reversible. In reversible events, the agent and patient always differed in age, sex, or both. These stimuli were a subset of those used in our previous studies (Hall, Ferreira, & Mayberry, submitted; Hall, Mayberry, & Ferreira, submitted). There were 32 unique event clips: 16 reversible and 16 non-reversible. The stimuli are listed in the Appendix.

Pantomime clips showed a confederate producing a controlled vocabulary of gestures in either SOV, SVO, or OSV order, without using space to indicate the location or direction of agents, patients, or actions. Although this may compromise the naturalness of the stimuli, it is justified by the gains in experimental control. It is similar to the approach of Langus and Nespors (2010), except that we used connected gestures

instead of filming each gesture in isolation and combining them digitally. This helped make the stimuli slightly more naturalistic, and was also necessary for our cover story. The gestures for each item were based on the most common gestures used by participants in our previous experiments on pantomime production. To avoid encouraging participants to use verbal encoding, we did not train the participants on the meaning of the gestures, relying instead on the gestures' iconicity to yield a clear indication of their referent. An experimenter monitored the participant's performance during practice trials to verify that they gave higher ratings to trials with correct semantic content than to trials with incorrect content. This indicated that inferring the referents of the gestures themselves was not a primary source of confusion.

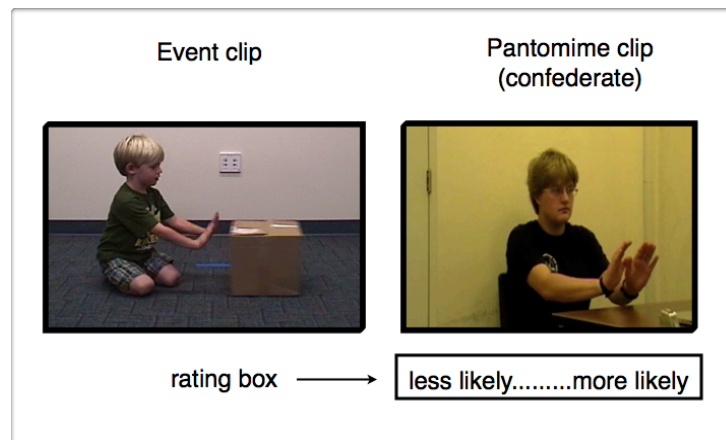


Figure 1. Schematic illustration of the experimental paradigm. Event clips and pantomime clips were not visible at the same time. Labels are included for illustration only; the only text displayed to participants was “less likely... more likely”, which was translated as “cok iliskisiz ... cok iliski” for Experiment 2.

The stimuli were presented on a 15-inch Macintosh laptop using PsyScope X software (Cohen, MacWhinney, Flatt, & Provost, 1993). Figure 1 provides a schematic view of the laptop screen during a typical trial, with the event clip on the left and the

pantomime clip on the right. A trial began with an event clip, which participants were free to replay as often as desired before viewing the pantomime clip. When ready, participants pressed the spacebar, which triggered the pantomime clip. At this point, participants could no longer replay the event clip, and only the pantomime clip was visible. Participants could replay the pantomime clip as often as they wished before responding. Participants' responses triggered the beginning of the next trial. (Our initial design had presented the pantomime clip first, followed by the event clip, but pilot subjects reported that they couldn't remember the content or order of the pantomime clip without covertly or overtly vocalizing it in words. Because we wanted to minimize the extent to which participants would rely on verbal coding using their native language, we presented the event clips first, which are episodically rich enough to be remembered without assistance from verbal coding. When tested with the event clip first, pilot subjects did not report relying on verbal coding to remember the event clip.)

Design and procedure. We used a 3 x 2 factorial design with constituent order (SVO, SOV, OSV) and reversibility (non-reversible, reversible) as within-subjects factors. This yielded 96 critical trials (16 event clips x 3 constituent orders x 2 levels of reversibility). These trials gave participants an opportunity to provide judgments about constituent order when all of the semantic content of the utterance was correct. However, we also paired each of the 16 non-reversible event clips with a pantomime clip that contained an incorrect gesture. (The incorrect gestures were distributed as evenly as possible across agents, patients, and actions as well as ordinal position in the utterance.) We analyzed these trials separately from those in which the semantic content was correct.

Participants were told that they would make judgments about the behavior of a participant in a previous experiment in our lab. An experimenter explained that in a previous experiment, we showed the event clips to other naïve participants and asked them to describe those events in pantomime. The current participant's job was to make judgments about those pantomimed descriptions. Specifically, current participants were asked to judge how likely it was that a given pantomime clip is what the "previous participant" (who was in reality a confederate) actually produced to describe the event clip that it was paired with. To increase the plausibility of the cover story, we also told the current participants that the previous participant had been asked to describe events that would not be presented to the current participant, and that they (the current participant) would also see events that the previous participant had not been asked to describe.

Participants made their judgments by using the computer's mouse to click in a rating box on the screen (see Figure 1), where the left side of the rating box indicated "less likely" judgments and the right side of the box indicated "more likely". Rather than imposing a categorical or ordinal scale on what could be a continuous measure, we instructed participants to click anywhere in the box, allowing their judgments to be as gradient as possible. Our dependent measure was the horizontal position of their mouseclick within the box. These pixel values were then standardized such that the extreme left pixel (lowest possible likelihood rating) was 0 and the extreme right pixel (highest possible likelihood rating) was 100.

After the cover story and task instructions, participants began 8 practice trials. The first four varied in constituent order but had all of the correct semantic content,

which introduced the participants to the gestural vocabulary. The second four contained incorrect semantic content, for which the correct gesture had already been presented. An experimenter monitored each participant's performance to verify that their judgments on the four incorrect trials were in the "unlikely" response region, and explained the task again if it appeared that the participant did not fully understand the task or the meaning of the gestures. No feedback was given as a function of participants' responses to different constituent orders, and very few participants needed any feedback after practice.

Following practice trials, the experimenter left the room. The testing session proceeded in two blocks, with a break halfway through. The only difference between the blocks was that for a given subject, half of the event clips appeared before the break and half after. Which clips appeared in which half was counterbalanced across subjects. Within each half, trials were presented in one of two fixed random orders (forward or backward). Reversible and non-reversible events were mixed throughout the testing session. After test trials, participants were shown each gesture in isolation and asked to provide a one-word translation for the meaning of that gesture, to verify that comprehension was ultimately successful. Upon completion, an experimenter asked the participant several debriefing questions. Most participants guessed that we were interested in something about constituent order. Some also reported that they gave higher ratings to orders that were easier for them to understand. Because those judgments are still informative about constituent order preferences in pantomime comprehension, we did not exclude those participants' data from analysis.

Results

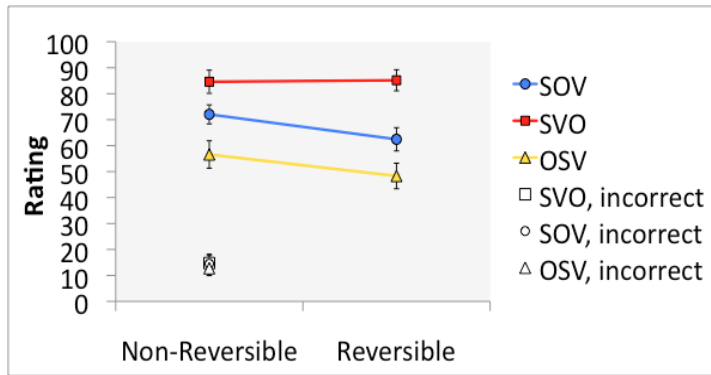


Figure 2. Results of Experiment 1, which tested native English speakers. Error bars represent SEM.

Filler trials (semantically incorrect). When the pantomime clips contained gestures that were not semantically appropriate for the event clip, participants gave low likelihood ratings (see Figure 2), which were not influenced by constituent order [$F(2,30) = .6, p = .56$].

Constituent order trials (semantically correct). When all of the gestures in the pantomime clip were semantically appropriate for the event clip, likelihood judgments were sensitive to both constituent order and reversibility. These data are displayed in Figure 2. A 3 x 2 repeated-measures ANOVA revealed a main effect of constituent order [$F(2,30) = 21.09, p < .001$] and a main effect of reversibility [$F(1,30) = 10.82, p < .01$]. There was no constituent order x reversibility interaction [$F(2,30) = 1.19, p < .32$]. We used planned contrasts to test the four specific predictions that we derived from previous research on constituent order preferences in pantomime production.

First, the concordance hypothesis predicts that SOV should receive the highest likelihood ratings for non-reversible events. This was not the case. Rather, ratings for SOV (72.0) were significantly lower than those for SVO (84.6: $F(1,30) = 6.15, p < .02$).

Second, the concordance hypothesis predicts that SOV should receive significantly lower ratings for reversible than for non-reversible events. Although the trend was in the predicted direction (72.0 for non-reversible vs. 62.4 for reversible), this difference was only marginally significant [$F(1,30) = 3.60, p = .07$].

Third, the concordance hypothesis predicts that SVO should receive higher ratings for reversible than for non-reversible events. Instead, the ratings for SVO did not differ between reversible (85.2) and non-reversible events (84.6: $F(1,30) = .01, p = .91$).

Fourth, the concordance hypothesis predicts that ratings for OSV should either remain constant or increase for reversible as compared to non-reversible events. Results here were somewhat equivocal; there was no statistical difference in the ratings for OSV orders [$F(1,30) = 2.67, p < .11$], but the numerical difference went in the opposite of the predicted direction: 56.58 for non-reversible versus 48.30 for reversible.

Discussion

The results of Experiment 1 illustrate the importance of measuring constituent order preferences when semantic content is held constant. On such trials, we found a very different pattern of constituent order preferences in comprehension than has been previously found in studies of pantomime production (Gibson et al., 2011; Goldin-Meadow et al., 2008; Hall, Ferreira, & Mayberry, submitted; Hall, Mayberry, & Ferreira, submitted; Langus & Nespors, 2010; Meir et al. 2010).

All six of these studies found that for non-reversible events, SOV was the most common order used in pantomime production, regardless of whether the participants spoke English, Turkish, Spanish, Italian, Chinese, Japanese, or Hebrew. In contrast, the

(English-speaking) participants in our comprehension task gave the highest likelihood ratings to SVO.

Four of these studies (Gibson et al., 2011; Hall, Ferreira, & Mayberry, submitted; Hall, Mayberry, & Ferreira, submitted; Meir et al. 2010) tested both reversible and non-reversible events. All four found that participants robustly avoided using SOV to describe reversible events. In contrast, the current results for pantomime comprehension found only a marginally significant trend toward lower likelihood ratings for SOV in reversible events. This could be easily attributed to the main effect of reversibility, where all reversible events received lower ratings.

Increases in SVO among reversible events for pantomime production were reported by Gibson et al. (2011), Hall, Mayberry, and Ferreira (submitted), and Hall, Ferreira, and Mayberry (submitted). In contrast, no such increase was observed in the current data from pantomime comprehension. This is unlikely to be due to a ceiling effect, because the means for both the non-reversible condition (84.6) and the reversible condition (84.7) were well below the maximum value of 100.

The fourth study, Meir et al. (2010), reported an increase in OSV for reversible events in pantomime production, a pattern that was also evident in Hall, Mayberry, and Ferreira (submitted) as well as Hall, Ferreira, and Mayberry (submitted). Yet again, we failed to observe this behavior in the current data from pantomime comprehension. If anything, the data seemed to indicate the opposite pattern: numerically, OSV descriptions were rated as being less likely for reversible than for non-reversible events, although this difference was not statistically significant. It is interesting to note that semantic reversibility affects SOV and OSV similarly in comprehension (a marginal decreases for

reversible events with no interaction [$F(1,15) = .02, p = .89$]), whereas reversibility affected SOV and OSV differently in production (i.e. large decrease for SOV, small increase for OSV).

Experimental design and statistical non-independence prevent us from testing for a three-way interaction (constituent order x reversibility x comprehension/production), but the convincing lack of a two way interaction (constituent x reversibility) is unexpected under the hypothesis that production and comprehension have similar preferences. Thus, this pattern of results suggests that constituent order preferences may diverge between production and comprehension.

In particular, SVO seems to be especially favored in comprehension, even for non-reversible events. It is possible that SVO order confers particular advantages to comprehenders, and that this may in turn contribute to diachronic shifts from SOV toward SVO. However, this aspect of the results is also compatible with an alternative hypothesis: namely, that participants gave high likelihood ratings to pantomime clips whose constituent order matched their native language, in this case SVO for English speakers. To test this hypothesis, we conducted Experiment 2, which replicates Experiment 1 but tests Turkish speakers, whose native language uses SOV order. If the patterns we observed in Experiment 1 reflect a native language bias, participants in Experiment 2 should give the highest ratings to SOV for both non-reversible and reversible events.

Experiment 2

Method

Participants. All testing was performed by native Turkish speakers, who recruited and tested 16 participants in Sariyer and Istanbul, Turkey. Of these, experimenter error resulted in the data from one participant not being recorded. The final dataset comes from 15 participants, one of whom did not complete the last 21 trials (19%) due to computer error. An ideal participant would have no contact with or knowledge of any SVO language. Since that is highly unlikely, we asked all potential participants about their experience with other languages prior to selection. Potential participants were excluded if an SVO language was spoken in their home. We also asked potential participants to use the multiple-choice scale in Table 1 to self-report their level of proficiency in all the languages they knew. Potential participants were excluded if they reported “3” or above in any SVO language. All participants gave consent to be videotaped as part of the study, and were paid for their participation. Informed consent was obtained from all participants, who received payment for their participation.

Table 1: Language screening questions for Experiment 2.

Language screening questions
5 = I am comfortable using this language in all situations.
4 = I am mostly comfortable using this language, but I prefer another one in some situations.
3 = I can use this language when necessary, but prefer not to.
2 = I can usually make myself understood, but it’s difficult.
1 = I would only use this language if I had no other options.
0 = I don’t remember any of this language.

Materials. We used the same materials as in Experiment 1.

Design and procedure. The design and procedure were identical to Experiment 1 except that the spoken and written instructions were delivered in Turkish instead of English.

Results

Filler trials (semantically incorrect). When the pantomime clips contained gestures that were not semantically appropriate for the event clip, participants gave low likelihood ratings which were not influenced by constituent order [$F(2,28) = .27, p = .77$], as displayed in Figure 3.

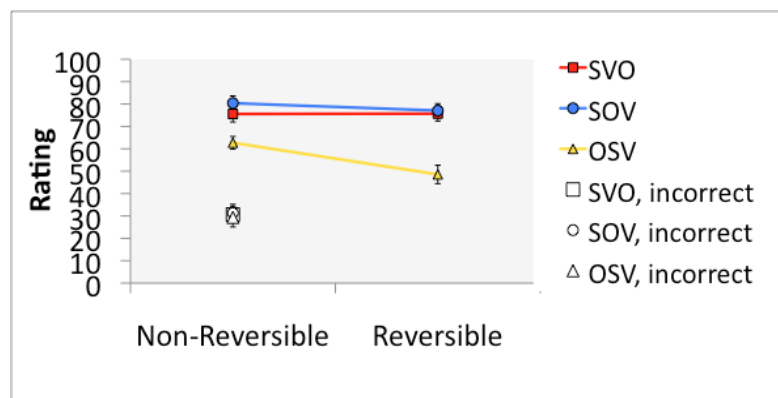


Figure 3. Results of Experiment 3, which tested Turkish speakers. Error bars represent SEM.

Constituent order trials (semantically correct). As in Experiment 1, likelihood judgments were sensitive to both constituent order and reversibility when all of the gestures in the pantomime clip were semantically appropriate for the event clip. These data are displayed in Figure 3. A 3 x 2 repeated-measures ANOVA revealed a main effect of constituent order [$F(2,28) = 29.61, p < .001$], a main effect of reversibility [$F(1,28) = 28.80, p < .001$], and a constituent order x reversibility interaction [$F(2,28) =$

10.74, $p < .001$]. We then used planned contrasts to test the four specific predictions that we derived from previous research on constituent order preferences in pantomime production.

First, the concordance hypothesis predicts that SOV should receive the highest likelihood ratings for non-reversible events. In fact, SOV did receive higher ratings (80.4) than SVO (75.6) for non-reversible events [$F(1,28) = 4.5$, $p < .05$]. Thus, this prediction did find statistical support, but the effect was not large.

Second, the concordance hypothesis predicts that SOV should receive lower ratings for reversible than for non-reversible events. Again, although the trend was in the predicted direction (80.4 for non-reversible vs. 77.1 for reversible), this difference did not reach significance [$F(1,28) = 2.12$, $p = .16$].

Third, the concordance hypothesis predicts that SVO should receive higher ratings for reversible than for non-reversible events. Instead, the ratings for SVO did not differ between reversible and non-reversible events (75.7 vs. 75.6, respectively: $F(1,28) = .01$, $p = .94$).

Fourth, the concordance hypothesis predicts that ratings for OSV should either remain constant or increase for reversible as compared to non-reversible events. This too was not the case. Participants gave significantly lower ratings to OSV orders for reversible as compared to non-reversible events (48.6 vs. 62.7, respectively: $F(1,28) = 38.58$, $p < .001$). It is clear from Figure 3 that this change in OSV ratings is primarily responsible for the constituent order x reversibility interaction in the omnibus analysis.

Discussion

The primary motivation for conducting Experiment 2 was to test whether participants' behavior in this pantomime comprehension task was strongly influenced by the constituent order of the participants' native language. The results of Experiment 1 were consistent with the hypothesis that participants gave the highest likelihood ratings to orders that matched the structure of English: namely, SVO. If consistency with the native language were the reason for this observation, we would have expected to see the Turkish speakers in Experiment 2 display a strong preference for SOV for both reversible and non-reversible trials. This was not the case. As can be seen in Figure 3, there is very little evidence of a generalized SOV preference. Only for non-reversible events did SOV pantomime clips receive significantly higher likelihood ratings than SVO pantomime clips. For reversible events, SOV and SVO pantomime clips were rated equally likely [$F(1,28) = .36, p = .56$]. This is not the pattern that would be expected if participants simply gave high ratings to pantomime clips that matched the constituent order of their native language. Therefore, we next consider how this pattern of behavior in pantomime comprehension compares to previous studies of pantomime production.

In a study conducted concurrently with the present experiments, Hall, Ferreira, and Mayberry (submitted) asked another group of native Turkish speakers to describe these same events in pantomime. For non-reversible events, the rate of SOV production was 86.6%, whereas the rate of SVO production was 3.7%. The magnitude of this discrepancy in production stands in contrast to the slight advantage that SOV had over SVO in the comprehension ratings for non-reversible events (80.4 for SOV vs. 75.6 for SVO). Given these results, it is difficult to take the present finding of an SOV advantage as evidence of concordance between comprehension and production in terms of

constituent order preferences. As previously mentioned, the ideal analysis would directly test for a three-way interaction of constituent order x reversibility x comprehension/production, but such an analysis would violate statistical assumptions of independence and introduce confounding factors. For now, we must make do with the descriptive differences, and exercise prudence in interpreting them.

Furthermore, several studies of pantomime production have found that even SOV speakers nevertheless avoid SOV for reversible events (Gibson et al., 2011; Hall, Ferreira, & Mayberry, submitted; Meir et al., 2010). That is, pantomimers were far less likely to produce SOV descriptions for reversible as compared to non-reversible events. As noted in the introduction, some explanations for why participants avoided using SOV to describe reversible events focus on the idea that such PERSON-PERSON-ACTION sequences would be potentially confusable to an addressee (Gibson et al., 2011; Meir et al., 2010). The current findings do not support this explanation, because participants did not show an analogous pattern in their ratings of SOV for reversible and non-reversible events. At the same time, these findings are compatible with the hypothesis that participants' avoidance of SOV is driven by constraints on production, rather than by concerns about potential ambiguity in comprehension (Hall, Mayberry, & Ferreira, submitted). (Of course, it remains possible that producers' behavior is guided by a *faulty* mental model of comprehenders' preferences. However, this would still be a production-centered locus, rather than one that is informed by direct or indirect feedback from actual comprehenders.)

To describe reversible events in pantomime, several previous studies report increases in SVO (Gibson et al., 2011; Hall, Ferreira, & Mayberry, submitted; Hall,

Mayberry, & Ferreira, submitted). For example, whereas the Turkish speakers from Hall, Ferreira, and Mayberry (submitted) produced SVO descriptions for only 3.7% of non-reversible trials, that figure rose to 21.4% of reversible trials. However, the current participants show no evidence of a corresponding increase in their likelihood ratings in comprehension. Instead, the present participants indicated that SVO descriptions seemed equally likely for reversible and non-reversible events. This again suggests divergence in constituent order preferences between pantomime production and comprehension. The general discussion considers the specific possibility that there may be a general preference for SVO order in comprehension.

Another effect that has been reliably observed in pantomime production is that OSV either remains constant or becomes more common for reversible than non-reversible events (Hall, Ferreira, & Mayberry, submitted; Hall, Mayberry, & Ferreira, submitted; Meir et al., 2010). Because OSV orders, like SOV orders, consist of PERSON-PERSON-ACTION sequences (for reversible events), we have previously argued that increases in OSV for reversible events constitute evidence against confusability for addressees (Hall, Mayberry, & Ferreira, submitted). A novel contribution of the present study is that OSV descriptions of reversible events were rated as being less likely than OSV descriptions of non-reversible events. Once again, this suggests a substantial divergence between constituent order preferences in pantomime comprehension and production.

General Discussion

The experiments presented here demonstrate that constituent order preferences in comprehension, as measured by our novel task, are different from those in production, as

measured by previous studies of elicited pantomime. Comprehenders did not always prefer SOV for non-reversible events, whereas producers did. Comprehenders did not give reliably lower ratings to SOV descriptions of reversible events than of non-reversible events, whereas producers did. Comprehenders did not specifically boost SVO for reversible events, whereas producers did. Finally, comprehenders showed a dispreference for OSV descriptions of reversible events (marginal for English speakers, significant for Turkish speakers), whereas producers did not.

In considering the implications of these findings, we return to the questions that motivated the study: if left to their own devices, what constituent order(s) would production and comprehension prefer? To what extent could these processing preferences contribute to attested patterns of language typology and diachronic change?

Results from the earliest studies of pantomime production (Goldin-Meadow et al., 2008; Langus & Nespors, 2010) suggested that the production system might prefer SOV. This is consistent with the observation that homesign, which is developed almost entirely on the basis of production, uses constituent orders that are consistent with SOV (namely, SV and OV). It is also consistent with claims that the earliest stages of human language evolution were characterized by SOV order (Givón, 1979; Newmeyer, 2000).

However, we also know that languages change over time. In terms of constituent order, languages tend to change away from SOV but not toward it. Meanwhile, languages change toward SVO, which tends to be more stable (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977). There are likely to be numerous factors at work in these processes; our question is whether processing dynamics in production and comprehension are among these factors, and if so, how. One clue comes from the process of *creolization*,

in which young learners organize noisy linguistic input into a more regular, grammatical system. A significant feature of creoles is that they are overwhelmingly SVO (Bakker, 2008; McWhorter, 2001). This remains true even when the input languages use SOV order (Kouwenberg, 1992). Although previous studies have focused on the fact that homesign is manual while creoles are oral (Langus & Nespors, 2010), another difference between the two is that creoles are the byproduct of both comprehension and production, whereas homesign is grounded primarily in production alone. Could the involvement of comprehension be partially responsible for the shift toward SVO in creoles, and in languages more generally?

Two previous studies of pantomime production suggested so, based on the finding that producers avoided SOV for reversible events (Gibson et al., 2011; Meir et al., 2010). They reasoned that SOV was dispreferred for reversible events in production primarily because it would be confusable to a comprehender. However, they did not provide empirical evidence in support of that assumption. Meanwhile, Hall, Mayberry, and Ferreira (submitted) suggested that concerns about confusability in comprehension were unlikely to be the culprit, because participants preferred to use orders that included OSV and SOSV. These orders should in theory be just as confusable to a comprehender as SOV, because they also consist of PERSON-PERSON-(PERSON)-ACTION sequences. Instead, Hall, Mayberry, and Ferreira suggested that producers' behavior was driven by a need to avoid role conflict. Although we argued that this constraint operated primarily on production, it could still be the case that what is beneficial to the producer is also beneficial to the comprehender, in which case the systems would be concordant. We did not directly assess whether or not avoiding role conflict was also beneficial to the

comprehender. The present studies addressed these gaps by explicitly testing constituent order preferences in pantomime comprehension.

The first step was to characterize constituent order preferences for non-reversible events. Here, the present data revealed an SVO preference in English speakers, and higher than expected ratings for SVO in Turkish speakers, although slightly less than SOV. This pattern is very different from the strongly dominant SOV preference reported in production for both groups. This pattern also differs from the one previous study of pantomime comprehension (Langus & Nespors, 2010). It's possible that the discrepancy between the studies is that the previous task contrasted correct versus incorrect descriptions, whereas the present task contrasted sequences that differed only in constituent order, keeping semantic content the same. (It is also worth noting that according to Langus and Nespors's account, pantomime should not have shown an SVO preference at all, since gesture is claimed to belong to a non-linguistic computational system with a built-in preference for SOV. These data argue against such a modular view.) On the basis of these data, we suggest that the preference for using SOV to describe non-reversible events comes primarily from the production system, rather than from producers' mental models of what comprehenders would prefer.

The next step was to test whether comprehenders would disprefer SOV descriptions of reversible events, as producers do, and as assumed by both Gibson et. (2011) and Meir et al. (2010). If so, comprehenders should give lower ratings to SOV descriptions of reversible events than to SOV descriptions of non-reversible events. Although our participants showed slight numeric trends in this direction, these effects (9.6% difference in English; 3.3% in Turkish) stand in contrast to the 40-60% reduction

observed in studies of pantomime production using the same stimuli (Hall, Ferreira, & Mayberry, submitted; Hall, Mayberry, & Ferreira, submitted). The absence of a strong SOV penalty for reversible events in comprehension argues against a confusability-based account of SOV avoidance in production.

The question then becomes whether the role conflict account offered by Hall, Mayberry, and Ferreira (submitted) fares any better. These authors assumed that the constraint to avoid role conflict operated primarily on production, but did not rule out the possibility that comprehenders might also benefit from a lack of role conflict. The stimuli in the present experiments allowed us to test this possibility, because all of the SOV descriptions of reversible events were instances of role conflicts. The fact that participants did not rate these reversible SOV stimuli significantly lower than their non-reversible counterparts suggests that participants were not bothered by the presence of role conflicts. This, in turn, suggests that SOV avoidance in production was not driven by producers' mental models of comprehenders' preferences.

Based on these data, it appears that, left to its own devices, the production system prefers to use SOV to describe non-reversible events, but avoids SOV for reversible events. Of the many non-SOV alternatives, SVO seems to be particularly salient. Langus and Nespors (2010) suggest that this is because the computational system of grammar inherently prefers SVO, while Hall, Ferreira, and Mayberry (submitted) suggest that this is because SVO is an efficient way of avoiding role conflict in production while keeping the subject before the object.

Constituent order preferences in comprehension are more complex, but clearly different from those in production. There was no single order that was dominant across

both English and Turkish speakers, and it is possible that participants' native language background influenced their behavior in this task to some extent. However, that alone cannot fully account for the data. Instead, we propose that, relative to production, comprehension prefers SVO for both reversible and non-reversible events. We further suggest that participants' responses were influenced by this general SVO preference in comprehension combined with additional influence from their own native language, such that the end result was that English speakers boosted SVO (their native language order) above all other orders, while Turkish speakers boosted SOV (their native language order) up to (and slightly beyond) the level of SVO. (This pattern can be seen by visually comparing Figures 2 and 3.)

Taken together, these data suggest that the interplay between production and comprehension might be a factor that contributes to the variation in constituent order preferences across languages and how such preferences change across generations. Human language may have begun with SOV order (Givón, 1979; Newmeyer, 2000), but languages gradually shift away from SOV, toward SVO (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977). The general assumption made by approaches that view language patterns as the result of cognitive pressures has been that this shift is motivated largely by the potential for SOV utterances to be confusable to an addressee for reversible events, whereas SVO is less confusable. The present data call for a revision to this story. Our findings suggest that although comprehension does prefer SVO, this is equally true for both reversible and non-reversible events. Instead, we suggest that constraints on production play a greater role in shifting languages away from SOV. This has the ancillary benefit of also being preferred by comprehension, even though it is not

specifically intended for the benefit of comprehenders. This independent convergence on SVO may help to explain why languages generally change from SOV to SVO but not vice versa.

Experimental studies using elicited pantomime have provided useful insights into the cognitive foundations of language structure, but they –like all studies– have limitations. One is that our experimental paradigm is not a direct measure of comprehension: that is, it does not assess how well or how quickly participants understood the pantomimed stimuli. Rather, it asks participants to judge the likelihood that a given pantomime stimulus was actually what a previous participant had produced to describe a given event. Therefore, at minimum, these data indicate that pantomime comprehenders are poor judges of what pantomime producers are likely to do. And yet that very insight provides a useful clue about the cognitive forces that shape constituent order. It is surprising to realize that, had we asked these same participants to describe events in pantomime, their own behavior would probably have differed from what they judged to be likely. This in turn suggests a deep disconnect between the processes involved in production and comprehension. Still, it will be important for future studies to probe comprehension from a variety of angles. Traditional comprehension studies generally contrast participants' responses to correct versus incorrect semantic content (as in Langus & Nespors, 2010); our paradigm was motivated by the goal of contrasting constituent orders while hold semantic content constant. There is a way to accomplish that goal in a more traditional paradigm, but it requires using only reversible events as stimuli, which would have been too large of a departure from the previous literature on pantomime comprehension (which considered only non-reversible events).

In such a study, participants would see a pantomimed stimulus in any order, and would be asked to choose between two pictures that are *both* plausible interpretations, such as a woman pushing a boy, and a boy pushing a woman. This would be a very straightforward measure of comprehension: that is, how do people tend to interpret sequences of PERSON-PERSON-ACTION (SOV or OSV?), PERSON-ACTION-PERSON (SVO or OVS?), and ACTION-PERSON-PERSON (VSO or VOS?). The same methods could be used with spoken stimuli, or even non-speech stimuli (e.g. slide whistle tones, or arbitrary images), which would help to clarify the extent to which constituent order preferences are modulated by modality. If production and comprehension do have congruent preferences, then participants should have no clear preference for PERSON-PERSON-ACTION sequences or for ACTION-PERSON-PERSON sequences, but should reliably interpret PERSON-ACTION-PERSON as SVO. However, based on the findings above, I would predict that comprehenders would reliably interpret PERSON-PERSON-ACTION sequences as SOV, suggesting that production and comprehension have different constituent order preferences. At present, this remains an empirical question.

Another limitation of the studies conducted thus far is that they have focused on production in isolation or comprehension in isolation. Although these are reasonable starting points for empirical study, it will be important to test whether the same preferences are observed when production and comprehension interact in the service of communication. This should be a rich area for future research.

Conclusions

The goal of this paper has been to contribute to our understanding of patterns of constituent order distributions and how they change within and across generations.

Drawing on insights from typological surveys, functional linguistics, homesign, creoles, emerging sign languages, and previous empirical studies of pantomime production, we hypothesized that different processing dynamics in production and comprehension might contribute to the observed distributions of constituent orders in the world's languages over time. Specifically, languages might begin with SOV order in part because that order is favored by the production system, as evidenced by the robust preferences for SOV (including SV and OV) in both homesign and in previous studies of pantomime production. Likewise, languages might gravitate away from SOV toward SVO because of two independent but converging factors. The first is the production system's dispreference for using SOV to describe reversible events. Although this pattern had previously been attributed to producers' mental models of comprehender preferences, the current data argue against this view, and instead support a production-centered locus. The second factor is that the comprehension system seems to prefer SVO more than the production system does, for both reversible and non-reversible events. Importantly, both factors were attested across speakers of English (SVO) and Turkish (SOV). We suggest that over time, the joint impact of these processing preferences conspire such that languages move away from SOV, toward SVO, even in the absence of external influences. More broadly, these observations suggest that the preferred constituent orders of the world's languages are a product of psychological tendencies through the interplay of the preferences that mechanisms of production versus comprehension have for representing and processing dynamic events.

Chapter 4, in part, has been submitted for publication of the material as it appears in 2012, Hall, M. L., Mayberry, R. I., & Ferreira, V. S. Comprehenders prefer SVO, but

producers prefer SOV: Evidence from elicited pantomime. The dissertation author was the primary investigator and author of this paper.

Appendix

Agent	Action	Patient	Reversible
boy	pet	ball	non-reversible
boy	pet	bike	non-reversible
boy	pet	box	non-reversible
girl	pet	cat	non-reversible
woman	pet	ball	non-reversible
woman	pet	bike	non-reversible
woman	pet	box	non-reversible
woman	pet	cat	non-reversible
boy	push	ball	non-reversible
boy	push	bike	non-reversible
boy	push	box	non-reversible
girl	push	cat	non-reversible
man	push	ball	non-reversible
man	push	bike	non-reversible
woman	push	box	non-reversible
woman	push	cat	non-reversible
boy	kiss	girl	reversible
boy	kiss	woman	reversible
woman	kiss	man	reversible
woman	kiss	boy	reversible
boy	lift	girl	reversible
girl	lift	woman	reversible
man	lift	woman	reversible
woman	lift	boy	reversible
boy	pet	girl	reversible
boy	pet	woman	reversible
man	pet	woman	reversible
woman	pet	boy	reversible
boy	push	woman	reversible
girl	push	boy	reversible
woman	push	girl	reversible
woman	push	man	reversible

References

- Bakker, P. (2008). Pidgins versus creoles and pidgincreoles. In S. Kouwenberg & J. V. Singler (Eds.), *The handbook of pidgin and creole studies* (pp. 130-157). Malden, MA: Wiley-Blackwell.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a complex adaptive system. *Language Learning*, 59, s1-26.
- Bybee, J. L. (1988). The diachronic dimension in explanation. In J. A. Hawkins (Ed.), *Explaining language universals* (pp. 350-379). New York: Basil Blackwell.
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25, 257-271.
- Dayley, J. (1985). Why all languages aren't SOV or SOV, or how competing motivations lead to natural inconsistency. *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 52-64.
- Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptation in language production: Limiting the role of the "Model of the Listener". In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore festschrift for Lila Gleitman* (pp. 105-129). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dryer, M. S. (1992). The Greenbergian word order correlations. *Language*, 68, 81-138.
- Dryer, M. S. (2008). Order of subject, object, and verb. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures online* (<http://wals.info>). Munich: Max Planck Digital Library.
- DuBois, J. W. (1985). Competing motivations. In J. Haiman (Ed.), *Typological Studies in Language, Vol 6* (pp. 343-365).
- DuBois, J. W. (2003). Argument structure: Grammar in use. In J. W. DuBois, L. E. Kumpf, & W. J. Ashby (Eds.), *Preferred argument structure: Grammar as architecture for function* (pp. 11-60). Philadelphia: John Benjamins Publishing Company.
- Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation*, 49, 209-246.
- Gell-Mann, M. & Ruhlen, M. (2011). The origin and evolution of word order.

Proceedings of the National Academy of Sciences, 108, 17290-17295.

- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1-76.
- Gibson, E., Brink, K., Piantadosi, S. & Saxe, R. (2011, April). *Cognitive pressures explain the dominant word orders in language*. Presented at the CUNY Conference on Human Sentence Processing, Palo Alto, CA.
- Givón, T. (1979). *On understanding grammar*. New York: Academic Press.
- Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.
- Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication in deaf children: Noneffect of parental input on language development. *Science*, 221, 372-374.
- Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, 391, 279-281.
- Goldin-Meadow, S., Özyürek, A., Sancar, B., & Mylander, C. (2008). Making language around the globe: A crosslinguistic study of homesign in the United States, China, and Turkey. In J. Guo, E. Lieven, N. Budwig, & S. Ervin-Tripp (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 27-39). New York: Psychology Press.
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 105, 9163-9168.
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language: Report of a conference held at Dobb's Ferry, NY, April 13-15, 1961* (pp. 73-113). Cambridge: MIT Press.
- Hall, M. L., Ferreira, V. S., & Mayberry, R. I. (2012). *Investigating constituent order change with the hands: Consistency promotes the emergence of SVO in pantomime*. Manuscript submitted for publication.
- Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2012). *Cognitive constraints on constituent order: Evidence from elicited pantomime*. Manuscript submitted for publication.
- Haspelmath, M. (1999). Are there principles of grammatical change? *Journal of*

Linguistics, 35, 579-595.

Haspelmath, M. (2002). On understanding word order asymmetries (Comments on John A. Hawkins, "Symmetries and asymmetries: Their grammar, typology, and parsing'). *Theoretical Linguistics*, 28, 159-170.

Hawkins, J. (1983). *Word order universals*. New York: Academic Press.

Hawkins, J. A. (1994). *A performance theory of order and consistency*. Cambridge: Cambridge University Press.

Hawkins, J. A. (2002). Symmetries and asymmetries: Their grammar, typology, and parsing. *Theoretical Linguistics*, 28, 95-149.

Hawkins, J. A. (2004). *Efficiency and complexity in grammar*. Oxford: Oxford University Press.

Jaeger, T. F., & Tily, H. (2011). On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 323-335.

Kirby, S. (1999). *Function, selection, and innateness: The emergence of language universals*. New York: Oxford University Press.

Kouwenberg, S. (1992). From OV to VO: Linguistic negotiation in the development of Berbice Dutch Creole. *Lingua*, 88, 263-299.

Ladefoged, P. (1982). *A course in phonetics* (2nd ed). New York: Harcourt Brace Jovanovich.

Langus, A., & Nespors, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, 291-318.

Lehmann, W. P. (1973). A structural principle of language and its implications. *Language*, 49, 47-66.

Li, C. N. (1975). *Word order and word order change*. Austin, TX: University of Texas Press.

Li, C. N. (1977). *Mechanisms of syntactic change*. Austin: University of Texas Press.

Malkiel, Y. (1968). The inflectional paradigm as an occasional determinant of sound change. In W. P. Lehmann & Y. Makiel (Eds.), *Directions for Historical Linguistics* (pp. 23-64).

- Maurits, L., Perfors, A., & Navarro, D. (2010). Why are some word orders more common than others? A uniform information density account. *Advances in Neural Information Processing Systems*, 23, 1585-1593.
- McWhorter, J. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology*, 5, 125-166.
- Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. A. (2010, April). *The interaction of animacy and word order in human languages: A study of strategies in a novel communication task*. Paper presented at the 8th International Conference on the Evolution of Language, Utrecht, Germany.
- Newmeyer, F. J. (2000). On the reconstruction of "proto-world" word order. In C. Knight, M. Studdardt-Kennedy, & J. R. Hurford, (Eds.), *The evolutionary emergence of language* (pp. 372-388). New York: Cambridge University Press.
- Newmeyer, F. I. (2002). The SO/OS asymmetry and two interpretations of the MaOP: Commentary on John A. Hawkins, 'Symmetries and asymmetries: Their grammar, typology, and parsing'. *Theoretical Linguistics*, 28, 171-176.
- Slobin, D. I. (1980). The repeated path between transparency and opacity in language. In U. Bellugi & M. Studdart-Kennedy (Eds.), *Signed and spoken language: Biological constraints on linguistic form* (pp. 229-243).
- Tomlin, R. S. (1986). *Basic word order: functional principles*. London: Croom Helm.
- Vennemann, T. (1973). Explanation in syntax. In J. P. Kimball (Ed.), *Syntax and semantics, Vol 2* (pp. 1-50). New York: Seminar Press.
- Vennemann, T. (1974). Topics, subjects, and word order: from SXV to SVX via TVX. In J. M. Anderson & C. Jones (Eds.), *Historical linguistics, Vol 1* (pp. 339-376). Amsterdam: North Holland Publishing Company.
- Vennemann, T. (1975). An explanation of drift. In C. N. Li (Ed.), *Word order and word order change* (pp. 269-305). Austin, TX: University of Texas Press.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort*. New York: Hafner Publishing Company.

Author note

The authors gratefully acknowledge members of the Laboratory for Multimodal Language Development, the Language Production Laboratory, the Sign Language Research Group, and the Center for Research in Language at UC San Diego for helpful feedback. We are also indebted to Professor Aylin Kuntay and Koç Üniversitesi for their facilities and support. This research would not have been possible without assistance with data collection and coding from Ecem Çoban, Rakel Duenyas, and Youngju (Danbi) Ahn. Funding for this research was provided by NIH grant HD051030, NIH grant 5T32DC000041-19, and by the Rita L. Atkinson Graduate Fellowship.

CHAPTER 5

GENERAL DISCUSSION

Together, these three studies have yielded new insights about the cognitive foundations of language structure. When hearing people who are naïve to sign language describe events in gesture without speech, there is a surprising amount of regularity in what they produce, both across individuals and across speakers of very different languages, such as English and Turkish. Following Goldin-Meadow et al. (2008), I use these regularities as clues to the functional pressures that operate on human language, spoken or signed. In natural language contexts, the effects of these pressures are typically observable only over very long time scales, perhaps because natural languages are subject to many different pressures, some of which are in opposition to one another, thereby slowing the pace of change. However, I have shown that through empirical manipulation, some of these hidden pressures can become readily observable, suggesting that they are latently present in the minds of language users. The findings of these studies have sometimes supported and sometimes challenged previous theories of cognitive influences on language structure.

The present research was inspired by the observation that SOV order seemed to be uniquely preferred when participants with diverse linguistic backgrounds were asked to describe transitive events in pantomime (Goldin-Meadow et al., 2008). These authors hypothesized that SOV order is especially compatible with the way that humans organize our conceptual representations of dynamic transitive events. This accorded well with the observations that SOV is the single most common order across the world's languages today (Dryer, 2008; Greenberg, 1963; Hawkins, 1983; Tomlin, 1986), and that SV and OV order are prevalent in homesign systems created by linguistically isolated deaf children (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, Özyürek, Sancar, &

Mylander, 2008). However, two additional linguistic facts seemed to be at odds with this observation. First, although SOV is the single prevalent constituent order in the world's languages, it is not the only one. Indeed, roughly half of the world's languages use a non-SOV order, with SVO being the most common of these (Dryer, 2008; Greenberg, 1963; Hawkins, 1983; Tomlin, 1986). Second, when constituent order changes over time, languages move away from SOV, not toward it (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977). If SOV is ideally suited for expressing transitive events, we must wonder why there is such a proliferation of alternative constituent orders, and why SOV becomes less dominant, rather than more dominant, over time.

The results of Study 1 may offer a way to resolve these apparent discrepancies. Experiments 1-3 discovered that constituent order preferences change dramatically depending on semantic properties of the to-be-described events. We replicated previous observations that SOV is strongly preferred for describing semantically non-reversible events, but showed that participants avoided using SOV to describe reversible events. Two other studies observed similar findings (Gibson et al., 2011; Meir et al., 2010), which increases confidence in the reliability of these data. However, those studies suggested that producers avoided SOV for reversible events because of the potential for confusability on the part of an addressee. In contrast, the results of Study 1 suggest that constraints on production may play a greater role than has been previously thought. Specifically, I introduced the concept of *role conflict* in production. Because participants are motivated to be in the agent role at the time they produce an action gesture, and because they typically adopt the patient role to describe a human object, the sequence O-V would present a conflict between these two roles for events with both a human agent

and a human patient. I claim that it is primarily for this reason that participants avoid describing reversible events with SOV order.

A somewhat unexpected finding of Study 1 was that participants sometimes produced gestures that had some of the features of the case marking systems found in natural languages. For example, these gestures did not refer to semantic content, but instead were functional units that differentiated the roles of the nominal gestures. Furthermore, the earlier reversible events began to appear in the testing session, the more common these gestures became. Although I do not claim that the participants in Study 1 have created a true case marking system, I do suggest that their behavior shows evidence of a cognitive drive to mark reversible events specially. In natural languages, this drive is often realized by case marking systems. This is especially evident in languages that use differential object marking, where case marking is increasingly likely for reversible events.

When the participants in Study 1 avoided using SOV to describe reversible events, they gravitated toward three main alternatives. One was to move the object to the end of the utterance, yielding SVO. Another was to move the object to the beginning of the utterance, yielding OSV. The third was to repeat a gesture referring to the subject before producing the action gesture, yielding SOSV. Although SVO, OSV, and SOSV are all effective ways to avoid role conflict in production, they are not all attested in equal measure among the world's languages. In fact, of these three solutions, only SVO is attested to a substantial degree as a basic constituent order in the world's languages. OSV is the least common basic word order across languages, and no languages yet described have a basic constituent order that involves the kind of repetition found in SOSV

utterances. This raises questions about the extent to which the results of Study 1 are informative about constituent order in natural languages. First, why were OSV and SOSV more prevalent in Study 1 than would be expected based on their infrequency in natural languages? Second, should the prevalence of SVO in Study 1 be taken as evidence for a cognitive drive toward SVO specifically, or might that simply reflect the fact that the participants in Study 1 were native speakers of an SVO language (English)?

Study 2 was designed to address these questions. We hypothesized that pantomime might not be subject to all of the same pressures that operate on natural language. For example, natural languages generally mention subjects before objects, and are highly efficient. These pressures may be less relevant for pantomime. If so, that might explain why some patterns that are relatively uncommon in natural language (e.g. OSV, SOSV) were relatively common in pantomime. To test this hypothesis, we modified the pantomime paradigm to more closely approximate the features of natural languages. A first step was to simulate the emergence of a lexicon by asking some participants to create and use consistent gestures for the people, objects, and actions in our stimulus set. A further step was to stimulate the emergence of shared linguistic knowledge by asking a subset of the lexicon users to also teach their gestures to an experimenter. We then compared the constituent orders produced by participants in these more language-like conditions to those produced by participants in a baseline condition (which was essentially identical to Study 1). In addition, to assess whether the constituent order of the participants' native language would influence their behavior, we tested native speakers of both an SVO language (English) and an SOV language (Turkish). Results indicated that under more language-like conditions, participants were

more likely to be efficient and to mention subjects before objects. As in Study 1, participants avoided using SOV to describe reversible events. Crucially, when describing reversible events under more language-like conditions, SVO became a more prevalent solution than orders like OSV and SOSV, for both English and Turkish speakers. This suggests that SVO has become increasingly prevalent in human language in part because it is an efficient way to describe reversible events while mentioning the subject before the object and avoiding the use of SOV to describe reversible events. This view offers a more motivated explanation for the emergence of SVO than a previous account that attributed the emergence of SVO to an innately-specified grammatical preference (Langus & Nespors, 2010).

The idea that the rise of SVO is related to the need to communicate about reversible events is familiar from the linguistic literature (at least, from those linguists who are concerned with identifying cognitive mechanisms for linguistic change). These linguists commonly adopt the intuitive assumption that the potential for confusability in comprehension is what ultimately drives these changes. Because producers presumably know the intended meaning of an utterance, it stands to reason that it is comprehenders who are at risk for misunderstanding who is the agent and who is the patient in an utterance of the form PERSON-PERSON-ACTION. However, very little evidence has been brought to bear on this assumption. Aside from transfer effects in second-language acquisition, psycholinguistics in general has had surprisingly little to say about processing preferences for various constituent orders. A notable exception is Hawkins (1994, 2002, 2007), who has argued that constituent orders that are most prevalent in the world's languages are those that are easiest to process. However, even Hawkins has not

been explicit about the source of this processing ease, although at first glance it appears that he too assumes that comprehension is likely to be the culprit.

Some of the most relevant claims about the connections between production, comprehension, reversibility, and constituent order come from studies of elicited pantomime. However, these claims are in conflict, with some emphasizing the potential for confusability in comprehension (Gibson et al., 2011; Meir et al., 2010), and others emphasizing role conflict in production (Study 1 and Study 2 of this dissertation). Notably, all of these studies have focused exclusively on production, without directly assessing constituent order preferences in comprehension. Study 3 was designed to fill this gap, and its results speak to both the specific question of whether SOV-avoidance in pantomime production is due to constraints on comprehension or production, and to the more general question of how the processing dynamics of comprehension and production might shape constituent order in the world's languages over time.

Regarding the question of whether SOV avoidance in pantomime production is due to constraints on comprehension or constraints on production, the results of Study 3 support the production-based accounts of Study 1 and Study 2. If constraints on comprehension were what drove SOV avoidance in previous studies of pantomime production, we would have expected comprehenders in Study 3 to give markedly lower ratings to SOV descriptions of reversible events than to SOV descriptions of non-reversible events. In fact, there was only a marginal trend in this direction for English speakers, and no difference at all for Turkish speakers. Likewise, whereas producers treated OSV as a potential solution for describing reversible events, comprehenders rated OSV as being worse for reversible events (again a trend for English speakers, but a

significant difference for Turkish speakers). Therefore, the present results provide little support for the intuition that PERSON-PERSON-ACTION sequences are especially problematic for comprehenders. Their behavior in Study 3 suggests that they interpret such sequences as SOV, and reject OSV interpretations. It is therefore all the more surprising that producers avoid using SOV, indicating that they do not do so based on an accurate assessment of what would be preferred by a comprehender. However, it would be prudent to replicate these findings using a paradigm that more directly measures how comprehenders semantically interpret PERSON-PERSON-ACTION descriptions, with stimuli that are signed, spoken, or other.

Regarding the question of how comprehension and production might jointly shape constituent order more broadly, the results of Studies 1-3 together offer a new explanation for the attested shifts away from SOV, toward SVO. According to this account, the drive away from SOV is sparked not by comprehension, but by production (Study 1). Of the many possible alternatives to SOV, producers gravitate toward SVO because it allows them to avoid role conflict while still being efficient and mentioning subjects before objects (Study 2). This shift toward SVO meets with little resistance in comprehension, because comprehenders are quite willing to accept SVO descriptions of both reversible and non-reversible events (Study 3). Although this shift was not motivated specifically for the benefit for comprehenders (who would also have accepted SOV descriptions for both types of events), it is nonetheless compatible with comprehension preferences. Thus, in the presence of strong evidence of cognitive pressures leading languages away from SOV, and in the absence of cognitive pressures leading languages away from SVO, it seems reasonable to expect that, all other things

being equal, the distribution of constituent orders in human languages should continue to shift such that SOV becomes less common and SVO becomes more common.

One limitation of the above account is that the role conflict hypothesis, while useful for understanding constituent order preferences in pantomime and sign languages, finds no clear analog in spoken languages. There may be other constraints on production in spoken languages, such as the monitoring hypothesis proposed at the end of Chapter 2 of this dissertation. Constraints on comprehension may play a greater role than detected with the paradigm presented in Chapter 4 of this dissertation, and of course there may also be constraints on child language acquisition that allow some constituent orders to be acquired more efficiently or robustly than others. These questions remain open for further research. However, regardless of the mechanism behind SOV-avoidance for reversible events, the findings reported in Chapters 3 and 4 of this dissertation should remain intact.

Through this work, I hope to have demonstrated that gesture provides a window into the cognitive pressures that have been hypothesized to shape human language over centuries and millennia. These methods allow us to see those forces at work on remarkably short time scales, and in the minds of people whose native language may not (yet) show overt evidence of those processes. Furthermore, elicited pantomime may also prove to be a particularly useful tool for testing theories derived from observations of newly emerging sign languages. Constituent order is currently under-studied in these languages, and I hope that the work presented here will inspire further investigation into how constituent orders emerge and evolve in more naturalistic contexts. If we wish to

arrive at a full understanding of the patterns that characterize human language, I humbly submit that we would do well to keep the hands in mind.

References

- Dryer, M. S. (2008). Order of subject, object, and verb. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures online* (<http://wals.info>). Munich: Max Planck Digital Library.
- Gell-Mann, M. & Ruhlen, M. (2011). The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, 108, 17290-17295.
- Gibson, E., Brink, K., Piantadosi, S. & Saxe, R. (2011, April). *Cognitive pressures explain the dominant word orders in language*. Presented at the CUNY Conference on Human Sentence Processing, Palo Alto, CA.
- Givón, T. (1979). *On understanding grammar*. New York: Academic Press.
- Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, 391, 279-281.
- Goldin-Meadow, S., Özyürek, A., Sancar, B., & Mylander, C. (2008). Making language around the globe: A crosslinguistic study of homesign in the United States, China, and Turkey. In J. Guo, E. Lieven, N. Budwig, & S. Ervin-Tripp (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 27-39). New York: Psychology Press.
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 105, 9163-9168.
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language: Report of a conference held at Dobbs's Ferry, NY, April 13-15, 1961* (pp. 73-113). Cambridge: MIT Press.
- Hawkins, J. (1983). *Word order universals*. New York: Academic Press.
- Hawkins, J. A. (1994). *A performance theory of order and consistency*. Cambridge: Cambridge University Press.
- Hawkins, J. A. (2002). Symmetries and asymmetries: Their grammar, typology, and parsing. *Theoretical Linguistics*, 28, 95-149.
- Hawkins, J. A. (2007). Processing typology and why psychologists need to know about it. *New Ideas in Psychology*, 25, 87-107.
- Langus, A., & Nespors, M. (2010). Cognitive systems struggling for word order.

Cognitive Psychology, 291-318.

Li, C. N. (1977). *Mechanisms of syntactic change*. Austin: University of Texas Press.

Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. A. (2010, April). *The interaction of animacy and word order in human languages: A study of strategies in a novel communication task*. Paper presented at the 8th International Conference on the Evolution of Language, Utrecht, Germany.

Tomlin, R. S. (1986). *Basic word order: functional principles*. London: Croom Helm.