Title
Mouthings in American Sign Language : : biomechanical and representational foundations

Author
Udoff, Jonathan Andrew

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Mouthings in American Sign Language: Biomechanical and Representational Foundations

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

in

Language and Communicative Disorders

by

Jonathan Andrew Udoff

Committee in charge:

San Diego State University

Professor Karen Emmorey, Chair
Professor Ignatius Nip, Co-Chair

University of California, San Diego

Professor Rachel Mayberry
Professor Carol Padden
Professor Sharon Rose

2014

The dissertation of Jonathan Andrew Udoff is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____

____

____

____
Co-Chair

____
Chair

University of California, San Diego

San Diego State University

2014

TABLE OF CONTENTS

Signature Page iii

Table of Contents iv

Notation Conventions viii

Abbreviations for Foreign Sign Languages ix

List of Figures x

List of Tables xii

Acknowledgements xiii

Vita xv

Abstract of the Dissertation xvii

Introduction 1

Chapter 1: Theoretical Background 3

1.1. Mouth Actions: The Mouth as an Articulator in Sign Language 3

1.1.1. Mouthings 4

1.1.2. Mouth gestures 21

1.1.3. Mouth actions: further considerations 24

1.1.3.1. Representation of mouth actions in signed languages 24

1.1.3.2. Are the hands the head of the mouth? 27

1.1.3.3. Are mouthings and mouth gestures the same? 30

1.1.3.4. Methodological pitfalls and new techniques 32

1.2. Interactions Between the Hands and Mouth: Articulatory and Neurological Considerations ... 38
1.2.1. Kinematics of the coordination of the hands and mouth ... 39
1.2.1.1. Reaching with arbitrary mouth movements ... 39
1.2.1.2. Reaching with speech ... 41
1.2.1.3. Gesturing with speech ... 42
1.2.2. Directionality of the hand–mouth interaction ... 43
1.3. Motivation for Current Work ... 46
Chapter 2: Coordination of the Hand and Mouth for Fingerspelled Nonce Words (Experiment 1) ... 49
2.1. Introduction ... 49
2.2. Methods ... 52
2.2.1. Participants ... 52
2.2.2. Materials ... 53
2.2.3. Procedure ... 56
2.2.4. Data capture ... 56
2.2.5. Data post-processing ... 58
2.2.6. Missing data ... 59
2.3. Results ... 60
2.4. Discussion ... 66
Chapter 3: Coordination of the Hand and Mouth between Manual Homophone Pairs (Experiment 2) ... 71

3.1. Introduction 71
3.2. Methods 73
3.2.1. Participants 73
3.2.2. Materials 74
3.2.3. Procedure 74
3.2.4. Data capture 76
3.2.5. Post-processing 76
3.2.6. ELAN coding 77
3.2.7. Missing data 79
3.3. Results 79
3.4. Discussion 90
Chapter 4: The Motor Representation of Mouthings and Speech (Experiment 3) 95
4.1. Introduction 95
4.2. Methods 100
4.2.1. Participants 100
4.2.2. Materials 100
4.2.3. Procedure 103
4.2.4. Data Capture and Post-Processing 104
4.3. Results 104
4.3.1. Item Stability Analysis 105
4.3.2. Mouthing vs. Voiced Model Analysis 107
4.3.3. Silent Condition Analysis 108
4.4. Discussion 110

Chapter 5: General Discussion 114
5.1. Hand-Mouth Coordination 114
5.2. The Motor Representation of Mouthings 119
5.3. Visibility and Visemes 120
5.4. The Relationship between Speech Skills and the Production of Mouthings 122
5.5. Conclusion 123
5.6. Future Directions 124
Appendices 129
Appendix A: Speech Background Questionnaire 129
Appendix B: Graphs of Fingerspelled Items 133
References 139

Notation Conventions

HOME	The closest English translation of the ASL sign is written as a gloss in small caps. To indicate a foreign sign language, the gloss is written in the surrounding spoken language whenever possible (e.g., Hebrew for Israeli Sign Language).
EVERY-DAY	If more than one English word is required to gloss a single sign, hyphens connect the words.
J-O-H-N	Fingerspelled sequences are written with dashes between each letter.
"vote" or /voʊt/	Mouthing components of ASL production derived from spoken English are either written between quotation marks or as IPA transcriptions between slashes.
deaf	Written in lowercase, *deaf* refers to the audiological condition of hearing loss.
Deaf	When capitalized, *Deaf* refers to membership in the Deaf community, most often associated with use of sign language.

ABBREVIATIONS FOR FOREIGN SIGN LANGUAGES

ABSL	Al-Sayyid Bedouin Sign Language	
ASL	American Sign Language	
BSL	British Sign Language	
DGS	German Sign Language	(*Deutsche Gebärdensprache*)
DSGS	Swiss German Sign Language	(*Deutschschweizer Gebärdensprache*)
IPSL	Indo-Pakistani Sign Language	
ISL	Israeli Sign Language	
IUR	Inuit Sign Language	(*Inuit Uukturausingit*)
LIS	Italian Sign Language	(*Lingua dei Segni Italiana*)
NGT	Dutch Sign Language	(*Nederlandse Gebarentaal*)
NSL	Nicaraguan Sign Language	
SSL	Swedish Sign Language	

LIST OF FIGURES

Figure 2.1: Marker placement for data capture. Participants wore a total of ten reflective markers: four on a rigid body attached to the forehead, two on the lips, one on the chip, one on the index finger, one on the pinky finger, and one on the wrist.......... 58

Figure 2.2: The spatial coordination for items in which the hand and mouth have the same or different number of movements. Error bars in this and all subsequent graphs reflect ±1 standard error of the mean. 63

Figure 2.3: The spatial coordination by number and direction of movements. The data only reflect the subset of items in which the hand and mouth share the same number of movements. 63

Figure 2.4: The temporal coordination by items in which the hand and mouth have the same or different number of movements 65

Figure 2.5: The temporal coordination by number and direction of movements. The data only reflect the subset of items in which the hand and mouth share the same number of movements. 65

Figure 3.1: A manual homophone pair. The left panel shows the sign CANCEL with the corresponding mouthing /ˈkænsəl/. The right panel shows the sign CRITICIZE, which has the same manual features as CANCEL, but with the oral mouthing /ˈkɹɪ.tɪ.saɪz/.. 73

Figure 3.2: The sign MY is produced with the open-B handshape, which touches the central chest. Each repetition began and ended with this sign................................ 76

Figure 3.3: The ELAN annotation window 79

Figure 3.4: The duration results for both participant groups by component and mouthing length condition 81

Figure 3.5: Proportions of which component of the target item began its articulation first 84
Figure 3.6: The proportion of items by number of frames in which the beginning of the manual component precedes the beginning of the mouthing component 85
Figure 3.7: Proportions of which component of the target item finished its articulation first 88
Figure 3.8: The proportion of items by number of frames in which the termination of manual component precedes the termination of the mouthing component 89
Figure 4.1: Within-item consistency for each group across conditions 106
Figure 4.2: Correlations between mouthed items and voiced average models 108
Figure 4.3: Silent condition modeled against averages from the other two production conditions 110
Figure B.1: Individual items graphs of simultaneous hand and mouth movements when each articulator moves once in the same direction. Graphs are averaged across participants. 134
Figure B.2: Individual items graphs of simultaneous hand and mouth movements when each articulator moves once in the opposite directions 135
Figure B.3: Individual items graphs of simultaneous hand and mouth movements when each articulator moves twice and in the same direction 136
Figure B.4: Individual items graphs of simultaneous hand and mouth movements when each articulator moves twice in opposite directions 137
Figure B.5: Individual items graphs when the hand moves twice, but the mouth only moves once 138

LIST OF TABLES

Table 2.1: Experiment 1 stimulus list 55

Table 3.1: Criteria used to identify the temporal beginning and end of the manual component of each target sign item in Experiment 2 77

Table 4.1: Experiment 3 stimulus list 103

ACKNOWLEDGEMENTS

This dissertation is the culmination of seven years of work, and as such I owe a great debt of gratitude to the many people who have helped me both professionally and personally in bringing this project to fruition. First and foremost I would like to thank my advisor and thesis chair, Dr. Karen Emmorey for her continual support. I am humbled by her passion for research and her capacity to always take on new challenges with ease. Despite her hectic schedule, she always had time to give me comments on my latest draft or pilot a study. I could not have asked for a better advisor to supervise me during my graduate career. I would also like to thank my co-chair, Dr. Ignatius Nip, for his invaluable expertise regarding speech motor control as well as providing access to his motion capture equipment and helping me to understand how to interpret the massive amount of data it spewed out. I would like to thank Dr. Rachel Mayberry, Dr. Carol Padden, and Dr. Sharon Rose for their helpful feedback and warm encouragement throughout this process. I would also like to Dr. Martha Tyrone for her advice and support, particularly in the early stages of the experimental design.

I would like to thank my research assistants, Travis Cox, Jackilyn Escosio, Nancy Mariscal, Matt Sampson, Desiree Sasunian, Mahie Sumathipala, for their work in cleaning and editing the raw motion capture data. Behind every scientist is a team of dedicated students, and I am lucky to have worked with such a great group of young scholars.

I would like to thank my friends and family for supporting me throughout this entire process. My deepest gratitude goes out to Max, Brian, Eric, and Ronni for their love and support. I would also like to thank my graduate program cohort, Erica Ellis, Roberto Gutierrez, Amy Pace, and Lara Polse, who have kept me sane all these years. And to the members of LLCN, in particular Allison Bassett, Jill Weisberg, Jen Petich, and Kristen Secora – thank you for being there for me, day after day, and most of all, for making me laugh.

This project could not have happened with out extramural financial support, and I wish to acknowledge and thank Jeff and Sheila Lipinsky for their Lipinsky Family Scholarship as well as the National Institutes of Health for RO1 grant DC010997, awarded to Karen Emmorey.

Lastly, I would also like to extend my deepest thanks to the members of the hearing and Deaf communities for their participation. This dissertation would not have been possible without their cooperation and enthusiasm, and I hope this work benefits their respective communities.

Chapters 2 and 3, in part, are currently being prepared for submission for publication of the materials. Udoff, Jonathan; Nip, Ignatius; and Emmorey, Karen. "Interarticulatory Compensation between the Hands and Mouth during the Production of Mouthings in American Sign Language". The dissertation author was the primary investigator and author of these materials.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Udoff, Jonathan; Nip, Ignatius; and Emmorey, Karen. "The Motor Representation of English Speech and American Sign Language Mouthings". The dissertation author was the primary investigator and author of this material.

VITA

EDUCATION AND WORK HISTORY

2007	B.A. *magna cum laude*	Brown University Cognitive Science – Psycholinguistics & Hispanic Studies – Spanish Language and Linguistics
2007–2014	Graduate Research Assistant	Laboratory for Language & Cognitive Neuroscience, Karen Emmorey, director
2014	Ph.D.	San Diego State University & University of California, San Diego Language and Communicative Disorders

CONFERENCE PRESENTATIONS

8 July 2014	"Coordination between the Hands and Mouth: A Kinematic Experiment in ASL" 6th Conference of the International Society for Gesture Studies University of California – San Diego
11 July 2013	"Mouthings and their Reductions: A Kinematic Approach" 11th Theoretical Issues in Sign Language Research Conference University College London, London, United Kingdom
1 October 2010	"The Phonological Representation of the Non-Dominant Hand: Evidence from Articulatory Compensation in ASL" 10th Theoretical Issues in Sign Language Research Conference Purdue University, West Lafayette, Indiana
9 January 2009	"Put Your Hands Together: Phonological Constraints on Handshape Mapping in ASL" 83rd Annual Meeting of the Linguistic Society of America San Francisco, California
8 November 2008	"Put Your Hands Together: Phonological Constraints on Handshape Mapping in ASL" 8th High Desert Linguistic Society Conference University of New Mexico, Albuquerque, New Mexico

FIELDS OF STUDY

Major field: Linguistics

Studies in sign language – phonology, phonetics, and psycholinguistics
Professor Karen Emmorey, SDSU

Studies in motor control for language – kinematics, articulatory compensation, and multimodal production
Professor Ignatius Nip, SDSU

ABSTRACT OF THE DISSERTATION

Mouthings in American Sign Language:
Biomechanical and Representational Foundations

by

Jonathan Andrew Udoff

Doctor of Philosophy in Language and Communicative Disorders

University of California, San Diego, 2014
San Diego State University, 2014

Professor Karen Emmorey, Chair
Professor Ignatius Nip, Co-Chair

This dissertation explores the phenomenon of co-sign mouthings in ASL. Mouthings are silent movements of the mouth that are derived from English and accompany manual signs, yet it has not been established how closely mouthings resemble voiced speech. Furthermore, it is unclear how mouthing productions are influenced by

interarticulatory coordination between the mouth and hands. These two issues are investigated in a series of three experiments. Chapter 1 lays out the theoretical groundwork of the form and function of oral articulations in sign language as well as reviews previous findings concerning hand–mouth motoric interactions. Chapter 2 investigates the factors that contribute to hand–mouth coordination for mouthed fingerspelled items produced by Deaf participants, using a motion capture recording paradigm (Experiment 1). Results indicate that both spatial and temporal coupling are greater for hand–mouth combinations that are shorter rather than longer and in which the two articulators share the same number of movements. Direction of the articulators' movements, however, did not affect coordination. Moreover, items mismatched in number of movements coordinate by one articulator's single movement lengthening its duration to match the temporal domain of the other articulator's two movements. Chapter 3 extends the previous experiment by investigating the same issues but during the production of mouthed lexical signs, with the inclusion of a hearing non-signer comparison group (Experiment 2). Results suggest that while Deaf participants initiate the movements of the hand and mouth simultaneously, hearing participants begin articulating the manual component before the mouthing. However, Deaf and hearing groups do exhibit similar effects of item length. Chapter 4 again uses motion capture to investigate the motor representation of mouthings as they compare to speech in both Deaf signers and hearing non-signers through a comparison of mouth movement profiles (Experiment 3). Results suggest that Deaf signers use a single shared representation to produce mouth movements for both co-sign mouthings and voiced speech. Chapter 5 considers the findings revealed in the experimental data and offers a general discussion of the major issues that underlie the production, representation, and interarticulatory coordination of mouthings in ASL. Finally, suggestions to extend this work and apply its findings to new domains are offered.

INTRODUCTION

The language use of Deaf[1] signers of ASL is quite varied. Even for Deaf individuals born into Deaf families, which account for less than 10% of the Deaf population (Schein & Delk, 1974), signers in the United States are by and large highly bilingual between American Sign Language (ASL) and English, having to interact and communicate with the hearing world on a daily basis. ASL utilizes the visual and gestural modalities to communicate. Although the manual signs produced by the hands have received the most attention in the study of the structure of sign language, non-manual markers such as the eyebrows, mouth, cheeks, head, and torso also contribute to the production of ASL. These non-manual markers serve a variety of communicative functions in the language, conveying both affective and grammatical information. This dissertation is primarily concerned with the role of the mouth in the production of ASL. Additionally, when interacting in the hearing world, deaf individuals may choose to communicate in English, obviously requiring them to move their mouth to produce speech. Therefore, deaf individuals produce language using their mouths in two different,

[1] Following cultural and linguistic conventions, *deaf* will be used to refer to the audiological condition of hearing loss whereas *Deaf* will be used to refer to the cultural identity of Deafhood, most often associated with the use of sign language.

yet functionally overlapping ways. It is the goal of this dissertation to better understand mouthings, an oral feature of sign language, as they are situated in the mental motor representation of native, Deaf signers of ASL as well as how these mouthings coordinate their articulatory movements with the simultaneous movements of the hands.

To this end, the dissertation is divided into the follow chapters: Chapter 1 lays out the theoretical groundwork of the form and function of oral articulations as used by Deaf individuals alongside the manual production of sign language. Additionally, a review of hand–mouth motoric interactions is also discussed from a number of domains. Chapter 2 presents an experiment investigating the factors that contribute to hand–mouth coordination for mouthed fingerspelled items produced by Deaf participants, using a motion capture recording paradigm. Chapter 3 extends the previous experiment by investigating the same issues but during the production of mouthed lexical signs, with the inclusion of hearing non-signer comparison group. Chapter 4 again uses a motion capture technique to investigate the motor representation of mouthings as they compare to speech in both native, Deaf signers and hearing non-signers through a comparison of mouth movement profiles. Chapter 5 considers the findings revealed in the experimental data and offers a general discussion of the major issues that underlie the production, representation, and inter-articulatory coordination of mouthings in ASL. Finally, suggestions to extend this work and apply its findings to new domains are offered.

Chapter 1: Theoretical Background

It is the goal of this chapter to review the various ways in which Deaf signers use the mouth across various language production modes and to explore what factors determine how the mouth articulates. First, in Section 1.1, the form and function of mouthings and mouth gestures during the production of sign language are examined. Next, in Section 1.2, studies of the motoric and neural interactions between the hands and mouth in spoken language are brought to bear on the multimodal nature of sign language production. The chapter concludes in Section 1.3 by posing questions that have motivated the research conducted for this dissertation.

1.1. Mouth Actions: The Mouth as an Articulator in Sign Language

Though the hands are viewed as the primary articulators in sign language, non-manuals, such as the head, eyebrows, and mouth, play an active role in language production. In the following section, the role of the mouth in sign language will be examined through two related phenomena. The first phenomenon, *mouthings,*[2] refers to mouth movements that originate from the surrounding spoken language, usually the oral articulation of the manual sign's spoken language translation. For example, a signer may approximate the English articulation of /kræb/ "crab" while producing the manual sign CRAB. The second, *mouth gestures,* are non-manual mouth configurations that accompany signs and may modify the manual sign much like an adjective or adverb. Mouth gestures, unlike mouthings, do not bear any relation to spoken language. For example, the manual sign WRITE may be combined with the mouth gesture of a protruding tongue to mean *to write carelessly*. Collectively, mouth gestures and mouthings are referred to as *mouth actions.* In this section, I will review the literature on mouthings and mouth gestures across many different sign languages. As ASL is of particular interest in this dissertation, I will draw connections to similar phenomena in ASL when available and make conjectures about whether a phenomenon is likely to occur in ASL where it is appropriate.

1.1.1. Mouthings

Early reports from established European sign languages initially suggested that all sign languages use mouthings, though the prevalence of the phenomena varies between languages. However, more recent observations from rural village sign languages challenge this generalization. Signers living in remote and newly developed sign

[2] Mouthings have also been called spoken pictures, and *mundbilder*. Mouth gestures have also been called oral adverbials, mouth arrangements, and oral components. The terms *mouthings* and *mouth gestures* are used in this paper, following the conventions standardized in Boyes Braem and Sutton-Spence (2001).

language communities usually exhibit no mouthings whatsoever. This has been reported for Al-Sayyid Bedouin Sign Language, a young sign language used in the Negev region of Israel (ABSL; Tkachman, 2012); Kata Kolok, a relatively young language used in northern Bali (de Vos & Zeshan, 2012); and Nicaraguan Sign Language, a very young language used throughout Nicaragua (NSL; Kegl, Senghas, Coppola, & DeGraff, 1999). But there also exists at least one village sign language, Inuit Sign Language, which does make use of mouthings, originating from both Inuktitut and English (IUR; Schuit, 2012). IUR is used by Deaf Inuit throughout the Nunavut territory in northern Canada, and it likely developed in the 18th century from a contact sign system used when Inuit from different bands needed to facilitate communication. It is unclear why village sign languages tend not to employ mouthings given that this type of signing communities generally have more contact with the surrounding spoken language because the majority of the language users are hearing signers. The answer may have something to do with the educational system of the communities or the (lack of) influence of an established national sign language, though the mechanism of such an interaction has yet to be described. In the remainder of this dissertation, mouthings will only be analyzed as they appear in established sign languages.

Although many factors have been observed to contribute to the frequency of mouthings that typically occur in sign language discourse, most reports agree that, overall, about one-half of all manual signs are produced with mouthing (e.g., Schermer, 2001). However, this rate is greatly influenced by a variety of linguistic, cultural, sociopragmatic, and individual factors. General characteristics of a sign language have been argued to affect the amount of mouthing observed in that language. Signed languages that do not have a robust and frequently used fingerspelling system are reported to exhibit more mouthing (Boyes Braem, 2001; Schermer, 2001). Signers of Italian Sign Language (LIS), which only uses fingerspelling to represent foreign words, have been observed to produce mouthings more often than signers of American Sign Language, which makes heavy use of its one-handed manual alphabet (Ajello, Mazzoni, & Nicolai, 2001). For instance, proper names are mouthed in LIS rather than fingerspelled. Presumably, the inverse relationship between the frequency of fingerspelling and the frequency of mouthing in a signed language highlights two methods of borrowing lexical items in signed languages (cf. Brentari, 2001). While languages like ASL can manually spell out a word or proper name of English, a language like LIS only has the option of using mouthings to represent Italian; therefore, users of these languages without a manual alphabet will mouth more often.

The social status of the surrounding spoken language also plays a role in the prevalence of mouthing. In Switzerland, for instance, the various spoken languages of the country (German, French, Italian, Romansh) are emblematic of cultural identity; the pressure to associate oneself with one's regional spoken language has been argued to result in a strong tendency for Deaf Swiss Germans to mouth frequently while signing Swiss German Sign Language (DSGS) (Boyes Braem, 2001). Furthermore, the national philosophies and methodologies employed to educate the deaf children of a country can also put greater emphasis on spoken language within the Deaf community. For instance, Germany has historically embraced the oral method of deaf education, and as a result signers of German Sign Language (DGS) frequently mouth while signing (Keller, 2001).

Other factors that influence the propensity to mouth within a single language user often relate to their personal background and language environment. Age and home language use has been suggested to play a role in the frequency of mouthing in British Sign Language (BSL) (Sutton-Spence & Day, 2001). While the data were not conclusive, Sutton-Spence and Day claim that BSL signers from younger generations as well as signers born into hearing families show a pattern of increased mouthings as compared to older signers and signers born into signing (i.e., Deaf) families. However, the authors admit that the data are equivocal and may alternatively be interpreted as indicating that older signers and native signers show greater variability in their use of mouthing because they are able to manipulate the rate of mouthing across different signing registers (e.g., informative texts, descriptive and narrative texts, child-directed signing). Non-native signers of BSL, in contrast, sign with a maximum rate of mouthing across all register types because they are unable to either perceive or control these subtle differences in register with regards to mouthing (Sutton-Spence & Day, 2001). Boyes Braem (2001) also observes that age of acquisition affects the rate of mouthing in DSGS. In contrast to the findings from BSL, Boyes Braem's analysis reports that late learners of DSGS mouth just as much as early learners (76% vs. 80%). However, like Sutton-Spence and Day (2001), this data is inconclusive due to the greater individual variability among the late learners and the small sample size (n = 3 for each group).

Collapsing across early and late signers, Sutton-Spence and Day (2001) did find substantial register effects in BSL. Texts that were intended to inform the audience had a higher incidence of mouthing (77% of signs) than narrative texts (50%). Child-directed signing also showed increased mouthing (79%), an interesting finding in and of itself because children would be expected to know less about the phonological structure of the surrounding spoken language, yet mouthings appear more often in the signing that is directed towards them. Since mouthings are part of the language input to which children are exposed, it is not surprising that young children also produce mouth gestures in their early speech – around the second year of life in ASL – before they have been exposed to English as part of their schooling (Reilly, McIntire, & Bellugi, 1990). As seen in BSL, register effects have also been documented in ASL. Zimmer (1989) finds that there is less use of non-manual articulations, including obligatory mouthings and mouth gestures, in very formal signing registers. Nadolske and Rosenstock (2007), however, find *more* mouthing in formal registers, but only for pronouns. Sixty percent of pronouns are mouthed when giving a lecture, but only 30% and 10% of pronouns are mouthed during conversations and storytelling, respectively. Overall, the storytelling register was associated with fewer mouthings than the other two signing registers (42% vs. 60%).

Education and other cultural factors have also been documented to affect the use of mouthing. In her study of Indopakistani Sign Language (IPSL), Zeshan (2001) finds that signers without any formal education tend to mouth considerable less than signers who have attended school. However, schooling did not reliably predict that a signer would mouth often; there is significant individual variation in the amount of mouthing among educated signers of India and Pakistan. Contact with other signers, though, did have a significant effect on the rate of mouthing, decreasing the propensity to mouth. In particular, women in these communities are often restricted from leaving the home in most circumstances, and these women tend to use more mouthings than men, who were more likely to have social contact with other signers. This effect was observed more often in Pakistani communities than Indian ones, presumably due to the more restrictive attitudes towards women in Pakistani culture. On the other hand, contact with other signers is particularly likely to induce greater rates of mouthing when the signers use different regional dialects of a sign language, such as the five varieties of DSGS (Boyes Braem, 2001). In these cases, mouthing a word in a shared alternate language, Swiss German, facilitates the comprehension of unfamiliar manual signs. Finally, as mentioned before, the rate of mouthing is often subject to great individual variability, and it is possible for a signer to barely produce mouthings at all (Hohenberger & Happ, 2001 for DGS; Zeshan, 2000 for IPSL).

Lastly, linguistic features of the utterance or target sign can influence whether mouthing will appear. Lexicographers of Sign Language of the Netherlands (NGT), for example, have noted that while a sign may be required to have mouthing in its citation form, once the sign is embedded in a sentence, the mouthing often does not appear, either replaced by a mouth gesture or completely absent (Schermer, 2001). This suggests that the added contextual information provided by the sentence reduces the necessity of mouthing, and that for these signs the mouthing primarily serves a disambiguating role. Finally, the grammatical class of a sign also interacts with the likelihood of mouthing accompanying that sign. Open class lexemes (nouns, verbs, adjective, adverbs) tolerate mouthings in DGS whereas close class lexemes “generally resist mouthings, especially when they are realized as bound morphemes” (Hohenberger & Happ, 2001, p. 165). Similar effects are note for ASL (Nadolske & Rosenstock, 2007). Sutton-Spence and Day (2001) also report low incident rates of mouthing in BSL for function words, though signers with hearing parents are significantly more likely to mouth these words as compared to signers with Deaf parents (62% vs. 38%). Inflected signs, which mostly consist of verbs, also tend to resist mouthing, both in ASL and DGS (Hohenberger & Happ, 2001 for DGS; Nadolske & Rosenstock, 2007 for ASL). These inflected forms often represent multiple arguments simultaneously in one manual articulation.[3] Because the surrounding spoken language usually must translate this complex predicate into multiple words, mouthings do not appear with these signs. IPSL, however, is a notable exception to this finding: Zeshan (2001) finds no effect of lexical class on mouthings, reporting equal rates of mouthing for arguments, predicates, and modifiers alike.

While most transcriptions of mouthings suggest that these words are phonologically and articulatorily identical to words spoken by the surrounding hearing community, closer examination reveals significant differences between speech and mouthing. For instance, Vogt-Svendsen (2001) reports that in Norwegian Sign Language VENTE is mouthed as "v(en)te". In this example the /t/ is produced with the tongue visible between the teeth; however, in spoken Norwegian, a /t/ would never surface with a dental place of articulation in this word. Moreover, Vogt-Svendsen posits that many vowels in her data do not seem to be realized as full-fledged segments. Instead, vowels may simply be represented as the transition between one consonant to the next. Along with the

---

[3] First noted in ASL, two types of verb inflections are frequently reported across sign languages: spatial and agreeing verbs (Padden, 1988). Spatial verbs begin at the spatial locus of the referent for the origin and end at the locus associated with the destination: NEW-YORK INDEX_a CALIFORNIA INDEX_b ${}_a\text{MOVE}_b$ means "(I) moved from New York to California." Agreeing verbs use referential loci of space as well, but they instead refer to subjects and objects: MARY INDEX_a JOHN INDEX_b ${}_a\text{ASK}_b$ means "Mary asks John." Both spatial and agreeing verbs use the spatial start- and end-point of verb to simultaneously incorporate its arguments. In this way, a single manual sign expresses three semantic units: the verb, and two arguments.

optional vowel, the VENTE example also shows that the nasal coda /n/ can be deleted in the mouthing. The deletion of segments from the phonology of a spoken word when the word is mouthed during the production of sign language is referred to as *phonological reduction*, and is a very frequent phenomenon across signed languages, e.g., mouthing "educa" with the manual sign EDUCATION in ASL (Davis, 1989; Lucas & Valli, 1989). Though some researchers have claimed that the form of phonological reductions is not standardized because the data contain too much individual variability to extract any reliable patterns (Schermer, 2001), most sign linguists report some basic trends in how mouthed words are reduced. A formal model that both accounts for different types of reductions described below and makes predictions across the entire lexicon, however, has yet to be developed.

The most commonly reported trend is that mouthed words often are reduced so as it adhere to a maximum word length (Ajello et al., 2001; Bergman & Wallin, 2001; Schermer, 2001; Zeshan, 2001). Mouthings in IPSL are always one or two syllables in length (Zeshan, 2001), and Swedish Sign Language mouthings tend to consist of just one mouth movement, either opening or closing (Bergman & Wallin, 2001). Similarly, in LIS reductions usually affect the phonemes that occur after the first tonic vowel, suggesting that stress, rather than syllable length, may have a significant role in how mouthed words are reduced (Ajello et al., 2001). It should be mentioned, however, that the pattern reported by Ajello et al. was not consistent, and other reduction patterns were observed. Bergman and Wallin (2001) also emphasis that the more visually salient phonemes, e.g., bilabials, will resist deletion during phonological reduction and the resultant mouthing will be shaped by linguistic pressure to fit the "native [sign] pattern" (p. 51), suggesting an interaction between the movements of the hands and mouth for reduced mouthings. Schermer (2001) also points out the potential for the manual component of the sign to impose restrictions on mouthings. For example, the NGT sign KOKEN usually is mouthed as “koko”. Because the spoken Dutch form and reduced mouthed form are both disyllabic, Schermer hypothesizes that the reduplication of the initial syllable is the result of the manual component of KOKEN, which is also a reduplicated disyllabic lexeme (i.e., made up of two distinct movements). Articulatory effects between the hand and the mouth for mouthings are corroborated by other accounts of such an interaction with mouth gestures (see Section 2.1 for a discussion of Woll’s concept of echo phonology) and in non-sign tasks, both linguistic and non-linguistic (see Section 1.2.1).

Patterns have also been observed with regard to which signs are more likely to trigger mouthing reductions. Frequency has been suggested to interact with reduction as the most frequent signs tend to be reduced, e.g., DSGS GEHÖRLOS “gehörlos” (“deaf”) reducing to “gelos” (Boyes Braem, 2001). Lexical class also plays a role, as reductions are more prevalent for verbs than nouns in DSGS, with the mouthing of verbs only appearing with simple inflections (e.g., first person, infinitive, past participle) regardless of the inflection of the manual sign or how the word would be inflected in the surrounding spoken language. Mouthings of ASL also tend to delete bound morphemes such as the past tense marker (Lucas & Valli, 1989). In contrast, grammatically appropriate inflections have been observed in the mouthings of IPSL verbs, though only among signers who lost their hearing late in life or still retain some level of residual hearing (Zeshan, 2001). Among the profoundly deaf and early-deafened signers, IPSL mouthings bear minimal inflection. Likewise, Boyes Braem (2001) also hypothesizes that early signers will phonologically reduce their mouthings more than late signers across the lexicon, however the video resolution of her signing data of DSGS was too low in quality to provide any evidence to support this claim. More sensitive recording techniques, such as motion capture, would be well suited to address this hypothesis regarding the exact nature of phonology reductions of mouthings in sign language.

Broadening the treatment of mouthings to look at their role in an utterance as a whole, several unique combinations of signs with mouthings are possible. The most prevalent correspondence by far between manual signs and mouthings is a one-to-one relationship: 42%–47% of mouthings in DSGS exist as a single mouthed word co-occurring with a single sign (Boyes Braem, 2001). But other combinations do exist. One mouthing may be produced across multiple signs, a situation referred to as *stretching* or *spreading* (Bank, Crasborn, & van Hout, 2013; Boyes Braem, 2001; Crasborn, van der Kooij, Waters, Woll, & Mesch, 2008; Hohenberger & Happ, 2001; Schermer, 2001). Early signers of DSGS tend to stretch their mouthings more than late signers (49% vs. 25%), and early signers in particular, accomplish this stretching by lengthening the vowels rather than breaking up the stretched mouthing into its component syllables (Boyes Braem, 2001). Most often, the stretched mouthing is a direct translation of one of the manual signs, but its timing domain is independent of the manual articulation so that it lasts for multiple manual signs (Hohenberger & Happ, 2001). For example, a common BSL phrase consists of the manual signs DEAF YOU with the mouthing "deaffffff" lasting the entire utterance (per Hohenberger & Happ, 2001). Here, the hands and mouth articulate distinct lemmas, with different phonological forms and timing domains. As the hands and mouth appear to behave independently from one another, these instances of stretched mouthings beg the question of which articulator is primary in the multimodal utterance. Hohenberger and Happ (2001) conclude that the hands are still considered the "baseline" of signing because the mouthing component means little without the accompanying manual signs. Hohenberger and Happ also identify another, less common, type of stretched mouthing in which the mouthed component is not a direct translation of any manual sign of the utterance. Specifically, they report the DGS sentence with the mouthing "breakfast" stretched over the manual signs EAT DRINK-COFFEE EAT. In this example, the mouthing serves to provide global information about the discourse topic while the manual signs relate more specific information. But again, even though the mouthing may represent a "higher" discourse level, its semantic content is dependent on the manual signs, as "breakfast" by itself communicates very little, underscoring the primacy of the hands over the mouth. Bank et al. (2013) suggest that stretching in NGT may be instigated by a mismatch in the durations of the manual and oral components as stretched mouthings are long than non-stretched mouthings and sign sources of stretching are shorter than signs without stretched mouthing. Stretching has not been attested in ASL, and it seems unlikely that signers of ASL would exhibit stretching as in the DGS "breakfast" example. However, it is possible that mouthings may stretch over neighboring indexicals[4] to bind prosodic constituents (see below).

[4] In signed languages, indexicals are pointing gestures to establish referents or refer back to them. Indexicals may point directly at a present referent (IX_2 "you") or to an arbitrary region of empty space to establish a locus associated with a non-present referent (JOHN IX_a, MARY IX_b, IX_a LIKE IX_b "John likes Mary").

The inverse correspondence also occurs: multiple mouthings may appear during a single manual sign. In DSGS, this situation is much more characteristic of late learners' signing than of early signers (26% vs. 6%; Boyes Braem, 2001). These kinds of mouthing patterns may be more common for late learners because they are also more likely to mouth function words of the surrounding spoken language. Because these function words either do not have corresponding manual signs or the manual signs are not usually used in most registers of sign language communication, there are more words to mouth than manual signs, so each manual sign may be mouthed with multiple words. Early signers, in contrast, tend to use multiple mouthings with a single sign only for certain set German phrases, such as *nicht schlimm* ("not bad") or *am Abend* ("in the evening") (Boyes Braem, 2001).

It is also possible for mouthings to occur without any manual articulation at all. In BSL, only about one percent of all mouthings occur without a sign, usually produced by either young signers (younger than 30 years old) or Deaf signers with hearing parents (Sutton-Spence & Day, 2001). Mouthings without a manual component are more frequent in IPSL than in other sign language, and are most likely to be used when there is no manual sign in the lexicon for the semantic target, e.g., proper names and kinship terms (Zeshan, 2001). Mouthings with no accompanying manual sign are also typically reported in ASL, but only for contact signing (Lucas & Valli, 1989; Lucas, 1992). This is the result of the influence of English on the sign production. The signer mouths grammatically full and correct English sentences; however, because ASL does not contain most of the function words that English has, no manual sign can be produce while the signer is mouthing them. Schermer (2001) reports similar findings for NGT.

Finally, mouthings and signs may appear in a one-to-one relationship but with the oral and manual lexemes attaching out of order. Examples include signing QUIET REMAIN in DSGS while mouthing “remain quiet” (Boyes Braem, 2001) or, in DGS, MUTTER VATER with “vater mutter” (“father mother”) (Hohenberger & Happ, 2001). It is not clear why these misaligned sign–mouthing pairs occur, though at least in one of these two examples (MUTTER VATER), it seems that the misalignment was a slip, and it was immediately corrected. It does suggest that the mouthing and manual components of sign production are retrieved independently, but the evidence is quite limited. In a related example from BSL, the sign for *foot and mouth disease* is realized as a point to the mouth and then down towards the feet, but the mouthing matches the English order of “foot and mouth” (Woll, personal communication). In this case, it does not appear that the lexemes foot and mouth are retrieved separately for the signed and mouthed components; rather, the two body parts are likely selected as an amalgam in each language, but the phonotactics of BSL prefers signs with higher locations to precede those with lower locations within a phonological phrase. This hypothesis, however, is purely conjecture. Because these types of sign–mouthing inversions are quite rare – though not completely unattested – it is difficult to draw many conclusions.

Finally, I shall turn to the question of the function of mouthings. Certainly, this is an important question as mouthings expressing information on the mouth that is usually already encompassed by the manual sign. Indeed, the majority of mouthings are semantically redundant (e.g., Hohenberger & Happ, 2001 for DGS). In DSGS, early signers’ mouthings are proportionally more redundant than late signers (74% vs. 51%) (Boyes Braem, 2001). The mouthings of late signers more often are the sole bearer of meaning for two reasons. First, late learners of DSGS mouth full spoken language sentences more often than early signers, so lexical items such as function words do not co-occur with a manual sign (for a similar case in ASL, see Marschark, LePoutre, & Bement, 1998). Second, these late learners are also more likely to substitute a semantically equivalent mouthing for a sign, even if that mouthing has a signed translation, e.g., signing SAY and mouthing "report". Even though the two components are close in meaning, the signer could have matched the sign with the mouthing: either mouthing "say" or signing REPORT (Boyes Braem, 2001). Because late learners have smaller sign vocabularies than native signers – Boyes Braem reports that in the first four minutes of signing, late learners used an average of 99 different signs whereas early learners used 135 – late learners are more likely to make a substitution error in the selection of a manual sign while still mouthing their intended lexical target.

In other cases, both native and late signers exploit a difference in the mouthed and signed components to supplement the meaning of a sign or fill in a lexical gap in the sign language. Schermer (2001) finds that mouthing can be used in this way in NGT, and Ajello et al. (2001) note this effect of using mouthing to specify hyponyms in LIS as well. For example, the broad semantic field represented by the manual sign FISH can be specified with the mouthing "tuna" or POLITICIAN with the mouthing of the proper name "Andreotti". IPSL is even more lenient in these types of constructions, allowing adjectival modifiers mouthings to accompany signs: FRIEND "best" and CLOTHES "clean" (Zeshan, 2001). Interestingly, however, IPSL does not permit substitution of homonymic mouthings. The sign BAD must always be mouthed as "bad", even when the intended meaning is *wicked* or *evil*, despite these words existing spoken Hindi and Urdu. As mentioned above, mouthing is an especially important method of lexical borrowing in signed languages that do not have a robust manual alphabet system. For languages that do exhibit prominent use of a manual alphabet, like ASL, these kinds of specifying mouthings would be predicted to be less common, but still present in the language. For instance, the ASL HOPE is mouthed as "o" to disambiguate it from EXPECT, as the two signs share the same manual specification (Marschark et al., 1998). However, the type of hyponym specification that occurs in NGT and LIS does not typically happen in ASL; instead, the manual alphabet would be used to fingerspell the more specific English word.

For all the functions discussed so far – disambiguating, specifying, and filling lexical gaps – mouthing can be described as method to code-switch and borrow lexemes from the surrounding spoken language (cf. Brentari, 2001). The mouthings of late learners of sign language are considered particularly indicative of code-switching, as their signing and mouthing shows a greater influence from the surrounding spoken language (Boyes Braem, 2001 for DSGS; Lucas, 1992 and Lucas & Valli, 1989 for ASL). Moreover, Deaf communities embedded in multilingual communities have an even larger lexical source from which to borrow. As English is spoken and understood by a significant proportion of Indian society, IPSL also borrows mouthings from the English language to accompany a sign, e.g., HEARING mouthed with the English "normal" (Zeshan, 2001).

In addition to mouthing being used to borrow from spoken language, Boyes Braem (2001) identifies four other lexical functions of mouthing in DSGS. Mouthings may behave like adverbs (e.g., signing GOOD and mouthing "very"), be the sole grammatical marker of negation (i.e., mouthing "no" to negate the manual sign), denote the nominal in a classifier construction (e.g., mouthing the agent “man” while depicting a person classifier walking uphill), and indicate a referent of possession (e.g., signing MY and mouthing “father’). Mouthing in ASL does not appear to be used for any of these functions described by Boyes Braem. Mouthing may also be used as a strategy in contact signing. Contact signing occurs when interlocutors do not share the same native language. It may be that one or both of the interlocutors is hearing or a late learner of sign (e.g., Lucas, 1992; Lucas & Valli, 1989 for ASL), or as discussed above, when signers use different dialects of the same sign, as often happens in DSGS (Boyes Braem, 2001). When signing with people from neighboring regions, DSGS signers increase the amount of mouthings they produce in order to facilitate communication.

Finally, mouthing can serve stylistic and discourse functions as well. In DSGS, referents are more likely to be mouthed when first introduced in the discourse, and mouthing has also been argued to be used to emphasis or stress a sign (Boyes Braem, 2001). Prosodically, mouthings may also be stretched across several manual signs in order to bind constituents into a unit, e.g., a noun with a classifier, a noun and an index, or a noun with affixes (Boyes Braem, 2001 for DSGS; Bank, Crasborn, & van Hout, 2013 and Crasborn, van der Kooij, Waters, Woll, & Mesch, 2008 for NGT; Sandler, 1999 for Israeli Sign Language).

Mouthings are also used to denote reported speech in DSGS, indicating that the reconstructed communication was oral in nature (Boyes Braem, 2001). In these cases, the mouthing patterns tend to appear slightly differently: there is less stretching and more instances of multiple mouthings occurring during a single manual sign. In these respects, this type of mouthing resembles the way late learners mouth.

In summary, ASL exhibits many of the properties of mouthing that have been attested in other signed languages of the world. Though ASL researchers had historically relegated mouthings as only occurring in contact signing situations, it is now accepted that mouthings play a role in the production of native ASL (Lucas, 1992; Nadolske & Rosenstock, 2007). As with IPSL and the various sign languages of Europe, mouthing in ASL serves multiple purposes. While mouthings primarily are redundant in nature, i.e., the mouthed and manual components refer to the same lexeme, mouthings may also disambiguate manual homophones and elements of hyponyms that have no unique manual sign. ASL also is subject to effects of grammatical class and register on the production of mouthing, and the mouthings of ASL also phonologically reduce, just as in other sign languages reviewed here. However, mouthing in ASL is different in certain ways than it is in other languages. ASL has one of the lowest rates of mouthing among established sign languages of the western world, and ASL's rich fingerspelling system subsumes some of the functions mouthing plays in other languages, such as denoting proper names and kinship terms. Finally, since sign linguists of ASL have spent less attention on mouthing than other sign languages have received, it remains an open question whether certain phenomena occur with ASL mouthings. For example, stretching and multiple mouthings on a manual sign have not been reported for ASL, but there is no reason to suspect that ASL would resist these constructions while they are permitted in European sign languages. Nor is it known if lexical frequency plays a role in the production of mouthings in ASL. Only with further research into the use of the mouth in the production of ASL will these questions be answered.

1.1.2. Mouth gestures

Mouth gestures, bearing no relationship to spoken language, are categorized by their function in the sign language. Ajello and colleagues (Ajello et al., 2001) differentiate LIS mouth gestures based on their grammatical status. Lexical mouth gestures carry no meaning of their own and appear when they disambiguate manual homophones, like mouthings also do, or are otherwise obligatorily encoded in the phonological representation of the sign. ASL, while not using mouth gestures to disambiguate, does have signs that require a mouth gesture to be considered well-formed: for example, SUCCESS must by accompanied with an opening mouth gesture, frequently transcribed as "pah!" In contrast to lexical mouth gestures, other mouth gestures are more like bound morphemes and add additional semantic information to the sign when produced in conjunction with the manual component. In LIS these kinds of mouth gestures can intensify the base sign, add modal information (usually emotional: self-satisfaction, astonishment, sorrow, delusion), or comment on the "actionality" of a verb (e.g., if a situation is impossible because of a temporary hindrance or a permanent impediment). Ajello et al. also mention a third type of mouth gestures, semantic reinforcement, but provide no explanation of their role.

Woll (2001) also divides BLS mouth gestures into three categories – adverbials, enactions, and echoes – which appear to fall in line with Ajello et al. (2001), but she bases these categories on form and semantic function rather than grammatical role. Adverbial mouth gestures modify the manual sign by adding semantic information regarding manner and degree via a metaphorical link between the aperture and shape of the lips (for adverbial mouth gestures in ASL, see Liddell, 1980). Enactions are mimetic gestures in which the mouth depicts a literal mouth, with its movement and form faithfully mapping onto the representation of a mouth in the discourse, e.g., the manual sign CHEW while the mouth reenacts the precise nature of the chewing action (Woll, 2001). Finally, echoes (also referred to as echo phonology), unique to Woll's classification system, are obligatory mouth gestures in which the shape of the mouth "mirrors" the movement of the hands (see Section 2.1). Like the lexical mouth gestures of Ajello et al. (2001), echoes are lexically specified in the citation form of the sign.

Although both Woll (2001) and Ajello et al. (2001) report on the variety of functions of mouth gestures, there is very little data on the distribution of these forms and the environments in which they occur. Mouthing is a frequent occurrence in the sign languages of many Deaf communities, however mouth gestures appear in a much smaller portion of data sets (e.g., Ebbinghaus & Hessmann, 2001 for DGS). Furthermore, Sutton-Spence and Day (2001) observe that BSL mouth gestures more commonly attach to verbs and pronouns than nouns and adjectives, but this pattern is poorly understood. Finally, accounts of the acquisition of ASL suggest that the hands and mouth are decoupled in the early stages of learning adverbial mouth gestures (Anderson & Reilly, 1998). Although young children make few errors in the production of mouth gestures, the majority (83%) of the errors they do make are misalignments in the temporal domain of the mouth and hands, so that the start- and end-points of the two articulators do not match. This observation suggests that lexical signs and adverbial mouth gestures are separate components of the language, and children must learn to coordinate the two articulators to produce fluent ASL.

Whereas most accounts of mouth gestures discussed above treat these mouth gestures as an embedded component of the grammar of sign languages, other linguists instead view mouth gestures as the sign language homologue of manual co-speech gestures (Fontana, 2008; Sandler, 2009). Though mouth gestures occasionally are required to disambiguate homophonous manual signs (e.g., ASL LATE VS. NOT-YET), Fontanta argues this does not refute the status of mouth gestures as co-linguistic gestures because the same situation is observed in spoken language. Particularly with spatial language, co-speech gestures can be important components of conveying meaning in spoken language discourse, rendering the auditory component incomplete without the gesture. For example, the phrase, "You left the television remote in *there*," with no other contextualizing information in the conversation, depends on a deictic point to a location (e.g., the refrigerator) for the interlocutor to make sense of the statement. Sandler (2009), however, avoids this issue by only analyzing iconic mouth gestures as being the sign analogue of co-speech gesture. For Sandler, adverbial and disambiguating mouth gestures are part of the grammar of Israeli Sign Language, while iconic gestures are holistic and spontaneously created by the signer. Similarly, Fontana argues that mouth gestures are more like co-speech gestures than a phonological feature of a manual sign because mouth gestures contain no segmentable units and are meaningless without the manual sign to which they attach. Given this evidence, Fontana (2008) proposes that mouth gestures – along with silent mouthings and voiced mouthings – exist on a gestural continuum comparable to Kendon's (1988) continuum of co-speech gestures (gesticulation, pantomime, and emblems) seen in spoken language.

1.1.3. Mouth actions: further considerations

In the following section I will address several question that relate to both mouthings and mouth gestures at once. First, how are mouth actions represented and accessed in sign language production? Second, do the hands hierarchically dominate the mouth in the production of mouth actions? Third, are mouthings distinct from mouth gestures, or can mouth actions be treated one phenomenon? And finally, what limitations have traditional methodologies used in the study of mouth action placed on the interpretation of the data?

1.1.3.1. Representation of mouth actions in signed languages

The phenomenon of mouth actions in sign language production raises new questions in the nature of human language representation and processing, given that most models of language production are based solely on spoken language data. However, mouthings, mouth gestures, and other non-manual articulators, coupled with manual signs, demonstrate the highly multidimensional and simultaneous nature of sign production (see Ebbinghaus & Hessmann, 2001).

Recent experimental work suggests that, at least for mouthings, the oral and manual components of a sign are represented and retrieved separately (Vinson, Thompson, Skinner, Fox, & Vigliocco, 2010). In this study, BSL signers named pictures with signs or performed an English-to-BSL translation task (i.e., producing the sign translation when presented with the written English word). An analysis of the error data showed that slips of the manual component and slips of the oral component of a single sign did not always co-occur. That is, often a signer would produce the correct sign while mouthing a different word or would produce the correct mouthing with an erroneous manual sign. If the sign lexeme were specified with both the manual and mouthing components bundled together, these two components should always slip together in the incorrect trials. Therefore, Vinson and colleagues interpret this result as evidence that mouthings and manual signs are stored and retrieved separately. Presumably, mouthings and manual signs are stored separately because of their phonological differences.

Though the degree to which deaf individuals represent the phonology of English is not fully understood, this experiment provides some limited insight into the relationship between mouthing in sign production and the retrieval of English. More errors in the production manual signs were recorded in testing blocks in which all the test items were semantically related (the semantic similarity effect) for both the picture-naming and the translation task; however, mouthings only showed a semantic similarity effect in the picture-naming task, suggesting that the English orthography provided resistance to the mouthing component. That is, the presence of the written English word was easily transformed to a mouth articulation, and performance did not suffer when the written words were semantically related within a block. This further suggests that BSL signers have a strong sublexical mapping between orthography and mouth articulations. It is possible that this mapping between orthography and mouth articulations also involves a link between orthography and phonology and then another link between phonology and oral articulation, but the data do not speak to this possibility. Regardless, this orthography-to-articulation mapping is sufficient to allow the signer to mouth a written word without accessing its lexical representation, thereby avoiding any semantic interference effects in the task. However, this is not to say that signers never accessed the

English word's lexical representation to complete the task. Indeed, in order to produce the manual sign translation, the signer must access the semantics of the English word. This is why a semantic similarity effect was still seen for the manual signs of the translation task. Nonetheless, the mouthing component of the sign appears not to rely on lexical access for this task, and instead makes use of a direct pathway between orthography and articulation.

The findings reported by Vinson et al. (2010) have significant ramifications to language production models, as mouthing and signing doubles the processing load as compared to spoken language. Hohenberger and Happ (2001) describe the cognitive burden:

> If we take a model-theoretical view on language production as is done in contemporary cognitive psycholinguistics…we immediately become aware of the extravagance of simultaneous signing and mouthing, purely on psycholinguistic grounds. In no other language use do we find a comparable mixed input/output and thus a doubly charged language comprehension and production device. (p. 169)

Nevertheless, recent work by Emmorey, Petrich, and Gollan (2010) suggests that multimodal language production does not result in a processing cost. Bimodal bilinguals, who are native signers of ASL and speakers of English, were no slower to name pictures in "code-blended" English and ASL than they were to name the pictures using just ASL. It was observed, however, that during the code-blend trials, subjects waited for their hands to get into position before beginning the speech. Even though the spoken word had been retrieved, these bimodal bilinguals waited for the hands to catch up so that both the sign and speech could be produced at the same time. This suggests that while there is no cognitive cost to code-blending, there is an articulatory coordination cost. Presumably, these same costs apply to unvoiced mouthings during ASL production as well. Further research is required in order to better understand the cognitive demands of sign language production, motor execution, and comprehension.

1.1.3.2. Are the hands the head of the mouth?

In many of the analyses of mouth actions, the question arises regarding the interaction of the hands and mouth. Because spoken language, from which most linguistic theories are grounded, only involves one articulatory system, the multimodal nature of sign production raises the question of the hierarchical relationship between manual and oral articulators. Most sign linguists argue that the hands dominate the mouth, often summarized by the maxim, “the hands are the head of the mouth” (Ajello et al., 2001; Boyes Braem, 2001; Hohenberger & Happ, 2001; Keller, 2001; Vogt-Svendsen, 2001; Zeshan, 2001). Other researchers propose that the two modalities are autonomous and neither one dominates the other (Ebbinghaus & Hessmann, 2001; Vinson et al., 2010). Currently, no one claims that the mouth dominates the hands.

Evidence in support of the hands dominating the mouth in sign language production primarily comes from observations of mouth actions altering their formational structure to match the structure of the manual components of sign. Vogt-Svendsen (2001) reports instances of mouth gestures in Norwegian Sign Language reduplicating when the manual sign contains repetition, mouthings being temporally shortened or extended to coincide with the temporal start- and end-points of the manual sign (see also Ajello et al., 2001), matching in the number of movements of the hands and mouth, and effects of the intensity and rhythm of a sign on the production of mouth actions. The stretching of mouthings over multiple manual signs has also been interpreted as the mouth conforming to the parameters of the hands as these mouth actions are still defined by the prosody of the manual utterance (Bank et al., 2013; Boyes Braem, 2001; Crasborn et al., 2008; Sandler, 1999) and are semantically dependent on the manual signs as well (Hohenberger & Happ, 2001). The behavior of phonological reductions also underscores the dependence of the mouth on the hands. As these reduced mouthings, especially reduced uninflected verbs, do not contain enough phonological information to be well-formed, they are only permissible because they are dominated by the simultaneous manual sign (Hohenberger & Happ, 2001). Zeshan (2001) also notes that mouthings are subservient to manual signs because although mouthings may modify the sign, signs never modify the mouthed element. (However, mouth actions can be produced without any accompanying manual sign, though these cases are exceptional. For instance, in ASL only three such mouth gestures are attested [Marschark, LePoutre, & Bement, 1998].) Finally, Keller (2001) proposes that the mouth is dependent on the hands and mouth actions must conform to the kinematic properties of the hands (see Sections 2.3 and 2.4).

On the other hand, Ebbinghaus and Hessmann (2001) are adamant in their argument that mouth actions are independent linguistic structures that are produced in tandem with manual signs without being dominated by the mouth. The strongest evidence to support this claim comes from the behavior of mouth gestures. One mouth gesture can have a variety of meanings dependent on the manual sign to which it attaches. For example, the DGS mouth gesture “ff” can mean either *relaxed* or *continuously*, depending on which manual sign with which it co-occurs. Similar patterns have been reported for ASL, with the “mm” mouth gesture behaving much like the DGS “ff”

(Liddell, 1980). The mouth gesture "mm" can mean roughly *relaxed enjoyment* when accompanying the manual verb FISH, but the same gesture means *normal and proper* with the verb ENGINE-RUNNING. Additionally, in DGS and ASL, one manual sign does not consistently occur with just one mouth gesture. Many different mouth actions may appear with a lexical sign with no predictable pattern. In ASL, the manual utterance LAST-NIGHT I READ can appear with several different mouth gestures, each one contributing a different meaning to the sentence. With the "cs" gesture (mouth cinched to one side and head tilt to ipsilateral shoulder), it would mean *I just read (it) last night;* with the "th" gesture (tongue protruding between the teeth), it would mean *Last night I was mindlessly reading (but didn't really pay attention to the material);* and with the "mm" gesture (protruding but unpursed lips), the sentence would mean *Last night I leisurely read (and enjoyed the experience)*. Crucial to Ebbinghaus and Hessmann, there is nothing in the manual structure of the sign that requires one mouth gesture over the others. Each gesture contributes its own meaning to the signed construction and is chosen as the signer's discretion. Therefore, the authors argue that mouth gestures have a semantic domain that does not align with the semantic domain of lexical signs because the two systems are independent of each other. That is, the semantic content of manual signs and mouth gestures do not have a one-to-one, one-to-many, or many-to-one mapping; rather, many mouth gestures may augment the meaning of many lexical signs. Furthermore, mouth actions, though infrequent, can be produced without any corresponding manual sign: IPSL allows mouthings to be the sole bearer of kinship terms (Zeshan, 2001), and ASL allows for mouth gestures to stand alone (Marschark et al., 1998). Mouth actions could only stand alone if they are independent channels of sign communication rather than dependencies of the hands. Finally, new empirical evidence from errors observed in BSL suggests that mouthings and manual signs are retrieved separately during sign production (Vinson et al., 2010). If these two components are stored and retrieved independently from each other, it suggests that mouth actions and manual signs are modular and independent linguistic components of sign language.

1.1.3.3. Are mouthings and mouth gestures the same?

Another important question is whether mouth gestures and mouthings are fundamentally the same phenomenon or are linguistically distinct. For the purposes of many analyses, researchers do not find any significant differences in form or function between mouthings and mouth gestures (Keller, 2001; Schermer, 2001; Vogt-Svendsen, 2001). Keller (2001) formalizes this equivalence by stating that all mouth actions in DGS can be best described by their kinematic properties (see Section 2.3), and from this perspective any differences between mouthings and mouth gestures disappear. Schermer (2001) also reports in her NGT corpus that reduced mouthings are indistinguishable from mouth gestures, though there was one instance of a functional difference between the two mouth actions to specify the grammatical class of a lexical item. Specifically, the type of mouth action may disambiguate noun–verb pairs on occasion, with the noun being mouthed whereas the corresponding verb is produced with a mouth gesture. Vogt-Svendsen (2001), while conceding that mouthings and mouth gestures of Norwegian Sign Language differ in their etymology, attitudes, and distribution, argues that through a perceptual analysis of these two mouth actions, mouthings and mouth gestures are similar in their form, function, and interaction with the hands.

Other linguists make either an implicit or explicit distinction between mouthings and mouth gestures (Ebbinghaus & Hessmann, 2001; Hohenberger & Happ, 2001; Sutton-Spence & Day, 2001). Sutton-Spence and Day (2001), among others, use separate coding systems to transcribe mouthings and mouth gestures in BSL: mouthings are written out in the orthography of English whereas mouth gestures are coded with a system of their own design that distinguishes mouth configurations based on their visual and articulatory properties (e.g., mouth open or cheeks puffed). Though the authors do not provide an explanation as to why they chose to code these two mouth actions differently, this methodology presupposes that mouthings and mouth gestures are fundamentally distinct phenomena. Ebbinghaus and Hessmann (2001) argue that mouthings, mouth gestures, and manual signs of DGS are three separate and autonomous dimensions of communication in sign language, each having its own mode of articulation and contributing overlapping, but non-identical meanings to the sign utterance. The same evidence used to refute the notion that the hands dominate the mouth can be applied here to support the idea that mouthings and mouth gestures are distinct processes (see Section 1.1.3.2).

Lastly, Hohenberger and Happ (2001) propose that mouth gestures and mouthings are not the same phenomenon because mouth gestures dominate mouthings in the data of DGS. For instance, even in utterances in which most signs are produced with mouthing, signs that use mouth gestures block the mouthing of that manual sign. Additionally, non-manual facial markers, such as pursed lips to mark the subjective mood and counterfactual clauses also prevent mouthings to appear during the entire prosodic domain of the non-manual marker. Even in the extremely rare cases in which mouthings and mouth gestures are both produced for one sign, the mouth gesture temporally aligns with the manual sign while the mouthing is produced later with no accompanying sign. These examples illustrate how mouth gestures are distinguished from mouthings in the grammar of DGS as mouth gestures take precedence over mouthings.

As for ASL, it remains unclear whether mouth gestures and mouthings are the same phenomenon. While most researchers who have studied the mouth in ASL treat mouth gestures separately from mouthings, more data, especially concerning the form of phonological reductions of mouthings, are needed to be able to draw better comparisons between the two types of mouth actions. Functionally, mouthings and mouth gestures have been reported to serve different, though somewhat overlapping roles. While mouth gestures are primarily discussed in the ASL literature as only being adverbials, mouthings are described as being redundant or disambiguating representations of spoken English. However, both mouthings and mouth gestures can appear as obligatory non-manual markers that disambiguate manual homophones in ASL. Regardless of the functions of the two types of mouth actions, it is my prediction that both mouth gestures and mouthings are subject to the same kinematic principles during their articulation, and it is this line of inquiry that needs further exploration.

1.1.3.4. Methodological pitfalls and new techniques

Though research, especially that done in the past decade, has made huge strides in describing and analyzing the mouth actions observed in sign language production, almost every study discussed so far – with the notable exception of Vinson (2010) – make several assumptions regarding the nature of mouth actions, leading to problems of data collection and coding, ultimately contaminating the results and any conclusions inferred therein. Referring primarily to mouthings, Keller (2001) argues that mouth actions are fundamentally kinematic in nature, reflecting the "patterns of articulatory actions that are prominent in visual perception of voiced speech" (p. 191). Although studies of the speech of deaf individuals suggest that their phonological knowledge of spoken language is significantly different than that of hearing speakers (see Section 4.1), sign linguists continue to base their analysis of mouthings and mouth gestures using the phonological system of the surrounding spoken language. This method of coding implicitly assumes that deaf and hearing individuals share the same phonological repertoire, though all evidence suggests otherwise.

Keller (2001) identifies two main issues with the standard methodologies used to study mouth actions in sign language. The first issue, the segmental-phonology bias, is the assumption that mouthings are composed of phonological segments in the same way that speech is organized. However, there is no evidence that this is the case for mouthings. Given that the majority of mouthings are produced without any audible sound, it is impossible to accurately identify embedded segments from the movement of the mouth alone. Instead, researchers introduce their own biases of what the supposed segments are, based on the translation of the manual sign. Although this bias *potentially* would not have a significant impact in the analysis of mouthings in late deafened signers and sign-native hearing children of Deaf adults (CODAs), the transcription of prelingually deafened signers will suffer the most because they have not developed a native-like mapping between mouth articulations and their acoustic consequences (see Guenther & Vladusich, 2012; Guenther, 1995; Tourville & Guenther, 2011 for details of this mapping process), unlike the other two groups of signers. Moreover, analyzing mouthings phonologically appears to be inappropriate because no phonological rules seem to apply to mouthings. Despite more than twenty years of mouthing research, especially prevalent among European sign linguists, Keller states that there have been no rule-governed processes that apply to the phonological representation of mouthings. Phonological reduction in particular would be a fertile area of exploration of phonological rules, yet researchers consistently struggle to provide a compelling segmental account of reduced mouthings. In contrast, spoken language production is characterized by pervasive phonological effects resulting from the interaction of neighboring segments, thus supporting the psychological reality of segments in spoken language. The lack of any phonological rules discovered that govern the production of mouthing suggests that mouthings may not be composed of segments at all (Keller, 2001). Additionally, patterns seen in the "phonological" reductions of mouthings present another challenge to the segmental assumption of mouth actions. While several researchers find that the beginning portion of a word is more likely to be preserved than the end, no reliable rules accurately predict which supposed segments would be deleted. Furthermore, examples like the ASL sign FINISH being mouthed as "fsh" are particularly challenging to explain by a purely phonological account; what principle would motivate the deletion of two medial segments while preserving the segments near the left and right edges of the word?

Second, Keller (2001) argues that the segmental bias for analyzing mouthings highlights another important issue: the problem of transcribing mouth actions using the orthography of speech. Because the people who transcribe these data see the manual component of the signed utterance and are familiar with the phonological system and orthographic conventions of the surrounding spoken language, it is easy to make assumptions about what the mouthed word should be and therefore what must have been faithfully produced by the signer. Therefore, orthographic transcriptions merely represent how speakers might produce the same word, rather that providing an accurate account of how signers actually move their mouths. For example, Vogt-Svendsen (2001) provides one example of how a mouthed /t/ was articulated in a mouthing of Norwegian Sign Language that would be considered incorrectly articulated in spoken Norwegian. Without her explicitly describing this difference, the orthographic transcription would imply that the mouthed word looked just like the spoken word. Nonetheless, almost every sign language linguist uses alphabetic orthographies to code for mouthings, though Bergman and Wallin (2001) are a notable exception. A better research practice, therefore, would include having another coder who does not know the sign or surrounding spoken language perform a second analysis of the data. This would ensure that the transcription reflects what is visually present in the data without introducing the coder's own linguistic biases into the analysis. Additionally, transcriptions of mouthings could follow Bergman and Wallin's example and code the actions of the mouth in gestural or perceptual terms rather than phonologically.

This problem of transcription is not unique to the study of signed languages. Spoken language research has also encountered the problem of listener bias affecting the identification and coding of spoken data. For example, covert contrasts frequently reported in child acquisition data illustrates how two distinct classes may be miscategorized as one, based merely on perceptual impressions. Macken and Barton

(1980) discovered that while children seemed to produce all stops as voiced tokens, acoustic analysis revealed that these children did indeed produce a reliable VOT contrast between voiced and voiceless target. This contrast was masked, however, because both of the children's VOT lengths were within adults' VOT target range for voiced stops. Examples like this one underscore the necessity to objectively quantify linguistic data rather than analyzing categorical data based on human judgments. In this way, kinematic analysis, as measured by a motion capture system, can be viewed as the sign language methodology that is analogous to the spectrogram in spoken language research. Although IPA transcriptions are adequate for a gross representation of a spoken word, any sublexical analysis of spoken data is best done using quantitative methods. Investigations into the structure of mouthings and mouth gestures should also strive to quantify their data rather than relying on the categorical nature of orthographic transcriptions.

Additionally, orthographic transcription may give the illusion of articulatory differences that do not actually exist in the mouthing data. Many phonemes, though acoustically distinct, are visibly indistinguishable on the mouth (see Section 4.1 for a discussion on visemes). For example, /p, b, m/ all look the same in the absence of voicing because they all involve a bilabial closure. The articulatory features that distinguish these three sounds – the relative timing of voice onset, lowering of the velum, and nasal release of air – are not visible to an interlocutor. However, when a signer mouths the word "banana", for example, an orthographic transcription assumes these invisible articulatory actions occurred, as the transcriber does not record "manana" or "panana". But in reality, the transcriber does not have the requisite information to distinguish between these possibilities, but they are never considered. The use of orthographic transcription creates the appearance of a level of articulatory accuracy on par with hearing speakers without verifying if mouthings are truly so precise. Again, evidence from deaf speech would suggest otherwise. Moreover, the coding of these examples with phonological units rather than visemes obscures the similarity between different mouthings, which may provide insight into the nature of certain processes like mouthing reductions. In fact, where Keller (2001) claims there are no phonological rules that apply to mouthing data, viseme-based rules may emerge.

Finally, while almost every transcription system codes mouthings orthographically, sign linguists differ on whether they transcribe mouth gestures using alphabetical approximations or visual–gestural scores. For researchers who use different coding systems for mouthings and mouth gestures (e.g., Sutton-Spence & Day, 2001), this has the disadvantage of implying that these two types of mouth actions are categorically different in form, though a kinematic analysis may find little difference between the two. The implication is that mouthings are easily mapped onto a phonological form whereas mouth gestures are more difficult to transcribe orthographically. Before such a distinction is made, evidence must be gathered to justify coding mouthings and mouth gestures by different means. I would argue, instead, that mouthings, without any knowledge of the surrounding spoken language, would be just as difficult to code orthographically. Given the inherent limitations that traditional methods of coding and analyzing mouth actions impose on the data, Keller (2001) therefore proposes a kinematic investigation of the mouthings and mouth gestures of Deaf signers. Analyzing mouthings and mouth gestures in terms of the articulators' motion in space (as measured by distances, velocities, accelerations, etc.) would avoid the issues of segmental and transcription biases and would make meaningful contributions to the motoric processes involved in mouthing and its so-called phonological reductions. However, in the more than ten years since Keller first made this proposal, very few studies have examined the kinematic properties of sign language production, and none so far have investigated the mouth as a sign articulator.

1.2. Interactions Between the Hands and Mouth: Articulatory and Neurological Considerations

Though linguistics has traditionally addressed language production in terms of speech, that is the motoric actions and acoustic consequences of the vocal tract, linguists have begun to include the gestures of the arms and hands as extra- or para-linguistic information. Together with the speech stream, these two simultaneous pathways form a multimodal and complete picture of language production. New models of language production are therefore needed to incorporate both the hands and the mouth as articulators and explain the behavioral and neurological interactions between the two. As sign language has already enjoyed a long history of multimodal analysis, any insights relating to the neural structure and mechanics of these links between the hands and mouth will be of particular value to the burgeoning field of sign language “speech” science and phonetics.

1.2.1. Kinematics of the coordination of the hands and mouth

1.2.1.1. Reaching with arbitrary mouth movements

Kinematic studies of the hand and mouth of human participants have begun to investigate the precise nature of how these anatomically disparate effectors interact with each other both in their underlying neural substrates and their physical actions. Based on the hypothesis that the links between the hand and the mouth are the result of common motor commands used for feeding, many studies find strong kinematic effects between these two effectors in grasping tasks. Without involving language production at all, simply grasping with one effector while making a similar aperture change in the other effector will cause the kinematics both effectors to be affected by other factors of the task in the same way (Gentilucci, Benuzzi, Gangitano, & Grimaldi, 2001). In one experiment of this study, participants were told to open their mouths by an arbitrary but consistent degree while grasping large or small objects (rectangular blocks) with their hand (requiring power and precision grasps, respectively). Unsurprisingly, maximal finger aperture, the greatest distance between the thumb and index finger, is larger for trials with the large object; however, maximal mouth aperture, the greatest distance between the upper and lower lips, is also larger for manual grasps towards the larger object. This motoric connection between the mouth and the hands is also present when the mouth is the effector that grasps the target object. In a second experiment, the participants grabbed suspended bread cubes, small and large, while simultaneously opening their fingers from a closed position. Here too, oral grasps towards the larger bread cubes evoke greater maximal apertures of not only the mouth, but the hand as well. Gentilucci and colleagues refer to the evolution of eating behaviors to explain these motoric connections: the physical properties, such as size, of a wanted food item will affect both the manual and oral movements. The motor command for reaching a food item will send commands to the hands so the item is grasped properly as well as send commands to the mouth, as it will soon have to open accordingly to ingest the food. When reaching for an apple, as compared to a cherry, these motor commands must encode for a larger aperture grasp for both the hand that retrieves the fruit and the mouth that must open wide enough to fit it between the upper and lower teeth to bite into it, a type of grasp. Bolstering these claims that grasping actions signal this common motor command, control experiments failed to find a connection between the hands and mouth when no grasping task was required of the participants. Presenting the object blocks of different sizes without requiring them to be grabbed or altering the distance of the target objects, but not their size, so the kinematics of the reach component like velocity and acceleration were increased for the farther object, did not have any affect on the aperture of the mouth. Similarly, experiments in which the participants grasp bread cubes but the fingers of the hand spread (abduct) rather than open or they open but while the arm is held behind the participant's back also fail to show any kinematic effect of the mouth on the hands. These control experiments indicate that it is the grasping action, rather than the object's properties or any movement of the hand or arm, that cause the hand and mouth to share kinematic features. Moreover, the grasp must have the potential to achieve its target to be controlled by this shared motor command; when the arm is in a position that physically precludes it from reaching its target, as it is when held behind the back, then the mouth no longer exhibits a connection with the "grasping" hand.

1.2.1.2. Reaching with speech

Other studies have moved one step closer to language by observing similar effects between the hands and mouth when uttering nonsense syllables, e.g., /ba/ (Gentilucci et al., 2001; Gentilucci, Santunione, Roy, & Stefanini, 2004; Gentilucci, Stefanini, Roy, & Santunione, 2004; Gentilucci, 2003). In these experiments, not only are the kinematics of the mouth and hands analyzed as effectors, but also spectral analyses of the acoustic consequence of the syllable production have been interpreted too to further demonstrate a link between the hands and the mouth in grasping tasks. Specifically, analyses of the first two resonant formants of speech suggest that F1 and F2 tend to correlate to the size of the object of a grasping action. F1 increases when grasping a large object compared to a small one. The authors attribute the increase of F1 to an increase in jaw aperture – which was never measured directly – and conclude that reaching for and grasping larger food will require a larger jaw opening when the item is subsequently ingested. F2, in a similar manner, increases with larger objects; however, this formant is more like to increase in frequency when a food item is brought to the mouth rather than the previous action of reaching for and grasping the food item. The authors attribute the increase in F2 to the forward displacement of the tongue – again without directly observing or measuring the tongue's position itself – and conclude that this displacement aligns with the necessary coordination of oral muscles required for ingestion and mastication of the food, though further details to support this claim are lacking (Gentilucci & Corballis, 2006).

Though the results from acoustic analysis of nonsense speech have led the authors to make some provocative claims, these conclusions should be evaluated with caution. The acoustic effects found for F1 and F2 are interpreted in articulatory terms. Though much work has been done in English phonetics, with corroboration in Italian, the language used in the above experiments, that relate articulatory gestures to their acoustic outcomes, there is also a large body of work that demonstrates that there is not a one-to-one correspondence between the two, e.g., quantal effects (Stevens, 1989) and motor equivalence (e.g., Brunner et al., 2011). A variety of different articulatory strategies may be employed across speakers with the equivalent acoustic results. It is therefore premature to claim that increases in F1 absolutely correspond to lower jaw positions or that increases in F2 absolutely correspond to more fronted tongue displacements. As these studies primarily investigate the kinematics of gestures – linguistic utterance and manual grasps – conclusions regarding jaw height and tongue position should be corroborated by direct observation rather than indirect inference. Moreover, the conclusions about tongue displacement during food-to-mouth actions are made without any literature support or logical explanation of the phenomenon. There is no reason why a larger piece of food imminently entering the mouth would induce the tongue to displace farther forward in the oral cavity than a smaller one. If such manual movements and the corresponding increase in F2 are the result of shared neural circuitry used in feeding movements, this detail needs elucidation.

1.2.1.3. Gesturing with speech

Looking further into the role that the hands may play in speech production, another set of studies investigated if hands gesturing during single word production can have any effect on the mouth without any grasping actions involved. Indeed, simultaneous speech and emblematic gestures production affects both effectors: for the mouth, both F2 and pitch increase, and for the hands, the gestures decrease in duration, maximal height, and number of velocity peaks (Bernardis & Gentilucci, 2006). These results are seen if the participants produce the speech and gesture themselves or if they watch an actor perform them instead. Importantly though, these effects only occur if both the spoken word and gesture are meaningful and co-referential. If either the spoken word is meaningless, i.e, a pseudoword, or the gesture is meaningless, then no kinematic or acoustic effects are observed. Similar results have been found in a repetitive transcranial magnetic stimulation (rTMS) paradigm (Gentilucci, Bernardis, Crisi, & Volta, 2006). When producing words with emblematic co-speech gestures that shared the same meaning as the spoken word, F2 increased. Then when Broca's area was induced with a "functional lesion" via rTMS pulses, this effect disappeared. However, when the spoken word and manual gesture did not share the same meaning, though both were meaningful in and of themselves, no difference was observed between the stimulation of Broca's area as compared to the other stimulation conditions. Together, these two studies indicate that manual gestures may activate a neural system that connects the hands and mouth, but this can only occur when the hands and mouth both produce meaningful and co-referential gestures. Bernardis and Gentilucci (2006) conclude that "gestures reinforce speech, whereas words inhibit gesture" (p. 12).

1.2.2. Directionality of the hand–mouth interaction

A final consideration with regard to the linkage between the mouth and the hands during the tasks described above is the directionality of the effects. Can the kinematics of the mouth effect changes in the kinematics of the hands just as easily as the hands can affect the mouth, or does the interaction only occur in one direction? Several pieces of evidence suggest that the hands exert greater influence on the kinematics of the mouth than vice versa. The rTMS results suggest that Broca's area is responsible for translating such aspects of the arm gesture like goal and intent into mouth gestures of words (Gentilucci et al., 2006). That is, movements of the arm are mapped onto the movements of the mouth. Similar conclusion are reached in grasping tasks: because the effects are driven by the physical size of the object being grasped, the hand's kinematics are influenced by the outside world, which in turn, acting through an intermediary step, the kinematics of the mouth are altered to match the gestures of the grasping hand (Gentilucci et al., 2001). In an experiment in which subjects were asked to synchronize repetitive utterances of a nonsense syllable to the tapping of their own finger, Smith and colleagues (1986) found that requiring the participant to alternate between large and small finger taps resulted in alternation of speech intensity of the spoken syllable; however, requiring the participants to alternate between loud and quite utterances of the target syllable did not cause any alternations of the finger tapping amplitude. Taken together, it appears that the relationship between the hands and mouth is that the mouth conforms to match the kinematics of the hands rather than reciprocal interactions between the two effectors. Similar conclusions are found in the interaction between simultaneously produced mouth actions and manual signs in the sign language literature: the mouth movements follow the movements of the hands (e.g., echo phonology). There is, however, one case in which the hands follow the kinematics of the mouth: when the mouth reaches to bite bread cubes (Gentilucci et al., 2001). It is notable that for this task the mouth is the effector that makes the grasping action, in contrast to all other tasks outlined above, and that the hands are not needed to bring the food to the mouth. As much of the grasping literature has been described in models of feeding behavior, grasping with the mouth without involving a grasp with the hands circumvents the prototypical action sequence, yet the hands show evidence that they are primed to execute a grasping action nonetheless. Further research is needed to understand why such motor commands are executed when they would serve no purpose (the bread has already reached the mouth so the hands are not needed), considering that grasping with the mouth does not affect other non-grasping movements of the hand.

In sum, the kinematics of the hands and mouth have been shown to affect one another under certain conditions, namely when grasping an object or producing speech and a manual gesture simultaneously. In addition, these same effects are present when one observes such scenarios but does not execute them oneself. These interactions within the domain of spoken language remain to be observed in sign language production. As sign language makes use of multiple effectors as simultaneous articulators, such as the hands and mouth, investigating sign language production could provide useful insights into the kinematic and neural links between the hand and the mouth. Of particular interest is the coordination of oral and manual movements during fingerspelling with accompanying mouthing. As the model drawn from spoken language data supports the idea that the hands affect the movements of the mouth, this hypothesis could be tested in an instance where the sequential handshapes of fingerspelling serve as a representation of spoken language orthography, while spoken language phonology is simultaneous mimicked on the mouth of the signer. Additionally, in all the experiments discussed from the spoken language literature, the spoken utterances and manual gestures have been monosyllabic (for gestures this means having only one principal movement). However, in the case of fingerspelling in ASL, these words are polysyllabic. This provides the opportunity to expand upon the findings presented in the spoken language research to better understand how movements are coordinated in their timing and kinematics when the task demands are more complex and naturalistic.

1.3. Motivation for Current Work

In this section I have explored the variety of ways that deaf people use their mouths when producing sign language. In addition, I have considered how the mouth interacts with other effectors, with special emphasis on the links observed between the hands and mouth. These links are evident both in the general coordination of the two body parts in non-linguistic contexts of hearing individuals and in the multimodal production of sign language among Deaf individuals.

While the literatures discussed here all have helped inform the nature of the mouth and its interaction with the hands during sign language production, further work is still required in order to make more specific conclusions. In particular, the theoretical frameworks and methodologies that have long been used in spoken language research need to be applied to the study of sign language production. As motion capture technology gains traction in the sign linguistics community, a greater understanding of the phonetic and kinematic principles that guide sign production is sure to follow. Currently, the field of sign language phonetics is especially sparse, largely due to methodological limitations. Just as speech science has placed a growing emphasis on quantitative methods over impressionistic judgments and categorization, sign language research must embrace new methods of inquiry and analysis. Techniques like motion capture provide the ability to measure the articulators of sign language with fine-grained temporal and spatial resolution, and with this increased precision, subtle articulatory effects and contrasts will be discovered.

To conclude this section, I pose a series of questions that lie at the intersection of the ideas discussed above, which the framework of "speech" motor control and the technique of motion capture are well suited to answer. It is the goal of this dissertation to address these questions. First, it remains unclear what the relationship is between the mouthings of sign language production and the speech of hearing individuals. Although qualitative evidence suggests that mouthings are not faithful reproductions of the mouth movements involved in speech production, no model exists to account for the occurrence and articulatory form of these differences. Exactly how do mouthings differ from speech? What phonological properties of the utterance or kinematic features of the articulators drive the differences between mouthing and speech? What role, if any, do the visuoperceptual features of speech (i.e., visemes) play in the acquisition and production of mouthing? How consistent are the movements of mouthings across tokens in comparison to speech?

Second, while it is known in the spoken language literature that the hands exert articulatory pressure on the movements of the mouth, as it has been suspected in sign language, the precise nature of such an interaction has not been empirically demonstrated. Is the mouth as affected by the hands during the production of mouthings as it is for the production of mouth gestures? Do mouthings adhere to the temporal and kinematic characteristics that are taken from speech, or do the movements and timing of the mouth more closely match those of the signing hands? Though the answers to these questions are not yet known, we have the tools available to us to begin to explore these issues regarding the mouth as an articulator in the production of sign language.

Chapter 2: Coordination of the Hand and Mouth for Fingerspelled Nonce Words (Experiment 1)

2.1. Introduction

The production of sign language requires the integration of multiple articulators (e.g., the two hands, mouth, eyebrows), but it is unclear how the motor system coordinates the simultaneous movements of a signer's body. Though research from co-speech gesture may be brought to bear on signing, gesture researchers predominantly employ coarse-grained analyses that do not consider the contribution and coordination of individual oral and manual movements; rather, these studies are more concerned with identifying each gesture's "lexical affiliate," the spoken word or phrase that the gesture refers to or modifies (e.g., Wagner, Malisz, & Kopp, 2014). Moreover, co-speech gestures tend to contain fewer and less complex movements than signs (Kita, van Gijn, & van der Hulst, 1998), limiting the application of gesture results to the type of movement concatenations necessary for sign language production.

In the domain of grasping and reaching studies, there have been some experiments conducted on non-signers that are relevant to the coordination of the hand and mouth in sign language. Specifically, the size of a hand grasp towards an object affects the size of the lip opening gesture of a simultaneously uttered syllable. Reaches and grasps for larger objects not only require larger maximal hand apertures but also result in larger oral movements for utterances spoken at the same time (e.g., Gentilucci, Benuzzi, Gangitano, & Grimaldi, 2001). This result suggests that whatever coordinative effects signers exhibit for sign language production may reflect general principles that underlie the human motor system rather than being the product of specific experience with a signed language. The proposed principle that the "hands are the head of the mouth" and therefore oral movements are influenced by manual movements, appear to apply to signers and non-signers alike (Boyes Braem & Sutton-Spence, 2001).

The phenomenon of echo phonology also provides insights into coordinative processes for sign language production. It was observed that for the set of meaningless, obligatory mouth gestures which accompany some signs, the oral movement must "mirror" the movement of the hands by copying a movement feature (Woll & Sieratzki, 1998; Woll, 2001). This conclusion comes from the observation that signs in which the hands open, the mouth also tends to open, and vice versa. For example, the ASL sign SUCCEED[5] is accompanied by an obligatory mouth gesture often transcribed as "pah!" Woll argues that the feature of the hands moving away from the body is "mirrored" by the opening mouth gesture; thus the same manual sign with a closing mouth gesture would be considered ill-formed. Indeed, it is even difficult to produce ill-formed combinations like SUCCEED with a closing mouth gesture. Woll codes these mouth

---

[5] SUCCEED is signed with each hand in the 1 handshape and the tips of the forefingers near the corners of the mouth. The hands twist outward as they simultaneously move out and away from the mouth. The obligatory mouth gesture resembles a bilabial plosive resembling the production of the sound "pah!"

gestures as being either CV (opening) or VC (closing) syllable structures. Echoes with CV syllables are associated with exhalation, hand opening, movement away from the body and hand separation; VC echoes are associated with inhalation, hand closing, movement towards the body, and the hands moving towards each other. Using the LM model proposed by Sandler (1989), Woll labels locations as being equivalent to consonants in the syllable while movements are more like vowels. Following this reasoning, the more sonorous (i.e., proximal movements because they are the most visually salient) manual components of the sign attach to equally sonorant mouth gesture (i.e., the "vowel" or open position). This match between the sonority of the hands and mouth is a motivating force underpinning the echo phenomenon, which drives the two articulators to match the form of their movements (Woll, 2001). This finding suggests that perhaps all hand and mouth movement combinations are subject to the same congruency constraint, regardless from which domain of the lexicon a sign or non-manual feature originates, but this has yet to be confirmed.

The present experiment assesses the degree of coordination present in a sign language production task. Rather than investigating mouth gestures produced with lexical signs (see Section 1.1.2), this study investigates fingerspelled items produced with mouthings because these movements are more complex than those that occur in mouth gestures and most lexical signs. As with co-speech gestures, the majority of lexical signs are produced with just one path movement (Coulter, 1982). Likewise, mouth gestures are typically produced as a static mouth posture or as one movement (e.g., Anderson & Reilly, 1998). Mouthings and fingerspellings, however, frequently require multiple articulatory movements because they are derived from the surrounding spoken language.

Nonce words were created for the experiment so the number and direction of movements across the hands and mouth could be directly manipulated in order to determine how these factors influence the degree of coordination between the two articulators. I hypothesize that coordination effects are sensitive to the absolute number of articulatory movements of an utterance and whether the hands and mouth share the same number of movements. Therefore, I predict that fingerpelled items with fewer movements will exhibit greater coordination than items with more movements and items with the hands and mouth moving the same number of times will exhibit greater coordination than items with mismatched number of movements. Finally, I hypothesize that the congruency of the direction of hand and mouth movements will affect coordinative efforts. Items for which the hands and mouth move in a similar fashion should exhibit greater coordination than items in which the two articulators move in opposite directions.

2.2. Methods

2.2.1. Participants

Ten Deaf signers participated in this experiment (2 male, 8 female). All of the participants were native signers of American Sign Language and were profoundly deaf, reporting a hearing loss of 90 dB HL or greater in their better ear. All participants were right-hand dominant signers and reported no history of cognitive or language impairments. The participants had a mean age of 30.3 years with a standard deviation of 9.3 years.

2.2.2. Materials

The stimuli were selected to create a diverse array of hand-mouth movement combinations. For fingerspelling, hand movements were defined by the transition between one handshape to the next during a fingerspelling sequence. Letters that have inherent movement to them, such as the tracing path movements of -Z- or the orientation change of -H-, were not used to create the set of nonce words. Expected mouth movements derived from the phonological structure of the item. The nucleus was taken to be the most open point of the mouthed syllable. Therefore, the transition from an onset – simple or complex – to a nucleus was marked as an opening movement, and the transition from a nucleus to a coda was marked as a closing movement. A list of these items and their features are detailed in Table 2.1. Two items were created for each possible one- and two-movement combination represented in the stimuli set. There are no combinations for two mouth movements co-occurring with one hand movement, however, because such a combination is impossible given the orthography-to-phonology mapping of English.

Background questionnaire – Background information regarding participants' history and use of speech and speechreading was collected prior to test. From the questionnaire's 11 responses, a single composite speech score was computed using the formula

$$composite = \sum_i z_i \times w_i ,$$

where i is the index for each question, z is the z-score transformation, and w is the weight. Higher scores indicate that the participant was more familiar with speech based on their formal speech instruction and day-to-day use of voiced speech and speechreading. The questionnaire is reproduced in Appendix 1 along with each item's weight for the composite score calculation.

Table 2.1: Experiment 1 stimulus list

			Hand Movements		Mouth Movements		Same	Same
Item	Pronunciation	English guideword	Number	Direction	Number	Direction	Number?	Direction?
o-s	/ɑs/	Oscar	1	close	1	close	yes	yes
u-n	/ʌn/	under	1	close	1	close	yes	yes
b-a	/bɑ/	bar	1	close	1	open	yes	no
v-a	/vɑ/	Nevada	1	close	1	open	yes	no
n-u	/nʌ/	numb	1	open	1	open	yes	yes
s-i	/si/	casino	1	open	1	open	yes	yes
a-b	/æb/	absent	1	open	1	close	yes	no
o-b	/ɑb/	obscene	1	open	1	close	yes	no
u-m-b	/ʌm/	thumb	2	close-open	1	close	no	-
u-s-k	/ʌsk/	tusk	2	close-open	1	close	no	-
k-n-u	/nʌ/	knuckle	2	close-open	1	open	no	-
s-c-a	/skɑ/	scar	2	open-close	1	open	no	-
s-l-a	/slɑ/	Slavic	2	open-close	1	open	no	-
a-l-s	/ælz/	pals	2	open-close	1	close	no	-
o-b-s	/əbs/	obscene	2	open-close	1	close	no	-
o-s-o	/o.so/	so-so	2	close-open	2	close-open	yes	yes
o-t-o	/o.to/	photo	2	close-open	2	close-open	yes	yes
b-a-b	/bæb/	babble	2	close-open	2	open-close	yes	no
v-o-l	/vɑl/	volume	2	close-open	2	open-close	yes	no
o-b-o	/o.bo/	oboe	2	open-close	2	close-open	yes	no
o-l-a	/o.lə/	cola	2	open-close	2	close-open	yes	no
s-u-s	/suʃ/	sushi	2	open-close	2	open-close	yes	yes
t-u-t	/tʌt/	stutter	2	open-close	2	open-close	yes	yes

2.2.3. Procedure

Participants were told that they would see a series of letters projected onto a screen in front of them, and they were instructed to fingerspell and mouth the item when prompted by a head nod from the experimenter, repeating the task until prompted again to stop. Each item was presented along with a real English word as a guide to pronunciation. For example, the screen would display "V-O-L as in V-O-L-U-M-E" with the letters V, O, and L marked in red. The participants were instructed only to fingerspell and mouth the red portion; the full word "volume" only served to provide a contextual guide to the desired pronunciation. The participants repeated each item 12 times before moving to the next item. If a participant made an error, such as forgetting to mouth at all or mouthing the full guideword, the experimenter interrupted the trial, gave feedback to the participant, and re-started the trial from the beginning. If the participant's pronunciation deviated from the guide pronunciation enough to cause a different movement pattern of the lips (e.g., /nu/ instead of /nə/ resulting in significant lip rounding at the expense of lip opening), then the trial stopped, and the experimenter demonstrated the desired pronunciation. If the participant repeated a pronunciation error three times in a row, the item was thrown out.

2.2.4. Data capture

The data was recorded using both a traditional digital video camera as well as an optical motion capture system. The video was collected with a Sony HD HDV Handycam® HDR-HC9 camcorder, recording 720×480 pixels at 29.97 frames per second. Motion capture data was recorded using an 8-camera Hawk Digital RealTime System from Motion Analysis Corporation. The motion capture system calculates the three-dimensional position for every passive marker at 120 frames per second. Participants wore a total of 10 markers, small reflective spheres (2 mm in diameter) that were attached to the skin with double-sided adhesive tape. A rigid body device consisting of 4 markers attached to wooden dowels fixed in a cross arrangement was attached to the participant's forehead and served as reference markers to later control for head translation and rotation. Three additional markers were placed on the face: one on the vermillion border of the upper lip, one on the vermillion border of the lower lip, and one on the chin at the left mental tubercle. Finally, three markers were place on the right hand: one on the radial side of the distal interphalangeal joint of the index finger, one on the ulnar side of the distal interphalangeal joint of the pinky finger, and one in the middle of the palmar wrist crease (Figure 2.1).

Figure 2.1: Marker placement for data capture. Participants wore a total of ten reflective markers: four on a rigid body attached to the forehead, two on the lips, one on the chip, one on the index finger, one on the pinky finger, and one on the wrist.

2.2.5. Data post-processing

Using Cortex (Motion Analysis Corp.), the first step in post-processing the raw movement signal was to go frame-by-frame to ensure that all ten markers were correctly labeled, as the software occasionally mistakes an illusory point for a marker of interest. Next, data gaps were addressed, as the data capture may lose track of a marker for a few frames. If a gap was found with fewer than 30 consecutive frames missing (250 ms at 120 fps), a cubic spline algorithm was used to interpolate the missing data points. Gaps longer than 30 frames were left unfilled, as the algorithm becomes less reliable as length increases. Next, the movement signal underwent a Butterworth filter (F_{LP} = 6 Hz) to smooth out noise in the data collection. Finally, the first 10 stable repetitions of the item –

repetitions with no unfilled data gaps, disfluencies made by the participant, or any other movement anomaly – were identified and exported as separate files.

Movement traces for the mouth and hand apertures were derived as the Euclidean distance between the upper and lower lip markers and the distance between the wrist and index finger markers, respectively. In the case of item S-I, the hand movement signal was instead measured as the Euclidean distance between the wrist and pinky finger markers. Using SMASH (Speech Movement Analysis and Spatial Histograms; Green, Wang, & Wilson, 2013), a set of custom algorithms for Matlab, the beginning and end points for each repetition were delineated. The duration of a production was defined as the smallest time window that included the complete gestures of both the hand and mouth. (Table 2.1 lists these gestures for each item.) An opening gesture begins at a local minimum and ends at a local maximum in the movement trace; a closing gesture begins at a local maximum and ends at a local minimum.

2.2.6. Missing data

A total of 317 out of 2300 total repetitions were excluded from the final analysis (13.8%). Due to experimenter error, 10 repetitions were not recorded of the item B-A-B for one participant. An additional 5 repetitions were recorded of one participant's item N-U, but the movement trace revealed no discernable mouth movements. For 31 repetitions across three participants, portions of the signal were missing data for the hand, and the data gaps could not be filled using the algorithm described above. These instances were largely due to the hand moving out of the capture window because the hand gesture was completed, but the mouth was still moving. Finally, 271 repetitions across eight participants were excluded because the mouth movements did not conform to the gesture pattern predicted in Table 2.1. For example, 70 repetitions (all 10 repetition for each of seven participants) of the item V-O-L /vɑl/ were excluded because the signers produced only an opening gesture instead of the expected opening-closing gesture. This is likely a result of velarization of the /l/ in the coda position, which eliminates the need to close the jaw for a coronal articulation. Across six participants, an additional 120 of the 271 non-conforming repetitions were produced with an erroneous rounded vowel, in which the protrusion action of the lips reduced the vertical movement of the mouth to such an extent that the measurement of the opening and closing gestures was obscured.

2.3. Results

Graphs for the results from each item (averaged across all 10 participants) are given in Appendix 2. Based on visual inspection, the main qualitative result presented in these graphs is that the two articulators match each other in their overall timing. Even for the items for which the two articulator movements are the most distinct, i.e., when the hand moves twice while the mouth only moves once (e.g., S-L-A "slɑ"), the mouth movement stretches out to temporally align its start- and end-points with the beginning and end of the hand movements (see Figure B.5).

Based on the derived aperture signals, the coordination between the hand and mouth movement gestures were assessed using a cross-correlation analysis. The cross-correlation analysis returns not only the Pearson product-moment correlation coefficient (hereafter *r* or correlation coefficient) for the two signals as they actually appear in time, but it also time-shifts one signal against the other and calculates *r*-values for all possible time-shifts (also called lags). For signals A and B, both of length N, the cross-correlation returns a 2N-1 × 2 matrix; the first column contains the lag values, from −N+1 to N−1, and the second column contains the correlation coefficient for that lag. The maximum correlation coefficient and its associated lag are then taken to be the optimal interpretation of the relationship between signals A and B. In other words, the cross-correlation function reports two measures: what amount of time-shifting is needed to produce the best fit between the two signals (*lag*) and what the correlation coefficient would be for such a fit (*r*).

The correlation coefficient, *r*, is a metric of spatial coordination, of how well the shape of one signal matches the shape of second signal. In these data, the correlation coefficient specifically quantifies how much the opening and/or closing of the mouth matches the opening and/or closing of the hands. Values close to 1 indicate strong spatial coordination, values close to 0 indicate little to no spatial coordination, and values close to −1 indicate an anti-correlation.

Lag is a metric of temporal coordination, of how much delay there is between one signal and the coordinated response of the second signal. Lag values close to 0 indicate synchronization between the two signals, and therefore strong temporal coordination, with coordination decreasing the father away the lag value is from 0 (lags can be positive or negative).

For this study, the cross-correlation analysis was performed in Matlab using custom written scripts. The mean of each aperture signal of the hand and the mouth was first subtracted out in order to normalize for the size difference between the two articulators (this is mathematically equivalent to the cross-covariance). Because the two movement signals could either be positively or negatively correlated, such as when the two articulators move in opposite directions, the maximum absolute *r* value was selected for each cross-correlation. Lag values were also converted into absolute values because only the general degree of temporal alignment was of interest rather than determining which articulator preceded the other.

Analysis of spatial coordination, as measured by the correlation coefficient, *r*, revealed that items in which the hand and mouth both required the same number of movements were more coordinated than items with different number of movements across articulators: $r = .828$ vs. $r = .729$, $t(9) = 6.5$, $p < .001$ (see Figure 2.2). For the items with the same number of movements between the two articulators, a 2 × 2 (1 movement vs. 2 movements of the hand and mouth; same vs. different direction of movements between the hand and mouth) ANOVA was conducted to assess the effect of both the number and direction of movements on spatial coordination, *r*. The ANOVA revealed that items in which the hand and mouth both move once had a higher degree of spatial coordination than items in which the articulators both move twice: $r = .862$ vs. $r = .782$, $F(1, 9) = 56.0$, $p < .001$. The direction of movements did not affect the correlation values ($F(1, 9) = 1.7$, $p = .224$); items in which the hand and mouth both move in the same direction are no more coordinated than items in which the hand and mouth move in opposite directions. There was also no significant interaction between the direction of movement and the number of movements: $F(1, 9) < 1$, $p = .522$ (see Figure 2.3).

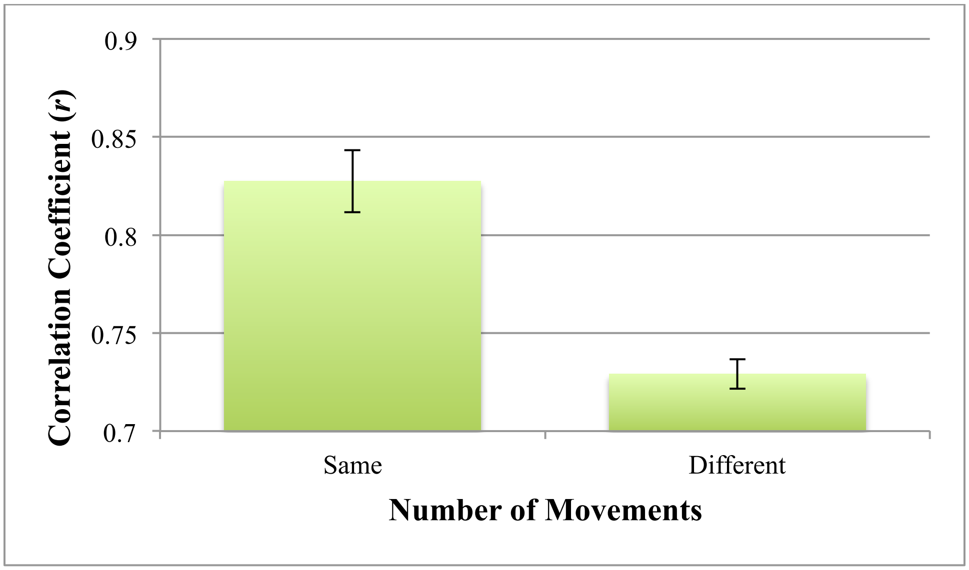


Figure 2.2: The spatial coordination for items in which the hand and mouth have the same or different number of movements. Error bars in this and all subsequent graphs reflect ±1 standard error of the mean.

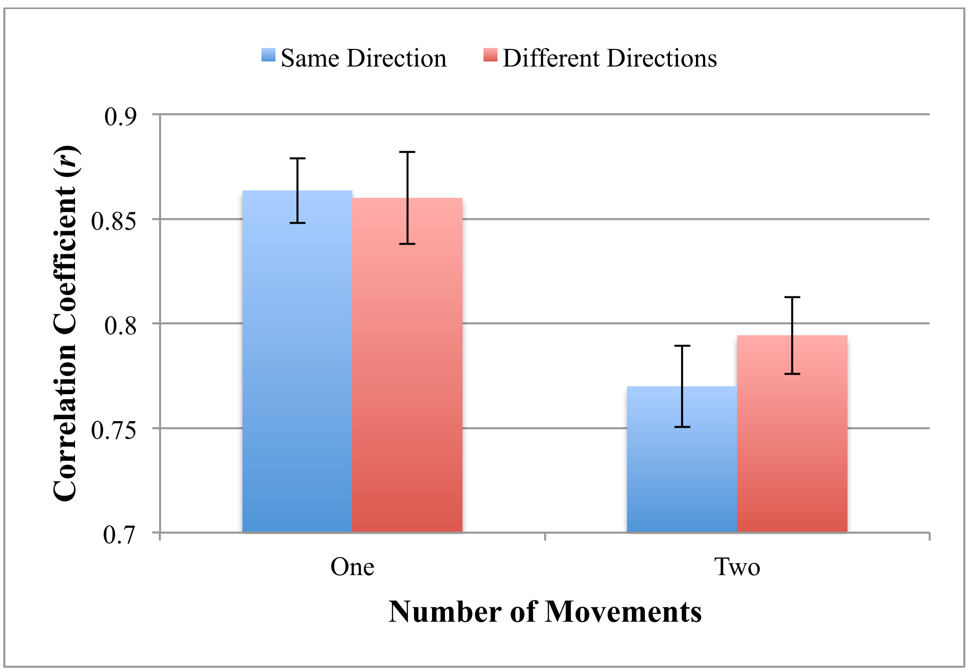


Figure 2.3: The spatial coordination by number and direction of movements. The data only reflect the subset of items in which the hand and mouth share the same number of movements.

Analysis of temporal coordination, as measured by the lag values, revealed the same pattern of results as the spatial coordination analysis. Items in which the hand and mouth both required the same number of movements were more coordinated (i.e., had smaller lags) than items with different number of movements across articulators: 57 ms vs. 106 ms, $t(9) = 6.4$, $p < .001$ (see Figure 2.4). Among the items with the same number of movements between the two articulators, a 2 × 2 (direction: same, different; number of movements: 1, 2) ANOVA was conducted to assess the effect of the number and direction of movements on spatial coordination. The ANOVA revealed that items in which the hand and mouth both move once had a higher degree of temporal coordination than items in which the articulators both move twice: 37 ms vs. 88 ms lag, $F(1, 9) = 26.6$, $p = .001$. The direction of movements did not affect the temporal coordination of the hand and mouth ($F(1, 9) = 1.1$, $p = .330$); items in which the hand and mouth both move in the same direction are no more coordinated than items in which the hand and mouth move in opposite directions. There was also no significant interaction between the number of movements and the direction of the movements: $F(1, 9) = 1.6$, $p = .243$ (see Figure 2.5).

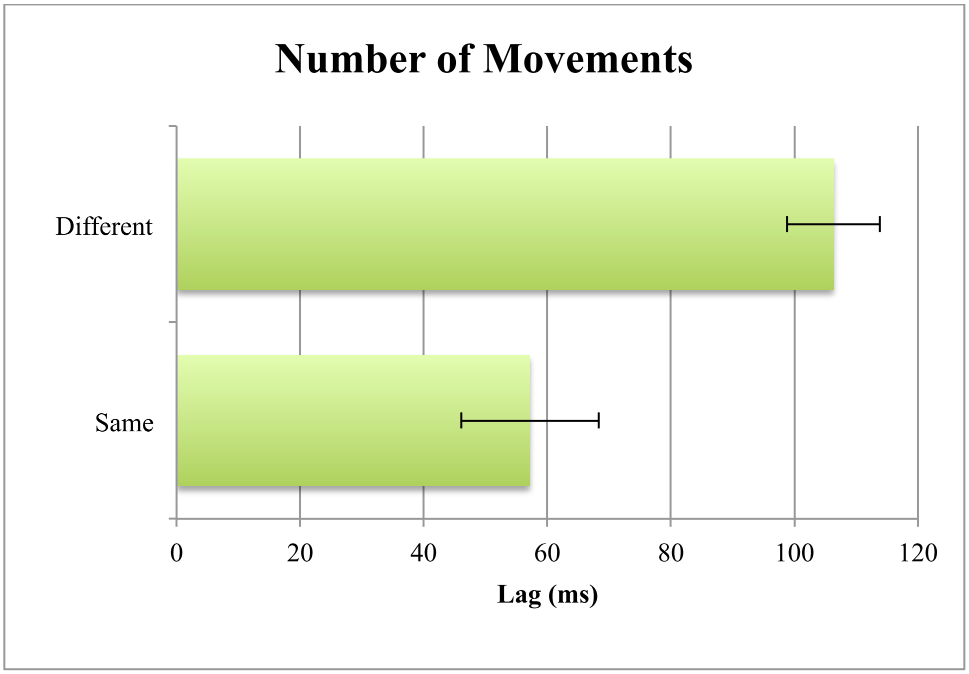


Figure 2.4: The temporal coordination by items in which the hand and mouth have the same or different number of movements

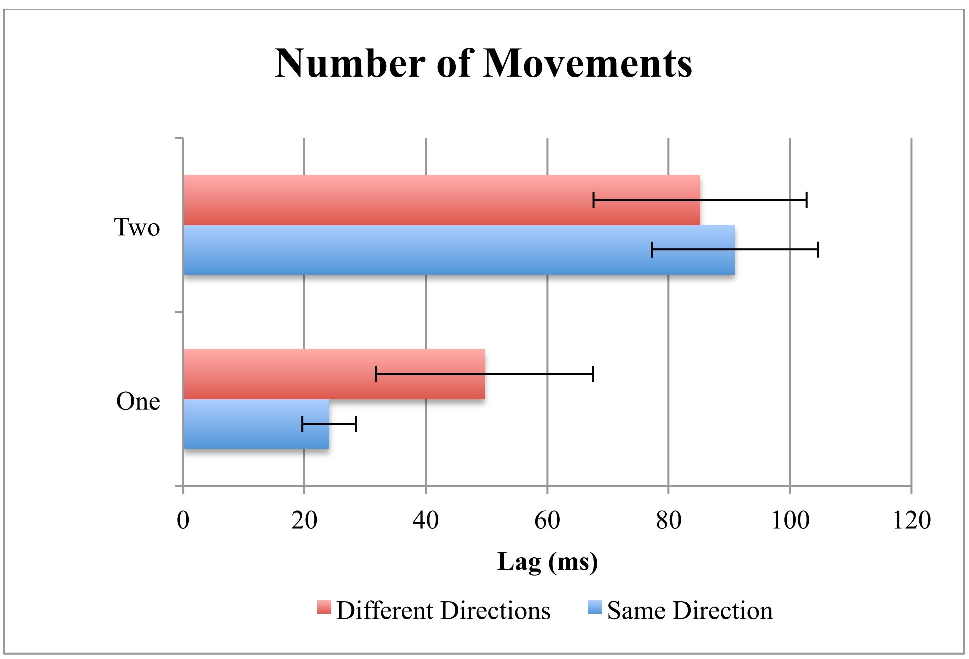


Figure 2.5: The temporal coordination by number and direction of movements. The data only reflect the subset of items in which the hand and mouth share the same number of movements.

Lastly, a regression analysis was performed to determine whether experience with speech affects a participants' ability to coordinate the movements of the hand and mouth. However, the analysis did not reveal any significant relationship between the composite speech score and the measures of spatial coordination ($r = .426$, $p = .22$) or temporal coordination ($r = .582$, $p = .077$). The regression between the composite speech score and the lag values approaches significance, and because high lag scores indicate reduced temporal coordination, this finding would indicate that better speech skills correlate with less temporally coupled productions. If this marginal result were indeed a real effect, it would suggest that participants who have more speech experience have a greater tolerance for temporal delays between sign articulators. However, the regression does exceed the standard .05 alpha level, so such an interpretation must be taken with caution. To test whether this regression reaches significance within particular levels of the number of movements factor and thus create a marginal effect overall, three additional regression analyses were performed, using the Holm-Bonferroni method to adjust for multiple comparisons. Regressions on the one movement level, two movement level, and the remaining cases in which the hand and mouth were mismatched in number of movements all failed to reach significance ($r = .583$, $p > \alpha_{.025}$; $r = .213$, $p > \alpha_{.05}$; and $r = .65$, $p > \alpha_{.017}$, respectively).

2.4. Discussion

Overall, this experiment demonstrates that the hands and mouth are tightly coordinated in both the spatial and temporal domain when producing fingerspelled items with simultaneous mouthings, regardless whether the hands and mouth match in the number or direction of their movements (grand mean *r* = .794 and lag = 74 ms). This coordination occurs both in the spatial dimension of the shape of the movement gestures as well as in the temporal dimension of the alignment of each articulator's movement. Moreover, the two types of coordination appear to go hand in hand. Items with strong spatial coordination between the hand and mouth also exhibit strong temporal coordination; items with weaker spatial coordination also exhibit weaker temporal coordination.

However, a deeper investigation also indicated that not all productions were equally well coordinated; rather, a set of motoric principles emerged from the data. First, movements articulated simultaneously across the hands and mouth exhibit greater spatial and temporal coordination if the two body parts are matched in number of movements produced. This principle is fairly intuitive: it is easier for the motor system to coordinate equal number of movements across the body; if the hand has to make twice as many movements than the mouth, then the system must compensate for the temporal disparity, usually by lengthening the mouth movement, but at a motoric cost that ultimately reduces the coordination between the two parts of the body.

Second, the hands and mouth show a stronger degree of spatial and temporal coordination for simpler productions than more complex productions. As an utterance grows in length and requires more movements – even if the hands and mouth both share the same number of movements – the increased load reduces the coordinative effect of the articulation due to increased motor demands (Maner, Smith, & Grayson, 2000).

Third, hand and mouth movements are capable of coordinating movements in opposing directions just as well movements in the same direction. An opening movement of the mouth coordinates just as well in space and time with a closing movement of the hand as it does with an opening movement of the hand. The two body parts need not match their movements; so long as the hand and mouth follow the first two principles described above, their movements will still be tightly coordinated. This finding was unexpected given the pattern of movement combinations described in the domain of echo phonology, in which mouth gestures mirror the direction of hand movements (Crasborn et al., 2008; Woll & Sieratzki, 1998; Woll, 2001, 2009).

Two possibilities may explain the different results between Woll's work and the current investigation. One, the tasks studied in this dissertation differ from the sign that Woll and colleagues describe. The data presented here come from nonce fingerspelled items that are accompanied by mouthings, whereas Woll and colleagues investigated lexical signs combined with mouth gestures, which are not derived from the surrounding spoken language. It may be that the fingerspelling system itself, as part of the periphery of the sign language lexicon (Brentari & Padden, 2001; Padden, 1998), might allow for movement combinations that are illegal within the lexical sign system. Or, manual movement coordination with mouthings may be the root of the difference because mouthings have a spoken language phonological form that must be produced regardless of corresponding movement of the hands. In contrast, echoic mouth gestures may be selected from a set of meaningless oral gestures so that the hand and mouth movement similarity is maximized. However, it is currently unknown precisely how echoic oral gestures come to be associated with their corresponding lexical sign. Two, the differing results may reflect a distinction between motoric restrictions and linguistic preferences. Work in echo phonology suggests that oral movements must match the direction of the corresponding manual movement, but this may be a constraint imposed by the linguistic system rather than a physiological strategy to optimize cross-articulator coordination. Indeed, the present study demonstrates that the motor system is quite capable of coordinating such incongruent movement combinations; sign languages may simply disprefer such combinations as being ill-formed.

The principles described above provide new insights into hand-mouth gesture coordination, and could be applied to offer new accounts of both co-speech gesture as well as mouthing reductions in sign language. For instance, in attempting to provide a more complete description of mouthing reduction, the inclusion of coordination effects could help elucidate exactly when a mouthing reduction will occur as well as which specific portion of the spoken word will be eliminated instead of relying solely on the phonology of the spoken language. For example, Bank, Crasborn, and van Hout (2011) note that for NGT signs with two-syllable Dutch translations (all with stress on the first syllable in spoken Dutch), sometimes the mouthing surfaces as two syllables but other times the mouthing is reduced to one syllable (always the stressed syllable). It is particularly telling that though there is variation between one or two syllable mouthings within the whole set, for each sign there is usually a strong preference for one production style. The authors do not describe the phonological form of the NGT signs, but it may be the case that number of manual movements of the sign (or even the number of movements in the larger prosodic phrase) determines whether the corresponding mouthings are produced as one or two syllables. Coordination effects should also be taken into account when computer scientists develop sign-speech translation applications so that signing avatars appear more natural, and signing recognition systems are better able to capture the signing stream.

This chapter and Chapter 3, in part, are currently being prepared for submission for publication of the materials. Udoff, Jonathan; Nip, Ignatius; and Emmorey, Karen. “Interarticulatory Compensation between the Hands and Mouth during the Production of Mouthings in American Sign Language”. The dissertation author was the primary investigator and author of these materials.

Chapter 3: Coordination of the Hand and Mouth between Manual Homophone Pairs (Experiment 2)

3.1. Introduction

Chapter 2 explored the coordination of hand and mouth movements in a controlled experimental paradigm in which the number and direction of movements for each articulator was directly manipulated. Using the fingerspelling system in order to achieve this fine-grained manipulation, it was necessary to invent nonsense pseudowords that may not be entirely representative of fingerspellings that occur in naturalistic signing conditions. Furthermore, the fingerspelling system is only a subsystem of the lexicon of American Sign Language, and effects observed in fingerspelling may not necessarily hold true for lexical signs. This chapter presents an extension to the experiment discussed in Chapter 2, in which the signed items are real lexical signs of ASL. It is difficult to manipulate one component of a sign while keeping the other components constant within the lexical sign system (e.g., the minimal pair FLY and AIRPORT differ in their manual components by movement alone; however, the mouthings of these two signs also differ). Therefore, Experiment 2 exploits ASL's small set of manual homophones, in which a pair of signs is distinguished only by their non-manual components, such as their mouthings (see Figure 3.1). Therefore, any difference observed in the coordination between the hands and mouth within a pair must be attributed to the non-manual components, as the hands have the same phonological specifications in each sign. Furthermore, coordination effects between the motor systems may cause alterations of the articulations of both the hands and mouth (see Section 1.2). Even though the hands may share the same representation between the two manually homophonous signs, the different mouthing components could create unique coordinative effects between the two signs, resulting in different manual productions. Such a finding would demonstrate that not only do the hands influence the movements of the mouth, but the mouth can also influence the movements of the hands. Additionally, in order to distinguish between effects of sign language experience and general neuromotor capabilities, the performance of native signers must be compared to sign-naïve participants.

Figure 3.1: A manual homophone pair. The left panel shows the sign CANCEL with the corresponding mouthing /ˈkænsəl/. The image frame was taken at the peak of the first mouth opening gesture, assossiated with the vowel /æ/. The right panel shows the sign CRITICIZE, which has the same manual features as CANCEL, but with the oral mouthing /ˈkɹɪ.tɪ.saɪz/. The image frame was taken at the maximal lip protrusion, assossiated with the production of /ɹ/.

3.2. Methods

3.2.1. Participants

Ten Deaf, native signers of ASL participated in this experiment and were the same participants from Experiment 1. In addition, ten hearing non-signers also participated (5 male, 5 female). The hearing participants had no knowledge of ASL beyond a passing acquaintance with the fingerspelled alphabet. All participants were right-handed and reported no history of language or cognitive disorders. The mean age of the hearing participants was 27.2 years with a standard deviation of 4.6 years. This was roughly similar to the Deaf participants, who had a mean age of 30.2 years with a standard deviation of 9.3 years.

3.2.2. Materials

The stimuli for the experiment comprised 12 ASL signs, 6 sign pairs that are manually homophonous, yet differ in their mouthings: BURGER/HAMBURGER, CANCEL/CRITICIZE, DORM/DORMITORY, PICTURE/PHOTOGRAPH, SWISS/SWITZERLAND, and POP/SODA-POP. As such, each pair contained a shorter item and a longer item. This length difference (hereafter, mouthing length condition) is based on the number of syllables required for the mouthing component; between the two signs of each pair, the specified length of the manual component is the same. The pairs DORM/DORMITORY and PICTURE/PHOTOGRAPH are produced as location-movement-location sequences. The remaining 4 pairs are produced as two sequential path movements with an intervening transition movement. The signs with two path movements are phonologically disyllabic, whereas the location-movement-location items are monosyllabic; however both types of signs in the experiment had a doubled feature – either two locations or two path movements – and are phonologically complex enough to potentially be partitioned into smaller units for coordination with the oral component of the sign.

3.2.3. Procedure

Participants were told that items came in pairs in which the manual component was the same but the mouthing component was different; therefore, it was important to mouth each item to distinguish it from its pair. Hearing non-signers saw the stimulus list before test to learn and practice the target signs. Videos of a Deaf, native signer demonstrated each sign with the appropriate mouthing, and the hearing participant copied the sign. English glosses of the signs were displayed about the video. If the participant made any errors (e.g., producing the sign with an incorrect phonological parameter, forgetting to mouth, or switching hand dominance), the experimenter explicitly corrected the participant, and the participant signed the item again. The entire set of stimulus videos were shown to hearing non-signers twice before the testing phase, in the same order as they appeared at test.

For both the hearing non-signers and Deaf participants, the testing procedure began with the presentation of the target sign as an English gloss projected onto a screen. The item was embedded in the carrier phrase "MY (target sign) MY," also presented on-screen, because the sign MY provides a clear three-dimensional landmark to mark the beginning and end of the manual movement, and the bilabial closure that begins the spoken word "my" also provides a clear landmark for the lip movement. This is in contrast to Experiment 1, in which no carrier phrase was used because the mouthing and manual components were specifically chosen for movement properties that could be easily identified by their kinematic properties. In the present experiment, however, the specific form of the mouth and hand movements were largely ignored beyond the fact that manual movements were identical within homophone pairs, so the sign MY provided a clear beginning and end boundary for each utterance. Figure 3.2 illustrates the sign MY. Once prompted, the participant signed and mouthed the target item within the carrier phrase 12 times in succession, until prompted to stop. Any error in production – typically only occurring within the hearing group – was corrected by the experimenter and the trial was repeated with the correct production form.

Figure 3.2: The sign MY is produced with the open-B handshape, which touches the central chest. Each repetition began and ended with this sign.

3.2.4. Data capture

Data from the hearing and Deaf participants was recorded using motion capture and a digital video camera, identical to the data capture methods from Experiment 1 (Section 2.2.4).

3.2.5. Post-processing

Motion capture data were post-processed in the same manner as Experiment 1 (Section 2.2.5). Unfortunately, the manual markers were occluded too often to recover reliable motion capture data of the hands, so only the mouth markers were processed. Therefore, the manual sign data was instead coded frame-by-frame in ELAN as described below (Crasborn & Sloetjes, 2008; The Language Archive, 2010).

Table 3.1: Criteria used to identify the temporal beginning and end of the manual component of each target sign item in Experiment 2

Item	Beginning	End
BURGER/HAMBURGER	Beginning of straight path movement of the two hands coming together	Second time hands touch
CANCEL/CRITICIZE	Beginning of straight path movement for the first stroke of the "x"	End of straight path movement of second stroke of "x"
DORM/DORMITORY	First contact of hand to face	Second contact of hand to face
PICTURE/PHOTOGRAPH	Contact of hand to face	Contact of hand to non-dominant hand
SWISS/SWITZERLAND	Beginning of downward path movement	End of lateral path movement
POP/SODA-POP	Beginning of downward path movement towards non-dominant hand	Contact of B handshape with non-dominant hand

3.2.6. ELAN coding

ELAN was used to align annotations of the hand and mouth movements to derive temporal measures of the two components. Table 3.1 details the gestures that define the temporal beginning and end of each manual sign. Transitional movements before and after the target sign were excluded from the analysis; therefore, any changes occurring on the handshape or orientation specification had to be completed in order to be included in the annotation. Motion capture information from the lip makers was imported into ELAN to delineate the beginning and end of the mouthing components. The Euclidean distance between the upper and lower lip markers defined the opening and closing of the mouth, and this aperture signal was aligned with the video and its annotations in ELAN (see Figure 3.3). A mouthing began at the first local minimum in the aperture signal that occurred after the maximum for the vowel of the first “my” (/maɪ/); a mouthing ended at the last local minimum or maximum before the production of the second “my,” which in turn was identified as starting with a local minimum. If the subject took any breaths or otherwise silently moved his mouth in anticipation of the next sound to articulate, this movement was excluded from the mouthing annotation. Most often this postural resetting of the mouth occurred between the end of the target mouthing and the mouthing of the second “my” rather than between the first “my” and the target. Rarely a local extremum needed to define a mouthing’s beginning or end was obscured because two oral gestures had partially coalesced together. In these circumstances, the inflection point of the gesture (i.e., the local extremum of the corresponding velocity signal) was used to distinguish the two underlying gestures and replaced the local extremum as the kinematic landmark to delineate the mouthing. Once the timing of the oral and manual components was annotated, the temporal information was extracted from ELAN and imported to Microsoft Excel for statistical analysis.

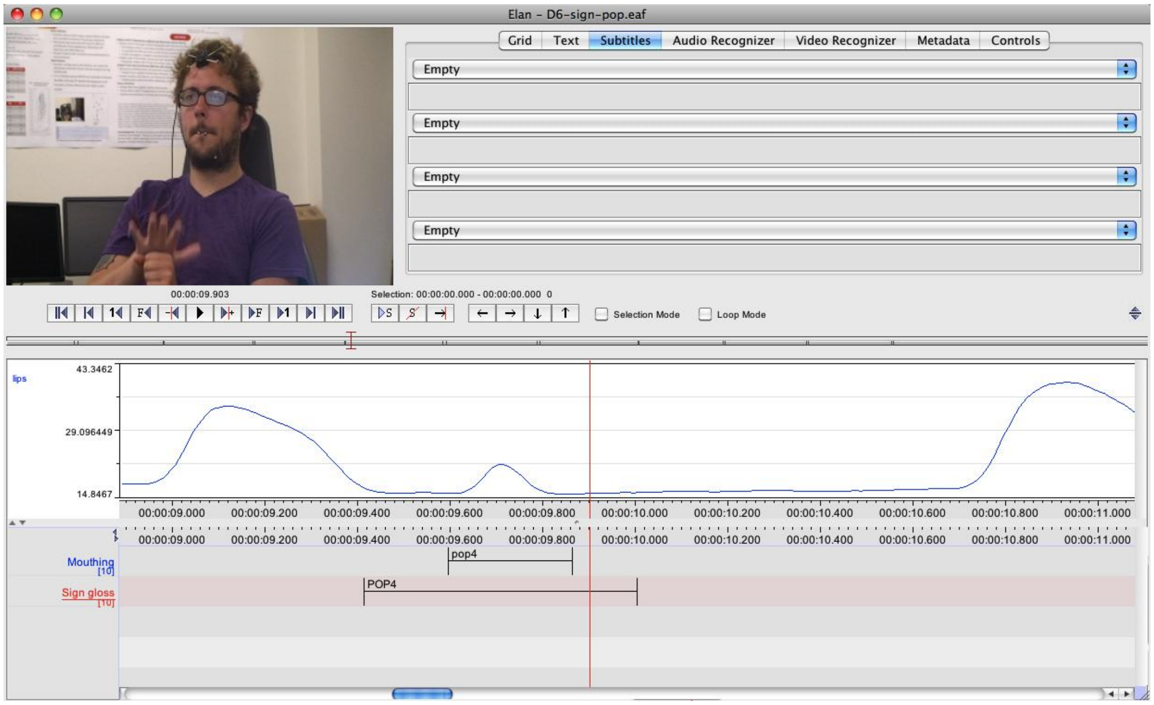


Figure 3.3: The ELAN annotation window. A reference video is in the upper left corner, the lip aperture signal is displayed in the middle of the window, and below it are the two annotation tiers to mark the duration of the oral mouthing and manual components of the target sign.

3.2.7. Missing data

A total of 50 out of 2400 repetitions were not included in the final analysis. Due to experimenter error, all ten repetitions of 4 items were not recorded, resulting in a loss of 40 repetitions. An additional 10 repetitions were excluded because one of the Deaf participants mouthed "pop" instead of "soda pop" for all repetitions of the item SODA-POP, making it indistinguishable from the item POP in both its manual and oral components. In total, 2.1% of the planned data set was excluded from the final analysis.

3.3. Results

A 2×2×2 ANOVA (participant group: Deaf or hearing, mouthing length: short and long, and articulator: hand and mouth) was performed in order to assess not only the effects of group and length on each articulator's duration but the influence one articulator's duration has on the other articulator as well. Figure 3.4 illustrates these results. Overall, hearing participants took longer to produce their utterances than the Deaf participants: 701 ms vs. 532 ms, $F(1, 18) = 17.7$, $p = .001$. Utterances within the long mouthing length condition were also longer than utterances within the short condition: 683 ms vs. 549 ms, $F(1, 18) = 128.6$, $p < .001$. This length effect was observed within the oral mouthing component (770 ms vs. 542 ms, $F(1, 18) = 184.9$, $p < .001$), which validates that the condition labels of short and long mouthings were accurately applied to the homophone pairs, as well as within the manual sign component (597 ms vs. 558 ms, $F(1, 18) = 16.9$, $p = .001$). There was no significant group × length condition interaction: $F(1, 18) < 1$, $p = .54$.

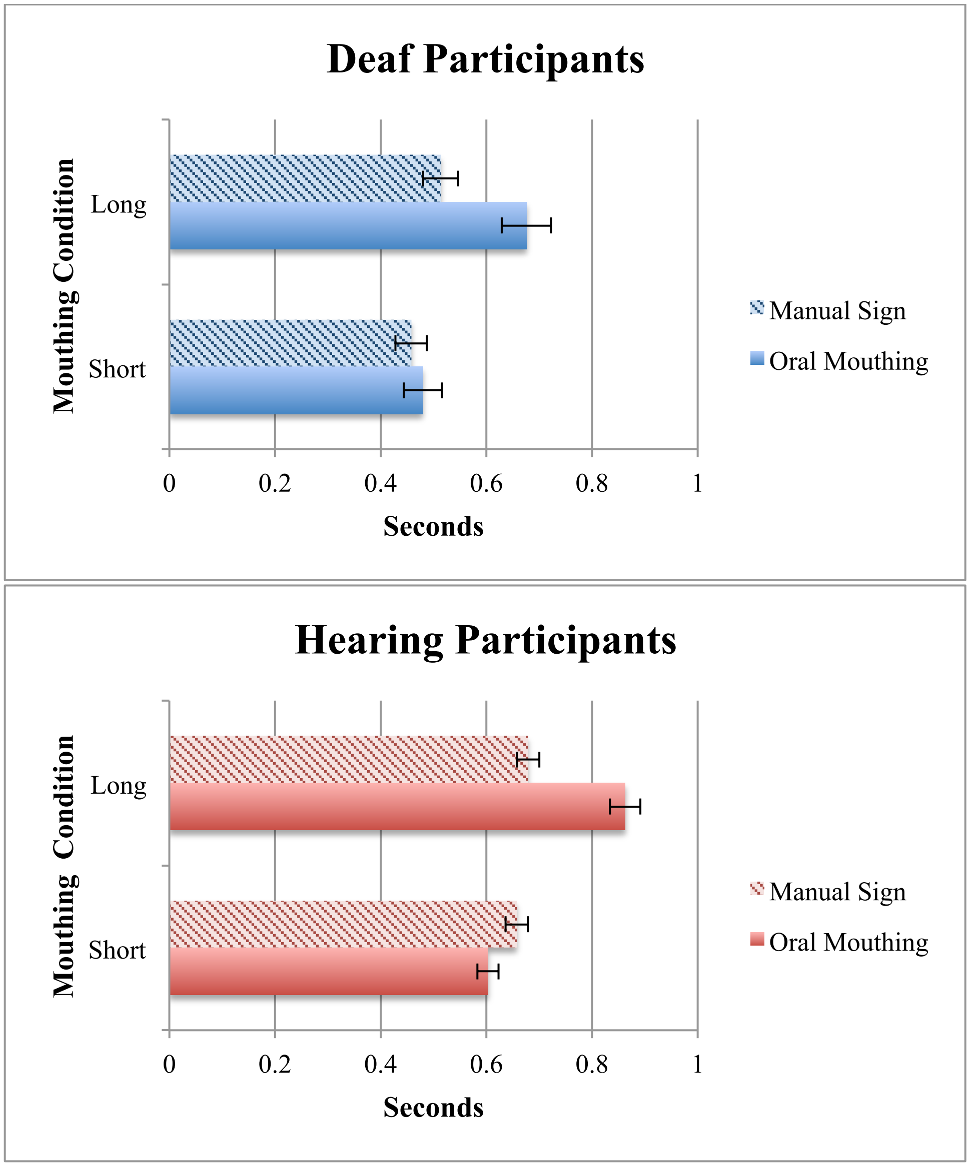


Figure 3.4: The duration results for both participant groups by component and mouthing length condition

Oral mouthing durations were revealed to be significantly longer than the durations of manual sign components: 655 ms vs. 577 ms, $F(1, 18) = 57.7$, $p < .001$. Significant interactions were also revealed for articulator × length condition ($F(1, 18) = 194.8$, $p < .001$) and the three-way interaction of group × articulator × length condition ($F(1, 18) = 13.0$, $p = .002$). The articulator × length interaction is driven by the long condition: Mouthings last longer than the manual sign within the long condition ($t(19) = 14.1$, $p < .001$), but in the short condition the two articulations are not significantly different in duration ($t(19) = 1.1$, $p = .277$). This articulator effect, however, is modulated by participant group. While both Deaf and hearing groups exhibit a significant difference between the duration of the articulations in the long condition ($t(9) = 9.6$, $p < .001$ and $t(9) = 10.2$, $p < .001$, respectively), Deaf participants show no significant difference in the short condition ($t(9) = 1.6$, $p = .151$), and hearing participants show a significant difference in the opposite direction in the short condition, with the manual sign lasting longer than the oral mouthing ($t(9) = 2.7$, $p = .026$), resulting in a group × component × length interaction. This three-way interaction seems to be a result of the hearing participants' tendency to produce longer manual signs regardless of the length of the mouthing component. There was no significant group × component interaction: $F(1, 18) = 1.7$, $p = .20$.

In order to determine the temporal ordering between the manual and oral components of the target productions, each item was classified based on which component started its articulation first. While the mouth movements were recorded using motion capture, which provides high temporal resolution (120 frames per second in this experiment), hand movements were only faithfully captured using a standard video camera, recording at 29.97 frames per second. Therefore, any absolute difference between the two start times of the articulators less than 33.37 ms (1000 ms / 29.97 fps) was treated as occurring in synchrony, as it was less than the margin of error.
Figure 3.5 summarizes the proportions of which articulator initiated its movement first, separated by participant group and mouthing length condition. Chi-squared analyses

revealed that Deaf participants' distribution significantly differ from hearing participants ($\chi^2(2) = 243.8$, $p < .001$), and across both participant groups, the distribution among short mouthing items significantly differed than that of long mouthing items ($\chi^2(2) = 113.2$, $p < .001$). Hearing participants were more likely than Deaf participants to begin a trial with hand movements whereas Deaf participants were more likely to begin the utterance with mouth movements or synchronize the two articulators than hearing participants were. Overall, while both short and long mouthing conditions had the same proportion of items that began with synchronized mouth and hand movements, short mouthing items were more likely to begin with the manual component whereas long mouthing items were more likely to begin with the mouthing component. Furthermore, this difference between short and long mouthing conditions was significant for both Deaf ($\chi^2(2) = 70.4$, $p < .001$) and hearing ($\chi^2(2) = 55.4$, $p < .001$) participants. Both participant groups exhibited a similar pattern between the two mouthing length conditions: while the proportion of synchronization remained constant between the short and long mouthing items, the proportion of items beginning with the mouthing component increased from the short to long condition as the manual component proportion decreased from the short to long condition. Figure 3.6 details the degree of temporal coordination between the two articulators by reporting the relative proportion of items in which the beginning of the manual sign component preceded the beginning of the mouthing component in number of frames.

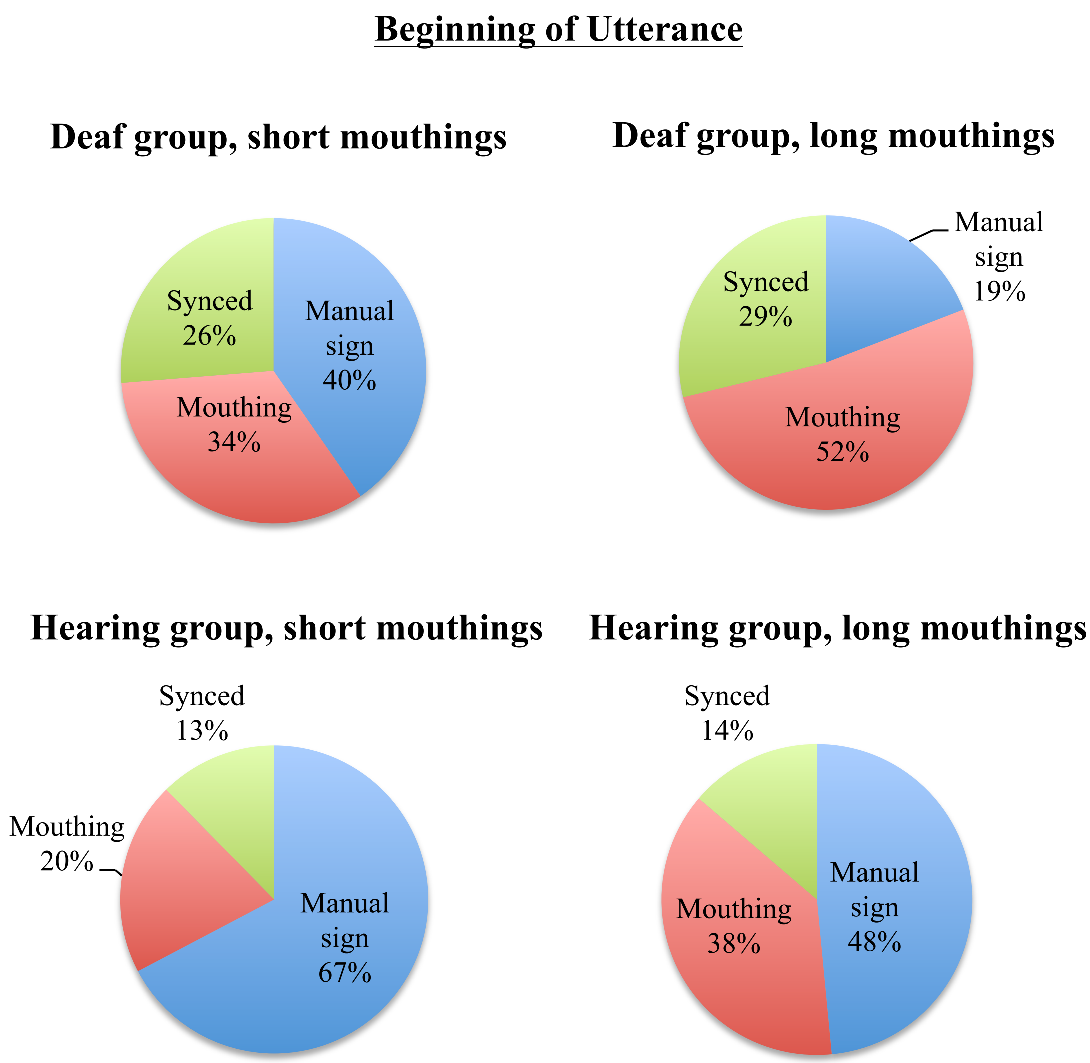


Figure 3.5: Proportions of which component of the target item began its articulation first

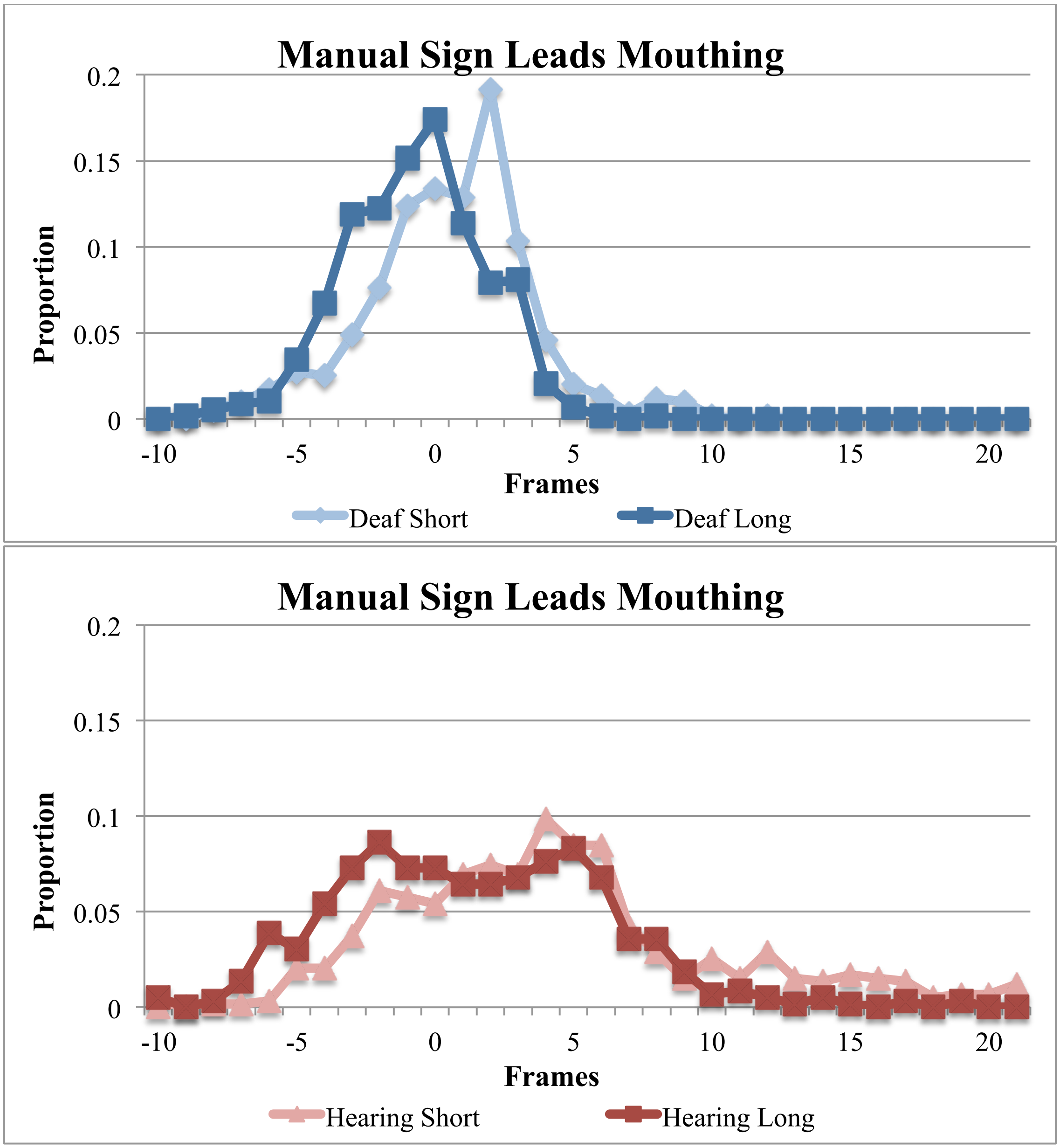


Figure 3.6: The proportion of items by number of frames in which the beginning of the manual component precedes the beginning of the mouthing component. Upper graph shows the data from the Deaf group, and the lower graph shows the data from the hearing group. In each, the data point farthest to the right represents the proportion of all instances in which the manual component's starting point preceded the mouthing component's starting point by more than 20 frames (667 ms).

Finally, with regard to the coordination of the end of an utterance, the results were more uniform, with the majority of items ending first with the conclusion of the manual sign and then the mouthing component concluding afterward. Figure 3.7 summarizes the proportions of which articulator concluded its movement first, separated by participant group and mouthing length condition. Chi-squared analyses revealed significant differences between hearing and Deaf participants ($\chi^2(2) = 46.9$, $p < .001$) and between short and long mouthing conditions ($\chi^2(2) = 205.2$, $p < .001$). Specifically, hearing participants more often concluded the manual component of the target item first than did the Deaf participants while hearing participants employed the other two timing alignments less often than Deaf participants did.

As for the length condition, both the short and long conditions had relatively similar proportions of synchronous utterances, but the long condition had more instances of the manual component concluding first and fewer instances of the mouthing component concluding first. Further examining these effects within each group, both the Deaf ($\chi^2(2) = 74.8$, $p < .001$) and hearing participants ($\chi^2(2) = 142.5$, $p < .001$) demonstrated significant differences between the two mouthing length conditions. Both groups exhibited a larger proportion of the manual component finishing first and a smaller proportion of the mouthing component finishing first in the long condition as compared to the short condition. However, the hearing participants also showed a decrease in synchronization in the long condition as compared to the short condition whereas synchronization proportions were relatively equal between conditions for the Deaf participants. Figure 3.8 details the degree of temporal coordination between the two articulators by reporting the relative proportion of items in which the end of the manual sign component preceded the end of the mouthing component in number of frames.

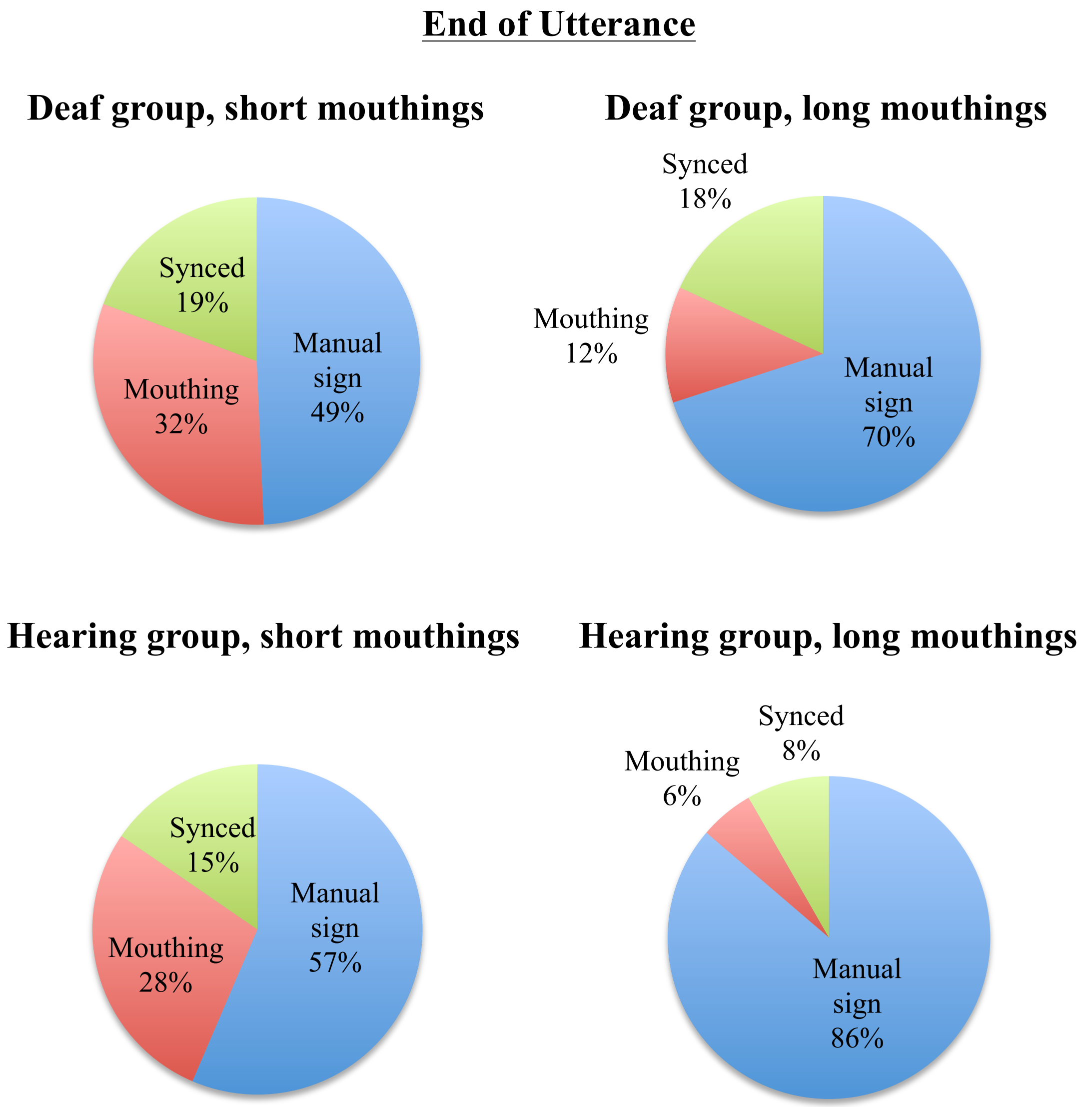
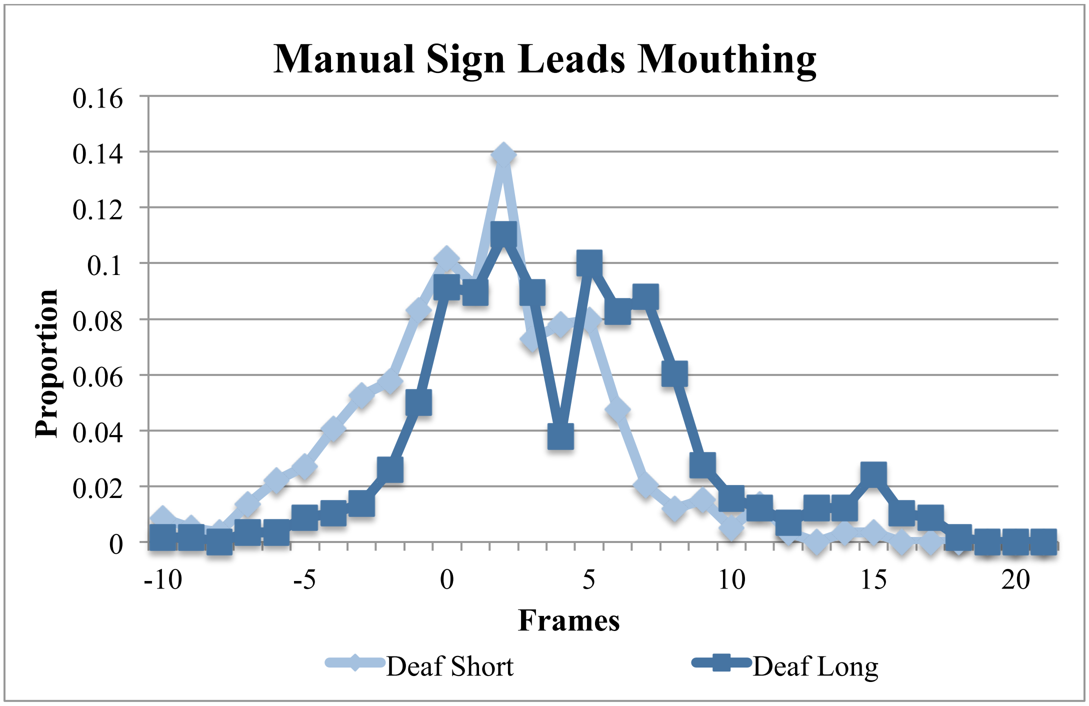


Figure 3.7: Proportions of which component of the target item finished its articulation first



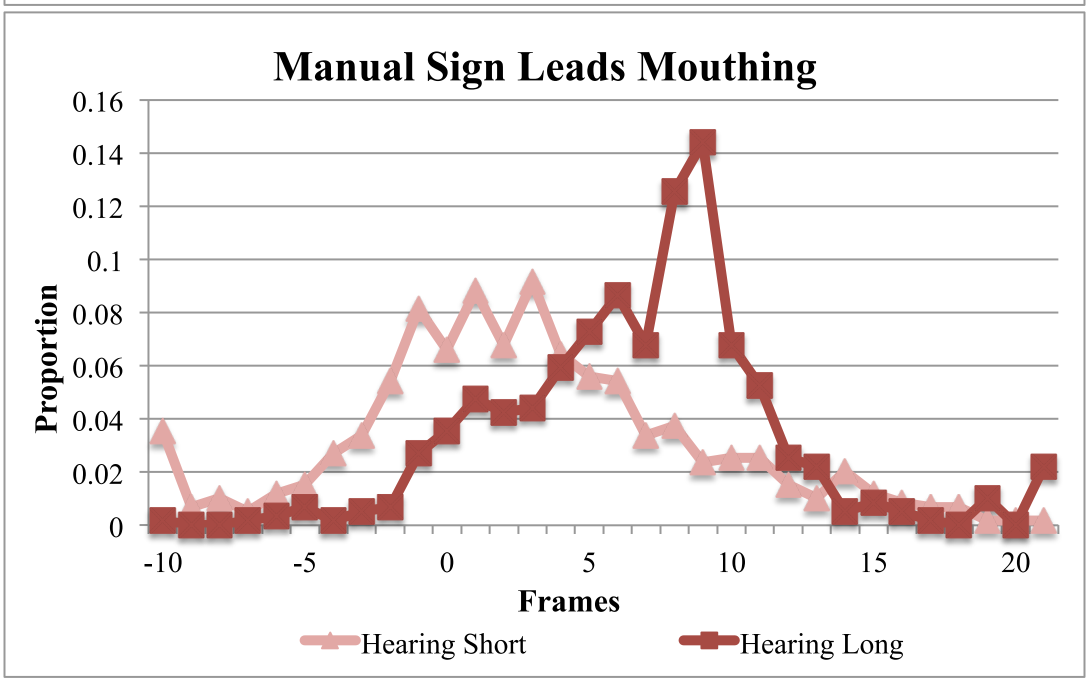


Figure 3.8: The proportion of items by number of frames in which the termination of manual component precedes the termination of the mouthing component. Upper graph shows the data from the Deaf group, and the lower graph shows the data from the hearing group. In each, the data point farthest to the right represents the proportion of all instances in which the manual component's endpoint preceded the mouthing component's endpoint by more than 20 frames (667 ms).

3.4. Discussion

The major contribution of this study is that we found evidence that the effect between the hands and the mouth is bidirectional. As numerous studies have found, using evidence from both sign language and other domains, the movements of the hands can influence the movements of the mouth. However, evidence for the mouth influencing the hands is sparse (Gentilucci et al., 2001; Smith et al., 1986). The key manipulation in this experiment, the length of the mouthing component between otherwise homophonous sign pairs, corroborates other findings that hand movements are affected by movements of the mouth. Participants' sign articulations were significantly longer when the accompanying mouthings were for the longer items of the homophone pairs. They accommodated the longer oral articulations by lengthening the manual articulations. Both the Deaf signers and hearing non-signers exhibited this articulation strategy, suggesting that this length synchrony effect between the hands and mouth reflects a general physiological principle of temporal coordination across the body rather than a consequence of sign language experience.

Though both hearing and Deaf groups exhibited the same effect between the two length conditions, overall the two participant groups differed significantly in the absolute time of their productions; hearing non-signers took longer to articulate both the manual and oral mouthing components as compared to the Deaf participants. This difference is likely due to the sign naïve group's unfamiliarity with signing. The task demands of this study require hearing non-signers to recall a recently learned manual gesture while simultaneously silently mouthing an English word. Without prior practice, these productions would certainly impose task demands that affect the non-signing group more than life-long signers and ultimately slow down their articulatory system. For the non-signers, the combination of (unvocalized) speech with manual signs may be interpreted as a concurrent motor task rather than a dual linguistic task. Motor “distractor” tasks have been previously documented to affect speech articulations (Dromey & Bates, 2005; Dromey & Benson, 2003), coordination between the speech and motor tasks has not been directly investigated. In contrast, native signers treat both the manual and oral articulations as part of the same linguistic task, resulting in different articulatory coordination strategies between the two participant groups. To verify this hypothesis, it would be beneficial to extend this study to include a group of hearing signers, either adult children of Deaf adults, who would have acquired sign language from birth, or late learners of sign, who studied ASL as adolescents or adults. The results from these hearing signers would then differentiate whether the differences seen between the hearing and Deaf groups in this experiment are the result of familiarity with signing (in which case both native hearing signers and hearing late learners of sign should both perform like the native Deaf group) or if instead the differences are a result of native language proficiency (in which case native hearing signers should perform like native Deaf signers but late learners of sign should perform like the hearing non-signers or somewhere between the native and naïve groups).

While the durational analysis demonstrates that the production of one component of a sign will lengthen when the other articulator becomes longer, this does not result in signed articulations in which the hands and mouth are always perfectly synchronized and aligned in time. As the analyses of the utterances’ start and end coordination indicate, there still remains a temporal pattern of how the two components of the target items are ordered. Overall, Deaf participants exhibited no discernible preference for how to coordinate the beginning of an utterance, with roughly equal proportions among starting the two components in synchrony, starting the sign component first, and starting the mouthing component first. Hearing participants, however, preferred to begin their utterances with the production of the manual component first, roughly two thirds of all instances. Again, unfamiliarity with sign language production may be driving these results for the hearing non-signers; they first attempt to recall the manual gesture, and only after they begin executing the necessary manual movements do they then begin to mouth the more familiar English word. This is not merely an effect of language dominance but rather of complete unfamiliarity. Hearing individuals bilingual in ASL and English, with English being more dominant, do not stagger the two articulations during the simultaneous production of (voiced) speech and sign ("code-blends"); rather, they synchronize the onset of speech with the onset of the sign (Emmorey, Petrich, & Gollan, 2012). Though code-blending and mouthing production are separate phenomena, the coordination between the hand and mouth articulation by experienced sign language users appears to be similar.

Despite the differing baseline preferences between how the hearing and Deaf participant groups initiate an utterance, the effect between mouthing length conditions was remarkably similar between the two groups. Both hearing non-signers and Deaf signers showed little difference in the proportion of synchronized repetitions between the two conditions, but instead they both were more likely to begin with the oral mouthing component and less likely to begin with the manual sign component for the longer item of the pair rather than the shorter item. This pattern indicates that even for the hearing group, which overall preferred to begin with the manual component, when the mouthing grew in length, their production system was sensitive to extra time such a production would require, and therefore initiated the mouthing production earlier on in the utterance.

In contrast, the coordination of the end of an utterance was much more consistent: across all participants and conditions, the majority of utterances (65%) ended with the manual component concluding first while the mouthing completing its articulation afterwards. This pattern was observed in both participant groups and in both mouthing length conditions, and in conjunction with the analysis of the difference in the two components' durations, this suggests that mouthings end later because they last longer than the manual sign component. This effect is only enhanced in the long mouthing condition as compared to the short condition; especially long mouthings are even more likely to cause the participant to complete the manual sign while still needing additional time to finish mouthing the oral component. Taking both the start and end coordination analyses into account, particularly from the information presented in Figure 3.6 and Figure 3.8, a pattern emerges of the coordination of the oral and manual components of the test items: as the duration of a mouthing grows longer, it does not only spread rightward with respect to the manual component, but there is also a tendency for the oral component to spread slightly to the left as well.

Overall, the findings drawn from this experiment suggest that the motor systems of the hands and mouth require a minimum level of familiarity with a given inter-articulatory task in order to achieve a level of articulatory flexibility in which the order of hand–mouth coordination may vary. Otherwise, if the body is not accustomed to the task involving concurrent hand and mouth articulations, it appears that the motor systems' strategy to coordinate the movements is to start with the less familiar task first (sign generation) and after it is initiated, then resources (be they motoric or attentional) can be reallocated to also execute the more familiar task (silent speech). Though the coordination of the initiation of a multi-effector utterance seems to be mediated by task familiarity, compensation for increasing the length of the oral articulation is unaffected by experience. The motor control systems of the hand and mouth exhibit bidirectional influences so that changes in hand movements can affect the movements of the mouth just as movements of the mouth can affect alterations to the movements of the hands, as seen in this experiment between the two length conditions. Given this diversity of effects, the nascent field of sign phonetics must pay special attention to the underlying source of articulatory phenomena: while some articulatory regularities may be codified as part of the linguistic system, still others may instead be general physiological consequences of the involved motor systems. It is crucial that future studies distinguish between these two possibilities so that extraneous effects do not contaminate our understanding of sign phonetics as well as motor control.

This chapter and Chapter 2, in part, are currently being prepared for submission for publication of the materials. Udoff, Jonathan; Nip, Ignatius; and Emmorey, Karen. "Interarticulatory Compensation between the Hands and Mouth during the Production of Mouthings in American Sign Language". The dissertation author was the primary investigator and author of these materials.

Chapter 4: The Motor Representation of Mouthings and Speech (Experiment 3)

4.1. Introduction

Mouthings have been documented and studied extensively, and the phenomenon has been reported in every established sign language known (Boyes Braem & Sutton-Spence, 2001). However, the precise status of mouthings remains controversial, and there exists a continuum of opinions regarding the position mouthings hold within sign language structure. On one end of the continuum, some maintain that mouthings are a foreign intrusion of spoken language onto the native sign language structure (Lucas & Valli, 1989; Lucas, 1992). As such, mouthings are indicators of non-native sign proficiency or are merely a contact language strategy for communicating with inexperienced signers.

Others take a more moderate stance and posit that mouthing, while still not a part of the sign language, are more akin to code-switches (Bank et al., 2011; Mohr, 2012), which is a bilingual phenomenon that occurs in spoken languages, and have thus described mouthings as "code-blends" (Emmorey, Borinstein, & Thompson, 2005). The only difference between code-switches and code-blends is the timing: unimodal bilinguals can only produce one language at a time and must abandon one language for another (i.e., switch) in order to mix languages; in contrast, bimodal bilinguals can divide up their language selection across the body due to the modality difference between spoken and signed languages. The hands produce sign language while the mouth simultaneously produces spoken language, resulting in a blend of the two lexical forms. Most researchers distinguish between these two types of co-sign mouth movements: silent utterances are mouthings and voiced utterances are code-blends. Crucially, Bank et al. and Mohr propose that these two types of mouth movements are more alike than different and therefore treat them as the same phenomenon.

Further along the continuum, Ebbinghaus and Hessmann (2001) argue that mouthings are native to the linguistic system of signed languages, but mouthings are represented as units that are distinct from lexical signs, interacting with manual signs during sign production to add to the overall meaning of an utterance. Finally, others still insist that mouthings are an inherent part of a signed language's lexicon and represent a distinct phenomenon from voiced speech that is produced without accompanying signs (Boyes Braem, 2001; Rainò, 2001). Proponents of this view argue that mouthings are used, at least in some languages, to disambiguate otherwise homophonous signs and as such mouthings cannot be excluded from the linguistic system.

Despite this debate persisting for 25 years, there is very little experimental work to bring to bear on the representation of mouthings and their status in sign language. Vinson and colleagues (2010) recently found that mouthing errors and lexical sign errors do not necessary occur together, suggesting that the two representations are not "bundled" together in a signer's mind (see Section 1.1.3.1). Neuroimaging work has also suggested that perceiving sign and mouthing combinations activate cortical regions in the brain that are distinct from the activation for perceiving sign and mouth gesture combinations, again suggesting that mouthings may not be fully incorporated into the mental representation of sign (Capek et al., 2008).

The goal of the present study is to extend this previous work regarding the status of the mental representation of mouthings. If mouthings are not an intrinsic part of sign language representations, then one alternative is that they are derived online from representations of spoken language. Though this hypothesis has been assumed to be true in previous accounts of mouthings (e.g., Ajello et al., 2001; Bank et al., 2011; Vogt-Svendsen, 2001), there has been no empirical work to support the claim. Therefore, this study directly compares the motor representation of mouthings to the representation of speech by examining the articulatory gestures used in their production. Measures such as inter-item stability and correlations between conditions may reveal differences in the motor execution of items when they occur in different production conditions and thereby provide evidence for distinct representations.

In addition, the articulatory gestures of silently produced speech in the absence of signing were included in the comparisons. Co-sign mouthings require movement of both the hands and mouth whereas speech only requires movement of the mouth. Speech requires audible vocalization whereas mouthings are produced silently. To discern possible differences between the confounding effects of vocalization and the participation of multiple articulators, the silently non-signing condition provides an opportunity for participants to still produce silent mouth movements but without the need to coordinate movements of multiple articulators. Also of interest is the role of speech and language experience on the task. In order to distinguish between effect that may arise from general principles of motor control rather than an effect of linguistic representation or audiological status, both native Deaf signers and hearing non-signers participated in this experiment.

It was hypothesized that Deaf participants maintain motor representations of co-sign mouthings that are distinct from the motor representation of speech whereas hearing participants only have speech representations. It was therefore predicted that Deaf speech articulations would exhibit less internal consistency than mouthings and hearing participants would show no difference between the two conditions. Likewise, it was also predicted that Deaf participants would exhibit differences in the movement patterns of the mouth between speech and co-sign mouthing productions whereas the hearing group would produce more similar oral movements between these two conditions.

Lastly, an additional manipulation of speech visibility was include in order to determine if Deaf participants are more reliable in their speech productions if the item being spoken contains more visible elements, which are more easily accessible to a deaf individual. Thus, it was predicted that Deaf participants would produce more visible items with greater internal consistency, but this effect would not be observed in the hearing group.

This last prediction is based on the hypothesis that Deaf signers' motor representations of speech are significantly different in form than hearing non-signers' representations, irrespective of their relationship to the motor representations of mouthings. Numerous studies on the speech of deaf individuals supports the idea that deaf vocal productions differ from normally hearing individuals[6] as the result of deficits in motor planning rather than deviant phonology (Lane & Perkell, 2005; McGarr & Osberger, 1978; Tourville & Guenther, 2011; Weismer, 1984). However, studies of the activity of oral muscles have found that certain classes of sounds by deaf speakers are more privileged during production than others. In particular, deaf speakers are more consistent among themselves and look more similar to hearing speakers when producing bilabial consonants /p, b/ than coronal consonants /s, ʃ, t/ (Huntington, Harris, & Sholes, 1968; McGarr & Harris, 1983). Similar observations were made indicating a privileged status of open vowels over closed vowels as well as rounded vowels over unrounded vowels (McGarr & Gelfer, 1983). Moreover, there is a general trend for lip movements to show more consistency than tongue movements among deaf speakers (McGarr & Harris, 1983).

This recurring finding that deaf speakers exhibit greater accuracy and stability for the production of speech segments with highly visible features suggests that deaf individuals' knowledge of English may be based upon viseme categories. Visemes are the basic unit of visible speech, and the visual counterpart of acoustic phonemes (Chen & Rao, 1998; Fisher, 1968). The relationship between phonemes and visemes is a many-to-one mapping. For instance, /p, b, m/ are acoustically distinct, so they represent three separate phonemes. However, these three phonemes are visually indistinguishable, so

[6] For a review of the various differences documented between deaf and hearing speech, see Osberger & McGarr, 1982

they form one viseme. In contrast, phonemes that are acoustically similar, like /m/ vs. /n/, may fall into separate viseme classes, as /m/ is produced with the mouth closed and /n/ with the mouth open. If a particular set of sounds can be distinguished, in part, based on visual properties, such as pursed lips always occurring with /w/, then certainly a deaf speaker would take advantage of this cue, especially given the inaccessibility of acoustic information during the acquisition of a phonemic inventory. Although almost all of the studies mentioned above appeal to the idea that speech visibility may play a role in producing certain sounds, the present experiment is the first to directly test such an effect of speech visibility in the production of English by deaf speakers.

4.2. Methods

4.2.1. Participants

Ten Deaf, native signers of ASL and ten hearing non-signers participated in this experiment. The Deaf participants were the same as the participants in Experiment 1 and Experiment 2 (Section 2.2.1), and the hearing participants were the same from Experiment 2 (Section 3.2.1).

4.2.2. Materials

The stimulus list consisted of 10 items. With regard to the items' English phonology, half of the words contained a high percentage of visible segments (75% or more) and half with low percentage of visible segments (50% or less). Visible segments were defined as the set of speech sounds that have visibly salient articulations: visible contact of the lower lip either with the upper lip in /m, p, b/ or with the upper teeth in /f, v/; visible protrusion or constriction of the lips for rounded vowels in /o, ɔ, u/; and significant and visible lowering of the jaw in the English open vowel /ɑ/. Because the word list was also designed so that each item begins and ends with a visible segment, the percentage visible items was inversely correlated with the item's length. Due to the constraints of speech visibility on the stimuli selection, many of the target words were low frequency items (Brysbaert & New, 2009a, 2009b). Therefore, to ensure that Deaf participants would recognize the English words, items were only selected if they also had a familiarity rating no lower that one standard deviation below the source corpus's mean (Wilson, 1988).

Table 4.1 summarizes the items' spoken language characteristics. With regard to the items' signed phonology, each word was chosen so that its ASL sign translation had just one path movement in order to minimize variability introduced by hand–mouth coordination effects (see Experiments 1 and 2).

Table 4.1: Experiment 3 stimulus list

Item	Phonology	Percentage of visible segments	Viseme group	Frequency	Familiarity[7]
become	/bɪ.ˈkʌm/	40%	low	115.57	603
bloom	/blum/	75%	high	5.51	426
boredom	/ˈbɔɹ.dəm/	50%	low	2.14	527
freedom	/ˈfɹi.dəm/	33%	low	33.10	568
move	/muv/	100%	high	418.14	572
palm	/pɑm/	100%	high	13.24	515
paragraph	/ˈpɛ.ɹə.gɹæf/	25%	low	2.82	559
platform	/ˈplæt.fɔɹm/	50%	low	6.14	498
pope	/pop/	100%	high	10.71	489
proof	/pɹuf/	75%	high	34.39	546

4.2.3. Procedure

As in Chapter 3, participants were presented English glosses of signs within the phrase "MY (target sign) MY" and asked to repeat each item 12 times in each of three conditions. Hearing non-signers were taught the signs in the stimulus list using the same procedure from Experiment 2 (see Section 3.2.3). The mouthing condition required the participants to produce the manual signs with the accompanying silent mouthings. Next, in the silent condition, participants were asked to silently move their lips as if they were speaking, but without producing sound. The participants did not produce manual signs in

[7] Familiarity ratings refer to print familiarity and were derived from merging three corpora: Paivio (unpublished), Toglia and Battig (1978), and Gilhooly and Logie (1980). Values range from 100 to 700 with a mean of 488 and a standard deviation of 99. Retrieved from http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm

this condition. Finally, in the voiced condition, participants spoke the target word and carrier phrase out loud, again without producing manual signs.

4.2.4. Data Capture and Post-Processing

Data was recorded using motion capture and digital video recording, using the same equipment and set-up as Experiments 1 and 2 (see Section 2.2.4). The post-processing of the markers' movement signals was also the same as in Experiments 1 and 2 (see Section 2.2.5). As this experiment was only concerned with comparing lip movements between production conditions, only the lip markers were subjected to post-processing and analysis. A kinematic landmark from the spoken/mouthed word "my", when the distance between the upper and lower lips were at a minimum for the bilabial closure, defined the start and end points of each repetition. 61 of the planned 6000 repetitions were not recorded and thus not included in the final analysis due to experimenter error. Overall, this accounted for 1.0% of the total data set.

4.3. Results

This experiment uses a similar statistical technique, cross-correlation, as was performed in Experiment 1. However, there are three slight differences in the precise methodology. First, because the present experiment makes comparisons of the same articulator across different conditions, rather than comparing multiple articulators that were moving in the same trial, lag scores were not analyzed from the analysis. The cross-correlation still calculated correlation coefficients for all possible lag values, but because one repetition may have began or ended with longer transition movements, the precise lag value was unimportant; rather it was the maximal *r*-value that was of sole interest.

Second, the correlation coefficient, *r*, was also chosen in a different manner than in Experiment 1. Unlike in that experiment in which articulators sometimes moved in opposing directions and yielded negative correlations, in the present experiment the overall movement signals of an utterance between conditions should not vary so greatly that one movement trace is the inverse of the other. Therefore, the maximum *r*-score was taken from the cross-correlation analysis, rather than the maximal absolute value of the *r*-scores. Third, because the cross-correlation analysis prefers signals of equal length, all repetitions were time-normalized before undergoing the analyses. Spline interpolation was used to transform each movement trace into 1000 data points.

4.3.1. Item Stability Analysis

The data were first subjected to an item stability analysis to access the consistency of each participant's productions across all ten repetitions of a single item. Cross-correlations were performed on each repetition in comparison to the mean movement signal of the remaining 9 repetitions from the same item and participant. Higher average *r*-values across all 10 repetitions of an item are interpreted as participants being more consistent with their productions as a result of a more robust motor representation of the item (Green, Moore, Higashikawa, & Steeve, 2000).

A 3 × 2 ANOVA (condition: mouthing while signing, silent speech without signing, voiced speech without signing; participant group: Deaf, hearing) was performed on the *r*-values and revealed a main effect of condition ($F(1, 18) = 34.2$, $p < .001$) and significant condition × group interaction ($F(1, 18) = 7.2$, $p = .015$). There was no significant difference between participant groups: $F(1,18) = 2.2$, $p = .152$ (see Figure 4.1). Post-hoc comparisons revealed the interaction is the result of the hearing group exhibiting increasing stability among conditions as the task became more like spoken English ($F(2, 27) = 6.8$, $p = .004$) whereas the Deaf participants were equally stable across all three conditions ($F(2, 27) < 1$, $p = .541$).

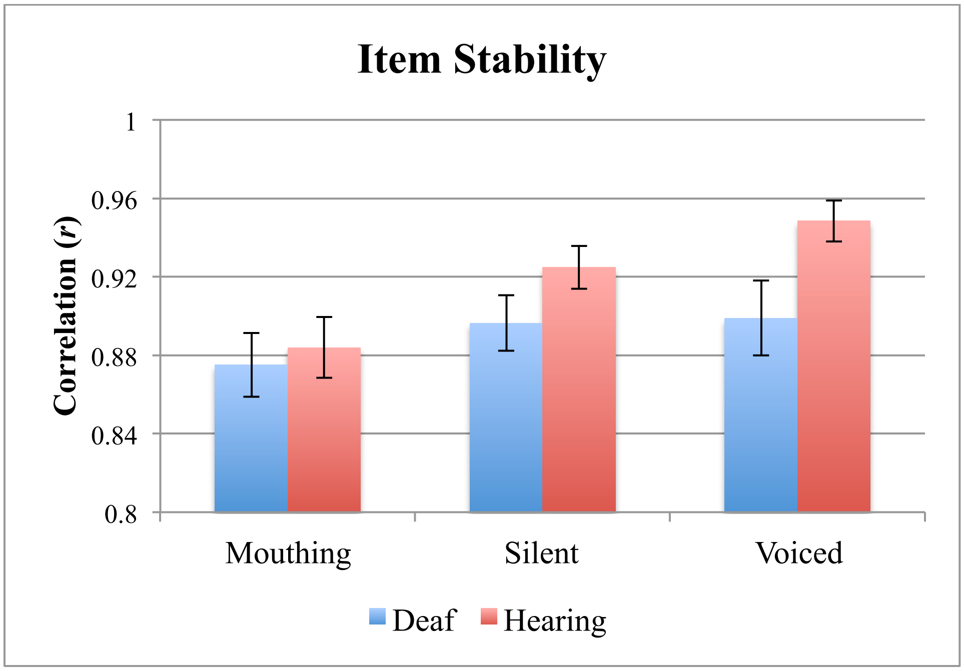


Figure 4.1: Within-item consistency for each group across conditions

To determine whether the visibility of speech affected the stability of mouth productions, a second ANOVA was performed. The design was 2 × 3 × 2 (proportion of visible segments: high, low; condition: mouthing, silent, voiced; group: Deaf, hearing). The results of factors and interactions not involving the new factor of visibility remained the same as above (main effect of condition, no effect of group, but a significant condition × group interaction; *F*- and *p*-values same as above). The ANOVA also revealed that items with a high proportion of visible segments were produced with greater consistency than items with a low proportion: ($r = .917$ vs. $r = .892$, $F(1, 18) = 15.0$, $p = .001$). No interactions involving visibility were significant: visibility × group

($F(1, 18) < 1$, $p = .37$); visibility × condition ($F(1, 18) = 2.7$, $p = .115$); and visibility × condition × group ($F(1, 18) = 1.0$, $p = .328$).

Additionally, a linear regression analysis was performed to determine whether Deaf participants' speech skills were correlated with the consistency of their oral productions. The analysis revealed no relationship between speech scores (see Section 2.2.2) and item stability ($r = .116$, $p = .750$).

4.3.2. Mouthing vs. Voiced Model Analysis

In order to determine whether co-sign mouthings used the same motoric representations as voiced speech, repetitions from the mouthing condition were compared to models of the voiced condition (see Figure 4.2). In this analysis, a model movement signal of the voiced condition for each item and participant was calculated by averaging across all ten repetitions. A cross-correlation was then performed for each repetition within the mouthing condition against the voiced model created for the same item and subject. The results found no evidence that Deaf and hearing participant groups differed in the relationship between these two conditions ($t(18) < 1$, $p = .504$); however, the power of this test was quite low ($\pi = .094$) and the addition of more participants to the study might yield significant results, given that a difference between participant groups is suggested by the item stability results above.

These data were also subjected to an ANOVA that included the factor of the items' proportion of visible segments. The analysis failed to find any significant effects of visibility ($F(1, 18) < 1$, $p = .675$), participant group ($F(1, 18) < 1$, $p = .506$), or a visibility × group interaction ($F(1, 18) < 1$, $p = .764$). Additionally a regression analysis was performed on the mouthing-to-voiced correlations against the Deaf participants' speech scores, but no significant relationship was discovered: $r = .445$, $p = .197$.

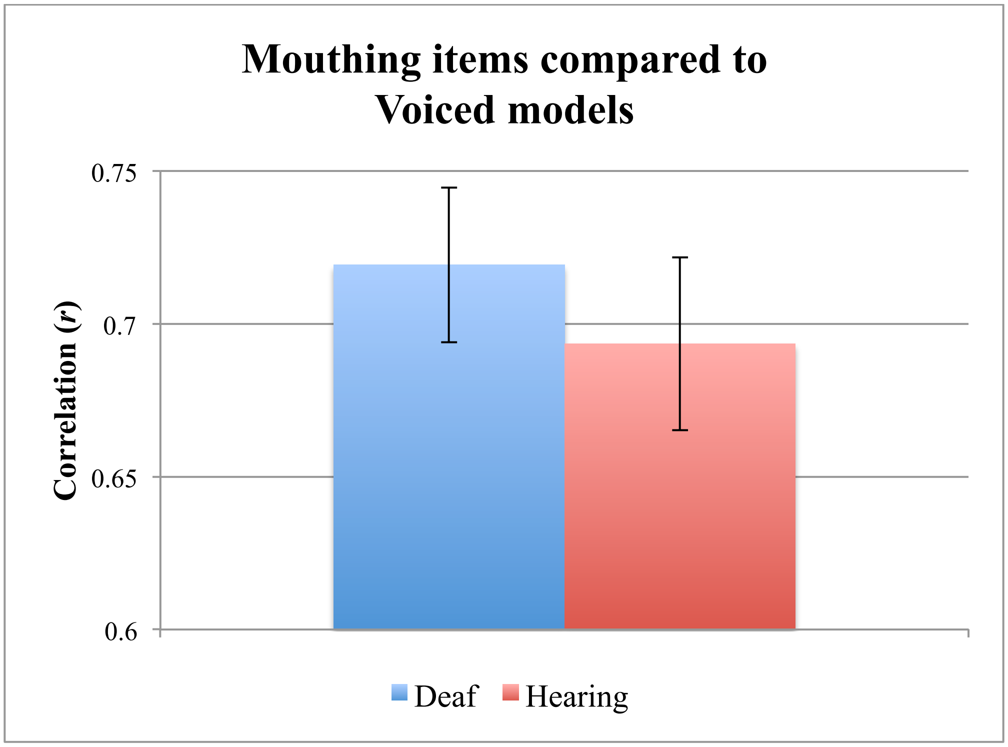


Figure 4.2: Correlations between mouthed items and voiced average models

4.3.3. Silent Condition Analysis

Last, an analysis was performed in order to determine whether the silent speech without signing condition was better associated with the mouthing condition (as both conditions are silent) or the voiced speech condition (as both conditions are produced without sign). Cross-correlation analyses compared repetitions from the silent condition against average models from both the mouthing and voiced speech conditions (see Figure 4.3). A 2 × 2 (comparison model: mouthing condition, voiced condition; group: Deaf, hearing) ANOVA revealed that the silent productions were more strongly correlated with the voiced speech condition than the mouthing condition ($r = .814$ vs. $r = .740$; $F(1, 18) =$

13.3, $p = .002$). There was no effect of participant group ($F(1, 18) < 1$, $p = .77$), and the comparison model × group interaction was marginal ($F(1, 18) = 4.1$, $p = .058$). The marginal interaction appears to be driven by the hearing participant group exhibiting a stronger correlation between the silent and voiced conditions than between the silent and mouthing conditions ($r = .839$ vs. $r = .723$; $t(9) = 3.8$, $p = .004$), but Deaf participants' silent utterances fit equally well against both the mouthing and voiced comparison models ($t(9) = 1.2$, $p = .258$).

These data were also subjected to a 2 × 2 × 2 ANOVA that was the same as above but included the additional factor of the items' proportion of visible segments (high, low). The results of factors and interactions not involving the new factor of visibility remained the same as above (main effect of comparison model, no effect of group, and a marginal interaction; *F*- and *p*-values same as above). The ANOVA also revealed that items with a high proportion of visible segments were more strongly correlated with their models than items with a low proportion of visible elements ($r = .797$ vs. $r = .756$; $F(1, 18) = 5.9$, $p = .026$). No other interactions involving visibility were significant: visibility × group ($F(1, 18) < 1$, $p = .564$); visibility × comparison ($F(1, 18) = 3.4$, $p = .083$); and visibility × comparison × group ($F(1, 18) < 1$, $p = .821$).

Finally, another linear regression analysis was performed to assess if Deaf participants' speech skills were related to the correlation between the silent and other two conditions. The analysis revealed no relationship between the speech scores and correlations ($r = .212$, $p = .557$).

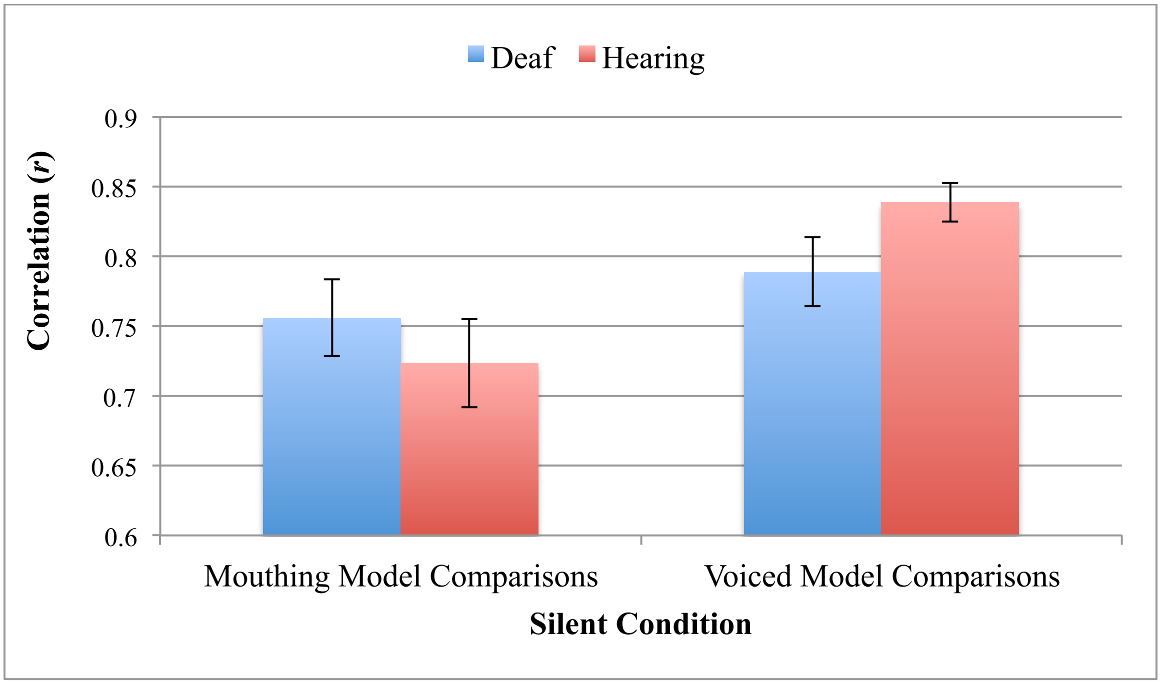


Figure 4.3: Silent condition modeled against averages from the other two production conditions

4.4. Discussion

Overall, the Deaf participants behaved similarly across all three experimental conditions, suggesting that they use one common motor representation for all tasks involving mouth movements regardless of target language. When producing spoken English or articulating co-sign mouthings, Deaf individuals appear to rely on the same motor plans to execute these oral movements. However, these findings cannot distinguish if one linguistic system originates the motor representations while the other system merely relies on the representations of the first system. It could be that because mouthings derive from the surrounding spoken language, Deaf individuals draw on their knowledge of English – regardless of how closely these representations conform to those of native, hearing speakers of English – in order to produce mouthings while signing in

ASL. But it is also possible that as native signers, these Deaf participants have representations of mouthings that are inherent to their sign system and were acquired prior to their exposure to the phonological system of English, as has been claimed for Finnish Sign Language (Rainò, 2001; Tapio & Takkinen, 2012). However, from the limited work on the psycholinguistics of mouthings, evidence suggests that mouthings are stored and retrieved separately from lexical signs and therefore are borrowed from spoken language representations rather than the other way around (Vinson et al., 2010). Furthermore, the English words used in this study are too infrequent to assume that children would acquire them in ASL before their exposure to written English.

In contrast to the Deaf participants, the hearing group showed larger differences between experimental conditions, especially between the mouthing condition and the other two conditions, which did not require them to sign. Although this could be interpreted as the hearing participant having two distinct motor representations for mouth movements, one for co-sign mouthings and another for speech, this is unlikely as sign-naïve participants come to the tasks with no representations specific for mouthings and cannot create such representations in the span of the task. Instead, the findings from the hearing group likely reflects a novelty effect in which hearing people are unskilled in combining manual and oral movements in such tight temporal alignment (as in Experiment 2). Therefore, the mouthing condition is the most dissimilar of the three conditions for the hearing participants because they lack the motor finesse to produce consistent co-sign mouthings, but as experienced speakers of English, their mouth movements are more stable during the non-signing conditions. There may also be an effect of increased cognitive load in the co-sign mouthing condition, as the hearing participants had to recall the signed form, though was likely ameliorated by the fact that all ten repetitions occurred in succession, so the burden on short-term memory should have only taxed the first repetition.

The experiment also spoke to the role of visibility of mouth movements in that items with a high proportion of visible segments exhibited greater correlations across production conditions than items with a low degree of visible elements. This may suggest that Deaf participants exploit the visibility of the speech produced around them and organize their motor representations of speech based on these visual properties. However, such an interpretation must be tempered with two important caveats. First, the effect of visibility on the correlation scores was not mediated by participant group; none of the visibility × group interactions was statistically significant. Since hearing participants also exhibit a visibility effect, it seems more likely that the effect is a result of the recording procedure. Because the motion capture system only records the visible articulators of the body (the lips and jaw), items that recruit these articulators in greater proportion may result in more robust and consistent articulations even across distinct production conditions. Second, because item visibility was confounded with item length, it is also possible that an item's length drives the observed effect. Longer items, which are here coded as low visibility, may introduce greater variability in the movement signal and result in utterances that are less correlated across conditions (Maner et al., 2000). Further study with more carefully controlled stimuli is necessary to resolve this unfortunate confound in the experimental design.

The greatest contribution of this study is that it supports the long-held assumption that mouthings and speech rely on the same representation, which has never been directly tested before. Mouthings are frequently reported using spoken language transcriptions and analyses of processes that affect mouthings – such as mouthing reductions – implicitly presume that a Deaf individual's target mouth movements should be the same for mouthings and speech, and therefore any deviation from speech observed in mouthings must be explained in terms of the spoken language phonology. While Experiment 1 offers evidence that other articulatory factors may contribute to these differences, the present study validates the application of spoken language phonology on mouthing productions. While the findings from the hearing group suggest that they might have distinct representations of speech and sign, the findings are difficult to interpret without also obtaining data from hearing signing groups in order to distinguish a novelty effect from an audiological effect or even a dual task effect (Dromey & Bates, 2005; Dromey & Benson, 2003).

Finally, the role of speech visibility warrants further investigation. This experiment offers limited support for the role of visibility in Deaf oral representations and productions, this area of study must await more carefully controlled investigations to corroborate the effect and to better understand how hearing individuals may also be sensitive to this variable. If speech visibility is indeed discovered to affect the strength and precision of a Deaf individual's representation and production of speech sounds, this could possibly be used to improve speech therapy intervention techniques.

This chapter, in part, is currently being prepared for submission for publication of the material. Udoff, Jonathan; Nip, Ignatius; and Emmorey, Karen. "The Motor Representation of English Speech and American Sign Language Mouthings". The dissertation author was the primary investigator and author of this material.

Chapter 5: General Discussion

Through the three experiments presented in this dissertation, two broad issues concerning co-sign mouthings were explored. The coordination between the hands and mouth was investigated in Experiment 1 (fingerspelling) and Experiment 2 (lexical signs). Then, in Experiment 3, the motor representation of mouthings was investigated in comparison to voiced speech.

5.1. Hand-Mouth Coordination

Hand and mouth coordination was assessed in two experiments of this dissertation. In Experiment 1 (Chapter 2), precise three-dimensional data of the mouth and dominant hand were recorded from native Deaf signers of ASL. The participants mouthed and fingerspelled nonce words in which the number and direction of articulatory movements of the hand and mouth were manipulated. The results suggest that fingerspelled forms exhibit tight spatial and temporal coupling of manual movements and oral movements when producing simultaneous mouthings. In particular, the hand and mouth demonstrated greater spatial and temporal coordination when the two articulators share the same number of movements and when they each produced fewer movements as opposed to numerous movements. Moreover, the direction of an articulator's movement in relation to the other did not affect the coordinative effect. For example, a closing movement of the hand combined with an opening movement of the mouth (e.g. B-A + /bɑ/) shares the same degree of inter-articulator coordination as a closing movement of the hands combined with a closing movement of the mouth (e.g. U-N + /ʌn/). Lastly, in the cases in which the hand moves twice while the mouth only moves once, the oral movement appear to stretch over the entire duration of the manual movements rather than aligning with only one of the two movements of the hands.

Work on the coordination of grasping actions with simultaneous syllable productions by hearing participants suggests that these types of coordinative effects observed in the fingerspelling data may generalize from signers to non-signers alike (Bernardis & Gentilucci, 2006; Gentilucci et al., 2006; Gentilucci, Dalla Volta, & Gianelli, 2008; Gentilucci & Corballis, 2006; Gentilucci, Santunione, et al., 2004; Gentilucci & Volta, 2007). However, because hearing sign-naïve individuals struggle to produce fingerspelling sequences – they tend to "bounce" each letter instead of smoothly transitioning between handshapes, and they also tend to mouth the name of the individual letters rather than the phonological word as a whole – the results cannot offer direct evidence whether these effects apply only to experienced signers or if they generalizes to all people with an intact neuromotor system and articulators. Coordination effects, whether for speech combined with a non-linguistic motor task or for production of a multi-effector language like ASL, are likely the neural consequences of the cortical proximity of hand and mouth motor areas as well as the common circuitry involving Broca's area, rather than the results of prolonged use of a signed language.

The second hand-mouth coordination study substituted nonce fingerpelled items with real lexical signs of ASL in order to increase the ecological validity of the study. Given the nature of the sign stimuli, it was not possible to capture the data using the same technique as in the previous study, so the data analysis differed considerably for this second experiment. However, the task demands were also decreased from the first experiment, so it was possible to directly compare native signers to sign-naïve participants. In comparing mouthing instances between two manual homophonous signs of ASL, three basic findings emerged from the data. First, evidence was presented in which the movements of the mouth affected the simultaneous movements of the hands, as longer mouthing specification durations induced longer sign durations, for both hearing and Deaf groups. Although there is considerable research that emphasizes how the mouth is subordinate to and affected by the movements of the hands (e.g., Boyes Braem & Sutton-Spence, 2001; Gentilucci & Corballis, 2006; Gentilucci et al., 2008), this finding supports the idea that the influence between hand and mouth motor regions is in fact bidirectional (Gentilucci et al., 2001). This finding is particularly important within the domain of sign language research, as the dominant view is that the "hands are the head of the mouth" (Boyes Braem & Sutton-Spence, 2001), a maxim that reflects the belief that the oral component of sign is a dependent of the manual component and cannot exert influence on the hands. However, Experiment 2 provides evidence that alternations in mouthing productions can indeed induce concomitant changes in manual production.

Second, hearing participants used a different overall strategy to coordinate the productions of the oral and manual components of a sign as compared to Deaf signers. Specifically, hearing non-signers usually began an utterance with the manual component alone and the oral articulation lagged behind whereas Deaf signers were just as likely to use this strategy as they were to start instead with the mouthing component or to start with both components synchronized. This difference between the two participant groups is interpreted as hearing non-signers experiencing a novelty effect. Because they are not accustomed to having to produce communicative gestures across multiple parts of the body in tight synchrony, these hearing participants may start each repetition with the more unfamiliar component, the manual sign, and once the motor plan for the hands began its execution, then the more familiar oral motor plan was executed. Hearing non-signers are not wholly unaccustomed to producing hand movements in concert with speech, however, as is the case in the production of co-speech gestures. But co-speech gesture is crucially different from signing in that these manual gestures do not necessarily align with a specific movement or movements of the vocal tract (Wagner et al., 2014). A co-speech pointing gesture produced with the voiced utterance "Go to the back of that line," is not better or worse when the manual gesture is initiated with the word "go", "back", "that", or "line". Co-sign mouthings (as well as grammatical non-manual gestures), in contrast, have specific temporal domain and must align to the corresponding word (or phrase, in the case of non-manual markers) (Baker-Schenk, 1983; Liddell, 1980; Neidle, Kegl, MacLaughlin, Bahan, & Lee, 1998). Therefore, the motor strategies developed by hearing gesturers to handle interarticulatory coordination may not be adequate to produce fluent co-sign mouthings given its stricter timing constraints.

Third, despite the differences between the baseline strategies of coordination between the two participant groups, in which hearing participants are more likely to begin an utterance with just the manual articulation than Deaf participants, both groups exhibited the same effect between the short and long mouthing length conditions. As the mouthing component grew in length, all participants became more likely to start an utterance with the mouthing component and less likely to start an utterance with the manual sign component. This observation suggests that though language experience may play a role in how an individual approaches the problem of coordinating the movements of multiple motor systems during sign language production, factors intrinsic to the linguistic form of the utterance, like length, have universal effects on the signer as mediated by a common neuromotor system shared by native Deaf signers and hearing non-signers alike.

In comparing the results between these first two experiments, the question arises of why the findings seem to be different. Given that the hands and mouth are so tightly coupled in time for fingerspellings, why do we then find that the mouth movements can start before and finish after the manual articulation? The difference here can be explained by the differing analyses across the two experiments. Since the movements of the hands for the manual homophone experiment were measured only through videotape, the data lacked the necessary precision to directly compare to the fingerspelling data, with its high temporal resolution (29.97 vs. 120 frames per second). Therefore, the homophone analysis primarily focused on whether the manual or oral components were offset from each other by more than 33.4 ms, the video recording's instrument margin of error. Taking this difference in precision into account, a direct comparison of the temporal coordination effects in each experiment yields remarkably similar results. For the Deaf participants, the manual movements of fingerpelled items lagged behind the oral movements by 8.7 ms (signed lag scores; SD = 116.7 ms), and the manual movements of homophones lagged behind hand movements slightly after their mouth movements, by 15.2 ms (SD = 93.6 ms). In contrast, hearing non-signers produced their utterances with the hand movements leading the mouth movements by 88.4 ms (SD = 183.5 ms). Though these findings must be interpreted with caution due to the issue of instrument precision, they imply that the temporal coordinative effects are indeed similar between the two experiments. This result further suggests that the hands and mouth coordinate their movements in time in the same fashion among all substrata of the sign language, including lexical signs and fingerspellings, during the production of mouthings.

5.2. The Motor Representation of Mouthings

Based on the findings from Experiment 3, it appears that Deaf individuals do not maintain separate motor representations of co-sign mouthings that are distinct from their motor representation of voiced English. Although it remains an open question whether mouthings are borrowed from speech representations or if instead speech production draws on mouthing representations that were established before the acquisition of spoken language, the common assumption in the mouthing literature is that mouthings and their reductions can be analyzed in relation to the phonology of speech appears to be well founded, as Deaf participants' oral movements for silent co-sign mouthings were highly correlated to their oral movements for voiced speech, regardless of the direction of this relationship.

This interpretation is limited by the findings from the hearing non-signer group. In contrast to the shared motor representation demonstrated by the Deaf signers, sign-naïve participants appear to produce mouthings differently than spoken English, counter to predictions. It is safe to assume that this difference in productions does not reflect differing underlying representations because the hearing group has no previous knowledge of ASL. Instead, it is suggested that the non-signers are unskilled at sign production, like due to the strict timing requirements between articulators, which introduces variability into their co-sign mouthing articulations. The inclusion of hearing participants who have some command of ASL is needed to further address this issue. Novice hearing signers have adequate skill in producing sign language, but do not have the life-long experience of communicating in ASL nor the audiological depravation of Deaf signers. Therefore, if novice signers differ from the sign-naïve group while performing similarly to the Deaf group, it would support the conclusion that unfamiliarity with sign language production is the cause of the hearing non-signers' pattern of results. However, if hearing signers pattern like hearing non-signers it would suggest that Deaf individuals' representations of oral articulations are fundamentally different that hearing individuals' representations, regardless of signing ability.

5.3. Visibility and Visemes

Findings from Experiment 3 suggest that the degree of visibility of a mouthing affects the stability of its internal motor representation as well as how correlated its production was with the production of the analogous spoken word. Because deaf individuals, by and large, do not have access to the acoustic properties of spoken language, the visual modality is their primary means of acquiring spoken language phonology. The degree to which speech visibility plays in developing and retrieving mental representations of spoken language, however, remains an open question. Results presented in Chapter 4 offer limited evidence that Deaf individuals' motor representations of words with highly visible content were more robust and more closely matched their hearing peers' representations than words with low visibility: correlations between mouthing and speech conditions as well as inter-item stability measures were higher for high visibility items than for low visibility items. Therefore, the visibility of a spoken item may play a large role in determining how well a deaf individual will acquire, understand, and reproduce speech (Huntington et al., 1968; McGarr & Gelfer, 1983; McGarr & Harris, 1983). Further investigation is warranted to explore how manipulation of speech visibility may impact outcomes for speech therapy and intervention.

Furthermore, as the formal division of phonemes into visemes, the categorization of speech sounds based on their visual properties alone, is currently being refined, these findings shed light on how visemes may play a functional role in the lexicon. Though neglected in the literature for many decades by speech scientists, sign language and deafness researchers have rediscovered the utility of applying a viseme perspective (see Elliott, Braun, Kuhlmann, & Jacobs, 2012), and visemes-based model have aided the development of automatic sign language recognition systems (e.g., Leszczynski, Skarbek, & Badura, 2005; Schmidt, Koller, Ney, Hoyoux, & Piater, 2013). However, there is still little agreement how the phonological units of spoken English are grouped into visemes, so comparisons across studies must be scrutinized with special attention to these categories (see Koller, Ney, & Bowden, 2014 for an example of different viseme groupings of German). The discrepancies between viseme categories, however, did not affect the findings from Experiment 3 because phonemes were merely labeled as highly visibility or not, without the need to define its exact corresponding viseme.

The findings concerning speech visibility from Experiment 3 must also be interpreted with caution. Though speech visibility had an effect on Deaf participants' productions, these effects were observed for hearing participants too. This observation may indicate that the finding was an artifact of the data capture technique (see Section 4.4), but it could also reflect a confound that exists in the design of the stimuli. Item length was inversely co-varied with speech visibility so the findings described above may be the result of length effects (such as those discussed in Chapter 2) rather than a direct result of speech visibility. There is some support that length should not play a role in analyses performed in Chapter 4 (see Sadagopan & Smith, 2008), but the evidence is mixed (see also Kleinow & Smith, 2006; Maner, Smith, & Grayson, 2000)

5.4. The Relationship between Speech Skills and the Production of Mouthings

Because co-sign mouthings are derived from a surrounding spoken language, it stands to reason that a signer's spoken language training and abilities may influence his mouthing productions. If a signer has had extensive overt instruction in the phonological structure of speech, that skilled Deaf speaker may show strong correlations of their lip movements between the co-sign mouthing and voiced speech conditions in Chapter 4's experiment. The same skilled Deaf speaker may also have built a more robust motoric representation of mouthings that are more resistant to the coordination effects demonstrated in Experiments 1 and 2.

However, this pattern was not observed in the data. All regression analyses between correlation scores and the composite speech score failed to reach statistical significance. Instead, it appears that the number of years spent in speech training and a participant's self-reported use and comfort with speech and speechreading have no affect on his oral articulations in ASL or English nor on how those articulatory gestures coordinate with the articulations of the hands. This is further evidence that these coordination effects reflect physiological and neuromotor generalizations rather than being mediated by linguistic experience.

5.5. Conclusion

This dissertation has addressed two issues surrounding co-sign mouthings. First, by comparing the movements of the mouth during mouthing production against mouth movements during voiced speech production, it appears that Deaf signers use one common motor representation for both mouthings and speech. Second, in comparing the movements of the mouth with simultaneous movements of the hands, new evidence is offered to the field demonstrating the tight temporal and spatial coordination of mouthing components with their manual counterparts, be they lexical signs or sequences of fingerspelled letters.

This new account of mouthings in which productions are shaped in part by interactions with manual articulations has many implications across multiple fields of research. In particular, models of sign language phonology and phonetics must address these kinds of inter-articulatory coordination effects by adopting some formal mechanism that allows for bidirectional interaction between the hands and non-manual components of sign production. This mechanism should also be able to account for coordination effects of movement shape as well as the timing alignment between the two articulators. Such a model may ultimately resemble the framework of autosegmental phonology in speech, in which different articulators occupy separate tiers that interact to align their productions (Goldsmith, 1990). Machine translation also can benefit from this new perspective of the hands and mouth in the production of sign language. Speech-to-sign translation using animated signing avatars can make their avatars appear more naturalistic and human if they account for these kinds of interactions between different parts of the body. Moreover, automatic sign language recognition systems can provide more reliable and accurate computations if they incorporate how inter-articulatory coordination affects a signer's movements.

5.6. Future Directions

The experiments described in this dissertation are a first step in exploring the representation of mouthings and speech in Deaf individuals and the coordination of the mouth and hand in sign language productions. To strengthen the claims made here, I propose several modifications to the experiments' design as well as suggest new avenues of investigation. The manual homophone experiment presented in Chapter 3 should be redone with a motion capture system that does not rely on line-of-sight measurement so that the system can track the markers on the hands as well as the face even during periods of visual occlusion. Once reliable three-dimensional data is obtained, kinematic analyses can be performed on the sign data, as was done in Experiment 1, and a more direct comparison can be made between that experiment and the fingerspelling experiment from Chapter 2. Hopefully, this new study will corroborate the principles of motor coordination discussed in Section 2.4, demonstrating the universality of these principles across participant groups and signing tasks.

Next, it would be beneficial to redesign the stimuli list from Experiment 3 so that the factor of speech visibility is not confounded with item length. This confound in the current form of the experiment hinders the interpretation of speech visibility on Deaf individuals' representation and production of speech. With a redesigned elicitation protocol that includes a carrier phrase with the necessary kinematic landmarks, there is more freedom in stimuli selection so that items with lower proportions of visible speech elements do not necessarily have to be longer than items with higher visibility proportions.

Additionally, I propose extending all three of the dissertation experiments to include a participant group of hearing signers. As the experiments are designed now, the fingerspelling experiment has no manipulation of participant group, and the participant group manipulation for the other two experiments confounds language experience with audiological status. Where the two groups differ, then, it is not clear whether the difference is driven by inexperience with sign by auditory deprivation. I suggest using hearing, recent learners of sign to tease apart these two issues. These hearing signers would be proficient in ASL, but also have a command of spoken English equal to the non-signing group so that, between the hearing signers and hearing non-signers, only sign knowledge differs, and between the hearing signers and Deaf signers, only audiological status differs. I suggest using recent learners of ASL because not only are they easier to recruit for participation, but they also speak to my hypothesis that the differences observed in the hearing non-signer group's performance is caused by the novelty of the signing task. If indeed the performance of new learners of sign patterns with Deaf, native signers, it would indicate that only minimal exposure to ASL is enough to induce the body to reorganize and coordinate its movements across articulators rather than requiring life-long, native experience with the language. Such a finding would remove the burden of explaining these coordinative effects within linguistic models of sign language, relying on models of motor control to account for interarticulatory coordination. However, if this hypothesis does not bear out, and novice signers pattern in a different way – either similarly to hearing non-signers or in a way unique to themselves – I would then recruit a second group of hearing signers, who would be native users of ASL. Hearing (adult) children of Deaf adults are exposed to sign language from birth, but as hearing individuals they also acquire the surrounding spoken language. The performance of this group could then elucidate whether native sign language abilities or audiological status is the crucial element of the Deaf signers' results.

Additionally, I would like to extend my line of inquiry by applying the principles developed in this dissertation onto a corpus of spontaneous signing, increasing the ecological validity of the findings presented here. In particular, I believe mouthing reductions would be a fruitful domain of research in which to apply the findings from the coordination experiments. Though this dissertation did not directly manipulate or account for mouthing reductions, it should be noted that the phenomenon of mouthing reduction inspired the experiments presented in this dissertation and reductions were observed in the data by both Deaf signers and hearing non-signers alike. By examining the potential influences of manual movements on the production of mouthings, it is possible that we can develop a better account of mouthing reductions that will more accurately predict when a reduction will occur and how the reduction will be shaped. For instance, the alterations observed in NGT mouthing reductions may relate to the phonology of the sign

(Bank et al., 2011): a two-syllable Dutch-derived mouthing may only reduce to one syllable when the accompanying sign is made with one path movement (monosyllabic) whereas another two-syllable mouthing will not reduce because its accompanying sign has two path movements (disyllabic). The factors of hand–mouth coordination explored in this dissertation, namely absolute number of movements and congruency of number of movements between articulators, may dictate the incidence and specific shape of mouthing reductions.

In order to corroborate and extend the findings presented in this dissertation, several avenues of investigation are possible. To pursue the finding that Deaf signers represent mouthings as the same as speech, other methodologies may be employed to measure the vocal tract. EMMA (electromagnetic midsagittal articulography) is particularly attractive because it can record speech articulators inside the vocal cavity and will be able to address the role speech visibility on oral productions. Such a technique would elucidate if the visibility effect observed in Experiment 3 was caused by instrumental bias due to the optical recording paradigm of the motion capture system or if Deaf signers are more consistent in their production of visible items even when measured at a more dorsal site of the vocal tract.

Finally, pursuing the findings of interarticulatory coordination, it would be worthwhile to continue to use motion capture techniques to study other articulators involved with sign language production. By modeling the interactions of multiple parts of the body – head, eyebrows, eye gaze, hips, to name a few – a more complete model can be established for the production system of sign languages. In addition to the benefits such a detailed model would have for machine translation of sign, this body of knowledge would have implications for non-signing situations as well. A better understanding of how the neuromotor system coordinates and plans simultaneous motor activity across the body can be applied to numerous fields, from kinesiology to the diagnosis and treatment of movement disorders, both in hearing and Deaf populations. Research in co-speech gesture may also benefit from this type of work and may initiate new analyses of how speech and gesture interact at a more fine-grained temporal scale. Additionally, it is unknown how these coordinative skills develop so it would be particularly valuable to study these questions, for instance, with young children acquiring ASL as well with adults who are just learning to sign. These kinds of research programs will then begin to glean new insights into how the neuromotor system multitasks the various demands of simultaneous interarticulatory movement with apparent ease, which in turn will facilitate the development of new technologies and therapeutic interventions.

APPENDICES

Appendix A: Speech Background Questionnaire

Questions from the background questionnaire relating to speech production and speechreading that were used for the composite score calculation are reproduced below. Only Deaf participants completed this questionnaire. Data from questions 3 and 10 were converted to a Likert scale ranging from 0 (never) to 6 (all the time). The values for each row were then summed before the z-score transformation and weighting. Data from question 5 was compiled by multiplying the two column values of each row and then summing the row products. In parentheses and to the left of each question number, the weights used in the composite score calculation are listed. The weights were chosen so they reflect the amount of data each question contributes to the whole questionnaire. For example, questions 3 and 10 ask for multiple responses, so each data point was weighted double. Because formal, explicit speech instruction was hypothesized to have the most significant impact in a Deaf individual's knowledge and skill with speech, question 5 was weighted so that its data contributed to half of the total composite score. A regression analysis determined that the data obtained from the speech instruction question (#5) was highly correlated with the remaining, more subjective, measures ($r = .78$, $p = .008$).

QUESTIONS ABOUT USE OF SPEECH

In this section we would like you to think about the extent to which you use speech and/or lipreading to communicate.

(1) 1. How comfortable do you feel using speech?

1 (not at all)	2	3	4	5 (very much)

(1) 2. How much do you think that your speech is understood?

1 (not at all)	2	3	4	5 (very much)

(2) 3. How often do you currently use speech when communicating with the following people:

	All the time	Once a day	Few time a week	Once a week	Once a month	Less than once a month	Never
Family							
Friends							
Coworkers							
Strangers							

(1) 4. How long have you been using speech (circle one)?

Since before I started school

Since elementary school

Since middle or junior high school

Since high school

Since I left high school

(12) 5. How much speech training/speech therapy have you had?

	Hours per week (approx.)	Number of years
Pre-school		
Elementary school		
Middle school/Junior high		
High school		
After high school		

(1) 6. How often do you mouth English words while signing to:

A deaf person? 1 (never) 2 3 4 5 (always)

A hearing person? 1 (never) 2 3 4 5 (always)

(1) 7. How many of your signs are accompanied with English mouthing when communicating with:

A deaf person? 1 (none) 2 3 4 5 (every sign)

A hearing person? 1 (none) 2 3 4 5 (every sign)

(1) 8. How much do you currently use lipreading to communicate with hearing people?

1 (not at all) 2 3 4 5 (very much)

If not at all, when did you start/stop?___________________________________
(If you circled not at all, please skip remaining questions)

(1) 9. On average, how much do you understand when lipreading:

	1	2	3	4	5
A deaf person?	1 (not at all)	2	3	4	5 (very much)
A hearing person?	1 (not at all)	2	3	4	5 (very much)

(2) 10. How often do you currently use lipreading when communicating with the following people:

	All the time	Once a day	Few time a week	Once a week	Once a month	Less than once a month	Never
Family							
Friends							
Coworkers							
Strangers							

(1) 11. How long have you been using lip-reading (circle one)?

Since before I started school

Since elementary school

Since middle or junior high school

Since high school

Since I left high school

Appendix B: Graphs of Fingerspelled Items

In order to create average signals for the hand and mouth, each participant's raw aperture signal was first time-normalized into 1000 equally spaced data points, using spline interpolation. The average movement signal was then calculated as the average point-by-point value across participants. Because the distance between the lip markers is so much smaller than the distance between the hand markers, it was also necessary to amplitude-normalize the two movement signals so that each articulator would be visible on the same graph. The amplitude normalization was calculated by the feature scaling formula

$$f'(t) = \frac{f(t) - \min[f(t)]}{\max[f(t)] - \min[f(t)]},$$

so that each transformed signal has a global minimum of 0 and a global maximum of 1 while preserving the general shape of the signal.

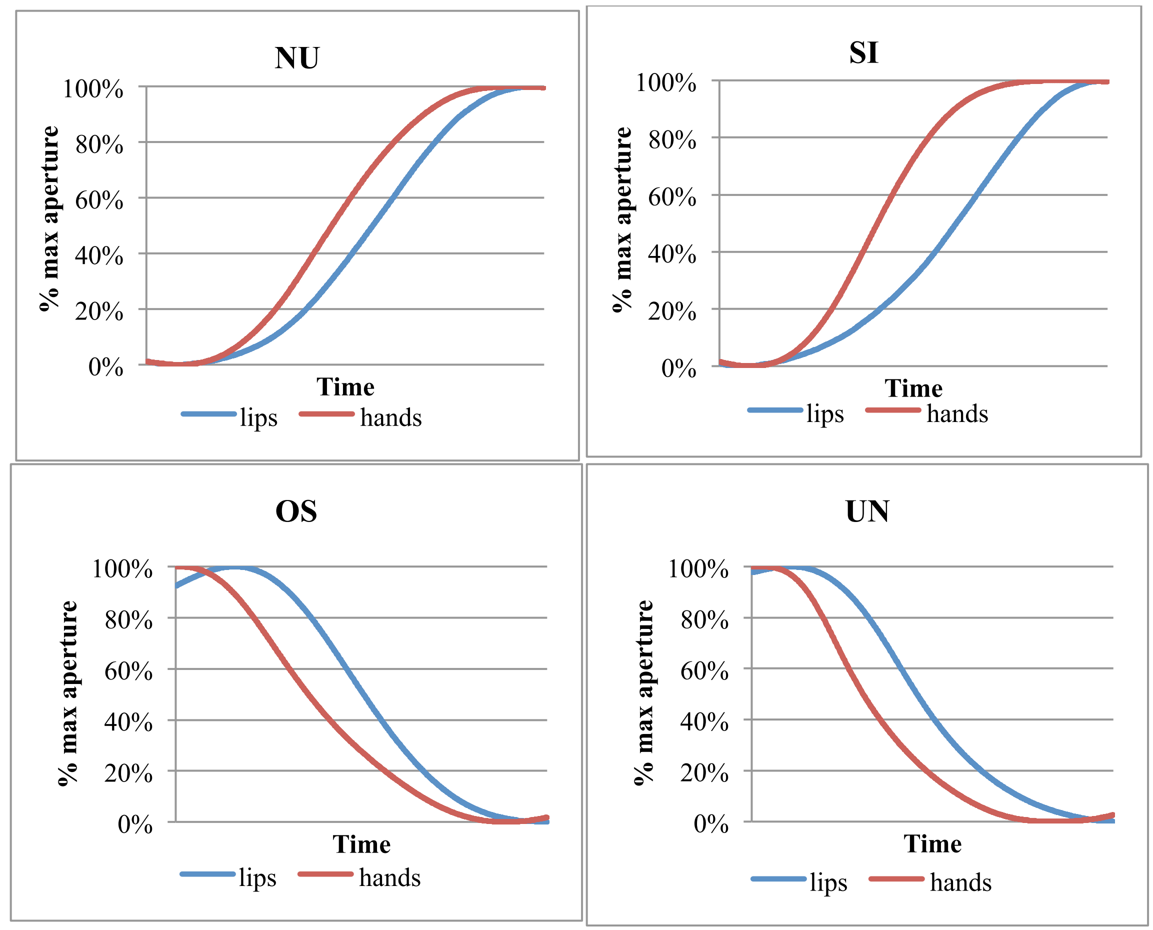


Figure B.1: Individual items graphs of simultaneous hand and mouth movements when each articulator moves once in the same direction. Graphs are averaged across participants.

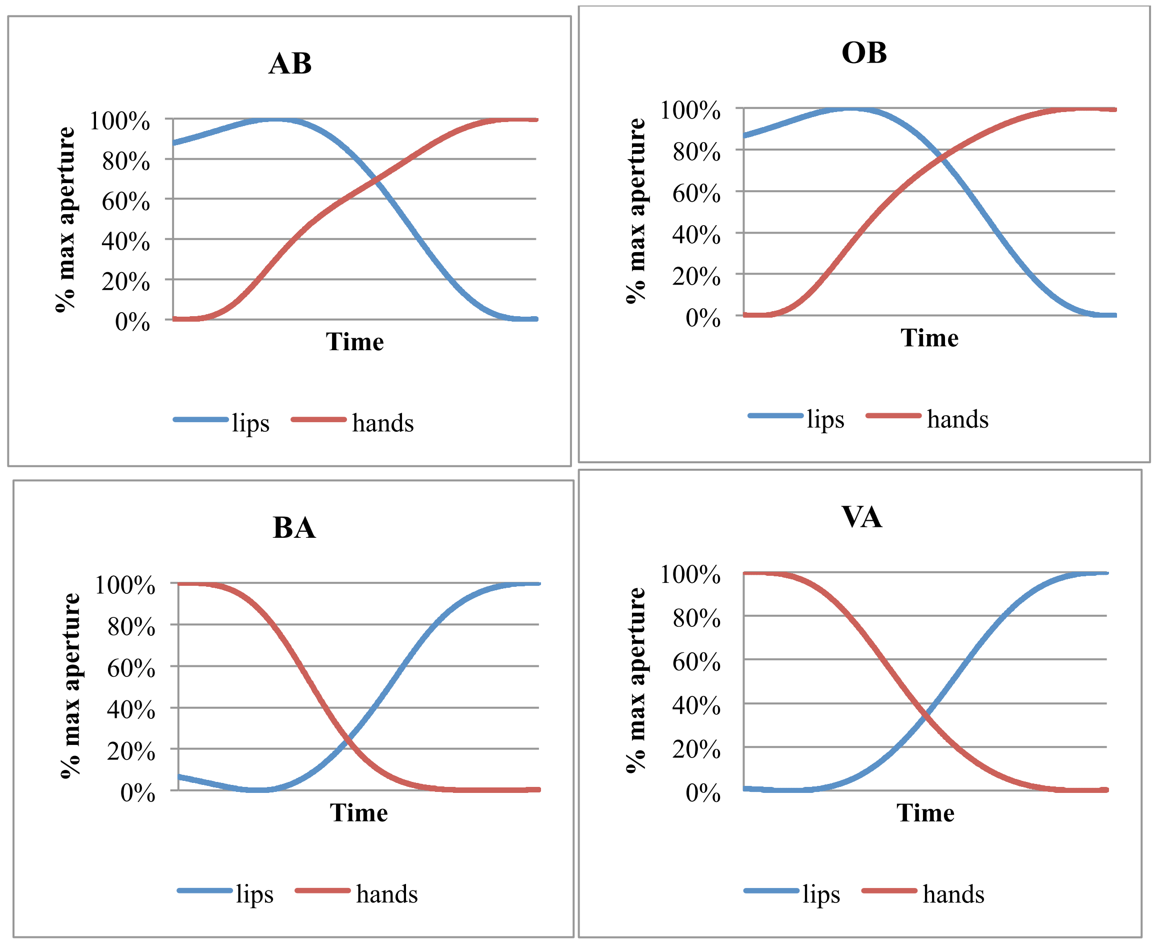


Figure B.2: Individual items graphs of simultaneous hand and mouth movements when each articulator moves once in the opposite directions

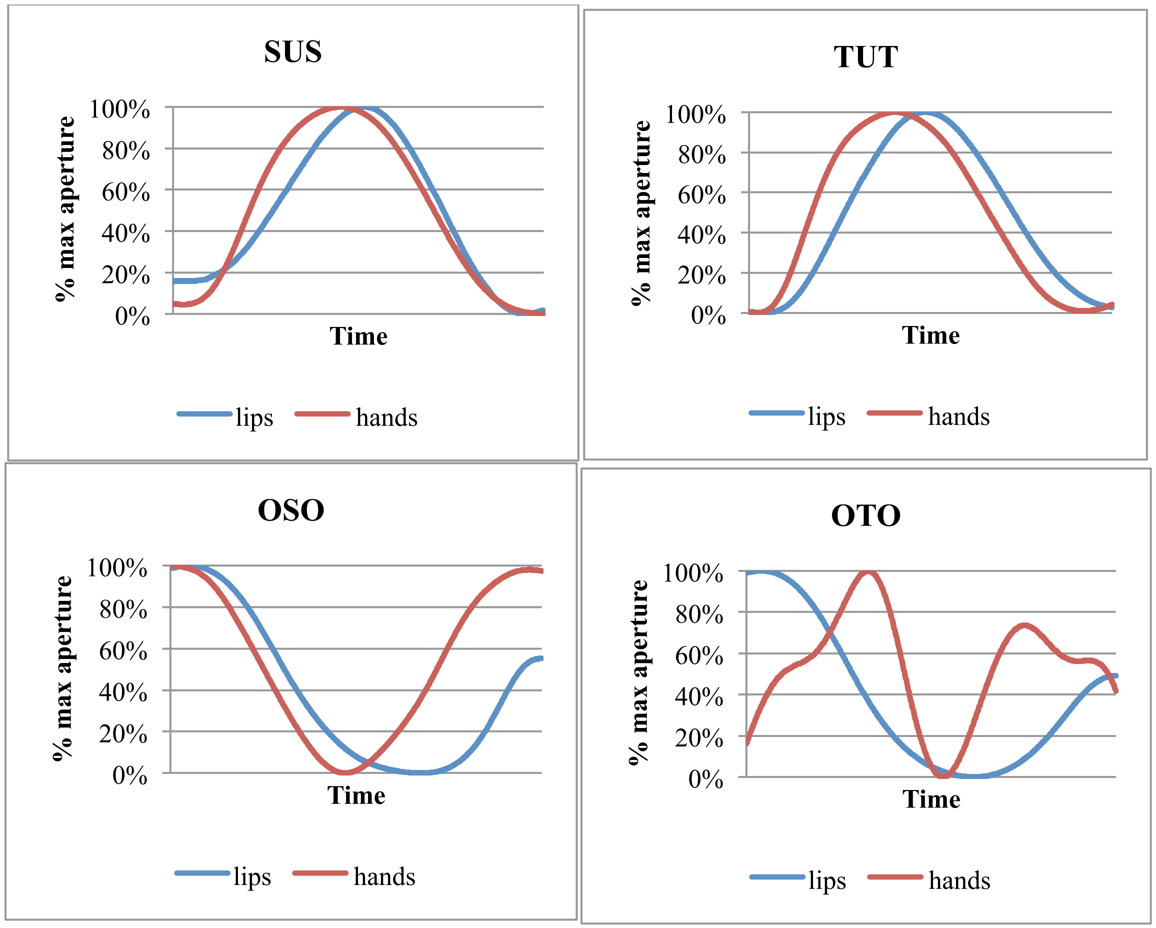


Figure B.3: Individual items graphs of simultaneous hand and mouth movements when each articulator moves twice and in the same direction

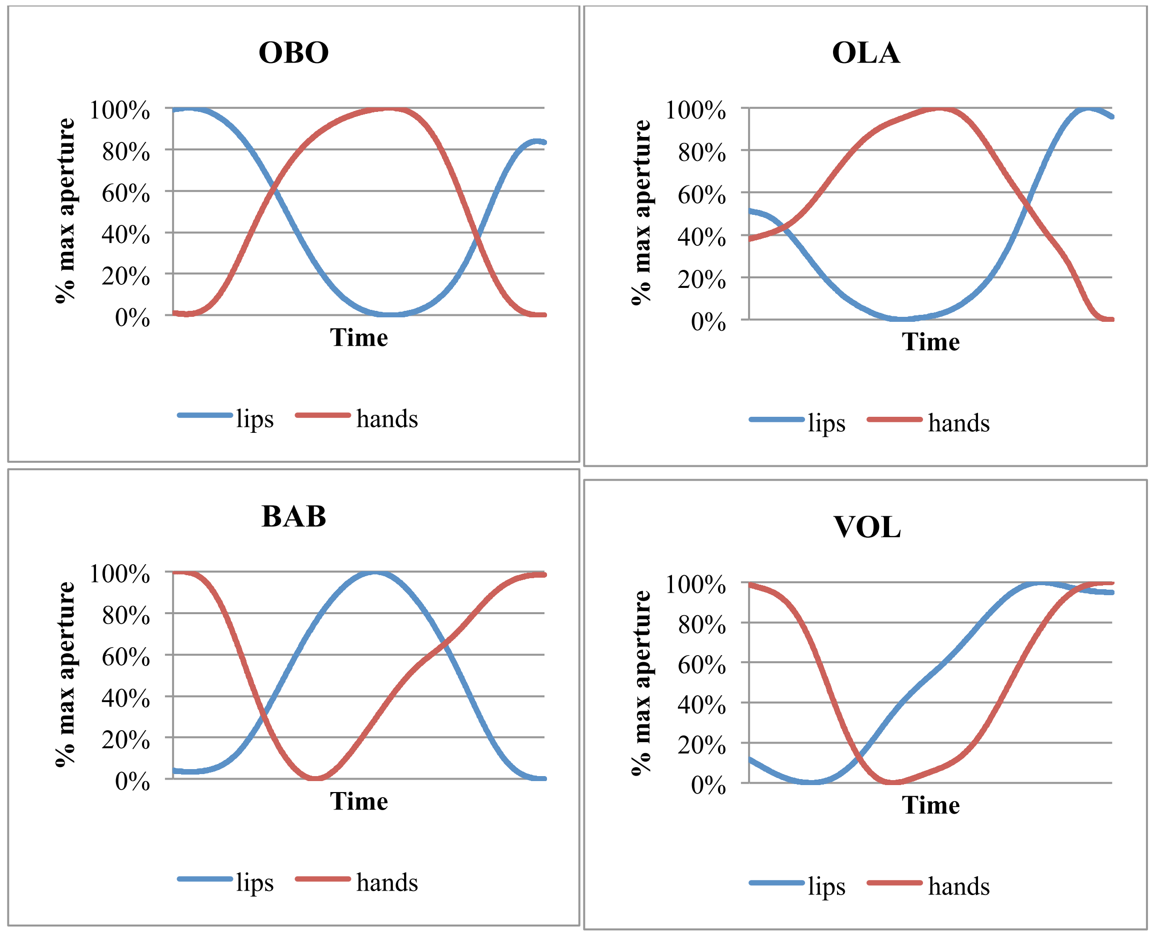


Figure B.4: Individual items graphs of simultaneous hand and mouth movements when each articulator moves twice in opposite directions

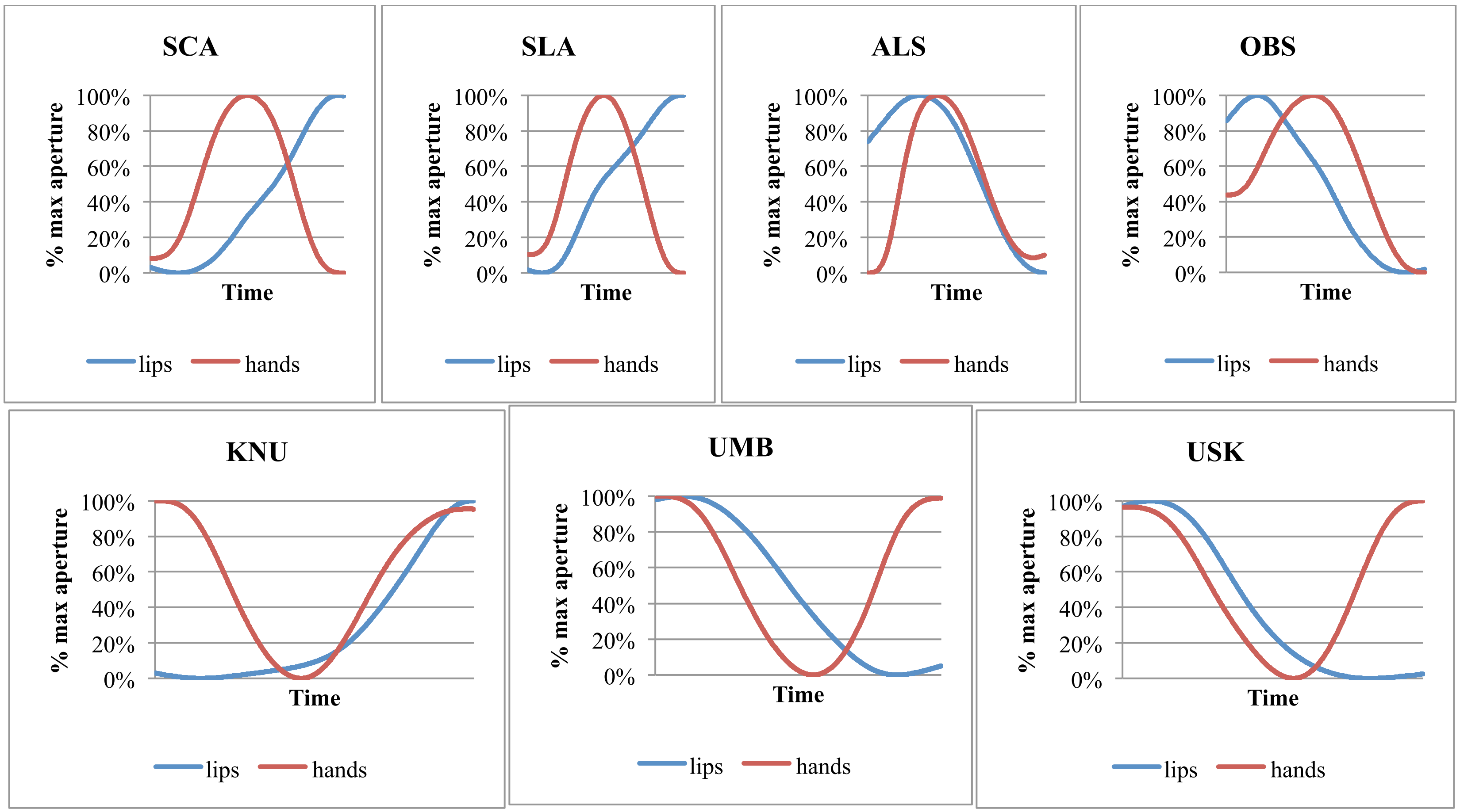


Figure B.5: Individual items graphs when the hand moves twice, but the mouth only moves once

References

Ajello, R., Mazzoni, L., & Nicolai, F. (2001). Linguistic gestures: Mouthing in Italian Sign Language. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 231–246). Hamburg: Signum-Verlag.

Anderson, D. E., & Reilly, J. S. (1998). PAH! The Acquisition of Adverbials in ASL. *Sign Language & Linguistics*, *1*(2), 117–142. doi:10.1075/sll.1.2.03and

Baker-Schenk, C. L. (1983). *A Microanalysis of the Nonmanual Components of Questions in American Sign Language* (Ph.D dissertation). University of California, Berkeley.

Bank, R., Crasborn, O., & van Hout, R. (2011). Variation in mouth actions with manual signs in Sign Language of the Netherlands (NGT). *Sign Language & Linguistics*, *14*(2), 248–270. doi:10.1075/sll.14.2.02ban

Bank, R., Crasborn, O., & van Hout, R. (2013). Alignment of two languages: The spreading of mouthings in Sign Language of the Netherlands. *International Journal of Bilingualism*. doi:10.1177/1367006913484991

Bergman, B., & Wallin, L. (2001). A preliminary analysis of visual mouth segments in Swedish Sign Language. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 51–68). Hamburg: Signum-Verlag.

Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia*, *44*(2), 178–190. doi:10.1016/j.neuropsychologia.2005.05.007

Boyes Braem, P. (2001). Functions of the mouthings in the signing of Deaf early and late learners of Swiss German Sign Language. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 99–132). Hamburg: Signum-Verlag.

Boyes Braem, P., & Sutton-Spence, R. (Eds.). (2001). *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages*. Hamburg: Signum.

Brentari, D. (Ed.). (2001). *Foreign Vocabulary in Sign Languages: A Cross-Linguistic Investigation of Word Formation*. Mahwah N.J.: Lawrence Erlbaum Associates.

Brentari, D., & Padden, C. (2001). A lexicon of multiple origins: Native and foreign vocabulary in American Sign Language. In D. Brentari (Ed.), *Foreign vocabulary* *in sign languages: a cross-linguistic investigation of word formation* (pp. 87–119). Mahwah N.J.: Lawrence Erlbaum Associates.

Brunner, J., Ghosh, S., Hoole, P., Matthies, M., Tiede, M., & Perkell, J. S. (2011). The influence of auditory acuity on acoustic variability and the use of motor equivalence during adaptation to a perturbation. *Journal of Speech, Language, and Hearing Research*, *54*, 727–739.

Brysbaert, M., & New, B. (2009a). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. doi:10.3758/BRM.41.4.977

Brysbaert, M., & New, B. (2009b). *SubtlexUS: Word frequency in American English*. Retrieved from http://subtlexus.lexique.org/

Capek, C. M., Waters, D., Woll, B., MacSweeney, M., Brammer, M. J., McGuire, P. K., … Campbell, R. (2008). Hand and mouth: cortical correlates of lexical processing in British Sign Language and speechreading English. *Journal of Cognitive Neuroscience*, *20*(7), 1220–1234. doi:10.1162/jocn.2008.20084

Chen, T., & Rao, R. R. (1998). Audio-visual integration in multimodal communication. *Proceedings of the IEEE*, *86*(5), 837–852.

Coulter, G. (1982). On the nature of ASL as a monosyllabic language. Presented at the 56th Annual Meeting of the Linguistic Society of America, San Diego, CA.

Crasborn, O., & Sloetjes, H. (2008). Enhanced ELAN functionality for sign language corpora. In *Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora* (pp. 39–43).

Crasborn, O., van der Kooij, E., Waters, D., Woll, B., & Mesch, J. (2008). Frequency distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Language & Linguistics*, *11*(1), 45–67.

Davis, J. (1989). Distinguishing language contact phenomena in ASL interpretation. In C. Lucas (Ed.), *The Sociolinguistics of the Deaf Community* (pp. 85–102). New York: Academic Press.

de Vos, C., & Zeshan, U. (2012). Introduction: Demographic, sociocultural, and linguistic variation across rural signing communities. In U. Zeshan & C. de Vos (Eds.), *Sign Languages in Village Communities: Anthropological and Linguistic Insights* (pp. 2–23). Boston/Berlin: Walter de Gruyter.

Dromey, C., & Bates, E. (2005). Speech Interactions With Linguistic, Cognitive, and Visuomotor Tasks. *Journal of Speech, Language, and Hearing Research*, *48*(2), 295–305. doi:10.1044/1092-4388(2005/020)

Dromey, C., & Benson, A. (2003). Effects of concurrent motor, linguistic, or cognitive tasks on speech motor performance. *Journal of Speech, Language, and Hearing Research*, *46*(5), 1234–1246.

Ebbinghaus, H., & Hessmann, J. (2001). Sign language as multidimensional communication: Why manual signs, mouthings, and mouth gestures are three different things. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 133–152). Hamburg: Signum-Verlag.

Elliott, E. A., Braun, M., Kuhlmann, M., & Jacobs, A. M. (2012). A dual-route cascaded model of reading by deaf adults: evidence for grapheme to viseme conversion. *Journal of Deaf Studies and Deaf Education*, *17*(2), 227–243. doi:10.1093/deafed/enr047

Emmorey, K., Borinstein, H. B., & Thompson, R. (2005). Bimodal bilingualism: Code-blending between spoken English and American Sign Language. In J. Cohen, K. T. McAlister, K. Rolstad, & J. MacSwan (Eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism* (pp. 663–673). Somerville, MA: Cascadilla Press.

Emmorey, K., Petrich, J. A. F., & Gollan, T. H. (2012). Bilingual processing of ASL–English code-blends: The consequences of accessing two lexical representations simultaneously. *Journal of Memory and Language*, *67*(1), 199–210. doi:10.1016/j.jml.2012.04.005

Emmorey, K., Petrich, J., & Gollan, T. (2010). Simultaneous production of ASL and English costs the speaker but benefits the listener. Presented at the 10th Theoretical Issues in Sign Language Research conference, West Lafayette, Indiana.

Fisher, C. G. (1968). Confusions among visually perceived consonants. *Journal of Speech and Hearing Research*, *11*(4), 796–804.

Fontana, S. (2008). Mouth actions as gesture in sign language. *Gesture*, *8*(1), 104–123. doi:10.1075/gest.8.1.08fon

Gentilucci, M. (2003). Grasp observation influences speech production. *European Journal of Neuroscience*, *17*(1), 179–184. doi:10.1046/j.1460-9568.2003.02438.x

Gentilucci, M., Benuzzi, F., Gangitano, M., & Grimaldi, S. (2001). Grasp with hand and mouth: A kinematic study on healthy subjects. *Journal of Neurophysiology*, *86*(4), 1685–1699.

Gentilucci, M., Bernardis, P., Crisi, G., & Volta, R. D. (2006). Repetitive transcranial magnetic stimulation of Broca's area affects verbal responses to gesture observation. *Journal of Cognitive Neuroscience*, *18*(7), 1059–1074.

Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: a gradual transition. *Neuroscience and Biobehavioral Reviews*, *30*(7), 949–60. doi:10.1016/j.neubiorev.2006.02.004

Gentilucci, M., Dalla Volta, R., & Gianelli, C. (2008). When the hands speak. *Journal of Physiology-Paris*, *102*(1-3), 21–30. doi:10.1016/j.jphysparis.2008.03.002

Gentilucci, M., Santunione, P., Roy, A. C., & Stefanini, S. (2004). Execution and observation of bringing a fruit to the mouth affect syllable pronunciation. *European Journal of Neuroscience*, *19*(1), 190–202.

Gentilucci, M., Stefanini, S., Roy, A. C., & Santunione, P. (2004). Action observation and speech production: study on children and adults. *Neuropsychologia*, *42*(11), 1554–67. doi:10.1016/j.neuropsychologia.2004.03.002

Gentilucci, M., & Volta, R. D. (2007). The motor system and the relationships between speech and gesture. *Gesture*, *7*(2), 159–177.

Gilhooly, K. J., & Logie, R. H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behaviour Research Methods and Instrumentation*, *12*, 395–427.

Goldsmith, J. A. (1990). *Autosegmental and Metrical Phonology*. Cambridge, MA: Basil Blackwell.

Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The Physiologic Development of Speech Motor Control: Lip and Jaw Coordination. *Journal of Speech, Language, and Hearing Research*, *43*, 239–255.

Green, J. R., Wang, J., & Wilson, D. L. (2013). SMASH: A tool for articulatory data processing and analysis. *Interspeech 2013*, 1331–1335.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, *102*(3), 594–621.

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, *25*(5), 408–422. doi:10.1016/ j.jneuroling.2009.08.006

Hohenberger, A., & Happ, D. (2001). The linguistic primacy of signs and mouth gestures over mouthings: Evidence from language production in German Sign Language. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 153–190). Hamburg: Signum-Verlag.

Huntington, D. A., Harris, K. S., & Sholes, G. N. (1968). An electromyographic study of consonant articulation in hearing-impaired and normal speakers. *Speech and Hearing Research*, *11*(1), 147.

Kegl, J., Senghas, A., Coppola, M., & DeGraff, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In *Language Creation and Language Change: Creolization, Diachrony, and Development* (pp. 179–237). Cambridge, MA: MIT Press.

Keller, J. (2001). Multimodal representations and the linguistic status of mouthings in German Sign Language. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 191–230). Hamburg: Signum-Verlag.

Kendon, A. (1988). How gestures can become like words. In F. Potyatos (Ed.), *Cross-Cultural Perspectives in Nonverbal Communication* (pp. 131–141). Lewiston, NY: Hogrefe.

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and Sign Language in Human–Computer Interaction* (Vol. 1371, pp. 23–35). Berlin, Germany: Springer Berlin Heidelberg.

Kleinow, J., & Smith, A. (2006). Potential interactions among linguistic, autonomic, and motor factors in speech. *Developmental Psychobiology*, *48*(4), 275–287. doi:10.1002/dev.20141

Koller, O., Ney, H., & Bowden, R. (2014). Weakly supervised automatic transcription of mouthings for gloss-based sign language corpora.

Lane, H., & Perkell, J. S. (2005). Control of voice-onset time in the absence of hearing: a review. *Speech, Language, and Hearing Research*, *48*(6), 1334–43. doi:10.1044/ 1092-4388(2005/093)

Leszczynski, M., Skarbek, W., & Badura, S. (2005). Fast viseme recognition for talking head application. In A. Campilho & M. Kamel (Eds.), *Image Analysis and Recognition: Second International Conference, ICIAR 2005, Toronto, Canada, September 28-30, 2005, Proceedings* (pp. 516–523). Berlin, Germany: Springer-Verlag. Retrieved from http://dx.doi.org/10.1007/11559573

Liddell, S. (1980). *American Sign Language Syntax*. The Hague: Mouton.

Lucas, C. (1992). *Language contact in the American deaf community*. San Diego: Academic Press.

Lucas, C., & Valli, C. (1989). Language contact in the American Deaf community. In C. Lucas (Ed.), *The Sociolinguistics of the Deaf Community* (pp. 11–40). New York: Academic Press.

Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: Study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *7*(1), 41–74.

Maner, K. J., Smith, A., & Grayson, L. (2000). Influences of utterance length and complexity on speech motor performance in children and adults. *Journal of Speech, Language, and Hearing Research: JSLHR*, *43*(2), 560–573.

Marschark, M., LePoutre, D., & Bement, L. (1998). Mouth movement and signed communication. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by Eye II: Advances in teh Psychology of Speechreading and Auditory-visual Speech* (pp. 245–266). Hove, UK: Psychology Press.

McGarr, N. S., & Gelfer, C. E. (1983). Simultaneous measurements of vowels produced by a hearing-impaired speaker. *Language and Speech*, *26*, 233–46.

McGarr, N. S., & Harris, K. S. (1983). Articulatory control in a deaf speaker. *Speech of the Hearing-Impaired (I. Hochberg, H. Levitt, and M. J. Osberger (Eds.)*, 75–95.

McGarr, N. S., & Osberger, N. J. (1978). Pitch deviancy and intelligibility of deaf speech. *Communication Disorders*, *11*(2-3), 237–47.

Mohr, S. (2012). The visual-gestural modality and beyond: Mouthings as a language contact phenomenon in Irish Sign Language. *Sign Language & Linguistics*, *15*(2), 185–211. doi:10.1075/sll.15.2.01moh

Nadolske, M. A., & Rosenstock, R. (2007). Occurrence of mouthings in American Sign Language: A preliminary study. In P. M. Perniss, R. Pfau, & M. Steinbach (Eds.), *Visible Variation* (Vol. 188, pp. 35–62). Berlin, New York: Mouton de Gruyter.

Retrieved from http://www.reference-global.com/doi/abs/10.1515/9783110198850.35

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (1998). *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure*. Cambridge, MA: MIT Press.

Osberger, M. J., & McGarr, N. S. (1982). Speech production characteristics of the hearing impaired. In N. J. Lass (Ed.), *Speech and Language: Advances in Basic Research and Practice 8* (pp. 221–283). New York: Academic Press. Retrieved from http://www.haskins.yale.edu/SR/SR069/sr069_13.pdf

Padden, C. (1988). *Interaction of Morphology and Syntax in American Sign Language*. New York: Garland.

Padden, C. (1998). The ASL lexicon. *Sign Language & Linguistics*, *1*(1), 39–60. doi:10.1075/sll.1.1.04pad

Rainò, P. (2001). Mouthings and mouth gestures in Finnish Sign Language. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 41–49). Hamburg: Signum-Verlag.

Reilly, J. S., McIntire, M. L., & Bellugi, U. (1990). Faces: The relationship between language and affect. In V. Volterra & C. J. Erting (Eds.), *From Gesture to Language in Hearing and Deaf Children* (pp. 128–141). New York: Springer-Verlag.

Sadagopan, N., & Smith, A. (2008). Developmental changes in the effects of utterance length and complexity on speech movement variability. *Journal of Speech, Language, and Hearing Research*, *51*(5), 1138–1151. doi:10.1044/1092-4388(2008/06-0222)

Sandler, W. (1989). *Phonological Representation of the Sign*. Dordrecht, Netherlands: Foris.

Sandler, W. (1999). The Medium and the message: Prosodic interpretation of linguistic content in Israeli Sign Language. *Sign Language & Linguistics*, *2*(2), 187–215. doi:10.1075/sll.2.2.04san

Sandler, W. (2009). Symbiotic symbolization by hand and mouth in sign language. *Semiotica*, *2009*(174), 241–275. doi:10.1515/semi.2009.035

Schein, J. D., & Delk, M. T. (1974). *The deaf population of the United States*. Silver Spring, MD: National Association of the Deaf.

Schermer, T. (2001). The role of mouthings in Sign Language of the Netherlands: Some implications for the production of sign language dictionaries. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 273–284). Hamburg: Signum-Verlag.

Schmidt, C., Koller, O., Ney, H., Hoyoux, T., & Piater, J. (2013). Using viseme recognition to improve a sign language translation system. In *International Workshop on Spoken Language Translation* (pp. 197–203). Heidelberg, Germany.

Schuit, J. (2012). Signing in the Arctic: External influences on Inuit Sign Language. In U. Zeshan & C. de Vos (Eds.), *Sign Languages in Village Communities: Anthropological and Linguistic Insights* (pp. 181–208). Boston/Berlin: Walter de Gruyter.

Smith, A., McFarland, D. H., & Weber, C. M. (1986). Interactions between speech and finger movements: An exploration of the dynamic pattern perspective. *Journal of Speech, Language, and Hearing Research*, *29*, 471–480.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, *17*, 3–45.

Sutton-Spence, R., & Day, L. (2001). Mouthings and mouth gestures in British Sign Language. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 69–86). Hamburg: Signum-Verlag.

Tapio, E., & Takkinen, R. (2012). When one of your languages is not recognized as a language at all. In J. Blommaert, S. Leppänen, P. Pahta, & T. Räisänen (Eds.), *Dangerous Multilingualism: Northern Perspectives on Order, Purity and Normality* (pp. 284–308). Basingstoke, UK: Palgrave Macmillan.

The Language Archive. (2010). *ELAN (EUDICO Linguistic Annotator)*. Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics. Retrieved from http://tla.mpi.nl/tools/tla-tools/elan/

Tkachman, O. (2012). *The Basic Noun-Verb Distinction in Two Young Sign Languages* (Unpublished MA thesis). University of Haifa, Israel.

Toglia, M. P., & Battig, W. R. (1978). *Handbook of Semantic Word Norms*. New York: Erlbaum.

Tourville, J., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language & Cognitive Processes*, *26*(7), 952–981. doi:10.1080/01690960903498424

Vinson, D. P., Thompson, R. L., Skinner, R., Fox, N., & Vigliocco, G. (2010). The hands and mouth do not always slip together in British Sign Language: Dissociating articulatory channels in the lexicon. *Psychological Science*, *21*(8), 1158–1167. doi:10.1177/0956797610377340

Vogt-Svendsen, M. (2001). A comparison of mouth gestures and mouthings in Norwegian Sign Language. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 9–40). Hamburg: Signum-Verlag.

Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, *57*, 209–232. doi:10.1016/ j.specom.2013.09.008

Weismer, G. (1984). Control of the voicing distinction for intervocalic stops and fricatives: Some data and theoretical considerations. *Journal of Phonetics*, *8*, 427–438.

Wilson, M. D. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary, version 2. *Behavioural Research Methods, Instruments and Computers*, *20*(1), 6–11.

Woll, B. (2001). The sign that dares to speak its name: Echo phonology in British Sign Language. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 87–98). Hamburg: Signum-Verlag.

Woll, B. (2009). Do mouths sign? Do hands speak?: Echo phonology as a window on language genesis. In R. Botha & H. de Swart (Eds.), *Language Evolution: The View from Restricted Linguistic Systems* (pp. 203–224). Utrecht: LOT.

Woll, B., & Sieratzki, J. S. (1998). Echo phonology: Signs of a link between gesture and speech. *Behavioral and Brain Sciences*, *21*(4), 531–532.

Zeshan, U. (2000). *Sign Language in Indo-Pakistan: A Description of a Signed Language*. Philadelphia: John Benjamins.

Zeshan, U. (2001). Mouthing in Indopakistani Sign Language: Regularities and variation. In P. Boyes Braem & R. Sutton-Spence (Eds.), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages* (pp. 247–272). Hamburg: Signum-Verlag.

Zimmer, J. (1989). Toward a description of register variation in American Sign Language. In C. Lucas (Ed.), *The Sociolinguistics of the Deaf Community* (pp. 253–272). New York: Academic Press.