REVISITING FIRST-PERSON SINGULAR PRONOUNS AND SOCIAL STATUS USING ASL DATA: A CROSS-LINGUISTIC REPLICATION ATTEMPT

KIVA MARJORIE BENNETT

A DISSERTATION

Submitted to the Department of Linguistics
and the Graduate School of Gallaudet University
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

November 2024

This is to certify that the dissertation entitled:

REVISITING FIRST-PERSON SINGULAR PRONOUNS
AND SOCIAL STATUS USING ASL DATA:
A CROSS-LINGUISTIC REPLICATION ATTEMPT

prepared by

Kiva Marjorie Bennett

is approved in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Linguistics.

Approved by:

Julie Hochgesang, Ph.D. (Chair)

Regina Nuzzo, Ph.D.

Deanna Gagne, Ph.D.

Miako Villanueva, PhD

Department of Linguistics
Gallaudet University
November 2024

Acknowledgements

Dissertating is a deeply solitary journey. Fortunately, mine has not been a lonely one. I mention individuals who were involved with my work in Chapter 1, but throughout this process I have been surrounded by people who have helped shape my research, and who have also shaped me.

First, I want to acknowledge Dr. Julie Hochgesang, the chair of my dissertation committee. Julie somehow perfected the balance of providing the right guidance and the right amount of space for me to figure out what to do with it. As an MA student in the Linguistics department, I presented Julie with my list of ideas for a dissertation topic and asked which she would rather work on. She laughed and told me that if I was going to get a PhD, I should choose a topic that was interesting to me, not her. I clarified that they were all interesting to me, so my most pressing task was to convince her to be my advisor. I am forever grateful that she agreed to embark on this journey with me. She has shaped my understanding of what it means to be a linguist, a responsible researcher, a teacher, and a friend.

I am also grateful to Dr. Regina Nuzzo, who served on my committee and provided far more support than was reasonable to ask for. When I enrolled for my first statistics course, taught by Regina, I only did so because I knew I was woefully underprepared to do anything quantitative with my research. I did not even like numbers at the time. Fast-forward several years, and my degree will say “Linguistics,” but Regina tutored me to what I think of as an unofficial minor in Applied Practical Statistics and Science Communication. My dissertation would not be what it is without her guidance and her expertise. She has shaped how I think about research (quantitative or otherwise), how I ask questions, and how I write about all of it.

I am also appreciative of the support provided by the other members of my committee, Dr. Deanna Gagne and Dr. Miako Villanueva. Their experience with research and linguistic thinking helped to shape this project. Then their suggestions helped reshape it when everything got turned upside down. Their perspectives were particularly helpful in striking a balance between the language and the numbers.

I would also like to thank Dr. James Pennebaker for being willing to engage via email with me as an unknown graduate student, and then for being willing to share his data set with me. This has greatly enhanced my ability to contribute to the larger conversation about social status and FPS pronoun use.

Thank you also to my friends and colleagues who have supported me along the way, including Shane Blau, Ardavan Guity, Heather Hamilton, and Wink Smith for years of group work sessions. And to all of those who have indulged my linguistic thought experiments and my stream-of-consciousness rants about how "status is not even real!" Those conversations have helped shape my work. Shane also helped me with data collection, various aspects of annotation, and wrestling with assorted ideas. Ardavan was also the sign model for the demonstration forms of the seven ASL FPS pronouns. Additionally, Wink and Shane interpreted my dissertation defense into spoken English—dream team.

I am especially grateful to Wink, who has been a true partner in every sense. We had shaped a life together that allowed us to begin PhDs simultaneously, and we have both continued to dissertate for much longer than either of us expected. I am thankful to have done this together, and I look forward to what comes next.

Prior to all of this, I was shaped by my parents, Gail Bennett and Mike Bennett, and my sisters, Dayna Bennett and Myka Turley. It is because of them that I am who I am, curiosities and

all. They have continued to support me through this journey, providing me with love and encouragement, as well as valuable feedback on my ideas and on drafts of my presentations. My continued personal development is tied to them. Additionally, Dayna took on the task of providing alt text for the numerous figures in this dissertation, and Myka cross-checked and formatted all of my references. In other words, they helped me preserve my sanity near the end of this process.

Finally, I owe a deep debt of gratitude to the 24 participants who allowed me to record and analyze their interactions. This work exists because of you. Thank you for trusting me.

Abstract

Research over the past two decades has reported a robust relationship between relative social status and first-person singular (FPS) pronoun use in English. For my dissertation study, I wanted to test the replicability of those findings using American Sign Language (ASL) data that I collected for this purpose. In alignment with previous work, I hypothesized an inverse relationship between relative social status and FPS pronoun use in ASL: people with higher relative social status were expected to use proportionally fewer FPS pronouns. I chose an influential paper in the field as a methodological guide for replication and adapted its study design and data analysis approach to fit the context of my study, which included the signed modality and notions of social status specific to American deaf communities. During my investigations, however, I found that the chosen paper appeared to contain several inconsistencies. Moreover, I could not computationally reproduce some results even using the original English data shared by the authors. I refined their methods, applying the refinements to their English data as well as the ASL data I collected, and applied my adapted data analysis methodology to the original English data. My results in ASL hint at the possibility of an inverse relationship between relative social status and FPS pronoun use, but with insufficient evidence to allow me to conclude with confidence that such a relationship exists. FPS pronoun use in ASL may indicate something about attention, but it does not appear to indicate relative social status. Additionally, after a closer examination of the prior work done in English, I am skeptical that such an inverse relationship truly exists in either language. My work for this dissertation also highlights the critical importance of research reproducibility. While achieving reproducibility requires substantial effort, its role in advancing reliable scientific knowledge is indispensable.

Table of Contents

List of Figures

Chapter I: Introduction

In this dissertation, I investigate the claim that the use of first-person singular (FPS) pronouns tells us something about social status. Research, I am told, rarely happens in a linear fashion. This project certainly has not been linear. I will begin describing my work here by providing an overview of my motivation and personal interest in the topic of social status and FPS pronoun use, along with the general trajectory of how this dissertation came to be.

My interest in the topic of social status and FPS pronoun use began with a general-audience book called *The Secret Life of Pronouns* (Pennebaker, 2011). The author's overarching claim is that function words like pronouns reveal things about people's psychological states. The book describes English usage almost exclusively. As a linguist specializing in American Sign Language (ASL), I wondered if and how the pronoun patterns Pennebaker described would show up in ASL. As an ASL-English interpreter, I wondered how knowledge of these patterns could be applied to the field of interpreting. It seemed that patterns had been identified for communication in English, but if I wanted to know what patterns existed in ASL, I would need to conduct that research myself.

In *The Secret Life of Pronouns*, Pennebaker mentions some of the studies that were later described by Kacewicz et al. (2013), a paper second-authored by Pennebaker, titled "Pronoun Use Reflects Standings in Social Hierarchies." Various elements of Kacewicz et al. (2013) are described throughout this dissertation. I used one of the studies reported in that paper in designing my own study, which would become the first investigation into a potential relationship between social status and first-person singular (FPS) pronoun use in ASL. The data I collected is the ASL data described in this dissertation. My preliminary data analysis ("Round 1," described in Chapter 3) showed similar results as those reported in the literature: a moderately strong inverse correlation between social status and FPS pronoun use.[1] I then designed the next phase of my research to be an experimental study. My intention was to temporarily manipulate perceived social status in order to see what effect this would have on FPS pronoun use. Before I began on that next phase, though, I completed the annotation and analysis for the remainder of my original data set. The complete results, even at first glance, were quite different from the preliminary ones. They were different enough that I began to question whether the data truly supported an inverse relationship. Seeing how nearly I overlooked this extra complexity, I also began to look more closely at the literature that made claims about an inverse relationship in English.

This dissertation is the result of my investigations into the potential relationship between social status and FPS pronoun use. The goal of this study was to begin developing an understanding of the potential relationship between these two variables in ASL. My primary analysis uses the ASL data set I collected, and my secondary analysis includes one of the English data sets used by Kacewicz et al. (2013). Working from the reported results and conclusions in prior work done on this potential relationship in English, I expected similar results for the ASL data that I collected. Specifically, I hypothesized an inverse relationship between relative social status and FPS pronoun use in ASL. That is, I expected that participants of relatively lower social status would use relatively higher rates of FPS pronouns. This was not borne out in my data set. Additionally, after a closer examination of the prior work done in English, I am skeptical that such an inverse relationship truly exists in either language.

Before providing an overview of what to expect in each chapter, I would like to address a few topics that I hope will provide some context, setting the stage for the rest of this dissertation:

[1] An inverse relationship describes the situation where an increase in one variable tends to be associated with a decrease in the other variable. This can often be quantified as a negative correlation with a particular strength. Here, the assumption was that when social status was lower, FPS pronoun use would be higher.

social status as a sensitive topic, positionality and its importance in this work, whether to capitalize “d” in “deaf”, and my approach to data sharing and access.

Social Status is Sensitive

Before getting any further into this writing, I would like to acknowledge that social status can be a thorny subject. Social status hierarchies allow elitism to exist. Differences in status can be used to pass judgment, to hold others in contempt, and to treat people as less-than. Because of this, openly discussing social status can feel uncomfortable. However, these negative framings are not the only ways to understand social status.

Social status is a multifaceted concept, the intersection of multiple factors scaled to reflect the context at hand. And it can be treated as a neutral topic, free from any negativity, and perceived as a theoretical object for study. We are all aware of differences in social status, though this awareness often does not register in our conscious thoughts. This awareness is enacted in what we expect from others, and in how we address and otherwise interact with others. Differences in social status can be acknowledged without passing judgment. We often acknowledge that someone holds relatively lower social status in a particular context without looking down on them. For example, a tenured employee does not automatically feel disdain for a new hire simply because of their workplace status difference. Tenured employees who do treat new hires as less-than are often considered rude and unkind. To be of “higher status” than someone does not equate to being “better than” them. I readily defer to those who have more experience than I do in a given endeavor, without feeling that they are holding me in contempt.

While all of this may be true, it is also true that some people use status differences to put others down. This explains why openly discussing social status is considered somewhat taboo in many cultures, including in the United States, where my study took place. I hope that my approach throughout this dissertation will assure readers of my treatment of social status as a sociological reality, free from any requisite presumptions.

Positionality

When I read work done in signed languages, I am always curious about the authors and their connections to the language communities they are writing about. This is not just blanket curiosity. I have read a good deal of signed language linguistics literature written by people who I later found out were not fluent in the language they were making claims about. Using a language informant/consultant during annotation is not the same as having fluent users of the language as an integral part of the research team. The “research team” for this dissertation is mostly just me, so for readers who share my curiosity, I will provide some information about my positionality with regard to this work. Because I leaned on others for various types of research support (e.g., two research assistants who did some annotating, a few people who did phonological checks), I will provide brief positionality statements for those who expressed a willingness to have me share them. For more on the importance of transparency regarding positionality, specifically in signed language and deaf[2] research, see Hochgesang (2022).

I am a hearing, cisgender, white woman, raised in Florida. I learned some ASL vocabulary as a child through Girl Scouts and from my mother, who had gone back to college and was taking ASL as a foreign language. During my senior year of high school, I transferred to a new school, and found that there was a mainstream program there for deaf and hard-of-hearing students. I had learned just enough ASL vocabulary by then to periodically embarrass myself while interacting with my deaf classmates during gym. I did not begin to learn the language in earnest until my early 20s, when I began taking ASL classes in college. While pursuing a degree

[2] See section below on capitalization conventions for the word “deaf.”

in English Education at the University of Central Florida (taking my time, I might add), I worked my way through all of the ASL and interpreting classes available between the university and the nearby interpreter training program at Valencia College. By then, I knew I wanted to be an interpreter. I began interpreting in 2012, and have held the National Interpreter Certification (NIC) from the Registry of Interpreters for the Deaf (RID) since 2014. My training in linguistics began at Gallaudet University when I entered the MA program in Linguistics in 2015, where I have been a student and periodic adjunct instructor since. I took the American Sign Language Proficiency Interview (ASLPI) in 2016, and received a score of 4. My partner of over 10 years, Wink Smith, is a child of deaf adults (coda), and I am grateful to him and his parents for lovingly accepting me and my L2 signing into their family. Since beginning my formal training in ASL, I have maintained active relationships within deaf and interpreting communities, both personally and professionally. At the time of this writing, I plan to publish the video of my dissertation defense as a supplemental file to this PDF (as well as including it in the publicly available Figshare folder described below[3]). This video will provide an example of my ASL use in a relatively nerve-wracking academic setting.

Below I provide brief positionality statements for some of the other individuals who provided research support on this project, as written by the individuals themselves.

Julie A. Hochgesang assisted with phonological checks and reliability checks, in addition to serving as the chair of my dissertation committee:

[3] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

"Deaf North American woman, grew up using ASL and English as primary languages, mainstreamed with hearing family who could sign (with varying degrees of fluency), currently professor of linguistics at Gallaudet University."

Regina Nuzzo was my statistics advisor, in addition to serving on my dissertation committee: *"North American cis woman; grew up hard of hearing using listening, lip-reading, and spoken English; now an oral deaf cochlear implant user; first learned ASL in graduate school."*

Faith M. Sanders assisted with annotation on one of the chats:
"Hearing Black woman, grew up using English as a primary language, began learning and using ASL in middle school, currently an ASL interpreter in the DMV area."

Wink Smith assisted with phonological checks:
"A hearing, white, cis man who grew up using ASL and English as his primary languages in North America. He was homeschooled by his deaf parents, who are fluent signers. Wink is currently an instructor of interpretation and translation at Gallaudet University."

Shane Blau assisted with phonological checks, a status indicators annotation pass, and filming logistics:
"A hearing, white, transmasculine individual who grew up as a monolingual English user and first learned ASL as a teenager. He is currently a linguistics instructor at Gallaudet University and Ohlone College, and a researcher with the Family ASL Project."

Deanna L. Gagne assisted with my exploratory FPS rate stabilization analysis, in addition to serving on my dissertation committee:

"North American woman of mixed Latine and Eastern European descent, hearing child of deaf parents ("Coda"), grew up in a trilingual household with ASL, English, and Spanish as primary languages, parent of deaf and coda kids, currently professor of linguistics at Gallaudet University."

Capitalization Conventions for "Deaf"

There are some who distinguish between "deaf" as referring to audiological deafness and "Deaf" as referring to those who have some degree of audiological deafness and also identify with one or more cultures of deaf people who value a signed language as one of their primary/preferred languages. Some people find this distinction useful for various reasons. A non-signing hearing person who becomes deaf in old age will have some shared experiences with someone who is born audiologically deaf to audiologically deaf parents who sign. Yet, referring to one's grandmother as "Deaf" rather than "deaf" may provide a shorthand within signing communities, avoiding the need for a lengthy mid-story sidebar. While this may be true, some find the Deaf/deaf distinction unnecessarily divisive (Kusters et al., 2017). Writing conventions have changed over time, with some writing "d/Deaf" or "D/deaf," and others only using either the lowercase or capitalized version, abandoning the other entirely. Outside of the forced decision in writing, this distinction is also made in spoken and signed languages as well (e.g., "big-D deaf" and "little-D deaf" in spoken English). At the time of this writing, these conversations are ongoing. Indeed, these conventions may remain in flux indefinitely. They are also beyond the scope of this dissertation. For more on the discussion surrounding these capitalization conventions, see Fisher et al. (2021). Out of respect for the various opinions regarding this issue, as well as for the sake of simplicity, I will only capitalize "Deaf" when quoting sources where it was originally capitalized by the writer. Otherwise, I will default to lowercase "deaf."

Data Sharing and Access

This work was made possible through the use of data: primary data I collected in ASL and primary data collected by others in English, along with various codings, arrangements, and calculations thereof. I aim to provide adequate transparency around this work, up to and including making it reproducible. While ideas and terminology around "research reproducibility" include related terms (e.g., replicability, repeatability, reliability) without consensus (Goodman et al., 2016), I am a proponent of making research as transparent as possible. Ideally, interested parties could dig deeper into shared data to better understand how authors came to their conclusions. Some of those interested parties could also find ways to improve various elements of the original research design in order to move the scientific process forward. I am on board for both.

Best practices for data citation and sharing in the field of linguistics are laid out in *The Austin Principles of Data Citation in Linguistics* (Berez-Kroeker et al., 2018). According to these principles, linguistic data should be "as open as possible, in order to facilitate reproducibility; and as closed as necessary, to honor relevant ethical, legal and speaker community constraints" (p. 2).

In the spirit of "as open as possible," I have created a publicly available folder using Figshare.[4] As mentioned earlier, at the time of this writing, I plan to include with this publication

---

[4] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

a video of my dissertation defense as a supplemental file. The video is likely to be more accessible than this dissertation, if only because watching it will take less time than reading this dissertation. The presentation is in ASL with English voiceover and captions, and it will be available in my Figshare folder. Also, in case the alt text for the figure descriptions gets lost in the formally published version of this dissertation, my Figshare folder will also contain a version with the alt text intact.

In the spirit of keeping the data "as closed as necessary," I have been judicious about which parts of the data set to include in the publicly available Figshare project folder. Given the sensitive nature of social status as a concept, I want to protect the identities of the participants who made this work possible. In addition to obtaining consent for their data to be analyzed by myself and those assisting with my research, I asked participants for permission to share their chat videos only in the form of screenshots and video clips in papers and presentations. The identities of those who did not provide that consent are not discoverable from my work. Even for those who provided consent for their images to be used, I have made their identifying information, including the information I used as their status markers, as opaque as possible. Accordingly, any assumptions (related to status or anything else) made about a participant recognized in a screenshot or video clip would be just that—assumptions.

Overview

This dissertation is divided into five chapters, of which this is the first. In Chapter 2, the literature review, I discuss aspects of both social status and pronouns that are relevant to this work. I also provide an overview of the current state of the literature on social status and FPS pronouns. Chapter 2 contains a fair amount of detail about Kacewicz et al. (2013) because I draw upon their methodology and because it is the most widely cited work on this topic. In Chapter 3, I describe my own methodology. This includes the demographics of the participants whose data comprises the ASL data set, along with descriptions of my procedures laid out with a level of detail that hopefully makes them easily accessible as well as easily reproducible. Chapter 4 contains the results of my analyses. Finally, in Chapter 5, I discuss the implications of these results, along with describing a handful of exploratory analyses.

Chapter II: Review of the Literature

Introduction

As mentioned in Chapter 1, my interest in the topic of FPS pronouns and social status in ASL began with a general-audience book called *The Secret Life of Pronouns* (Pennebaker, 2011). Much of the work that has been published about FPS pronouns and social status builds upon work done by Pennebaker and his students and colleagues, using the LIWC software (described later in this chapter). The earlier work linking pronouns to social status appears to have used various proxies for status, including age, gender, and physical and mental health factors (e.g., Pennebaker et al., 2003; Tausczik & Pennebaker, 2010). Later work attempted a more direct connection between social status and pronouns, the most influential of which has been Kacewicz et al. (2013). This chapter sets the stage for my contributions to this conversation.

Understanding the relationship between two variables first requires an understanding of each of those things independently. I will thus begin this chapter with a broad description of pronouns in general, then narrow the discussion to pronouns in signed languages, which leads to a more detailed look at first-person singular (FPS) pronouns in ASL. That section closes with a brief discussion of FPS construal without the use of pronouns.

I then provide an overview of social status, beginning with a broad description of various concepts that contribute to notions of status, followed by a discussion of social status within deaf communities. The chapter closes with a review of the literature on the relationship between FPS pronouns and social status, including a critical look at Kacewicz et al. (2013).

Pronouns

General.

The prefix ‘pro’ means to take the place of, or to substitute for. In language, pro-forms are schematic substitutions for more specified words and phrases (Evans & Green, 2006). Pro-forms are deictic and schematic, relying on context to provide whatever level of specified meaning is necessary. This context includes shared knowledge, including but not limited to overt antecedents uttered in the current discourse (Langacker, 1991).

Pro-forms include substitutions for verbs and verb phrases (e.g., *do*, as in “He likes cats, and I do too”), locative expressions (e.g., *there*, as in “I like California, but the cost of living there is high”), and entire clauses (e.g., *so*, as in “They think it’s almost done, but I don’t think so”), to name a few (Kullavanijaya, 2004).

The most well-known subgroup of proforms, though, are pronouns, which stand in for nouns and noun phrases. Pronouns constitute a small closed class, meaning new forms are exceedingly rare when compared with more open word classes such as nouns and verbs. Since each of the relatively few pronoun forms is highly schematic, each one can stand in for a wide array of more specified forms (i.e., nouns and noun phrases). This leads to high frequency of use, which often results in phonological reduction, leaving many pronouns with a relatively simple phonological structure (e.g., “I” in English, “yo” in Spanish). High frequency across contexts also makes them interesting subjects for analysis with regard to distribution patterns, including patterns relevant to cross-linguistic comparisons (e.g., Cormier et al., 2010), child acquisition (e.g., Reynolds, 2016), overt use of pronouns in pronoun-optional contexts (e.g., Wulf et al. 2002), and second language acquisition (e.g., Thompson et al., 2009). And, of course, in relation to social status (e.g., Pennebaker’s work, this dissertation).

Pronouns can be marked for grammatical categories such as person, number, gender, and case. The two relevant categories for this work are person (first) and number (singular). As one example, the English word “my” is marked for both: first person and singular number. In contrast, the English personal pronoun “we” is marked for person (first) and number (plural); the English demonstrative pronoun “this” is marked for number (singular), but not person; and the English interrogative pronoun “who” is marked for neither person nor number.

In semantic terms, first-person pronouns are distinct from all non-first-person pronouns, regardless of language. Both first-person and non-first-person pronouns are schematic and rely on context to provide more specified meaning. But that more specified meaning is typically more readily recoverable for first-person pronouns than for non-first-person pronouns. This is especially true in face-to-face conversations. In any language, there is a much higher likelihood that someone would be confused about which third person referent someone meant than which first-person referent (e.g., “she” is more potentially ambiguous than “me”).

Similarly, first-person plural “we” can be safely assumed to include at least the person who uttered it, even if the other members of the “we” are not immediately obvious. Even a “my” written in a message in a bottle evokes a more specified referent (the writer of the message), than a “his” (someone who is either male or serving as a generic stand-in for a person, or perhaps a creature), and a “they” would be even more schematic. “They” is more readily construed (than even “he”) as evoking a referent without a specified gender, and “they” can also be construed as either singular or plural.[5]

[5] See Balhorn (2004) for use of singular *they* in English since the 1400s; see Bodine (1975) for a discussion of how prescriptivism artificially disrupted the use of singular *they*; see Moulton et al. (2020) for experimental evidence of the acceptability of singular *they*, even in contexts where the referent has a known and binary gender.

Second-person pronoun forms are also more varied in their possible referents than first-person pronouns. As one example from English, the form "you" can be used for singular and plural, and can be construed as either the intended receiver, or as an indeterminate person semantically understood as an indeterminate or third-person reference (e.g., "when *you* name a cat, *you* do so knowing that the cat will never respond to their name"). Second person singular rarely requires the utterance of an overt antecedent to be fully specified, instead relying on other cues, such as simultaneously produced eye gaze,[6] or even simply the existence of the participants in the discourse event. The broad availability of possible referents for second and third person pronouns is not limited to English, but is an inherent part of the meaning structure of each type of pro-form. First-person singular pronouns, by their very nature, have but one possible referent—the one who uttered the first-person singular pronoun.[7]

FPS pronouns are used at a higher rate than other pronouns in English, and are among the most frequently used of all English words (see "Cross-Linguistic Frequency" section later in this chapter). There is literature that suggests similar usage frequencies for pronouns in other languages, including Chinese (Huang et al., 2012), Spanish (Ramirez-Esparza et al., 2007), German (Wolf et al., 2008), and Dutch (Zijlstra et al., 2004).[8]

---

[6] Could eye gaze be considered an antecedent? I could probably be convinced.

[7] This is still true in instances of constructed dialogue that include first-person singular pronouns. The one who is constructed as uttering the pronoun would still be the referent, even when the person doing the construction is someone else. See Tannen (1986) for more on constructed dialogue, and Metzger (1995) for constructed dialogue and constructed action in American Sign Language.

[8] Frequency studies tend to be written in the language they are analyzing, so I have had a difficult time finding more examples than these four papers, which are predominantly European. These are each written in the language they are discussing, with only the abstracts available in English, so I can not be certain of the exact nature of the usage rate comparability. The abstracts report using translated versions of the LIWC software, which is described later in this chapter, to calculate word category correlations that are "reliable" (Chinese), "high" (Spanish), "high to very high" (Dutch), and of "high equivalence" (German), when compared to English. A similar comparison paper exists in French (Piolat et al., 2011), which will be mentioned with the other four later in this chapter, but the English abstract does not provide sufficient information to speculate about word

These two facts, the high rate of FPS usage and the unwavering referent, make FPS pronouns quite a remarkable bit of language. The uniqueness of FPS pronouns is not surprising, since we cannot prevent our own immediate experience from framing our take on the world. It only makes sense that our first-person construal[9] of our "I-here-now" experience (Langacker, 2007) would come through in our language use.

In signed languages.

As discussed above, many pronouns have a relatively simple phonological structure, in both signed and spoken languages. Some of these phonological forms are nearly identical across most signed languages. This is especially true when comparing personal pronouns, which most often take the form of pointing with the index finger (see Cormier et al., 2013 and Fenlon et al., 2019 for discussion).[10] Perhaps these overlaps with gestures used by non-signers, along with cross-linguistic similarities, help to explain why there are a number of contested claims around pronouns in signed languages, not the least of which is the very existence of pronouns in signed languages.

Starting from the beginning of ASL research, Stokoe, Casterline, & Croneberg (1965) discuss what they call "indicative signs" as including words that function much like the English words "I" and "me," but they fall short of explicitly identifying any ASL "pronouns." A decade later, Friedman (1975) was the first to explicitly assert that the ASL lexicon contained no signs

---

category frequency comparisons. It stands to reason that the nature of FPS pronouns would make them high-frequency in many languages, though this is difficult to measure without access to frequency studies for more languages.

[9] For more on construal, see the later section called "FPS construal without pronouns."

[10] The prototypical first-person singular pronoun usually directs the tip of the index finger to the signer's chest. One notable exception is the prototypical first-person singular personal pronoun in Japanese Sign Language (JSL): the difference is that the focal site for the JSL word is the nose rather than the chest, which gives this word the same form as the prototypical self-referencing gesture in mainstream Japanese culture.

classifiable as pronouns. This is somewhat in contrast to much of the early work in signed language linguistics, which was largely focused on justifying the linguistic status of signed languages, fitting the tools and trappings of spoken language linguistics to the analysis of signed languages. Much of this early work asserted that signed languages were more like spoken languages than one might assume, which helped to justify calling signed languages "real" languages. Pronouns seem to be one area where early signed language linguists were willing to claim deviations from spoken language, perhaps because of the phonological overlaps mentioned above. Cormier et al. (2013) provide a useful overview of the work during this period.

As the linguistic status of signed languages shifted from questionable to given, the research shifted as well, and signed language studies began to challenge existing assumptions about the nature of human language. Johnston (2013) provides a review of the theoretical debates of this period. One of those debates is around the linguistic status of pronominal pointing in signed languages. Some analyze these forms as true pronouns, citing a lack of evidence supporting any other analysis.

Other proposals have included analyzing pronominal points as demonstratives (e.g., Swedish Sign Language: Ahlgren, 1990; ASL: McBurney, 2002) or as half linguistic and half gestural (e.g., Liddell, 2003; Cormier et al., 2013). The latter proposal presupposes a hard distinction between language and gesture that is contested by Kendon (2008) and Occhino & Wilcox (2017), among others. The debate about the linguistic status of pronominal pointing in signed languages is far from over, as are debates about whether and how signed languages encode person in these forms.[11] See Nilsson (2004) and Fenlon et al. (2019) for further

[11] The debate over whether ASL has distinct first-, second-, and third-person pronouns is outside the scope of this paper, but for more on this, see: Stokoe et al., 1965 (first-, second-, and third-person pronouns); Meier, 1990 (first and non-first); Lillo-Martin & Klima, 1990 (not encoded for person at all); Neidle, et al., 2000 (first and many subdivisions of non-first); Liddell, 2003 (first and non-first, with a slight reframing and elaboration

discussion on these topics, as well as on the types of referents that first-person singular personal pronouns account for.

It is worth pointing out that much of the work published on pronouns in signed languages, including those referenced here, prioritizes discussion of personal pronouns (or person-as-reference points) with animate or human referents, while spending less time and attention on other pronouns (e.g., possessives, reflexives). See Nilsson (2004), Johnston (2013), and Sloan (2013) for discussion.

Theoretical claims like those listed here are often situated to promote or counter a particular theoretical framework, as in the conclusion by Johnston (2013) about Australian Sign Language (Auslan) that “Auslan probably does not have a type of sign that constitutes an identifiable grammatical class of pronoun” (Johnston, 2013, p. 111), which is framed as corroborating evidence in support of the argument made by Evans & Levinson (2009) against the universality of pronouns (e.g., Pinker & Bloom, 1990). While these theoretical matters continue to be debated, there does not seem to be any objection in the literature thus far to acknowledging that people using signed languages incorporate points that can be classified as functioning pronominally, be they called “pronouns,” “pronominal points” as in Johnston (2013), “pronominal signs” à la Cormier et al. (2013), or something else. It is also quite obvious that some of these forms refer to a first-person singular referent, no matter one’s opinion on how the grammatical category of person is encoded in these forms. For these reasons, as well as for the sake of simplicity, I will continue referring here to the form-meaning pairings in question as first-person singular pronouns (FPS). While I am not particularly interested in participating in the

regarding discourse roles); Lillo-Martin, 2018 (updating her stance from 1990, using child acquisition data as evidence for a three-person system); and Cooperrider & Mesh, 2022 (overview of the arguments). For whatever it may be worth, I lean heavily toward a three-person system.

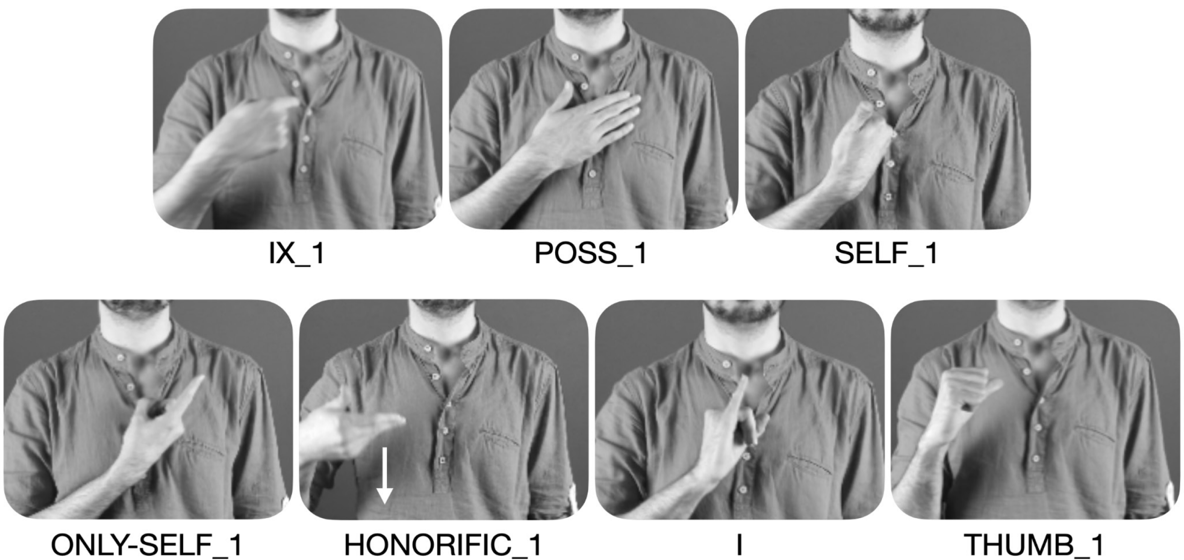
ongoing debates around the existence or classification of pronouns in signed languages, my work may provide evidence for theorists working in these areas to work with.

First-person singular (FPS) pronouns in ASL.

There are seven FPS pronouns in ASL, stills of which can be seen in Figure 2.01, labeled with ID glosses following the SLAASh conventions (Hochgesang, 2020) used in the ASL Signbank (Hochgesang et al., 2024).[12]

Figure 2.01

First-person singular (FPS) pronouns in ASL



[12] Videos of these seven pronouns are available in my Figshare project folder (https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932) and in ASL Signbank. Throughout this dissertation, the first time I mention an ASL word (FPS pronouns and a few others), I provide the Signbank ID gloss and link to the entry. Note that while ASL Signbank is intended to become truly publicly available, there is no set date for this, as of the time of this writing. For now, some entries are only visible to those who have registered and are logged in, though registration is free and open to anyone. The entries for the FPS pronouns are as follows:
IX_1: https://aslsignbank.haskins.yale.edu/dictionary/gloss/1711
POSS_1: https://aslsignbank.haskins.yale.edu/dictionary/gloss/2154
SELF_1: https://aslsignbank.haskins.yale.edu/dictionary/gloss/539
ONLY-SELF_1: https://aslsignbank.haskins.yale.edu/dictionary/gloss/2813
HONORIFIC_1: https://aslsignbank.haskins.yale.edu/dictionary/gloss/1762
I: https://aslsignbank.haskins.yale.edu/dictionary/gloss/1751
THUMB_1: https://aslsignbank.haskins.yale.edu/dictionary/gloss/2775

Phonological variation of FPS in ASL.

Studying language in use calls for a descriptive, rather than prescriptive, approach to variation. Attempts to preserve ASL are often well-intended, but can too easily become pernicious, with the dismissal of language use that falls outside of dictionary-worthy production as somehow erroneous, or worse, as "sloppy" (see Hill, 2012 for discussion of language attitudes). Not only is this kind of language-purification stance oppressive, it fundamentally misconstrues the true nature of language, both the language in question, and language at large. I seek to describe language as it is naturally used, and there is a great deal of phonological variation in natural language use. Thus, a discussion of phonological variation is warranted, specifically, phonological variation of FPS pronouns in ASL.

The hand configuration seen in the demonstration form of IX_1 in Figure 2.01 above, a loose fist with a fully extended index finger, is often referred to as a 1 handshape. This handshape, while present in citation and demonstration forms of IX_1, is seen far less often during natural language use.

In fact, quantitative work has shown that IX_1 exhibits a hand configuration other than that of the citation form more often than not. Bayley et al. (2002) found this happened in ASL at a rate of 81% (1,637 out of 2,019 tokens). Fenlon et al. (2013) found the same pattern in British Sign Language (BSL),[13] at a rate of 88% (324 out of 370 tokens). This means that the overwhelming majority of the IX_1 forms produced during natural language use likely do not adhere to the hand configuration that most people would produce if asked to demonstrate IX_1 in isolation. Pay close attention to utterances in any language/modality, and you will see that the phonetic details of surface forms regularly deviate from those of citation forms (Börstell, 2002).

[13] The citation forms of IX_1 are identical in ASL and BSL.

In addition to variation in hand configuration, there is also a high degree of variability in the location (i.e., placement) of IX_1. Some of this variability can be attributed to assimilation of handshape and/or location. Assimilation is the phonological process by which a word's form becomes more similar to an adjacent word through a change to one or more features. IX_1 easily assimilates to adjacent words due to the central location, short path movement, and relatively simple handshape[14] of the citation form. There are similar types of phonological variation in the other six FPS forms as well. To illustrate this point, Figure 2.02 below shows several instances of IX_1 with handshape and/or location variance that occurred in the data I collected for this study.

Figure 2.02

Form variations of IX_1

[14] See McIntire (1977), Boyes-Braem (1990), and Whitworth (2011) for more on the continuum of markedness of various hand configurations.

Frequency of FPS in ASL.

The FPS pronoun IX_1 is the most frequent of the seven words in use, both anecdotally and in Morford & MacFarlane (2003). In their study on subjective word frequency in ASL, 27 signers produced 4,111 words, and the most frequent FPS pronoun was IX_1. In fact, looking at all 4,111 words, IX_1 came in second most frequent overall, with the only more frequently used word being a non-first-person index (which includes all instances of pointing that are not FPS pronouns). The average use of the non-first-person index was 7.9% (or 7.9 out of every 100 words used), followed by IX_1 at 5.6%. The next most frequent word was another FPS pronoun, POSS_1, at just 1.5%. This aligns with the work mentioned earlier[15] that shows the most frequent FPS pronouns to be the personal ones (e.g., “me” in English), followed by possessive (e.g., “my” in English).

At the other end of the FPS frequency spectrum are THUMB_1 and I, which are less prototypical and also less frequent in use. They are included in this work despite existing prescriptive notions about what is “linguistic” or “pure ASL.” As mentioned earlier in the section on phonological variation, my aim is to describe language as it is actually used in American deaf communities,[16] thus these forms are included here. The lower-frequency FPS pronouns (i.e., the five besides IX_1 and POSS_1) may not contribute heavily to total FPS pronoun use, but all seven are included in my study because they function pronominally and have first-person singular referents. See Chapter 4 for frequencies for each of these words in my study.

---

[15] Chung & Pennebaker (2007) in English, Wolf et al. (2008) in German, Huang et al. (2012) in Chinese, Ramirez-Esparza et al. (2007) in Spanish, and Zijlstra et al. (2004) in Dutch.

[16] For more on the use of the plural “deaf communities,” see Hochgesang et al. (2023).

Cross-linguistic frequency.

The idea to research the relationship between first-person singular pronouns and social status in ASL came from work done in English, and the present work compares results in ASL and English. Given this, a brief comparison of FPS pronoun usage frequency in the two languages is warranted.

Since ASL exhibits more frequent pro-drop than English does, it would be reasonable to wonder if there is a comparatively lower rate of pronoun use in ASL. A lower overall rate would then suggest that patterns related to FPS pronoun usage frequency in ASL might be harder to determine, or even hypothetically less reliable, due to more potential variability with a lower total incidence.

However, a look at the most frequent type of FPS pronouns used in ASL and English (i.e., personal; e.g., IX_1) shows comparable frequencies in the two languages. The rates reported by Morford & MacFarlane (2003) for ASL, discussed above, can be seen in Figure 2.03 below, compared to four different English sources: The British National Corpus (BNC), the Corpus of Contemporary American English (COCA), and two of the studies described in Kacewicz et al. (2013).

The British National Corpus (BNC)[17] is designed to represent a broad range of British English usage, but it is heavily skewed to written English, with 90% of the corpus representing written sources.[18] Written English can be expected to contain fewer instances of pro-drop, such as diary drop (i.e., subject pronoun drop), as well as phonological reduction that renders some spoken FPS pronouns phonologically null. Written English also drives up the use of passive

[17] http://www.natcorp.ox.ac.uk/

[18] http://www.natcorp.ox.ac.uk/corpus/index.xml

Figure 2.03

Comparison of personal FPS pronouns in ASL and spoken English corpora

Data source	Total words	Modality & language	Word(s) searched	Usage rate
Morford & MacFarlane (2003)	4,111	Signed ASL	IX_1	5.6%
BNC *Spoken portion only*	10,376,198	Spoken English	"I" + "me"	3.2%
COCA *Spoken portion only*	127,396,932			2.1%
Kacewicz et al. (2013) *Study 1*	123,496			4.5%
Kacewicz et al. (2013) *Study 3*	91,789			6.2%

voice, depersonalized content/subject matter, and an academic reluctance to use FPS pronouns in so-called formal writing. Consequently, written English can be expected to include lower rates of FPS pronoun use than spoken English. Overall, work on the BNC indicates "I" and "me" as being two of the most frequent words in use (Leech et al., 2014b).

If there were a widely used orthographic system for writing ASL,[19] it would perhaps have some similar effects on written ASL production, resulting in lower FPS usage. As it stands, when it comes to distribution of pronouns, ASL use is more comparable to spoken English than written, as both are ephemeral (i.e., not written). Thus, in order to draw a more appropriate comparison to Morford & MacFarlane (2003), Figure 2.03 includes only the portion of the BNC

[19] There are several proposed orthographies for signed languages, but as of this writing, none are in widespread use, nor do any have a cross-platform font available, limiting their portability. Examples include ASLwrite (http://www.aslwrite.com/about_aslwrite), si5s (https://en.wikipedia.org/wiki/Si5s), and SignWriting (http://www.signwriting.org).

that represents spoken usage, which is from the period of 1991 - 1994.[20] In this subset of over 10,000 words, the personal FPS pronouns "I" and "me" are used at a combined rate of 3.2% (or 3.2 times for every 100 words used).[21]

The Corpus of Contemporary American English (COCA), includes a spoken language subset of over 127 million words, with a balanced representation from 1990 to 2019.[22] The spoken transcripts are of "unscripted conversation from more than 150 different TV and radio programs."[23] This spoken portion of the COCA shows "I" and "me" at a combined rate of 2.1%. Again, this represents a relatively high frequency.

The remaining two English examples in Figure 2.03 are from Kacewicz et al. (2013), which is discussed in greater detail later in this chapter. Of the five studies described in their paper, Studies 1 and 3 were the only two that used spoken language data, as opposed to written. While both of these data sets are considerably smaller than the BNC and COCA, they are also comprised entirely of language that was spontaneously produced. The Study 1 data transcripts are from an experiment in which 164 undergraduates were divided into 41 groups of four to complete a collaborative problem-solving task. Out of 123,496 total words spoken, 4.5% of them were personal FPS pronouns. Study 3 included 100 people paired into 50 dyads who each had a 10-minute get-to-know-you conversation. Out of the 91,789 words spoken, 6.2% of them were personal FPS pronouns.

---

[20] Leech et al. (2014a), Leech et al. (2014b), http://www.natcorp.ox.ac.uk/corpus/index.xml?ID=intro

[21] From the "Rank frequency order: spoken English (not lemmatized)" list at https://ucrel.lancs.ac.uk/bncfreq/flists.html.

[22] Davies (2008), https://www.english-corpora.org/coca/

[23] See https://www.english-corpora.org/coca/help/coca2020_overview.pdf for a breakdown of the included genres.

As Figure 2.03 shows, the frequency rates for personal first-person singular pronouns are reported to be similar in ASL and English, at least in the ephemeral language used in the five included data sets; ASL at a rate of 5.6%, and English with a range from 2.1% - 6.2%. The takeaway here is not necessarily that ASL definitely uses all FPS pronouns at a comparable rate as English—this only compares IX_1 to "I" and "me," the ASL data set is quite small, and issues with word counting still remain in both languages (e.g., how to count depiction in signed languages, the issues with word-counting in any language discussed in Chapter 3 under "Variable: FPS pronoun use"). My point here is simply that looking for similarities between FPS pronoun usage rates in ASL and English begins from a commensurate starting point.

FPS construal without pronouns.

First-person singular construal, the inclusion of oneself in the conceptualization of a scene, can be communicated with or without the use of FPS pronouns. Content can be construed in different ways. A person's mental conceptualization will always include their own construal of the content, as no point of view can be entirely objective. A person's particular construal guides their linguistic decisions (consciously or not), which in turn indicates where their attention is focused in that moment. For more on the relationship between construal and attention in general, see Langacker (2008); for an application of this logic to analyses of ASL, see Villanueva (2010). A situation where someone is stumped by a computer that is not working properly may result in Person A focusing their attention on the computer and saying "something is wrong with this computer." This person construes the situation without a self-focus, thus there is no FPS pronoun. Meanwhile, Person B in the same situation might include themselves in the construal, and say "I can't figure out what's wrong with this computer." Person B has framed their experience with an FPS construal, and so uses an FPS pronoun. FPS pronouns are one way of evoking FPS construal, but they are by no means the only way.

Take the grammatical Spanish sentence *Estoy despierto*, which literally translates in English to “Am awake.” The FPS construal is seen in the verb conjugation of the copula *estar* into its first-person singular form, comparable to English’s “am.” The verb evokes a primary participant, in this case the experiencer of being awake (Langacker, 1987, 2008), without the overt use of a subject. In Spanish, the overt naming of this primary participant is not obligatory. In fact, including the nominal personal FPS pronoun *yo* (comparable to English’s “I”) would be acceptable only with an intended emphasis on the speaker. *Estoy despierto* signaling FPS construal without an overt pronoun is an example of subject dropping, a type of pro-drop common in languages around the world (Duguine, 2017).[24]

Complex verbal morphology, as in the Spanish example above, can create opportunities for dropping pronouns without sacrificing clarity, since verbs can take over the role of providing information about the referent. In addition to verb conjugation, referents can also be accessed in other ways, including through context. In any pro-drop language, pronouns in subject and/or object positions may be optional, emphatic, or even ungrammatical (Sloan, 2013). When pronouns are dropped in ASL, the necessary information is conveyed through context, as well as constructions like depicting verbs, constructed action, and indicating verbs[25] (Liddell, 2003; Wulf et al., 2002; see McKee et al., 2011 for incidence in Auslan and New Zealand Sign Language).

---

[24] Rather than thinking of pro-drop as a language-specific binary parameter, it may be more useful to think of it as a construction-specific feature. See Duguine (2017) for discussion.

[25] Indicating verbs are also sometimes referred to as agreement verbs or directional verbs.

I hope the present work will be extended to pronoun-less FPS construal in the future. For my current study, I have begun with a more controlled (i.e., finite and countable) set of units for analysis: FPS pronouns occurring in ASL use. In Chapter 5, I provide a very brief exploratory analysis extending the present work to pronoun-less FPS construal. Any future analyses of the relationship between social status and the broader concept of first-person singular construal will benefit from the context provided here.

Social Status

Social status in general.

Defining, and in turn measuring, social status is something I have agonized over since I began thinking about this work. Social status is socially constructed and ever-changing during interactions; thus a person's social status in a given moment is not something inherent to them, but a temporary state they are occupying. As a result, we can never capture an absolute status for any person, much less predict it, even within a specific context. To get at an individual's perceptions of status, which are often largely subconscious, we can only observe their behavior and ask them about their experiences—then analyze the constructed responses we receive. The only possible conclusion about the objective measurement of social status is what I already knew before beginning this study: social status is not objectively measurable. Additionally, the concept of social status need not be a distasteful one. Acknowledging social status differences between people is not the same as passing judgment on them. See Chapter 1 for more on this concept.

For the purposes of this dissertation, the terms "status" and "social status" will be used interchangeably, with other types of status (e.g., socioeconomic status) specified as needed. That said, there are several concepts related to social status that tend to be used interchangeably, though there are distinctions to be made between them. These include social class, dominance, leadership, power, and prestige, among others. There are so-called “fixed” status relationships between individuals in certain contexts, such as job interviews, prisons, and schools. However, when people engage in those contexts, they are not actually locked into pre-determined status relationships. And even during interactions between people who are objectively considered “equals,” their status relationships will fluctuate (Cheepen, 1988). Many sociolinguistic theorists echo this sentiment, including Johnstone (2008), who says “power is also negotiable...in this sense, power is not necessarily dominance, but rather more like agency: an individual’s ebbing and flowing ability to shape the activity at hand” (p. 130). And as Tannen (1989) points out, “the symbols that display power (unequal status) and solidarity (human connection) are often the same, so every utterance is potentially ambiguous as to whether it is establishing power or solidarity” (p. 144).

Features of social status can include aspects of group membership, age, socioeconomic status, level of education, leadership, and dominance understood as control or influence over others. Status can also be thought of in terms of prestige, power (realized as the ability to influence others to fulfill one’s own needs), respect and admiration from others, or access to unequally distributed resources (Schooler, 2013). There are also the concepts of charisma, clout, and what might be best described as “being one of the cool kids.” These concepts as a whole are considered by some to be the driving force behind all of human nature (e.g., Nietzsche 1901/1967). For more on the intersection of these concepts, see Cheng et al. (2013).

All of these related concepts include the notion of rank, which is inherently relative. An individual may be a certain age, with a certain level of education and prestige, and may have access to a particular set of resources; but translating any of this to “social status” requires rank, which can only be determined in context.

Research in the US has historically included identifying participant "status" based on measures such as race, level of education, and income, while research in Europe has traditionally collapsed education and income into "class," also including factors such as type of employment, and largely ignoring race and ethnicity. This has changed in more recent decades, with American researchers taking a more European approach to considering the role of social class in their work (Grusky, 2019). I avoid the term "social class," as it presupposes a fixed-category view of social hierarchy, based largely on income, essentially asking "who are you?" In contrast, I am interested in the relationships we build and modify in real time during interactions, essentially asking "what role are you playing?" Still, operational definitions vary widely, reducing generalizability of findings and opportunities for cross-study comparisons.

Another important consideration lies in the distinction between objective and subjective conceptualizations of status. Health outcomes have been linked more strongly to self-reported subjective socioeconomic status (SES) than to objective SES (e.g., Cohen et al, 2008; Adler et al., 2000), suggesting that objective SES is insufficient when it comes to predicting practical impact. In other words, an individual's subjective beliefs about their status may sometimes be a better predictor for certain outcomes than external measures, such as those on a demographic survey.

Subjective social status, while potentially important, can be more difficult to access than more objective measures. The concept of status is often metaphorically conceived vertically (in western cultures: Lakoff & Johnson, 1980; in ASL: Taub, 2001), meaning any discussion of a difference in social status (i.e., relative status) amounts to a claim about how high or low each person falls on a vertical scale. Status in general can hypothetically be thought of as something that each person "has" to varying degrees in different contexts, meaning two people may both have relatively high status in a given context; this is a variable-sum view. Relative status, on the other hand, is zero-sum. Thus, reporting yourself to be of higher status than another person includes an indirect report that the other person is of lower status. In this way, claiming higher status than another person can come across as impolite or arrogant in many cultures, including in American society. Such claims about relative status run counter to the default desire to save face, for oneself and for others, which is an ingrained part of social communication (Brown & Levinson, 1987; see Hoza, 2007 for application to deaf Americans). People are often hesitant to blatantly defy social norms, so directly asking for self-reports of relative social status may yield responses with varying levels of honesty, with reports staying closer to "we are equal" than might otherwise be expected. These ideas, that relative status is zero-sum, and that cultural expectations around politeness can impact self-reports, influence my interpretation of some of the literature discussed below, and guide some of my methodological decisions discussed in Chapter 3.

There are innumerable factors that contribute to notions of social status. Some factors are inherently quantifiable, such as age and income. But most factors, including most of the ones mentioned in this section, are not naturally quantifiable (e.g., power, type of employment, the status relationship between a teacher and a student), so they can only be quantified in the ordinal sense. For example, a teacher is considered to have higher status than their student, but the teacher does not inherently possess a countable number of additional "status points," though such a measurement could be developed to represent relative degrees of status. For defining status in the present work, any of the factors mentioned thus far could be included, but there is a need for another set of social status markers, specific to the communities in question. These are addressed in the following section.

Social status in deaf communities.

The factors mentioned above in reference to status in general (e.g., age, education, income, occupation) also contribute to notions of social status within American deaf communities (Croneberg, 1965). Other contributing factors are tied to community membership, largely centering around a respect for the community's shared experiences and shared language of ASL (Padden, 1989; Kannapell, 1989; Lillo-Martin et al., 2023). While ASL is considered a minority language, there is covert prestige in being a member of the ASL community — covert prestige that is overtly acknowledged (Hill, 2015). Hall (1983) reports that some other factors contributing to a person's status within American deaf communities include "whether or not a person has Deaf parents, whether he was deafened before or after acquiring speech, and whether he is hard-of-hearing or deaf" (p. 83).

It is obvious that audiological deafness is not sufficient to claim membership in a deaf community, let alone to claim status within one. K-12 schooling provides one way in. Deaf and hard-of-hearing American students[26] have access to the same K-12 schooling options that hearing students do (e.g., public, private, homeschool). These options may result in a range of experiences. A deaf student may be immersed in a robust program designed for deaf and hard-of-hearing students, with deaf teachers, formal courses in ASL, and qualified ASL interpreters in the classes taught by hearing teachers. These environments are generally considered to be the exception. On the other end of the spectrum, a student may find themselves to be the only non-hearing student in their school, or even in their school district, with few practical resources for introduction to ASL or deaf communities. Deaf students are also afforded additional options in

[26] In the context of K-12 education, deaf and hard-of-hearing students are often referred to as "DHH students" and programs designed for them in mainstreamed public school settings are often referred to as "DHH programs."

the form of deaf schools, which can be residential (students live in on-campus dorms), day schools (students commute to school daily, much like traditional American public schools), or some combination thereof. Sometimes deaf schools partner with local public schools for students to be able to take classes at both locations.

For those who have not been exposed to signed language, gaining entry to a local deaf community can be challenging. Ladd (2003) shares that in the British deaf community, having attended a deaf school is akin to automatic membership. The author quotes one participant as saying "It feels like a private club membership — 'Been Deaf school? Here's a badge. Been Deaf school? That's it, finish, in'. Very difficult for others to get in; can do but not straightaway" (Ladd, 2003, p. 276). The situation is similar in America. Dorminy (2013) describes the experiences of American deaf college students who grew up in non-signing environments and then decided to integrate into American deaf communities by learning ASL and attending Gallaudet University, which Dorminy identifies as "a predominantly signing Deaf university" (p. 1). As suggested in Ladd (2003), these students were indeed able to integrate, but not without challenges from some existing community members.

For those with community membership, the collectivist nature of deaf communities encourages feelings of solidarity within those communities (Lane, 1992), which can obscure, or even override, the finer details of otherwise hierarchical social relationships (see Friedner & Kusters, 2015 for discussion). This phenomenon is sometimes anecdotally referred to as "DEAF-SAME." For example, Ladd (2003) was unable to elicit commentary from British deaf community members regarding "social class" until reframing his prompt in terms of the degree to which community members were "in with" the community's religious leaders (Ladd, 2003). This reframing resulted in responses that documented "for the first time the role of class differences within Deaf culture" (Ladd, 2003, p. 366), although this was presented with the caveat that it was "too early to conclude whether 'class' is the most appropriate framework for this experience" (Ladd, 2003, p. 366).

Indeed, categorizing this nuanced experience is difficult. Quantifying it is even more treacherous. Some attempts to quantify deaf identity have been undertaken in the form of the Deaf Identity Development Scale (Glickman & Carey, 1993) and the Deaf Acculturation Scale (Maxwell-McCaw & Zea, 2011). Both are modeled after similar scales used with other populations and both are based on subjective responses to various scenarios and preferences. Neither have been picked up for regular use by researchers as a reliable measure. This may be due in part to the scales themselves, but it is also likely due to the optics of such quantifications. In a marginalized community where being called "not deaf enough" is an entrenched insult,[27] measuring numerical levels of "deafness" would likely not go over well with prospective participants. If respondents feel they are being judged, they will feel a loss of face (Brown & Levinson, 1987; Hoza, 2007), so it would be unreasonable to expect them to answer completely transparently. Rather than measuring deaf identity using a scale, researchers often account for community-based factors by categorizing them in the form of demographic markers such as "language use in the home," "age of sign language acquisition," and "type of education" (Hill, 2014).

One advantage of categorizing community-based factors in the form of demographic markers for research purposes is that these factors are all external, and thus objectively

---

[27] See Bauman (2009) for a description of how the concept "not deaf enough" played a role in the 2006 Gallaudet protests. See Fernandes & Myers (2010) for a critical look at the need for more inclusive Deaf Studies programs, including the one at Gallaudet, and a more broad look at the need for more inclusivity within the American deaf community. Of note, Fernandes & Myers (2010) was first-authored by the person criticized as being "not deaf enough" during the events described in Bauman (2009).

identifiable. It is important to note that ranking such factors can not be done objectively; objective identification is not the same as objective ranking. There is no “type of education,” “language used in the home,” or “age of sign language acquisition” that objectively holds more or less value than another. Rather, the rank/value assigned to factors like these is subjectively defined at the community level, albeit nebulously. However, compared with subjective self-reports of how participants experienced status relationships during interactions, the externally identifiable factors are the most “objective” social status markers available for research purposes. Paradoxically, these community-based factors would almost certainly still need to be collected subjectively, via self-reporting.

Hill (2012) attempted to get at the possible effects of some objectively identifiable status markers, ones specific to the American deaf community. He asked deaf and hard-of-hearing people to rate signers’ language use on a scale from “strong ASL” to “non-ASL,” testing the effect of background information on rater perceptions. Importantly, the background information he provided was not factual, but assigned randomly to avoid correlation. Two of the categories of background information provided were family hearing status (deaf family or hearing family) and school setting (deaf school or mainstream school). Compared to a control group who were given no background information, raters who believed a signer was from a deaf family or had attended a deaf school were not significantly more or less likely to categorize that signer’s language as “strong ASL” or “non-ASL.” However, if raters believed a signer was from a hearing family or had attended a mainstream school, they were slightly more likely to categorize the signing as “Non-ASL.” Again, the school and family labels were not factual, but randomly assigned. In other words, the labels “hearing family” and “mainstream school” may have a negative effect on the perception of a signer’s ASL fluency, regardless of their actual language use. While the study was small, this work sheds light on some of the biases people hold even when they are unwilling to directly claim or discuss them.

The above considerations, especially the combination of the work done by Ladd (2003) and Hill (2012), reflect one of the general concerns with research related to status. Namely, people's willingness to discuss such matters frankly, no matter how real the effects may be, especially when being recorded and in the context of a research study. This is compounded when the study is being conducted in a signed language, as complete anonymity is compromised. These concerns can be avoided if status is measured using only externally measurable factors typical of sociolinguistic research conducted in signed language communities. This includes factors such as gender, age, ethnicity, and socioeconomic status, as well as community-based factors, such as language use in the home (during childhood and at the time of the study), age of signed language acquisition, and type of education (e.g., deaf school, oral program housed in a deaf school, mainstream program without interpreters) (Hill, 2014).

For these reasons, I have elected to define status using externally measurable (i.e., "objective") factors. As alluded to in the previous section, most of these factors are only measurable in the ordinal sense. The selected status markers are discussed in detail in Chapter 3. I also decided to collect information about participants' subjective experiences, both to draw comparisons with Kacewicz et al. (2013) (described later in this chapter), and to compare their subjective experiences of status to the objective measures.

Pronouns and Social Status

The role of attention.

At the beginning of this chapter, I named Kacewicz et al. (2013) as the paper that led me into this work. In their paper, the authors frame the relationship between pronouns and social status as the linguistic expression of the inverse relationship between social status and attentional focus. They say their work “suggests that status is associated with attentional biases, such that higher rank is linked with other-focus whereas lower rank is linked with self-focus” (p. 1). They propose that a person’s use of personal pronouns can be used to track where that person’s attention is focused, because they say these pronouns “reflect where people are paying attention” (p. 4). In short, their theory appears to be that status directs attention, and attention directs language, though they are careful not to claim causality. This can be thought of as a one-way chain, as seen in Figure 2.04.

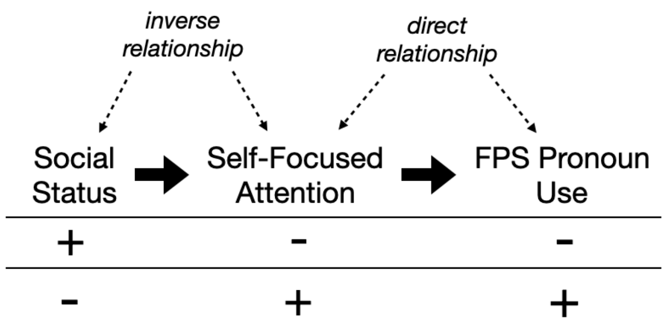
Figure 2.04

Relationships between social status, attention, and language use, as implied by Kacewicz et al. (2013)

Following this logic, if lower social status increases self-focused attention, and self-focus leads to FPS pronoun use, then people with lower social status would use more FPS pronouns. Conversely, higher social status would decrease self-focus (increasing other-focus), which would decrease FPS pronoun use. This would create an inverse relationship between social status and FPS pronoun usage rates; high status people with low FPS use, and low status people with high FPS use. Their theory as applied to self-focused attention can be seen below in Figure 2.05.

Figure 2.05

Nature of the relationship between social status, self-focused attention, and FPS pronoun use, as implied by Kacewicz et al. (2013)



The idea that a person's attentional focus will motivate their linguistic choices is not new; the direct relationship seen in Figure 2.05 has a solid foundation. See the previous section, "FPS construal without pronouns," for more on how language use encodes attentional focus, regardless of which language is being used.

The part of Figure 2.05 that is a bit harder to cleanly explain is the inverse relationship—the relationship between social status and attentional focus. We know that attention is finite, can be directed inward or outward, and can shift quickly (see Duval & Wicklund, 1972 and Carver & Scheier, 1981 for more). However, it is difficult to directly measure attentional focus, making it a tricky subject matter to study directly.

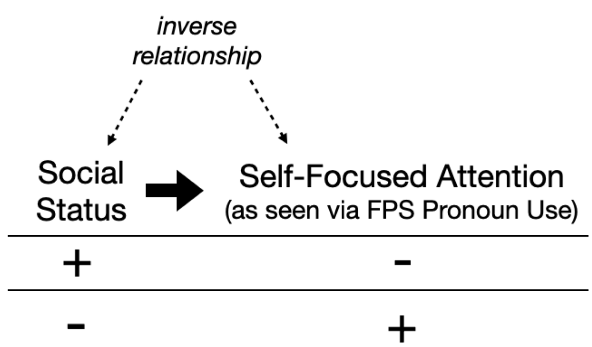
Instead, the well-understood connection between attention (one element of construal; see earlier section "FPS construal without pronouns" for more) and FPS pronoun use supports the idea of FPS pronoun use serving as a proxy for self-focused attention. In fact, this is a common proxy variable in research on mental health and psychology. See Brockmeyer et al. (2015) for a review of relevant literature.

There is also experimental evidence to support the idea of FPS pronoun use as a proxy for self-focused attention. Davis & Brock (1975) report that participants who were facing a mirror or a camera used[28] higher rates of first-person pronouns than did those performing the same task without any prompts to increase self-attention.[29] Sitting in front of a mirror or a camera objectively increases awareness of one's self. Although they did not differentiate singular (e.g., “me”) from plural (e.g., “we”), the significant first-person pronoun frequency difference they report lends evidence to the proxy relationship between self-attention and first-person pronoun usage frequency.

Updating Figure 2.05 in light of this information would result in a collapsing of “Self-Focused Attention” and “FPS Pronoun Use,” as in Figure 2.06 below.

Figure 2.06

FPS pronouns as a proxy for self-focused attention, as implied by Kacewicz et al. (2013)



[28] Pronoun “use” in their study was not typical language use. Instead, participants were given foreign-language texts and asked to guess the meanings of underlined pronouns.

[29] Interestingly, this effect disappeared when participants were given negative feedback prior to the pronoun activity. This tempers, but does not automatically invalidate, their other results, though it does add another layer of complexity.

There is currently no definitive answer to why lower social status seems to increase FPS pronoun use in the literature. In theory, those who are comparatively disadvantaged in a particular situation may focus their attention on themselves as a mechanism of self-preservation (e.g., Tausczik & Pennebaker, 2010), or perhaps as a submissive strategy (e.g., Sloman et al., 2003). For more on the potential connection between attention and status, see Kacewicz et al. (2013).

Linguistic Inquiry and Word Count (LIWC).

Much of the work linking pronoun usage rates to social status has made use of a word-counting software program called Linguistic Inquiry and Word Count, or LIWC (pronounced "Luke"), that was developed in the 1990s in the psychology department at the University of Texas at Austin (Pennebaker et al., 2015).[30] The software has been adapted for use in languages other than English, including Chinese (Huang et al., 2012), French (Piolat et al., 2011), Spanish (Ramirez-Esparza et al., 2007), German (Wolf et al., 2008), and Dutch (Zijlstra et al., 2004), but there is currently no version for ASL.[31] The following discussion will refer specifically to the English version of the software.

Essentially, LIWC is a frequency calculator, counting how many times words from predetermined categories occur in a text. The word "text" as used here refers not only to the orthographic representation of language, but more broadly to any captured (e.g., recorded)

[30] https://www.liwc.app/help/howitworks

[31] The lack of a standard writing system for ASL (orthographic or otherwise) would render an ASL version of LIWC only minimally useful. Additionally, the lack of automated transcription technology for signed languages (equivalent to auto-captions) means each signed event would need to be transcribed one unit at a time by someone trained in a coding system (e.g., using the ASL Signbank (Hochgesang et al., 2024) with SLAASh annotation conventions (Hochgesang et al., 2020)), which is a prohibitively time-consuming process. Hopefully, the slow building of various sign language corpora, coupled with future advances in technology, will make an ASL version of a LIWC-style software a practical option someday.

language event, be that written, spoken, or signed (including tactile languages). For LIWC to generate word frequencies, however, these texts must be machine-readable, i.e., in digital written form. Any text that originates in another form, such as a signed or spoken conversation, must be transcribed.[32] LIWC then generates word-type frequencies based on nearly 100 hierarchical word categories, called "dictionaries," designed with the aim of detecting the language user's social and psychological states. LIWC includes dictionaries for categories such as feelings, future focus, auxiliary verbs, and first-person singular pronouns, among others. Individual users can also create their own dictionaries. With some categorizations, a certain level of analysis is built into the dictionary. For example, the "affiliation" dictionary includes over 350 entries (e.g., words, word stems) that are presumed to reflect a person's desire to connect with others, such as "community" and "together." Other dictionaries are based on grammatical categories, without any inherent assumptions about usage implications, such as the pronoun dictionary, which is nested in the function word dictionary.

For conversations between two or more people (i.e., email threads, transcriptions of face-to-face conversations), the text of the entire conversation can be uploaded as a single text to generate word-type frequencies for the conversation as a whole, but more often, each interlocutor's language use is analyzed separately. When a text or group of texts is run through LIWC, each word is checked for membership in each of the dictionaries, resulting in a usage percentage for each dictionary. The word "me" belongs to the first-person singular dictionary, which is a subgroup of personal pronouns, which is a subgroup of pronouns, which is a subgroup

[32] See Johnston (2010) and Hochgesang (2014) for discussion of the differences between transcription and annotation with ID glossing. For my purposes here, I am using transcription as an overarching term.

of function words. Thus, every instance of "me" would count toward the total for each of these dictionaries.

FPS pronoun usage rates and social status.

To the best of my knowledge, this dissertation is the first work to be done linking social status and pronoun usage rates in any signed language, and there is little to no research of this kind done in non-English languages that has been published in English.[33] For these reasons, the rest of this section will focus on the work that has been done on the relationship between social status and FPS pronoun usage rates in English.

As mentioned in the introduction to this chapter, I came to wonder about the possible relationship between social status and FPS pronoun usage rates in ASL after reading about work done by Dr. James Pennebaker, with his students and colleagues, on the relationship between status and pronouns in English. I wanted first to identify the origin of the claim about FPS pronouns to better understand it. In this search for provenance, I found that as early as 1997, Pennebaker wrote about an inverse relationship between first-person pronoun use (singular and plural) and perceived dominance. Berry et al. (1997), second-authored by Pennebaker, asked college students to watch short videos of other students and then rate the students in the videos on various factors, including "dominance." This became their measure of "perceived dominance," and they reported a weak correlation between this measure and first-person pronoun use of the students in the videos. In Pennebaker & King (1999), first-person pronouns were separated out into singular and plural for an analysis of types of language that might mark "immediacy" (i.e., minimal psychological distance). This separation was a result of first-person

[33] See earlier section on LIWC for examples of languages that have a version of LIWC, though the publications are written in the language being analyzed.

plural pronouns overlapping with LIWC's "social words" category. Ten years after the initial claim of a connection between status and first-person pronoun use, Pennebaker had narrowed the discussion to a relationship between social status (rather than "perceived dominance" or "immediacy") and first-person singular pronouns (rather than all first-person pronouns), though still without citing any evidence from work that had been peer-reviewed. Specifically, Chung & Pennebaker (2007) claim an inverse relationship between FPS pronouns and social status, describing two studies as evidence. The first, Pennebaker and Davis (2006), appears in the reference list as "Unpublished data," and the second is a study—that is briefly described but not cited—analyzing tapes from the 1970s Watergate scandal.

A few years later, Tausczik & Pennebaker (2010) make the same claim about social status having an inverse relationship with FPS pronoun usage rates, this time citing what appears to be the paper that would become Kacewicz et al. (2013), second-authored by Pennebaker, though in the form of a 2009 manuscript submitted for publication.[34] When she first-authored Kacewicz et al. (2013), Ewa Kacewicz was a graduate student in the Department of Psychology under the guidance of Dr. James Pennebaker at the University of Texas at Austin. She graduated with her PhD in 2013,[35] the same year that Kacewicz et al. (2013) was first published online,[36] and her other two academic publications, both prior to 2013, include Pennebaker as a co-author.

---

[34] This was cited in their reference list as a "manuscript submitted for publication," with the same authors, but a slightly different title, "The language of social hierarchies," as opposed to the final title, "Pronoun use reflects standings in social hierarchies."

[35] In her dissertation, Kacewicz (2013), she references the Kacewicz et al. (2013) paper (listed as "under review" at the time) as support for two of her three hypotheses.

[36] It was later published in the same journal in print in 2014. This later version, Kacewicz et al. (2014), is identical in every way, except for the year, volume (later updated in the original), number (also later updated), and page numbers.

Tausczik & Pennebaker (2010) also include a table titled "Summary Table Linking LIWC Word Categories to Published Research Studies," in which they list 38 studies connecting FPS pronouns to the following "psychological correlates": "Honest, depressed, low status, personal, emotional, informal" (p. 39). This suggests a robust correlation between FPS pronoun use and status, so I was surprised to find that none of the 38 studies directly addresses a potential connection between FPS pronouns and social status. It is possible that the reported findings were re-interpreted with this lens by Tausczik & Pennebaker (2010) before inclusion in their chart (e.g., by using factors like gender or age as proxies for social status), but this is unclear from their paper. Other publications that directly claim an inverse relationship between FPS pronoun usage rates and social status tend to cite some combination of the aforementioned work as foundational evidence. Four such studies are described later in this chapter, all of them also making use of Pennebaker's LIWC software.

As it stands, Kacewicz et al. (2013), is the most widely-cited paper to claim a relationship between FPS pronoun usage rates and social status, with Elsevier's Scopus database showing 357 citations and Google Scholar showing 668.[37] Kacewicz et al. (2013) has also been referenced in the media, and the conclusions drawn in their paper have been shared by Pennebaker and others in multiple popular culture mediums.[38] The prominence of Kacewicz et al. (2013) makes it a

[37] Both counts are as of June 2024.

[38] e.g., Pennebaker (2011); 2011 article in the New York Times:

https://www.nytimes.com/2011/08/28/books/review/the-secret-life-of-pronouns-by-james-w-pennebaker-book-review.html; 2011 article in Harvard Business Review:

https://hbr.org/2011/12/your-use-of-pronouns-reveals-your-personality; 2011 article in NewScientist:

https://www.sciencedirect.com/science/article/abs/pii/S0262407911621672; 2012 article in Yale Scientific:

https://www.yalescientific.org/2012/03/the-secret-life-of-pronouns/; 2013 TEDx Talk by Dr. James Pennebaker: https://www.youtube.com/watch?v=PGsQwAu3PzU; 2017 HuffPost article:

logical place to begin for understanding what is currently known about the relationship between FPS and social status.

Looking into the five studies in Kacewicz et al (2013), my initial goal was to identify explicit methodological details that could aid in potential replication. As I worked through each study, I began to encounter obstacles that led me to consider their results and claims more carefully. During this process, I attempted to give the authors the benefit of the doubt whenever possible, including acknowledging the challenges of presenting work within the limited space of a journal article. Unfortunately, there are a number of things that pose potential challenges to replicating their studies, including methodological decisions made by the authors, inconsistencies, and omitted details. Below I will provide an overview of their work as described in their paper, along with elements that explain my decisions to deviate from their methodology in my own work. A deeper dive into Study 3 can be found in the following chapters, related to my work with their data from Study 3.

Kacewicz et al. (2013).

In Kacewicz et al. (2013), the authors describe five studies connecting social status to usage rates of various pronouns. Two of these studies used transcripts from spoken language interactions and the other three used texts generated in written form. The studies included many types of pronouns, both personal (first-person singular and plural, second-person, and third-person singular and plural), and impersonal (e.g., it, that). The FPS pronouns they counted

---

https://www.huffpost.com/entry/how-your-words-reveal-you_b_9144168; 2019 Quartz article: https://qz.com/work/1602069/the-history-of-the-royal-we-and-why-its-so-annoying.

included “I,” “me,” “my,”[39] “myself,” and “mine.”[40] They used LIWC to calculate pronoun usage frequencies (see previous section for details), and defined status overarchingly as “position/rank within a social hierarchy” (Kacewicz et al., 2013, p. 2). The authors concluded, among other things, that these five studies provided evidence for an inverse relationship (i.e., negative correlation) between FPS pronouns and social status, suggesting that higher status is associated with lower FPS usage. They posited that a difference in attentional focus was the cause, rationalizing that lower-status individuals likely focus more attention on themselves, generating FPS pronouns more frequently.

As discussed earlier in this chapter, social status is a somewhat nebulous concept, and tricky to pin down. Kacewicz et al. (2013) decided to handle this by operationalizing status slightly differently in each of their five studies. This attempt to triangulate an approximation of “status” may seem to have its benefits; getting similar results across different contexts appears to suggest robustness. On the other hand, changing the definition of a variable across studies also allows for confirmation bias to aid in the selection of a definition for each specific data set that is most likely to lead to the desired conclusions. This exploitation need not be a conscious decision; it may appear to be a happy accident. Still, with each statistical test that is run, there is a corresponding increase in the likelihood of arriving at a conclusion that looks significant, but that is actually the product of chance. If definitions and statistical tests are determined prior to data collection, the likelihood of these false findings decreases (Nuzzo, 2014). To look more closely

[39] The English word “my” is alternatively considered a possessive determiner, dependent possessive pronoun, and possessive adjective (e.g., Biber et al., 1999). As “my” serves to signal first-person singular construal as much as words like “me” and “myself,” it and other similar forms (e.g., our) are considered pronouns for the purposes of the present work.

[40] The full list of forms in the FPS pronoun dictionary at the time was: “i, i’d, i’d’ve, i’ll, i’m, i’ve, id, icd, idgaf, idk, idontknow, idve, ikr, ily*, im, ima, imean, imma, ive, me, methinks, mine, my, myself” (J. Pennebaker, personal communication, November 6, 2017).

at how status was defined in each of their studies, the summary of each study below describes the aspects that are relevant to first-person singular pronoun use, defining social status, and reporting potential relationships between the two.

The authors preface the descriptions of their five studies by saying, "For the purpose of this article and simplicity sake, unless describing a specific operationalization of rank, we will refer to position in the hierarchy as status" (p. 5). For my descriptions below, I will follow this convention. Exceptions include direct quotes from their paper that use different terms (e.g., power).

While I designed my analyses differently than Kacewicz et al. (2013), I still wanted to be able to compare patterns between ASL and English. This meant selecting English data that could be most directly compared with the ASL data set. The asynchronous nature of most written communication, including the emails and letters used in Studies 4 and 5, allows for each writer to hold the floor as long as they like without interruption or influence from the other person. This creates a different dynamic from that in the real-time conversations. While the data from Study 2 was from real-time conversations, it was also generated in written form (i.e., instant messenger chat). For reasons described in the section above on cross-linguistic frequency of FPS pronouns, conversational data in ASL is more directly comparable to conversational data in spoken English (as opposed to written English), when it comes to pronoun use. As a result, I am particularly interested in Studies 1 and 3 from Kacewicz et al. (2013), as those are the two studies that used data from spoken English, rather than data originally generated in written form. Though my aim was not to replicate their methods exactly, my data collection design was based most closely on their Study 3. In an attempt to compare patterns between ASL and English, I wanted to analyze my data using the approach from their Study 3, as well as analyzing their Study 3 data using my approach. Fortunately, the senior author, Dr. James Pennebaker, was kind enough to share the data from their paper. The work I did to this effect is described in the following chapters. The information below represents the five studies in Kacewicz et al. (2013) as reported by the original authors.

Study 1 – Kacewicz et al. (2013).

Study 1 made use of data collected by Burris et al. (2009) for an experiment conducted by the McCombs School of Business at the University of Texas at Austin. The experiment was designed to "examine the effects of inner-circle membership on group dynamics and interpersonal influence in hierarchical teams" (Burris et al., 2009, p.1244). Data was collected from 41 groups of four participants each. The authors of the original study specified that there were two friends and two strangers in each group, with the group leader randomly chosen from one of the two friends. The participants were deliberately misled though, being told that leaders had been selected based on the results of the personality questionnaire they had all taken. Groups were given a decision-making task, and their discussions were recorded and transcribed. Leaders were given the final say in the group's submission of their completed task, while also being paid a higher base rate ($7, compared to $4 for non-leaders). If bonuses were received for group performance, leaders were responsible for assigning amounts to non-leaders (one $3, one $6, one $9). Thus, status discrepancies were binary, defined as leaders and non-leaders,[41,42] and these distinctions were externally assigned, with all participants being objectively aware of them.

---

[41] Burris et al. (2009) refer to the participants who are not given the role of leader as "subordinates," with one exception where they are called "non-leaders" in a figure. Kacewicz et al. (2013) use both "subordinates" and "nonleaders." I have chosen to use "non-leaders" here. All refer to the same participants.

[42] In their paper, the original authors provide compelling evidence for an intermediary level of status. This would make the three resulting status designations: leaders (one per group), non-leader friends (one per group), and non-leader strangers (two per group). This hierarchy was not explored in Kacewicz et al. (2013).

When Kacewicz et al. (2013) analyzed the resulting transcripts, they compared the FPS usage rate for the total group of non-leaders (n = 123) to the average rate for the group of leaders (n = 41). The reported mean FPS pronoun usage rate for leaders was 4.30, meaning FPS pronouns represented 4.30% of the words spoken by leaders. The average rate for non-leaders was 5.60. In other words, FPS pronouns made up, on average, 4.3% of leaders' word use, and 5.6% of non-leaders' word use. These rates represent a discrepancy of 1.3 percentage points between leaders and non-leaders.

In testing for statistical significance of the difference between these means, paired *t*-tests were used to obtain the *p*-values, which I found surprising, given the different numbers of participants in these two groups. It is not clear from their paper how the data was "paired" for this test, but one way this could be finagled is by averaging the three non-leaders in each group into a single value. This violates the assumptions of the paired *t*-test, strictly speaking. The results would thus be comparing the leader in each group to a nonexistent "average" non-leader, obscuring variations in the status relationships between group leaders and each of the individual non-leaders. The effect size for FPS pronoun usage rates in Study 1 is reported in their Table 3 (p. 12) as a Cohen's *d* of -1.02. This makes it appear that, on average, leaders used considerably fewer FPS pronouns than non-leaders. However, it is unclear how this effect size was calculated, given that the reported group means and standard deviation suggest a different effect size, *d* = -.70.

Studies 2 and 3 – Kacewicz et al. (2013).

Studies 2 and 3 used data collected by the authors themselves, and status was measured subjectively, defined as self-reported "power." Study 2 analyzed text-only online chats conducted via instant messenger, with dyads working together to solve a series of complex problems. Study

3 consisted of dyads having get-to-know-you chats in person. After their interactions, each participant was asked to complete an Interaction Rating Questionnaire (IRQ; see Appendix) to capture their subjective sense of the status dynamic during the interaction. The authors report using a 7-point Likert scale, and the questions they report using for this measure are “To what degree did you control the conversation?” and “To what degree did you have power in the conversation?” where 1 = *not at all* and 7 = *a great deal* (p. 7). The authors explain that, for both studies, dyad-level status discrepancies were calculated “by subtracting the self-rating of power of the higher rated member of the pair from the member with the lower rated power” (p. 7). See Chapters 3 and 4 for compelling evidence that status discrepancies may instead have actually been calculated without regard for who was reported to be higher or lower in status, a decision which would drastically alter the outcome. Regardless as to whether the subtraction was ordered according to participant self-reports, calculating dyad-level discrepancies by subtracting one person’s self-report from the other’s appears to contradict the authors’ framing of status as relative and zero-sum (as opposed to variable-sum, a distinction discussed earlier in this chapter). Additionally, the reported correlations for dyad-level agreement about relative status were moderate (-.38 in Study 2 and -.46 in Study 3) suggesting the need for a closer look at agreement between chat partners. Despite this, no explanation is given for how dyad-level discrepancies were calculated for dyads whose contradictory self-reports make it impossible to objectively determine who was of higher or lower status. Both the calculation of dyad-level status discrepancies and the nature of dyad-level agreement are discussed in more detail in Chapter 3.

Pronoun discrepancies for each dyad were reported as being similarly calculated, by subtracting the higher-status participant’s FPS usage rate from the lower-status participant’s rate. The authors then correlated the status discrepancies with the pronoun usage discrepancies. It is unclear from their paper how the authors determined which person's FPS usage rate would be subtracted from the other's not only in those instance when dyads disagreed about who had higher status, but also when both members of a dyad self-reported the same status rating (e.g., both reported themselves as a 3 out of 7). While the absolute value of the pronoun usage discrepancy would be the same regardless of whose pronoun usage rate was subtracted from whose, arbitrarily assigning an order would make the pronoun discrepancy artificially appear to be positive or negative. This suggests that including dyads who either disagreed about their relative status and/or who self-reported the same status rating would be akin to introducing random noise into the data set, as the sign of the value (i.e., whether it is positive or negative) is critical in the context of the quantitative analyses of these studies. This concern is addressed further in Chapter 3.

Kacewicz et al. (2013) report the relationship between status discrepancy and FPS pronoun use in Study 2 with a correlation coefficient of -.55, which is considered moderate-to-large, measured on a scale of -1 to 1. The associated *p*-value of $p \leq .01$ is considered statistically significant. For Study 3, they report the relationship between status discrepancy and FPS pronoun use with a correlation coefficient of -.30, which is considered weak-to-moderate. The associated *p*-value of $p \leq .05$ is generally considered to be statistically significant. The correlation coefficients being negative indicates that lower-status individuals tended to use higher rates of FPS pronouns than higher-status individuals. Despite these apparently interesting results, the methodological issues mentioned here (and described further in Chapter 3) make it difficult to ascertain how the results reported by Kacewicz et al. (2013) should be interpreted.

Study 4 – Kacewicz et al. (2013).

In Study 4, status discrepancies were also subjectively defined, but with only one person per dyad contributing a rating. Participants provided emails they had exchanged with a number of correspondents, and also provided self-reports of the relative "status" for each correspondent, relative to the participant themselves. These ratings were on a scale from 1 (*other has much lower status*) to 7 (*other has much higher status*). The participants' correspondents did not provide ratings, so they may or may not have agreed with the participants' ratings.

The authors reported calculating pronoun usage discrepancies by subtracting the pronoun usage rates of participants from the pronoun usage rate of the correspondents; that is, "received" pronouns minus "sent" pronouns for each participant-correspondent pair. Based on the descriptive statistics provided for word count, it appears that many of the emails were very brief, with some extreme exceptions. For each participant, correlation coefficients were calculated for each of their correspondents. Each participant's correlation coefficients were then averaged at the level of the participant. The authors report a negative correlation between relative status and FPS pronoun use, with a correlation coefficient of -.43, meaning that lower-status individuals tended to use higher rates of FPS pronouns than higher-status individuals.

Study 5 – Kacewicz et al. (2013).

In Study 5, objective status measures were used, namely, military rankings. This study analyzed 40 letters translated from Arabic between Iraqi military personnel associated with the regime of Saddam Hussein. The letters were translated into English as part of a document archive project by the Department of Defense (Woods, 2007), and the English translations were used for Study 5. The data set was divided equally into two groups, creating a binary status variable: those of lower status writing to someone of higher status, and those of higher status writing to someone of lower status. No indication is given that any of the letters were written in response to each other, so as in Study 1, pronoun frequency was not compared within conversations. Rather, pronoun use by the group of lower status writers was compared to pronoun use by the group of higher status writers.

The authors report a correlation coefficient of -.27 for FPS pronouns, which is a weaker correlation than those reported in the first four studies. The *p*-value for this correlation ($.06 < p \leq .1$) is not generally accepted as statistically significant. As in Study 1, *p*-values were obtained through paired *t*-tests, which are inappropriate for comparing the means of two independent groups of people who were in no way "paired."

Of note, the nature of the translations suggests that preservation of meaning likely took precedence over preservation of form, including pronoun use (e.g., types, frequency). Thus, this data set is likely a better representation of English-target translated texts than a representation of either Arabic or English.[43]

Meta-analysis – Kacewicz et al. (2013).

Kacewicz et al. (2013) define status in their five studies variously as "leadership," "self-reported power," "self-reported status contingent on actual position," and "objective military rank" (p. 5). In their meta-analysis of the five studies, the authors compress these status variables, explaining, "to simplify the discussion, position in social hierarchy will be referred to as status" (p. 12). As discussed in my introduction to their paper, there are pros and cons to changing the definition of a variable across studies. That aside, the authors report effect sizes in their Table 3, titled "Meta-Analysis Combing [sic] Effect Sizes (Cohen's *d*) of Language Effects From All Five Studies" (p. 12), interpreting the results by saying "lower status individuals used

[43] See Baker (1993, 1997) for discussion of norms specific to translated texts.

overwhelmingly more ‘I’ across all five studies (*d* = .85) as compared with higher status individuals” (p. 12). It is unclear how *d* = -0.21 for Study 5 warrants the use of “overwhelmingly,” as this value is considered a small effect size. Additionally, as discussed earlier, it is unclear how the *d* value for Study 1 was calculated, given the reported values. For Studies 2 and 3, their Table 3 Cohen’s *d* values appear to have been converted from the correlation coefficients in their Table 2 (p. 8) using standard methods. However, see the section above on Studies 2 and 3 for reasons why the correlations they started from are questionable. The Cohen’s *d* values for Studies 4 and 5 do not appear to be computationally reproducible using standard methods, and the authors do not provide any insight as to how these values were calculated.

At face value, the FPS pronoun results reported for these five studies appear to be similar, suggesting that lower status individuals tended to use higher rates of FPS pronouns. Due to the issues I have raised in this and the preceding sections, it is unclear how reliable these results are.

Other studies purporting a negative correlation between FPS pronouns and status in English.

An inverse relationship between FPS pronoun use and status has been discussed in other papers besides Kacewicz et al. (2013). All appear to tie back to the work done by Pennebaker and his colleagues. Some of these reports are simply loose interpretations of results from studies involving proxies for status. Others are harder to determine.[44] Once a single paper overgeneralizes or mischaracterizes results, that misinformation can then easily be passed on as fact, leading to undeserved amplification (Greenberg, 2009).

---

[44] See examples in the section titled “FPS pronoun usage rates and social status,” including Chung & Pennebaker (2007).

What follows is an example of such a downstream effect, which I refer to as a "cascaded misrepresentation," specific to the claim that FPS pronoun use is inversely related to status. Later in this section, I will describe four papers that make this claim, and all four include similar cascaded misrepresentations. For this example though, I will trace the citation trail for the version that appears in Markowitz (2018). I will start by sharing Markowitz's claim and his citation for it. Then I will present the relevant claim from the paper he cited, which is itself a secondary source, including that paper's citations, which are the primary sources. Finally, I will summarize what the primary sources say about FPS pronouns and status. Note that all of the papers in this example include Pennebaker as a co-author, except for Markowitz (2018), which has a different connection to Pennebaker, described in the later section devoted to that paper. Note also that the findings reported in the primary sources are inconsistent with the claims made at the secondary and tertiary levels.

- Tertiary source: Markowitz, David M. (2018). Academy Awards Speeches Reflect Social Status, Cinematic Roles, and Winning Expectations. *Journal of Language and Social Psychology*.

  > "Leaders, or people of high social status, often speak for a collective by using 'we' words compared with nonleaders, or people of relatively low social status, who often speak for the self and use 'I' words (Hancock et al., 2010)" (p. 377).

- Secondary source: Hancock, Jeffery T., David I. Beaver, Cindy K. Chung, Joey Frazee, James W. Pennebaker, Art Graesser, & Zhiqiang Cai. (2010). Social language processing:

A framework for analyzing the communication of terrorists and authoritarian regimes. *Behavioral Sciences of Terrorism and Political Aggression*.

> “There are sex, age and social class differences in function word use (e.g. Newman, Groom, Handleman, and Pennebaker, in press; Pennebaker & Stone, 2003). For example, first-person singular pronouns (e.g. I, me, my) are used at higher rates among women, young people and among people of lower social classes” (pp. 113-114).

- Primary source #1: Newman, Matthew L., Carla J. Groom, Lori D. Handelman, & James W. Pennebaker. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*.

  The authors do not discuss status; in fact, the word “status” does not appear in this paper. They do report women using more FPS pronouns in their data samples, which suggests this citation better supports the claim above related to gender rather than status.

- Primary source #2: Pennebaker, James W., & Lori D. Stone. (2003). Words of wisdom: Language use over the lifespan. *Journal of Personality and Social Psychology*.

  The authors do not discuss status directly here either, but they do report FPS differences for age and gender. If a reader were inclined to interpret age and gender as proxies for status, they would need to consider greater age as equating to greater status, and presume that men are of greater status than women.

It may seem unusual to include such a detailed report investigating the provenance of a single claim. I have chosen to include it because the claim in question is central to my research, and is what inspired the present work. Crucially, the above is but one example of a cascaded misrepresentation of primary data used to assert an inverse relationship between FPS pronoun use and social status. For more, see the discussion of Tausczik & Pennebaker (2010) in the earlier section titled "FPS pronoun usage rates and social status." The misrepresentation of primary data, including presenting unpublished data in lieu of peer-reviewed data (e.g., Pennebaker & Davis, 2006, discussed earlier), is a problem not only for the fields of linguistics and psychology (Gawne & Styles, 2022), but for any scientific research (Task Group on Data Citation Standards and Practices, C-I., 2013). When these claims are amplified through heavy citation, the problem is likewise amplified.

While an inverse relationship between FPS pronoun use and social status is frequently presented as a given, I wanted to find research presenting evidence, or actual data, for this claim. Despite hitting many dead-ends, I have found four papers in addition to Kacewicz et al. (2013) that present evidence for the theory that there is an inverse relationship between FPS pronoun use and social status. These four papers still all start with the assumption that this theory is a given. Although Pennebaker is not on the author list, each of them cites between 3 and 10 of Pennebaker's earlier papers, in addition to his LIWC software. Another commonality within these four papers is that all of the first authors have limited research experience. I myself have limited research experience, so it would be hypocritical of me to suggest the early work of all researchers is not to be trusted. That said, all research should be able to stand on its own merit. See below for a description of each of these papers, presented in the order of their publication.

My descriptions of these four papers are brief, in part because these studies differ from mine in two important ways. First, they are conducted mostly on written language samples, with the exception of Academy Awards speeches, which are likely to have been heavily scripted. See earlier section titled “Cross-linguistic frequency” for more on the greater alignment of signed language data with spoken language data, rather than written. Second, none of these studies used conversational language data; in addition to the previously mentioned speeches, the other studies use tweets (i.e., posts on Twitter) or posts to online message boards with only one post per user included in the data set. Despite these differences from my own approach, as well as some apparent inconsistencies within these papers, I feel it is important to include summaries of these studies, to provide an overall sense of the current state of the literature on FPS and status. I would like to do a deeper dive on each of these to try and make sense of the apparent inconsistencies, but that is a project for a future publication. For now, I will provide overviews, along with observations that are relevant to either comparing their work with my own, or to evaluating the contribution their work makes to our understanding of the potential relationship between FPS pronouns and social status.

Dino, Reysen, & Branscomb (2009). Online interactions between group members that differ in status.

As mentioned in the earlier section titled “FPS pronoun usage rates and social status,” Kacewicz et al. (2013) appears to be Pennebaker’s first peer-reviewed paper providing evidence of the claim that FPS pronoun use is negatively correlated with status. Dino et al. (2009) cites Pennebaker’s earlier work on FPS and status, and was published around the same time that Kacewicz et al. (2013) was originally submitted for publication (see section “FPS pronoun usage rates and social status” above for details). At the time of publication, all three authors were affiliated with the University of Kansas; the first author, Amanda Dino, was an undergraduate student, the second author was doctoral student Stephen Reysen, and the third author was professor of psychology Nyla R. Branscombe.

Dino et al. (2009) analyzed data from online message boards dedicated to various interest groups. The posts they selected fit into one of two status categories: a lower-status group member writing to a higher-status member (which I will refer to as "low-status members") or a higher-status member writing to a lower-status member (which I will refer to as "high-status members"). Status was measured using the displayed post-count for each member, which is consistent with social norms on message boards. The authors looked for differences between status groups: the content of their communications (e.g., requests, instructions) and the language they used (e.g., FPS pronouns, words longer than six letters). The content categories were coded by two independent raters, and the language categories were counted using the 2001 version of Pennebaker's LIWC software.

In an effort to avoid any individual poster having an outsized effect on the results, the authors chose to include only a single post from any username, meaning that each person would only be included once. This decision also has the effect of preventing the results from providing insight on the kinds of patterns that might emerge over the course of a conversation. Their stated purpose was different than mine, however, and it was "to examine intragroup interactions between high and low status members on Internet forums" (p. 86). With this in mind, it is important to remember that the results are not intended to represent patterns within conversations.

Unfortunately, the results reported by the authors are difficult to interpret due to some apparent inconsistencies.[45,46] However, the authors' interpretation of the relationship between FPS pronoun use and status is clear: they believe their data set supports the pattern reported in Pennebaker's work. Specifically, they say, "supporting previous research (Chung & Pennebaker, 2007; Pennebaker & Davis, 2006[47]), low status members did use first person singular pronouns (e.g., *I, my, me*) more extensively than high status group members" (p. 90).[48] Indeed, their Table 2 (p. 89) shows a statistically significant difference between low-status and high-status members in their use of FPS pronouns. Low-status members averaged a 4.51% FPS rate, and high-status members averaged a 2.27% FPS rate. This represents a difference of 2.25 percentage points between the two groups. It is difficult to put these numbers into perspective, as the authors do not report any information about the length of posts in their data set, though they did state that they only included posts that had a minimum of two complete sentences. To see what the results might look like in use, we can calculate FPS pronouns used in a hypothetical 50-word post from each group: the mean of 4.51% for low-status members would represent 2.3 FPS pronouns used, and 2.27% for high-status members would represent 1.1 FPS pronouns used. This would be a difference of approximately 1 FPS pronoun within approximately 50 words.

---

[45] As one example, they report n = 251 posts (p. 87), which represent 251 people (p. 88). This is inconsistent with the n implied by the degrees of freedom reported under their Tables 1 and 2 (p. 89), which would be 242. Both of these are also inconsistent with the n values inferred in the results section: 154 for status communication direction ($n - 2 - 8 = 144$), and 1029 ($n - 2 - 8 = 1019$) for type of group (p. 88).

[46] The interaction between status communication direction and type of group shows the total degrees of freedom for the error term in the MANOVA as 1019 (p. 88), which implies a non-integer result for the number of dependent variables, or p ($1019 = (251 - 16) \times (p - 1)$ suggests $p = 5.34$). A non-integer number of dependent variables would be nonsensical, so this appears to be an error.

[47] [footnote mine] As previously mentioned, Pennebaker and Davis (2006) is a reference to unpublished data first cited in Chung & Pennebaker (2007).

[48] Both citations in this quote are discussed earlier in the section titled "FPS pronoun usage rates and social status." Neither present data to support the claim that FPS pronoun use is correlated with social status.

Be that as it may, the authors include an endnote clarifying that only three of their seven interest group categories showed a significant difference in FPS use between low-status and high-status members. While initially surprising, this is actually in line with another of their results, namely, that the language differences between the interest group categories were statistically significant. Given this, it should come as no surprise that the interest groups could have significantly different patterns across multiple variables. It appears that FPS pronoun use was one such variable. With this information we can surmise that if each of the seven interest group categories were analyzed independently (or treated as separate studies), only three out of the seven would have been statistically significant for FPS by status; a rate that could be reasonably explained by chance. Additionally, it is unclear from their paper how many participants were in each of the three significant groups, what the magnitude of the FPS difference was in each of those groups, and what level of significance was found for each. The inconsistencies in their paper, along with the details (additional and missing) of the FPS results, make it difficult to know how best to interpret the authors' reported findings.

Reysen, Lloyd, Katzarska-Miller, Lemker, & Foss (2010). Intragroup status and social presence in online fan groups.

This paper, first-authored by the second author of the previous paper, uses a similar methodology to the previous one, though to address a slightly different topic. The other authors being students or other recent graduates, Reysen appears to have been the senior author, having completed his PhD in 2009.[49,50] The stated purpose of Reysen et al. (2010) was "to examine the relationship between intragroup status and the use of social presence cues in online fan message forums" (p. 1315). FPS pronouns were included as indicators of the social presence feature of "immediacy" (i.e., minimal psychological distance), per Pennebaker & King (1999). Like Dino et al. (2009), Reysen et al. (2010) collected data from multiple online message boards for various interest groups, never including more than a single post from any username, and status levels were defined by user post counts. Some improvements were made to the methodology, such as raising both the minimum number of posts per group and the thresholds for determining status levels. Another change was that, rather than selecting posts that were addressed to a specific member, such as a lower-status member addressing a higher-status member, Reysen et al. (2010) selected posts that were addressed to the group at large. This decision supports the authors' treatment of the data as being non-dyadic (i.e., not representative of conversational data, p. 1315).

Like Dino et al. (2009), Reysen et al. (2010) report a statistically significant difference between high-status and low-status members in their use of FPS pronouns—in the overall data set. Unlike Dino et al. (2009), Reysen et al. (2010) do not report any analyses or results that look at the individual interest groups, instead treating the posts from all four interest groups as a single data set. While there is nothing wrong with this decision per se, reading this paper after reading Dino et al. (2009) makes me wonder if reporting the group-level data in this otherwise

[49] At the time of publication, Jason Lloyd appears to have been a graduate student, Iva Katzarska-Miller had also completed their PhD in 2009 (though with fewer publications at the time), Brett Lemker had received his BA in 2008, and Russell Foss appears to have received his BA in 2010.

[50] Reysen's dissertation is titled "Identity Theft: Moral Antecedents, Moral Anger, and Impression Management," and it does not appear to be related to the subject matter of the article in question.

similar study would have weakened the claim about FPS pronouns. That aside, Reysen et al. (2010) report the mean FPS rate for low-status members as 9.24% and the mean for high-status members as 4.07%. This represents a difference of 5.17 percentage points between the two groups. As in Dino et al. (2009), it is difficult to put these numbers into perspective, as the authors do not report any information about the length of posts in their data set, though they did state that they only included posts that had a minimum of two complete sentences. As with Dino et al. (2009), in order to see what the results might look like in use, we can calculate FPS pronouns used in a hypothetical 50-word post from each group: the mean of 9.24% for low-status members would represent 4.6 FPS pronouns used, and 4.07% for high-status members would represent 2.0 FPS pronouns used. This would be a difference of approximately 2.6 FPS pronouns within approximately 50 words.

As previously alluded to, one salient methodological similarity with Dino et al. (2009) is the starting assumption that FPS pronouns are inversely correlated with social status. Reysen et al. (2010) use this assumption to justify including FPS pronouns as indicators of immediacy, and then conclude that their results, in turn, justify using post counts as their measure of status. "The greater use of first person singular pronouns (e.g., I, me, my) by low status members supports previous research regarding language use and intragroup status (Chung & Pennebaker, 2007; Dino et al., 2009; Newman et al., 2003; Pennebaker & Davis, 2006). This finding supports our operational definition of using post counts as indicators of status" (p. 1316). Note that, for each of the studies they reference in this quote, I have already discussed limitations regarding the claims therein about FPS and status. Taken together, these observations make me question the foundation on which this paper was built.

Beach, Brownlow, Greene, & Silver (2016). The "I"s have it: Sex and social status differences on Twitter.

Unlike the previous two papers, Beach et. al (2016) does not contain overlapping authors, and the data did not come from online message boards. Similar to those previous two papers, the first author was a student at the time of publication; Jennifer Beach and Melissa Greene were both undergraduate students working under the guidance of Dr. Sheila Brownlow. Also in line with the previously described studies, Beach, et al. (2016) report an inverse relationship between FPS pronoun use and status, citing Pennebaker's work,[51] and they take this inverse relationship as a given. Their aim was to test how this effect interacted with sex and age in data from famous people's tweets (i.e., posts on Twitter). Tweets were collected from the top 50 most popular actors (25 female and 25 male)[52] during a one-month period in 2015. For each tweet, an FPS rate was calculated as a percentage of total words used. At the time, Twitter still imposed a 140-character limit per tweet, meaning each FPS percentage was likely calculated from a total of no more than about 35 words. The authors operationalized status using the number of followers each person had, split into two groups at the median number of followers. As a result, actors with fewer than 1,063,500 followers were considered to be of lower status, and those with more were considered to be of higher status.

Their data show lower-status famous actors used FPS pronouns at slightly higher rates than did higher-status famous actors. The mean for the higher-status group was 2.00%, and the lower-status group mean was 3.12%. This represents a difference of 1.12 percentage points

[51] This includes seven citations for Pennebaker's general-audience pop-psychology book, *The Secret Life of Pronouns* (2011).

[52] The criteria of "most popular" was measured using www.ranker.com for 2015, with separate lists for male and female actors.

between the two groups. To put these numbers in perspective, one FPS pronoun in a tweet of 35 words (an approximated maximum for 140 characters) would be a rate of 2.86%. For shorter tweets, one FPS pronoun would represent higher FPS rates. As an example, the tweet “Hey guys, wanna feel old? I'm 40. You're welcome,”[53] which contains one FPS pronoun in its nine total words, resulting in an FPS rate of 11.11%. Recall that FPS rates were calculated for each tweet, rather than each person’s collected body of tweets. In effect, the range of possible non-zero FPS rates begins around 2.86% and runs to 100%. The fact that the group means (2.00% and 3.12%) fall near or below the non-zero minimum suggests an extremely high percentage of tweets with no FPS pronouns at all. The authors acknowledge that their data are not normally distributed, and the standard deviations reported for each group (subgrouped by sex and status) are quite a bit larger than the means, suggesting a level of variability where means start to become less useful.

While Beach et al. (2016) do not provide an analysis of the content of the tweets, they do propose that the content may provide a more robust explanation than status does for the different rates of FPS pronoun use. Namely, they say “Twitter users with a higher social status may have used their following as a platform to Tweet about various causes, thus removing themselves from their tweets” (p. 20). They also acknowledge that some famous people have others post to their accounts for them (e.g., publicists), which would invalidate the results from those individuals.

While this is an interesting paper, the results do not represent a compelling argument in support of the theory that FPS pronouns are inversely related to social status in a meaningful way.

[53] This was tweeted by Macaulay Culkin (@IncredibleCulk) on August 26, 2020, and it received 2.9 million likes (Culkin, 2020).

Markowitz (2018). Academy Awards speeches reflect social status, cinematic roles, and winning expectations.

This paper is where the earlier example of cascaded misrepresentation came from. Recall that Markowitz (2018) cites Hancock et al. (2010), which includes Pennebaker on the author list, when stating as fact the theory that FPS pronouns are inversely related to status. This is followed by citations from Kacewicz et al. (2013), which also includes Pennebaker. Both Dr. Jeffrey Hancock and Dr. James Pennebaker feature prominently in Markowitz's reference list.[54] This is not surprising, as David Markowitz completed his dissertation the same year Markowitz (2018) was published,[55] and Hancock was Markowitz's PhD advisor, with Pennebaker serving on his dissertation committee. There is nothing inherently wrong with any of this, though it could potentially explain how the cascaded misrepresentation in this paper may have occurred.

Markowitz (2018) analyzes transcripts of Academy award acceptance speeches (n = 220) made by actors and directors. He considers award-winning actors to be of relatively lower social status than award-winning directors, via a "leader versus non-leader" lens. I can see how directors occupy more of a leadership role than do actors, though equating this with being of higher social status is harder for me to justify. This apparent logical leap may be explained by the author's statement that, in his study, "social status was assessed through pronouns" (p. 377). This appears to run counter to Markowitz's hypotheses about pronouns, which make it appear as though status was set as the independent variable prior to data collection and analysis. A related

[54] Hancock appears three times in the reference list, and also receives a personal acknowledgement at the end of the article. Pennebaker appears 11 times in the reference list, including his general-audience pop-psychology book, *The Secret Life of Pronouns* (2011), which is cited in-text four times.

[55] Markowitz's dissertation is titled "The Media Marshmallow Test: Psychological and Physiological Effects of Applying Self-Control to the Mobile Phone," and it does not appear to be related to the subject matter of the article in question.

and interesting point is that Markowitz reports that the rate of "power words" (e.g., "approve," "win") calculated by the LIWC software did not show a significant difference between actors and directors. He says this suggests that "the effects reported here are more consistent with status differences than power" (p. 385).

Regardless of the definition of the status variable, Markowitz found that the actors in his sample used a higher average rate of FPS pronouns than did the directors. The mean FPS rate for actors was 8.41%, and the mean FPS rate for directors was 6.81%. This represents a difference of 1.60 percentage points between the two groups. To put these numbers in perspective, we can calculate FPS pronouns used in a hypothetical 198-word speech (the mean speech length from this study) from each group; the mean of 8.41% for actors would represent 16.7 FPS pronouns used, and the mean of 6.81% for directors would represent 13.5 FPS pronouns used. This would be a difference of approximately 3 FPS pronouns within approximately 200 words.

The numbers presented in this article show a clear, though minimal, difference in average FPS usage rates in speeches by actors and directors. In my opinion, the question of which group is of higher social status, or even whether there is a clear difference in social status between them, is still up for debate.

Conclusion

This chapter has served to provide background information on both FPS pronouns and social status, as well as a foundational understanding of the literature relevant to connecting them. FPS pronoun use indicates self-construal, i.e., self-focused attention, and there is work that points to a connection between self-focused attention and social status. This dissertation is designed to further test the hypothesis that FPS pronoun usage rates may be an indicator of social status.

Social status is a multifaceted concept, and any attempt to measure it requires some amount of artificial delineation. As presented above, previous research has defined status variously in order to investigate the potential relationship between status and FPS pronouns. The most widely-cited work correlating status and FPS pronouns in English is Kacewicz et al. (2013), the relevant details of which are described above. I also included a critical overview of some of the methodological choices from that paper in order to provide context for presenting my own methodology in the following chapter. Other studies claiming the same pattern are also discussed above. This dissertation is the first work to consider a potential relationship between social status and FPS pronouns in ASL. In Chapter 3, I will describe the methodology for my study.

Chapter III: Methodology

On the surface, it may seem simple to ask about the relationship between FPS pronouns and social status. It should just be a matter of whether people with more status use fewer FPS pronouns, or more of them. Or whether there seems to be any pattern at all. While FPS pronouns are countable, and thus comparably straightforward, social status is a bit trickier to define, let alone to measure. This is due in part to the complex and ever-changing nature of social status. Status can be looked at both objectively and subjectively, and has no truly quantifiable definition. Additionally, social status is not absolute, but relative (see "Social status in general" in Chapter 2 for more). That said, my aim is to begin addressing whether a relationship exists between relative social status and FPS pronoun use in ASL, and I believe this is best done with both quantitative and qualitative approaches.

Here I will describe how I attempted to wrangle, and ultimately operationalize, the simple yet abstract potential connection between social status and FPS pronoun use. I will begin by discussing my recruitment strategies and the demographic information I used to group potential participants. Then I will describe the pool of participants in my data set, followed by descriptions of my filming procedures and annotation protocols. I close this chapter with a description of how the relevant variables are defined. With there being previous work to potentially build upon, namely Studies 1 and 3 from Kacewicz et al. (2013), I will also explain where and why I chose to deviate from those previous studies in my methodology, both in data collection and analyses.

Recruiting Participants

Because I aim to identify usage patterns in ASL, my goal was to include participants who reflect the diverse communities of deaf ASL users. This desire may seem obvious, but many linguistic studies seek to limit their participants to "native" users, defining native users as those who acquired the language before the age of five, or even from birth, potentially including a requirement that they grew up with parents who used the language natively. I agree with Hill (2014), who points out that "focusing only on native signers does not reflect the reality of language practices in the Deaf communities" (p. 203). Learning ASL after the age of five may well influence pronoun use.[56] However, limiting research to participants with early ASL exposure obscures the true nature of ASL as it exists in the real world.

That said, I used the word "native" in my recruiting (after acquiring IRB approval), in both the flyer and the questionnaire described below, because of its generally understood anecdotal implication of a deaf or hard-of-hearing person who is not a "new signer."[57] I wanted a wide range of representation across other social factors as well (e.g., race, age, education), so I got permission from various Facebook groups to post my call for participants on their pages, in addition to recruiting using paper flyers around Gallaudet University's campus and an e-flyer in Gallaudet's email newsletter (which at the time was called the Daily Digest). Figure 3.01 below shows the flyer used in all of these recruitment efforts.

---

[56] While I think this argument warrants further scrutiny, McKee et al. (2011) provide one take on how pronoun use varies in bilinguals: "Sorace et al. (2009) summarized research indicating that bilingual children who acquire a null subject language as a second language are significantly more likely to produce subject NPs in contexts in which native speaker monolinguals prefer a null subject. Similarly, the study of variable subjects in ASL by Wulf et al. (2002) found English influence in a clause to be the strongest constraint on variable subject presence in ASL" (McKee et al., 2011, p. 383). Reynolds (2016) found results similar to Sorace et al. (2009), but in "native-signing, bimodal bilinguals" in ASL.

[57] Upon reflection, I realize that using the word "native" may have discouraged some people from signing up to participate, even those who would have been perfectly qualified (e.g., a deaf person who grew up in a deaf family that frequently used spoken language in addition to signed, and who felt they did not find their deaf identity and/or learn what they think of as "real ASL" until taking ASL classes in high school or college).

Figure 3.01

Recruitment material

Background Questionnaire

Interested individuals were directed to complete a pre-enrollment questionnaire[58] that collected information to determine respondents' eligibility and availability. The questionnaire included both demographic information that allowed me to select a diverse group of participants and status-related information that I used to group participants for data collection sessions.

For most items on the questionnaire, I avoided multiple choice in favor of open-ended text boxes where respondents could answer how they saw fit.[59] This was in hopes of encouraging more detailed responses and allowing for greater nuance and more accurate self-identification

---

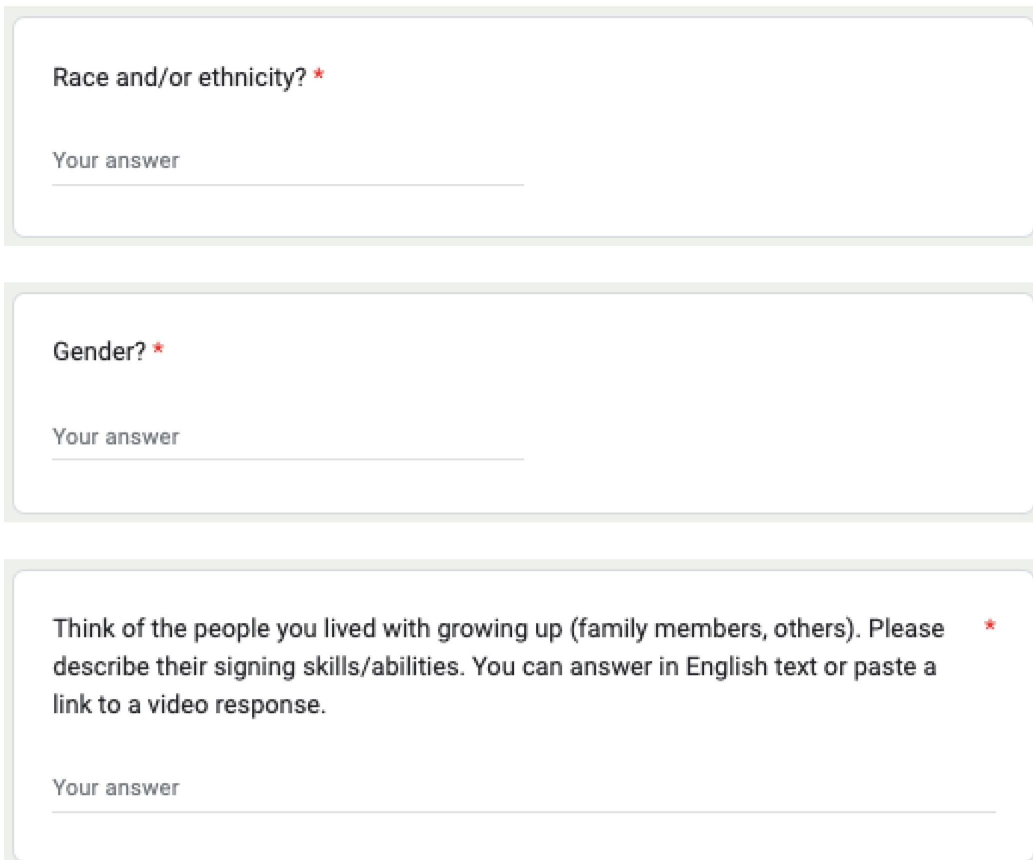
[58] A PDF of the survey can be found in the appendices as well as in my Figshare project folder (https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932), and is viewable in its original live form at tinyurl.com/ASL-Chat-Study.

[59] This was a good idea, and in retrospect, making all of the questionnaire items open-ended would have been an even better idea, though it may have decreased participation. See the "Participant Demographics" section below for examples of information that I learned from the chat content that I would have liked to have known during the participant grouping process.

than if I had asked for conformity to predetermined categories. The background information I collected included age, gender, race/ethnicity, age of ASL acquisition, K-12 school environment, formative home language, education level, employment, and states/countries lived in. Figure 3.02 shows a few examples.

Figure 3.02

Three open-ended questions from the initial Background Questionnaire


Race and/or ethnicity? *

Your answer

Gender? *

Your answer

Think of the people you lived with growing up (family members, others). Please describe their signing skills/abilities. You can answer in English text or paste a link to a video response. *

Your answer

A total of 42 people expressed interest in participating by responding to the questionnaire, and 24 people ended up being filmed. Some of this attrition was due to scheduling conflicts (e.g., people who were only available to film at times that no one else was available), and there were also some respondents who never replied to my request to schedule filming. Others confirmed but then cancelled, with or without advance notice. These were rescheduled when possible. One respondent was disqualified because they reported having very little experience with ASL.[60]

Grouping Participants

For data collection sessions, I scheduled multiple groups of three participants with the goal that their relative statuses be distinct: one relatively highest (H), one in-between (B), and one relatively lowest (L). From each triad, three dyads were filmed chatting, named for their social status composition: HL, HB, BL. This design was to allow the in-between (B) participant in each triad to fill two different status roles between their two chats, meaning the FPS pronoun rates of each B participant could be compared between their two chats.

It is worth emphasizing that these social status roles are always relative; one person might be relatively lowest in status within one triad, but the highest in another triad. For this reason, members of each triad are only considered ranked relative to each other, not to the members of other triads, regardless of their H/B/L designations. In order to create triads, I needed to rank participants relative to one another, despite knowing there is no objectively accurate way to quantify social status. Since social status itself can not be measured, social status markers can instead be selected to serve as collective proxies for social status. Ironically, this means that no study—this one or any other—can actually make direct claims about social status. We can only make claims about the markers used in a particular study. See both of the subsections under "Social Status" in Chapter 2 for more on the quantification of status and the "objective" nature of certain factors.

[60] This respondent reported acquiring ASL at 14, and was 19 at the time of the survey, giving them 5 years of ASL experience. In retrospect, the phrasing on my questionnaire may have been to blame in this instance; see earlier "Recruiting participants" section for a related discussion of the term "native."

As anyone who has coordinated research participants can imagine, there were filming sessions where not all three people showed up as planned, so some adjustments were made, which are explained in the Filming section below. Here I will describe the triad plan, since 12 of the 18 dyads I filmed and included in my study were part of a triad. The remaining 6 dyads followed the same procedures, but with only a single chat being filmed. The total of 18 dyads resulted in a total of 36 "participants" for analysis purposes, though there were only 24 unique individuals. From each triad, three dyads were filmed chatting, named for their social status composition: HL, HB, BL.

Since each chat had only two people, the chats had only two social status roles (in contrast to the triads having three): relatively higher and relatively lower, which I will refer to as *HighRole* and *LowRole*, respectively. See Figure 3.03 for a summary of the H/B/L designations and the in-chat roles.

Figure 3.03

Summary of H/B/L designations and in-chat roles

Participant designations	
H	**H**ighest predicted relative status within the triad
B	**B**etween the other two participants in the triad, in terms of predicted relative status
L	**L**owest predicted relative status within the triad

In-chat roles	
HighRole	Higher predicted status within the chat
LowRole	Lower predicted status within the chat

Recall that the reason for creating disparate triads in the first place, as opposed to dyads only, was to have the in-between person (B) be of relatively lower social status in one chat (*LowRole*) and of relatively higher social status (*HighRole*) in the other. I could then compare each B participant to themselves between their two chats, asking whether each B person used FPS pronouns at different relative rates when in *LowRole* versus *HighRole*. The person who was predicted to have the relatively highest social status of the triad (H) was also predicted to have the relatively higher social status in both of their chats (*HighRole*); thus, H participants were always *HighRole*. Similarly, the person with the predicted lowest social status of the triad (L) was *LowRole* in both of their chats; L was always *LowRole*. Within-participant FPS usage rates were also compared for H and L participants, though their relative rates were hypothesized to be different only in degree, rather than in direction. See Figure 3.04 for a visual representation of who filled which role during each of the three chat configurations.

Figure 3.04

Participant roles during chats, with emphasis on the in-between (B) participant's two different roles

Dyad	LowRole	HighRole
HL	L	H
HB	***B***	H
BL	L	***B***

When designing methods for collecting data that is potentially correlatable, it is tempting to create two quantitative variables. Of my two variables, only one is naturally quantitative: rates of FPS pronoun use are measured on a continuous scale where the distance between integers is even and meaningful, and there is a true value of zero. This makes rates of pronoun use a type of quantitative, ratio-level data. Admittedly, social status feels somewhat quantitative in the sense that we can describe someone as having “more” or “less” status than someone else, but any social status measurement system that was truly quantitative would be an exercise in false precision. That said, social status markers can be selected and measured in various practical ways for the purposes of finding patterns, which could then be used to create predictions to varying degrees. While it is more accurate to conceive of social status as ordinal, instead of quantitative, at least two difficulties remain for predictive purposes: (1) the factors that can potentially influence social status relationships are so multifaceted as to be nearly infinite, and (2) status relationships are not fixed; rather, they are co-created in an ongoing manner, changing in response to people’s behavior, changing as people learn relevant details about one another, and changing as different factors become more or less salient during an interaction. So rather than trying to predict “the” status relationship of a given dyad, a more useful question for grouping participants is: *Who is more likely to have higher/lower status more often in a given context?*

The given context for this study was casual chats between signing deaf Americans, filmed in the Linguistics department on Gallaudet University's campus. This context supports the use of social status factors that are likely to come up in a casual conversation and to be relevant among deaf Americans who have likely attended college. Thus, I included factors that align with the broader American population (i.e., age, education level, job), as well as factors reflecting American deaf community norms, meaning special attention was paid to linguistic background (i.e., type of K-12 environment, formative family language environment, age of ASL acquisition) and positions within the community (e.g., community-bound employment). Each factor I selected is described in its own section below. Each factor was considered to be ordinally ranked, even if roughly so, in a socially-constructed hierarchical sense. While all of the selected factors have the potential to impact status within American society at large, certain elements can be gauged somewhat differently in American deaf communities.

Below is a description of how each of the selected factors was "measured" in my study, both for grouping participants based on their responses to the initial Background Questionnaire, and for evaluating potential effects of additional relevant information. Such information included details that participants mentioned during their chats (e.g., behavior, updates to the original factors) and participant responses on the Post-Chat Survey. The relationship between each potential dyad was considered during the grouping process. Ideally each dyad would have clear status differences for each factor. For example, when compared to the *LowRole* participant, the *HighRole* participant would be older, their formative language experiences (i.e., family and school) would be more ASL-centric, and they would have a higher level of education, a more prestigious job, and a lower age of ASL acquisition. However, holding out for clear discrepancies in every category to determine who was “different enough” in status would not be realistic. Still, triads were only formed when there were clear discrepancies in multiple categories, and contradictions across categories were mitigated as much as possible. For example, a dyad with otherwise clear and consistent discrepancies (e.g., in the formative family environment, higher education, age of ASL acquisition, and employment) might still be included despite a minor contradiction. Examples of minor contradictions could include the *LowRole* participant being a year older than the *HighRole* participant, or the *LowRole* participant having attended a deaf school for a few years longer than the *HighRole* participant. For each section below, a different factor is described, beginning with the exact wording of the corresponding item from the Background Questionnaire.

Age at the time of filming.

Background Questionnaire item: "Age?"

This is the only selected social status factor that is truly quantitative. Put simply, greater age generally equals greater status, outside of potential extremes at either end of the age spectrum. However, this measurement may be less straightforward than one might think in the context of deaf communities. Anecdotally, there is a greater degree of multi-generational socializing among culturally deaf people than among the general American population. While I am not currently aware of any literature referencing this tendency directly, those with experience socializing in deaf and hearing communities may be able to attest. Such socialization patterns suggest that the effect of age on social status may be diminished. This is not to say, however, that age can have no potential effect on relative status. Thus, I included age in my grouping criteria, considering greater age as equating to greater status.

Level of higher education.

Background Questionnaire item: "What is the highest level of education you have completed?"

Higher education has contradictory indications for social status within deaf communities, making it potentially tricky to use as a marker of social status. On one hand, higher levels of education generally signal higher social status, in deaf communities or otherwise. On the other hand, higher education for deaf Americans is generally associated with a focus on English proficiency and more "English-like" signing. Higher education, thus, is considered by some to be a marker of ability and willingness to engage in the hearing world on hearing terms (e.g., using

English).[61] It is conceivable that this view may be somewhat tempered for those who attend Gallaudet, considering the cultural significance of the University (Friedner & Kusters, 2015). As my study took place on Gallaudet's campus, it came as no surprise that all of the participants in my study had some level of post-secondary education,[62] and so it did not seem likely that greater levels of education would be seen as lower status by the participants in my study. There is a clear, naturally-occurring ordinal rank associated with the various levels of higher education. This is not to say that a college freshman is somehow "lesser than" a graduate student or a professor with a doctoral degree, but that there is a relatively objective way to measure experience with higher education, and thus to measure status in that context. With this in mind, I treated a higher level of education as equating to greater status.

Employment.

Background Questionnaire item: "Where do you work and what is your current job title?"

There are various ways to conceive of social status in terms of employment. On the surface, someone with a high-prestige job, such as the CEO of a large and successful company, would have higher status than someone who is currently interviewing for an entry-level position at that same company. In the context of my study though, there are at least two other considerations. The first is how a lack of employment is viewed. Outside of the post-secondary context, not having a job is often looked down upon, with certain exceptions. In the context of a

---

[61] See Hill (2012) for a discussion of "grassroots" deaf Americans in contrast to deaf Americans who have a college education, and how signing is perceived in relation to higher education.

[62] The deaf population in America, indeed the general American population, does not have a 100% rate of higher education (Bloom et al., 2024). This is but one way that the sample in my study is not truly representative of the target population. See primary analysis limitations in Chapter 5 for more.

university, students who are not employed may actually be exhibiting greater privilege than those with jobs, since they appear to have access to sufficient means to afford their education and life expenses without needing to work. However, considering the wide availability of student loans, in addition to Vocational Rehabilitation services providing financial support for many deaf students (Gallaudet University, 2023), a Gallaudet student's personal or familial means are likely not easily correlated with employment status. In light of this, I did not consider employment itself, or lack thereof, as being a strong marker of status.

The other consideration specific to my study is that employment related to the deaf community (i.e., community-bound employment) is likely to hold greater prestige, meaning that someone who has an entry level position working for the National Association of the Deaf[63] may be seen as having higher employment-related status than someone who is a manager at a retail store. In such a scenario, the proximity to deafness may take precedence over the fact that management positions are generally considered higher-status than entry-level positions. Thus, I gauged employment as it would generally be gauged in American society at large, while considering community-bound employment as having greater sway. For example, I considered positions like "professional in the Center for Bilingual Teaching and Learning on Gallaudet's campus" to be of higher-status than positions like "student ambassador at Gallaudet's Visitor's Center," and I considered positions like "dorm assistant at a residential school for the Deaf" as being of higher-status than positions like "sales associate at Target."[64]

[63] https://www.nad.org/

[64] The four examples here are modified abstractions of actual responses I received.

Type of K-12 environment.

Background Questionnaire item: "What kind of K-12 school(s) did you attend?"

As described in Chapter 2, attending a deaf school is a marker of status. All of the other K-12 environments were ranked according to this benchmark. In theory, the highest status in terms of K-12 environment might be something along the lines of: "attended a Deaf school, and was mainstreamed to the local public school for select advanced courses." This would solidify the status conferred by the deaf school environment, as well as indicate that the individual was capable of handling more difficult content than their peers, even in the more challenging context of a public school, where they would need to access the content through interpreters or other non-traditional means.

Formative family language environment.

Background Questionnaire item: "Think of the people you lived with growing up (family members, others). Please describe their signing skills/abilities. You can answer in English text or paste a link to a video response."

As described in Chapter 2, ASL is a core component of American deaf communities. Correspondingly, early exposure to ASL (one component of more fluent ASL), signals greater status, but a simple distinction between "deaf families" and "hearing families" is insufficient to determine early ASL exposure. Similar to K-12 schooling environments, formative family language environments vary widely among audiologically deaf individuals. In theory, families with a multigenerational history of deaf community membership (e.g., "5th-generation Deaf") would confer the highest family-related status. However, this does not address perceived status differences for most deaf Americans, as the vast majority of audiologically deaf children grow up in hearing families.[65] Some hearing parents who have deaf children learn to sign, with varying degrees of immersion into ASL communities. Some hearing parents have multiple deaf children,[66] and some hearing siblings learn to sign along with their deaf siblings. The high value that American deaf communities place on ASL suggests that the factor of "family language environment" would likely be perceived along a continuum with "spoken language only" at the lower end, and variations of "regular use of ASL by skilled signers" at the higher end. This is the rough continuum I used when considering the factor of formative family language for the purpose of grouping participants.

Age of ASL acquisition.

Background Questionnaire item: "At what age did you first learn ASL?"

As previously discussed, and in oversimplified terms, earlier acquisition of ASL confers higher social status. However, age of acquisition (AoA) does not exist in isolation, as the previous two factors (K-12 environment and formative family language environment) largely determine a child's language acquisition. On the surface, a later age of acquisition may appear to indicate language deprivation, which can have lasting cognitive effects (see Hall et al., 2019 for a review), as well as effects on later language acquisition. A clear example of this would be a child who was audiologically deaf at birth who was not exposed to a signed language until adulthood. In the case of someone who was born hearing, and then became deaf in early or later childhood, their language acquisition likely followed the trajectory expected for a hearing child, until the onset of the hearing loss, at which point, the trajectory may have continued to be normal, simply

---

[65] See Mitchell & Karchmer (2004) and Lederberg et al. (2013) for estimates, and Lillo-Martin et al., 2023 for an overview of research related to deaf children in hearing families.

[66] See Gagne (2017) for discussion of the impact of linguistic peers on language development.

switching to a different language in a different modality (i.e., ASL or another signed language). On the other hand, their access to language may have been interrupted, disrupting their language acquisition and potentially their cognitive development.

Regardless of an individual's language acquisition experience, their age of acquisition for a given language is likely to have an effect on their abilities in that particular language, though their abilities in any L2 may be influenced by the age of their L1 acquisition. Later acquisition has measurable effects on an individual's language abilities (Mayberry & Kleunder, 2018; Hall et al., 2019), making age of ASL acquisition a factor with the potential to be directly visible, even if it is not explicitly recognized as "age of ASL acquisition." Rather than attributing "fluent" or "native-like" signing to being the result of "early ASL acquisition," it is often anecdotally attributed to the individual having grown up in a deaf family or having attended a deaf school. In reality, deaf schools and deaf families do not always prioritize ASL, especially with children who have some access to sound. Therefore "age of ASL acquisition" is a factor worth including, despite the potential for overlap with the factors of K-12 schooling environment and formative language environment. This redundancy within my selected factors could have become a problem (e.g., "double dipping") if I were quantifying these social status factors on a point system, with point values for each category being totaled to create numeric comparisons between people. Since I did not quantify these factors using an aggregated point system, the extra details from including K-12 schooling environment, formative family language environment, and age of ASL acquisition only served to provide more robust background information for me to work with while grouping participants.

To this point, I have discussed "ASL" acquisition, but there are many signing varieties used by deaf Americans, with at least twice as many corresponding language attitudes about them.[67] In light of this, asking "At what age did you first learn ASL?" on the Background Questionnaire was not ideal. In retrospect, I would have liked to have asked at what age they started signing. The question about K-12 education could then have elicited more details about the type(s) of signing they may have encountered at various ages, perhaps with a similar approach to the formative family language question above. Working with the data set as it was, I considered lower age of ASL acquisition as equating to greater status.

Other grouping considerations.

In addition to the factors described above, I took several other things into consideration when grouping participants. One was gender. Of the 42 people who expressed interest in participating, all self-identified as either female or male. Figure 3.05 below shows how each of the actual responses was categorized.

Figure 3.05

Responses from all prospective participants to the question "Gender?" on the Background Questionnaire

Responses categorized as "female"	Responses categorized as "male"
"F" "Female" "female" "Woman" "Lady"	"M" "m" "Male" "male"

I composed one triad of all female-identifying participants, one triad of all male-identifying participants, and one triad of deliberately mixed genders. All other triads were

[67] See Chapter 2 for a discussion of the work done on language attitudes in American deaf communities (Hill, 2012); See Ann (2001) for a discussion of language contact in deaf communities.

assigned without regard to gender. The four complete triads (i.e., when all 3 people showed up for filming) included 2 that were all female and 2 that were of mixed gender (in HBL order: MFF and MFM). The breakdown of gender configurations for the final 18 dyads was as follows:

Figure 3.06

Dyad makeup by gender, with the first letter signifying the gender of the person expected to have higher status at the time of grouping

Dyad-level gender configuration	Number of dyads
FF	8
MM	3
FM	3
MF	4
Total	**18**

Sexual orientation, race, and ethnicity were deliberately left out of these ranking predictions because of the variability of the impact of these factors. While context-dependent variability occurs with most, if not all, status indicators mentioned in this work (Tannen, 1989), the impact on status from sexual orientation, race, and ethnicity may be less predictable than most, making them suboptimal indicators for my initial placement calculations. On one hand, potential chat partners who share a minority identity in relation to one or more of these factors (e.g., both self-identify as Latinx or queer) may be more likely to perceive they are of comparable status despite other differences. That is, sharing a minority identity may override other differences that would otherwise have greater impact. This is similar to the earlier discussion of how deaf identity may override other disparate factors (e.g., “DEAF-SAME;” Friedner & Kusters, 2015; Ladd, 2003). On the other hand, there is also the possibility that those same two hypothetical participants with a shared minority identity may feel a stronger sense of social status disparity due to their expression of that identity (deliberate or not) or their relative roles within that minority community. For these reasons, I did not include sexual orientation in the Background Questionnaire. I did include a question about race/ethnicity, so that I could keep an eye on racial/ethnic diversity among the participants as a whole, but I did not consider race/ethnicity when I grouped participants into triads.

Another consideration was whether participants already knew each other. Ideally, participants would not have been acquainted prior to the chats, allowing dyads to co-construct their social status relationships from scratch over the course of the chats, rather than building on existing foundations. If chat partners were already acquainted prior to the chat sessions, they would have additional knowledge about each other that I would not have access to, potentially making it more difficult for me to identify their relative social status relationship (i.e., my independent variable).

That said, the collective American deaf community as a whole is small, and I was collecting data on Gallaudet's campus, so I knew there was a good chance some participants would already know each other. Still, the more I knew about what they knew about each other, the better I would be able to assess my independent variable of relative social status, so social media was used as a source of identifying obvious acquaintances. I discovered that two of my prospective participants were dating, so I made sure not to pair them as a dyad—not due to a lack of interest in how their FPS pronouns would pattern, but due to a lack of knowledge about the shared history (i.e., their social status relationship) that they would bring to the chat. I also considered contacting participants before assignment into dyads to see if, or how well, they knew their potential chat partners. But this would have likely led to participants looking up their potential chat partners on social media, defeating my purpose in asking. This added risk of participants learning potentially relevant status information about their chat partners did not seem to be worth the comparatively lower risk of inadvertently pairing people who already knew each other. Assuming I could not completely avoid pairing acquaintances, I captured information in the Post-Chat Survey about how well they knew each other before and after the chat, so I would at least be aware of when participants had already been acquainted. As it turned out, the chat content made it evident which dyads were already acquainted (e.g., greetings/introductions, discussion of previous encounters).

Participant Demographics

Figures 3.07–3.15 represent the 24 people who participated in my study, with each figure representing the entire pool of participants in terms of a particular demographic (e.g., gender, level of education). Half of these 24 people participated in two chats (as members of a complete triad, discussed in “Grouping Participants” above), meaning they contributed twice as much data as those who only participated in one chat. Regardless as to whether a participant had one or two chats, each participant is accounted for only once in the visualizations below, meaning that the visualizations total to 24 participants (rather than 18 chats x 2 people = 36 participants). The section reporting which states/countries participants had lived in includes a visualization with counts showing how many participants said they had lived in each location, so totaling the numbers does not produce a meaningful result.

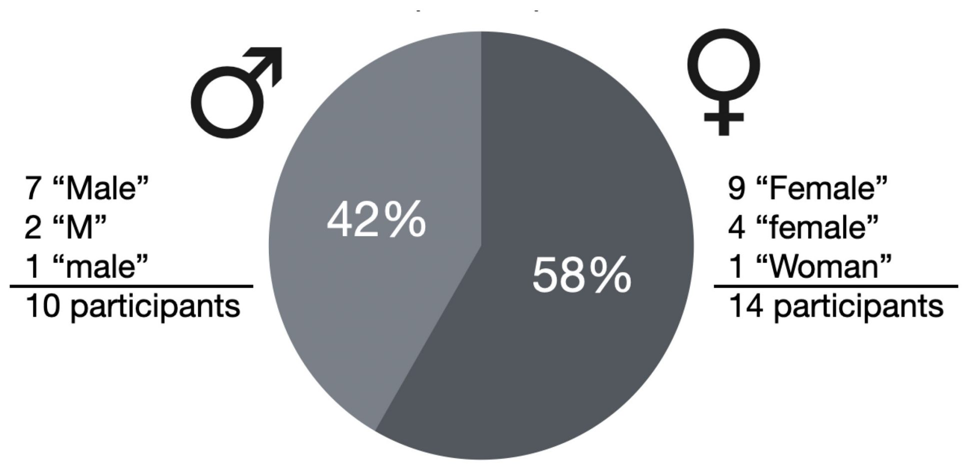
Each data visualization was designed with the goal of best representing the type and shape of the data for that demographic feature. This data is based primarily on participant responses to the initial Background Questionnaire that they completed to express interest in participating in the study. Any chat content that provided additional information (e.g., another state they lived in, a more detailed description of their parents' signing abilities) is also accounted for here, and discussed as needed. A brief description accompanies each figure.

"Gender?"

Figure 3.07 below shows participant responses to "Gender?" on the Background Questionnaire. Of the 24 participants, 10 identified as "male" (or some equivalent thereof) and 14 as "female" (or some equivalent thereof). Their actual responses are listed in quotes, with counts for each. None of the chats included any further detail or contradictory information.

Figure 3.07

Participant demographics: Gender

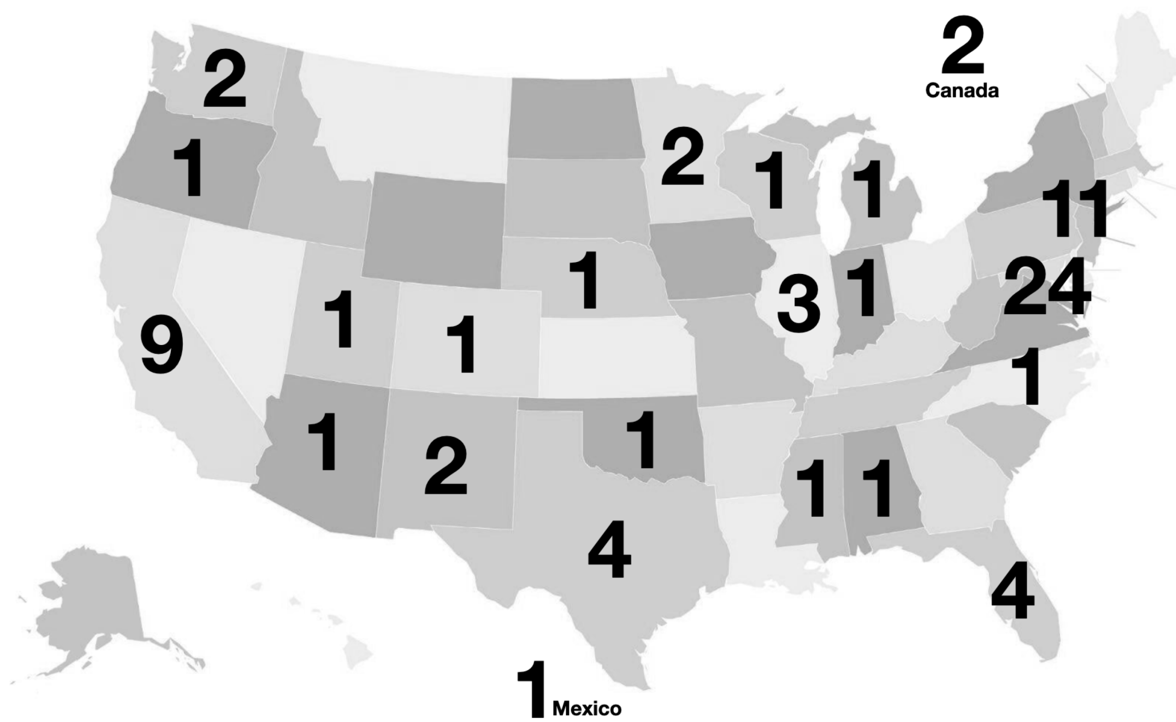


"What states/countries have you lived in?"

Figure 3.08 below shows participant responses to "What states/countries have you lived in?" on the Background Questionnaire. During the chats, many participants mentioned places they had lived, including some they had not listed in their questionnaires; all are included here. The purpose of this map is to provide a sense of the regional linguistic influences that may be present in this data set. All of the participants in my study had lived in the Washington DC metro area, either at the time of filming or at an earlier time. Most participants had lived in multiple states, and the distribution of their locations on the map in Figure 3.08 generally reflects population density across the country, in that more densely populated areas had greater representation in my study (e.g., nine participants had lived in California at some point, four had lived in Florida, one in Nebraska), though not every state was represented. In addition to the locations listed in Figure 3.08, one participant lived in Germany for a short time, and another lived in India for a short time.

Figure 3.08

Participant demographics: Places lived



"Where do you work and what is your current job title?"

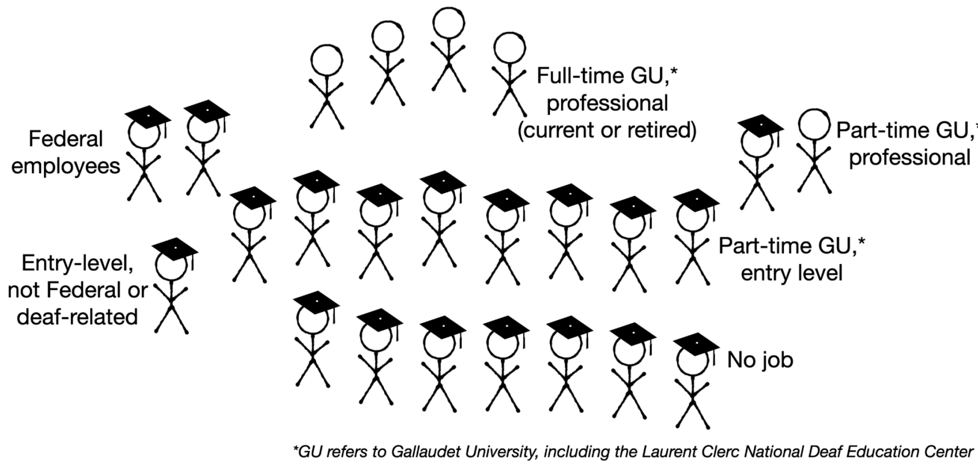
Figure 3.09 below shows participant responses to "Where do you work and what is your current job title?" on the Background Questionnaire. During the chats, many participants mentioned their current and former occupations, including some they had not listed in their questionnaires. None of the additional information would have affected their chat placements. The figure below represents the types of employment held by the participants at the time of filming.

The 24 participants are represented in Figure 3.09 in relative approximate hierarchical placement to one another. The within-level variation (i.e., each level is not perfectly aligned) comes from the hierarchical variation of their individual job titles. The category labels in the figure, which are intentionally broad in order to obscure potentially identifying participant information, are "no job" (seven participants), "entry-level, not federal or deaf-related" (one participant), "part-time GU, entry level" (eight participants), "part-time GU, professional" (two participants), "federal employees" (two participants), "full-time GU, professional (current or retired)" (four participants). Participants who were current students at the time of filming, either undergraduate or graduate, are represented with academic caps (19 participants). It is worth noting that most of the employment seen in Figure 3.09 is at Gallaudet University (GU),[68] and all current students were attending Gallaudet. Of the three participants employed elsewhere, two worked for the federal government. As my study was conducted on Gallaudet's campus in Washington, DC, this group-level employment profile is not surprising, but it is obviously not representative of the larger American deaf population.

[68] Gallaudet University (GU) includes the Laurent Clerc National Deaf Education Center, which houses K-12 programs for deaf and hard-of-hearing students: https://gallaudet.edu/clerc-center/overview/.

Figure 3.09

Participant demographics: Current occupation

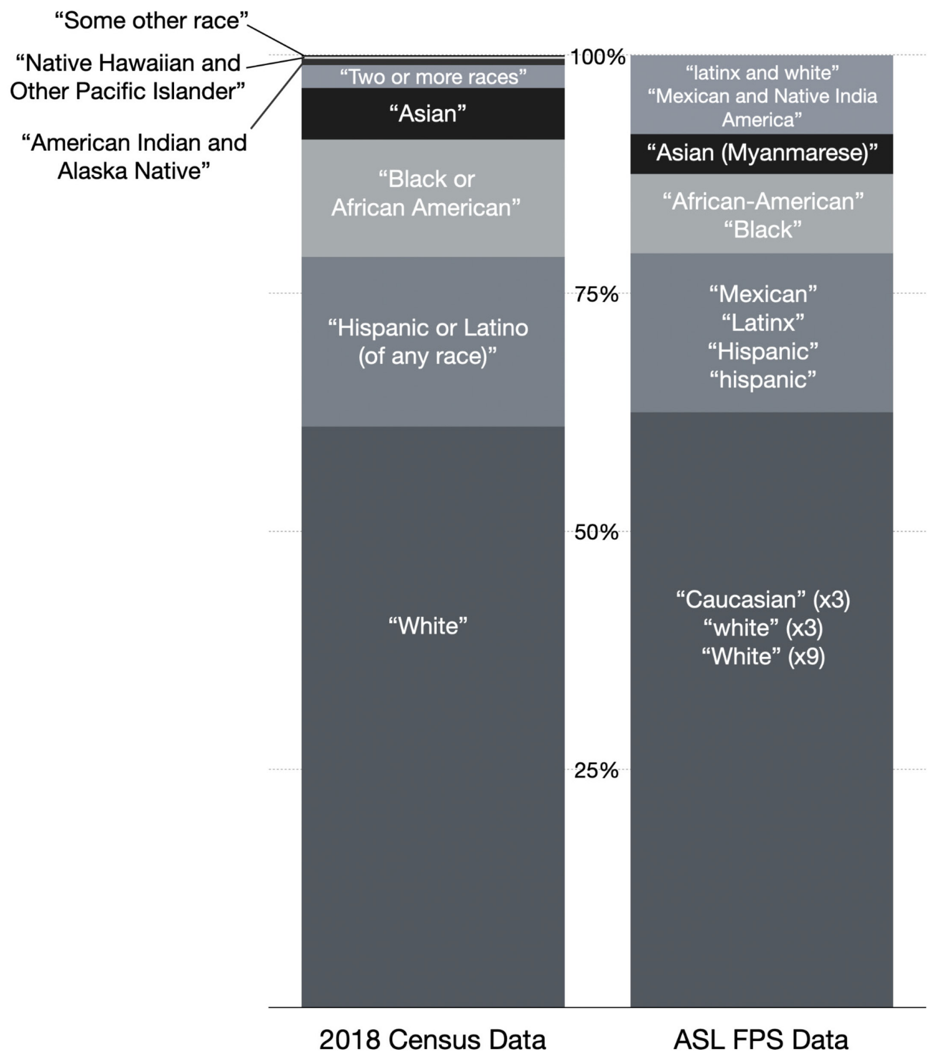


"Race and/or ethnicity?"

Figure 3.10 below shows participant responses to "Race and/or ethnicity?" on the Background Questionnaire, contextualized next to the U.S. Census data on race and ethnicity from 2018 (U.S. Census Bureau, 2019), when the chats were filmed. For the "2018 Census Data" bar, category names, in quotes, come from the Census data set. For the "ASL FPS Data" bar, actual participant responses are listed in quotes, with counts for responses that occurred more than once. From the 24 participants, 15 of the responses fall under the census category "White," four fall under the census category "Hispanic or Latino (of any race)," two fall under the census category "Black or African American," one falls under the census category "Asian," and two fall under the census category "Two or more races." None of the chats included any further detail or contradictory information.

Figure 3.10

Participant demographics: Race/ethnicity



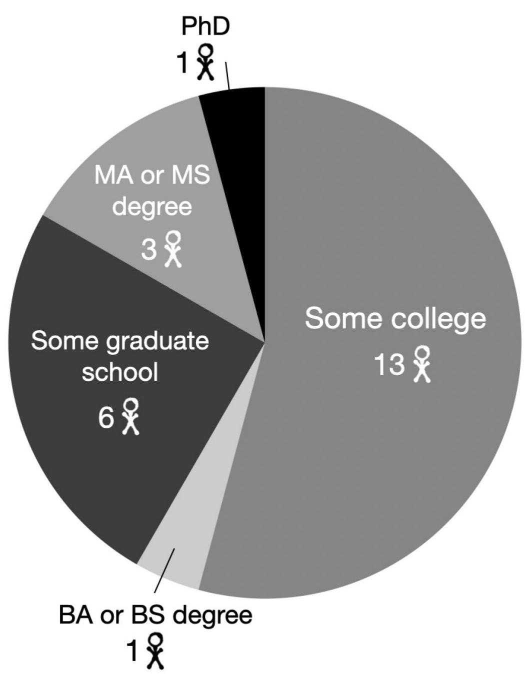
"What is the highest level of education you have completed?"

Figure 3.11 below shows participant responses to "What is the highest level of education you have completed?" on the Background Questionnaire, along with further detail provided during the chat sessions. Of the 24 participants, 13 had completed some college without yet receiving a bachelor's degree, one had completed a BA or BS degree (but no graduate school), six had completed some graduate school, three had completed an MA or MS degree, and one had a PhD. Given this topic came up in almost every chat, likely because filming took place on

Gallaudet University's campus, I was able to add "some graduate school" as a category even though this was not actually a provided category on the Background Questionnaire. Unfortunately, if I had included this category on the Background Questionnaire, there is one dyad who I may not have considered disparate enough to qualify for grouping together (more on this later in "Dyads warranting special consideration"). Any information from the chats that appeared to contradict Background Questionnaire responses (e.g., participant who told chat partners about being in an MA program, and chose "PhD" on the questionnaire) was cross-referenced with online sources (e.g., LinkedIn, current graduate student lists on Gallaudet's website) when publicly available. This figure represents the aggregate of all the information that was collected.

Figure 3.11

Participant demographics: Education

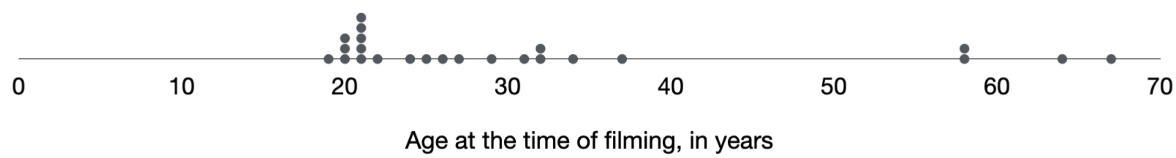


"Age?"

Figure 3.12 below shows participant responses to "Age?" on the Background Questionnaire. Each dot represents one participant. Their ages ranged from age 19 to 67, with about half being 20-30 years old. The majority of the participants (20 of 24) were between 19 and 37 years old, with eight of them being 20 or 21. The other four participants' ages were 58, 58, 64 and 67. None of the chats included any further detail or contradictory information.

Figure 3.12

Participant demographics: Age



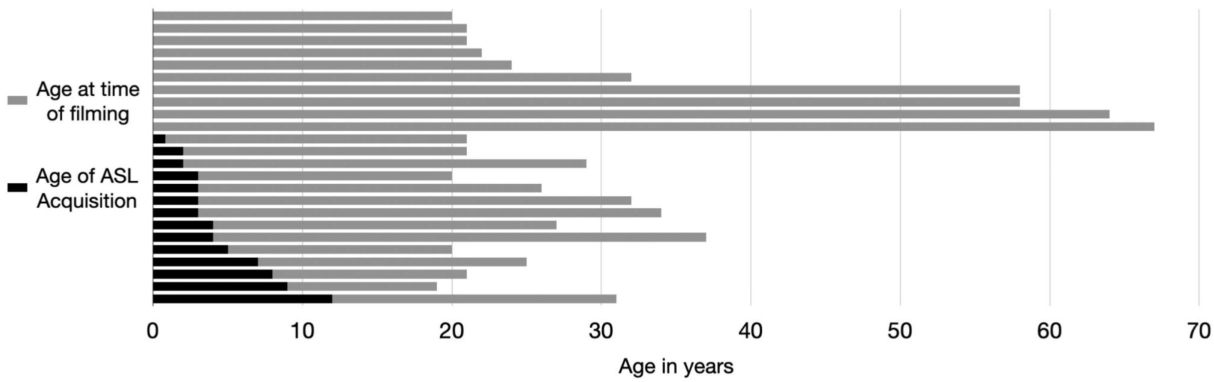
"At what age did you first learn ASL?"

Figure 3.13 below shows participant responses to "At what age did you first learn ASL?" on the Background Questionnaire (black sections) in the context of the participant's age at the time of filming (gray sections). Gray-only bars indicate that the participant was exposed to ASL from birth. The ages at which participants acquired ASL ranged from birth (age 0) to 13 years old. Two of the chats included additional details that indicated two participants had understood the question as referring to so-called "real ASL" as opposed to either (1) the signing they learned at home from their deaf parents, or (2) the signing done in a DHH pre-K program and a mainstream elementary school DHH program that used a language style the participant now felt was "more English-like." See the earlier sections, "Recruiting Participants" and "Age of ASL acquisition" to see why I included this range of participants, and to see how the corresponding

Background Questionnaire item could have been improved. Figure 3.13 reflects the aggregate of the available information.

Figure 3.13

Participant demographics: Age of ASL acquisition



"What kind of K-12 school(s) did you attend?"

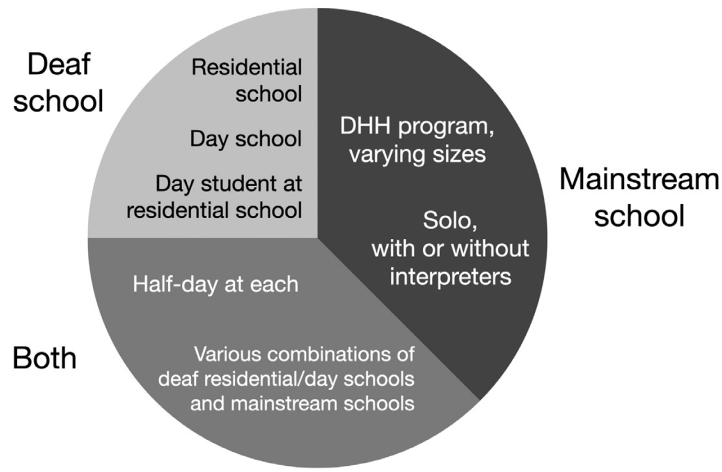
Figure 3.14 shows participant responses to "What kind of K-12 school(s) did you attend?" on the Background Questionnaire, along with further relevant details provided during the chat sessions. Of the 24 participants, six (25%) attended some type of deaf school, without any experience in a mainstream environment. The rest of the participants were evenly split between having attended only mainstream programs and attending some combination of deaf schools and mainstream schools. This topic came up in all but one chat (17 out of 18 chats), ranging from some participants mentioning it in passing, to some dyads discussing it at length. Most of the conversations about K-12 environments confirmed participant responses from the Background Questionnaire, often providing greater detail. There were some instances, however, when a participant's chat content did not line up exactly with their response on the Background Questionnaire. Some of those additional details may have affected the makeup of certain triads, had those details been available to me while grouping participants.

For example, the *LowRole* participant in one dyad selected “Mainstream school” on the Background Questionnaire, while the *HighRole* participant selected “Deaf residential school.” During their chat, they each brought up details that might have impacted their chat placement (more on this later in “Dyads warranting special consideration”). The *LowRole* participant described attending a deaf day school for kindergarten through 6th grade, which was a feeder school for the mainstream DHH program they attended for grades 7-12, including roughly 50 deaf and hard-of-hearing students and some deaf teachers. The *HighRole* participant described attending a mainstream school for kindergarten through 5th grade in a DHH program, then attending a deaf school for seven years before finishing their last two years of high school in a mainstream school. This participant repeated the same information to their other chat partner as well. These details make their K-12 environments much more comparable than they appeared from their selections on the Background Questionnaire. I might have learned some of this information, and similarly relevant information about other dyads, from the Background Questionnaire, had I made the relevant question an open-ended text box response, instead of a multiple choice question with an option to provide more detail within an “Other” selection.

Working within the confines of the information I had at the time, the survey responses from these two participants appear to justify assigning the *HighRole* and *LowRole* as I did. The chat content, however, suggests their K-12 environments were actually relatively comparable. In light of this information, Figure 3.14 includes both of these participants in the “Both” section at the bottom left of the pie chart. Figure 3.14 represents the aggregate of all the information that was collected from all participants.

Figure 3.14

Participant demographics: K-12 settings



"Think of the people you lived with growing up…"

Figure 3.15 shows participant responses to "Think of the people you lived with growing up (family members, others). Please describe their signing skills/abilities. You can answer in English text or paste a link to a video response"[69] on the Background Questionnaire, along with further detail provided during the chat sessions. In the figure, on the bar section representing the hearing families who used ASL, the "L2" designation refers to the fact that ASL was a second language. Of the 24 participants' responses, nine could be classified as some version of "deaf family," and fifteen could be categorized as some version of "hearing family." Though the

---

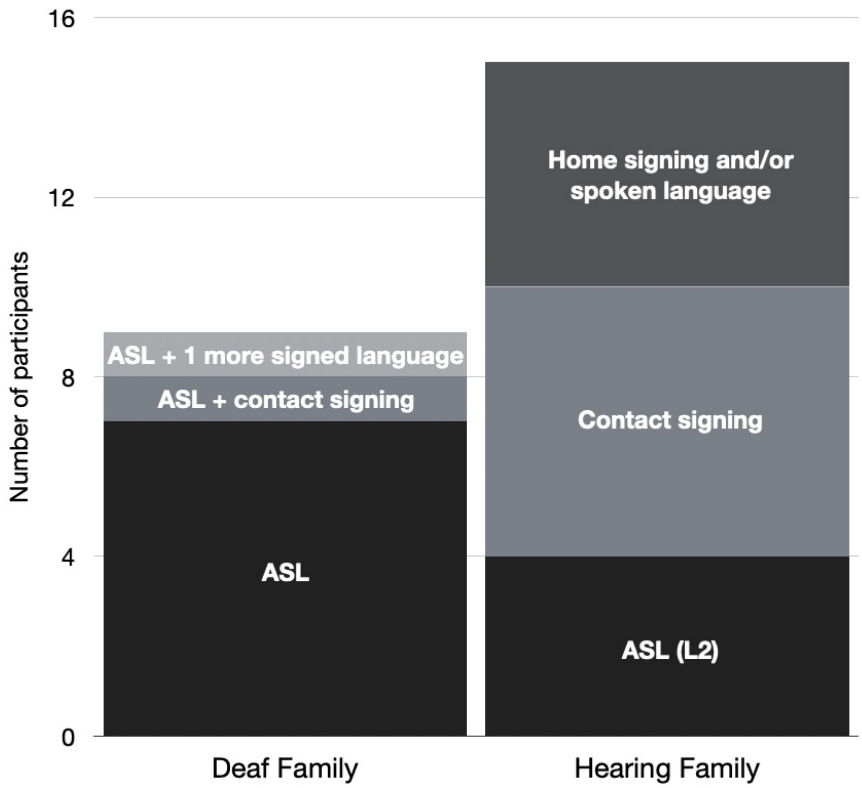
[69] Out of the 42 people who responded to the Background Questionnaire, none chose to respond in ASL. One potential reason for this is the fact that the platform I used, Google Forms, does not allow for recording video responses as a built-in option. A video response would thus require the respondent to record a video outside of Google Forms, upload it somewhere, generate a link, and come back to paste the link into the Google Form. These extra steps likely seemed to be more effort than they would be worth. I do think there is a good chance I could have gotten more detailed information up front if respondents had been given a simpler way to respond in ASL. Given the chance to do it all over again, I would look for a platform with integrated video recording capabilities.

Background Questionnaire and Figure 3.15 do not distinguish between parents and siblings, some participants mentioned the audiological status of their siblings: of the nine participants who came from deaf families, one person mentioned having a hearing sibling; of the 15 who came from hearing families, three mentioned having deaf siblings.

The topic of formative family language came up in all but two chats (16 out of 18), some dyads discussing it at length, and others mentioning it in passing. Most of the conversations about formative family language confirmed participant responses from the Background Questionnaire, often providing greater detail. There were some instances, however, when a participant's chat content did not line up exactly with their responses on the Background Questionnaire. Some of these instances may have impacted their chat placements (more on this later in "Dyads warranting special consideration"). For example, one participant's response to this item on the Background Questionnaire was "My family sign is slow and ASL." During one of their chats, this person explained to their chat partner that they had used "home signing" with their deaf younger sibling, but that their mother had only been convinced to start learning to sign when this participant was in high school. Another participant in a different group responded to this item with "I naturally alternate between ASL and PSE." During one of their chats, their chat partner asked if their parents signed, and this participant said no, before adding that their mom signed a little, but their dad did not. Looking only at the survey responses from the two participants described in these examples, it seems both of them might fit into the middle section of the Hearing Family bar in Figure 3.15, labeled "Contact signing." The chat content, however, suggests they are more appropriately included in the top section of the Hearing Family bar, labeled "Home signing and/or spoken language," which is where I have included them. Figure 3.15 represents the aggregate of all the information that was collected.

Figure 3.15

Participant demographics: Formative family language



Filming

As discussed above, my intention was to film three chats from each group, dividing the group of three people into three dyads: HL, HB, BL. Triads were never filmed as a group, but all three participants were asked to come in for the same time slot in order to minimize issues that could occur if each participant was asked to come at two separate times (e.g., participant missing their second chat time, different mood on different days). If one of the scheduled participants was unable to attend, and did not provide sufficient notice, I still filmed the other two as a single dyad. This was not only out of respect for their time (and I could then pay them the promised stipend), but it was also a practical decision. This decision became even more obvious after the first few chat sessions. The information I had to work with in determining relative status was limited by what I had chosen to include in the initial Background Questionnaire, as well as how participants had chosen to respond. Because of these limitations, I felt fairly certain there would be some instances where the status predictions did not play out as expected. After the first few sessions, I found that not only did the chat content provide additional insights that would have altered some of the original chat placements (as discussed in the "Participant Demographics" sections above), but those first few Post-Chat Survey submissions also confirmed that subjective responses were not going to be a good way to measure status discrepancies (as discussed in Chapter 4). In other words, things got messy. It would have been nice to film three clearly differentiated participants participating in two chats apiece. But with no guarantee that I had accurately grouped participants in the first place, I continued to film single dyads when only two of the three scheduled participants showed up, rather than pass on a perfectly good dyad. When only one of the three scheduled participants showed up, sessions were canceled and rescheduled when possible. I continued to reschedule prospective participants until their schedules no longer permitted, they stopped responding, or their relative objective status discrepancies no longer met the study criteria.

Chats were filmed in a closed studio with two chairs and three cameras positioned at different angles. The two chairs were angled around a low table with the wheels on the chairs taped to prevent rolling, minimizing participants inadvertently changing the camera angles. The table was there to help anchor the chairs at the slightly turned-out angle that worked best for filming, and a low table was chosen to prevent participants from resting their arms on it. See Figure 3.16 for layout.

Figure 3.16

Filming layout

The only exception to this setup was due to one *HighRole* participant who wanted to participate remotely, asking if they could join via Skype. I decided to allow this as an exception. I asked that participant's *LowRole* chat partner to come to the studio, and I connected the two of them on Skype from there. I screen-recorded the virtual session, and filmed in the studio from two angles. Otherwise, I followed the same procedures as with the other dyads. See the "Dyads warranting special consideration" section for the implications of this exception.

When participants arrived, I kept them mostly separate until they entered the studio for filming, and I asked them to refrain from chatting or introducing themselves until I started the cameras and left the room. Participants were then filmed for 20 minutes of free, unstructured chatting. I considered attempting to keep the chat content somewhat comparable across dyads by using various prompts or activities, which could also stimulate conversation, but I ultimately decided against it. I came to realize stimulating conversation was a nonissue. In my experience, people who have volunteered to be in a research study tend to be willing to follow researcher instructions, and my instructions—to chat freely about whatever they liked—were not difficult to follow. Because of this, I did not anticipate any issues with participants being willing to engage with one another. But, although periodic short lulls are a natural part of conversational flow, I wondered if participants might run out of things to talk about. Morris (2016) shared this concern for her study, in which she also asked dyads of deaf Americans to chat freely. She provided the game *Guess Who?* as a relatively hands-free activity for participants to engage in if they chose, or if they ran out of things to talk about. Out of the ten 40-minute sessions in her study, no one used the game. I opted not to provide an activity, and things turned out similarly for me: none of the dyads in my study had more than a few seconds at a time without conversation. While my annotating-self often wished for more lulls, my analysis-self was thrilled. Additionally, the short lulls that did occur provided interesting insight once I began to see which participants would tend to take the lead in those moments. See "Behavioral considerations" in Chapter 5 for more on this.

The idea of keeping the chat content somewhat comparable is appealing on the surface, as data being collected under more similar conditions suggests more reliable results, but the drawbacks significantly dampen the appeal. First and foremost, imposing conversational topics or tasks would restrict how participants conversed. This conflicted with my desire to allow participants to co-construct their own interactions to the greatest degree possible, in order to see what patterns could be found in how they intuitively interacted. Restricting their interactions in this way would also add further unnaturalness to the already somewhat artificial condition of being asked to converse freely while being filmed from three angles in an unfamiliar studio space. Additionally, narrowing the topics of conversation would narrow the generalizability of the results. Because of these potential drawbacks and the apparent lack of significant benefits,[70] I decided against assigning tasks.

---

[70] Had I used the results from the five studies reported in Kacewicz et al. (2013) to help with this decision, I might have concluded that assigning a task potentially allows greater discrepancies to emerge in relative FPS pronoun use. However, it would have been prudent to reanalyze their original data before drawing any such conclusions, given the issues with

Once a dyad was in the filming space, the cameras were on, and instructions were given, I left the room and closed the door. And wished I had thought to set up a monitor to watch them on. At the end of the 20 minutes, I came back into the room to tell them that the time was up, and I turned off the cameras.

Post-Chat Survey

Once the cameras were off, participants were asked to complete the Post-Chat Survey in separate rooms. Because I did not provide participants with any background information about their chat partners, each person was only privy to the other person's status-related information that (1) they knew prior to the study, or (2) they learned during the chat, either explicitly shared or gleaned from appearance or behavior. The goals of the Post-Chat Survey, then, were to find out if (or how well) participants were already acquainted, and to elicit each participant's subjective experience of the dynamic between themselves and their chat partner (i.e., their perceived social status relationship).

Participant responses to the Post-Chat Survey have been taken into consideration for some of the analyses discussed in Chapter 4, but are not intended to replace the objective measures for two reasons: (1) the potential effects of social expectations for politeness discussed earlier, and (2) this work is most focused on identifying potential patterns that can be used to predict behavior, which come from social status factors comprised of more objective measures and directly observable behaviors.

When designing the Post-Chat Survey, I wanted to work from the Interaction Rating Questionnaire (IRQ) used by Kacewicz et al. (2013) in Studies 2 and 3 to elicit participants'

---

these studies that I described in Chapter 2 (and further issues with their Study 3 that I describe in this and the following chapters). That said, such reanalyses would not have been guaranteed to produce any insight as to task assignment.

subjective takes on their experiences during the chats, so I requested the document from the senior author, Dr. James Pennebaker. There were some inconsistencies between the IRQ questions described in their paper and the IRQ questions I received, so I circled back, and received a slightly different version of the IRQ. Both of the IRQs that I received can be found in the appendices and my shared project folder.[71] One of the inconsistencies had to do with the wording of the items that were used to measure participants' status. These two items were listed in their paper as "To what degree did you control the conversation?" and "To what degree did you have power in the conversation?" where 1 = *not at all* and 7 = *a great deal* (p. 7).

The first of these questions was identical in the paper and both of the IRQs I received. However, the second question was phrased slightly differently in their paper than in either of the two IRQs. The IRQs included this question as "To what degree did you have more status than your partner in the conversation?" in one, and "To what degree did you have more status or power than your partner in the conversation?" in the other. Without any further information, it is unclear which version of the second question was answered by participants.

Additionally, the phrasing of the first question does not explicitly ask participants to compare themselves to their chat partners. This makes it unclear whether participants viewed power as zero-sum (i.e., power as relative) or variable-sum (i.e., power as akin to confidence or competence) when responding. See "Social status in general" in Chapter 2 for more on this distinction. This uncertainty presents a problem for determining how to calculate dyad-level discrepancies. This problem also exists for the second question if it was presented to participants as phrased in the Kacewicz et al. (2013) paper, but not if it was phrased as on either of the IRQs that were shared with me. For a discussion of why and how zero-sum discrepancies should be

[71] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

calculated differently from variable-sum discrepancies, see the section titled “Methodological consideration from Kacewicz et al. (2013), Study 3: Participant agreement” later in this chapter. To avoid any potential ambiguity around whether questions were intended as zero-sum or variable-sum, I phrased the items in the Post-Chat Survey for my own study to explicitly ask participants to compare themselves with their chat partners.

Modifications for my Post-Chat Survey continued with rephrasing questions into plain language, as is recommended in the literature on survey development (e.g., Fowler, 1995; Tourangeau et al., 2000; DeVellis, 2003). Most of the questions from the original survey used a 7-point Likert scale, and I kept that format for those questions in order to maintain some level of comparability with the methods used by Kacewicz et al. (2013). I also added questions, including asking about how they and their chat partner might hypothetically score on the American Sign Language Proficiency Interview (ASLPI) (Gallaudet University, 2024), and about the relative status of their chat partner and themselves in various contexts. Most of the items I added asked participants to compare themselves with their chat partner.

A copy of the Post-Chat Survey I used can be found in the Appendix. There were a total of 23 items, designed to glean each participant’s subjective experience of the chat as a whole. Five of these items were used in my conceptual replications, described later in this chapter. These included questions about how much relative control and relative status participants felt they had during the chat, and questions about their relative rankings in relevant contexts (i.e., ASLPI, “the Deaf community,” and Gallaudet University). These questions and their response choices can be seen later in this chapter, in the section titled “Conceptual replication: Using my data, applying the refined approach from above.”

Annotation

The video data collected during the chats was annotated in ELAN,[72] a program that allows video data, even multiple videos, to be time-synced to annotation tiers containing whatever types of annotations are needed for a particular project. I started with each participant in their own file, but later switched to having both participants in a single file when I realized there were data processing advantages to doing so, along with the convenience of having all of the annotations for a single chat together in one file. Early in the annotation process, I did a reliability check with my advisor to make sure I was maintaining consistency in how I determined turn boundaries and potentially ambiguous FPS pronouns (both issues discussed in the sections below). As mentioned in "Positionality" in Chapter 1, two trained research assistants assisted with one chat each, and another did a reliability check on one of my status indicators passes.

For each chat, I time-synced the three video angles in the ELAN file. Then I marked when I left the room and when I returned, using the sound of the studio door opening and closing to guide this timing. These became the bounds of what I annotated, as I did not want my data to include any stray participant interactions that had happened in front of me. I was not only "the researcher," but also a hearing signer, meaning they would be more likely to unconsciously alter their signing style to adapt to a mostly-unknown hearing person in the room. In fact, one participant explicitly asked me, before filming began, about whether they should slow down or otherwise alter their signing style to make it easier for me to understand and analyze later. I assured this participant that they could just sign how they normally signed, because I was

---

[72] ELAN (Version 5.8) [Computer software]. 2019, October 09. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan.

interested in how they normally signed with other deaf people, not how they signed with hearing people like me. Interestingly, this was an H participant (i.e., I expected them to be of highest relative status in the triad), so I was happy to defer to their apparent assumptions about our relative language proficiencies.

Before annotating for FPS pronouns and turns, I began by simply watching for content. I then did multiple passes, annotating for a different facet each time to keep my focus narrow in an attempt to reduce errors. The sections below describe each of the facets I annotated for during these separate passes, including a set of tiers for each participant to capture their FPS pronouns and turns, and dyad-level tiers to capture status indicators. I also created tiers to capture observations on phonological forms, to record notes about chat content to aid searchability, and to document other notes as needed. Before describing these, however, I will first describe how and why I split the data set into two rounds of annotation.

A motivated division of my data set.

In this section, I will describe how and why I divided the data set into two subsets, creating two rounds of annotation and analysis. I am taking the time to detail this process because, while these decisions resulted in patterns that differed from those of the complete data set, they also ended up providing useful insights and supported my plan to reanalyze my data using some of the methods from Kacewicz et al. (2013).

During the semester when I collected data, I had a tight deadline for presenting my preliminary findings. Knowing how time-consuming annotation is, I decided to split the data set in half, saving half of the annotation for after my initial presentation. Though I planned to include all 18 dyads in my final analysis, I wanted to prioritize analyzing complete triads. This is because complete triads include the in-between (B) participants holding different roles in their two chats (i.e., they were *LowRole* in their chat with a high (H) participant and *HighRole* in their chat with a low (L) participant), allowing for intra-participant analysis. Additionally, I was most eager to see the relative FPS rates from a few specific chats.[73] What follows is a description of the factors that influenced my selections for which dyads to annotate in Round 1, which would become the data set for my preliminary findings.

As discussed throughout this dissertation, status relationships are not tidy nor fully predictable. So from the beginning of this work, I knew that there was variability in how disparate the dyads would be, even in the objective "on paper" sense. Then through the data collection process, I sensed some dyads had been more or less disparate in status than I had predicted. Some of this sense came from my interactions with participants leading up to and on filming days. For example, I was taken by surprise when one participant unexpectedly reached out to me on the video messaging app Marco Polo to offer advice on my project. There was another instance when, upon meeting the two participants scheduled for a chat session, I immediately felt that I had assigned the status designations backwards (i.e., I felt the person assigned to *LowRole* should have been assigned to *HighRole*, and vice versa). This dyad is discussed in Chapter 5. I also observed interactions between some participants after the filming sessions that provided clues to perceived status dynamics. All of these observations were captured in my session notes, discussed later in the "Session notes" section.

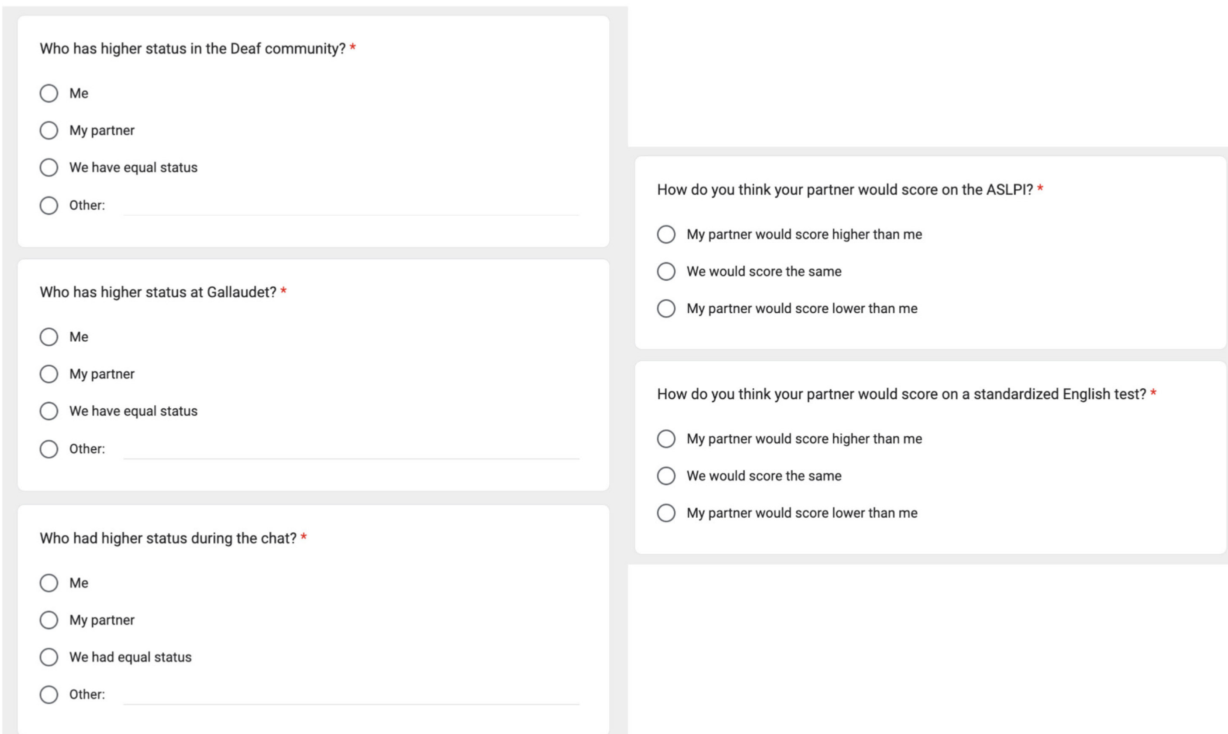
My initial review of the Post-Chat Surveys also provided some clues as to which dyads might have perceived a larger status discrepancy. Despite the overwhelming sense of equality imparted by participant responses as a whole, some response patterns did suggest perceived

[73] It did not occur to me until well after the fact that this division of the data set was likely to yield two different patterns. Such is the eager ignorance of a novice researcher.

status discrepancies in certain dyads, many in the predicted direction. I used five of the Post-Chat Survey questions to gauge participants' subjective experiences of their social status relationships. These five questions all asked participants to compare themselves to their chat partner, either in a particular social status context or in terms of hypothetical language proficiency test scores. Figure 3.17 shows the five Post-Chat Survey questions I used for this division.

Figure 3.17

The five Post-Chat Survey items used to divide the data set into two rounds of annotation


Who has higher status in the Deaf community? *

- Me
- My partner
- We have equal status
- Other:

Who has higher status at Gallaudet? *

- Me
- My partner
- We have equal status
- Other:

Who had higher status during the chat? *

- Me
- My partner
- We had equal status
- Other:

How do you think your partner would score on the ASLPI? *

- My partner would score higher than me
- We would score the same
- My partner would score lower than me

How do you think your partner would score on a standardized English test? *

- My partner would score higher than me
- We would score the same
- My partner would score lower than me

As I worked through dividing the data set, I began to wonder if there might be a different subset of questions that could be more reasonably expected to provide clues as to participants' subjective sense of their status dynamics. I pursued this line of thinking to create a different subset of questions for determining subjective status discrepancies once it came time to calculate (i.e., quantify) subjective status discrepancies in the conceptual replication of Study 3 from

Kacewicz et al. (2013). I discuss those decisions in their corresponding section, but the goal of the current section is simply to describe how I divided the data set into two rounds of annotation.

Using the questions in Figure 3.17, I calculated scores for each dyad according to how their responses aligned with the status prediction I had made during the participant grouping process. Individual participant responses were coded as one of the following:

- ***Fit*** the prediction (i.e., the participant's response aligned with my prediction about their status relative to their chat partner's)
- ***Contradicted*** the prediction (e.g., I predicted that the participant would be of higher status than their chat partner, but the participant's response indicated that they perceived themselves as being of lower status)
- Reported being of ***Equal*** status

For each of the five questions in Figure 3.17, I counted the number of participant responses that were categorized as ***Fit***, ***Contradicted***, or ***Equal.*** I then used the results of this coding, which can be seen in Figure 3.18 below (and also in my shared Figshare project folder[74]), to determine which dyads I would annotate during my first round of annotation.

For the language questions, each dyad had four responses: two participants each responding to two questions. These language questions were considered in the selection of dyads for first-round annotation, but the status questions took precedence.

For the status questions, each dyad had six responses: two participants each responding to three questions. The two dyads who fit the prediction most often on these three status questions, both with five points, were 4HL and 10HL. I included both in my first round, and because 10HL was part of a complete triad, I also included the other two dyads from Group 10. The next-

[74] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

Figure 3.18

Results from the initial coding used to determine which dyads would be included in first-round annotation

Group	Dyad	3 status Qs, out of 6pts			2 language Qs, out of 4pts			Why was this dyad chosen for first-round annotation?
		Fit	Contradict	Equal	Fit	Contradict	Equal	
1	**HL**	**4****	**0**	**2**	**2**	**0**	**2**	High score for "Fit"
	HB	**1**	**0**	**5****	**1**	**1**	**2**	H from this group displayed some behaviors indicating a high-status mindset
	BL	**2**	**0**	**4****	**1**	**0**	**3**	Prioritizing complete triads; this group already had 2 dyads selected
2	HL	3	2	1	1	1	2	
3	HL	2	0	4**	1	0	3	
	HB	3	0	3	3	0	1	
	BL	**0**	**0**	**6****	**0**	**2**	**2**	Only dyad in perfect agreement about being equal on the status Qs
4	**HL**	**5****	**0**	**1**	**2**	**0**	**2**	Tied (with 10HL) for highest score for "Fit" on status Qs
5	**HB**	**0**	**4****	**2**	**0**	**0**	**4****	Most extreme contradiction of the predicted status relationship; this tracked with my impressions
6	BL	3	0	3	0	0	4**	
7	HB	0	1	4**	1	1	2	
8	HL	2	0	4**	1	0	3	
	HB	1	0	5**	1	0	3	
	BL	4**	0	2	1	0	3	
9	HL	2	0	4**	2	0	2	
10	**HL**	**5****	**0**	**1**	**2**	**0**	**2**	HL tied (with 4HL) for highest "Fit" score on status Qs, and was the only top-scoring dyad to be part of a complete triad. Additionally, this group seemed to fit the prediction better than the other complete triads, overall.
	HB	**2**	**1**	**3**	**2**	**0**	**2**	
	BL	**2**	**1**	**3**	**1**	**1**	**2**	

Key
** Values greater than or equal to 4
Bold dyads: Selected for first-round annotation
H: Participant who was predicted to be of relatively highest status in their triad
B: Participant who was predicted to rank between the H and L participants in their triad
L: Participant was were predicted to be of relatively lowest status in their triad
Fit: How many times participant responses "fit" the relative status prediction
Contradict: How many times participant responses "contradicted" the relative status prediction
Equal: How many times participants responded that they were equal

highest scoring dyads, in terms of fitting the prediction for the status questions, with four points each, were 1HL and 8BL. With only five slots remaining in this first-half round of annotation, I could not include all dyads from both Group 1 and Group 8. The H participant in Group 1 had displayed some behaviors outside of the chats that indicated they felt themselves to be of high status, so I prioritized the three dyads of Group 1 over those from Group 8 for my first round of annotation. This left two slots, which I chose to fill with a dyad who had the highest rate of disagreement with my predictions (5HB) and the only dyad who responded with complete agreement that they had been of equal status within the three status-related questions (3BL).

Splitting the data set this way resulted in a pattern that appeared to generally support, on the surface, the notion of an inverse relationship between FPS pronoun use and status. However, as noted above, two of the selected dyads were selected specifically because they did not appear to fit my predictions about participants' relative status (5HB and 3BL). Additionally, as can be seen in Figure 3.18, there were other selected dyads (e.g., 1HB, 10BL) who also had relatively low Fit scores. These preliminary results provided useful insights about the nature of the data set and further justified reanalyzing this data using methods modified from Kacewicz et al. (2013).

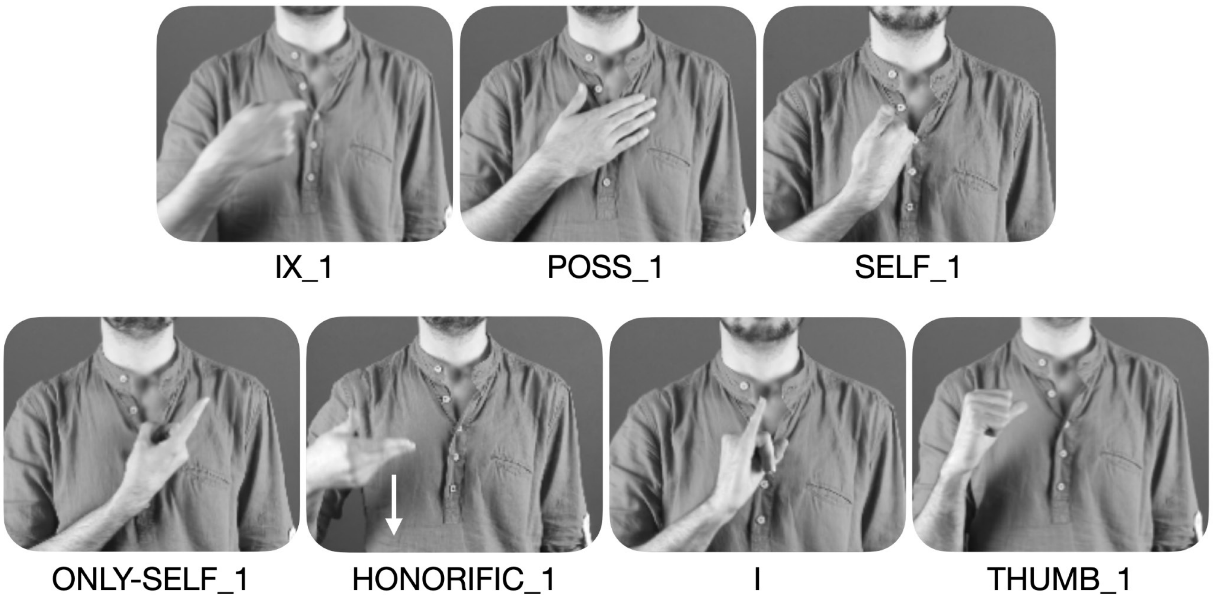
FPS tiers.

Each participant's FPS pronouns were annotated on one of three separate tiers, depending which hand(s) each instance was produced with: left, right, or both. The files were linked to ASL Signbank (Hochgesang et al., 2024) through an external controlled vocabulary (ECV) so that annotations could be selected from the ID glosses available through ASL Signbank, reducing human error, and thus improving the accuracy of FPS counts. Total FPS pronoun use for each participant was then obtained from ELAN. The FPS pronouns I annotated for are seen in Figure 2.01 from Chapter 2, repeated here as Figure 3.19 for convenience.[75]

[75] Videos of these seven pronouns are available in my Figshare project folder (https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932) and in ASL Signbank. As a reminder, the ASL Signbank entries for the FPS pronouns are listed with the first instance of this image, as Figure 2.01.

Figure 3.19

First-Person Singular (FPS) Pronouns in ASL



Only instances of the seven words shown in Figure 3.19 were annotated as FPS pronouns. However, as discussed in Chapter 2, there are multiple ways to convey FPS construal without the use of an FPS pronoun. As the previous work done on FPS and status includes only FPS pronouns, and not FPS construal without the use of pronouns, I do the same in this work. Figure 3.20 below shows three examples from my data set of FPS construal without the use of pronouns. The first image is a lexicalized fingerspelling of "own" produced with the final letter ending on the chest, which can be translated as "my own." The second is the word TURN[76] directed towards the signer's body ending with the base of the thumb on the chest, which can be translated as "my turn." The third is the word TEACH[77] produced as an indicating verb, which

[76] TURN (uninflected, so not incorporating first person):
https://aslsignbank.haskins.yale.edu/dictionary/gloss/308

[77] TEACH (uninflected, so not incorporating first person):
https://aslsignbank.haskins.yale.edu/dictionary/gloss/291

can be translated in this instance as “I taught her.” These and similar instances that convey FPS construal were not coded as FPS pronouns in my work.

Figure 3.20

Examples from the data of FPS construal without the use of a pronoun

As with all high-frequency words, FPS pronouns often exhibit phonological variation, including not always being fully phonologically produced (see “Phonological variation of FPS in ASL” in Chapter 2 and Börstell et al., 2024). This makes annotation tricky. At times, the imperfect frame rate of the recordings (approximately 30 frames per second) made it difficult to determine whether an FPS pronoun had actually been produced. For example, an FPS pronoun might have hypothetically been warranted, grammatically speaking, though without having been realized. In other cases, it is clear a word has been produced, but the phonological form is insufficient to identify which word it is. In these cases, the context sometimes provides enough information to identify the appropriate annotation label. This is the case in the two examples from my data seen in Figure 3.21 below, both of which appear to be the result of phonological assimilation. The first image in Figure 3.21 is an instance of an FPS pronoun that appears to have been influenced by the handshape of the following form. The word preceding the FPS pronoun is

RIGHT[78] and the word after it is a name sign with a K/P handshape (as seen on the signer's right hand). This instance could be interpreted as "Right, I am affiliated with [name sign]." The second is an instance of an FPS pronoun that was preceded and followed by words with similar handshapes, a version of WHY[79] and a version of PREFER.[80] This instance can be translated as "why I prefer." Again, in both instances, context was sufficient for deciding the appropriate annotation labels.

Figure 3.21

Examples from the data of IX_1 with phonological variation where the target word is identifiable based on context

On the other hand, sometimes the context is not sufficient to determine the word being produced. This is the case in the two examples seen in Figure 3.22 below. In the first image, the

[78] RIGHT: https://aslsignbank.haskins.yale.edu/dictionary/gloss/533

[79] BECAUSE-WHY8: https://aslsignbank.haskins.yale.edu/dictionary/gloss/36

[80] (prefer) FAVORITE: https://aslsignbank.haskins.yale.edu/dictionary/gloss/225

word was produced as part of a false start, and could arguably have been an instance of HAVE[81] or IX_1, based on context. The form in the second image could have been PAST,[82] HAVE,[83] or IX_1 based on context. When the evidence was insufficient to determine the word being produced, as in the examples in Figure 3.22, these forms were not coded as FPS pronouns, and therefore not included in my FPS calculations.

Figure 3.22

Examples from the data of potential FPS pronouns where the target word is unidentifiable based on phonological form or context

Other times, it was clear that an FPS pronoun had been produced, though it was difficult to determine which FPS pronoun it was. I came across many such instances while annotating my data set. This was most frequently the case with forms that, based on phonological form and context, could have been IX_1 or POSS_1. Unlike the previous examples, these types of coding

[81] HAVE: https://aslsignbank.haskins.yale.edu/dictionary/gloss/1779

[82] PASTbk: https://aslsignbank.haskins.yale.edu/dictionary/gloss/212

[83] HAVE: https://aslsignbank.haskins.yale.edu/dictionary/gloss/1779

decisions did not affect my primary analysis, only the word-level frequencies that I calculated. Figure 3.23 below provides three examples illustrating my deference to phonological form, supported by context as a secondary measure. The first image is an FPS pronoun produced in an introduction, between the words HELLO[84] and NAME.[85] This can be translated as "Hi, my name is." Due to the slightly extended index finger, this form was coded as IX_1. The second image comes before FATHER,[86] and is IX_1 in form, though the signer is talking about their father, making "my father" an accurate translation, rather than "I am a father." These two are not isolated examples of IX_1 being used to indicate possession. In each case, I prioritized form whenever possible. The third image in Figure 3.23 is of a form that comes after BECAUSE[87] and before a depiction that uses a claw handshape on the side of the signer's head. It can be translated as "because that part of my brain." It is possible that an additional camera angle may have captured an extended index finger that was not visible from the three angles I had. But without this evidence, and with the logical possibility of possession from the context, this was coded as POSS_1.

[84] HELLO: https://aslsignbank.haskins.yale.edu/dictionary/gloss/482

[85] NAME: https://aslsignbank.haskins.yale.edu/dictionary/gloss/511

[86] FATHERstr: https://aslsignbank.haskins.yale.edu/dictionary/gloss/764

[87] BECAUSEb: https://aslsignbank.haskins.yale.edu/dictionary/gloss/34

Figure 3.23

Examples from the data of FPS pronouns that are identified based primarily on phonological form, using context as a secondary measure

HELLO [FPS] NAME	IX_1 FATHER	BECAUSE [FPS] [claw-shaped depiction on side of head]
"Hi, my name is"	"my father"	"because that part of my brain"
Coded: IX_1	Coded: IX_1	Coded: POSS_1

These and other phonological realities raised uncertainties in determining the most appropriate annotation labels. I consulted with my advisor about such instances. In cases when participants had granted permission to allow other researchers to view their chats, I consulted with other researchers (described in Chapter 1). For each instance, I took the consulted opinions into account, and made the final decision as to the annotation label.

Turns tiers.

For the purposes of this study, a turn begins when a participant "takes the floor" and ends when they "give up the floor." This means that there are times when both interlocutors are engaged in overlapping turns. This also means some turns are quite short (though not every instance of backchanneling counts as a turn), and turns do not need to include any prototypical "words."

The concept of a turn is fairly straightforward, generally speaking. However, when it comes to measuring turn time with enough precision to calculate per-minute usage rates, some decisions need to be made; specifically, determining what types of communication count as a "turn," and determining turn boundaries. Turn boundaries have been defined variously in the literature, depending on the theory and aims of a given study. Some researchers include preparatory and retractive movements, such that any movement that transitions into or out of the utterance is considered part of the turn (e.g., Coates & Sutton-Spence, 2001). Others include only "stroke-to-stroke" or "content-bearing" manual movements (e.g., Casillas et al., 2015). For my work here, however, total turn time was needed for calculating FPS per minute (FPS/min) rates, meaning a participant's total turn time needed to represent the total time that the participant held the floor. Some preparatory movements signal an intent to take the floor, qualifying them for inclusion as part of the turn. Other preparatory movements do not signal such an intent. Similar distinctions can be made for retractive movements. This means that a strategy of automatically including or excluding all preparatory and retractive movements would be insufficient for my purposes. Additionally, because prototypical "words" are not necessary for taking the floor (e.g., "freeze-look" described by Manrique & Enfield, 2015), I could not rely on the "stroke-to-stroke" hypothesis. Instead, I developed a set of criteria for identifying and measuring turns for this study, presented below. Rather than an exhaustive list, this is a collection of examples and guiding questions.

Identifying a turn:

- Asking or answering a question

- If Person A had missed Person B's utterance (e.g., looked away as part of a depiction), would Person B have repeated it?
- Does it appear that the person intends to hold the floor (whether or not the other person gives up the floor)?
- Does the person's communication change the flow of the conversation?
- Questioning the other participant by raising or lowering the eyebrows, tilting the head, pronounced or rapid blinking, widening of the eyes, and/or leaning forward or backward

When a turn begins:

- Beginning communicative activity of a turn, manual or non-manual
- Does it appear that the person intends to take the floor (whether or not the other person gives up the floor)?
- Hand wave used to signal the beginning of a turn (glossed in ASL Signbank as i(hey)[88])

When a turn ends:

- Ending communicative activity, manual or non-manual, e.g.:
    - Hands rest or settle
    - Final word form is released (e.g., handshape relaxes or changes)

[88] i(hey): https://aslsignbank.haskins.yale.edu/dictionary/gloss/1820; note that this is often used mid-turn as well

- Any remaining phonological forms are perseveration (not fragment buoys[89])

Status indicators tiers.

While I did not meaningfully include this subset of data coding in my current study (other than exploratory behavioral considerations), I did a separate pass for status indicators, recording them on a dyad-level tier. I am describing this as part of my methodology for transparency about my research process, and to provide context for the relevant exploratory section in Chapter 5. Additionally, I am hopeful that I will return to this information in future studies. During this pass, I did thematic coding, looking for instances that could reflect something about the status relationship between paired participants. From these notes, several categories emerged, albeit grounded in my own perceptions, since that is where the notes came from. Some observations/behaviors fit more than one category. I prioritized looking for broad patterns over defining precise category boundaries. Below are the loosely bounded categories I identified, along with examples from each. These examples are not only thought-provoking, but they also provide a qualitative sense of the data in my study, as well as contributing to an overall understanding of what "social status" consists of and how it is constructed. Describing these interactions in written English does create quite a bit of distance from the ASL source data, but it also preserves the anonymity of the participants, to whom I owe a debt of gratitude for their involvement. Examples below are further anonymized by referring to individuals by their

---

[89] Perseveration and fragment buoys are both phonological remnants that linger after a form has been fully produced. Of the two, only fragment buoys appear to serve a semantic function (e.g., directing attention to a concept after naming it). In this way, a fragment buoy would signal the continuation of a turn, while perseveration is an example of when a phonological form might be present despite the signer having given up the floor. See Liddell (2003, pp. 248–250) for more on perseveration and fragment buoys.

participant role, rather than their participant number or HBL designation, and through the use of gender-neutral pronouns. Some examples are aggregated from multiple related annotations within a single chat.

- Overtly mentions hierarchical discrepancy
    - *HighRole* refers to *LowRole* as being part of a younger generation than themself; *HighRole* labels themself as from a "boring older generation" (they have not discussed their ages, but *HighRole* is 32, and *LowRole* is 20)
    - *LowRole* says they went to a deaf school K-12, unlike *HighRole*, adding "so we're different"
    - *HighRole* says, "when I was your age…"
- Advising
    - *HighRole* repeats that *LowRole* should email the coach about joining the team; *LowRole* says yes they'll join this fall, *HighRole* says they'll tell Coach to keep an eye out; *HighRole* later reminds *LowRole* that they'll be looking for them on the team in the fall
    - *LowRole* advises *HighRole* to learn to do some maintenance on their own car to save money like *LowRole* does, stresses how much money *HighRole* could save
- Physical/Posture
    - *HighRole* says *LowRole* doesn't have to explain why they got a cochlear implant, then crosses their arms with one hand at chin/mouth and waits to see how *LowRole* responds

    - After *LowRole* confirms that the only instructions they have been given are "just chat" *(LowRole*'s 2nd, *HighRole*'s 1st), *HighRole* does some large posturing to accept and opens with over-sized production of "ok what's up, what's your name" with arm rested on the back of the chair
- Challenging/Competition
    - *HighRole* won't let *LowRole* get away with "I dunno" when asked why they transferred to GU: "oh c'mon (SICK[90]) there has to be a reason" and gives a few examples
    - *HighRole* tells *LowRole* to come up with something to talk about next; *HighRole* says starting with basic introductions isn't a normal/natural conversation, *LowRole* counters asking how they're supposed to have a normal/natural conversation without knowing who the other person is
    - *LowRole* says they've been to more countries than *HighRole*, covers mouth in mock scandal, *HighRole* plays along, asks where...first is Canada, which *HighRole* good-naturedly blows off, saying they've also been there, *LowRole* playfully chastises that they should have included it
    - *LowRole* is a bit disappointed that so much emphasis is placed on assignments in ASL at Gallaudet, when students will need English to communicate with hearing people, *HighRole* counters that everyone gets

[90] ASL Signbank does not include this exact sign, but it is similar to these two entries, though directed at the signer's own forehead:

THAT'S-SICKneut: https://aslsignbank.haskins.yale.edu/dictionary/gloss/4206/

ANNOYINGstr: https://aslsignbank.haskins.yale.edu/dictionary/gloss/3621/

English their whole lives, so GU needs to get you caught up on the ASL side

- Challenging: ASL
    - *HighRole* friendly-challenges *LowRole*'s initialized sign for TRUCK,[91] and they discuss other options
    - *HighRole* "corrects" *LowRole*'s initialized sign for DEGREE,[92] but then "forgives" it
- Understanding each other
    - *LowRole* misunderstands *HighRole*'s question as a statement, *HighRole* attempts to clarify, but *LowRole* runs with it
    - *HighRole* misunderstands *LowRole*'s reasoning and also that they're talking about deaf schools in Canada, *LowRole* has to clarify twice
- Filling the gaps, carrying the convo
    - After another long pause *HighRole* asks *LowRole* if they have anything else they wants to say; *HighRole* responded on survey that they would have liked less control in the convo
    - A few longer-than-normal pauses in this chat (4 sec, 2 sec) following *HighRole* holding a turn for emphasis with energy that *LowRole* doesn't reciprocate; *HighRole* is really carrying the conversation here, so this may be making them feel some pressure to "perform"

[91] TRUCKt: https://aslsignbank.haskins.yale.edu/dictionary/gloss/1320/

[92] DEGREEd: https://aslsignbank.haskins.yale.edu/dictionary/gloss/3524/

    - It takes a 3–4 sec lull to get *LowRole* to ask their first question at 10:35 (that and *HighRole* giving a “I could sit here all day” look)
- More knowledgeable, explaining
    - *HighRole* (not from *LowRole*’s state) explains to *LowRole* some differences between *LowRole*’s state and another state that *HighRole* is also not from
    - *LowRole* explains how much more beautiful/wonderful a particular European country is than America, *HighRole* asks about towns/people there, *LowRole* explains they were only there for an extended layover
    - *LowRole* will stay with a friend in Silver Spring “near Maryland,” *HighRole* corrects that Silver Spring is *in* Maryland, *LowRole* blows it off
- Turn-taking
    - They start at the same time, *LowRole* defers to *HighRole*; *LowRole* tries to interrupt, *HighRole* finishes then explicitly turns the floor over; Both start to talk same time, both stop, *HighRole* nods to show not going to continue, and *LowRole* takes the floor
    - Been a lot of *LowRole* repeating their backchanneling/feedback/interruptions until *HighRole* sees/acknowledges
- Other
    - *HighRole* compliments *LowRole*’s eyebrows, saying they “look real,” though *LowRole* did not offer that they were not real to begin with

  - *LowRole* didn't remember meeting *HighRole* right off saying they tend to "easy forget"; In *LowRole*'s other chat they forgave the other *HighRole* for not remembering them in a self-depreciating way ("look a mess") and says they only remembered because they're generally good with faces
  - *HighRole* is impressed with *LowRole*'s math courses, because *HighRole* says they couldn't make it past any calculus stuff (STUPIDITYv[93])
  - Both have CIs, but don't use them anymore/in deaf environments; They discuss deaf with and without Cis

Session notes.

I also made notes when I was running sessions, noting my own impressions from my interactions with participants before and after filming. Below are examples from each of the session notes categories:

- Pre-session
  - *HighRole* was a little late, and interrupted my explanation of the paperwork to change the settings on my laptop's trackpad
- How they handled who filled out the Post-Chat Survey where
  - *HighRole* suggested they be the one to stay in the studio, and *LowRole* agreed to go to the other room
- Post-session
  - *HighRole* told me that *LowRole* finished their (*HighRole*'s) sentences, which appeared to have annoyed *HighRole*

[93] STUPIDITYv: https://aslsignbank.haskins.yale.edu/dictionary/gloss/606/

Calculating Variables

Below I will introduce how the variables were defined for my analyses. As previously mentioned, I chose to deviate from the previous studies conducted in spoken English (Kacewicz et al., 2013, Studies 1 and 3) in my methodology, so deviations to defining the variables will be explained and described below as well. Recall from Chapter 2 that I was most interested in their Studies 1 and 3, as they are the two that analyze spoken English rather than written. From the files shared with me by the senior author, Dr. James Pennebaker, I attempted to recreate both the calculations described and the results reported in their paper for Study 3. A similar approach for Study 1 will have to wait for a future paper. I then took an exploratory approach, necessitating other methodological considerations. Below I will introduce how the variables were defined for my analyses.

Variable: Relative status.

As discussed earlier, social status is not truly quantifiable. That said, it is possible to create approximations of status by selecting and aggregating factors that contribute to relative social status. Those approximations can then be used to create and test predictions. My main analysis avoided granular measurements of status, instead defining relative status between participants as simply "relatively higher" and "relatively lower." This approach treats relative status as a binary and ordinal variable.

In addition to my main analysis, I wanted to compare the ASL data I collected to existing English data. In order to do this, I used the data from Study 3 in Kacewicz et al. (2013), as this is one of the only studies I have found that analyzes FPS pronouns in ephemeral (i.e., not written), conversational language use (see the Chapter 2 section titled "Kacewicz et al. (2013)"). In Chapter 2, I touched on some elements of Study 3 that warranted further examination, centering around participant agreement. Below, I will go into more detail about those elements, and explain how they informed my own methodology.

Methodological consideration from Kacewicz et al. (2013), Study 3: Participant agreement

In this section, I will discuss the implications of participant agreement within this work. In their description of Study 3, Kacewicz et al. (2013) report that “there was high agreement within each pair regarding who had higher power r(48) = -.46, *p*< .01” (p. 9). While the reported correlation coefficient of -.46 is considered a medium-to-large correlation in certain contexts,[94] in this context it also reveals that there was likely a good deal of disagreement between participants about who had higher status. Disagreement between participants about their relative status impacts the calculation of both dyad-level status discrepancies and pronoun usage discrepancies. Throughout their paper, the authors frame power/status as relative, suggesting that they view power/status zero-sum, in that any power/status held by one side is, by definition, unavailable to the other. Additionally, as mentioned in Chapter 2, the two survey questions they report using to determine relative status are: “To what degree did you control the conversation?” and “To what degree did you have power in the conversation?” where 1 = *not at all* and 7 = *a great deal* (p. 7). As mentioned above under “Post-Chat Survey,” when I reached out to the second and senior author, Dr. James Pennebaker, to ask for the survey they used (their “IRQ”), the survey he shared included the first question from their paper, but the second question was phrased differently: “To what degree did you have more status than your partner in the

[94] Correlation coefficients are measured on a scale of -1 to +1. A coefficient of -1 represents perfect negative (i.e., inverse) correlation, 0 represents the absence of a correlation, and +1 represents a perfect positive correlation. Interpreting numerical correlations in terms like “weak” or “strong,” however, requires context. Cohen (1988) offers the following rules of thumb for *r* values as starting points: “small” = .01, “medium” = .03, “large” = .05. Context is needed, however, to apply an interpretation to any result.

conversation?" When I followed up about this, he shared another version of the survey that he said may have been used. This one phrased the question as "To what degree did you have more status or power than your partner in the conversation?" Regardless as to the specific wording of the items, I agree with the view in their discussion, viewing power/status as relative (i.e., zero-sum). Unfortunately, this assumption was violated in their dyad-level status discrepancy calculations.

The authors chose to calculate dyad-level status discrepancies by subtracting one participant's self-reported power rating from the other's. In a zero-sum context, this calculation can only produce a viable dyad-level metric when both participants are in agreement. To illustrate with examples, I will use a Likert scale with response options of 1 through 7, as used by Kacewicz et al. (2013). A hypothetical question with variable-sum (as opposed to zero-sum; see earlier "Post-Chat Survey" section for more) scoring would be "How much did you enjoy the chat?" The absolute difference between these responses—one participant's response minus the other's—represents a dyad-level enjoyment discrepancy. For example, if the participants both responded with 2s, the enjoyment discrepancy between them would be 2 - 2 = 0, which would make logical sense, as they had the same reported level of enjoyment (i.e., no discrepancy in their enjoyment). In another dyad, if one person responded with a 7 and the other person with a 6, the enjoyment discrepancy would be 7 - 6 = 1, which would accurately reflect a greater difference in their levels of enjoyment.

However, if status is considered to be relative (i.e., zero-sum), as suggested in Kacewicz et al. (2013), then every participant's self-reported status rating represents a claim to some proportion of the available status. This then also ascribes the remainder of the available status to their chat partner. This means a self-report of 1 is also an "other-report" of 7; a self-report of 5 is also an "other-report" of 3. For a full list of other-reports based on self-reports, see the Appendix or my shared files.[95] To examine the logic of subtracting one participant's self-reported status from the other's in this context, Figure 3.24 below includes the following hypothetical scenarios:

Dyad A. Both participants agree they had equal status during their interaction, so they both self-report 4s. Their dyad-level status discrepancy, as calculated by Kacewicz et al. (2013) is: 4 - 4 = 0. A discrepancy of zero makes logical sense for these two people who agree that they had equal status, or no discrepancy.

Dyad B. Both participants agree that they have a disparate status relationship: one is considerably higher, the other is considerably lower, and they self-report a 6 and a 2, respectively. We can tell these two people are in agreement because a self-report of 6 is also an other-report of 2, and vice versa. Subtracting one of these self-reports from the other results in a dyad level discrepancy of 6 - 2 = 4. This is a higher discrepancy than Dyad A, who agreed they were equal, which also makes logical sense.

Dyad C. These two participants disagree about their status dynamic; they both believe themselves to have had higher status than their chat partner, so they both self-report 6s. Recall that a self-report of 6 equates to an other-report of 2. The two participants' reports are contradictory. However, their dyad-level status discrepancy, according to the method used by Kacewicz et al. (2013), is 6 - 6 = 0. The contradiction is masked by the discrepancy of zero, falsely making it appear as though these two people believe they are of equal status (i.e., no discrepancy).

Dyad D. This scenario is similar to Dyad C, in that the participants disagreed with each other about their relative status, each believing themselves to be of higher status than the other

---

[95] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

person. Dyad D is slightly different though, because the two individuals self-reported slightly different numbers, 7 and 6, resulting in a discrepancy of 1. This dyad demonstrates how non-zero dyad-level discrepancies can also represent a situation where the two participants were operating under contradictory assumptions about their relative status relationship.

Figure 3.24

Calculating hypothetical dyad-level status discrepancies (as absolute values) for use in examining the logic of subtracting one participant's self-reported status from the other's

Dyad		Participant 1 (P1)		Participant 2 (P2)		Dyad-level status discrepancy: P1 minus P2
		self-report	*other-report*	self-report	*other-report*	
A:	**Participants agree** they are of equal status	4	*4*	4	*4*	**0**
B:	**Participants agree** who is higher in status	2	*6*	6	*2*	**4**
C:	**Participants disagree;** both report being of higher status	6	*2*	6	*2*	**0**
D:	**Participants disagree;** both report being of higher status	7	*1*	6	*2*	**1**

Dyads A and B present scenarios in which subtracting P2's self-reported status from P1's produce logical results. In these scenarios, both participants agree about their relative status relationship. Dyads C and D present flaws in this approach. Dyad C's identical self-reports actually indicate they experienced the chat quite differently, both believing they were on the higher end of a large status discrepancy. Rather than revealing what the shared status dynamic was between these two participants, their self-reports reveal two contradictory realities about who had higher status. In light of this, reporting this dyad as having no status discrepancy at all would obscure these contradictory self-reports. Dyad D's similar self-reports also reveal a contradiction. In fact, there are multiple types of disagreement possible between chat partners (e.g., one participant reporting being of equal status, a 4, while their chat partner reported being a 7 out of 7). None of these disagreements are captured in the status discrepancies calculated by Kacewicz et al. (2013).

There is an additional problem introduced by including dyads who self-reported the same value: these dyads do not actually qualify for this study as designed. I introduced the basis for this problem in Chapter 2, stemming from the fact that the correlations the authors report are calculated from discrepancies that are dependent on sign (i.e., being positive or negative). FPS discrepancies can not be properly construed as positive or negative for dyads who self-reported the same status rating. The definition for FPS discrepancies set by Kacewicz et al. (2013) depends on being able to identify a lower-ranking participant and a higher-ranking participant. Without a difference in rank, the FPS discrepancy is impossible to calculate (i.e., undefined). Such FPS discrepancies could instead be construed as absolute values, but this would require a different methodology. Following their stated methodology, the FPS discrepancies of equal-ranking dyads would essentially be absolute values with arbitrarily assigned positive or negative signs. This would make the data from these dyads nonsensical for the purposes of a quantitative analysis, jeopardizing the integrity of results.

There is no indication in their paper that any of the agreement issues discussed here were considered during the original authors' calculations. This may be explained by the possibility that, as their data set suggests, the authors did not actually follow their reported procedure of subtracting the higher-rated person's values from the lower-rated person's. See the "Computational reproduction" section later in this chapter for more on this.

As seen in these examples, subtracting one participant's self-reported status from the other's does not always result in a meaningful dyad-level relative status discrepancy score. This subtraction method instead tells us how different their responses were, but not how relatively different they reported themselves to be at the level of the dyad, nor whether they were in agreement with one another. In essence, dyads who disagree about their relative status relationship do not meet the criteria for the independent variable in this study, as it is impossible to identify a relative status relationship that would reflect both participants' self-reports. For this reason, I wanted to avoid including such dyads in my analyses.

By using binary, objective status designations in my own data collection, I sidestepped the issue of participant agreement in my main analysis. However, in order to compare the ASL data I collected to the English work done by Kacewicz et al. (2013), I also collected subjective information from participants about their perceptions of their relative status relationships using my Post-Chat Survey. Understanding the logic of calculating discrepancies, as discussed above, I chose to filter out dyads who disagreed with one another when analyzing both data sets, as I do not consider these dyads as having met the criteria for the independent variable. My initial attempt to reproduce the results from Kacewicz et al. (2013) using their data included all of the dyads, as I wanted to understand how they had arrived at their reported results. For my next analysis, I identified and excluded the dyads in their data set who disagreed. I first attempted to replicate their analysis using the subset of dyads who had viable status discrepancy scores, and I then redefined status as binary[96] for applying the calculation from my main analysis. When running my data using their approach, I used the subjective information collected from my Post-

[96] Because no personal information was collected from their participants, there was no way to redefine status in their data set that would align with the binary, objective rankings in my main analysis. Instead, I used participants' self-reported subjective status to determine binary status designations.

Chat Survey to create quantitative social status measurements that echoed the ones used in their Study 3, excluding dyads who disagreed with one another on my Post-Chat Survey. All of these attempts are detailed in the sections dedicated to my secondary analysis later in this chapter. For more details on how I measured participant agreement, see my shared files.[97]

Variable: FPS pronoun use.

Compared to quantifying relative social status, quantifying FPS pronoun use is more straightforward. However, simply looking at the raw number of FPS pronouns used by each interlocutor does not account for imbalances in the total volume of their individual contributions. In other words, if one person dominates the conversation, they will be more likely to use more FPS pronouns along with more of every other kind of language feature you could think of. We can control for these differences in verbosity by measuring the target feature (i.e., FPS pronoun use) as a rate, quantifying language use in either time (e.g., minutes, seconds) or utterance units (e.g., words, sentences, turns).

One notable difference between the English studies in Kacewicz et al. (2013) and my work in ASL is the quantification of FPS pronoun usage. For all five studies in Kacewicz et al. (2013), the FPS pronoun usage variable was expressed as a percentage of total words used, rounded to two decimal places. So someone who used 50 FPS pronouns within 1000 total words would have an FPS value of 5.00,[98] meaning 5% of their words were FPS pronouns. This approach is aligned with the word-counting software they used, LIWC, discussed in Chapter 2.

---

[97] Figshare project folder: https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

[98] This rate is calculated as 50/1000 = 0.05, which, when written as a percentage rounded to two decimal places is 5.00%. This FPS rate would appear in Kacewicz et al. (2013) as 5.00.

Counting words may not seem complicated on the surface, but it becomes quite tricky when going beyond a small and tightly-bounded category (such as FPS pronouns). The notion of a “word” is a funny thing. Take "ice cream" and "in-laws," or "727" versus "seven hundred and twenty-seven". How should these be counted? Do the answers change when ignoring written forms and considering only ephemeral language use (i.e., language use that is not written, but happens “in the air”), such as a signed or spoken conversation? What about when ephemeral use is transcribed into text? See Taylor (2015) for more on this and related word-defining issues independent of modality; see Pennebaker (2022) for related downsides of word counting software.

Additionally, there are units specific to signed languages that are difficult to definitively count. Such units include many forms of depiction, a category which accounts for a good amount of ASL usage,[99] and yet currently defies consensus in terms of unit identification and countability. Beyond depiction, other features that challenge word counters include non-manual forms (e.g., mouth morphemes, non-manual negation, temporal head movement), numeral incorporation, buoys, multi-word expressions, and indexing. See Johnston & Schembri (1999) and Liddell (2003) for more.

In addition to deciding how to quantify the language features described above, there are further considerations for if and how to count certain things that occur primarily in ephemeral language that would likely not be included if the person were writing or typing. These occurrences include stutters, false starts, superfluous repetition, fillers (e.g., the finger-wiggle appearing within ASL use that is glossed in ASL Signbank as i(umm),[100] and English’s “um” and

---

[99] As one example, Thumann (2010) reports an average rate of 20 depictions per minute in her data set.

[100] i(umm): https://aslsignbank.haskins.yale.edu/dictionary/gloss/1793.html

“ah”), and the like. See Mahl (1987) for more on these in spoken English, and how much more frequent they are than most people realize.

Despite the drawbacks of word counting, I think it is arguably the better option for analyzing the use of written language, such as the instant messenger conversations in Kacewicz et al. (2013)’s Study 2, and the emails in their Study 4. In written interactions, word counting is not subject to the issues created when transcribing from ephemeral use. And the option of measuring turn time would be all but meaningless for analyzing written exchanges. Would we measure the time it took to type each message? To read? On the other hand, it would be very logical to use time to measure how much each person contributed during an ephemeral conversation. If two people used the exact same words at different paces, the person whose pace was slower would dominate a larger percentage of the conversation than the other. For example, 500 words uttered within five minutes represents a rate of 100 words per minute, while the same 500 words uttered over the course of ten minutes represents a rate of only 50 words per minute.

My data included ephemeral use of a signed language. For all of the reasons above (i.e., what is a word?, the unique questions raised by counting units in signed languages, issues with transcription of ephemeral usage, the relevance of time), I quantified how much each person talked in terms of time, rather than number of words. For my study, the FPS pronoun usage variable is expressed as the number of FPS pronouns used per minute of their total turn time, rounded to two decimal places. So someone who used 50 FPS pronouns and had 10 minutes of total turn time would have an FPS value of 5.00, meaning they used an average of five FPS pronouns per minute.

Primary Data Analysis: ASL Data

My aim with this work was to begin developing an understanding of the potential relationship between relative social status and FPS pronoun use in ASL. To do so, I operationalized FPS pronoun use as the rate of FPS pronouns used per minute, and compared those FPS rates within dyads made up of participants with disparate social status. Social status was conceived "objectively" (in the sense of being comprised of externally identifiable status markers), as described earlier in this chapter, rather than subjectively from the participants' stated perspectives. Of the 24 people who participated in my study, half were part of a triad, meaning they each participated in two chats. The other half participated in a single chat. This arrangement resulted in 18 dyads which, for the purposes of this analysis, equates to 36 total "participants." For more details, see the "Grouping Participants" section earlier in this chapter.

My goal in looking at dyad-level FPS discrepancies was to report whether there seemed to be a meaningful difference in FPS rates based on status discrepancy. When I say "meaningful," I am referring to a difference that can be detected not only mathematically (e.g., a negligible difference that happened to be statistically significant), but rather, a difference that might be useful in understanding our human social experiences. In other words, if there seems to be an effect, how might it translate to our everyday lives?

Before describing the details of my primary analysis, I will first address two topics that provide some additional context. The first is the existence of four dyads that warranted special consideration. The second is a statistic I report in Chapter 4 that many readers may be unfamiliar with, the S-value. While unfamiliarity can be off-putting to some, the S-value itself is quite user-friendly—it is a result that can be intuitively understood using a simple coin toss analogy. These two sections help set the stage for the descriptions of my main analysis. I then describe the process for arriving at the dyad-level FPS/min discrepancies, and provide more detailed information about the four complete triads, to see what can be gleaned from the 12 individuals who participated in more than one chat. Statistics were run using Google Sheets and StatsKingdom.com.

Dyads warranting special consideration.

This section describes my process for how I identified and treated data that warranted special consideration, including issues with incomplete information (three dyads) and technology (one dyad).

As mentioned in the participant demographics sections earlier in this chapter, there were some instances when a participant's chat content did not line up exactly with their responses on the Background Questionnaire. Had the more complete details been available to me while grouping participants, three of the 18 total dyads may not have been paired. Some of these misalignments between the Background Questionnaire responses and the chat content might have been avoided with changes to the Background Questionnaire itself, also discussed in the participant demographics sections. In at least one instance, though, the misalignment appears to be a result of how a participant viewed themselves versus how they described their experiences. If we assume that a person's identities influence their perceptions of their own relative status, then their reported identities (e.g., a participant reporting their K-12 experience as "Deaf residential school") likely reveal something important about their perception of their relative status. This may even outweigh their own narrative descriptions of their experiences (e.g., that same participant telling their chat partner that they went to both deaf schools and mainstream schools).

As discussed throughout this work, status is never tidy. Though I operationalized status using externally identifiable status markers, real people do not fit neatly into boxes. Had I chosen a single status marker to serve as a proxy for status, such as "formative family language," discrepancies would have been easier to categorize, though not without gray areas. This would have minimized, but not completely prevented, opportunities for chats to contain details that contradicted participants' Background Questionnaires. To attempt an even more clear-cut proxy for status, a single truly objective marker, such as age, could have been used and measured very precisely. However, the downside to these simpler approaches is that they fail to capture the nuance of social status relationships. The tradeoff is that, by including multiple status markers to triangulate a rounder approximation of status, I introduced complexity that then had to be considered. The potential effects of additional status markers (e.g., discrepancy in K-12 educational experiences) would have been present within the interactions whether I acknowledged them or not. By including multiple markers, I gained nuance. I also ended up with three dyads whose *HighRole*/*LowRole* designations could be questionable. Because status in the real world is indeed messy, I feel that this tradeoff was worth it, so I avoided the temptation to screen out these dyads post-hoc. Instead, in the spirit of transparency, I indicate these three dyads within the results, where appropriate, for the reader to be able to assess what their impact might be.

The second type of issue, that of technology, arose with the only dyad that included a participant who joined via Skype, described in the "Filming" section above. At times, the lower frame rate of the Skype recording made it difficult to determine whether an FPS pronoun had definitely been produced—even more difficult than for the rest of the data set, as described earlier in the "FPS tiers" section. Because all of the other participants were filmed at the same frame rate, the incidence of FPS pronouns not captured by any of the three cameras should be fairly consistent across participants. But because this one *HighRole* participant was recorded at a lower frame rate, their FPS/min rate may be artificially low. This dyad is also indicated in the results, where appropriate. In order to maintain the anonymity of this single dyad, their results will not be distinguished from the three described above.

S-values.

Most of the measures I report are values that are familiar to most researchers (e.g., mean, *p*-value). One that is less common, but that I find extremely useful, is the S-value. The S-value (also referred to as "surprisal value," or as "Shannon-information value" for its creator, Claude Shannon) is one way to reframe *p*-values into a more intuitively comprehensible format (Greenland, 2019). There is nothing wrong with the *p*-value, in and of itself, but it does not lend itself to easy, intuitive interpretation. As a result, many problems have been raised with *p*-values, many of which reflect issues that are more cognitive than statistical (Rafi & Greenland, 2020). Knowing this, we can choose a different value that is more intuitively comprehensible to the average human brain (as opposed to the brain of a statistician). To this end, *p*-values can be converted to S-values, which are less susceptible to the common misunderstandings and misinterpretations that often plague *p*-values. For the math people: The direct mathematical relationship between p and S can be summarized as $p = .5$^S.

For math people and non-math people alike: An S-value tells us how surprising something is. More specifically, it can be thought of as a measure of how surprising a finding is in the context of a particular study. This is calculated such that an effect with an S-value of 2.0 is about as surprising as correctly predicting that two coin tosses in a row will both be heads. That is, not extremely surprising. Two heads in a row could easily happen with a fair coin, though predicting both tosses correctly certainly feels just a bit lucky. This S-value of 2 is equivalent to a *p*-value of .25, which is (rightly) nowhere near typical statistical significance. An S-value of 5 ($p = .03$), on the other hand, would be more surprising—correctly predicting heads five times in a row would likely feel out of the ordinary, or perhaps feel as though you were dealing with something other than pure luck or "chance." This would provide some evidence that there may be something at work other than chance (e.g., perhaps the coin was double-headed). Not proof, only evidence. For a more direct comparison with the traditional *p*-value cutoff point, a *p*-value of .05 corresponds to an S-value of 4.3. In my results, I will report both of these values: *p*-values to align my reporting with commonly used statistical measures, and S-values to make my reporting more intuitively reader-friendly.

Counting FPS pronouns.

After annotating FPS pronouns in ELAN, I exported the data and calculated frequencies for each FPS pronoun for each participant. While these frequencies do not feature independently in my analysis, they are valuable as a description of my primary data, and they were used to calculate the dependent variable of FPS/min. In Chapter 4, I provide information about the FPS frequency data before I present the FPS/min rates. This includes FPS frequencies from the data set as a whole, as well as the number of FPS pronouns used by each participant, independent of how much turn time they used. These totals are accompanied by simple differences between chat partners, calculated by starting with the *LowRole* participant's frequency and subtracting the *HighRole* participant's frequency, so *LowRole minus HighRole*. This order of subtraction was chosen so that the hypothesized results (i.e., *LowRole* using FPS pronouns at a higher rate) would be in the form of positive numbers. Whether or not the results fit the hypothesis is not a question for the section on FPS pronouns though, as that section reports only the numbers of FPS pronouns used, not the rates used per minute. However, I establish this order of subtraction from the very beginning to ensure consistent and transparent treatment of the data. In that section I also provide descriptive statistics (e.g., mean, median), and results from statistical hypothesis testing of mean group differences for the two groups (e.g., S-value, effect size), along with a data visualization showing the FPS frequencies paired by dyad.

Calculating turn time.

To calculate each participant's turn time, I started with the total turn time in ELAN, shown in seconds, and converted those values to minutes. As with FPS pronoun frequencies, turn times do not feature independently in my analysis, but were required to calculate the dependent variable of FPS/min. In addition to the total turn time for each participant, I calculated the percentage of the total chat time each participant's turn time represents. For these percentage values, I am interested in similar information to what is described for FPS frequencies and turn time in minutes (i.e., participant-level values, descriptive and mean-comparison statistics), but with one addition: the sum of both participants' turn percentages for each chat. These sums provide a sense of the character of the turns, with turns defined according to the earlier "Turns tier" section. The turn percentages indicate the balance of each chat partner's participation in the conversation (e.g., did one person dominate the conversation?), and the sums hint at the prevalence of overlapping turns, despite the presence of occasional brief lulls (e.g., if the turn percentage sum is greater than 100%, then there was enough turn overlap to make up for any lulls) . Because the chats varied slightly in length, I used each chat's length to calculate each participant's turn percentage, rather than working from the intended general chat length of 20 minutes.

Comparing FPS/min rates.

I compared the FPS/min rates of *LowRole* and *HighRole* participants both at the group-mean level (e.g., “In the context of paired dyads, was the mean FPS/min rate of the *LowRole* group greater than that of the *HighRole* group?”) and at the dyad level (e.g., “What were the discrepancies between the *LowRole* and *HighRole* participants?”). This section lays out the methods I used for both of these inquiries.

Before I describe comparisons of FPS/min rates, I would like to reorient the reader’s attention back to the relative nature of social status. Status relationships are dynamic, being co-constructed and continually recalibrated throughout interactions. Social status only exists in a relative sense. See Chapter 2 for more on this. Working from these assumptions, and attempting to test the inverse-relationship hypothesis laid out by Pennebaker (e.g., Tausczik & Pennebaker, 2010; Kacewicz et al., 2013), we arrive at the following supposition: If participants’ FPS rates truly are a result of their relative status relationships, then a given FPS rate would be, at least in part, the product of the specific status dynamic that occurred within a particular chat. In other words, the borrowed hypothesis being tested presupposes status differentials are somehow driving FPS rate fluctuations. To be clear, I am not assuming a causal relationship between social status and FPS rates; I am making methodological decisions in response to the hypothesis I have borrowed for testing. At the risk of prematurely revealing my results, I will nonetheless provide one evidence-based justification for tying FPS rates to their chats of origin: individuals in the ASL data set who participated in two chats tended to have very different rates of FPS/min across their two chats. This evidence supports the idea that an FPS rate is not the result of something intrinsic to the individual in a vacuum, but may be the result of the context the individual is in—and contexts include many factors other than social status. For these reasons, I will be considering FPS rates within the context of the chats they occurred in. Exceptions are made only in exploratory analyses in Chapter 5, and are flagged as such.

To begin the process of comparing FPS rates, I calculated the rate of FPS/min for each participant, dividing their total number of FPS pronouns by their turn time in minutes. As in the previous two sections for the component parts of this variable, I calculated the simple difference between chat partners' rates as *LowRole minus HighRole*. As mentioned earlier, this order of subtraction was chosen so that the hypothesized results (i.e., *LowRole* using FPS pronouns at a higher rate) would be in the form of positive numbers; a negative difference indicates that the hypothesis was not supported by observations from that dyad's chat.

The traditional way to determine whether there is a difference between two groups is to compare the means of the two groups and check for a statistically significant difference between them. Because my data set of FPS/min rates is small ($n = 18$), does not follow a normal distribution (i.e., the points do not fit under a normal bell curve), has an ordinal independent variable (*LowRole* and *HighRole*), and has a paired dependent variable (FPS/min rates within dyads), the most appropriate test is the nonparametric Mann-Whitney/Wilcoxon Signed-Ranks test.

Looking at data in the form of data visualizations is important no matter the shape of the distribution, but it can be especially useful for non-normal distributions. To conceptualize the distribution of dyad-level FPS/min discrepancies, I am using a paired-dyads visualization and a scatterplot visualization. When analyzing discrepancy data, percentage differences have an advantage over simple differences because they allow for easy visual pattern recognition and also proportional standardized comparisons. For example, a simple difference of 5 FPS/min would be much more noteworthy for a dyad with lower rates of FPS/min (e.g., 5 and 10) than one with high rates of FPS/min (e.g., 15 and 20). The contextualized magnitude of the discrepancies is better captured using percentage differences.

Percentage difference calculations come in multiple flavors, and I have chosen four to compare my *LowRole* and *HighRole* data: two standard percentage differences and two symmetrized percentage differences. By their nature, all four of these reveal variations of the same underlying pattern with the data points in the same order, the only difference being that their distributions are stretched/compressed to different scales. Choosing just one approach would be akin to selecting those results to represent the "true" percentage differences between chat partners. Presenting the results of multiple valid approaches next to each other places the emphasis on the schematic underlying shape, rather than the precise values obtained by any one calculation method. Each of the four calculations comes with its own lens, and I will describe each of them below.

The standard percentage calculation is the type of percentage difference that most people are familiar with, and the results can be interpreted in terms like "A is 10% greater than B." The results of this calculation are directionally dependent though, meaning one participant's FPS/min rate must be considered the baseline (i.e., the reference group) for comparing the other participant's rate. The first two of my four percentage difference calculations take this approach. The first uses *HighRole* as the baseline, meaning the results can be understood as showing how much *LowRole*'s rate differs from *HighRole*'s. For this reason, I refer to this approach as "*LowRole*>/<*HighRole*" as shorthand for the way results are framed: "*LowRole* used X% more/less than *HighRole*." The opposite framing, using *LowRole* as the baseline, is the second standard percentage difference approach that I use. I refer to this second calculation approach as "*HighRole*>/<*LowRole*." Because the baselines are different, the denominators are different in these formulas, making the results from these approaches asymmetrical. In other words, if A is 10% greater than B, then B can not be 10% less than A. As someone who did not learn to love the magic of math until well into adulthood, I was initially astounded by the asymmetry produced by reversing the directionality of the framing here. An explanation is warranted.

As one example, in a dyad where the *LowRole* participant used 60% more FPS/min than the *HighRole* participant (*LowRole*>/<*HighRole*), it is also true that the *HighRole* participant used 38% fewer FPS/min than the *LowRole* participant (*HighRole*>/<*LowRole*). Notice that reversing the direction of the comparison does not produce a symmetrical result (i.e., *LowRole* using 60% more does not equate to *HighRole* using 60% fewer). Comparing usage in terms of percentage differences is useful in standardizing the unit of difference (i.e., percentage), making it easier to compare the differences between dyads and allowing patterns to be seen more easily. However, the asymmetrical nature of percentage comparisons can make them difficult to translate into real-world terms. From the previous example, a 60% difference and a -38% difference both accurately represent the same data, but are not intuitively construed as equivalent. That aside, these percentage differences are intuitively interpretable as something along the lines of "the *LowRole* participant used X% more FPS/min than the *HighRole* participant." This type of intuitive interpretation is useful, albeit incomplete. However, there is a more salient problem with using these standard percentage differences.

When comparing "before" and "after" data, there is a natural baseline, making traditional percentage differences appropriate (e.g., "the value increased by 15% after treatment"). However, I am comparing relative use of pronouns during conversations, where neither person's language use is inherently the "baseline" for the other's. It feels intuitive to frame percentage differences in terms of one participant's rate being a certain percentage higher or lower than the other's. Yet the directional dependance of such calculations forces a choice between two inherently biased options (i.e., the asymmetry described in the previous paragraph). Because of the familiarity of these percentage calculations, I will include both in my reported results. However, I can also remove the directional dependence altogether by calculating a symmetrized percentage difference for each dyad, creating a more neutral comparison. One drawback to using a symmetrized percentage difference calculation is the fact that these are relatively unfamiliar to many readers, which can make them less intuitive to interpret, at least initially, but I see this as a tolerable drawback that can be overcome. See Nuzzo (2018) for a useful discussion of simple differences, traditional percentage differences, and symmetrized percentage differences.

There are various ways to calculate symmetrized percentage differences, each of which yields different results. I have chosen two of the more commonly used calculations to include in my reported results. One can be summarized as "the difference divided by the sum," which I will refer to as "*diff/sum*" (Berry & Ayers, 2006). *Diff/sum* produces bounded results on a scale of -100% to +100%. The other calculation can be summarized as "the difference divided by the mean," which I will refer to as "*diff/mean*" (Cole & Altman, 2017). *Diff/mean* does not impose the same upper and lower limits as *diff/sum*, allowing for a wider possible distribution of percentages, depending on the data set. Both *diff/sum* and *diff/mean* standardize the denominator, and this standardization creates a comparison that is not directionally dependent. That said, the order of subtraction for the *diff* portion of these calculations (i.e., the "difference" or discrepancy) is still a decision that needs to be made. In a data set where the differences can be reported as absolute values, this decision would have no effect on the results. But when the sign of the results (i.e., being positive vs. negative) is meaningful, reversing the order of subtraction would reverse the order of the results; that is, the sign of each result would be reversed. For example, 5% would become -5%. Once an order of subtraction is decided, the values resulting from *diff/sum* and *diff/mean* can each be understood as a way to view "the difference" between two values. While these two sets of results will be different, they are each a valid way to calculate percentage difference without requiring one participant group to serve as a baseline for the other. As such, these types of calculations are a better fit for my data set.

While it is uncommon to report multiple calculations for percentage difference in academic publications, I am doing so for two reasons.[101] One, as I described earlier, is to force the reader to focus on the underlying shape of the data rather than any single set of numerical summaries. The other is to acknowledge that the choice of calculation is an element of methodological design that necessarily imposes a lens, whether or not a researcher intends or realizes it. Even pre-registering the choice of calculation does not remove the powerful influence of such decisions on the results (e.g., *diff/sum*'s bounded limits can create results that appear less extreme in both directions than results of *diff/mean* or of the traditional percentage differences). And even if a researcher is unaware of how this choice could be exploited, the fact is that there is no single "right" or "neutral" way to frame results in a study like mine (or perhaps any study). This means that the presentation of results from any single percentage difference calculation gives these results an artificial veneer of being "the results," rather than one interpretation of the data. See Silberzahn et al. (2018) for an excellent example of how different data analysis methodologies, including the choice of calculations for statistical summaries, can dramatically alter reported results.

[101] And one reason I am able to report multiple calculations and describe them in great detail is because, thankfully, the space restrictions imposed on journal articles largely do not apply to dissertations.

The formulas I used for all four calculations of percentage differences are as follows, with "*LowRole*" and "*HighRole*" in the formulas representing that participant's per-minute rate (e.g., "*LowRole*" = the *LowRole* participant's FPS/min rate). Each formula is accompanied by an example of how the results of that formula can be read. Note that the results from the two symmetrized percentage difference formulas can be read the same way.

- *LowRole>/<HighRole* = $\frac{LowRole - HighRole}{HighRole} \times 100$

  "*LowRole* used X% more/fewer FPS/min than *HighRole*"

- *HighRole>/<LowRole* = $\frac{HighRole - LowRole}{LowRole} \times 100$

  "*HighRole* used X% more/fewer FPS/min than *LowRole*"

- *Diff/sum* = $\frac{LowRole - HighRole}{LowRole + HighRole} \times 100$

  "Between *LowRole*'s and *HighRole*'s FPS/min rates, there was a difference of X%"

- *Diff/mean* = $\frac{LowRole - HighRole}{(LowRole + HighRole)/2} \times 100$

  "Between *LowRole*'s and *HighRole*'s FPS/min rates, there was a difference of X%"

There is an argument to be made for indicating symmetrized percent differences with a unique denotation (i.e., not "%") to distinguish them from the more commonly used traditional percentage differences. However, the percentage symbol alone is already ambiguous as to which value is being used as the baseline, requiring an explanation of which direction the comparison should be read (e.g., "Was it *HighRole* or *LowRole* that used 25% more?"). Because of this, I do not see any problem with using the same symbol for symmetrized percentage differences as well, so long as the choice of calculation is made clear in all cases.

Triads.

My original plan, prior to beginning data collection, had been to analyze the FPS/min rates of B participants as my primary focus. As described in the "Filming" section earlier in this chapter, once data collection sessions began, I made the decision to include single dyads along with complete triads. The previous sections have described my approach to analyzing the data set as a whole, and this section describes my approach to analyzing the subset of dyads who were part of complete triads.

Individuals in complete triads each participated in two chats, sometimes with different in-chat roles, and sometimes with the same role. For a refresher on the difference between in-chat roles (*LowRole* and *HighRole*) and triad-level designations (H, B, and L), please see the "Grouping Participants" section earlier in this chapter. By having these two-chat participants, I was able to do some intra-subject analyses. I compared each participant to themselves across their two chats, hoping to see if their FPS/min rates showed any patterns across different status dynamics. As each individual only participated in two chats, there is no way to use this data to suggest any kind of intra-subject correlation, but it is reasonable to look for preliminary alignment in the observation of the complete-triad participants as a group.

I chose not to calculate group-level statistics for this group of dyads because doing so would suggest that these dyads represent a cohesive group that is somehow meaningfully different from the other dyads. It is true that the presumed social status relationships between participants in these complete triads can be considered across the entire triad, which is not a feature of the other dyads. Still, FPS/min rates occurred within the context of each specific chat; see “Comparing FPS/min rates” above for more on the importance of considering rates within their own contexts. However, as described in the “Social Status” sections in Chapter 2, there is no way to objectively quantify social status, and so there was no way for me to standardize status discrepancies between H, B, and L across different groups. This means that the distances between any two participants can only be considered within-triad. In other words, we can assume that the status difference between the H participant and the L participant is greater than the difference between the H participant and the B participant—but only within the same triad. Nothing can be assumed about the potential social status relationships between participants in different groups based on their H/B/L designations (see the “Grouping Participants” section for more on this). Hence, there is nothing that sets the complete triads (i.e., the 12 dyads who were part of complete triads) apart in a group-level sense. This means that calculating summary statistics for the triads as a subgroup would be akin to arbitrarily selecting a subgroup for analysis. The same can be said for calculating summary statistics for the subgroup of B participants. Such subgroupings are an opportunity for inadvertently “finding hidden patterns” that are likely to be spurious.

Secondary Data Analysis: English data and ASL data

As previously mentioned, most of the studies that look at the relationship between FPS pronoun use and social status have been conducted using data generated in written English. As the language data I collected in ASL was both ephemeral and extemporaneously produced, the only two comparable English studies that I have found are Studies 1 and 3 from Kacewicz et al.

(2013).[102] As previously mentioned, the senior author from Kacewicz et al. (2013), Dr. James Pennebaker, shared with me the files he had from this paper, as the first author had been working under his guidance at the time of the writing. Some of the files are not accessible, as the .spo file format (a type of SPSS output file) is no longer supported by SPSS, the SPSS Legacy Viewer, or any other program as far as I can tell. Fortunately, there were also .sav files containing what appears to be the raw data for each of their five studies. While more detailed analyses of Studies 1, 2, 4, and 5 are warranted, they will have to wait for future publications. Because my methodology most closely aligned with that of Study 3, I chose to do a detailed analysis of Study 3 in order to attempt initial comparisons between their English and my ASL data. This process can be broken into four stages. Because ideas and terminology around research "reproducibility" are currently in flux (Goodman et al., 2016), the following list includes the name I assigned to each stage and a brief description of my intention for it. The following sections describe each stage in detail. Results from each stage are reported in Chapter 4 and discussed in Chapter 5.

1. Computational reproduction: Using their data, attempting to reproduce their reported results
2. Refining their approach: Using their data, revising their design
3. Conceptual replication: Using my data, applying the refined approach from above
4. Reanalysis: Using their data, applying the approach from my study

[102] While Markowitz (2018) used English data that was transcribed from Academy Award acceptance speeches, the percentage of those speeches that had been generated in written form, including those written by an author other than the speaker, is unknowable. Additionally, these speeches do not represent conversational language use.

1. Computational reproduction: Using their data, attempting to reproduce their reported results.

During the computational reproduction, I initially used the authors' description of Study 3 in their paper to guide my efforts. Upon beginning my initial attempt, I discovered a few apparent discrepancies between the description in their paper and the data set that was shared with me. Namely, the potential exclusion of two dyads, identifying which questions were used to measure participant status, and how dyad-level discrepancies were calculated. I will discuss each of these in turn, as well as how I addressed them.

Potential exclusion of two dyads.

Kacewicz et al. (2013) report that for Study 3, 50 dyads participated and two were excluded "because of a computer error" (p. 9). The data set I received included data for 50 dyads, and, for several reasons, I decided to use all 50 in my computational reproductions. First, it was unclear to me which dyads might have been excluded. I checked for missing data points, but there was no clear pattern. For each of the 50 dyads, there are 343 columns of data. Four dyads are missing some number of data points (in order of most-to-fewest missing points: Dyad 28, Dyad 13, Dyad 7, Dyad 6). These missing points are marked in the corresponding spreadsheet in my shared project folder, and they represent a combination of what appear to be survey question responses and other data points that are harder to identify. None of these missing points appear to be from categories that were used in the calculations reported in their paper (i.e., pronoun data or status data they might have used). As far as I can tell, there is no clear pattern that makes it obvious which two of these four might have been excluded.

I also considered the dyad numbering system as a clue to which dyads might have been excluded. Each dyad has an identifying number, but there is a break in the numbering system;

there are dyads numbered 1-25, and 31-55. However, this would suggest five missing dyads, not two. And it turns out that the division appears to be motivated by something other than dyad exclusion, as dyads of one sex (coded as "2") are numbered 1-25 and dyads of the other sex (coded as "1") are numbered 31-55. Given this, I believe it is unlikely that this numbering system helps reveal anything about the two excluded dyads.

The second reason I decided to include all 50 dyads in my computational reproductions is that there appear to be inconsistencies in the reported degrees of freedom, suggesting slightly different sample sizes for this study. The agreement calculation they provide shows the degrees of freedom as 48 (p. 9). Assuming that degrees of freedom = n - 2, the sample size can be inferred as 50, since 50 - 2 = 48. One interpretation of this is that perhaps they collected data from 52 dyads, but were only able to use 50. Though this is not entirely consistent with the explanation that 50 dyads "participated" (p. 9), it could be that the two exclusions happened before any data was actually collected from those potential dyads, or that the authors intended "participated" to mean "included." However, it appears that 49 dyads may have been used for calculating the regression model, based on the reported degrees of freedom for the F statistic (p. 9), so I do not feel comfortable claiming to know their actual sample size. That said, for a sample size in this range (n = 49 to n = 52), any two dyads that may have been excluded would represent a small percentage of the data set. Additionally, the absence of extreme outliers among the 50 included dyads suggests that removing any two of the 50 I had data for would not have a substantial effect on the outcome.

Identifying which questions were used to measure participant status: *2Q* vs. *4Q*.

Kacewicz et al. (2013) report using two Interaction Rating Questionnaire (IRQ) questions to determine participants' perceived status: "To what degree did you control the conversation?"

and “To what degree did you have power in the conversation?” where 1 = *not at all* and 7 = *a great deal* (p. 7). I refer to this approach as *2Q*. While the data set contains columns that appear to correspond with participant responses to these two questions, there were no columns that summed these two questions to create participant-level status scores. However, there were columns that summed these two questions along with two additional IRQ items (which I refer to as *4Q*). The two additional items asked participants to place an X on a Likert-style line to represent “the percentage of control you had in the conversation” and “the degree to which you dominated the conversation.” The four-question summation columns were labeled “domin1” and “domin2” (adjacent to the other columns for Participants 1 and 2, respectively). The existence of these *4Q* columns, combined with the nonexistence of *2Q* columns, suggests that the authors may have used these four questions, rather than just the reported two, to calculate participant status scores.

I attempted to reproduce the results reported in their paper by following the methods described in their paper (*2Q* status scores), but this attempt was unsuccessful. As my goal was to determine, to the extent possible, the actual methods used by the authors to achieve their reported results, I also ran analyses using the four-question paradigm as suggested by the organization of their data set (*4Q* status scores). Neither the *2Q* nor *4Q* status score calculations successfully reproduced their reported results when following the discrepancy calculations as described in their paper. However, following the discrepancy calculations implied in the organization of their data set, as discussed in the following section, yielded results that were nearly identical to their reported results, but only when using the *2Q* status scores. For this reason, when it came time to do my initial conceptual replication attempt for Study 3 using the ASL data I collected, I based one of my status score calculations on this *2Q* calculation.

How dyad-level discrepancies were calculated: *Lo minus Hi* vs. *P1 minus P2*.

In order to calculate dyad-level discrepancies, for both pronoun usage rates and status, Kacewicz et al. (2013) report having subtracted the higher-rated participants' score from the lower-rated participants' score; in other words, low minus high. I will refer to this calculation as *Lo minus Hi*. As discussed earlier in this chapter, in the section titled "Methodological consideration from Kacewicz et al. (2013), Study 3: Participant agreement," the authors appear to have made no distinction between those dyads who agreed about their relative status discrepancy and those dyads for whom there is no clear "lower" or "higher" designation—either due to conflicting self-reports or self-reports indicating perceived equality.

Similar to the issue of which IRQ questions were used, as discussed in the previous section, the organization of the data set suggests a different calculation than the methods described in their paper. In the data set, two of the columns are labeled "domindf" and "idif," presumably "dominance difference" and "i-word[103] difference," respectively. Both of these columns contain values that are the result of subtracting Participant 2's score from Participant 1's. In other words: *P1 minus P2*, not *Lo minus Hi*.

As previously mentioned, my initial attempt to reproduce the results reported in their paper by following the methods described in their paper (*2Q* and *Lo minus Hi*) was unsuccessful. Because the organization of the data set implies that dyad-level discrepancies were calculated as *P1 minus P2*, I tried this calculation as well. The most successful computational reproduction attempt used two-question (*2Q*) status scores and calculated discrepancies as *P1 minus P2*. That said, my reported results calculated using *P1 minus P2* should be understood as thought-

[103] Kacewicz et al. (2013) sometimes refer to first-person singular (FPS) pronouns as "I-words" (p. 3, p. 13). Pennebaker, the senior author, does the same in other publications, such as Pennebaker (2011) and Chung & Pennebaker (2007).

experiments in replication, rather than plausible results for evaluating a potential relationship between social status and FPS pronouns. See the various *P1 minus P2* results in Chapter 4 and related discussions in Chapter 5 for more on this.

2. Refining their approach: Using their data, revising their design.

For the next step of my secondary analysis, I reanalyzed the data from Kacewicz et al. (2013), to see how a specific change to the methodology would affect the results. While working through the computational reproduction, I was bothered by a specific element of the methodology described in their paper. Based on their dyad-level status calculations, there is a logical possibility that some dyads would disagree on their relative status. Based on their reporting of dyad-level agreement, disagreement appears to have been prevalent. In these cases, when chat partners disagreed about their relative status relationship, it seems impossible to assume whose perception of the relative status relationship should be used for the FPS rate calculations. Accordingly, I excluded these dyads from the original data set and reanalyzed the remaining dyads. To see a detailed explanation of my reasoning behind excluding dyads based on their reported status relationships, as well as a look at the other-scores implied by the self-scores, see the earlier section titled "Methodological consideration from Kacewicz et al. (2013), Study 3: Participant agreement."

For this step, I determined which dyads would be excluded by calculating agreement scores for each dyad. For each of the two selected questions (see earlier section, "Identifying which questions were used to measure participant status: *2Q* vs. *4Q*"), I summed chat partners' numerical responses. A sum of 8 was considered perfect agreement. One example of this would be both participants scoring themselves as 4s, implying that both believed themselves to be of equal status. Another example would be one person scoring themselves as a 6, and the other scoring themselves as a 2, meaning they agreed who had higher status, and they agreed that it was a sizable difference. Because perfect agreement seems like an unrealistically high expectation, I chose to also include dyads who disagreed slightly. In addition to considering a sum of 8 as perfect agreement, I considered sums of 7 or 9 to be a reasonable level of agreement. Dyads who scored within the range of 7–9 for both selected questions were included in this refinement stage. For more details on these calculations and comparisons, see the Appendix or my shared Figshare project folder.[104]

Using the categories from my computational reproduction attempts, I applied this refinement both to the methodology Kacewicz et al. (2013) reported having used (*2Q* and *Lo minus Hi*) and the methodological choices that most closely reproduced their reported results (*2Q* and *P1 minus P2*). For the latter, no further adjustments needed to be made. However, for calculating *Lo minus Hi*, dyads who reported being in perfect agreement that they were equal (e.g., they both reported being 4s), had to be excluded. This was because there was no way to determine who was relatively lower or higher, making a *Lo minus Hi* calculation impossible.

3. Conceptual replication: Using my data, applying the refined approach from above.

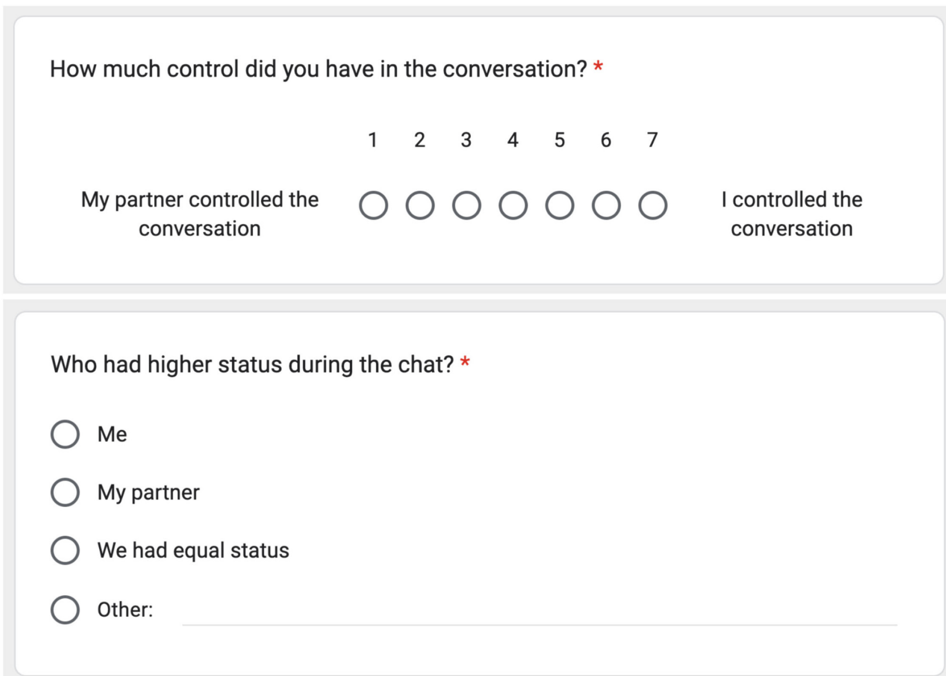
I performed two conceptual replications using the ASL data I collected, both using items from my Post-Chat Survey for the independent variable of social status discrepancy. For details on the quantification of these two paradigms, see my shared Figshare project folder. For the first of these, I measured status using the two questions from my Post-Chat Survey that most closely matched the two questions that appear to have been used by Kacewicz et al. (2013) in Study 3. These two questions and their response options can be seen in Figure 3.25 below. I refer to this

[104] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

as the two-question (*2Q*) paradigm. In this approach, I consider the status measurements to be "chat-internal," because these questions ask about participants' subjective perceptions of their relative status relationship constructed during the chat.

Figure 3.25

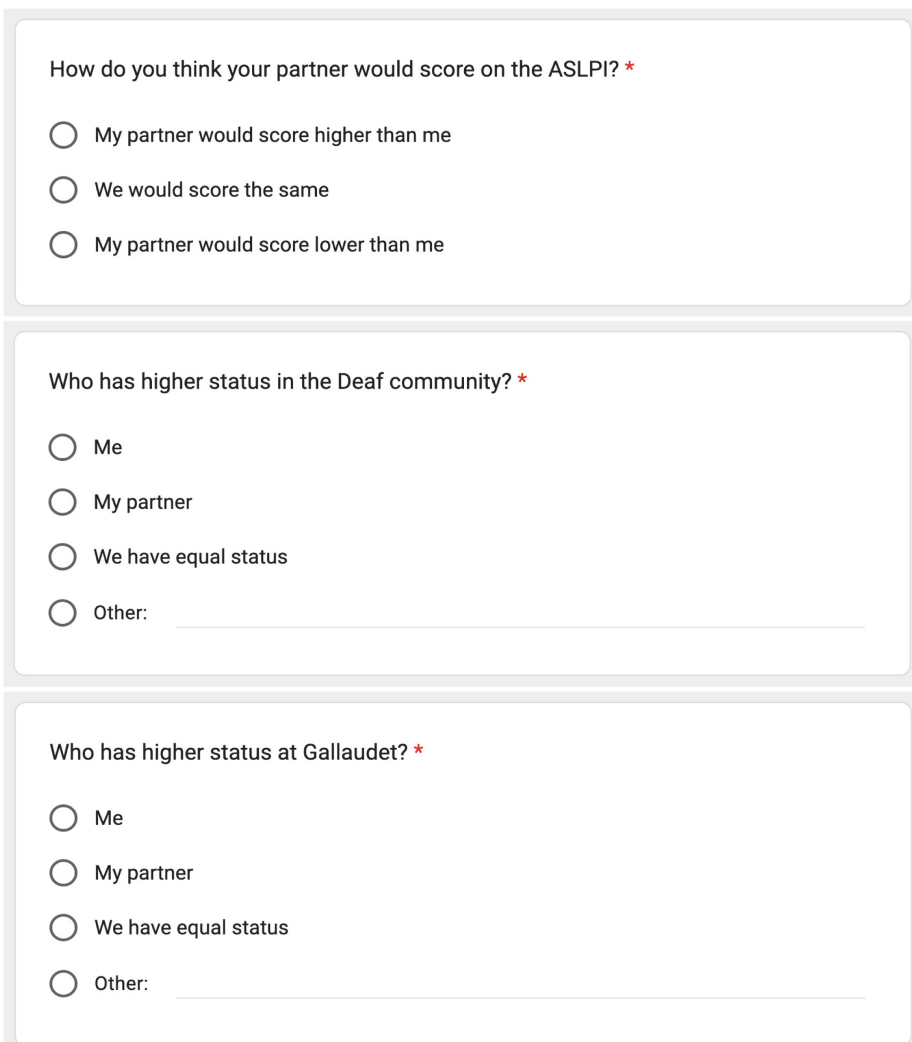
The two Post-Chat Survey questions used for conceptual replications under the chat-internal 2Q *paradigm (ASL data)*



For the second conceptual replication, I measured status using three questions from the Post-Chat Survey that I feel were likely to have captured the intent of Kacewicz et al. (2013). These three questions and their response options can be seen in Figure 3.26 below. I refer to this as the three-question (*3Q*) paradigm. As discussed in Chapter 2, directly claiming a status differential between yourself and another person can be uncomfortable, but it may be that people are more readily willing to acknowledge how a third party might compare you to that other person. I consider these questions to be "chat-external" because they still ask participants to rank themselves relative to their chat partners, but they shift the focus. Rather than asking about the respondent's experience during the chat, these questions ask about how the respondent thinks others would perceive them and their chat partner. My hope is that this framing allowed participants to acknowledge status differences without guilt or fear of being perceived unfavorably. See the "Social status in general" section in Chapter 2 for discussion of saving face as it relates to claiming status differences.

Figure 3.26

The three Post-Chat Survey questions used for conceptual replications under the chat-external 3Q *paradigm (ASL data)*



How do you think your partner would score on the ASLPI? *

- My partner would score higher than me
- We would score the same
- My partner would score lower than me

Who has higher status in the Deaf community? *

- Me
- My partner
- We have equal status
- Other:

Who has higher status at Gallaudet? *

- Me
- My partner
- We have equal status
- Other:

As discussed in the section on computational reproduction, Kacewicz et al. (2013) reported calculating discrepancies as *Lo minus Hi*, but appear to have actually reported results from calculating *P1 minus P2*. I included *P1 minus P2* in my conceptual replications using the ASL data I collected, not because I believe this is the most appropriate fit for the data, but in an effort to replicate the methods I believe the original authors used. For these calculations, I randomized participants to P1 and P2 designations. In addition to *P1 minus P2*, I also included the *Lo minus Hi* calculation, as it is the calculation they reported using, and because this approach makes logical sense. In sum, my conceptual replication results in Chapter 4 include four distinct approaches:

- *2Q + Lo minus Hi*
- *2Q + P1 minus P2*
- *3Q + Lo minus Hi*
- *3Q + P1 minus P2*

4. Reanalysis: Using their data, applying the approach from my study.

The final step of my secondary data analysis was to run the English data using the methods from my primary analysis with the ASL data. Previous steps of my secondary analysis included the quantitative status rankings used by the original authors, but my primary analysis treated status as a binary ordinal variable. To be able to code participants as relatively lower or higher in status, I used the refined data set from the *2Q + Lo minus Hi* attempt from the refining stage described above. I then followed the same procedures as in my primary analysis. My aim with this step was to provide a English results for a side-by-side comparison with my primary analysis of the ASL data.

Conclusion

This chapter presented my methodology for data collection and processing in my attempt to analyze a potential connection between relative social status and FPS pronoun use in ASL. I have detailed how I collected background information from prospective participants and how I used that information to determine chat placements. I have also described the participant pool in terms of various relevant demographics. The procedures for filming and annotating the data are presented, and my variables are defined.

I have laid out the procedures I followed, both for my primary analysis of the ASL data and the secondary analysis incorporating relevant existing English data. Chapter 4 will contain the results of my analyses. I then use these results to discuss my conclusions in Chapter 5 about what these data sets suggest about the nature of the relationship between relative social status and FPS pronoun use.

Chapter IV: Results

This chapter will address the question of a potential relationship between social status and FPS pronoun use in ASL by presenting the results of various quantitative analyses. I will also present the results of my work with the English data set used in Study 3 from Kacewicz et al. (2013). As discussed in Chapter 1, for the sake of transparency, as well as my belief in the importance of research reproducibility (e.g., Gawne & Styles, 2022), I have made my analyses and certain elements of the analyzed data sets available to anyone who would like to see them. In order to honor the privacy of those who participated in my study, especially considering the sensitivity of "social status" as a concept, I anonymized any data that could be used to identify non-transparent relative status markers or rates of FPS pronoun use. This means that the chat videos, annotation files, Background Questionnaire responses, and Post-Chat Survey responses are not included in the publicly available data. I promised participants that their anonymity would be preserved, with exceptions made only in accordance with their explicit permission (e.g., some participants granted permission to share screenshots and/or video clips in publications and presentations). Because I analyzed the data from Kacewicz et al. (2013) Study 3, I am also making available the relevant spreadsheets shared with me from their original data set, reshared with permission from Dr. James Pennebaker, along with the full details of how I analyzed it. All of this is available at the same link with my shared files mentioned earlier.[105] The spreadsheets are available as PDFs, .csv files, and links to Google Sheets.

I begin this chapter with my main analysis, that is, a look at the ASL data I collected through the lens I originally intended when I began this work. Through this lens, I seek to

[105] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

uncover whether there appears to be a relationship between social status and first-person singular pronoun use in the data I collected from 18 dyads of deaf American signers of ASL.

Next I present the results of my secondary analysis, which uses the English data from Study 3 in Kacewicz et al. (2013). I analyzed their data in a few ways, beginning with attempts to reproduce their results. I followed this with refinements to their methods that represent what I believe to be a more appropriate treatment of the data set. I then applied the refined methods to the ASL data I collected, and applied the approach from my primary analysis to the English data. This methodological triangulation is designed to provide a broader view of the potential relationship between status and FPS pronoun use across two co-existing languages.

I want to explicitly point out that the application of multiple methodologies can be used to "torture" the data into "confessing" nearly anything (a notion coined by Ronald Coase). Even in exploratory hypothesis-generating studies like this one, researchers should avoid running every analysis that comes to mind, because doing so only increases the likelihood of reporting a spurious correlation that appears meaningful. Running multiple analyses without good reason is inadvisable enough, and widespread enough, that it now goes by a few names, including *p*-hacking, data dredging, and significance chasing. A more responsible way of applying multiple methodologies, and the approach I take here, is to decide on logical methodologies that are designed to view the data in specific ways for specific reasons. Even this, though, does not guarantee that the results will not be spurious. See Silberzahn et al. (2018) for why we should not take results from any single study as the final word on a given matter.[106] When it comes to interpreting research and applying it to our lives, reproducibility of results—by different

[106] In short, Silberzahn et al. (2018) report results from a single data set (red cards issued in soccer games and skin tone of the carded players) analyzed by 29 different research teams. All pre-registered methods/analyses. Lots of variability in the results/conclusions. Fun!

researchers using different data sets—matters. I hope readers will keep this in mind when considering the results I present in this chapter.

Primary Data Analysis: ASL Data

The following sections present results for FPS pronoun frequencies, turn times, and FPS/min rates in my data. Each subsection includes descriptions and accompanying statistics (e.g., mean, range, S-value), along with data visualizations. When considering these results, recall that, while there were 24 unique human beings who participated in my study, some participated in two chats, making for 36 "participants" for the purposes of this analysis. As another reminder, the in-chat roles (*LowRole* and *HighRole*) are distinct from the triad-level designations (H, B, and L). For a refresher on these two concepts, please see the "Grouping Participants" section in Chapter 3.

Before getting into the numbers in the following subsections, I would like to briefly describe the ASL interaction data that the numbers come from. By and large, chat partners chatted amicably without any visible awkwardness or tension. Exceptions included things like lightheartedly challenging a sign choice (e.g., "that's your sign for that?") and discussing the unusualness of being put in a room and asked to chat without any agenda or prompts. In dyads who were meeting for the first time, all discussed to some degree whether they were from deaf families and/or whether they had attended a deaf school. Other topics varied, including where participants were from, the growing gig economy, the weather, sports, hangover pills, gun control, funny stories about old injuries, life advice, international travel, and interpreters vs. other accessibility options. Participants shared that they had enjoyed the chats, both on the Post-Chat Survey and in comments shared with me in person after filming.

While annotating, I periodically became aware that the participant whose language I was annotating had switched dominance at some point without my noticing. Because I had three separate annotation tiers for each participant's FPS pronouns (right hand, left hand, and two-handed), this meant frequent double-checking, especially at the beginning of my annotation process. I eventually learned not to assume handedness and to consciously consider which tier each FPS pronoun belonged on. Some participants exhibited right-hand dominance throughout their chats, while fewer appeared to be left-handed. I was initially surprised at the number of participants who switched dominance periodically over the course of their chats. My surprise was likely due to common prescriptive notions that handedness is a fixed trait. Some preliminary work has been done on dominance switching (also called dominance reversal, see Becker et al., 2023), but this is a research topic rife with possibility. The switches I observed in the ASL data set usually appeared to be motivated in the moment, but were often maintained longer than was necessary to simply fulfill the initial motivation. To provide an example, I looked at dominance switching in one participant for the first three minutes of one of their chats. The participant began the chat signing right-dominant, and at the three-minute mark, they were signing left-dominant. In the intervening three minutes, they switched dominance five times. Each switch happened either at the beginning of a new turn, during a depictive use of space, and/or as what appeared to be the result of a postural adjustment. An example of this participant switching dominance during a depictive use of space follows:

While signing left-dominant, the participant sets up a reference on their left to a non-present person. In describing that person, the participant leans back and to the right while looking toward the referenced space to the left, as though they are looking at the referenced *person. They then sign FEEL*[107] *with their right hand (dominance switch), followed by EQUAL*[108] *produced in the space between themselves and the referent to their left. They continue onto the next sentence signing right-dominant.*

The style and frequency of dominance switching in this three-minute clip are typical of this participant throughout the chat. They produced a total of 105 FPS pronouns during their chat: 58 produced with their left hand, 44 with their right hand, and 3 two-handed instances. In their other chat, their style of dominance switching was similar, but much less frequent. The total for that chat was 159 FPS: 5 produced with their left hand, 147 with their right hand, and 7 two-handed instances. Dominance switching seen in other participants' chats took on similar patterns and varied in frequency.

Next I will provide some FPS pronoun examples from the data in the form of images. These provide a sense of the phonological variation present in the data set as a whole, which is in line with previous work, as discussed in the "Frequency of FPS in ASL" section from Chapter 2. I had originally intended to make video clips of these examples publicly available, but in order to prevent unauthorized use of participant data, I have decided not to do so, though some example clips of IX_1 variation appear in the video of my dissertation defense (see Chapter 1 for access options). For demonstration forms of the seven FPS pronouns discussed below, see the "First-Person Singular (FPS) Pronouns in ASL" section in Chapter 2 or the "FPS tiers" section in Chapter 3.

[107] FEEL: https://aslsignbank.haskins.yale.edu/dictionary/gloss/624/

[108] EQUAL: https://aslsignbank.haskins.yale.edu/dictionary/gloss/835/

As described in the next section, the most frequent FPS pronoun in the data was IX_1. This was the most frequent FPS pronoun by far, with some chat partners even producing IX_1 simultaneously. Two examples of simultaneous FPS production can be seen in Figure 4.01.

Figure 4.01

Examples from the data of chat partners simultaneously producing IX_1

While I was looking for examples to screenshot for publication, I noticed that participants often had their eyes closed when producing IX_1. Some of these instances can be explained by the fact that signers must blink, and that IX_1 is frequent in use. Additionally, blinking is a well-known cue for sentence boundaries (Ormel & Crasborn, 2011). But time and again, I noticed the blink co-occurred with the signer's head either slightly dipped or jutted slightly forward. When I zoomed out to see these co-occurrences in context, it became obvious that these were frequently instances of the signer introducing a constructed action or constructed dialogue. I am unaware of any work that examines the frequency of various strategies for introducing/cueing a shift into constructed action or constructed dialogue, but any future work on this topic would do well to include a bundling of these three features (IX_1, blink, head dip/jut) as a strategy to investigate. Figure 4.02 below shows two examples of these three features co-occurring to introduce a constructed dialogue or constructed action.

Figure 4.02

Examples from the data of IX_1 produced simultaneously with a blink and a slight dipping/jutting out of the head during the introduction of a constructed action or constructed dialogue

The phonological variation seen in the forms shown in the previous figures (as well as the forms in the "FPS tiers" section in Chapter 3) is typical of the range of forms for IX_1 throughout the data set. Forms that most closely resemble the citation form tended to appear more often when IX_1 was being produced in initial or final syntactic positions (i.e., when IX_1 was the first or last word in a sentence). Figure 4.03 below, duplicated from Figure 2.02 in Chapter 2, provides examples of IX_1 variation in the data.

Figure 4.03

Examples from the data showing some of the phonological variation seen with IX_1

Like IX_1, POSS_1 is produced with a wide range of variability in phonological form. Figure 4.04 below shows two examples that both include a wider finger splay than is seen in demonstration forms, but is nonetheless likely more common, due to ease of articulation. The first image also shows a degree of wrist bend that is common in my data set, which leads to a minimized or complete lack of contact between the signer's palm and their chest. This tended to appear more often when POSS_1 was being produced in syntactic positions other than initial or final. The wrist bend is less pronounced in the second image, though it is hard to tell whether there is full contact between the signer's palm and their chest.

Figure 4.04

Examples from the data showing common configurations of POSS_1

Figure 4.05 below provides examples of other phonological variations of POSS_1 in the data. The first image shows POSS_1 produced prior to MOTHER.[109] The form in this image appears to have been the result of assimilation, as the orientation and active articulator of POSS_1 are more similar to that of MOTHER than that seen in demonstration forms of POSS_1. I saw this type of assimilation frequently with instances of POSS_1 in the data set. The second image in Figure 4.05 shows an instance of POSS_1 exhibiting phonological variation that appears to be a result of the signer's physical posture, rather than assimilation. This type of variation is also common in the data set.

[109] MOTHERstr: https://aslsignbank.haskins.yale.edu/dictionary/gloss/1590/

Figure 4.05

Examples from the data showing phonological variations of POSS_1

While the FPS pronoun SELF_1 is less frequent in the data set, I also saw variation in these forms. Figure 4.06 below shows four examples of SELF_1, each with a slightly different handshape and/or location. Two of the examples contain a second camera angle that shows the signer's handshape (i.e., configuration) in greater detail and better shows the active articulator (i.e., where the hand makes contact with the chest). All of these forms are typical of those in the data set as a whole.

Figure 4.06

Examples from the data showing phonological variations of SELF_1

Figure 4.07 below provides examples from the data of ONLY-SELF_1. This was used most frequently in the sense of someone being the only deaf person in a group (e.g., family), though there were other uses as well. The minor phonological variations seen in Figure 4.07 are typical of those in the data set as a whole.

Figure 4.07

Examples from the data of ONLY-SELF_1

Figure 4.08 below provides examples from the data of HONORIFIC_1. This was used most frequently in the sense of "who I am as a person," though there were other uses as well. Note that the two forms in Figure 4.08 show typical types of variation seen in the data set, with some exhibiting longer holds at the beginning (as in the first example below) and others exhibiting longer holds at the end (as in the second example below).

Figure 4.08

Examples from the data of HONORIFIC_1

The final two FPS pronoun types are THUMB_1 and I. I will not be sharing any examples of THUMB_1 because there are no instances of this word in the data set. I will not be sharing any examples of the FPS pronoun I for a different reason. As discussed in Chapter 2, I do not subscribe to prescriptive notions of "pure ASL." That said, there are those who would see an image of a person producing the initialized word I and assume that their language as a whole was somehow "less ASL." For this reason, I have chosen not to include images of the few instances of I that appear in the data. Suffice to say, at least one instance was used sarcastically, but the rest appeared to be subject pronouns, used unironically.

FPS pronoun frequencies.

I annotated a total of 410 minutes of turn time. In this data set, IX_1 was the most frequently used FPS pronoun by far, with 4,320 instances representing 87.6% of the total FPS pronouns produced. The next most frequent was POSS_1, representing 10.9% of the total with 537 instances. These two words made up a combined total of 98.5% of the FPS pronouns used, and they were also the only FPS pronouns to be used by all participants. Of the 18 chats, constituting 36 participants, 12 participants used only IX_1 and POSS_1, and no other FPS pronouns. Recall from Chapter 3 that there are many instances in my data set where it is unclear whether the FPS pronoun being produced is an IX_1 or a POSS_1. These instances do not affect the overall FPS pronoun totals, only the individual word frequencies.

Other than IX_1 and POSS_1, frequencies of FPS pronouns were minimal, with SELF_1, ONLY-SELF_1, HONORIFIC_1, and I all being used infrequently and by only a small percentage of people during their chat sessions. As previously mentioned, there were no instances of THUMB_1 in my data set. This tracks with my anecdotal impressions that third-person pronouns more frequently take on the thumb handshape than do first-person pronouns (i.e., the handshape in THUMB_1, but with the thumb directed towards someone or something indicated to the ipsilateral side of the signer). Figure 4.09 below shows frequency data for the entire data set.

Figure 4.09

An overview of FPS pronoun frequency across all 18 chats (ASL data)

FPS pronoun	FPS pronouns used		Average use, per participant		Range across participants		How many participants used it	
	number	% of total FPS	mean	median	min	max	number	%
IX_1	4320	87.6%	120.0	113.5	44	245	36	100%
POSS_1	537	10.9%	14.9	12.5	2	55	36	100%
SELF_1	49	1.0%	1.4	1.0	0	6	20	55.6%
ONLY-SELF_1	15	0.3%	0.4	0	0	7	6	16.7%
HONORIFIC_1	6	0.1%	0.2	0	0	2	5	13.9%
I	4	0.1%	0.1	0	0	2	3	8.3%
THUMB_1	0	-	-	-	-	-	-	-
All FPS Pronouns	**4931**	**100%**	**137.0**	**130.5**	**47**	**311**		

There are three tables in Figure 4.10 below. I have chosen to present them together in one figure to make it easier for the reader to compare the contents of all three without the visual/cognitive disruption of scrolling. Even for the rare soul who may read this printed on physical pieces of paper, the visual scanning of the triple-figure will be more efficient. The first table shows how many FPS pronouns were used by each participant in each chat, along with the simple differences between chat partners' totals (i.e., *LowRole*'s FPS count minus *HighRole*'s FPS count). These numbers do not take into consideration the differences in how much each person said during the course of the chat, only the number of FPS pronouns used, hence the designation of “frequency,” as opposed to “rate.” The four dyads marked with an asterisk (e.g., 1HB*) are those described in the “Dyads warranting special consideration” section in Chapter 3.

The other two tables in Figure 4.10 provide descriptive and inferential statistics for the FPS frequency data. The top-right table shows measures of central tendency for participants’ use of pronouns. Note that the “Simple difference” column in this table functions differently than the “Simple difference” column in the table on the left. This top-right table contains summary statistics across all dyads using the information from the table on the left. As an example, the reported simple difference of 20 in the top-right table represents the mean of the simple differences shown in the table on the left, not the simple difference between the means reported in the top-right table. This is easier to see with the reported value of 5 representing the median of the simple differences from the table on the left, rather than the simple difference of the medians reported in the top-right table. The final table in Figure 4.10 contains results from relevant statistical tests. Because the data were not normally distributed, I calculated statistical significance using the nonparametric Mann-Whitney/Wilcoxon Signed-Ranks test.

The average FPS frequency was greater for *LowRole* than *HighRole* when averages were calculated for each role-group. However, at the dyad level there is a good deal of variability in direction. This can be seen in the relatively balanced mix of positive and negative values in the simple difference column in the first table; recall that a positive discrepancy represents a *LowRole* participant using more FPS than *HighRole*. Also, keep in mind that the FPS frequency values can not be used to address the aim of my research, because they do not account for the turn time used by each participant. Instead, these values are used to calculate my dependent variable, FPS/min.

Figure 4.10

Summary data for FPS pronoun frequencies, counted for each participant in each chat, as well as descriptive and inferential statistics (ASL data)

Dyad	FPS (frequency)		*Simple difference*
	LowRole	HighRole	
1HL	159	154	*5*
1HB*	73	176	*-103*
1BL	105	101	*4*
2HL	101	81	*20*
3HL	311	47	*263*
3HB	86	130	*-44*
3BL*	199	73	*126*
4HL*	70	81	*-11*
5HB	167	99	*68*
6BL	105	169	*-64*
7HB*	168	187	*-19*
8HL	215	103	*112*
8HB	263	230	*33*
8BL	161	105	*56*
9HL	70	177	*-107*
10HL	153	107	*46*
10HB	100	131	*-31*
10BL	135	139	*-4*

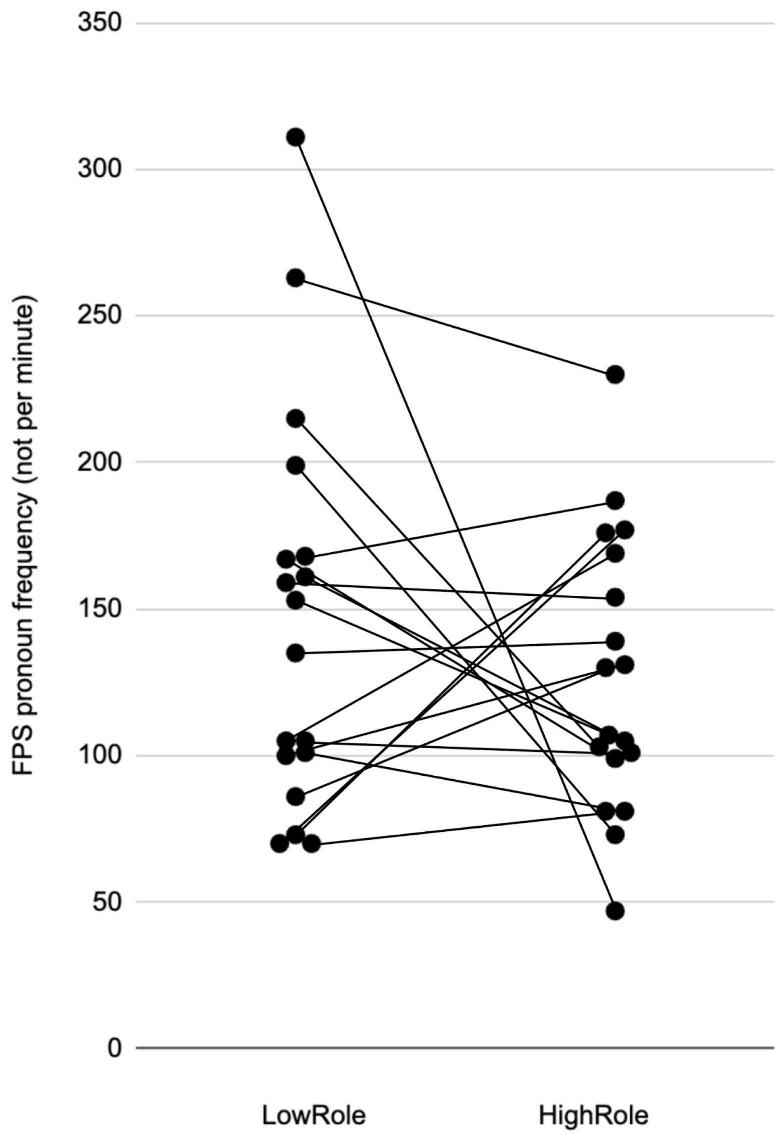
	FPS (frequency)		*Simple difference*
	LowRole	HighRole	
mean	147	127	*20*
median	144	119	*5*
range	241	183	*371*
min	70	47	*-107*
max	311	230	*264*
SD	68	47	*88*

	FPS (frequency)
p-value	0.446
S-value	1.165
Effect size (*r*)	-0.180

Seeing the individual FPS usage of each participant, as presented in Figure 4.10 above, is useful in terms of seeing specific numbers. Another useful way to visually conceptualize this information is by plotting the values on a graph, as in Figure 4.11 below. Lines connect chat partners to provide a sense of how the dyads patterned in their FPS usage, prior to the calculation of FPS/min rates.

Figure 4.11

Dyad-level frequencies of FPS pronouns used during chats, with dyads connected by lines (ASL data)



Turn times.

Similar to Figure 4.10 from the previous section, Figure 4.12 below contains three tables. The first shows how much turn time was used by each participant in each chat, reported in minutes (e.g., a turn time of 8.46 represents 8 minutes and 27.6 seconds), along with simple differences between chat partners' usage (i.e., *LowRole*'s turn time minus *HighRole*'s turn time). Note that some apparent inconsistencies between the reported turn times and simple differences are actually rounding errors, as a result of rounding these values to two decimal places for reporting purposes.[110] Also similar to the previous section, the other two tables in Figure 4.12 provide descriptive and inferential statistics for the turn time data. Because the data were not normally distributed, I calculated statistical significance using the nonparametric Mann-Whitney/Wilcoxon Signed-Ranks test. Again, note that the "Simple difference" column in the top-right table functions differently than the "Simple difference" column in the table on the left. The four dyads marked with an asterisk (e.g., 1HB*) are the ones described in the "Dyads warranting special consideration" section in Chapter 3.

[110] To use dyad 1HL as an example, *LowRole*'s turn time of 8.46 minutes minus *HighRole*'s time of 12.62 minutes equals -4.16, not the value in the table, -4.17. However, the same calculation with turn times carried out to three decimal places is 8.456 – 12.624 = -4.168, and this result rounds to -4.17.

Figure 4.12

Summary data for total turn time, counted for each participant in each chat, as well as descriptive and inferential statistics (ASL data)

Dyad	Turn time (in minutes)		*Simple difference*
	LowRole	HighRole	
1HL	8.46	12.62	*-4.17*
1HB*	6.15	15.61	*-9.46*
1BL	10.58	12.08	*-1.50*
2HL	13.30	9.44	*3.86*
3HL	15.65	8.75	*6.90*
3HB	9.93	13.08	*-3.14*
3BL*	13.44	8.09	*5.35*
4HL*	7.75	14.36	*-6.61*
5HB	11.40	9.98	*1.43*
6BL	10.20	13.27	*-3.07*
7HB*	14.26	13.08	*1.18*
8HL	16.68	8.72	*7.96*
8HB	12.55	10.99	*1.56*
8BL	14.43	10.47	*3.96*
9HL	7.48	14.46	*-6.98*
10HL	11.02	11.83	*-0.81*
10HB	9.59	11.81	*-2.22*
10BL	9.29	9.56	*-0.27*

	Turn time (in minutes)		*Simple difference*
	LowRole	HighRole	
mean	11.23	11.57	*-0.34*
median	10.80	11.82	*-0.54*
range	10.52	7.52	*17.41*
min	6.15	8.09	*-9.46*
max	16.68	15.61	*7.96*
SD	2.96	2.20	*4.85*

	Turn time (in minutes)
p-value	0.799
S-value	0.324
Effect size (*r*)	0.060

Similar to the information presented above in Figure 4.11 for FPS frequency, Figure 4.13 below shows the dyad-level view of turn time, shown in minutes. Again, lines connect chat partners to provide a sense of how the dyads patterned in their total amount of turn time. This figure shows that the dyads were split roughly in half in terms of which participant used more turn time. In eight of the 18 dyads, the *LowRole* participant used more turn time, a scenario that appears in Figure 4.13 as a line that is higher on the left and lower on the right. One dyad had very similar turn times between *LowRole* and *HighRole*: the difference between participant turn times for 10BL was only 16.3 seconds, making the line between them appear nearly horizontal.

The remaining nine dyads are connected by lines that are lower on the left and higher on the right, indicating that the *HighRole* person used more turn time.

Figure 4.13

Dyad-level turn times used during chats, with dyads connected by lines (ASL data)

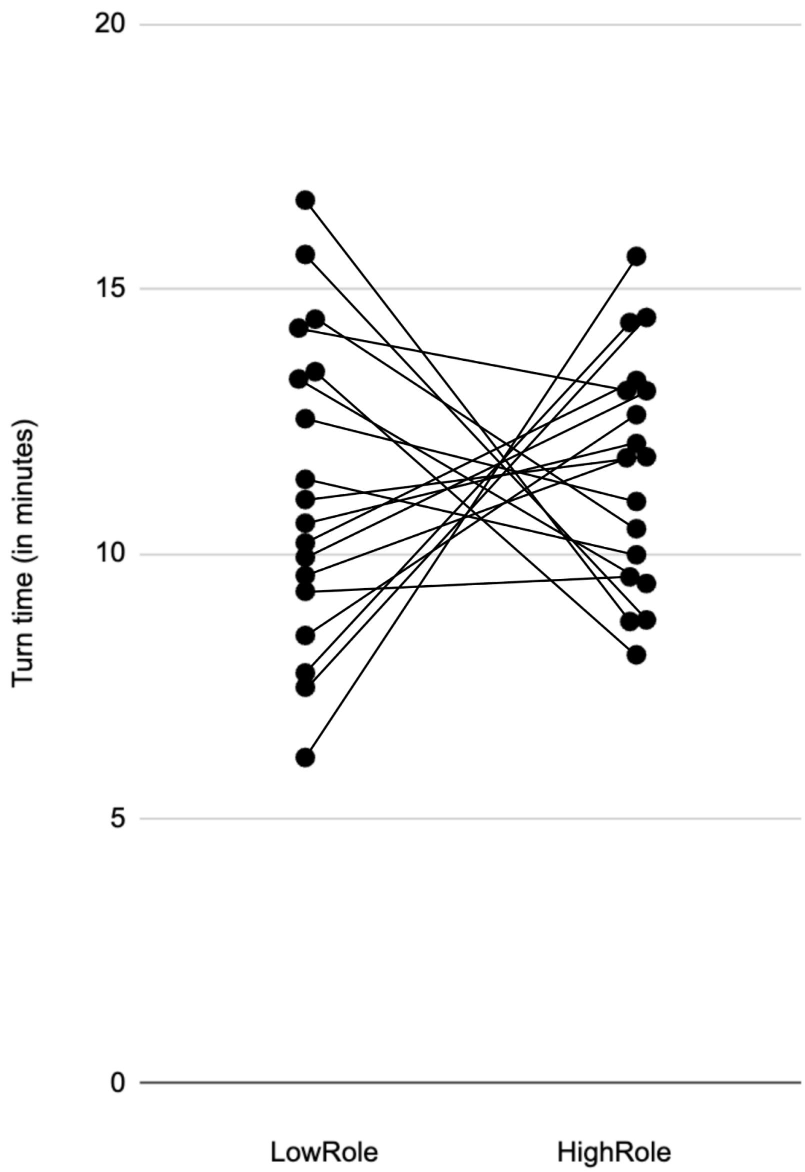


Figure 4.14 below reports information about how often each participant was holding a turn during their chat. Similar to Figures 4.10 and 4.12 above, Figure 4.14 below contains three

tables. The first shows the percentage of turn time used by each participant in each chat, along with simple differences between chat partners' usage (i.e., *LowRole*'s turn percentage minus *HighRole*'s). Again, the remaining two tables in Figure 4.14 provide descriptive and inferential statistics for the turn time data, the "Simple difference" column in the top-right table functions differently than the "Simple difference" column in the table on the left, and the four dyads marked with an asterisk (e.g., 1HB*) are the ones described in the "Dyads warranting special consideration" section in Chapter 3. Because the data were not normally distributed, I calculated statistical significance using the nonparametric Mann-Whitney/Wilcoxon Signed-Ranks test.

Chats varied slightly in length, so I calculated percentages using each chat's actual length, rather than the planned length of 20 minutes. Actual chat lengths ranged between 18.83 minutes to 22.70 minutes. The mean chat length was 20.63 minutes and the median chat length was 20.30 minutes.

Deviating from the previous figures, Figure 4.14 also includes a column for total turn time percentage. This is calculated as the sum of both participants' turn percentages. Note that the lowest total is 100%, occurring for only a single chat. This indicates that, even with occasional brief lulls between turns, turn overlap is prevalent in all 18 chats.

Figure 4.14

Summary data for participant turn time percentages (i.e., the percentage of the total chat time that each participant's' turn time represents), as well as descriptive and inferential statistics (ASL data)

Dyad	Turn time (% of chat)			*Simple difference*
	LowRole	HighRole	*Total*	
1HL	44%	65%	*109%*	*-21%*
1HB*	30%	77%	*108%*	*-47%*
1BL	52%	59%	*110%*	*-7%*
2HL	65%	46%	*111%*	*19%*
3HL	69%	39%	*108%*	*31%*
3HB	48%	64%	*112%*	*-15%*
3BL*	66%	40%	*105%*	*26%*
4HL*	39%	72%	*111%*	*-33%*
5HB	56%	49%	*106%*	*7%*
6BL	50%	64%	*114%*	*-15%*
7HB*	63%	58%	*120%*	*5%*
8HL	74%	39%	*113%*	*35%*
8HB	62%	55%	*117%*	*8%*
8BL	71%	52%	*123%*	*20%*
9HL	38%	73%	*111%*	*-35%*
10HL	50%	53%	*103%*	*-4%*
10HB	48%	59%	*106%*	*-11%*
10BL	49%	51%	*100%*	*-1%*

	Turn time (% of chat)			*Simple difference*
	LowRole	HighRole	*Total*	
mean	54%	56%	*110%*	*-2%*
median	51%	56%	*110%*	*-3%*
range	44%	39%	*23%*	*82%*
min	30%	39%	*100%*	*-47%*
max	74%	77%	*123%*	*35%*
SD	13%	12%	*6%*	*23%*

	Turn time (% of chat)
p-value	0.744
S-value	0.427
Effect size (*r*)	0.077

FPS/min rates.

This section reports results regarding the dependent variable in my research, rates of FPS pronouns used per minute, or FPS/min. As seen in the figures below, there was a fairly wide range of usage rates. See the exploratory section titled "FPS/min stabilization" in Chapter 5 for an analysis of the variability of these rates over the course of a single chat.

Similar to the corresponding figures above, Figure 4.15 below contains three tables. The first shows the rate of FPS/min for each participant in each chat, along with simple differences between chat partners' rates (i.e., *LowRole*'s rate minus *HighRole*'s). As before, the other two tables in Figure 4.15 provide descriptive and inferential statistics, the “Simple difference” column in the top-right table functions differently than the “Simple difference” column in the table on the left, and the four dyads marked with an asterisk (e.g., 1HB*) are the ones described in the “Dyads warranting special consideration” section in Chapter 3. Because the data were not normally distributed, I calculated statistical significance using the nonparametric Mann-Whitney/Wilcoxon Signed-Ranks test.

Compared to the previous two sections on FPS pronoun frequencies and turn times, there is less variability within the simple difference column in the table on the left. Similar to the “Turn times” section above, the apparent inconsistencies between the reported FPS/min rates and simple differences (e.g., 10BL) are actually rounding errors, as a result of rounding these values to two decimal places for reporting purposes.

Figure 4.15

Summary data for FPS/min rates, calculated for each participant in each chat, as well as descriptive and inferential statistics (ASL data)

Dyad	FPS/min (rate)		*Simple difference*
	LowRole	HighRole	
1HL	18.80	12.20	*6.61*
1HB*	11.86	11.27	*0.59*
1BL	9.93	8.36	*1.57*
2HL	7.59	8.58	*-0.99*
3HL	19.87	5.37	*14.50*
3HB	8.66	9.94	*-1.28*
3BL*	14.81	9.02	*5.79*
4HL*	9.03	5.64	*3.39*
5HB	14.64	9.92	*4.72*
6BL	10.29	12.73	*-2.44*
7HB*	11.78	14.30	*-2.52*
8HL	12.89	11.81	*1.08*
8HB	20.96	20.94	*0.02*
8BL	11.16	10.03	*1.13*
9HL	9.36	12.24	*-2.88*
10HL	13.88	9.05	*4.83*
10HB	10.43	11.09	*-0.66*
10BL	14.53	14.54	*0.00*

	FPS/min (rate)		*Simple difference*
	LowRole	HighRole	
mean	12.81	10.95	*1.86*
median	11.82	10.56	*0.84*
range	13.36	15.56	*17.38*
min	7.59	5.37	*-2.88*
max	20.96	20.94	*14.50*
SD	3.91	3.54	*4.28*

	FPS/min (rate)
p-value	0.142
S-value	2.821
Effect size (*r*)	-0.347

Recall from Chapter 3 that the S-value of 2.8 in Figure 4.15 equates to the level of surprise associated with correctly predicting heads on a coin toss an average of 2.8 times in a row. This is not a very surprising result. In other words, for the question of whether *HighRole* and *LowRole* participants used different rates of FPS pronouns, this data set does not provide enough evidence to say there is a pattern at work here beyond chance.

Similar to the information presented above in Figures 4.11 and 4.13, Figure 4.16 below shows the dyad-level view of FPS/min rates. As before, lines connect chat partners to provide a sense of how the dyads patterned in their use of FPS/min. Because the FPS/min data represents the dependent variable in my study, the directionality of the dyad-level discrepancies is relevant to the aim of my research in a way that it was not for the previous sections. For this reason, the lines connecting chat partners in Figure 4.16 are coded for directionality: solid lines connect chat partners whose results supported the prediction (i.e., the *LowRole* participant used a higher rate of FPS pronouns than did the *HighRole* participant), and dashed lines connect chat partners whose results contradicted the prediction. Although two of the lines appear to be horizontal, no dyads have equal FPS/min rates once the rates are carried out to three decimal places. The four dyads marked with an asterisk (e.g., 1HB*) in the previous figures (i.e., those described in the "Dyads warranting special consideration" section in Chapter 3) are marked in Figure 4.16 below by having their data points indicated with stars instead of dots. Figure 4.16 shows a good deal of variation, with results from 11 of the 18 dyads (or 61%) of the dyads supporting the prediction.

Figure 4.16

Dyad-level rates of FPS/min used during chats, with dyads connected by lines (ASL data)

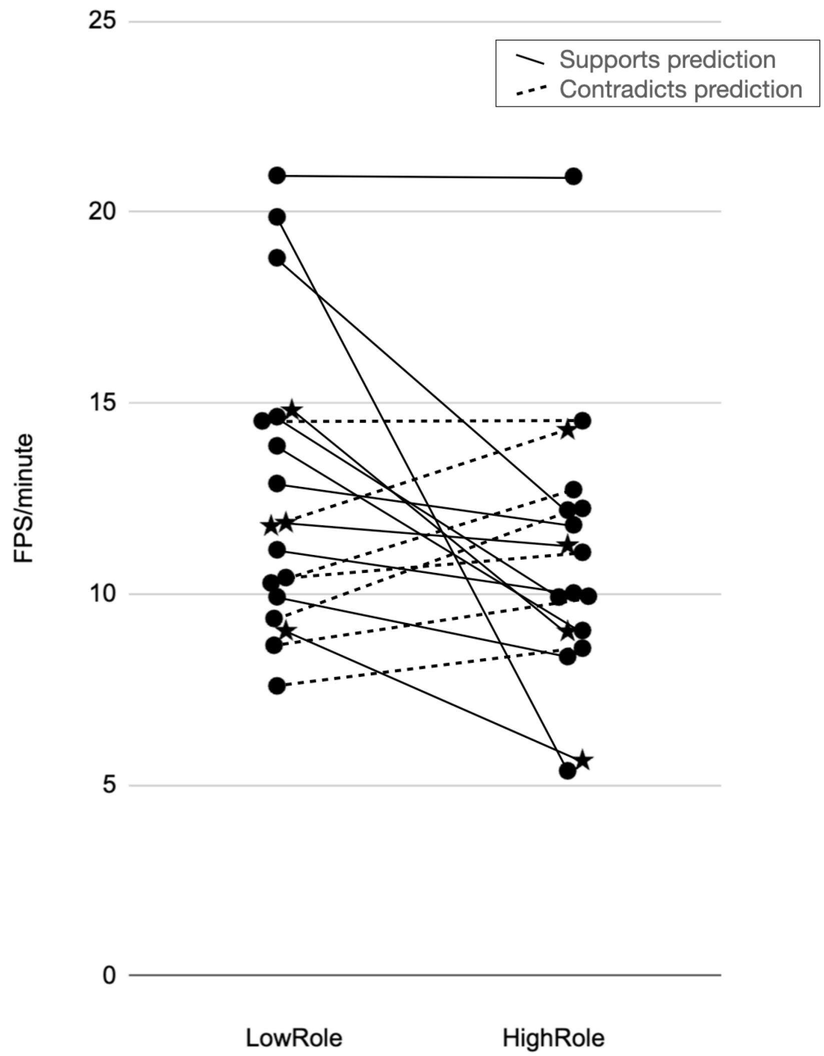
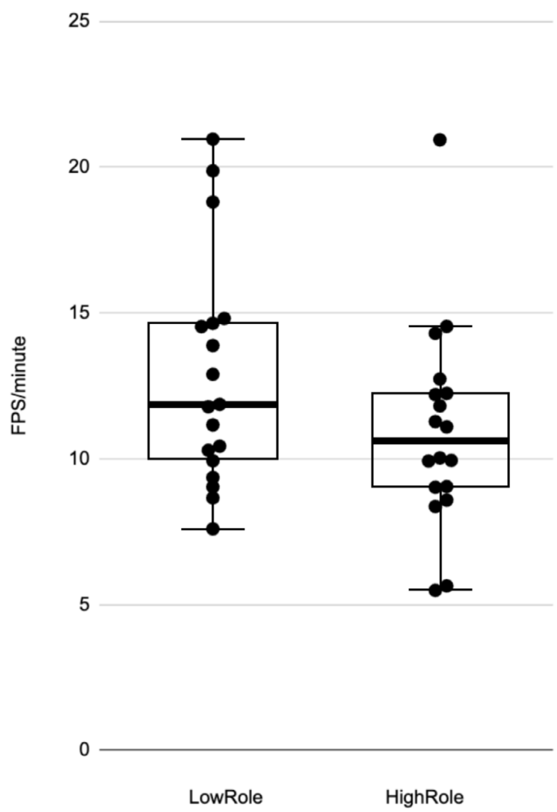


Figure 4.17 below shows FPS/min rates for *LowRole* and *HighRole* participants, unpaired, using box plots to show the group-level distribution of my dependent variable in greater detail. A thick line within each box shows the median for each group, upper and lower box limits indicate the 25th and 75th percentiles, whiskers extend 1.5 times the interquartile

range from the 25th and 75th percentiles, and the one data point not within the boxes or whiskers can be considered the single outlier.

Figure 4.17

Box plots of LowRole *and* HighRole *FPS/min rates (ASL data)*



In Figure 4.18 below, I have listed the dyad-level FPS/min percentage differences in the context of the individual FPS/min rates and simple differences. Chapter 3 includes an explanation of each of the percentage difference calculations, along with my reasons for including all four here. As a reminder, here are examples of how the percentage difference results can be read (note that the results from the two symmetrized percentage difference formulas can be read the same way):

- *LowRole>/<HighRole*:

  "*LowRole* used X% more/fewer FPS/min than *HighRole*"

- *HighRole>/<LowRole*:

  "*HighRole* used X% more/fewer FPS/min than *LowRole*"

- *Diff/sum*:

  "Between *LowRole*'s and *HighRole*'s FPS/min rates, there was a difference of X%"

- *Diff/mean*:

  "Between *LowRole*'s and *HighRole*'s FPS/min rates, there was a difference of X%"

Figure 4.18

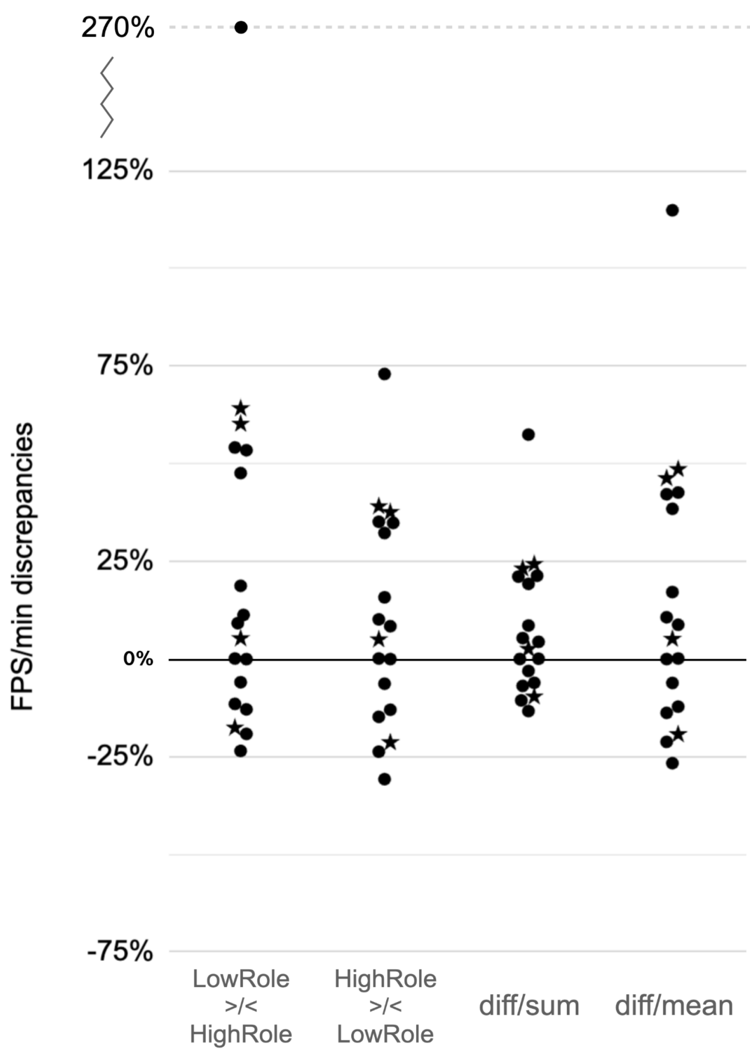
Dyad-level FPS/min percentage differences, alongside previously reported individual FPS/min rates and simple differences for context (ASL data)

Dyad	FPS/min (rate)		Simple difference	LowRole >/< HighRole	HighRole >/< LowRole	diff/sum	diff/mean
	LowRole	HighRole					
1HL	18.80	12.20	6.61	54%	-35%	21%	43%
1HB*	11.86	11.27	0.59	5%	-5%	3%	5%
1BL	9.93	8.36	1.57	19%	-16%	9%	17%
2HL	7.59	8.58	-0.99	-12%	13%	-6%	-12%
3HL	19.87	5.37	14.50	270%	-73%	57%	115%
3HB	8.66	9.94	-1.28	-13%	15%	-7%	-14%
3BL*	14.81	9.02	5.79	64%	-39%	24%	49%
4HL*	9.03	5.64	3.39	60%	-38%	23%	46%
5HB	14.64	9.92	4.72	48%	-32%	19%	38%
6BL	10.29	12.73	-2.44	-19%	24%	-11%	-21%
7HB*	11.78	14.30	-2.52	-18%	21%	-10%	-19%
8HL	12.89	11.81	1.08	9%	-8%	4%	9%
8HB	20.96	20.94	0.02	0%	0%	0%	0%
8BL	11.16	10.03	1.13	11%	-10%	5%	11%
9HL	9.36	12.24	-2.88	-24%	31%	-13%	-27%
10HL	13.88	9.05	4.83	53%	-35%	21%	42%
10HB	10.43	11.09	-0.66	-6%	6%	-3%	-6%
10BL	14.53	14.54	0.00	0%	0%	0%	0%

In Figure 4.19 below, the percentage differences from Figure 4.18 above are displayed in four scatterplots presented side by side. For those with a keen eye, I would like to point out that I inverted the *HighRole>/<LowRole* numbers for consistency of direction (i.e., the outlier is at the top for all 4). In this vein, when it came time to choose how to calculate the differences for *diff/sum* and *diff/mean*, I chose the subtraction order that preserved the same consistency of direction. See the "Comparing FPS/min rates" section in Chapter 3 for more on this.

Figure 4.19

Dyad-level discrepancies in FPS/min rates, presented as four side-by-side scatterplots showing results from four different percentage difference calculations (ASL data)



Triads.

In this section, I present FPS/min results from the subgroup of dyads who were part of a complete triad. Individuals in these complete triads each participated in two chats, sometimes with different in-chat roles, and sometimes with the same role. It has been some time since I introduced the in-chat roles (*LowRole* and *HighRole*) as distinct from the triad-level designations (H, B, and L). For a refresher, please see the "Grouping Participants" section in Chapter 3.

Figure 4.20 shows some of the same information from Figure 4.18 in the "FPS/min rates" section above, but this figure shows results only from dyads who were part of complete triads. Emphasis is placed on the FPS/min rates from B participants for easier visual comparison. As before, dyads marked with an asterisk (e.g., 1HB*) are described in the "Dyads warranting special consideration" section in Chapter 3. Figure 4.20 does not include any group-level descriptive or inferential statistics, only dyad-level differences. See "Triads" in Chapter 3 for an explanation of this decision.

Figure 4.20

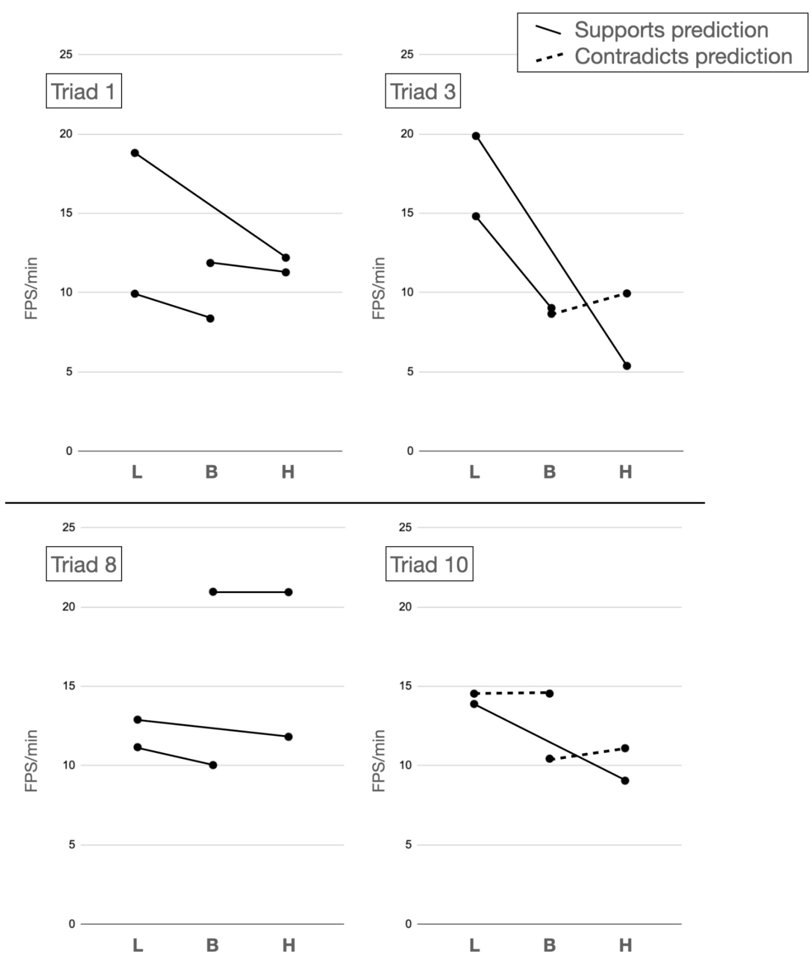
Results from complete triads: Individual FPS/min rates within each dyad (emphasis on rates from B participants), along with dyad-level simple differences and percentage differences (ASL data)

Triad	Dyad	FPS/min (rate)		Simple difference	LowRole vs. HighRole	HighRole vs. LowRole	Diff / Sum	Diff / Mean
		LowRole	HighRole					
1	HL	18.80	12.20	*6.61*	*54%*	*-35%*	*21%*	*43%*
	HB*	**11.86**	11.27	*0.59*	*5%*	*-5%*	*3%*	*5%*
	BL	9.93	**8.36**	*1.57*	*19%*	*-16%*	*9%*	*17%*
3	HL	19.87	5.37	*14.50*	*270%*	*-73%*	*57%*	*115%*
	HB	**8.66**	9.94	*-1.28*	*-13%*	*15%*	*-7%*	*-14%*
	BL*	14.81	**9.02**	*5.79*	*64%*	*-39%*	*24%*	*49%*
8	HL	12.89	11.81	*1.08*	*9%*	*-8%*	*4%*	*9%*
	HB	**20.96**	20.94	*0.02*	*0%*	*0%*	*0%*	*0%*
	BL	11.16	**10.03**	*1.13*	*11%*	*-10%*	*5%*	*11%*
10	HL	13.88	9.05	*4.83*	*53%*	*-35%*	*21%*	*42%*
	HB	**10.43**	11.09	*-0.66*	*-6%*	*6%*	*-3%*	*-6%*
	BL	14.53	**14.54**	*0.00*	*0%*	*0%*	*0%*	*0%*

Figure 4.21 below presents data visualizations for FPS/min rates within each of the four complete triads. As with the similar graphs in the FPS/min section above, the lines connecting chat partners are coded for directionality: solid lines connect chat partners whose results supported the prediction (i.e., the *LowRole* participant used a higher rate of FPS pronouns than did the *HighRole* participant), and dashed lines connect the few chat partners whose results contradict the prediction (3 out of 12). Although two of the lines appear to be horizontal, no dyads have equal FPS/min rates once the rates are carried out to two decimal places.

Figure 4.21

Dyad-level rates of FPS/min used by the four complete triads, with dyads connected by lines (ASL data)



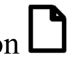
Secondary Data Analysis: English data and ASL data

The following sections present results from my secondary analysis, with the goal of using these results to compare the ASL data I collected to existing English data. The English data set is that from Kacewicz et al. (2013) Study 3, and is shared in my Figshare project folder,[111] with permission from Dr. James Pennebaker, along with my procedures for analyzing it. Chapter 3 contains an explanation for each of the steps in my secondary analysis.

Computational reproduction: Using the English data from Study 3 in Kacewicz et al. (2013) to reproduce their reported results.

In Chapter 3, I described some of the obstacles I encountered while attempting to reproduce the FPS results reported by Kacewicz et al. (2013) for Study 3. In this section, I will describe the results of my attempts. The two obstacles that I described in Chapter 3 that I will refer to here are the number of survey questions used to determine participant-level status (*2Q* vs. *4Q*) and the method used for calculating dyad-level status (*Lo minus Hi* vs. *P1 minus P2*). For more details on these methods, please see the "Computational reproduction" section in Chapter 3. I first attempted to reproduce the results reported by Kacewicz et al. (2013) by following the procedures laid out in their paper. These procedures are summarized as *2Q* and *Lo minus Hi*. This attempt was unsuccessful. Compared to the target results (i.e., those reported in their paper), this attempt produced a much smaller correlation coefficient that did not reach statistical significance, with a correspondingly smaller effect size. In the figures below, this attempt is indicated with a dialogue bubble icon  for easy recognition (i.e., "what they said they did").

[111] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

My second attempt followed the procedures implied by the organization of their data set that was shared with me. These procedures are summarized as *4Q* and *P1 minus P2*, and this attempt is indicated in the figures below with a file icon  (i.e., "according to the file they shared"). I report two sets of results under this attempt, one calculating the correlation coefficient using Pearson's *r* and one using Spearman's *rho*. I first used Spearman's *rho*, based on the distribution of the data (i.e., the data were not normally distributed), and the results from this calculation are indicated with , accordingly. These results were much closer to the target results than were the  results. Noticing that the  results included a correlation coefficient that was only somewhat larger than the target results (-.39 versus -.30), I was reminded that not everyone is diligent about making the distinction between Pearson's *r* and Spearman's *rho*. Spearman's *rho* is more robust against skewness and outliers, and while it is the more appropriate calculation for this data set (due to the non-normal distribution), I wondered if trying the more commonly-used Pearson's *r* calculation might provide a closer match to their lower reported correlation. To test this idea, I re-ran the  attempt, this time using Pearson's *r*. The results from this attempt are indicated in the figures below with . Carried out to three decimal places, the resulting correlation was -.293, which rounds to -.29, making it extremely close to the target results (-.30), while not quite hitting the mark. The associated *p*-value of .04 fits within the range they reported. The resulting effect size was also extremely close to their reported results, at *d* = -0.614 (compared to their *d = -0.62*). I tried recalculating the effect size with the truncated correlation result, a strategy borrowed from the  attempt discussed below, but this produced an effect size of -0.606, which represented a further deviation from the target results.[112] Thus, neither of the  attempts successfully reproduced the results reported by Kacewicz et al (2013).

My next two attempts took a mix-and-match approach, combining elements from the methods described in their paper and from the methods implied by the organization of their data set. One of the mix-and-match approaches was to use the number of questions from the methods implied in their data set (*4Q*) and the dyad-level calculations described in their paper (*Lo minus Hi*). This combination produced results that were even further from their reported results than those from the previous attempts. In the figures below, this attempt is indicated with a puzzle piece icon  (i.e., “an unsuccessful mix-and-match attempt”) .

Finally, my other mix-and-match approach combined the number of questions described in their paper (*2Q*) with the dyad-level calculations implied in their data set (*P1 minus P2*). This was the approach that most closely reproduced the results reported by Kacewicz et al. (2013). Because it was the most successful attempt at reproduction, it is represented in the figures below with a star icon, ★ (i.e., “the most successful attempt”). The difference of 0.01 between this attempt’s effect size (d = -0.63) and their reported effect size (d = -0.62) appears to be a rounding error, as their reported result can be achieved by redoing the effect size calculation using the truncated correlation coefficient, rather than the more precise value. Note that this attempt used Pearson’s r for the correlation coefficient, as it was normally distributed. In fact, it was the only combination of calculations that resulted in a data set that was normally distributed.

[112] There is one additional 0.01 discrepancy in this attempt, that of the correlation coefficient. It is possible that implementing other strategies within this attempt may more successfully reproduce their reported values, such as truncating further back into the calculations (e.g., truncating each participant’s rate) or experimenting with different theories for the possible exclusion of dyads as covered in Chapter 3 (though this would create a new discrepancy with their reported *n*). I did not attempt any of these, but want to acknowledge there are more avenues to explore.

The results from all of these attempts can be seen in Figure 4.22. They are presented in the context of the target results originally reported by Kacewicz et al. (2013), which are seen in the gray column. Bold numbers indicate reproduced results that match the target results. Note that the correlation coefficients are noted as being either Pearson's *r* or Spearman's *rho*, in accordance with statistical best practices (Lang & Altman, 2016).

Figure 4.22

Summary of my computational reproduction attempts (English data)

	Reported by Kacewicz et al. (2013)	**Computational reproduction** (*n* = 50 dyads)				
		2Q		**4Q**		
		Lo minus Hi	P1 minus P2	Lo minus Hi	P1 minus P2	
		Spearman	Pearson	Spearman	Spearman	Pearson
correlation coefficient	-0.30	-0.14	**-0.30**	-0.06	-0.39	-0.29
***p*-value**	0.01 < *p* ≤ 0.05	0.34	**0.03**	0.66	**0.01**	**0.04**
effect size (Cohen's *d*)	-0.62	-0.28	**-0.63***	-0.13	-0.85	-0.61
		As described in their paper	*Mix & Match V1*	*Mix & Match V2*	*As implied in their dataset*	

*Discrepancy of 0.01 appears to be the result of a rounding error

Refining their approach: Using the English data from Study 3 in Kacewicz et al. (2013), revising their design.

For the next step of my secondary analysis, I addressed the methodological issue of including dyads for whom it would be impossible to assume a hierarchical status relationship. For details on why the *2Q* attempts, but not the *4Q* attempts, were selected for refinement, see the "Refining their approach" section in Chapter 3. Because this approach began by filtering the data set, the results of these refinements are indicated in the figures below with funnel icons,[113] (i.e., "what they said they did, but filtered") and  (i.e., "what I think they did, but filtered"). Each of these compound icons contains the icon of the original attempt that is being filtered. Figure 4.23 below shows the results of both of these refinements in the context of the results from the respective attempts in my computational reproductions in the previous section. Note that the results appear to be the only suggestion thus far—in this dissertation or in the literature—of a positive correlation between FPS pronouns and status.

As shown in Figure 4.23, excluding dyads in these refinements reduced their sample sizes. For , I excluded dyads who disagreed with one another about their relative status relationship, reducing the original sample of 50 dyads by 72%, resulting in a sample of n = 14 dyads (i.e., 36 dyads disagreed with one another about who had higher status). For , I excluded three additional dyads based on the fact that there was no way to determine which participant in these dyads was of higher or lower status. Two of these dyads agreed that they were of equal status on both questions (sum of 8 for each). For the third dyad, one of the two questions showed one participant reporting they were equal, and the other participant saying they were slightly higher (sum of 9); on the other question, this response pattern was reversed (also summed to 9). For these three dyads, there was no way to determine which participant should be treated as having higher or lower status for the calculations. The sample size for  was thus reduced by 78% from the original 50 dyads, resulting in a sample size of n = 11 dyads. More details on determining dyad agreement can be seen in the "Refining their approach" section in

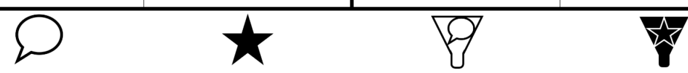
[113] While various funnel icons are widely used to indicate sorting or filtering (e.g., Apple's Mail app, Google Sheets), the function of a physical funnel is not to filter. This had never occurred to me until I began to type up my reasoning for choosing this icon here. That said, the English verb "funnel" does include the possibility of filtering, at least when used informally. I have learned much through my dissertation process.

Chapter 3 and in the appendices, with the full guide and accompanying charts available in my Figshare project folder.[114] Note that neither version of the refined results reached statistical significance (as defined by $p \leq .05$).

Figure 4.23

Results from the refined approach, presented in the context of the relevant results from the computational reproduction attempts (English data)

	Computational reproduction		**Refining their approach**	
	Lo minus Hi	P1 minus P2	Lo minus Hi	P1 minus P2
	Spearman	Pearson	Spearman	Spearman
sample size	50	50	11	14
correlation coefficient	-0.14	-0.30	0.32	-0.39
***p*-value**	0.39	0.03	0.14	0.17
effect size (Cohen's *d*)	-0.28	-0.63	0.68	-0.85



To look at the refined data sets in more detail, Figures 4.24 and 4.25 below are scatterplots of the  and  results, respectively. These figures show the FPS rate discrepancy for each dyad plotted against their status discrepancy. Trend lines indicate the direction (positive/negative) and degree of the linear relationships in these data sets; note that the more closely the points cluster around the trend line, the stronger the correlation. Note also that the  scatterplot is the same plot that would result from folding the  scatterplot in half at the *y*-axis, removing the three dyads on the *y*-axis (because they have no status discrepancy, as discussed above), and inverting the FPS discrepancy for a few of the remaining dyads. This is because the

---

[114] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

P1 minus P2 calculation essentially randomizes the dyads to have a positive or negative status discrepancy, which also affects the sign of the FPS discrepancy. This random assignment creates the appearance of meaningful spread, but is actually spreading the data out arbitrarily. In other words, while I used the P1 and P2 designations from the original data set , I could have randomized the participants to P1 and P2. Had I done this multiple times, the scatterplot would have "unfolded" each time into a scatterplot with a trend line at a different angle. The *Lo minus Hi* calculation, by contrast, does not allow for this level of tinkering, as it does not introduce the possibility of multiple randomizations. As mentioned in Chapter 3, I am including the *P1 minus P2* results here, not as plausible results for consideration, but as one example of the range of possible results that can be obtained through this randomized "unfolding."

Figure 4.24

Dyad-level results comparing status discrepancies and FPS rate discrepancies in the refined English data, using the Lo minus Hi *calculation (*n *= 11)*

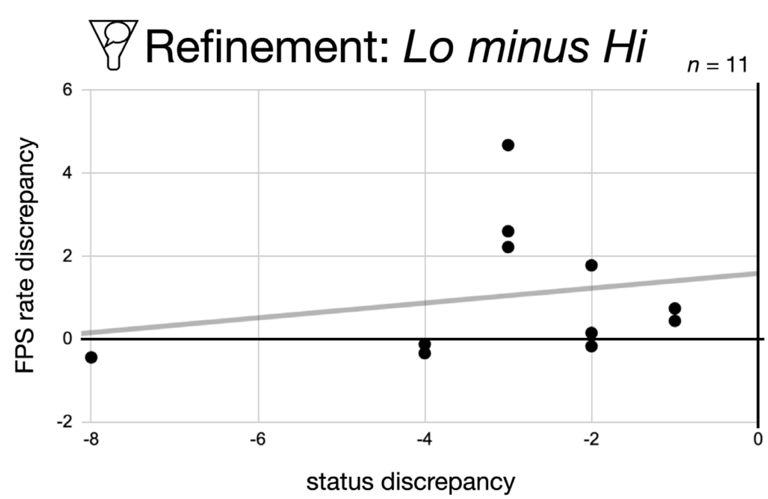
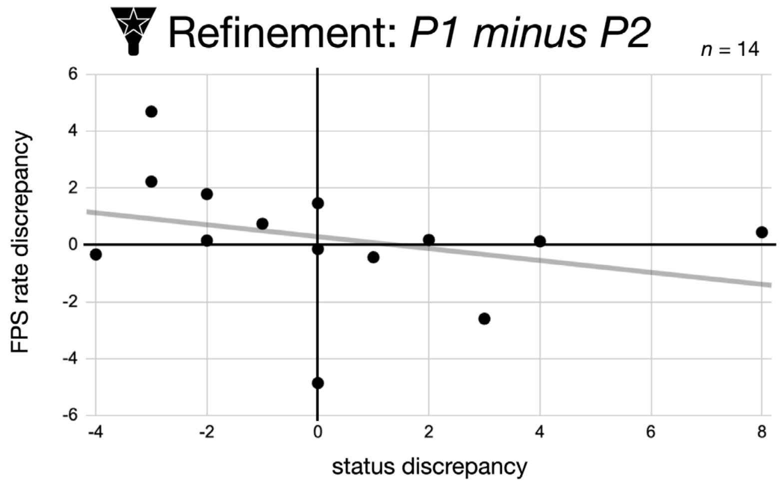


Figure 4.25

Dyad-level results comparing status discrepancies and FPS rate discrepancies in the refined English data, using the P1 minus P2 *calculation (*n *= 14)*



With the reduced sample size and the non-normal distribution of the refined data sets, the effect of potential outliers or anomalies is magnified. Fortunately, the smaller sample size also makes examination of individual dyads more feasible. To this end, Figure 4.26 below shows the FPS percentage results from my refinement calculations, similar to those reported for the ASL data I collected. The rate of FPS use for each participant is listed, along with simple differences between chat partners' rates at the dyad level.[115] For the *P1 minus P2* calculations, differences had a sign (i.e., were treated as positive or negative), according to the procedures determined during my computational reproduction process. However, note that simple differences are shown in Figure 4.26 below as absolute values, because the order of subtraction is not meaningful when the participants are not designated as being of higher or lower status. The three dyads discussed prior to the previous figure, for whom relative status discrepancies could not be determined, are

[115] The dyad numbers are from the original authors' data set that was shared with me, and their dyad numbering system (including that it goes beyond the *n* of 50) is discussed in the "Potential exclusion of two dyads" section in Chapter 3.

noted with asterisks in Figure 4.26 below. Keep in mind that the FPS rates in Figure 4.26 below can not be directly compared to those in the ASL data set; these FPS rates are reported as a percentage of the total words used, while the ASL data is reported as FPS/min rates. While the range of rates in the two data sets are similar, each represents a different type of quantification. For more on my decision to quantify this variable differently in the ASL data, see the "Variable: FPS pronoun use" section in Chapter 3.

Figure 4.26

Summary data for FPS percentage rates in the refined English data set

Refinement: 2Q &

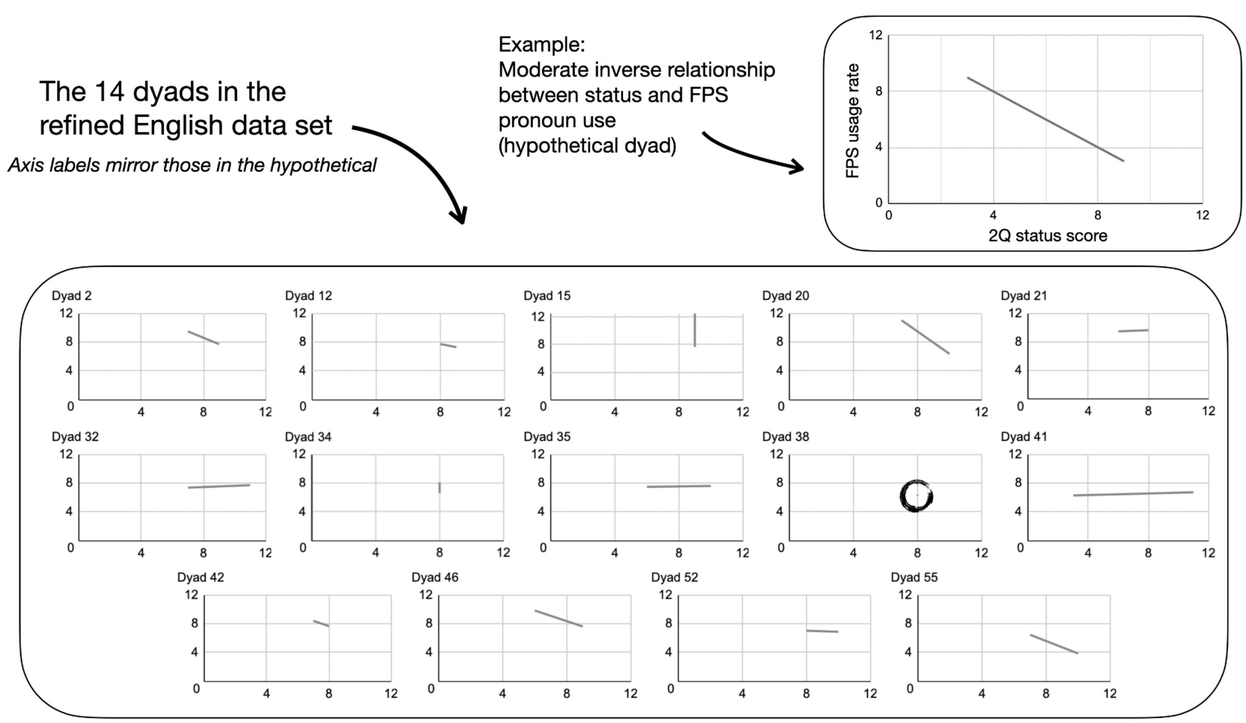
Dyad	FPS rate (% of words)		*Simple difference (absolute value)*
	P1	P2	
2	9.51	7.73	*1.78*
12	7.31	7.75	*0.44*
15**	7.67	12.53	*4.86*
20	11.06	6.38	*4.68*
21	9.64	9.47	*0.17*
32	7.38	7.72	*0.34*
34**	8.10	6.64	*1.46*
35	7.61	7.49	*0.12*
38**	6.24	6.39	*0.15*
41	6.73	6.29	*0.44*
42	8.42	7.68	*0.74*
46	9.85	7.63	*2.22*
52	7.04	6.89	*0.15*
55	3.87	6.47	*2.60*

**Excluded from the attempt

Figure 4.27 below provides a way to visualize the individual rates of FPS pronoun use for the dyads remaining in these refinements. The graph at the top of the figure shows a hypothetical example of the shape that would be seen with a dyad exhibiting a moderate inverse relationship between FPS use and status. The 14 smaller graphs below that one (i.e., small multiples) show the shapes of the 14 dyads included in . The axis titles of these smaller graphs have been removed for readability, but are identical to those in the hypothetical example at the top of the figure. Because these graphs do not require an assumption about which chat partner was of higher or lower status, all 14 dyads are included. The graphs for the three dyads that were excluded from  show vertical lines, reflecting that those dyads reported being of equal status. One of these is Dyad 38, the line for which is circled for easier visibility because their FPS rate discrepancy was so small that it is difficult to see.

Figure 4.27

Dyad-level relationships between status and FPS pronoun use in the refined English data set



Conceptual replication: Using the ASL data, applying the refined approach from above.

Figure 4.28 below shows the results from my conceptual replications in the context of the results I obtained from the refinement stage above. All four of the replications are modeled after either the  or the  refinements described above, in that they all use a data set filtered for participant agreement, and they then apply the dyad-level status calculation from either  (*Lo minus Hi*) or ★ (*P1 minus P2*). The compound icons for these replications include the icons from the refinement they are replicating, along with an image of a hand holding up either two or three fingers, depending on the paradigm (*2Q* or *3Q*). The resulting icons for the ASL replications under the chat-internal *2Q* paradigm are  (i.e., the filtered version of what they said they did, applied to the ASL data under the *2Q* paradigm) and  (i.e., the filtered version of what I think they did, applied to the ASL data under the *2Q* paradigm), and the resulting icons for the ASL replications under the chat-external *3Q* paradigm are thus  (i.e., the filtered version of what they said they did, applied to the ASL data under the *3Q* paradigm) and  (i.e., the filtered version of what I think they did, applied to the ASL data under the *3Q* paradigm). The mechanics of these replications are described in Chapter 3. Similar to my decision in the refining section above, I am including *P1 minus P2* results here not as an endorsement of this calculation, but for the sake of comparison with the originally published results.

Filtering the ASL data set for agreement reduced the sample sizes for these replications, though not as drastically as in the English data set. Excluding dyads who disagreed with one another about their relative status, based on the *2Q* paradigm, resulted in a sample size of 11 dyads for , down from the original 18 dyads (a 39% reduction). This number was further reduced when those who agreed they were equal were excluded, resulting in a sample size of *n* = 8 for (a 56% total reduction). In the *3Q* paradigm, the initial reduction was one of three dyads, taking the original 18 dyads down to 15 for (a 17% reduction), and the next reduction excluded two dyads who agreed they were of equal status, for a total of 13 dyads in (a 28% total reduction). Each set of results in Figure 4.28 below was calculated with Spearman's *rho*, due to the non-normal distributions of the data sets. Note that the results from the earlier refinement stage (included here for context) are still the only suggestion thus far of a positive correlation between FPS pronouns and status.

Figure 4.28

Results from multiple conceptual replications (ASL data), presented in the context of the refinements (English data)

	English data		ASL data			
	Refined approach (2Q)		*2Q* Paradigm: Chat-internal		*3Q* Paradigm: Chat-external	
	Lo minus Hi	P1 minus P2	Lo minus Hi	P1 minus P2	Lo minus Hi	P1 minus P2
sample size	11	14	8	11	13	15
correlation coefficient	0.32	-0.39	-0.24	-0.53	-0.53	-0.40
***p*-value**	0.14	0.17	0.56	0.10	0.06	0.14

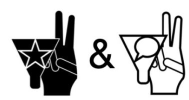
Figure 4.29 below presents dyad-level results from the *2Q* paradigm (i.e., and) in the style of similar figures from my primary analysis and the refinements in my secondary analysis. The rate of FPS/min use by each participant is listed, along with simple differences between chat partners' rates. As in my primary analysis, the apparent inconsistency between the reported FPS/min rates and the simple difference for 3HL is actually a rounding error, as a result

of rounding these values to two decimal places for reporting purposes. For the *P1 minus P2* calculations, differences had a sign (i.e., were treated as positive or negative). However, as with the refinements in the previous section, note that simple differences are shown in Figure 4.29 below as absolute values, because the order of subtraction is not meaningful when the participants are not designated as being of higher or lower status. The three dyads for whom relative status distinctions could not be determined were excluded from , since requires *Lo minus Hi* calculations. These dyads are noted with asterisks in Figure 4.29 below.

Figure 4.29

Dyad-level FPS/min results from my computational replications under the 2Q *paradigm (ASL data)*

Conceptual replication:
2Q Paradigm



Dyad	FPS/min		Simple difference (absolute value)
	P1	P2	
3HL**	19.87	5.49	*14.39*
3BL	9.02	14.81	*5.79*
4HL	9.03	5.64	*3.39*
10HL	9.05	13.88	*4.83*
5HB**	9.92	14.64	*4.72*
1BL**	9.93	8.36	*1.57*
8BL	10.03	11.16	*1.13*
1HB	11.27	11.86	*0.59*
10HB	10.43	11.09	*0.66*
2HL	7.59	8.58	*0.99*
7HB	11.78	14.30	*2.52*

**Excluded from the attempt

Figure 4.30 below presents dyad-level results from the *3Q* paradigm (i.e.,  and ), similar to the figure just above this. Again, the rate of FPS/min used by each participant is listed, along with simple differences between chat partners' rates, with the same treatment of sign and of dyads who were excluded from the *Lo minus Hi* calculations. As in my primary analysis, the apparent inconsistency between reported FPS/min rates and simple differences (e.g., 3HL) are actually rounding errors, as a result of rounding these values to two decimal places for reporting purposes.

Figure 4.30

Dyad-level FPS/min results from my computational replications under the 3Q *paradigm (ASL data)*

Conceptual replication:
3Q Paradigm 

Dyad	FPS/min		*Simple difference (absolute value)*
	P1	P2	
3HL	19.87	5.49	*14.39*
3BL	9.02	14.81	*5.79*
4HL	9.03	5.64	*3.39*
1HL	12.20	18.80	*6.61*
10HL	9.05	13.88	*4.83*
5HB	9.92	14.64	*4.72*
1BL	9.93	8.36	*1.57*
8BL	10.03	11.16	*1.13*
8HL	12.89	11.81	*1.08*
1HB**	11.27	11.86	*0.59*
8HB	20.94	20.96	*0.02*
10BL**	14.53	14.54	*0.00*
3HB	9.94	8.66	*1.28*
6BL	12.73	10.29	*2.44*
9HL	12.24	9.36	*2.88*

**Excluded from the  attempt

Reanalysis: Using the English data, applying the approach from my primary analysis.

The final step of my secondary data analysis was to run the English data using the methods from my primary analysis with the ASL data. As described in Chapter 3, this step used the refined data set from  described above, as I needed to be able to code participants as relatively lower or higher in status. Figure 4.31 below shows the results of this analysis. Similar to the corresponding figures from my primary analysis, Figure 4.31 below contains three tables. The first shows the FPS rate for each participant, along with simple differences between chat partners' rates (i.e., *Lo*'s rate minus *Hi*'s). The simple difference column shows that 7 of the 11 sets of chat partners (or 64%) used FPS rates that differed by less than one percentage point.

As before, the other two tables in Figure 4.31 provide descriptive and inferential statistics, and the "Simple difference" column in the top-right table functions differently than the "Simple difference" column in the table on the left. Because the data were not normally distributed, I calculated statistical significance using the nonparametric Mann-Whitney/Wilcoxon Signed-Ranks test. The effect size and the S-value are moderate, but the *p*-value does not reach statistical significance (as defined by $p \leq .05$).

Figure 4.31

Summary data for FPS percentage rates in the refined English data set, as well as descriptive and inferential statistics

Dyad	FPS rate (% of words)		*Simple difference*
	Lo	Hi	
2	9.51	7.73	*1.78*
12	7.75	7.31	*0.44*
20	11.06	6.38	*4.68*
21	9.47	9.64	*-0.17*
32	7.38	7.72	*-0.34*
35	7.49	7.61	*-0.12*
41	6.29	6.73	*-0.44*
42	8.42	7.68	*0.74*
46	9.85	7.63	*2.22*
52	7.04	6.89	*0.15*
55	6.47	3.87	*2.60*

	FPS rate (% of words)		*Simple difference*
	Lo	Hi	
mean	8.25	7.20	*1.05*
median	7.75	7.61	*0.44*
range	4.77	5.77	*5.12*
min	6.29	3.87	*-0.44*
max	11.06	9.64	*4.68*
SD	1.54	1.38	*1.61*

	FPS rate (% of words)
p-value	0.091
S-value	3.458
Effect size (*r*)	-0.510

Figure 4.32 below is modeled after figures presented in my primary analysis, showing the dyad-level view of FPS pronoun usage rates. As before, lines connect chat partners to provide a sense of how the dyads patterned in their FPS rates. Because the directionality of the dyad-level discrepancies is relevant, the lines connecting chat partners in Figure 4.32 are coded for directionality: solid lines connect chat partners whose results support the claim made by the original authors (i.e., the *Lo* participant used a higher rate of FPS pronouns than did the *Hi* participant), and dashed lines connect chat partners whose results contradict the claim. Note that most of the lines are closer to being horizontal than diagonal, indicating that there were few sizable discrepancies. This is in line with the observation that 7 of the 11 sets of chat partners (or 64%) used FPS rates that differed by less than one percentage point, as mentioned prior to the

previous figure. That said, Figure 4.32 shows that results from 7 of the 11 dyads (or 64%) support the claim made by the original authors.

Figure 4.32

Dyad-level rates of FPS pronouns used during chats, with dyads connected by lines (refined English data set)

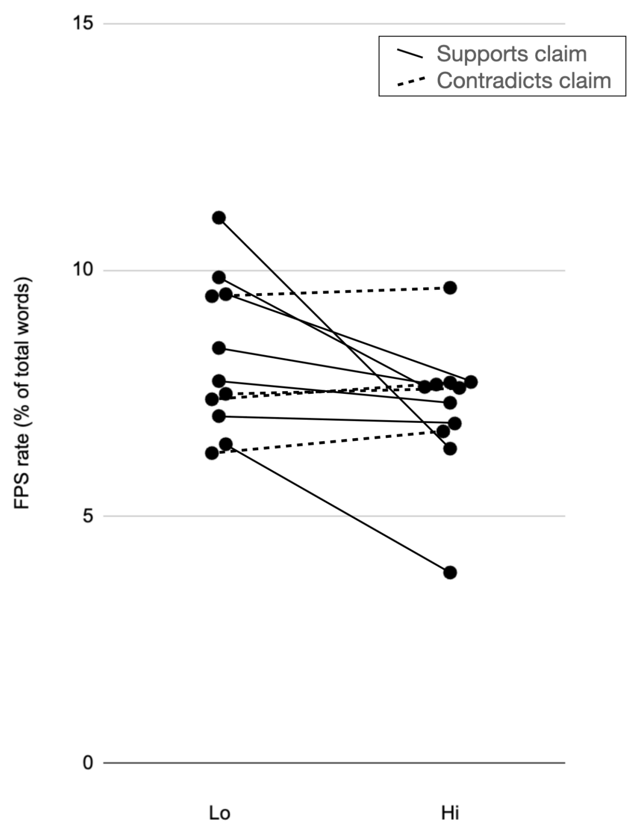
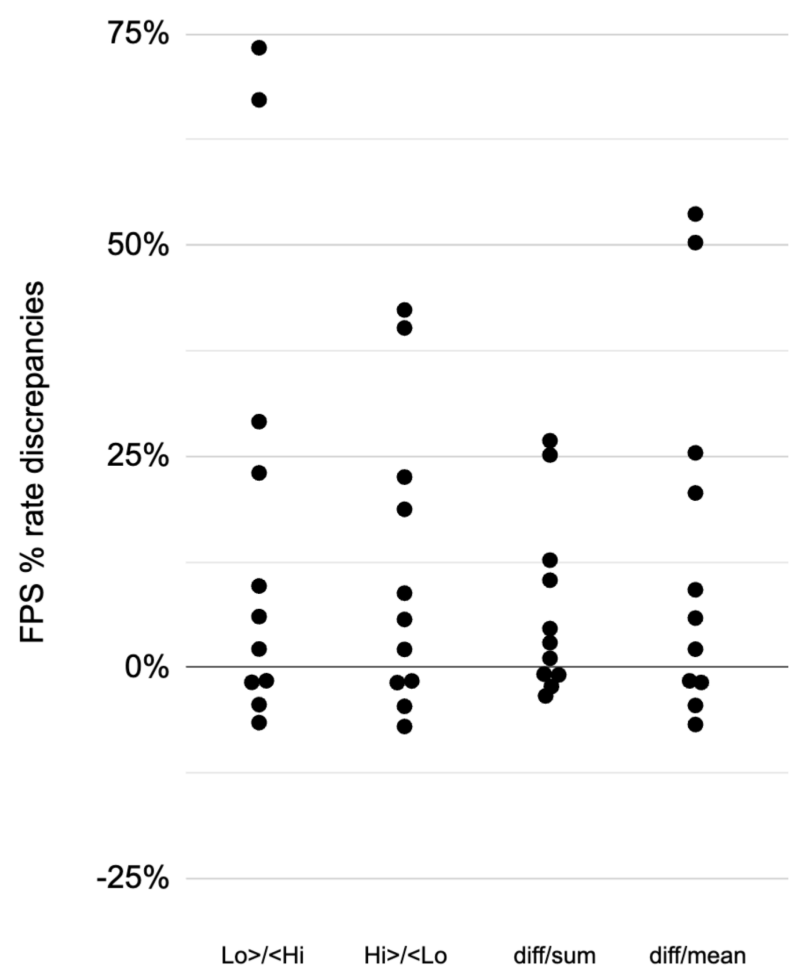


Figure 4.33 below is modeled after Figure 4.19 from my primary analysis, presenting the percentage differences from the refined English data set (calculated from the rates in Figure 4.26 above) displayed in four scatterplots presented side by side. As before, all four sets of results have been directionally aligned for visual comparability. See Chapter 3 for an explanation of each of the percentage difference calculations, along with my reasons for including all four in these analyses.

Figure 4.33

Dyad-level discrepancies in FPS percentage rates, presented as four side-by-side scatterplots showing results from four different percentage difference calculations (refined English data)



Conclusion

In this chapter, I have presented the results of my primary and secondary analyses. In the primary analysis, I worked with the ASL data I collected. This data set included rich phonological variation, with IX_1 and POSS_1 making up most of the FPS use. The results from 11 of the 18 dyads (61%) supported the prediction that *LowRole* would use higher rates of FPS/min than *HighRole*. That said, there were few dyads where the discrepancy was extreme, and differences between group means did not reach statistical significance (as defined by $p \leq .05$).

My secondary analysis focused on existing English data from Study 3 in Kacewicz et al. (2013). Working with their data set, my efforts at computational reproduction of their results were varied. I then applied some refinements to the methods reported by the original authors, which drastically reduced the sample size, and achieved results that were quite different from those they reported. I also applied the refined methods to the ASL data I collected, which produced interesting patterns. The final step in my secondary analysis was to apply the methods from my primary analysis to the English data set, the results of which were similar to those from my primary analysis. The results of both sets of analyses will be discussed in Chapter 5.

Chapter V: Discussion

Introduction

The aim of this study was to begin developing an understanding of the potential for a relationship between relative social status and FPS pronoun use in ASL. Working from the reported results and conclusions in prior work done on this potential relationship in English, I expected similar results for the ASL data that I collected. Specifically, I hypothesized an inverse relationship between relative social status and FPS pronoun usage rates in ASL. This was not borne out in my data set. After a closer examination of the prior work done in English, I am unable to say with confidence that an inverse relationship exists between social status and FPS pronoun use in English either.

In this chapter, I will discuss the results presented in Chapter 4, along with some exploratory analyses. In doing so, I will also present my conclusions about the current state of evidence for the claim that social status is inversely related to FPS pronoun use. As with the results presented in Chapter 4, additional details of the discussions and results below can be found in my Figshare project folder.[116]

Primary Analysis

Discussion.

For my primary analysis, my goal was to determine whether the ASL data I collected supported a particular relationship between social status and FPS pronoun use. The 18 dyads in my data set display varying degrees of status discrepancies, as measured using externally identifiable factors such as age, K-12 environment, and age of ASL acquisition. Among these dyads, most FPS pronoun rate discrepancies ranged from the *LowRole* chat partner using

---

[116] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

moderately more FPS/min (inverse relationship) to using moderately fewer (direct relationship). Although 10 of the 18 dyads supported the prediction of an inverse relationship, ultimately, there was insufficient evidence to conclude that a true inverse relationship exists in the population, especially since the average difference between the paired group means was not statistically significant (p = .14). One way to conceptualize the statistical significance of these results is to use the S-value, described in Chapter 3, which provides a measure of how surprising the results would be, assuming there was no true effect. The S-value for the difference in means was 2.8. This equates to the level of surprise associated with correctly predicting heads on a coin toss 2.8 times in a row. In other words, these results are not very surprising. While it is possible that there may be a true effect under certain conditions, these results do not provide enough evidence to claim anything other than chance is at work here.

The FPS usage rates across all participants ranged from 5.37 FPS/min to 20.96 FPS/min. Yet the differences in rates within each dyad were often quite small (see Figure 4.15 in Chapter 4). The rates were often so similar, in fact, that the discrepancies between them would likely not be readily noticeable to a participant or observer. This is represented in the effect size of r = -0.347.

As one way to understand the FPS/min results in real-world terms, Figure 5.01 below shows a breakdown of the FPS/min results for two dyads from the ASL data set. These two dyads are on either side of the median for the dyad-level FPS/min rate discrepancy percentages (with 18 dyads, the true median would actually fall between these two). This is one way to imagine a representation of an "typical" dyad from this data set. In Dyad 8HL, the *HighRole* participant used 103 FPS pronouns over the course of almost 9 minutes, and the *LowRole* participant used 215 FPS pronouns over the course of almost 17 minutes. This means *LowRole* used more than double the number of FPS pronouns. This difference is tempered when their turn times are considered, though, since *LowRole* used nearly twice as much turn time. Their resulting FPS/min rates end up only differing by approximately one FPS pronoun per minute; *HighRole*'s rate is 11.81 and *LowRole*'s rate is 12.89. In Dyad 1HB, the FPS rate pattern is opposite of that in 8HL, making these two dyads a nice example of the variability within this data set. In this dyad, *HighRole* used 176 FPS pronouns over the course of about 15.5 minutes, and *LowRole* used 73 FPS pronouns over the course of almost 6 minutes, meaning *HighRole* used nearly 2.5 times the number of FPS pronouns. Again, this difference is tempered when their turn times are considered, since *HighRole* used just over 2.5 times as much turn time. The resulting FPS/min rates for Dyad 1HB only differ by an average of slightly over one-half of one FPS pronoun; *HighRole*'s rate is 11.27 and *LowRole*'s rate is 11.86. These dyad-level differences (while representing discrepancies of 9% and 5%, respectively[117]), are unlikely to be noticeable in practical terms.[118]

---

[117] As described in Chapter 3, calculating percentage differences comes in many forms, none more correct than the others (though some are less appropriate for specific types of data). For the purposes of these examples, the dyad-level discrepancies were calculated using *diff/mean*, one of the symmetrized percentage difference calculations described in Chapter 3.

[118] To get a sense of what various rates "feel like" during use, see the following exploratory section later in this chapter: "Replication with publicly available ASL data: Primary replication, pronoun-less FPS construal, and an English translation."

Figure 5.01

Breakdown of the FPS/min rate results for the median-split dyads from the ASL data set

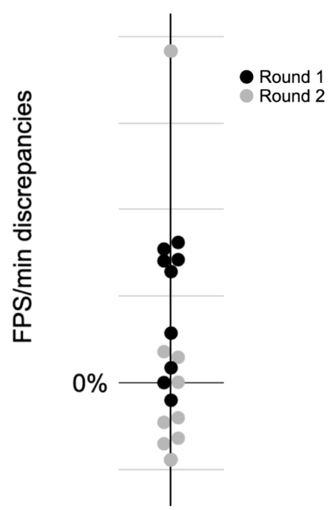
Median Dyad 1: **8HL**	**FPS**	**Turn time**	**FPS/min**	**% difference** (*diff/mean*)
HighRole	103	8 min, 43 sec	11.81	9%
LowRole	215	16 min, 41 sec	12.89	

Median Dyad 2: **1HB**	**FPS**	**Turn time**	**FPS/min**	**% difference** (*diff/mean*)
HighRole	176	15 min, 37 sec	11.27	5%
LowRole	73	6 min, 9 sec	11.86	

With these modest results in mind, I will now compare the results from my artificial division of the data set into two rounds of annotation, as described in Chapter 3. As mentioned in Chapter 1, analyzing only Round 1 of the ASL data set produced similar results to those reported in the literature for work done in English. I did not report the Round 1 results in Chapter 4 for the same reasons I described for not running statistics for other subsets (e.g., complete triads). That said, the Round 1 subset had a unique opportunity to replace the larger data set in a way that no other subset had. When I completed the results from Round 1, I was pleased—though not particularly surprised—to see that they supported the hypothesis of an inverse relationship between social status and FPS pronoun use in ASL. In fact, the Round 1 results showed a large effect that was statistically significant (r = -.74, p = .03). I could have stopped there with this data set and used the results as informed priors for the experimental data I planned to collect. Fortunately, I did not. The complete data set told a different story, as described above, namely, that the evidence does not support an inverse relationship between social status and FPS pronoun use in ASL. The contradictory conclusions of the Round 1 subset compared to the complete data set makes this an excellent example of how easily cherry-picking (intentional or not) can distort results. Additionally, this highlights the fact that statistical significance does not imply an adequate sample size. The different distributions of FPS/min discrepancies (using *diff/mean*) for Round 1 versus the complete data set can be seen in Figure 5.02 below. FPS/min discrepancies from Round 1 are in black, and those from Round 2 are in gray. The purpose of this figure is to show the general shape of the data in terms of the distribution (using *diff/mean*), particularly for comparing the positive values (that support the hypothesis) and negative values (that contradict the hypothesis). For this reason, the only value labeled on the vertical axis is 0%.

Figure 5.02

Distribution of dyad-level FPS/min discrepancies for the ASL data set, highlighting the different distribution of the Round 1 subset within



Limitations.

There are a number of limitations in the present study. My primary analysis includes 18 dyads, which is a relatively large sample size for an ASL study, but is nonetheless woefully small for making strong generalizations to the population. Additionally, everyone in my sample has completed at least some college. Since not all deaf (or hearing) people go to college, my sample was not representative in this way.[119] Similarly, individuals with connections to Gallaudet University are overrepresented in this sample, along with individuals with early exposure to ASL, individuals from deaf families, and individuals in their 20s and 30s. Beyond the limited participant pool, the Background Questionnaire and Post-Chat Survey left much to be desired, as described in Chapter 3.

Secondary Analysis

Discussion.

Working with the English data from Study 3 in Kacewicz et al. (2013) felt somewhat like doing a really well-thought-out (and very nerdy) escape room. Many puzzles had to be addressed, some logical and some mathematical, for me to keep making progress. And this was despite the fact that the senior author, Dr. James Pennebaker, was kind enough to share their data set and the IRQ documents with me. I am extremely grateful for the opportunity to work with their data myself, as this has greatly enhanced my ability to contribute to the larger conversation about social status and FPS pronoun use. In addition to the secondary analysis described here, I also worked with the original data from Study 1 in Kacewicz et al. (2013). For reasons of time and space, that work will have to wait for a future publication. Suffice to say, the inconsistencies I mentioned in my description of their Study 1 (in Chapter 2) are an incomplete list.

As for the work I did with Study 3 from Kacewicz et al. (2013), I believe the results presented in Chapter 4 paint a clear picture. Once the methods were refined (including using the discrepancy calculation in their paper) there does not seem to be adequate evidence to support any claim about a potential connection between social status and FPS pronoun use. For the

[119] According to the National Deaf Center on Postsecondary Outcomes, as of 2018, among people in the US ages 25-64, the percentage of individuals who had completed some college or higher was: Deaf 53.5%, Hearing 65.2% (Bloom et al., 2024).

discussion below, I will first discuss the analyses I did that used the refined methods from Kacewicz et al. (2013), applying them to both the English and ASL data sets. Then I will discuss how I applied the methods from my primary analysis (with the ASL data set) to the English data set.

While it appears that the original authors may have used the *P1 minus P2* approach to arrive at their reported results, the approach they reported having used was *Lo minus Hi*. I believe that the latter is the more appropriate choice. As discussed in Chapter 4, the *P1 minus P2* approach essentially randomizes the dyads to have a positive or negative status discrepancy. While randomization is generally considered desirable in research, this would be a randomization of values to be positive or negative, which arbitrarily distorts the data.

During the refinement stage, I screened the dyads for inclusion based on whether the more appropriate *Lo minus Hi* calculation could be reasonably applied to their results. Dyads who disagreed about who had relatively higher and lower status could not be included, nor could dyads who agreed that they were of equal status. Without “Lo” and “Hi” designations, *Lo minus Hi* is a non-starter. Applying this logic (described in more detail in Chapter 3) to the original English data set, only 11 of the original 50 dyads qualified for inclusion. This is a dramatic reduction in sample size, and it suggests that the methods used for determining relative status could be improved upon. In fact, this reduction introduces the question of whether the 11 “qualified” dyads are simply a random sample.

As for the *2Q* and *3Q* paradigms in my conceptual replication using the ASL data, the *2Q* chat-internal paradigm is useful in making more direct methodological comparisons with Kacewicz et al. (2013). However, in the “Conceptual replication” section in Chapter 3, I describe why I prefer the *3Q* chat-external paradigm, namely because it was designed to allow participants to more easily claim relative status differences while saving face. Accordingly, I believe the *3Q* paradigm is likely to better represent participants' subjective sense of their relative status.

Figure 5.03 below shows a pairing of results from my secondary analysis that can be used to evaluate the potential for a relationship between social status and FPS pronoun usage rates in both ASL and English. This figure is a reproduction of Figure 4.28 from Chapter 4, modified to include only the most suitable methodology used with each language's data set. As discussed above, the *P1 minus P2* calculation is not suitable for this type of data, so both sets of results in Figure 5.03 below use the *Lo minus Hi* calculation. And for reasons discussed above, the *3Q* paradigm is more useful here than the *2Q* paradigm for the ASL data. The same screening and discrepancy calculation methods were applied to both data sets to arrive at the results in Figure 5.03, making the results analogous; differences remain in the quantification of FPS usage rates and in some of the wording/questions used in quantifying social status (described in Chapter 3, under "Variable: FPS pronoun use" and "Conceptual replication," respectively). Figure 5.03 below retains the same icons used in Chapter 4 to indicate the various methods I followed for each approach.

Figure 5.03

Results that can be used to evaluate the potential for a relationship between social status and FPS pronoun usage rates in both ASL and English

	English Refined (*2Q*)	**ASL** *3Q*: Chat-external
sample size	11	13
correlation coefficient	0.32	-0.53
***p*-value**	0.14	0.06

The results in Figure 5.03 above do not support the hypothesis of an inverse relationship between social status and FPS pronoun use in either language. Ironically, the ASL data in  come closest to supporting this claim, though with a small sample size and a *p*-value that is only "nearly" significant (as defined by $p \leq .05$). While such a *p*-value can seem promising, ultimately, "nearly significant" is equivalent to "not significant." Recall that  used the chat-external questions eliciting what participants predicted others would say about their relative status relationship with their chat partner. This approach to eliciting subjective status perceptions might be advisable in future work of a similar nature. That said, with a *p*-value of .06, the probability of finding a correlation at least as extreme as this (i.e., a moderate-to-strong effect like -.53) would be 6%, if there was truly no relationship in the population (i.e., if only random chance were at work). In other words, if social status and FPS pronouns were not related at all, and we were to repeat this study over and over, then we would still see results at least as extreme as these in 6% of those studies. This equates to an S-value of 4.1, meaning that if there were no real relationship between the two variables, this result would be about as surprising as correctly predicting that four coin tosses in a row would all be heads.

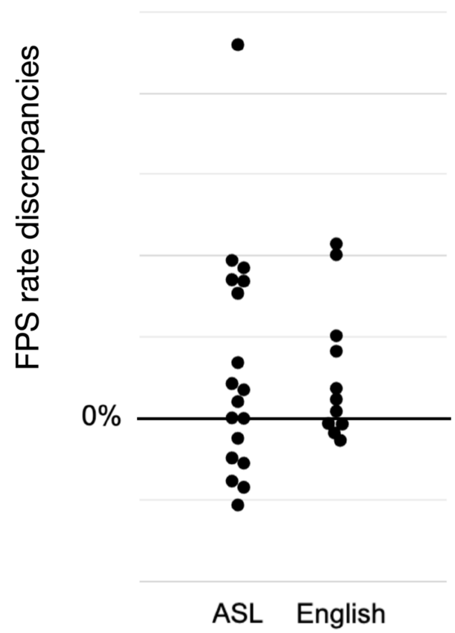
The English-data results from are incongruous with those published by the original authors. Their reported correlation coefficient was -.30, a weak-to-moderate negative correlation. The correlation coefficient from was .32, a correlation of comparable strength, but in the opposite direction. This positive correlation suggests a direct (rather than inverse) relationship between social status and FPS pronoun use. In sum, I started with the same data set and ended up with a nearly opposite conclusion. Additionally, the refined results from no longer reach statistical significance (as defined by $p \leq .05$). The p-value of .14 equates to an S-value of 2.84, meaning that if there were no real relationship between the two variables, this result would be about as surprising as correctly predicting that 2.84 coin tosses in a row would be heads. Taken together, the results in Figure 5.03 above appear to be fairly consistent with random chance. As discussed in the introduction to Chapter 4, different results can be expected from the same data set, given different methodologies—even when those methodologies are all pre-planned and well-reasoned. Such a lack of convergence in results suggests that the hypothesis is not true (Silberzahn et al., 2018).

Next, I applied the approach from my primary analysis to the existing English data, operationalizing status as binary. Under this approach, the results were somewhat consistent across languages, though not without their differences. Figure 5.04 below illustrates this similarity by showing the general shapes of the two data sets in terms of the distribution, both using the *diff/mean* calculation described in Chapter 3. As previously discussed, the choice of calculation is what determines the exact numerical values of the dyad-level percentage differences. For this reason, the only value labeled on the vertical axis is 0%. Similar to Figure

5.03 above, Figure 5.04 below can be used to evaluate the potential relationship between social status and FPS pronoun usage rates in both ASL and English, though this figure does not include inferential statistics, making it harder to determine the reliability of any apparent patterns.

Figure 5.04

Distributions of binary dyad-level FPS rate discrepancies for both the ASL and English data sets



Differences between these two sets of results include not only the visible differences in the spread of the distributions seen in Figure 5.04 above, but also differences in the quantification of the FPS rates. As discussed in Chapter 3, FPS rates were calculated as a percentage of total words in the English data and as a per-minute-of-turn-time rate in the ASL data. Additionally, for the English data, I was unable to consider the types of objectively identifiable status markers that I had used with the ASL data set. Many of the markers I used (e.g., age of ASL acquisition) are inapplicable to most hearing people. Those aside, even the markers that would have been applicable to this participant group (e.g., age) were unavailable; they were either not collected, or not included in the data set that was shared with me. With these facts in mind, I was bemused to see the fairly similar distributions shown in Figure 5.04 above. Differences aside, neither data set yielded a statistically significant difference (as defined by $p \leq .05$) between average group means for *LowRole* and *HighRole* (or *Lo* and *Hi*). This suggests that, while both data sets appear to lean toward an inverse relationship between social status and FPS pronoun use, the results align well with those we would expect purely from random chance. A more elegant, and more satisfying, explanation would include a third variable that more clearly drives FPS pronoun use, while also having some connection to social status. Such a variable has already been identified in the literature and addressed in this dissertation: Attention. I will further explore this possibility in the conclusion.

Limitations.

There are a number of limitations in my secondary analysis. Once the English data set was refined, the sample size became very small (n = 11 dyads). The ASL data set, which was quite small to begin with (n = 18), was also reduced in size (n = 13). Additionally, the files from Kacewicz et al. (2013) that I received from the senior author of the original paper included some files that I was unable to open, as the .spo file format (a type of SPSS output file) is no longer supported by SPSS, the SPSS Legacy Viewer, or any other program as far as I can tell. These may have contained useful information. Other limitations remain from the original English study. Examples include their IRQ being ambiguous as to the nature of status (i.e., zero-sum or variable-sum) for some questions, along with limitations described by the original authors, such as the participants all being undergraduate students, that questions of causality and perceptions of language use were not addressed, and that the word-counting approach precludes a more nuanced approach to language in use. Similarly, the limitations from my primary analysis also remain for the ASL data set in the secondary analysis.

Exploratory

The four dyads warranting special consideration.

I would now like to close the loop on the four dyads warranting special consideration, first described in Chapter 3 and indicated with asterisks (in tables) or stars (in charts) in the figures in Chapter 4. The results from these dyads ended up being fairly typical of the data set as a whole. Two of these dyads showed a moderate inverse relationship (*diff/mean*[120] of 46% and 49%) between social status and FPS/min rates. One dyad showed a negligible difference (*diff/mean* of 5%), and one dyad showed a weak direct relationship (*diff/mean* of -19%). This is not to say that these four dyads constitute a representative sample, only that, in a stroke of luck, they do not appear to have skewed the data in a different direction than the other 14 dyads trended in.

Potential reversal of status designation assignments.

Another dyad I singled out in Chapter 3 was the one instance where I suspected that I had assigned the status designations backwards, which was Dyad 5HB. The *HighRole* participant had arrived early, and we had a pleasant, brief conversation. When the *LowRole* participant arrived, I immediately felt certain that I had miscategorized these two. Vowing to double-check their Background Questionnaires, I began their session. When I came out of the studio, my colleague who was helping with sessions that day said something to the effect of "I thought you said the

[120] For this and some of the other exploratory analyses listed in the following sections, I will refer to dyad-level status discrepancies using the results from the *diff/mean* calculation described in Chapter 3. As previously discussed, these numerical values do not represent "the" difference between chat partners' rates. However, when using numerical values is useful to the discussion, I have chosen to use this version of symmetrized percentage difference for ease of comparison because it is not directionally dependent and the results constitute a larger and more nuanced spread than those of *diff/sum*.

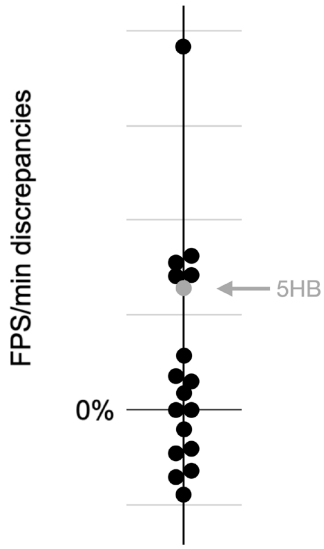
second person was the *LowRole* person…?" I was somewhat relieved to have my intuitive impressions validated, and even more curious about this dyad's Post-Chat Surveys.

Their survey responses suggested that they agreed with our impressions. While they were unaware of the status designations I had assigned, their responses indicated they both felt that the *LowRole* participant had been the one with higher relative status. I included them in Round 1 for early annotation for this reason. Using their status designations as originally assigned for my primary analysis, their results showed an inverse relationship between social status and FPS pronoun use. Yet, when I used their subjective responses during the secondary analysis, their results showed the opposite effect: a direct relationship between social status and FPS pronoun use. In fact, out of the 13 dyads in the  attempt, their FPS/min rate discrepancy is the second most extreme discrepancy in that direction.

Dyad 5HB's FPS/min rate discrepancy from my primary analysis is indicated in gray in Figure 5.05 below, in the context of the complete data set. As with Figure 5.04 above, the purpose of this figure is to present the general shape of the data in terms of the distribution (using *diff/mean*), this time to provide context for one particular dyad's results. For this reason, the only value labeled on the vertical axis is 0%. This dyad serves as an interesting case study in measuring status using subjective status markers versus the externally identifiable "objective" markers.

Figure 5.05

FPS/min discrepancy of Dyad 5HB, presented within the distribution of dyad-level FPS/min discrepancies for the entire ASL data set



Gender and age.

When I tell people about this project, I am often asked whether I found any effects for gender. As described in Chapter 2, claims have been made about people who identify as female, as well as younger people, tending to use higher rates of FPS pronouns (e.g., Pennebaker & Stone, 2003). And while conversations are changing around so-called “powerless language,” a category that tends to include FPS pronouns, this type of language use is still attributed to women and younger people more often than to men and older people, in addition to the claim that women talk more than men (Tannen, 1994; Kaplan, 2016). For these reasons, I am including this short commentary of how this data set responds to these assumptions. As mentioned in Chapter 3, all participants self-identified as either male or female.

Taking the participants out of the context of their chats (see “Comparing FPS/min rates” in Chapter 2 for more on this), and ordering them by their FPS/min rates, there are no clear patterns for either age or gender being good predictors for FPS pronoun usage rates. The people in my data set who were over 50 years old tended to use FPS/min rates that were toward the lower end of the range, though not exclusively. And some of the youngest people, in their early 20s, used comparably low rates. Once duplicate participants were considered (see “Participant Demographics” section in Chapter 3), there were 23 female “participants” and 13 male “participants.” Of these 36, the most extreme FPS/min rates (the four above 15 FPS/min, and the two below 6 FPS/min) were from female-identifying participants. Other than that, the distributions of FPS/min rates were comparable for both genders.

As for who used more turn time (i.e., who “talked more”), again, there were no clear patterns for gender or age. In mixed-gender dyads (n = 7), who used more turn time was slightly more closely related with status designation (i.e., *HighRole* or *LowRole*) than with gender, though even the status designation differences were minimal and not uniform. Age also did not appear to be a predictor for the amount of turn time used by participants. Taken together, these results do not suggest any effects related to gender or age with relation to FPS/min usage or turn time in this data set.

FPS/min stabilization.

When I was designing my study, I needed to decide how long to plan for the chat sessions to be. In Study 3 from Kacewicz et al. (2013), sessions were described as being 10 minutes long. However, in his general-audience book, Pennebaker (2011) provides a detail that I considered making use of. In describing a study that appears to be either the one that became Study 3 in Kacewicz et al. (2013) or a similar study, he says “within the first minute of the conversation, the social hierarchy was established” (Pennebaker, 2011, p. 182), adding “their relative use of pronouns was fixed almost immediately” (p.182). He then compares this to a text-only online chat study (not described in Kacewicz et al., 2013[121]), saying "just like the face-to-face natural conversation project, the status differences were apparent in the first three minutes of the interaction" (p. 183). Seeing this, I wondered if the potential benefits of shorter chat sessions (e.g., shorter annotation time, increased participant availability) would be worth losing data that could potentially change the results. On the other hand, I reasoned that if FPS rates did not stabilize quickly in ASL interactions, longer sessions might be a better choice. Ultimately, I decided to aim for 20-minute sessions, doubling the length of the sessions described in Kacewicz et al. (2013) for Study 3.

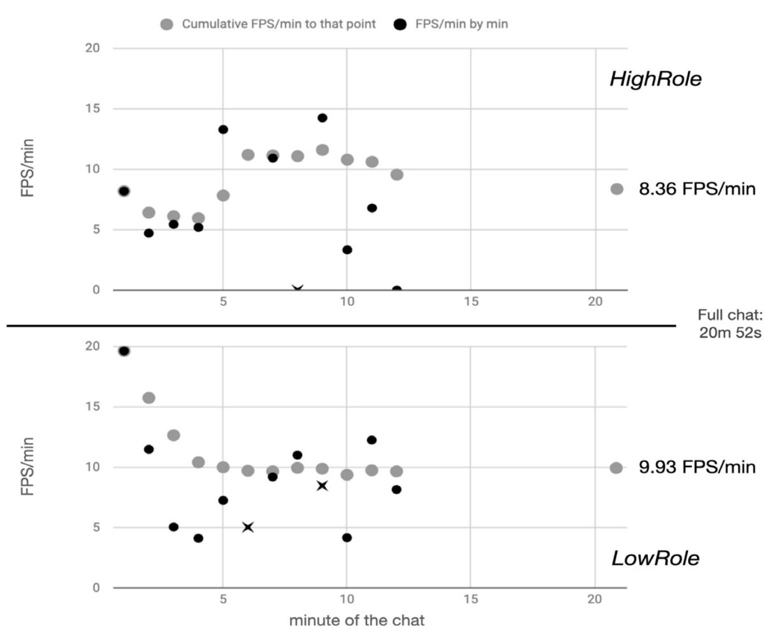
Still, I wondered if and/or when FPS/min rates may have stabilized in my data set. I randomly selected six of the dyads from Round 1 annotation, for a total 12 participants. I calculated each participant's FPS/min rate for each minute during the first 12 minutes of the chat (i.e., minute-by-minute rates). I also calculated the average FPS/min rate as if the chat had ended at each of those first 12 minutes (i.e., hypothetical chat-average rates, based on different hypothetical chat lengths). These 12 participants did not display any reliable stabilization trends within the 12-minute timeframe. As one example, Figure 5.06 below shows the results from this analysis for Dyad 1BL. Black dots show FPS/min rates for each individual minute (i.e., the number of FPS pronouns the participant used during that minute divided by the turn time they used during that minute). Rates during the minutes where the participant used less than 15 seconds of turn time are marked with a black X instead of a black dot. Gray dots (slightly larger than the black dots) show the hypothetical average FPS/min rates, had the analysis ended at that minute. Had I only included the first 3 minutes of this dyad's chat, as suggested in Pennebaker

[121] It is not immediately obvious if this study was published. There is nothing with the corresponding authors and a fitting title on Pennebaker's CV, which was last updated in February 2024 at the time of this writing (https://minio.la.utexas.edu/colaweb-prod/person_files/0/168/Pennebaker_CV_2024.pdf).

(2011), the results would have been very different from the results of the full chat time of 20 minutes and 52 seconds. This can be seen for each of the two participants by comparing the gray dot at the 3-minute mark to the gray dot at the end of the chat. The gray dots starting at the 6-minute mark show less variability than those for the first five minutes, though the within-participant variability is comparable to the between-participant variability. As such, this level of variability would still have been enough to have dramatically affected the dyad-level results. Using the *diff/sum* calculation described in Chapter 3, this dyad's full chat resulted in a FPS/min discrepancy of 17%. Had only the first three minutes been included, the *diff/sum* discrepancy would have been 200%; if only the first six minutes, the *diff/sum* discrepancy would have dropped (below the full-chat rate) to 14%; using the first 12 minutes, it would have dropped (further below the full-chat rate) to 1%. By the time the chat ended, the discrepancy was up to 17%. This suggests that there was a good deal of rate variability throughout the chat, including within the minutes I did not include in this analysis.

Figure 5.06

Results from the rate stabilization analysis for Dyad 1BL (ASL data set)



The results from these 12 participants in the ASL data set do not support the rate stabilization claim made for English by Pennebaker (2011). The reason for this is unclear, though it may be related to differences between the two languages, such as the fact that ASL is a pro-drop language. Unfortunately, I am unable to run corresponding stabilization analyses on the English data from Kacewicz et al. (2013), because the data set I received does not include time codes. As it stands, it appears that the variability in the FPS/min-by-minute rates contributes to a good deal of variability in the hypothetical chat-average rates (based on different hypothetical chat lengths). This degree of variability suggests that FPS rates may not be reliable enough to use as a measure for drawing comparisons with social status, at least not by the 20-minute mark and conceived as per-minute rates in ASL. If further investigation were to support this preliminary finding, that FPS rates are highly variable over time, then addressing the question of a relationship between social status and FPS rates would become moot.

Triads.

As described earlier, my original plan with this work was to focus my analysis on the B participants in each complete triad. I later changed my plan (described in Chapter 3), but remained curious about what could be gleaned from the complete triads. There were only four complete triads, so the sample sizes here are even smaller than those in my primary and secondary analyses. What follows is a brief exploratory indulgence.

As discussed in Chapter 3, running group-level statistics on the triads subgroup is not particularly meaningful. Nor are group-level statistics on further subdivisions, such as the B group. That said, some of the higher-order questions addressed by group-level statistics (e.g., group means) can also be addressed using the results presented in Chapter 4. Three such questions follow, accompanied by results. As previously described, the social status differences between H, B, and L are not standardized. This means that the distances between any two participants can only be considered within-triad, making the comparisons in Triad Question 1 below valid only within each triad, not across multiple triads. Additionally, all three Triad Questions below require FPS rates to be taken out of context, which calls for an additional caveat. See the “Triads” section in Chapter 3 for more on both of these points.

Triad Question 1: *Did a larger status discrepancy between chat partners coincide with a larger FPS rate discrepancy (*n *= 8 participants)?* In other words, for each chat with an H participant, was there a larger FPS rate discrepancy during the HL chat than the HB chat? Similarly, for each chat with an L participant, was there a larger FPS rate discrepancy during the HL chat than the BL chat? For seven out of the eight participants that this applies to (four Hs and four Ls), a larger status discrepancy between chat partners did coincide with a larger FPS rate discrepancy. For the eighth participant (8L), the FPS rate discrepancy was minimal, and in the opposite direction (i.e., a larger status discrepancy between chat partners coincided with a slightly smaller FPS rate discrepancy).

Triad Question 2: *Did B participants use higher FPS rates while in* LowRole *than while in* HighRole *(*n *= 4 participants)?* In other words, I am comparing each B participant’s FPS/min rate from their HB chat with their rate from their BL chat. This was the goal of my original study design, when I had planned to collect data only from complete triads. For the four participants this applies to, two used a higher FPS rate while in *LowRole* than while in *HighRole*. That is, half of the relevant data points supported the hypothesis, and the other half contradicted it.

Triad Question 3: *When non-B participants (Hs and Ls) were in the more extreme version of their role (the HL dyad) were their FPS rates correspondingly more extreme (*n *= 8 participants)?* In other words, when H was paired with L (i.e., at a greater distance from their chat partner), was H's FPS rate lower than when they were paired with B? When L was paired with H, was L's FPS rate higher than when paired with B? Note that this is more similar to Dyad Question 2 than Dyad Question 1, as this is a participant-level (i.e., within-participant) question, rather than a dyad-level one. Of the eight participants this applies to (i.e., four Hs and four Ls), six of them had FPS rates that were correspondingly more extreme when they were in the more extreme version of their role. With each participant contributing just two data points to this pattern, a correlation can hardly be assumed.

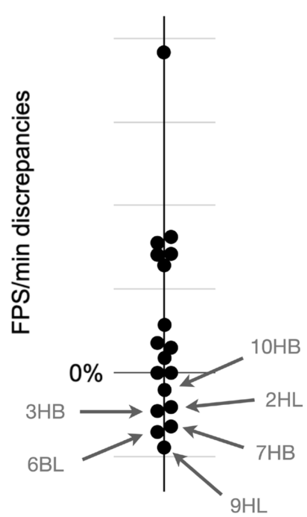
Behavioral considerations.

As I annotated the ASL data, my initial coding included notes on incidental behaviors that were not directly related to my research hypothesis but that nonetheless struck me as interesting. As I began to notice certain types of behavior showing up repeatedly, those notes became more frequent. I considered adding a separate pass for annotation of status-related behaviors, but decided against it, in order to keep my focus on my prespecified quantitative analyses (and to avoid letting the annotation process continue indefinitely). But as I worked on the quantitative analyses, I could not let go of my curiosities about behavior. I remembered noticing what had appeared to be a pattern in which of the participants tended to take the lead and initiate a new topic of conversation when there was a short lull. I also remembered some specific comments and behaviors that I wanted to look into. Ultimately, I did go back and complete a separate annotation pass (see "Status indicators tiers" in Chapter 3) specifically noting behaviors that might be somehow related to social status. Because I did not begin coding for behaviors in earnest until I had already calculated the FPS/min results, the following observations are more susceptible to confirmation bias and should be treated with skepticism. The results presented here serve simply to satisfy some of the many curiosities that emerged while I have been immersed in this work.

There were six dyads whose FPS/min discrepancies were negative, supporting a direct (rather than inverse) relationship between social status and FPS pronoun use. For each of these, there appears to be information that might be useful in generating hypotheses. Figure 5.07 below shows these dyads in the context of the data set as a whole. The exact value of the dyad-level percentage differences depends on the calculation used (as discussed ad infinitum throughout this dissertation), but the rank-order remains the same, as does the sign, so these dyads are labeled by name rather than percentage difference.

Figure 5.07

The six dyads with negative FPS/min discrepancies, presented in the context of the ASL data set as a whole



For each dyad with a negative discrepancy, I describe my relevant observations below. The descriptions begin with the dyad closest to zero, moving down the number line to more extreme negative discrepancies.

10HB: On the Post-Chat Survey, the *LowRole* participant indicated that they felt they had more than their fair share of control during the conversation, and that they would have liked to have had much less control. This may be a sign that they felt the need to keep the conversation going, or that they felt their chat partner had not contributed enough.

2HL: After their chat, the *HighRole* participant commented to me, off-hand, that they had been able to answer the survey question about relative status during the chat without any hesitation, because they recognized that their chat partner had been of much higher status. In short, *HighRole* told me they were lower than *LowRole*.[122] This participant responded accordingly on the Post-Chat Survey. While I can not venture to guess what created this impression, if their self-perception of status had been used, then their FPS pronoun use relative to their partner's would support the research prediction (i.e., result in an inverse relationship). *HighRole* also indicated on the Post-Chat Survey that they would have liked more control in the conversation than they had perceived having. Correspondingly, *LowRole* indicated that they would have liked slightly less control than they had perceived having.

3HB: During their chat, the *HighRole* participant took the lead, filling several short conversational lulls, eventually asking *LowRole* a general question along the lines of "is there anything else you'd like to say?" near the end of the chat. *HighRole* also indicated on the Post-Chat Survey that they would have liked less control during the conversation than they had perceived having.

[122] I can only hope that my face did not reveal my internal reaction.

7HB: After their chat, the *HighRole* participant told me, unprompted, that *LowRole* had interrupted them often, and finished their (*HighRole*'s) sentences. *HighRole* appeared to have been annoyed by this behavior. On the Post-Chat Survey, *HighRole* reported having slightly less control, and said they would have liked to have had equal control with *LowRole*. Correspondingly, *LowRole* reported having slightly more control, and reported that they too would have liked to have had equal control.

6BL: During their chat, the *LowRole* participant took a somewhat hands-off approach, not seeming to feel the need to actively engage in conversation. These two participants were already acquainted, and so did not have the easy fallback of exchanging get-to-know-you pleasantries. One of my notes from the behavior pass for this dyad was along the lines of "Wow, *HighRole* is really carrying the conversation here…I wonder if they feel some pressure to 'perform' for the cameras." On the Post-Chat Survey, *HighRole* reported having much more control than they would have liked, while *LowRole* reported being satisfied with their own perception that they had both shared equal control.

9HL: During their chat, the *HighRole* participant did most of the leading, asking questions and then supplying their own reciprocal answers even when *LowRole* did not actually reciprocate the question. One of the longest lulls in this entire data set happened more than halfway through this chat. It lasted nearly four seconds, and was finally broken by *LowRole* asking their first question of the chat, but not until *HighRole* gave a look that I interpreted as "I could sit here all day..." On the Post-Chat Survey, *HighRole* indicated that they would have liked to have had much less control in the conversation; ideally, they would have liked to have had equal control.

These observations, taken together, support the possibility that the negative dyad-level FPS/min discrepancies support a slightly different hypothesis than that addressed throughout the rest of this work. Namely, that FPS use is simply an indicator of attention, which entails much more than just social status. I will discuss this further in the conclusion.

Replication with publicly available ASL data: Primary replication, pronoun-less FPS construal, and an English translation.

As discussed in Chapter 1, I have not made the video data I collected publicly available. However, I would like to provide an example of what FPS/min rates look like in ASL use. When I first began this work, I became mildly obsessed with trying to get a sense of people's FPS pronoun frequencies in use during my daily interactions.[123] If someone is talking about themselves, it is not difficult to pick up on FPS use in real time. However, it is much more challenging to discern in real time whether someone is tending to frame things with an FPS construal (with or without FPS pronouns, see "FPS construal without pronouns" in Chapter 2). Recall the following example sentences from Chapter 2: "Something is wrong with this computer" (construed without self-focus, thus without an FPS pronoun), and "I can't figure out what's wrong with this computer" (construed with self-focus, thus with an FPS pronoun).

When I first began annotating the ASL data I collected for this dissertation, I would often come to the end of a chat and note my sense of who had used a higher rate of FPS pronouns for later comparison with the actual FPS/min results. After all, I had just annotated the chat (between 47 and 311 individual FPS pronouns per participant), so I thought there was a good chance that I had an accurate sense of who had used more FPS pronouns per minute, and perhaps even to what degree. This stopped being fun when I realized that I was nearly always wrong.

[123] Apologies to any readers who experience this distracting habit after reading about my work here.

In an effort to provide a sense of what FPS/min rates can look like in an ASL interaction (and to provide an opportunity for like-minded readers to test their own skills), I offer an example with a publicly available video. The video I chose is a clip of about six minutes from a longer video that was published on YouTube by The Daily Moth (as ASL news channel), and is titled "Interview with Marlee Matlin and Daniel Durant."[124] During this clip, Alex Abenchuchan from The Daily Moth is interviewing actors Marlee Matlin and Daniel Durant about their recent film, *CODA*. All three participants are deaf and using ASL. The clip includes three questions from Alex and responses from Marlee and Daniel.

I did three analyses of this video, each addressing a different curiosity raised earlier in this dissertation. The first analysis was a replication of my primary analysis, comparing FPS/min rates used by all three participants. For the second analysis, I added instances of FPS construal that occurred without the use of an FPS pronoun (e.g., pro-drop with indicating verbs; see "FPS construal without pronouns" in Chapter 2). For the third analysis, I worked with the English captions from the video. Figure 5.08 below contains a summary of the results from all three analyses. Following the figure are discussions of all three.

---

[124] The original video, posted on August 11, 2021, can be found here: https://www.youtube.com/watch?v=Ta6y10hxPcw&t=19s

The roughly six-minute clip I analyzed is between timecodes 3:42 and 9:54, and is available as a video file in my Figshare project folder, along with the spreadsheets I used for these analyses, here: https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

Figure 5.08

Results from three exploratory analyses using publicly available ASL data

		Alex	Marlee	Daniel
	Turn time	1m 55s	2m 48s	1m 25s
ASL	FPS pronouns	2	15	14
	FPSpro/min	**1.04**	**5.35**	**9.86**
	Total FPS (including construal)	5	16	19
	Total FPS/min (including construal)	**2.61**	**5.71**	**13.38**
English captions	Total FPS	6	14	15
	Total FPS/min	**3.13**	**4.99**	**10.56**

First, I applied the approach from my primary analysis, calculating and comparing rates of FPS pronouns used per minute. I selected the clip to contain complete conversation topics (questions and their answers), but did not consider the balance of turn time. As it happened, Marlee used the most turn time during this clip, at 2 minutes and 48 seconds, followed by Alex at 1 minute and 55 seconds, then Daniel at 1 minute and 25 seconds. The ranking of their FPS/min rates, in order of lowest rate to highest was Alex (1.04), Marlee (5.35), then Daniel (9.86). The FPS/min discrepancies between each pairing of these three participants are quite a bit higher than most of the discrepancies seen in my primary analysis; that is, participants' discrepancies were more disparate. Using the *diff/mean* calculation for percentage difference, the discrepancies were:

- Alex and Daniel: 162%
- Alex and Marlee: 135%
- Marlee and Daniel: 59%

For comparison, consider that the *diff/mean* values in my primary analysis ranged from -27% to 115%. What I find most interesting in this analysis though, is the ranked order. I will discuss this in more detail after I have described the two other analyses.

For the second analysis with this video, I coded for instances of FPS construal that occurred without the use of an FPS pronoun. Identifying and quantifying these instances is not always straightforward. I added these instances of pro-drop to the FPS count, for a “Total FPS/min” rate that includes both FPS pronouns and pronoun-less FPS construal. All three participants evoked FPS construal without the use of pronouns (e.g., pro-drop with indicating verbs or constructed action). Interestingly, the order of their relative rates remained the same: Alex (2.61), Marlee (5.71), then Daniel (13.38). Considering my decision to only count overt pronouns in my main study, seeing this felt like a small victory—despite the fact that this exploratory analysis uses exceptionally small language samples from only three individuals. As in the previous analysis, the FPS/min discrepancies between each pairing of these three participants are higher than most of the discrepancies seen in my primary analysis. Using the *diff/mean* calculation for percentage difference, the discrepancies are:

Alex and Daniel: 135%

Alex and Marlee: 75%

Marlee and Daniel: 80%

My final analysis with this clip involved analyzing the English captions. The original video includes open captions, meaning that the on-screen text is part of the video itself, rather than being a setting that can be toggled on or off by individual users. These captions provide a prepared English translation of the video. The purpose of the translation was not to represent FPS pronoun use in ASL, and so it should not be used to make any claims about ASL (cf. Study 5 from Kacewicz et al., 2013). However, analyzing this translation allows for a different type of FPS analysis, one that begins to address my interest in seeing how FPS rates might pattern in interpreted contexts. Similarly to the ASL analyses for this video, I counted overt FPS pronouns in the captions, and I also looked for instances of pro-drop. To my surprise, there were no instances of pro-drop in the captions. This may be due to the fact that English tends to require the overt use of pronouns. Additionally, some of the more frequent types of English pro-drop (e.g., diary drop) are more frequent in ephemeral use and are considered somewhat informal, which may have discouraged the translator(s) from using them.

When calculating FPS rates for the English captions, I decided not to calculate rates as percentages of the total words a participant used, deviating from the quantification method used with the previous English data sets. I opted instead to use the "per-minute" quantification that I used in my primary study with the ASL data. This makes the rates more comparable across languages within this exploratory analysis. All three sets of results from this analysis maintain the rank order of the participants' FPS/min rates: Alex (3.13), Marlee (4.99), then Daniel (10.56). While this is not completely unexpected, it is certainly interesting. The FPS/min discrepancies between each pairing of these three participants are within the range of discrepancies seen in my primary analysis, though they are still higher than average. Compared to the discrepancies in the refined English data set, they are higher than most. Using the *diff/mean* calculation for percentage difference, the discrepancies are:

Alex and Daniel: 109%

Alex and Marlee: 46%

Marlee and Daniel: 72%

As previously mentioned, the order of usage rates among the three participants is what interests me most here, as the order was maintained in all three analyses. Excluding instances of pro-drop meant excluding instances of FPS construal for all three participants' ASL use, but this did not affect the rank order of participants' FPS rates, only the magnitude of the differences between those rates. Also, the order of the dyad-level discrepancies remained the same between the second analysis (FPS construal in ASL, with or without FPS pronouns) and the third (FPS pronouns in the English captions). This order preservation supports the idea that a complete translation may preserve construal of self-focus without the need to preserve each word as a translation equivalent. Further research could explore related ideas. As for potential reasons for the rank order of these three participants' FPS rates, any order could be rationalized, given enough motivation from a rationalizer. Using concepts I have discussed in this dissertation so far, I will share some potential rationalizations here. The underlying sentiments could easily be applied to any results connecting social status and FPS pronoun use, including those from the ASL and English data sets I used in my primary and secondary analyses.

First, let us assume there is some minor effect for FPS pronoun use being inversely related to social status, perhaps as a function of attention. Alex may have used the lowest rate of FPS pronouns because he was in a position of control as the interviewer, and so directed the flow of conversation. He may have also had more of a say in what parts of the interview would be included in the published video. While Marlee may be more famous on a larger scale (see the Oscar statue in the background of her video feed), Alex may have had more power during the interview (see his comment that is captioned as beginning with "regardless of your opinion of the film CODA"). Daniel, being the least well-known of the three, may have felt deferential to both Alex and Marlee, which would explain his FPS/min rate being the highest.

Now, let us begin from a different premise: that there is no true relationship between social status and FPS pronoun use. Alex may have used fewer FPS/min simply because part of his job in the role of interviewer was to talk about and ask questions of the people he was interviewing, focusing on them more than on himself. Also, Marlee appeared to have been signing slightly slower than the other two, and she used more turn time. Thus, her average rate of FPS/min may have been lower than Daniel's simply due to having fewer utterances during her longer turn time. Had I analyzed a longer portion of the interview, all three rates may have fluctuated significantly (see "FPS/min stabilization" section for evidence supporting this idea).

Alternatively, the rates used by all three participants may have simply been due to random chance.

Future Directions

This work could be expanded and extended in any number of ways. Some would continue along the same path, asking similar questions with different data, perhaps expanding and otherwise improving the data sets to allow more refined questions to be asked. Others would make use of the ASL data set I collected, or a similar one, but would ask new questions. Both approaches have been touched upon in this dissertation. Below are lists of additional possibilities for both approaches.

<u>Continuing this line of inquiry</u>

- How would different operationalizations of status pattern with FPS pronoun usage rates?
- Are there FPS pronoun distribution patterns based on topic?
- Do patterns differ with fluent hearing signers? New signers who are deaf or hard-of-hearing?
- Are there FPS patterns that are specific to interpreted interactions?

- How accurately do individuals perceive their chat partners' FPS use?

- How would priming participants' subjective feelings of status affect intra-subject FPS use?[125]

Continuing with this data set (or similar)

- How does coding for pronoun-less FPS construal (e.g., indicating verbs, depicting verbs) affect rates of FPS/minute?

- How are different FPS pronouns used: to hold a turn? as fillers? perseverated? repeated?

- Are there pro-drop patterns in so-called "English-based" constructions with the translation-equivalent FPS pronoun (e.g., "sign me up," "I think so")?

- Distribution of FPS pronouns: What kinds of positions, constructions, and discourse types do different FPS pronouns appear in (e.g., subject vs. object position, redundantly with indicating verbs, in constructed action/dialogue)?

- When content can be construed with/without FPS (construal or pronouns), how do participants phrase things (see two-sentence example in Chapter 2 under "FPS construal without pronouns" and "super hero" example in my dissertation defense video[126])?

- What are some of the different hand configurations used for each FPS pronoun?

- In which environments does the distinction between IX_1 and POSS_1 become blurred? For example, indicating verbs with less marked handshapes such as TELL and GIVE, or assimilation with neighboring words like FEEL and HAVE. Also, how is our current understanding of the distribution of these words circular?

[125] This was the original plan for my dissertation, before I realized that there was so much more foundational work to do before embarking on that next stage.

[126] https://figshare.com/projects/First-person_singular_FPS_pronouns_and_social_status/221932

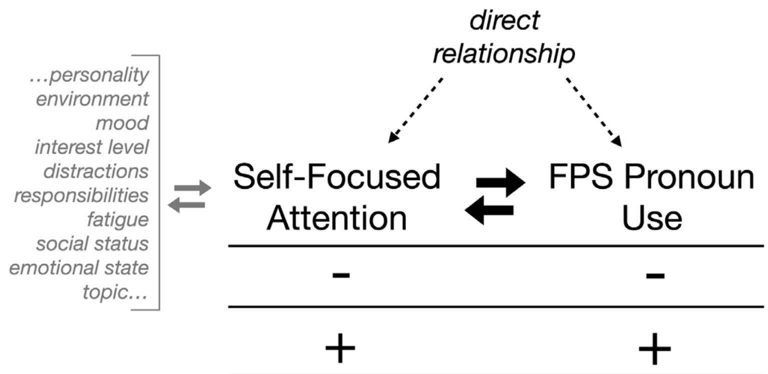
Conclusion

This dissertation was originally designed as a cross-linguistic study, testing whether a hypothesis found in the literature for English language data would be supported by ASL language data. After testing the hypothesis using ASL data I collected, as well as reexamining previous English data, I am skeptical of the original hypothesis. The connection between attention and language use is well-attested, as discussed in Chapter 2. However, the use of relative social status as a proxy for attention, appears to be an oversimplification. The data I have personally worked with is admittedly quite limited, in both ASL and English. That said, after looking more closely at the broader existing literature purporting an inverse relationship between social status and FPS pronoun use, I remain doubtful about the strength of the evidence presented therein.

Attention is influenced by numerous complex factors, and exploring each is outside the scope of this work. Suffice to say, relative social status is but one bundle of factors among many more that may influence attentional focus at any given time. Figure 5.09 below, a modification of Figure 2.05 from Chapter 2, provides a visual representation of this idea. It shows a reciprocal relationship between self-focused attention and FPS pronoun use, and also a reciprocal relationship between self-focused attention and various attentional influences.

Figure 5.09

Nature of the relationships between (1) various attentional factors and self-focused attention and (2) self-focused attention and FPS pronoun use



I hope this work underscores the importance of research reproducibility, particularly in the context of hypothesis-generating studies. It is encouraging to witness the growing emphasis on transparency and practices like pre-registration, which help mitigate publication bias in linguistics and related fields (Gawne & Styles, 2022). While much of the linguistic research involving signed languages tends to be exploratory rather than confirmatory, the need for transparency and rigor remains essential, and I have aimed to uphold these principles in my analyses. As to the potential for a relationship between social status and FPS pronoun use: The results of my analyses leave this matter unresolved, suggesting ample room for future inquiry.

References

Adler, Nancy E., Elissa S. Epel, Grace Castellazzo, & Jeannette R. Ickovics. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Health Psychology*, *19*, 586–592.

Ahlgren, Inger. (1990). Deictic pronouns in Swedish and Swedish Sign Language. In Susan D. Fischer & Patricia Siple (Eds.), *Theoretical issues in sign language research: Linguistics* (Vol. 1, pp. 167–174). University of Chicago Press.

Ann, Jean. (2001). Bilingualism and language contact. In Ceil Lucas (Ed.), *The sociolinguistics of sign languages* (pp. 33–60). Cambridge University Press.

Baker, Mona. (1993). Corpus linguistics and translation studies: Implications and applications. In Mona Baker, Gill Francis, & Elena Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 233–250). John Benjamins.

Baker, Mona. (1997). Corpus-based translation studies: The challenges that lie ahead. In Harold Somers (Ed.), *Terminology, LSP and translation studies in language engineering in honour of Juan C. Sager* (pp. 175–186). John Benjamins.

Balhorn, Mark. (2004). The rise of epicene they. *Journal of English Linguistics*, *32*(2), 79–104.

Bauman, H-Dirksen. (2009). Postscript: Gallaudet protests of 2006 and the myths of in/exclusion. *Sign Language Studies, 10*(1), 90–104. https://www.jstor.org/stable/26190680

Bayley, Robert, Ceil Lucas, & Mary Rose. (2002). Phonological variation in American Sign Language: The case of 1 handshape. *Language Variation and Change*, *14*, 19–53.

Beach, Jennifer C., Sheila Brownlow, Melissa D. Greene, N. Clayton Silver. (2016). The "I"s have it: Sex and social status differences on Twitter. *Social Behavior Research and Practice Open Journal*, *1*(1), 17–21. https://doi.org/10.17140/SBRPOJ-1-104

Becker, Amelia A., Julie A. Hochgesang, Meredith Tamminga, & Jami N Fisher. (2023). Sociophonetics and signed languages. In Christopher Strelluf (Ed.), *The Routledge handbook of sociophonetics* (pp. 467–488). Routledge. https://doi.org/10.4324/9781003034636

Berez-Kroeker, Andrea L., Helene N. Andreassen, Lauren Gawne, Gary Holton, Susan Smythe Kung, Peter Pulsifer, Lauren B. Collister, The Data Citation and Attribution in Linguistics Group, & the Linguistics Data Interest Group. (2018). *The Austin Principles of Data Citation in Linguistics*. Version 1.0. http://site.uit.no/linguisticsdatacitation/austinprinciples/ Accessed [October 14, 2024].

Berry, Donald A. & Gregory D. Ayers. (2006). Symmetrized Percent Change for Treatment Comparisons. *The American Statistician*, *60*(1), 27–31.

Berry, Diane S., James W. Pennebaker, Jennifer S. Mueller, Wendy S. Hiller. (1997). Linguistic Bases of Social Perception. *Personality and Social Psychology Bulletin, 23*(5), 526–537.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, & Edward Finegan. (1999). *Longman grammar of written and spoken English*. Longman Publications Group.

Bloom, Carrie Lou, Jeffery Levi Palmer, & Jonah Winninghoff. 2024. Deaf Postsecondary Data from the American Community Survey [Data visualization tool]. National Deaf Center on Postsecondary Outcomes, University of Texas at Austin. www.nationaldeafcenter.org/dashboard

Bodine, Ann. (1975). Androcentrism in prescriptive grammar: Singular ‘they’, sex-indefinite ‘he’, and ‘he or she’. *Language in society*, *4*(2), 129–146.

Börstell, Carl, Adam Charles Schembri, & Onno Crasborn. (2024). Sign duration and signing rate in British Sign Language, Dutch Sign Language and Swedish Sign Language. *Glossa Psycholinguistics*, *3*(1), 1–15. https://doi.org/10.5070/G60111915

Börstell, Carl. (2022). Searching and utilizing corpora. In J. Fenlon & J. A. Hochgesang (Eds.), *Signed language corpora* (pp. 90–127). Gallaudet University Press.

Boyes-Braem, Penny. (1990). Acquisition of the handshape in American Sign Language: A preliminary analysis. In V. Volterra & C. J. Erting (Eds.), *From gesture to language in hearing and deaf children* (pp. 107–127). Gallaudet University Press.

Brockmeyer, Timo, Johannes Zimmermann, Dominika Kulessa, Martin Hautzinger, Hinrich Bents, Hans-Christoph Friederich, Wolfgang Herzog, & Matthias Backenstrass. (2015). Me, myself, and I: Self-referent word use as an indicator of self-focused attention in relation to depression and anxiety. *Frontiers in Psychology*, *6*, Article 1564.

Brown, Penelope & Stephen C. Levinson. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge university press.

Burris, Ethan R., Matthew S. Rodgers, Elizabeth A. Mannix, Michael G. Hendron, & James B. Oldroyd. (2009). Playing favorites: The influence of leaders’ inner circle on group processes and performance. *Personality and Social Psychological Bulletin*, *35*, 1244–1257.

Carver, Charles S., & Michael Scheier. (1981). *Attention and self-regulation: a control-theory approach to human behavior*. Springer-Verlag.

Casillas, Marisa, Connie de Vos, Onno Crasborn, & Stephen C. Levinson. (2015). The perception of stroke-to-stroke turn boundaries in signed conversation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 37. Retrieved from https://escholarship.org/uc/item/192318fb

Cheepen, Christine. (1988). *The predictability of informal conversation*. Bloomsbury Academic.

Cheng, Joey T., Jessica L. Tracy, Tom Foulsham, Alan Kingstone, & Joseph Henrich. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*, 103–125.

Chung, Cindy K., & James W. Pennebaker. (2007). The psychological functions of function words. In Klaus Fiedler (Ed.), *Social communication* (pp. 343–359). Psychology Press.

Coates, Jennifer, & Rachel Sutton-Spence. (2001). Turn-taking patterns in deaf conversation. *Journal of Sociolinguistics*, *5*(4), 507–529.

Cohen, Jacob. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cohen, Sheldon, Cuneyt M. Alper, William J. Doyle, Nancy Adler, Josh J. Treanor, & Ronald B. Turner. (2008). Objective and subjective socioeconomic status and susceptibility to the common cold. *Health Psychology*, *27*, 268–274.

Cole, Tim J, & Douglas G Altman. (2017). Statistics notes: What is a percentage difference? *BMJ, 358*. https://doi.org/10.1136/bmj.j3663.

CODATA-ICSTI Task Group on Data Citation Standards and Practices. (2013). Out of cite, out of mind: The current state of practice, policy, and technology for the citation of data. *Data Science Journal*, *12*, CIDCR1–CIDCR7. https://doi.org/10.2481/dsj.OSOM13-043

Cooperrider, Kensy, & Katherine Mesh. (2022). Pointing in gesture and sign. In Aliyah Morgenstern & Susan Goldin-Meadow (Eds.), *Gesture in language: Development across the lifespan* (pp. 21–46). American Psychological Association. https://doi.org/10.1037/0000269-002

Cormier, Kearsy, Adam Schembri, & Bencie Woll. (2010). Diversity across sign languages and spoken languages: Implications for language universals. *Lingua*, *120,* 2664–2667.

Cormier, Kearsy, Adam Schembri, & Benice Woll. (2013). Pronouns and pointing in sign languages. *Lingua*, *137*, 230–247.

Croneberg, Carl G. (1965). Sign language dialects (Appendix D). In William. C. Stokoe, Dorothy. C. Casterline & Carl. G. Croneberg (Eds.), In *A dictionary of American Sign Language on linguistic principles*. Gallaudet College Press.

Culkin, Macaulay [@IncredibleCulk]. (2020, August 26). "Hey guys, wanna feel old? I'm 40. You're welcome" [Tweet]. Twitter. https://x.com/IncredibleCulk/status/1298730289737293824

Davies, Mark. (2008-). *The Corpus of Contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/

Davis, Deborah & Timothy C. Brock. (1975). Use of first person pronouns as a function of increased objective self-awareness and performance feedback. *Journal of Experimental Social Psychology*, *11*, 381–388.

DeVellis, Robert F. 2003. *Scale development: Theory and applications.* Sage Publications.

Dino, Amanda, Stephen Reysen, & Nyla R. Branscomb. (2009). Online interactions between group members that differ in status. *Journal of Language and Social Psychology*, *28*, 85–93.

Dorminy, Jerri Lyn. (2013). *The experiences of non-signing deaf and hard-of-hearing students and their academic and social integration into a primarily signing deaf university environment*. [Doctoral dissertation, University of Maryland]. ProQuest.

Duguine, Maia. (2017). Reversing the approach to null subjects: A perspective from language acquisition. *Frontiers in Psychology*, *8*, Article 27.

Duval, Shelley, & Robert A. Wicklund. (1972). *A theory of objective self awareness*. Academic Press.

ELAN (Version 6.8) [Computer software]. (2024). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan

Evans, Vyvyan & Melanie Green. (2006). *Cognitive linguistics: An introduction*. Edinburgh University Press.

Evans, Nicholas, & Stephen C. Levinson. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioural and Brain Sciences*, *32*, 429–492.

Fenlon, Jordan, Adam Schembri, Ramas Rentelis, & Kearsy Cormier. (2013). Variation in handshape and orientation in British Sign Language: The case of the “1" hand configuration. *Language and Communication*, *33*(1), 69–91.

Fenlon, Jordan, Kensy Cooperrider, Jon Keane, Diane Brentari, & Susan Goldin-Meadow. (2019). Comparing sign language and gesture: Insights from pointing. *Glossa: A Journal of General Linguistics*, *4(*1*)*, Article 2. http://doi.org/10.5334/gjgl.499

Fernandes, Jane K., & Shirley Shultz Myers. (2010). Inclusive deaf studies: Barriers and pathways. *Journal of Deaf Studies and Deaf Education*, *15*(1), 17–29. https://www.jstor.org/stable/43666251

Fisher, Jami N., Julie A. Hochgesang, Meredith Tamminga, & Robyn Miller. (2021). Uncovering the lived experiences of elderly Deaf Philadelphians. In Roland Pfau, Aslı Göksel & Jana Hosemann (Eds.), *Our lives – Our stories: Life experiences of elderly deaf people* (pp. 277–322). De Gruyter Mouton. https://doi.org/10.1515/9783110701906

Fowler, Floyd J., Jr. (1995). *Improving survey questions: Design and evaluation.* Sage Publishing.

Friedman, Lynn. (1975). Space and time reference in American Sign Language. *Language*, *51*(4), 940–961.

Friedner, Michele, & Annelies Kusters. (2015). DEAF-SAME and difference in international deaf spaces and encounters. In Michele Friedner & Annelies Kusters (Eds.), *It's a small world: International deaf spaces and encounters* (1st ed., pp. ix–xxix). Gallaudet University Press.

Gallaudet University. (2023, December 21). *Information for VR Counselors*. https://gallaudet.edu/undergraduate-admissions/information-for-vr-counselors/

Gallaudet University. (2024, January 3). *American Sign Language Proficiency Interview (ASLPI)*. https://gallaudet.edu/american-sign-language-proficiency-interview-aslpi/

Gagne, Deanna L. (2017). *With a little help from my friends: The contributions of a peer language network on the conventionalization of space in an emerging language* (Publication No. 1493) [Doctoral Dissertation, University of Connecticut]. Digital Commons at UConn. https://digitalcommons.lib.uconn.edu/dissertations/1493

Gawne, Lauren, & Suzy Styles. (2022). Situating linguistics in the social science data movement. In Andrea L.Berez-Kroeker, Bradley McDonnell, Eve Koller, & Lauren B. Collister

(Eds.), *The open handbook of linguistic data management* (pp. 9–26). MIT Press Open. doi.org/10.7551/mitpress/12200.003.0006

Glickman, Neil S., & John C. Carey. (1993). Measuring deaf cultural identities: A preliminary investigation. *Rehabilitative Psychology*, *38*, 275–283.

Goodman, Steven N, Daniele Fanelli, & John P. A. Ioannidis. (2016). What does research reproducibility mean? *Science Translational Medicine*, *8*(341). https://doi.org/10.1126/scitranslmed.aaf5027

Greenberg, Steven A. (2009). How citation distortions create unfounded authority: Analysis of a citation network. *BMJ: British Medical Journal*, *339*(7714), 210–213.

Greenland, Sander. (2019) Valid P-values behave exactly as they should: Some misleading criticisms of P-values and their resolution with S-values. *The American Statistician*, *73*(S1), 106–114. https://doi.org/10.1080/00031305.2018.1529625

Grusky, David B. (Ed.). (2019). *Social stratification: Class, race, and gender in sociological perspective* (4th ed.). Routledge.

Hall, Matthew L., Wyatte C. Hall, & Naomi K. Caselli. (2019). Deaf children need language, not (just) speech. *First Language*, *39*(4), 367–395.

Hall, Stephanie. (1983). Train-Gone-Sorry: The etiquette of social conversations in American Sign Language. *Sign Language Studies*, *41*, 291–309. https://doi.org/10.1353/sls.1983.0006

Hancock, Jeffery T., David I. Beaver, Cindy K. Chung, Joey Frazee, James W. Pennebaker, Art Graesser, & Zhiqiang Cai. (2010). Social language processing: A framework for analyzing the communication of terrorists and authoritarian regimes. *Behavioral Sciences of Terrorism and Political Aggression*, *2*(2), 108–132.

Hill, Joseph. (2012). *Language attitudes in the American Deaf community*. Gallaudet University Press.

Hill, Joseph. (2014). Data collection in sociolinguistics. In Eleni Orfanidou, Bencie Woll, & Gary Morgan (Eds.), *Research methods in sign language studies: A practical guide* (pp. 193–205). John Wiley & Sons.

Hill, Joseph C. (2015). Language attitudes in deaf communities. In Adam C. Schembri & Ceil Lucas (Eds.), *Sociolinguistics and deaf communities* (pp. 146–174). Cambridge University Press.

Hochgesang, Julie A. (2014). Using design principles to consider representation of the hand in some notation systems. *Sign Language Studies*, *14*(4), 488–542. https://www.jstor.org/stable/26191225

Hochgesang, Julie A. (2020). *SLAASh ID Glossing Principles, ASL Signbank and Annotation Conventions* (Version 3.0). Figshare. https://doi.org/10.6084/m9.figshare.12003732.v1

Hochgesang, Julie A. (2022, Jan 6). *Documenting signed language use while considering our spaces as a Deaf* linguist*. The 96th Annual Meeting of the Linguistic Society of America. https://doi.org/10.6084/m9.figshare.17911973.v1

Hochgesang, Julie A., Onno Crasborn, and Diane Lillo-Martin. (2024). *ASL Signbank*. New Haven, CT: Haskins Lab, Yale University. https://aslsignbank.haskins.yale.edu/

Hochgesang, Julie A., Ryan Lepic, & Emily Shaw. (2023). W(h)ither the ASL corpus?: Considering trends in signed corpus development. In Ella Wehyrmeyer (Ed.), *Gaining ground in sign language corpus linguistics* (pp. 287–308). John Benjamins. https://doi.org/10.1075/scl.108.11hoc

Hoza, Jack. (2007). *It's not what you sign, it's how you sign it: Politeness in American Sign Language*. Gallaudet University Press.

Huang, Chin-Lan, Cindy K. Chung, Natalie K. Hui, Yi-Cheng Lin, Yi-Tai Seih, Ben C.P. Lam, Wei-Chuan Chen, Michael H. Bond, & James W. Pennebaker. (2012). 中文版「語文探索與字詞計算」詞典之建立 [The development of the Chinese Linguistic Inquiry and Word Count dictionary]. *Chinese Journal of Psychology*, *54*, 185–201.

Johnston, Trevor. (2010). From archive to corpus: Transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics*, *15*(1), 106–131.

Johnston, Trevor. (2013). Formational and functional characteristics of pointing signs in a corpus of Auslan (Australian Sign Language): Are the data sufficient to posit a grammatical class of 'pronouns' in Auslan? *Corpus Linguistics and Linguistic Theory*, *9*(1), 109–159.

Johnston, Trevor A., & Adam Schembri. (1999). On defining lexeme in a signed language. *Sign Language & Linguistics*, *2*(2), 115–185.

Johnstone, Barbara. (2008). *Discourse analysis*. Blackwell Publishing.

Kacewicz, Ewa. (2013). *Language as a marker of CEO transition and company performance*. [Doctoral dissertation, The University of Texas at Austin]. Texas ScholarWorks. https://repositories.lib.utexas.edu/server/api/core/bitstreams/8d75c058-3d1a-421b-9a0c-41fde7c5d5b4/content

Kacewicz, Ewa, James W. Pennebaker, Matthew Davis, Moongee Jeon, & Arthur Graesser. (2013). Pronoun use reflects standings in social hierarchies. *Journal of language and social psychology*, *33*(2), 1–19. https://doi.org/10.1177/0261927X1350265

Kacewicz, Ewa, James W. Pennebaker, Matthew Davis, Moongee Jeon, & Arthur Graesser. (2014). Pronoun use reflects standings in social hierarchies. *Journal of language and social psychology*, *33*(2), 125–143.

Kannapell, Barbara. (1989). Inside the deaf community. In Sherman Wilcox (Ed.), *American deaf culture: An anthology*. Linstok Press.

Kaplan, Abby. (2016). *Women talk more than men.* Cambridge University Press.

Kendon, Adam. (2008). Some reflections on the relationship between `gesture' and `sign.' *Gesture*, *8*(3), 348–366.

Kullavanijaya, Pranee. (2004). Pro-forms. In Philipp Strazny (Ed.), *Encyclopedia of linguistics*. Routledge.

Kusters, Annelies, Maartje De Meulder, & Dai O'Brien. (2017). *Innovations in deaf studies: The role of deaf scholars*. Oxford University Press.

Ladd, Paddy. (2003). *Understanding deaf culture: In search of Deafhood*. Multilingual Matters.

Lakoff, George, & Mark Johnson. (1980). *Metaphors we live by*. University of Chicago Press.

Lane, Harlan. (1992). *The mask of benevolence*. Vintage Books.

Lang, Thomas A., & Doug G. Altman. (2016). Statistical analyses and methods in the published literature: The SAMPL guidelines. *Medical Writing*, *25*(3), 31–36.

Langacker, Ronald W. (1987). *Foundations of cognitive grammar: Vol. I. Theoretical prerequisites*. Stanford University Press.

Langacker, Ronald W. 1991. *Foundations of cognitive grammar: Vol. II. Descriptive application*. Stanford University Press.

Langacker, Ronald W. (2007). Constructing the meanings of personal pronouns. In Günter Radden, Klaus-Michael Kopcke, Thomas Berg & Peter Siemund (Eds.), *Aspects of meaning construction* (pp. 171–187). John Benjamins.

Langacker, Ronald W. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.

Lederberg, Amy R., Brenda Schick, & Patricia E. Spencer. (2013). Language and literacy development of deaf and hard-of-hearing children: Successes and challenges. *Developmental Psychology*, *49*(1), 15–30. https://doi.org/10.1037/a0029558

Leech, Geoffrey, Paul Rayson, & Andrew Wilson (2014a). *Word frequencies in written and spoken English: Based on the British National Corpus*. Routledge.

Leech, Geoffrey, Paul Rayson, & Andrew Wilson (2014b). *Companion website for: Word frequencies in written and spoken English: Based on the British national corpus* [Data set]. Pearson Education. https://ucrel.lancs.ac.uk/bncfreq/

Liddell, Scott. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press.

Lillo-Martin, Diane. (2018, Oct 22). *Development of pointing signs in Deaf and Koda children*. Presentation at Gallaudet University, Washington, DC.

Lillo-Martin, Diane, & Edward S. Klima. (1990). Pointing out differences: ASL pronouns in syntactic theory. In Susan D. Fischer & Patricia Siple (Eds.), *Theoretical issues in sign language research.* (Vol. 1: Linguistics, 191–210). University of Chicago Press.

Lillo-Martin, Diane C., Elaine Gale, & Deborah Chen Pichler. (2023). Family ASL: An early start to equitable education for deaf children. *Topics in Early Childhood Special Education*, *43*(2), 156–166. https://doi.org/10.1177/02711214211031307

Mahl, George F. (1987). *Explorations in nonverbal and vocal behavior*. Lawrence Erlbaum Associates, Inc.

Manrique, Elizabeth, & N. J. Enfield. (2015). Suspending the next turn as a form of repair initiation: Evidence from Argentine Sign Language. In Judith Holler, Kobin H. Kendrick, Marisa Casillas & Stephen C. Levinson (Eds.), *Turn-taking in human communicative interaction*. Frontiers in Psychology.

Markowitz, David M. (2018). Academy awards speeches reflect social status, cinematic roles, and winning expectations. *Journal of language and social psychology*, *37*(3), 376–387.

Maxwell-McCaw, Deborah & Maria Cecilia Zea. (2011). The Deaf acculturation scale (DAS): Development and validation of a 58-item measure. *The Journal of Deaf Studies and Deaf Education*, *16*(3), 325–342.

Mayberry, Rachel I., & Robert Kluender. (2018). Rethinking the critical period for language: New insights into an old question from American Sign Language. *Bilingualism: Language and Cognition*, *21*(5), 886–905. https://doi.org/10.1017/S1366728917000724

McBurney, Susan Lloyd. (2002). Pronominal reference in signed and spoken language: Are grammatical categories modality-dependent? In Richard P. Meier, Kearsy A. Cormier & David Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 329–369). Cambridge University Press.

McIntire, Marina L. (1977). The acquisition of American Sign Language hand configurations. *Sign Language Studies*, *16*, 247–266.

McKee, Rachel, Adam Schembri, David McKee, & Trevor Johnston. (2011). Variable "subject" presence in Australian Sign Language and New Zealand Sign Language. *Language Variation and Change*, *23*(3), 375–398.

Meier, Richard P. (1990). Person deixis in ASL. In Susan D. Fischer & Patricia Siple (Eds.), *Theoretical issues in sign language research* (Vol. 1: Linguistics, 175–190). University of Chicago Press.

Metzger, Melanie. (1995). Constructed dialogue and constructed action in American Sign Language. In Ceil Lucas (Ed.), *Sociolinguistics in Deaf Communities* (pp. 255–271). Gallaudet University Press.

Mitchell, Ross E., & Michael A. Karchmer. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, *4*(2), 138–163.

Morford, Jill P., & James Macfarlane. (2003). Frequency characteristics of American Sign Language. *Sign Language Studies*, *3*(2), 213–225.

Morris, Carla Dee. (2016). *Examining the associations between social network ties and linguistic production* (UMI Number: 10612715) [Doctoral dissertation, Gallaudet University]. ProQuest Dissertations and Theses Global.

Moulton, Keir, Chung-hye Han, Trevor Block, Holly Gendron, & Sander Nederveen. (2020). Singular *they* in context. *Glossa: A Journal of General Linguistics*, *5*(1), 1–13. https://doi.org/10.5334/gjgl.1012

Neidle, Carol, Judy Kegl, Dawn MacLaughlin, Benjamin Bahan, & Robert G. Lee. (2000). *The syntax of American Sign Language: Functional categories and hierarchical structure*. MIT Press.

Newman, Matthew L., James W. Pennebaker, Diane S. Berry, & Jane M. Richards. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, *29*, 665–675.

Newman, Matthew L., Carla J. Groom, Lori D. Handelman, & James W. Pennebaker. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes, 45*, 211–236.

Nietzsche, Friedrich. (1967). *The will to power* (Walter Kaufmann & R. J. Hollingdale, Trans.). Random House. (Original work published 1901).

Nilsson, Anna-Lena. (2004). Form and discourse function of the pointing toward the chest in Swedish Sign Language. *Sign Language & Linguistics*, *7*(1), 3–30.

Nuzzo, Regina. (2014). Scientific method: Statistical errors. *Nature*, *506*, 150–152. https://doi.org/10.1038/506150a

Nuzzo, Regina. (2018). Percent differences: Another look. *PM&R*, *10*(6), 661–664. https://doi.org/10.1016/j.pmrj.2018.05.003

Occhino, Corrine, & Sherman Wilcox. (2017). Gesture or sign? A categorization problem. [Peer commentary on “Gesture, sign and language: The coming of age of sign language and gesture studies," by Susan Goldin-Meadow & Diane Brentari]. *Behavioral and Brain Sciences*, *40*, 36–37. https://doi.org/10.1017/S0140525X15003015

Ormel, Ellen, & Onno Crasborn. (2011). Prosodic correlates of sentences in signed languages: A Literature Review and Suggestions for New Types of Studies. *Sign Language Studies*, *12*(2), 279–315. https://doi.org/10.1353/sls.2011.0019

Padden, Carol. (1989). The Deaf Community and the culture of Deaf people. In Sherman Wilcox (Ed.), *American Deaf culture: An anthology*. Linstok Press.

Pennebaker, James W. (2011). *The secret life of pronouns: What our words say about us*. Bloomsbury Press.

Pennebaker, James W. (2022). Computer-based language analysis as a paradigm shift. In Morteza Dehghani & Ryan L. Boyd (Eds.), *Handbook of language analysis in psychology*, (pp 576–587). The Guilford Press. https://www.dropbox.com/s/rj7lcoyq3dvyy0o/Pennebaker_ParadigmShift_2022.pdf?dl=0

Pennebaker, James W., & Laura A. King. (1999). Linguistic Styles: Language use as an individual difference. *Journal of Personally and Social Psychology*, *77*(6), 1296–1312.

Pennebaker, James W., Matthias Mehl, & Kate Niederhoffer. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*, 547–577.

Pennebaker, James W., Roger J. Booth, Ryan L. Boyd, & Martha E. Francis. (2015). Linguistic Inquiry and Word Count: LIWC 2015 [Computer software]. Pennebaker Conglomerates, Inc. Retrieved from http://liwc.net

Pennebaker, James W., & Lori D Stone. (2003). Words of wisdom: Language use over the lifespan. *Journal of Personality and Social Psychology*, *85*, 291–301.

Pinker, Steven & Paul Bloom. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, *13*(4), 707–727.

Piolat, Annie, Roger J. Booth, Cindy K. Chung, Morgana Davids, & James W. Pennebaker. (2011). La version française du LIWC : Modalités de construction et exemples d'utilisation [The French dictionary for LIWC: Modalities of construction and examples of use]. *Psychologie Française*, *56*, 145–159.

Rafi, Zad, & Sander Greenland. (2020). Semantic and cognitive tools to aid statistical science: Replace confidence and significance by compatibility and surprise. *BMC: Medical Research Methodology*, *20*(244). https://doi.org/10.1186/s12874-020-01105-9

Ramirez-Esparza, Nairán, James W. Pennebaker, Florencia Andrea Garcia, & Raquel Suria. (2007). La psicología del uso de las palabras: Un programa de computadora que analiza textos en Español (The psychology of word use: A computer program that analyzes texts in Spanish). *Revista Mexicana de Psicología*, *24*, 85–99.

Reynolds, Wanette. (2016). *Early bimodal bilingual development of ASL narrative referent cohesion*. [Doctoral dissertation, Gallaudet University].

Reysen, Stephen, Jason D. Lloyd, Iva Katzarska-Miller, Brett M. Lemker, & Russell L. Foss. (2010). Intragroup status and social presence in online fan groups. *Computers in Human Behavior*, *26*(6), 1314–1317.

Schooler, Carmi. (2013). *Social class and social status.* Oxford Bibliographies Online. https://dx.doi.org/10.1093/obo/9780199828340-0085

Silberzahn, Raphael, Eric Luis Uhlmann, Dan Martin, Pasquale Anselmi, Frederik Aust, Eli Awtrey, Štěpán Bahník, Feng Bai, Colin Bannard, Evelina Bonnier, Rickard Carlsson, Felix Cheung, Garret Christensen, Russ Clay, Maureen A. Craig, Anna Dalla Rosa, Lammertjan Dam, Mathew H. Evans, Ismael Flores Cervantes, & Brian A. Nosek. (2018). Many Analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337–356. https://doi.org/10.1177/2515245917747646

Sloan, Vivion Smith. (2013). *ASL nominal constructions involving signs that resemble pronouns* (UMI Number: 3576358) [Doctoral dissertation, Gallaudet University]. ProQuest Dissertations and Theses Global.

Sloman, Leon, Paul Gilbert, & Gary Hasey. (2003). Evolved mechanisms in depression: The role and interaction of attachment and social rank in depression. *Journal of Affective Disorders*, *74*(2), 107–121.

Sorace, Antonella, Ludovica Serratrice, Francesca Filiaci, & Michela Baldo. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, *119*, 460–477.

Stokoe, William C., Dorothy C. Casterline, & Carl G. Croneberg. (1965). *A dictionary of American Sign Language on linguistic principles*. Gallaudet College Press.

Tannen, Deborah. (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In Florian Coulmas (Ed.), *Direct and indirect speech* (pp. 311–332). Mouton.

Tannen, Deborah. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge University Press.

Tannen, Deborah. (1994). *Gender and discourse*. Oxford University Press.

Taub, Sarah F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge University Press.

Tausczik, Yla R., & James W. Pennebaker. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*, 24–54.

Taylor, John R. (2015). *The Oxford handbook of the word*. Oxford University Press.

Thompson, Robin L., Karen Emmorey, & Robert Kluender. (2009). Learning to look: The acquisition of eye gaze agreement during the production of ASL verbs. *Bilingualism: Language and Cognition*, *12* (4), 393–409.

Thumann, Mary Agnes. (2010). *Identifying depiction in American Sign Language presentations* (UMI Number: 3425466) [Doctoral dissertation, Gallaudet University]. ProQuest Dissertations & Theses Global.

Tourangeau, Roger, Lance J. Rips, & Kenneth Rasinski. (2000). *The psychology of survey response*. Cambridge University Press.

University of Oxford, Faculty of Linguistics, Philology and Phonetics. (2022). *British National Corpus* [Data set]. http://www.natcorp.ox.ac.uk/

U.S. Census Bureau. (2019). *2018: ACS 5-Year Estimates Data Profiles*. U.S. Department of Commerce. Retrieved from https://data.census.gov/table?tid=ACSDP5Y2018.DP05

Villanueva, Miako Noel. (2010). *Non-agent focused constructions in American Sign Language: A cognitive linguistic analysis* (UMI: 3441296) [Doctoral dissertation, Gallaudet University]. ProQuest Dissertations and Theses Global.

Whitworth, Cecily. (2011). Features and natural classes in ASL handshapes. *Sign Language Studies*, *12*(1), 46–71. http://www.jstor.org/stable/26190823

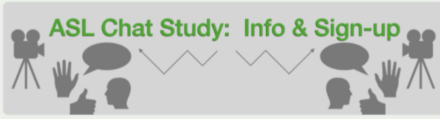
Wolf, Markus, Andrea Horn, Matthias R. Mehl, Severin Haug, James W. Pennebaker, & Hans Kordy. (2008). Computergestützte quantitative textanalyse: Äquivalenz und robustheit der deutschen version des Linguistic Inquiry and Word Count [Computer-aided quantitative text analysis: Equivalence and robustness of the German adaption of the Linguistic Inquiry and Word Count]. *Diagnostica*, *54*, 85–98.

Woods, Kevin M. (2007). *Iraqi Perspectives Project; Primary source materials for Saddam and terrorism: Emerging insights from captured Iraqi documents (Redacted)*. Institute for Defense Analysis. https://irp.fas.org/eprint/iraqi/index.html

Wulf, Alyssa, Paul Dudis, Robert Bayley, & Ceil Lucas. (2002). Variable subject presence in ASL narratives. *Sign Language Studies*, *3*(1), 54–76.

Zijlstra, Hanna, Tanja van Meerveld, Henriët van Middendorp, James W. Pennebaker, & Rinie Geenen. (2004). De Nederlandse versie van de 'Linguistic Inquiry and Word Count' (LIWC), een gecomputeriseerd tekstanalyseprogramma [Dutch version of the linguistic inquiry and word count (LIWC), a computerized text analysis program]. *Gedrag & Gezondheid*, *32*, 273–283.

Appendix A: Background Questionnaire

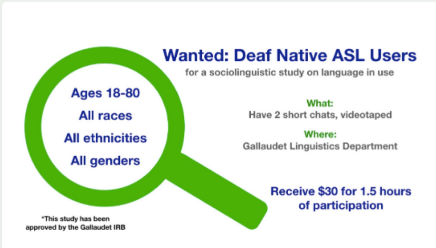


Sociolinguistic research - ASL in use

kiva.bennett@gallaudet.edu Switch account

Not shared

* Indicates required question



My name is Kiva Bennett, and I'm a PhD student in the department of linguistics at Gallaudet University. About my study:
Who: Deaf native ASL users ages 18-80 of various genders, races, ethnicities, and backgrounds.

What: Participants will have two short chats in pairs, both videotaped. I will analyze these videos under the guidance of my faculty advisor, Julie Hochgesang. They will be shared (as screenshots or short clips) in papers, presentations, etc. ONLY if the participants consent in writing.

When: Videotaped sessions will take place in January and February, 2018

Where: Gallaudet linguistics department

Time commitment: 1.5 hours per participant (including signing forms, setting up, etc.)

Compensation: $30 per participant

Feel free to contact me with any questions at Kiva.Bennett@gallaudet.edu or text, Glide, or Marco Polo me at

If you would like to participate, please answer the questions below, and I will contact you within 48 hours.
Your answers to the following questions will not be directly linked to your name in any papers or presentations resulting from this study.

Are you Deaf? *

- ○ Yes
- ○ No

Do you consider yourself a native signer of ASL? *

- ○ Yes
- ○ No

Age? *

Your answer

Gender? *

Your answer

Race and/or ethnicity? *

Your answer

Height? *

Your answer

At what age did you first learn ASL? *

Your answer

What kind of K-12 school(s) did you attend? *

- ○ Deaf residential school
- ○ Mainstream school
- ○ Other:

Think of the people you lived with growing up (family members, others). Please describe their signing skills/abilities. You can answer in English text or paste a link to a video response. *

Your answer

What is the highest level of education you have completed? *

- ○ High school
- ○ Some college
- ○ BA or BS degree
- ○ MA or MS degree
- ○ PhD
- ○ Other:

Where do you work and what is your current job title? *

Your answer

What states/countries have you lived in? *

Your answer

Please enter dates and times you are available between Sunday 2/18 and Thursday 2/22 *
Example: Tuesday 1/30 5pm-8pm, Wednesday 1/31 9am-11am, Saturday 2/3 10am-3pm

Your answer

Please enter dates and times you are available between Monday 2/26 and Saturday 3/3 *

Your answer

Please enter dates and times you are available between Sunday 3/4 and Saturday 3/10 *

Your answer

Any other comments and/or questions?

Your answer

Please enter your email address

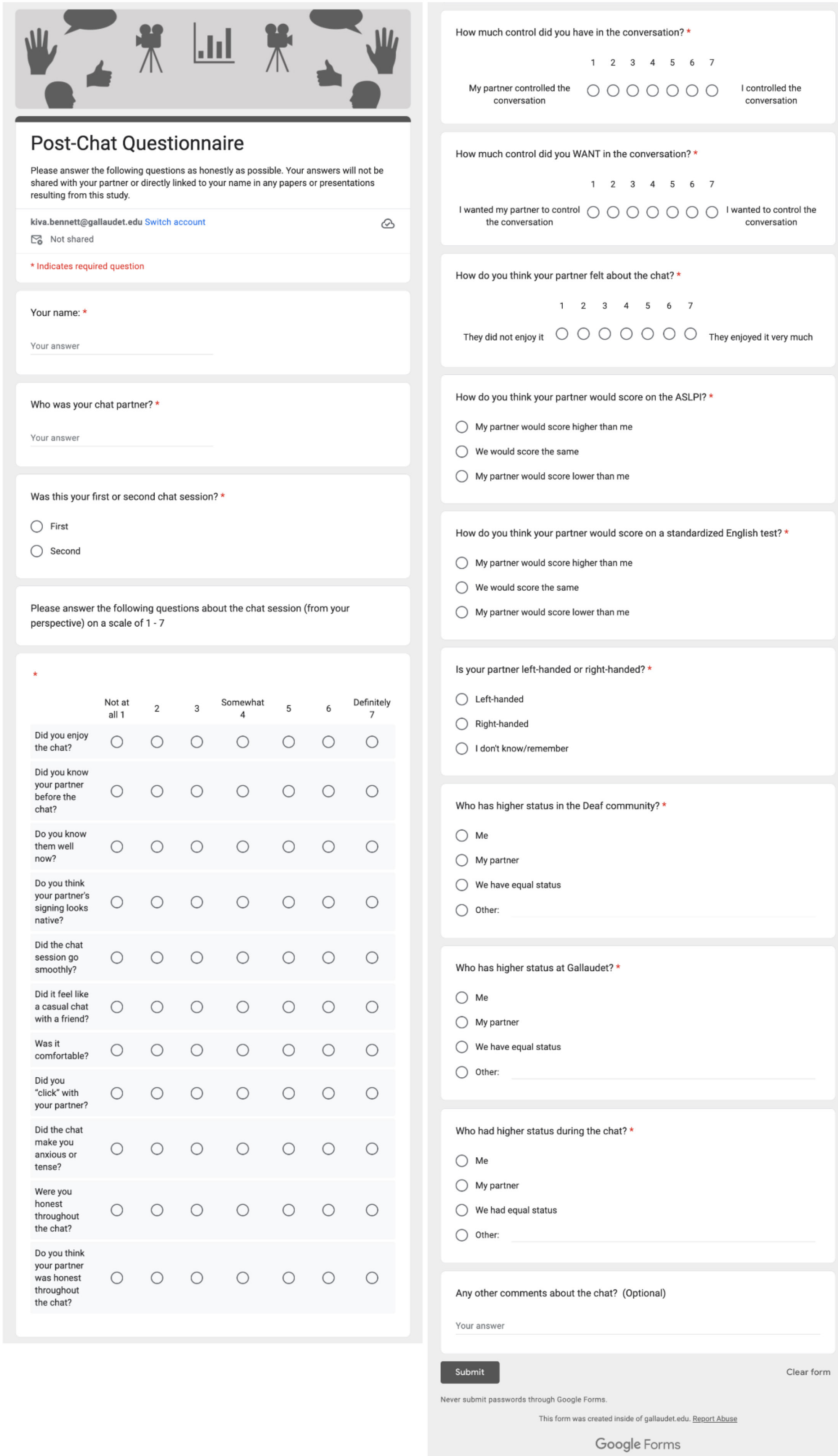
*

Your answer

Submit

Clear form

Appendix B: Post-Chat Survey



Post-Chat Questionnaire

Please answer the following questions as honestly as possible. Your answers will not be shared with your partner or directly linked to your name in any papers or presentations resulting from this study.

kiva.bennett@gallaudet.edu Switch account

Not shared

* Indicates required question

Your name: *

Your answer

Who was your chat partner? *

Your answer

Was this your first or second chat session? *

- First
- Second

Please answer the following questions about the chat session (from your perspective) on a scale of 1 - 7

*

	Not at all 1	2	3	Somewhat 4	5	6	Definitely 7
Did you enjoy the chat?	○	○	○	○	○	○	○
Did you know your partner before the chat?	○	○	○	○	○	○	○
Do you know them well now?	○	○	○	○	○	○	○
Do you think your partner's signing looks native?	○	○	○	○	○	○	○
Did the chat session go smoothly?	○	○	○	○	○	○	○
Did it feel like a casual chat with a friend?	○	○	○	○	○	○	○
Was it comfortable?	○	○	○	○	○	○	○
Did you "click" with your partner?	○	○	○	○	○	○	○
Did the chat make you anxious or tense?	○	○	○	○	○	○	○
Were you honest throughout the chat?	○	○	○	○	○	○	○
Do you think your partner was honest throughout the chat?	○	○	○	○	○	○	○

How much control did you have in the conversation? *

	1	2	3	4	5	6	7	
My partner controlled the conversation	○	○	○	○	○	○	○	I controlled the conversation

How much control did you WANT in the conversation? *

	1	2	3	4	5	6	7	
I wanted my partner to control the conversation	○	○	○	○	○	○	○	I wanted to control the conversation

How do you think your partner felt about the chat? *

	1	2	3	4	5	6	7	
They did not enjoy it	○	○	○	○	○	○	○	They enjoyed it very much

How do you think your partner would score on the ASLPI? *

- My partner would score higher than me
- We would score the same
- My partner would score lower than me

How do you think your partner would score on a standardized English test? *

- My partner would score higher than me
- We would score the same
- My partner would score lower than me

Is your partner left-handed or right-handed? *

- Left-handed
- Right-handed
- I don't know/remember

Who has higher status in the Deaf community? *

- Me
- My partner
- We have equal status
- Other:

Who has higher status at Gallaudet? *

- Me
- My partner
- We have equal status
- Other:

Who had higher status during the chat? *

- Me
- My partner
- We had equal status
- Other:

Any other comments about the chat? (Optional)

Your answer

Submit

Clear form

Never submit passwords through Google Forms.

This form was created inside of gallaudet.edu. Report Abuse

Google Forms

Appendix C: Interaction Rating Questionnaire (IRQ) from Kacewicz et al. (2013) – V1

INTERACTION RATING QUESTIONNAIRE

Please answer the following questions based on the interaction you just had with your partner using the following scale:

1	2	3	4	5	6	7
Not at all			Somewhat			A great deal

_____ 1. How much did you enjoy the interaction?

_____ 2. To what degree do you feel you got a good sense of the other person?

_____ 3. To what degree did the interaction go smoothly?

_____ 4. To what degree did you control the conversation?

_____ 5. To what degree was this interaction typical of one you would have with a friend?

_____ 6. To what degree was the interaction comfortable?

_____ 7. To what degree did you "click" with the other person?

_____ 8. To what degree did the interaction make you feel anxious or tense?

_____ 9. To what degree did you know the person before?

_____ 10. To what degree do you feel you now truly know the other person?

_____ 11. To what degree did your partner disclose personal information?

_____ 12. To what degree did you disclose personal information?

_____ 13. To what degree did you tell the truth to your partner throughout the interaction?

_____ 14. To what degree do you think your partner was honest with you throughout the interaction?

_____ 15. To what degree did you have more status than your partner in the conversation?

The following scales range from 1 (0%) to 7 (100%).

Please make two marks on each one: place an “x” to signify the actual amount, and an “o” to signify what the ideal amount would have been. Note, the “x” and the “o” can be in the same place.

On the first scale place an “x” to mark the percentage of control you had in the conversation. With an “o” mark what the ideal amount of control would have been. For example if you perceived yourself to be more in control than your partner place the “x” towards the left of the scale, but if you would have preferred more control, place the “o” further to the left. If there was no difference in the amount of control either you or your partner had, put the “x” in the middle of the scale (4).

1	4	7
I controlled the conversation	We had equal control	My partner controlled the conversation

On this next scale place an "x" to mark the degree to which you dominated the conversation. With an "o" mark what the ideal degree of dominance would have been.

1	4	7
I dominated conversation	Neither of us dominated	My partner dominated the conversation

Appendix D: Interaction Rating Questionnaire (IRQ) from Kacewicz et al. (2013) – V2

Participant ID:_________________

INTERACTION RATING QUESTIONNAIRE

Please answer the following questions based on the interaction you just had with your partner using the following scale:

1	2	3	4	5	6	7
Not at all			Somewhat			A great deal

_____ 1. How much did you enjoy the interaction?

_____ 2. To what degree do you feel you got a good sense of the other person?

_____ 3. To what degree did the interaction go smoothly?

_____ 4. To what degree did you control the conversation?

_____ 5. To what degree was this interaction typical of one you would have with a friend?

_____ 6. To what degree was the interaction comfortable?

_____ 7. To what degree did you “click” with the other person?

______ 8. To what degree did the interaction make you feel anxious or tense?

______ 9. To what degree did you know the person before?

______ 10. To what degree do you feel you now truly know the other person?

______ 11. To what degree did your partner disclose personal information?

______ 12. To what degree did you disclose personal information?

______ 13. To what degree did you tell the truth to your partner throughout the interaction?

______ 14. To what degree do you think your partner was honest with you throughout the interaction?

______ 15. To what degree did you have more status or power than your partner in the conversation?

Based on the conversation you just had, rate **your partner** on the following adjectives on the same scale of 1 (not at all) to 7(a great deal):

___1.Assertive

___2. Sympathetic

___3. Organized

___4. Nervous

___5.Deep

___6.Withdrawn

___7.Trustful

___8. Practical

___9. Irritable

___10. Creative

___11. Attractive

Rate **yourself** on the following adjectives using the same scale:

___1.Assertive

___2. Sympathetic

___3. Organized

___4. Nervous

___5.Deep

___6.Withdrawn

___7.Trustful

___8.Practical

___9. Irritable

___10. Creative

___11. Attractive

The following scales range from 1 (0%) to 7 (100%).

Please make two marks on each one: place an “x” to signify the actual amount, and an “o” to signify what the ideal amount would have been. Note, the “x” and the “o” can be in the same place.

On the first scale place an “x” to mark the percentage of control you had in the conversation. With an “o” mark what the ideal amount of control would have been. For example if you perceived yourself to be more in control than your partner place the “x” towards the left of the scale, but if you would have preferred more control, place the “o” further to the left. If there was no difference in the amount of control either you or your partner had, put the “x” in the middle of the scale (4).

1	4	7
I controlled the conversation	We had equal control	My partner controlled the conversation

On this next scale place an “x” to mark the degree to which you dominated the conversation. With an “o” mark what the ideal degree of dominance would have been.

1	4	7
I dominated conversation	Neither of us dominated	My partner dominated the conversation

Appendix E: Understanding IRQ Scores

Understanding IRQ Scores

With status understood as zero-sum and measured on a 7-point Likert scale

P1's self-report	...is an other-report of	Implied status discrepancy
1	7	-6
2	6	-4
3	5	-2
4	4	0
5	3	2
6	2	4
7	1	6

P2's self-report	...is an other-report of	Implied status discrepancy
1	7	6
2	6	4
3	5	2
4	4	0
5	3	-2
6	2	-4
7	1	-6

Appendix F: Understanding Participant Agreement

Understanding Participant Agreement

Colors are used here to visually connect the dyads between the charts; Shapes and/or labels also distinguish differences even when colors are unavailble to the viewer

As an example, look at the top chart, specifically the row for Agreement scores of 8. The **two blue &s** indicate Likert scores of 2 and 6.

The corresponding hypothetical dyad can be seen in the bottom chart, in the row with the **blue 8** in the "Ideal" section, along with the interpretation of the reported 2 and 6 scores

Hypothetical score combinations

For each possible total (2-14), possible combinations are shown as pairs of same-color symbols (i.e., #, X, &, @)

Likert Score >> / Agreement score	1	2	3	4	5	6	7	
14							XX	
13						X	X	<< How to read: One person reported 6 and the other reported 7
12					#	XX	#	
11				#	X	X	#	
10			&	#	XX	#	&	
9		&	#	X	X	#	&	**<< Reasonable agreement = 9**
8	@	&	#	XX	#	&	@	**<< Ideal agreement = 8** — These designations are explained in the chart below
7	&	#	X	X	#	&		**<< Reasonable agreement = 9**
6	&	#	XX	#	&			
5	#	X	X	#				<< How to read (green Xs): One person reported 1 and the other reported 4
4	#	XX	#					
3	X	X						
2	XX							

Interpreting the Likert scores and agreement scrores

11 is unacceptable, so not charting the more extreme scores (12-14)

	Is the status variable reliable?	Agreement score	A says: A =	B =	dom dif	B says: A =	B =	dom diff	Degree to which they disagree	Narrative: What's going on here?
	unacceptable	11	5	3	2	2	6	-4	6	Disgree who is higher, and to what degree // A thinks themself a bit higher; B thinks themself even higher
		11	4	4	0	1	7	-6	6	Equal vs extreme dif // A thinks they're equal; B thinks themself highest possible
Red 10s (both thought selves higher): Included in "sorta agreed"; excluded from "didn't disagree" >>	no	10	5	3	2	3	5	-2	4	They both think themselves a bit higher than the other
		10	4	4	0	2	6	-4	4	A thinks they're equal; B thinks themself significantly higher
		10	3	5	-2	1	7	-6	4	Agree who is higher, but to diff degrees // A thinks themself a bit lower, so B a bit higher; B thinks themself highest possible
	reasonable	9	4	4	0	3	5	-2	2	A thinks they're equal; B thinks themself a bit higher
		9	3	5	-2	2	6	-4	2	They agree who is higher, but B sees it as a bit bigger diff than A does
		9	2	6	-4	1	7	-6	2	They agree who is higher, but B sees it as a bit bigger diff than A does
	ideal	8	4	4	0	4	4	0	0	Agree that they're equal
		8	3	5	-2	3	5	-2	0	Agree who is higher and to what degree
		8	2	6	-4	2	6	-4	0	
		8	1	7	-6	1	7	-6	0	
	reasonable	7	3	5	-2	4	4	0	-2	Same as 9s, but flip A and B
		7	2	6	-4	3	5	-2	-2	
		7	1	7	-6	2	6	-4	-2	
Red 6s (both thought selves lower): Included in "sorta agreed"; excluded from "didn't disagree" >>	no	6	3	5	-2	5	3	2	-4	Same as 10s, but flip A and B
		6	2	6	-4	4	4	0	-4	
		6	1	7	-6	3	5	-2	-4	
	unacceptable	5	2	6	-4	5	3	2	-6	Same as 11s, but flip A and B
		5	1	7	-6	4	4	0	-6	

5 is unacceptable, so not charting more the extreme scores (2-4)