The Social Identity of Sounds

by

Savannah Jane Williams

(Under the Direction of Margaret E. L. Renwick)

Abstract

Some lexical classes are susceptible to the effects of sound symbolism, a hypothesized relationship where speech sounds represent non-phonetic properties. Sound symbolic principles are manifested in male and female personal names in English. While previous research found distinct differences between female and male names in English, the current study fills a gap in existing research by taking a perceptual approach. This thesis investigates potential phonological cues to name gender by employing an Internet survey asking participants to indicate the gender of a given name on a 5 point scale. Statistical analysis conducted via cumulative link model in R found that participants were more likely to assign a name to a female category if the name ended in a vowel; no other phonological characteristics were observed to be significant. This differs from results of previous corpus-based studies, and points both towards a gap between observed patterns in corpus data and the functional salience of English name phonology, and towards change over time in the phonological distribution of English first names.

Index words: [sound symbolism; onomastics; ordinal regression; language change]

The Social Identity of Sounds

by

Savannah Jane Williams

B.A., University of Georgia, 2022

A Thesis Submitted to the Graduate Faculty of the

University of Georgia in Partial Fulfillment of the Requirements for the Degree.

Master of Arts

Athens, Georgia

2023

The Social Identity of Sounds

by

Savannah Jane Williams

Major Professor: Margaret E. L. Renwick

Committee: John Hale

Jon Forrest

Electronic Version Approved:

Ron Walcott

Dean of the Graduate School

The University of Georgia

December 2023

Acknowledgments

This thesis would not have been possible without the consistent encouragement and guidance from the lovely people in the UGA Linguistics Department. Thanks specifically to both iterations of Ling Lab staff for their community and collaboration, and to Amy Smoler, who has improved many of my days with her kindness and excellent advice. I'd also like to thank the members of my committee, Dr. John Hale and Dr. Jon Forrest, whose exceptional instruction and guidance contributed greatly to the success of this project.

I owe a large debt of gratitude to my advisor Peggy Renwick, who has been a constant source of wisdom and encouragement throughout both my undergraduate and graduate degrees. Her careful and patient guidance has made me a better researcher and a better person, and the care and intentionality that she demonstrates in her advisement has opened up linguistic research to countless students, myself included. She has been and remains a pivotal role model for me as a linguist.

I am immensely grateful to my parents, Camilla and Michael Williams, and sister, Chloe Williams, for their continued support of my academic career, and for listening to me talk about sounds far more than is probably normal. Thanks to my beloved roommate Abby who has cheerfully listened to me practice many presentations on things widely divergent from her field. And thanks to our third roommate, Romeo the cat, who has provided many welcome distractions by sitting on my keyboard.

Contents

List of Figures

List of Tables

Listings

Chapter 1

Introduction & Background

1.1 Introduction

The idea that individual sounds can hold iconic meaning is a longstanding topic of debate in linguistics. Studies focusing on the iconicity of certain subsets of language have found that individual sounds can be associated with size, shape, emotion, and specific physical characteristics. This line of inquiry is called sound symbolism, and is observed at varying degrees of frequency throughout the world's languages. Sound symbolism is often most salient in character or product names, as they involve both the intentional act of naming and, crucially, prior knowledge of the characteristics of the thing being named. Studies on Pokémon and Disney characters have found associations between the sounds in a character's name and its physical characteristics; similar trends are observed in corpus-based studies of English, Cantonese and Mandarin names, where male and female names show distinct phonological distributions (Cutler et al 1988; Kawahara 2019; Klink 2014; Pitcher et al 2013; Shih 2012; Shih 2019; Starr 2018; Uno 2020).

This thesis extends previous corpus-based studies on iconicity in names by implementing a perceptual study assessing what factors impact a participant's judgment of a name as male or female. Using an online survey, participants were presented with a name and asked to rate it on a 5 point Likert scale ranging from "strongly male" to "strongly female". Findings show that names ending in a vowel are much more likely to be judged as female than names ending in a consonant, and that this pattern extends to novel data. The only significant predictor of name sex was whether its final phoneme was a vowel or consonant; this contrasts to the findings of earlier corpus-based studies which found that differences in name length, stress pattern and final phoneme all impacted a name's likelihood of being male or female. These findings show that trends in naming have changed with the US in the last few decades. Ultimately, the results of this study contribute to the arguments that subconscious iconic associations between sound and meaning have the potential to impact language perception, and that sound-symbolic data is a useful tool for both phonological and socio-linguistic research.

1.2 Sound Symbolism: History & Background

The question of a potential association between linguistic features and non-linguistic properties has been the subject of numerous studies investigating if and how we pair individual sounds with descriptive or social meaning. Sound symbolism, the name given to this direction of linguistic inquiry, posits that in a subset of language the association between sound and meaning is not entirely arbitrary. This hypothesis is supported by findings showing that it is possible for certain speech sounds to represent non-phonetic concepts such as size, shape, or emotion (Kawahara 2019; Ohala 1996; Sapir 1929; Shih 2019; Uno 2020). The goal of the study of sound symbolism is not to prove that there is a symbolic association attached to every sound, but rather to propose and investigate ways in which the perception of language is potentially impacted by non-phonetic factors.

The discussion of sound symbolism in language can be traced back to Plato's dialogue Cratylus, in which his characters discuss the nature of language and the possibility that there might be some relationship between the sign (speech sounds) and the signified (the objects they represent). Plato's dialogue concludes with an argument against the existence of sound symbolism, where he claims that because different languages use the same sounds to convey different meanings, language must be an arbitrary system. This argument was later adopted and built upon by Ferdinand De Saussure, who argued that different languages using different sounds to express the same meaning proves conclusively that language is entirely arbitrary (1916).

In their arguments, both Socrates and De Saussure focus on language as a whole, concluding that because the system as a whole is arbitrary so must be its constituent parts. Contrary to this belief, research from multiple fields including linguistics, cognitive science, anthropology, and marketing shows that for a small subset of a language, the relationship between sound and meaning may not be entirely arbitrary. It may seem like a radical claim in the face of a decades- (if not centuries-) long precedent of arbitrariness; however, when looking at the potential evolutionary roots of sound symbolism, its intuitiveness becomes evident.

1.3 Evolutionary Roles

One of the most basic demonstrations of sound symbolism is seen in intonation patterns. In many languages, a high fundamental frequency (F0) indicates a question, while a low F0 indicates a statement. This relationship has been named the frequency code hypothesis

(Ohala 1996), which posits that a high F0, most often found in the smaller voice boxes of smaller creatures, signifies the deference needed when making a request, while a low F0, found in the larger voice boxes of larger creatures, conveys the authority and confidence needed when making a statement.

Not only is this phenomenon cross-linguistic, it is cross-species as well. In many animal species, aggressive vocalizations take the form of a low growl, while surprise or pain is communicated by a high-pitched yelp. It is hypothesized that this association has an evolutionary root, as the ability to use vocalizations to appear either smaller or larger depending on the situation gives animals a significant survival advantage (Hinton, Nichols & Ohala 1994). In both human and animal cases, the pitch of the vocalization gives the listener an idea of what the vocalizer might mean, a concept that is evidenced in a slightly more nuanced way by ideophones.

Ideophones are sound-symbolic words that convey sensory perception. In English we see evidence of this in words such as *glimmer* and *twinkle*. Ideophones differ from onomatopoeia in that ideophones are not limited to only representing sound; instead, they can demonstrate a wide variety of sensory experiences such as taste, smell, and texture. While relatively rare in Western languages, ideophones are quite common in languages such as Japanese and Korean. Haiman (2012) hypothesizes that ideophones could be the missing link in understanding the evolution of language. At some point, language evolved from gestural coordination (showing) to linguistic vocalization (telling). Ideophones provide a sound-symbolic combination of showing and telling, offering one possible explanation for how humans bridged the gap between gestures and speech. The use of ideophones then became less frequent as vocalizations became more sophisticated (Haiman 2012).

In a similar vein, De Carolis and colleagues (2017; 2019) claim that the origins of language were iconic in nature. It is possible that the first human vocalizations were motivated by the desire to analogize a vocal sound to a particular sensation or emotion. This analogy is cited as a potential explanation for how humans were able to agree on shared speech sounds for referents that may not have a physical counterpart. This shared understanding marks the point where humans gained the cognitive ability to make the connection between a speech sound and someone else's emotional state.

1.4 Types of Sound Symbolism

Sound symbolism in present day language takes several different forms. Corporeal sound symbolism uses sounds and intonation to express internal states, as seen in the aforementioned concept of ideophones. Synesthetic sound symbolism, seen in sound/shape and sound/size correspondence, refers to the association of certain sounds with non-acoustic properties. For examples of onomatopoeia, which is perhaps the most culturally familiar example of sound symbolism, we must only turn to the language common in comic books, where a vocalization is used to represent a sound in the environment, such as *bang*, *zoom*, or *meow*.

The three categories mentioned so far have been shown to occur cross-linguistically, in contrast with the final category, conventional sound symbolism, which is language-specific. Conventional sound symbolism is the association of phonemes or groups of phonemes with a particular meaning, evidenced in English by the prevalence of the cluster "gl" in words meaning "light" (*glisten*, *glow*, *glimmer*, etc.) (Hinton, Nichols & Ohala 1994). These pairings are known as phonaesthemes, and act as units of meaning in between phonemes and contrastive morphemes. Phonaesthemes do not qualify as morphemes due to the fact that when removed from words they do not leave behind any sort of meaningful unit, as seen with the removal of "gl" from "glimmer" leaving behind "immer" (Bergen 2004).

1.4.1 Synesthetic Sound Symbolism

Synesthetic sound symbolism is most widely studied, perhaps due to the novelty of the idea that acoustic sounds can represent non-acoustic properties such as size or shape. Evidence from novel word recognition studies provides insight into the role that synesthetic sound symbolism might play in present day communication. Sapir (1929) concluded that some vowels "sound bigger" than others, namely that the vowel [ɑ] is judged as referring to larger objects, and the vowel [i] as referring to smaller objects. This is evidenced by the use of [i] in English to emphasize smallness, as seen in words like "teeny-tiny" and "mini".

This phenomenon is not unique to English, in fact, Lapolla (1994) demonstrated that there is a cross-linguistic association of acute segments with words denoting smallness, and grave segments with words denoting largeness. In a study comparing sensitivity to sound symbolism between native English speakers and native Chinese speakers, Lapolla (1994) found that both speaker groups reacted similarly to this type of sound symbolism, supporting the frequency code hypothesis and demonstrating in a more nuanced way that sound can be associated with the concept of size.

Sound-shape correspondence has been shown as well, in a phenomenon known as the "bouba-kiki" effect. Participants are more likely to pair pseudowords containing the velar plosive [k] with spiky shapes, and the bilabial plosive [b] with round shapes (Kohler 1929). Many different versions of this study have been conducted cross-linguistically, using a variety of pseudowords and shapes. The results are largely consistent throughout the studies: velar sounds are judged to represent spiky shapes and bilabial sounds to represent rounded shapes.

D'Onofrio (2013) expanded on previous research both by looking at phonetic features in isolation and by linking pseudowords to real-world objects instead of abstract shapes. Isolating potential variables such as vowel quality and consonant place of articulation allowed for the investigation of how the combination of multiple potentially sound-symbolic factors affects listener perception. These findings demonstrated that the "bouba-kiki" effect is potentially produced by a combination of multiple phonetic features, and that sound/shape association can extend to real-world objects in addition to abstract shapes.

1.5 Functional Use

Besides enriching the experiences of comic book readers, sound symbolism plays several functional roles within language. A study involving native English speakers' perception of Japanese haiku demonstrated that a high ratio of plosives to nasals (i.e. a less overall sonorant speech signal) is perceived as active and excited, while a low plosive to nasal ratio is perceived as calm and quiet. This association demonstrates that sound symbolism may play a role in conveying semantic meaning, particularly emotions, in poetry (Miller 2014). Additional research supports the idea that a shifting of mood or setting in literature could be indicated by a change in the contrasts between phonetic features which unconsciously signal the shift to the reader (Miall 2001).

In English, sound symbolism may play a role in the classification of nouns and verbs. Examination of a corpus of words with a high lexical frequency led to the finding that nouns are more likely to have back than front vowels, and verbs are more likely to have front than back vowels. This association of vowel quality with syntactic class impacts the processing speed of nouns and verbs: listeners process words more quickly if they have the expected vowel quality according to syntactic class, vs. words with an unexpected vowel quality (Sereno 1994).

1.6 Sound Symbolism & Proper Names

Within the broader field of synesthetic sound symbolism, a significant portion of the active research is focused on sound symbolism in proper names. Proper names serve as a repository for sound-symbolic characteristics due to the level of intentional design involved in their creation. Current research focuses on the forms sound symbolism in proper names can take, whether this use is conscious or unconscious, and what kinds of things can be symbolized.

1.6.1 Product & Character Names

When naming inanimate objects such as products, research from the field of marketing has shown that incorporating sound symbolism into product names leaves customers with a more favorable impression of the brand. Assuming the existence of sound symbolism, Klink (2001; 2003) showed that consistency between brand design and brand name better communicates the goal of a product than non sound-symbolic brand names. Most marketing-based research is language-specific, and assumes the existence of sound symbolism in order to investigate the practical ways it can be used to achieve the highest levels of consumer engagement.

Sound symbolism is much easier to study language-specifically rather than universally due to the difficulty involved in developing a cross-linguistic dataset. Researchers have found a way around this problem by using characters from globally recognized franchises such as Disney or Pokémon. Pokémon, though originating in Japan, is now a global fixture, with Pokémon character names being translated into dozens of different languages. Using Pokémon names to study sound symbolism comes with a multitude of benefits. Each name has a fixed, cross-cultural referent, and every Pokémon has its own unique list of characteristics such as height, weight, and strength. The only variation between Pokémon in different languages is linguistic; everything else remains constant.

The first research on Pokémon studied sound symbolism in the original Japanese. Kawahara and colleagues (2018) found a positive correlation between the number of voiced obstruents in a name and the corresponding Pokémon's size, weight, and evolution level; and the number of mora correlated with size, weight, evolution level, and strength. Vowel height was also found to play a small role in conveying the size and weight of Pokémon.

Further research done focusing specifically on the function of sibilants in Pokémon names found that, when given a choice between a name with sibilants and a name without, native Japanese speakers are more likely to associate flying-type Pokémon with names that contain sibilants. The authors hypothesize that this sound-symbolic relationship has a kinesthetic root. The large amount of airflow that accompanies the production of a sibilant could potentially cause listeners to associate sibilants with air and flying (Kawahara, Godoy & Kumagai 2020). Sapir (1929) hypothesized a similar explanation for his aforementioned discovery of sound/size correspondence, suggesting that either the volume or acoustic space of a vowel determines what non-acoustic characteristics are associated with it.

Several studies have shown that Pokémon have sound symbolic properties beyond the original Japanese. Shih and colleagues (2018) expanded Kawahara's original study both by increasing the number of phonological characteristics examined and by including English in the study in addition to Japanese. Results showed continuity across languages as well as language-specific features. For both English and Japanese, name length was found to have a positive correlation with size, power, and evolutionary stage. For English, this correspondence was only observed across evolutionary stages, while for Japanese the correlation was still observed within evolutionary stages. This variation in results is most likely attributed to language-specific structural differences, such as the difference in structural units used in English and Japanese (syllables vs. mora, respectively).

The association of name length with character size was further evidenced by a later study encompassing English, Japanese, Mandarin, Cantonese, Russian, and Korean. Using Pokémon names as the dataset, the authors examined each language individually as well as any commonalities they might share between them. Examination of Pokémon name formation across multiple languages suggests that the most salient sound symbolic patterns are those that are associated with important goals within the Pokémon universe. This is evidenced by the strong cross-linguistic sound-symbolic indicators of evolutionary stage and overall Pokémon power rating, two key aspects of gameplay (Shih et al. 2019).

In addition to concrete concepts such as size and shape, sound symbolism may be able to portray abstract concepts such as “villain”, as investigated by Uno and colleagues (2020) using two datasets, one of the English names of Disney villains and one of the Japanese names of Pokémon villains. Findings showed that villains in both Disney and Pokémon are more likely to have names that contain voiced obstruents than non-villains. The names of non-villainous Disney characters are also more likely to contain bilabial consonants. This trend can be seen in Pokémon names as well, but only at the extremes of the dataset when comparing the most evil set of characters to the least evil.

The association of voiced obstruents with villains and bilabials with non-villains can potentially be explained kinesthetically. Voiced obstruents require increased constriction in the airway, complicating their production, and it is posited that this increase in production difficulty (in relation to bilabial consonants) could lead to their association with evil characters. On the other hand, bilabial consonants are often among the sounds first produced by babies, leading to their association with innocent or good characters. The study of character names shows that while language as a whole may be arbitrary, there exists sound symbolic processes whose existence can be predicted but not universally determined (Uno 2020).

1.6.2 Sex-Specific Phonology

A similar dataset to Pokémon involving real-world subjects was studied by Shih and colleagues (2019) in their investigation of sound symbolism in baseball player names. Sound symbolic correlation was found between the physical characteristics of the players and their chosen names and nicknames. These results are similar to those of the Pokémon dataset: nicknames for smaller players are more likely to contain high vowels, and longer nicknames with sonorant consonants are more likely to belong to heavier players. No correspondence was found between physical characteristics and the given names of the baseball players (Shih & Rudin 2019), supporting the claim that sound symbolism emerges most strongly where there is prior knowledge of the characteristics of the referent.

Although given names may not hold sound symbolic characteristics with regard to variables like height and weight, studies on English, Mandarin, and Cantonese personal names have revealed that personal names may include sex-specific phonology. English has been demonstrated to contain patterns in the phonology of male and female names (Cassidy, Kelly & Sharoni 1999; Cutler, McQueen & Robinson 1990; Whissell 2001; Williams & Renwick 2022). Female names are, on average, longer, more likely to have an unstressed initial syllable, more likely to end in a vowel or sonorant consonant, and more likely to contain the vowel [i] than male names. Male names tend to be shorter, end in a stop or fricative, and generally take the unmarked case, meaning that female names can be derived from male names, but male names are not usually derived from female names. Gender-neutral names are more likely to be used as names for female babies and follow the phonology of female names more closely than they do male.

Female names are more likely to contain phonemes that are associated with characteristics such as pleasantness, passivity and softness, while male names are more likely to contain phonemes associated with activity, unpleasantness, and cheerfulness. These findings suggest that English personal names (like character names) are capable of conveying abstract emotional concepts through their phonology, and that the emotional content of female names is different from that of male names (Whissell 2001). Research using a connectionist network showed that this pattern is learned in English speakers and that it can be used to infer sex from unfamiliar names. This learning process is likely largely unconscious and facilitated by repeated exposure (Cassidy, Kelly & Sharoni 1999).

As noted by Shih and colleagues (2018) while studying Pokémon names, one of the only characteristics of Pokémon not significantly represented in name phonology is biological sex, most likely due to the fact that breeding was not introduced to the game until later versions, meaning that the ability to differentiate sex based on name was not crucial to original gameplay. This is in contrast to the real world where many languages employ names that are specifically male or specifically female. This name dichotomy could be explained by the evolutionary necessity of differentiating between male and female, as several Pokémon researchers have pointed out that the most salient sound symbolic characteristics are those that are most necessary for survival (Shih et al. 2019).

Using a dataset of Mandarin and Cantonese personal names, Starr (2018) demonstrated that female names are more likely to contain high and front vowels than male names, and that in general sounds with a high acoustic frequency are more likely to be in female names, while male names are more likely to contain sounds with a low acoustic frequency. This is congruent with the previously mentioned frequency code hypothesis; the association of maleness with lower frequencies and femaleness with higher frequencies can potentially be explained by the fact that males generally have deeper voices than females (Hinton, Nichols & Ohala 1994).

Building upon previously found phonological differences in male and female names in English, Williams & Renwick (2022) conducted a corpus-based study of sex-specific phonology using a dataset of 5600 names sourced from the United States Social Security Administration (https://www.ssa.gov/oact/babynames/decades/). This SSA data is provided publicly, and contains the most popular names for each decade from 1880-2010. With the goal of determining if sex-specific phonological differences (a) are present in the dataset, and (b) experience change over time, the top 400 names (200 male, 200 female) for every available decade (n = 13) were sourced from Social Security records. Statistical analysis of this dataset found that while the differences between male and female names are consistent over time, the extent of the contrast between phonological inventories experiences variation over time. These findings suggest that naming conventions may be experiencing variation due to changing cultural trends and increased rates of immigration. This leads towards the conclusion that, either implicitly or explicitly, our iconic associations are dependent on shifting cultural norms. This finding of name change over time is foundational to the design of the present study, where collecting perceptual data will provide a window into the characteristics perceived as defining name sex at the present moment. The results of studies on the sound symbolism of proper nouns, and particularly of names, points to the fact that some lexical classes might be more susceptible to the effects of sound symbolism than others. This hypothesis supports an argument in favor of the usefulness of employing known sound-symbolic data in other areas of linguistic research, as sound-symbolic data has the potential to provide a unique perspective on the linguistic and sociological factors that impact language perception.

Chapter 2

The Current Study

Considering that previous corpus-based studies have found distinct differences in the phonological distribution of male and female names, a logical extension of this line of inquiry is a perceptual study investigating if and how these characteristics are perceived in real-world language use. To that end, the following study makes use of an online survey to test participants' gender judgments of names via a 5 point Likert scale ranging from "strongly male" to "strongly female". A diverse selection of names is presented to participants, including traditional male and female names, androgynous names, Black-coded names, and artificially generated nonce names.

Using this selection of names, I test whether previously identified distinctive characteristics of male and female names are actually used by people to assign a gender judgment to names, and additionally whether this judgment is impacted by demographic characteristics. This chapter is structured as follows: §2.2 continues this discussion with an overview and justification of survey-based studies of linguistic intuition; §2.3 and §2.4 discuss the research questions and proposed hypotheses.

2.1 Survey-Based Studies of Linguistic Intuition

Data collection for this study was conducted using an online survey where participants' intuitions regarding the assignment of names as male or female were examined via a 5 point Likert-style judgment scale for name gender. Though Likert scale data are fairly uncommon in phonological studies, they are frequently used for collecting various forms of psychological judgments in experimental psycholinguistics, language acquisition (specifically for quantifying bilingualism, see Vérrisimmo 2021), assessment of machine-generated language in natural language processing, and grammaticality judgments in experimental syntax.

Though less commonly seen in the collection of phonetic or phonological data, survey-based intuition studies have been shown to provide accurate data on linguistic intuition below the sentence level, and have been used for studies of phonetic intuition in Catalan (Renwick & Nadeu 2018) and in English (Shitara 1993). Judgment scales are particularly common in research on sound symbolism, as perception data is a fundamental aspect of studies investigating potential sound/meaning association. Nearly all of the studies on the "bouba-kiki" effect (also known as the "maluma-takate effect"), which is perhaps the most well known formal study of sound symbolism, use a type of scaled judgment question to ascertain the extent of participants' associations (Barton 2016; Cuskly et al. 2015; De Carolis 2019; Knoferle 2017, among others).

The present structure of survey-based data collection parallels a common method of gauging syntactic intuitions on grammaticality, where participants are asked to rate their grammaticality judgments and confidence on a Likert scale ranging from "least grammatical" to "most grammatical". Use of this methodology in experimental syntax allows for the examination of linguistic features not commonly found in corpus data (Gross 2021), the assessment of grammaticality judgments over a certain population, and also for the identification of the psychological categories for grammaticality both unique to each participant and aggregate across respondents (Erlewine & Kotek 2016). This style of questionnaire allows for the possibility of gradient space between “grammatical” and “ungrammatical”, and results in judgments that are potentially more accurate to the reality of grammatical perception (Sprouse 2015). Such approaches attempt to distinguish between syntactic theory and syntactic reality by dividing theoretically binary categories into gradient scales of psychological judgment.

I propose that similar benefits are seen when using survey data to test phonological intuition, and that the allowance of gradient categories for phonological judgments will result in data that is more accurate to the reality of language perception and use. The choice in methodology for the present study is motivated by these previous results demonstrating not only that survey-based studies can provide accurate data regarding linguistic intuition, but also that this data can provide valuable insight into the psychological and sociological constructs that underlie language perception.

2.2 Research Questions

The formation and direction of specific research questions for this study is significantly impacted by the results of previous studies, namely Cutler et al (1986) and Williams & Renwick (2022), both of which identify specific phonological characteristics associated with English first names. These findings allow for the creation of specific research questions that take into consideration which phonological variables have been found to be significant predictors of name sex in previous studies. Thus, in addition to formalizing the research questions, this section will discuss and justify the inclusion of the selected phonological variables in this study.

Both Cutler et al (1986) and Williams & Renwick (2022) found final phoneme, name length, and initial syllable stress to be significant predictors of name sex. The present study adopts these same three variables as the phonological variables to control for and model in stimuli selection. Choosing to include these three variables in this study directly ties it to the results of previous corpus-based studies, with the goal of testing if and how these three previously-significant variables are perceived and used by English speakers. In short, the primary goal of this study is to determine whether the trends observed in corpus-based studies of English first names are mirrored in perceptual data collected from English speakers.

To this end, in the following study I use a survey-based data collection approach to investigate the functional use of sound symbolism as it pertains to the perception of the phonological differences found in first male and female names in US English. Using data collected from 414 native English speakers, the following paper investigates if and how the previously found differences in male and female name phonology are used synchronically by speakers, and if this use is impacted by speaker demographics. This line of inquiry can be formalized into the following research two research questions:

(1a) Are the findings of corpus-based studies on gendered phonology (Cutler et al 1989; Williams & Renwick 2022) seen in perceptual judgments?

- **(1b)** If (a) is true, do these trends hold when the participant is presented with novel data?

(2a) Do demographic characteristics impact participants' judgments of name sex?

- **(2b)** If (a) is true, what specific factors impact judgment?

2.3 Hypotheses

Given the above research questions, I propose the following hypotheses:

1. **Hypothesis 1**- Relationship to Corpus Data:
   (a) Participants will use the phonological cues from previous corpus-based name studies to differentiate between male and female names.
   (b) These trends will be extrapolated to novel data.
2. **Hypothesis 2**- Demographic Characteristics
   (a) Groups sharing similar demographic characteristics will behave in similar ways.

Chapter 3

Data & Methods

There are two primary sources of data in this study, the stimuli and the participant judgements of the stimuli. This chapter discusses the construction of stimuli, survey design, the recruitment and resulting demographics of survey participants, and the handling of the data collected from the survey. Chapters 4 and 5 analyze the collected survey data descriptively and quantitatively, respectively. The following section outlines the steps taken to prepare stimuli for the resulting perceptual study. The data collection and preparation processes for this study are divided into sections discussing stimuli construction (§3.1), survey construction (§3.2), participants (§3.3) and data processing (§3.4).

3.1 Stimuli Construction

Stimuli for assessment of gender judgments consisted of 96 names in 6 categories: modern androgynous names, such as *Fox* or *River,* traditionally male or female names such as *John* and *Mary*, traditionally androgynous names like *Casey* and *Taylor*, names traditionally given to Black babies in the United States such as *Journey* and *Latoya*, and computer-generated "pseudo-names" which adhered to English phonotactics (this is discussed further in §3.1.1). A range of categories was chosen in order to provide a diverse sampling of names across ethnicities and time period of popularity. Traditionally male and traditionally female names were included in the study for two reasons: (1) to act as a type of control in the survey where multiple judgments diverging from the expected response would indicate that closer examination of the validity of a respondents' choices was necessary, and (2) to provide a metric by which the characteristics of the other four name categories could be compared with. 14 total traditional names were included in the final stimuli list (7 male, 7 female), and these names were chosen from the publicly available Social Security records on name popularity (https://www.ssa.gov/oact/babynames/decades/). Names were chosen from the list of the top ten most popular names for the following range of decades: 1940, 1960, 1990, 2010. The goal of choosing from a range of decades was to provide a diverse selection of name popularity over time. The full list of all name stimuli included in the study can be found in Appendix A.

Traditional and modern androgynous names were included to ascertain what phonological characteristics could potentially influence a participant to assign an androgynous name to a male or female category. This category is divided into traditional and modern in order to account for frequency, where theoretically the participant has been exposed to traditional androgynous names more than modern androgynous names and might have differing preconceptions of the two categories. Names popular among Black babies in the United States were included for two reasons: (1) naming practices in Black communities often have a cultural and social significance that is a particularly clear window into social conventions and beliefs at a certain time period (Logan 2021); (2) Black names often display distinctive linguistic innovation while preserving gender distinction, leading to an interesting dataset in which to study the phonological indicators of name gender (Lieberson 1995).

3.1.1 Pseudo-Name Generation

In order to test whether potential sound-symbolic intuitions extended to novel names, a pseudo-name generator was created to generate nonce names that conformed to English's phonotactic constraints. The functionality of this program is most easily conceptualized as a series of cascading "buckets", where constraints (acting as the phonotactic constraints on English) dictate what kinds of characteristics can be passed from one bucket to another.

The generator was coded using Python (van Roussum 1995). The top 200 (100 male, 100 female) most popular names of the last two decades according to records sourced from the United States Social Security Administration (ssa.gov/OACT/babynames/decades/)were searched in the Carnegie Mellon Pronouncing Dictionary (CMU Speech Group 1995) and their pronunciations were imported into a Python file. The sounds from these names were separated into one list of consonants and three lists of vowels, one for each possible stress classification (primary, secondary and unstressed). The three lists of vowels were identical save for the differences in stress pattern.

The lists of phonemes were combined into a master list of all phonemes, but retained their original classification (consonant, vowels with varying stress assignments) in sub-lists. This master list serves as the phoneme bank. A syllable generator function iterates through the sounds in the phoneme bank twice, first to find all consonant and vowel combinations (CV), and second to find all possible CVC combinations. These CV and CVC syllables are then assigned to a syllable bank list, retaining their classification as either open (CV) or closed (CVC) syllables.

The most pertinent English phonotactic constraint for the purposes of name generation is stress pattern. In order to ensure that the generated names only had stress patterns possible in English, the stress patterns from the original 200 sample names from the CMU dictionary were identified and incorporated as constraints to name construction. Linguistic evaluation of the stress-based rhythm patterns in English traditionally distinguishes between ”strong” and ”weak” syllables. This binary is usually defined by a combination of vowel quality and stress. Strong syllables contain stressed, unreduced vowels, and weak syllables contain unstressed, reduced vowels. Opinions on whether syllable strength should be attributed to vowel quality or to stress vary (see Taft 1984 for an argument in favor of a stress-based definition; Cutler & Norris 1982 and Cutler & Butterfield 1992 for a vowel-based definition; Cutler et al 1995 for a combined approach).

Rather than defining an objective measure of a ”strong” and ”weak” syllable, Lieberman and Prince introduced the idea of ”relative prominence” in 1977, and this idea was further refined by Hayes (1980). Instead of assigning an absolute prominence to the vowel, that is, being able to categorize a ”strong” or ”weak” syllable in isolation, Lieberman and Prince proposed that stress should be represented using the relative prominence between the syllables of an individual word.

The analysis of name stress in this study follows this analytic system of relative prominence, where three possible stress levels are identified by examining the relative prominence of the syllables that make up an individual name. Primary stress, which can occur once per word, is the most prominent syllable in a word. Secondary stress, which can occur more than once, is the second most prominent. An unstressed syllable is the least prominent, and often contains a reduced vowel. Eight total stress patterns were identified from the sampled names, one for one syllable names, two for two syllable names, three for three syllable names, and two for four syllable names (3.1).

Table 3.1: List of possible stress patterns

Stress Pattern	Syllable Count	Example Name
P	1	John
P + U	2	Mary
U + P	2	Elise
P + U + U	3	Margaret
U + P + U	3	Marissa
U + U + P	3	Mirelle
U + P + U + U	4	Olivia
U + U + P + U	4	Alexander

The word builder function selected syllables at random and passed them through the stress constraints listed above, so that only a syllable marked "primary" could go into a slot marked "P", and so on. This resulted in 40 total generated names, 10 for each possible name length (1 - 4 syllables). The names were then passed through a final function to remove any real names that may have been generated. The generated names were converted from ARPaBet to IPA and orthography by hand. 15 of the 40 total generated names were included in the survey; this choice was based on criteria of ease of pronunciation and potential similarity to a real name. Any generated names that sounded very similar to a real name were excluded from the final list, as well as any names that seemed prohibitively difficult for an English speaker to pronounce. A full list of the generated names is included in Appendix B for reference.

3.1.2 Annotation

The stimulus names were coded for three variables: number of syllables, stress pattern and final phoneme (3.2). These variables were selected based on the results of both Cutler et al. (1986) and Williams & Renwick (2022), where these three variables were the most significant predictors of name sex. Stress was evaluated based on the stress of the first syllable of the name, and names were placed in one of the three following stress categories: primary stress, meaning that the initial syllable was stressed; secondary stress, meaning that the initial syllable had secondary stress; or unstressed, meaning that the initial syllable had no stress. In names with multiple pronunciations resulting in more than one option for the stress pattern, the most common pronunciation was selected. If the most common pronunciation was not evident, the stress patterns available on https://www.babynamewizard.com were used as a reference.

Table 3.2: Sample of annotated dataset with example name from each category: traditionally male (TM), traditionally female (TF), traditionally androgynous (TA), modern androgynous (MA), traditionally Black (TB), and generated (G).

Name	Name Category	Syllable Count	Final Phoneme	Initial Syllable Stress
John	TM	1	C	P
Mary	TF	2	V	P
Casey	MA	2	V	P
Dakota	MA	3	V	S
Jamar	TB	2	C	S
Vashaethee	G	3	V	U

3.2 Survey Construction

The survey was introduced to participants with a brief narration about a character named Nigel, a comic book artist who needed help picking names for some of his characters. This framing served two purposes: (1) to obscure the goal of the survey and (2) to increase participant interest and attention. After being introduced to Nigel, participants were presented with 50 names, with each appearing on its own page. Participants were asked if they would assign the given name to a male character, a female character or to either a male or female character. A five-point scale offered five choices: “this name is strongly male”, “this name is somewhat male”, “this name could be either male or female”, “this name is somewhat female”, and “this name is strongly female”. Choice order was randomized across participants, with approximately half of the participants viewing “strongly male” as the first option, and the other half viewing “strongly female” as the first option. A graphic was provided with each question for reference (3.1).

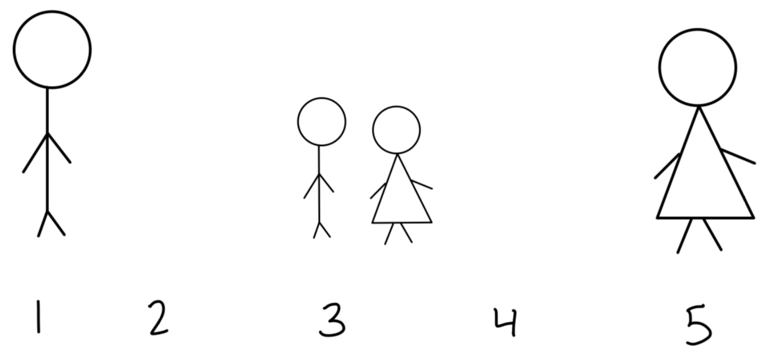

Figure 3.1: Illustrated rating scale for gender judgements

After making their judgment about the name, participants were asked to rate their familiarity with the name on a five point familiarity scale with the following options: “I have never heard of this name”, “I have heard of this name”, “I know someone with this name”, “I know several people with this name”, “this is my name”. The familiarity question was included for two purposes: (1) to see if participants responded to unfamiliar names in the same way as familiar names, and (2) to serve as a metric to judge the validity of a participant’s response (i.e., if a participant indicated unfamiliarity with a large number of common names, this would indicate a closer examination of their response patterns). A list of each name’s average familiarity score is included in Appendix C for reference.

Each name and its accompanying two questions received its own page in the survey. Participants were not allowed to view previous questions once they moved on from a name and its accompanying two questions. Each participant was shown a random selection of 50 of the total 96 names, with the number of appearances of each individual name standardized across survey responses. Generated names were accompanied by phonetic pronunciation guides, with spaces between syllables and stressed syllables represented by capital letters (ex: *Vashaethee* > ”vuh SHAY thee”).

3.2.1 Demographic Questionnaire

After being presented with 50 name questions, each participant was asked to complete an optional demographic questionnaire. 339 participants completed the entire questionnaire, 40 partially completed the questionnaire, and 62 participants chose not to complete the questionnaire. 14 of the 60 participants who chose not to complete the demographic questionnaire also had incomplete responses to the name judgment section and were excluded from analysis (see §3.3 for further discussion). Participants were presented with 14 total demographic questions, asking them for their age, gender, level of education, approximate yearly income, birth country (if outside the USA), birth state and region (if inside the USA), native language, and political affiliation.

3.3 Participants

After receiving approval from UGA’s Institutional Review Board, participants were recruited through advertisements on social media and word of mouth. A QR code and link to the web-based survey, administered via Qualtrics, was included in the distributed advertisements. Any person 18 years of age or older was allowed to complete the survey, though it specified in the consent form that participants should be native speakers of American English. 414 responses were obtained, collected over a period of approximately one month. Of these 414 responses, 27 were excluded due to incomplete response, 11 due to non-native English speaker status, and 2 for lack of consent, resulting in n = 374 analyzable responses. The survey took approximately 20 minutes to complete.

71% of respondents were female, 13% were male, 3% were non-binary, and 12% elected not to share their gender in the survey (3.2).

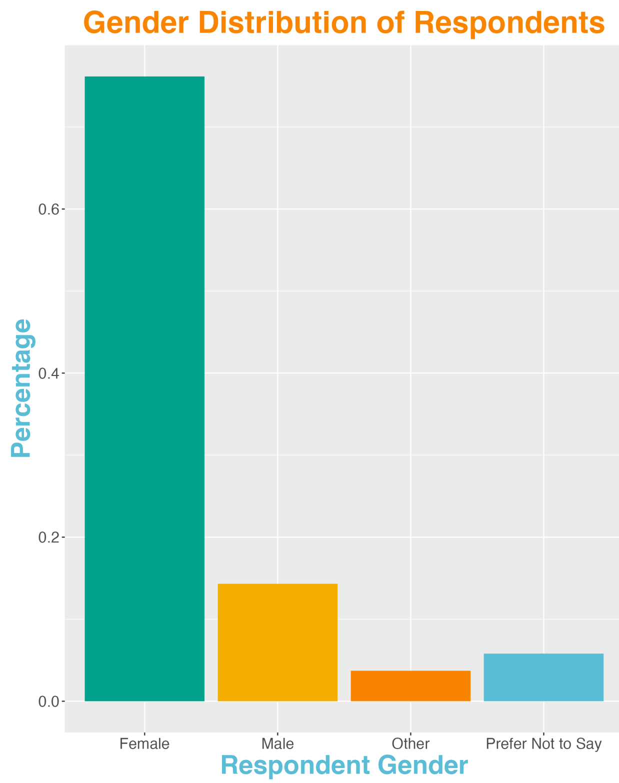

Figure 3.2: Age distribution of survey participants

18% of respondents were between the ages of 18 and 24, 10% were between the ages of 25 and 40, 51% were between the ages of 41 and 58, 5% were between the ages of 59 and 70, 2% were over the age of 71, and 12% did not report their age (3.3). The majority of participants (48%) were from the Southeastern US. Political affiliation was fairly evenly dispersed across conservative (n = 82), liberal (n = 68) and moderate (n = 92).

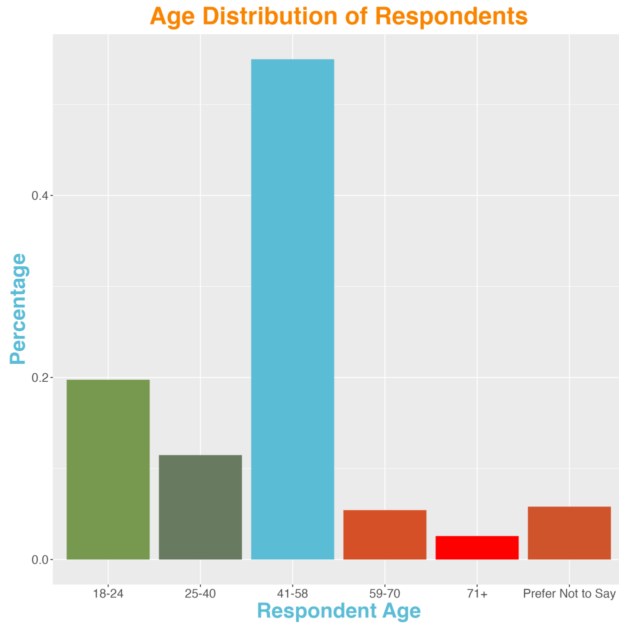

Figure 3.3: Age distribution of survey participants

3.4 Data Processing

Each participant was assigned a unique ID number, allowing for differentiation between individual responses while maintaining participant anonymity. Each token of a name received its own row in the final data set, and was coded for participant, name category, gender judgment, familiarity judgment, choice order, syllable count, final phoneme, initial syllable stress, respondent age, gender, ethnicity, level of education, income, birth country/state/region, and political affiliation. As the demographic questionnaire was not mandatory, there are instances where responses to some questions are missing. In these cases "prefer not to say" was used as a placeholder for the unavailable data [1].

[1]The use of *NA* was avoided due to the difficulty *NA* values pose in quantitative analysis.

Chapter 4

Descriptive Results

Visualization of response category treatment shows that the "either" category is selected more than any other category (Figure 4.1). Names were slightly more likely to be assigned to a male category than they were to a female category, and both "strong" categories were used slightly more often than the "somewhat" categories.

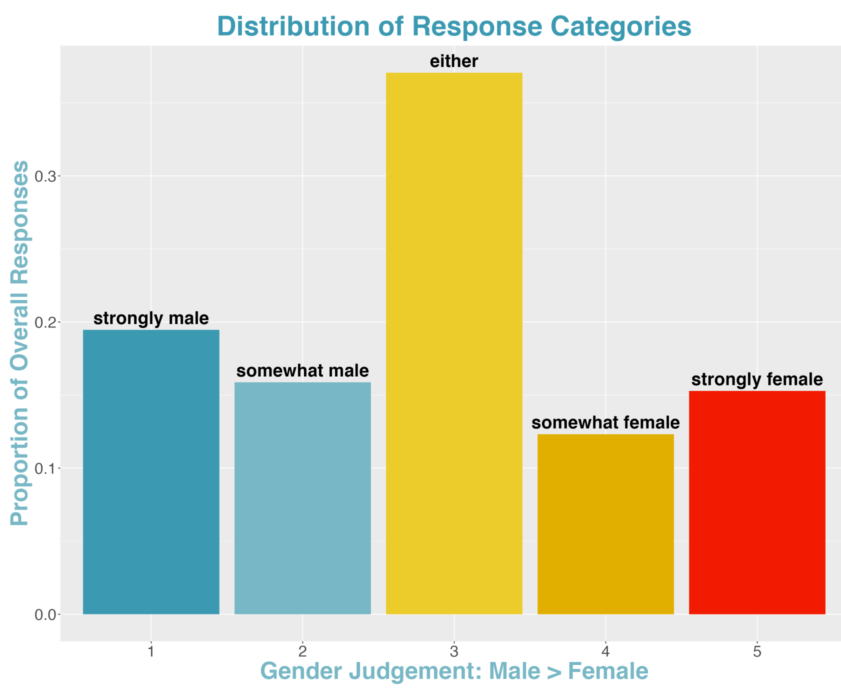

Figure 4.1: Distribution of participant responses across response categories

Examination of response distribution by demographic categories shows no significant trends individual to demographic groups, as is evidenced by Figures 4.2 and 4.3 below. Response distributions remained fairly consistent across categories of gender identity, pointing towards the conclusion that the gender identity of an individual may not impact their perception of name sex.

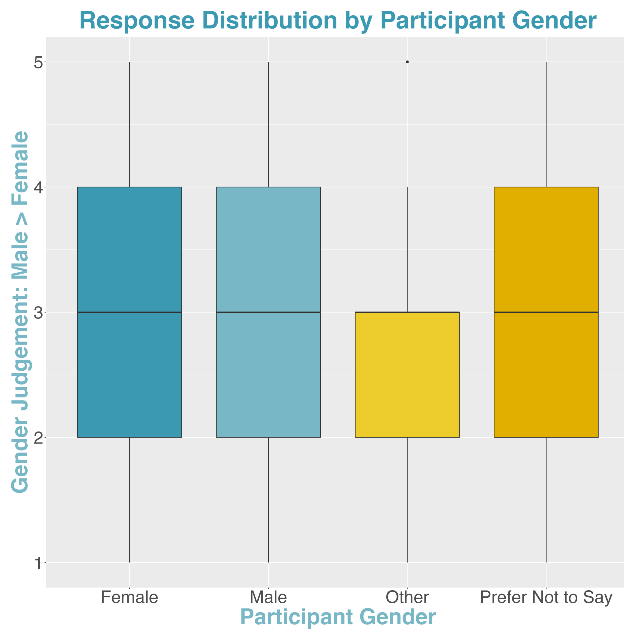

Figure 4.2: Distribution of participant responses by participant gender identity

A similar uniform trends across demographic categories is seen in Figure 4.3, where the age of a participant does not seem to impact their selection of name sex. This variable was included in particular to see if there was any generational variation in response, and, as Figure 4.3 shows, there is not a difference in response patterns between older and younger participants.

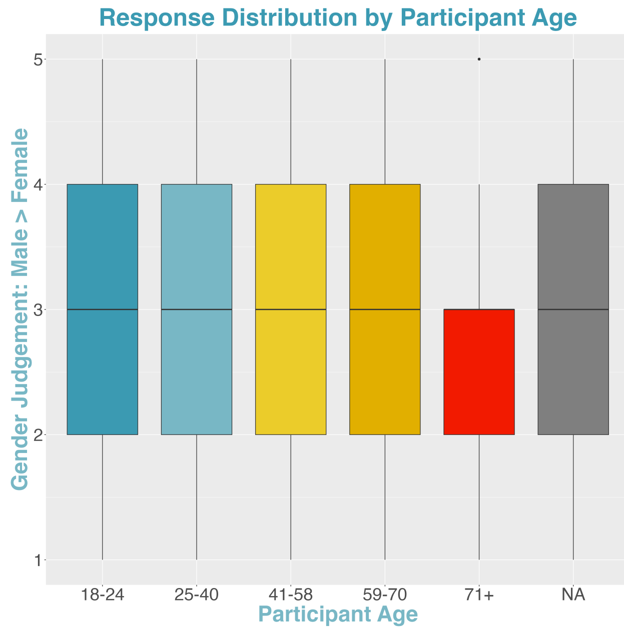

Figure 4.3: Distribution of participant responses by participant age

4.1 Linguistic Predictors

Looking at the distribution of responses based on the number of syllables in a name (Figure 4.4), and taking into account that two syllable names are the most popular name type overall, the only time that a "strong female" judgment occurred more frequently than a "strong male" judgment was when a name had 3 syllables; this is consistent with findings from corpus studies (Cassidy, Kelly & Sharoni 1999; Cutler, McQueen & Robinson 1990; Whissell 2001; Williams & Renwick 2022) where female names had an average higher syllable count than male names.

Interestingly, if a name has four syllables, it is slightly more likely to be assigned to a male, based on the fourth panel of Figure 4.4. This is contrary to previous findings, and suggests that it may no longer just be a longer name that is associated with female, but specifically a name with three syllables.

Figure 4.5 shows the differences in number of responses per category based on whether the name ends in a consonant or a vowel. The left panel of 4.5 shows that names ending in a consonant are rarely assigned to either of the two female categories, while the right panel confirms the inverse of this: names ending in a vowel are less likely to be assigned to one of the two male categories than to the female categories. In the case of both final phoneme options, the most common category is still "either"; this is consistent with the results of 4.4, which shows that "either" was chosen much more frequently that any other category option.

4.2 Descriptive Results by Name Category

One of the primary motivators behind including a range of name types in the survey was to investigate whether (a) these sets of name categories differ from each other phonologically, and (b), to see if, on average, participants respond differently to different categories. The following set of graphs looks at the descriptive results of how participants responded both to individual names and to the six different stimulus categories.

Figure 4.6 shows the percentage of female judgements for each name, judgments for each name, with the left panel displaying names ending in a consonant and the right panel displaying names ending in a vowel. Both extremes of Figure 7 look as expected; the traditionally male and female names have almost 100% accuracy of expected categorization. Every name with a response rate assignment to female categories that was over 60% ended in a vowel (such as Olivia, Sophia, Latoya, Quantavia, and Nova). This is consistent with the by-phoneme graph in Figure 4.5.

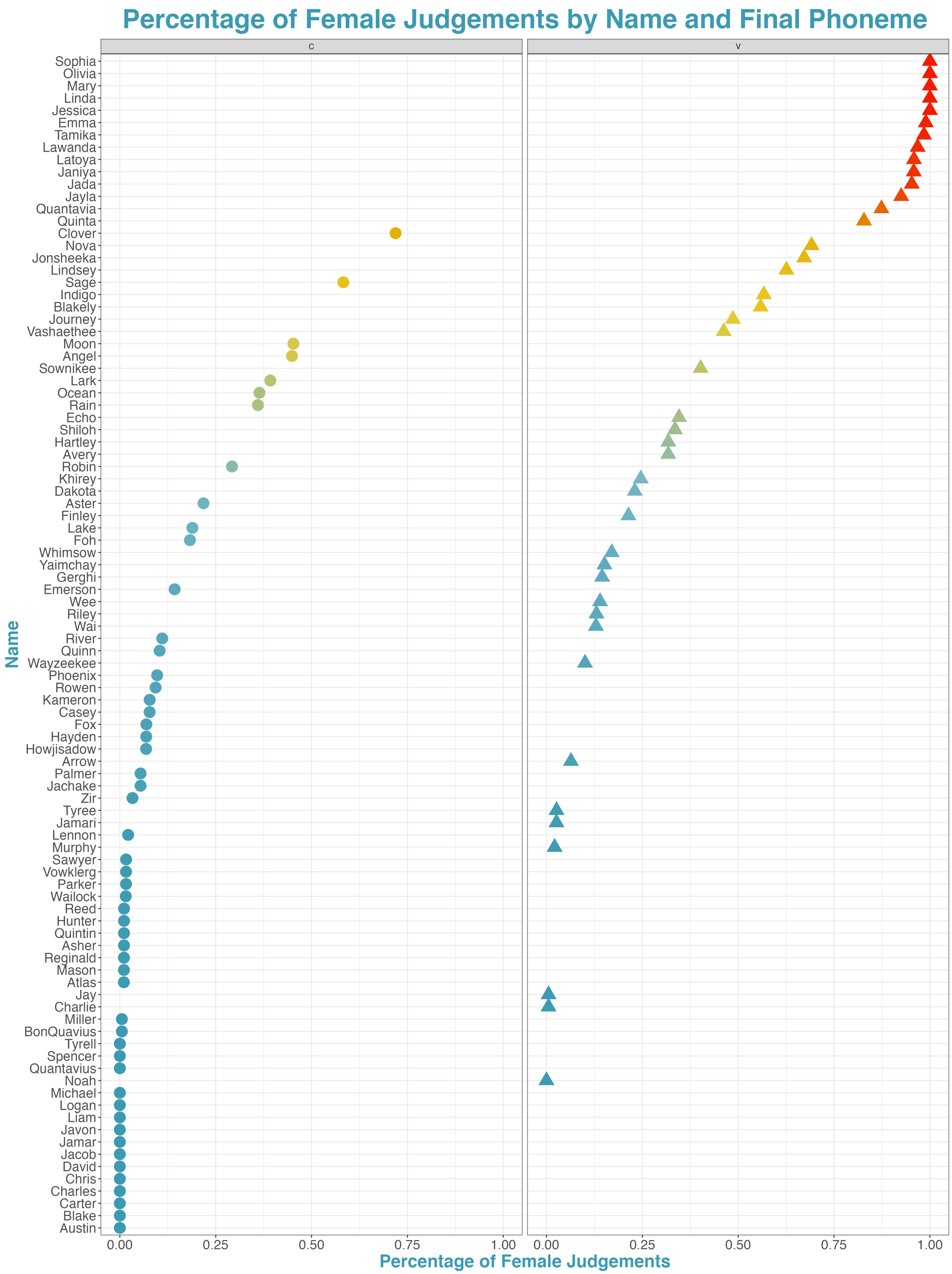

Figure 4.6: Distribution of participant responses by name and final phoneme

Figure 4.7 is the first of 5 graphs that isolate the response patterns for each category of stimulus names. Figure 4.7 shows judgments for with the generated nonce names, meaning that participants should be encountering each of these names for the first time. This provides us with interesting information about whether or not gender-specific phonological patterns are able to be extrapolated to novel data. Figure 4.7 shows that, of the generated names, on average the names ending in a vowel had a higher percentage of female assignment. The four names with the highest percentage of female judgments all ended in a vowel, and none had primary stress on the first syllable.

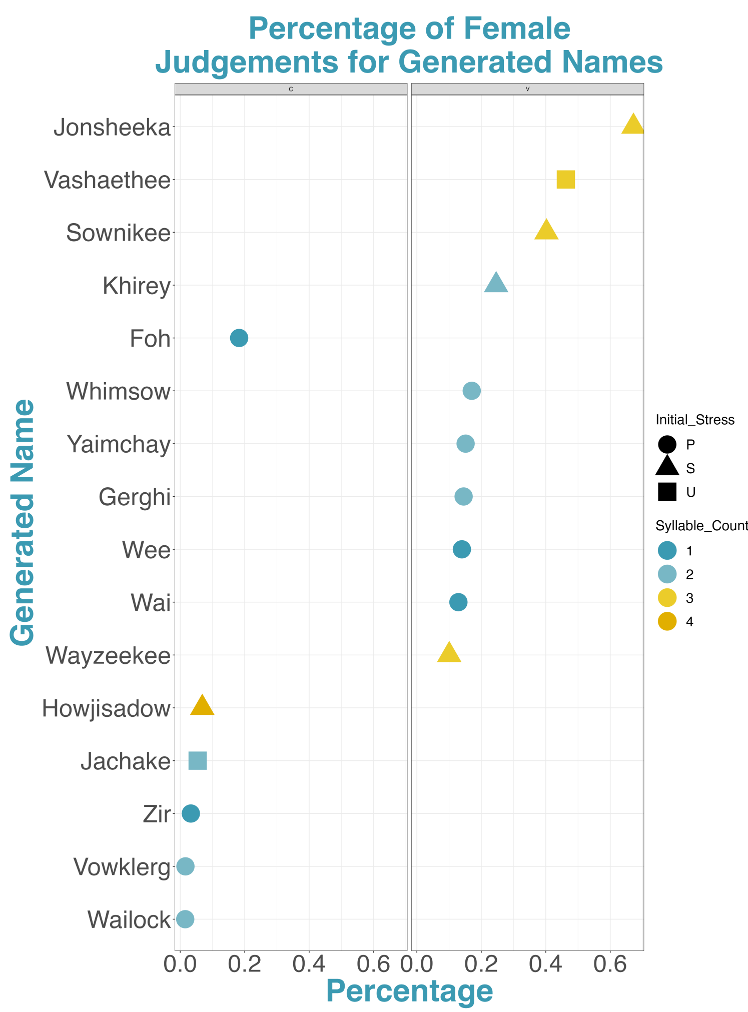

Figure 4.7: Percentage of female judgements for generated names

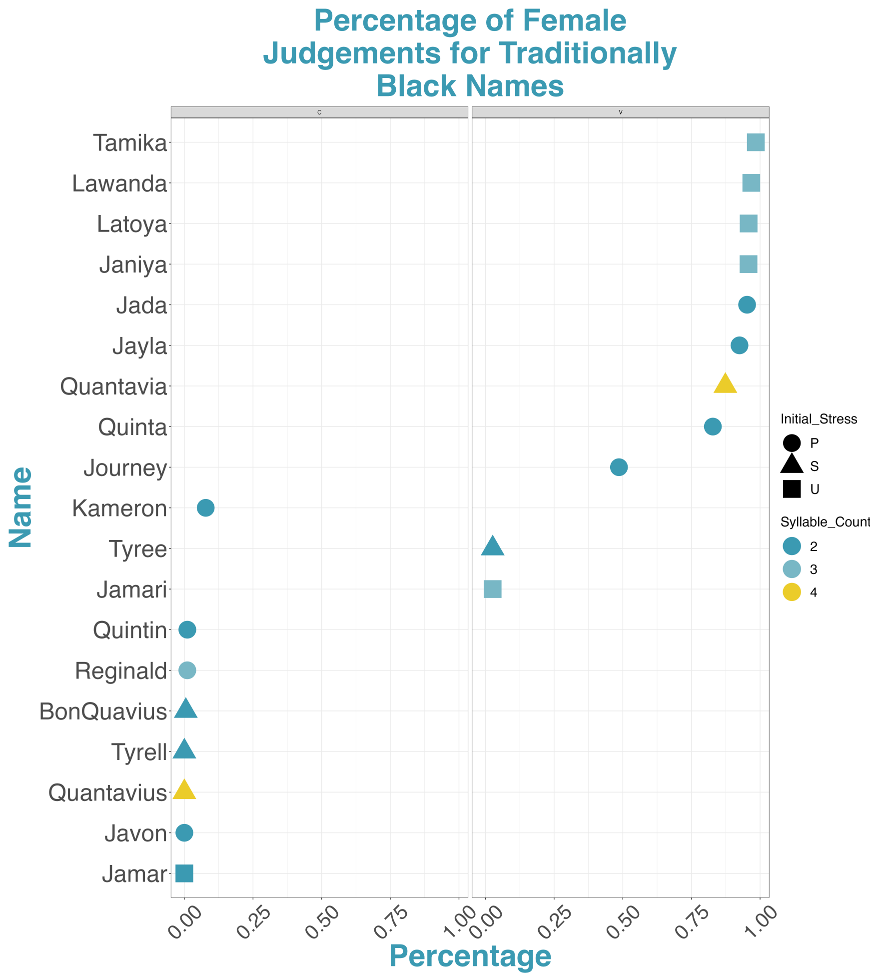

Figure 4.8: Percentage of female judgements for Black names

Figure 4.8 shows the distribution of gender assignment for names traditionally given to Black babies in the United States. Note the presence of a "minimal pair"-like set of names: *Quantavius* (usually assigned to male babies) and *Quantavia* (usually assigned to female babies). The only difference between these two names is their final phoneme, yet *Quantavius* was assigned to a male category 100% of the time, and Quantavia was assigned to a female category 80% of the time. Given that these names are not widely popular, this shows a strong association between gender assignment and final phoneme type.

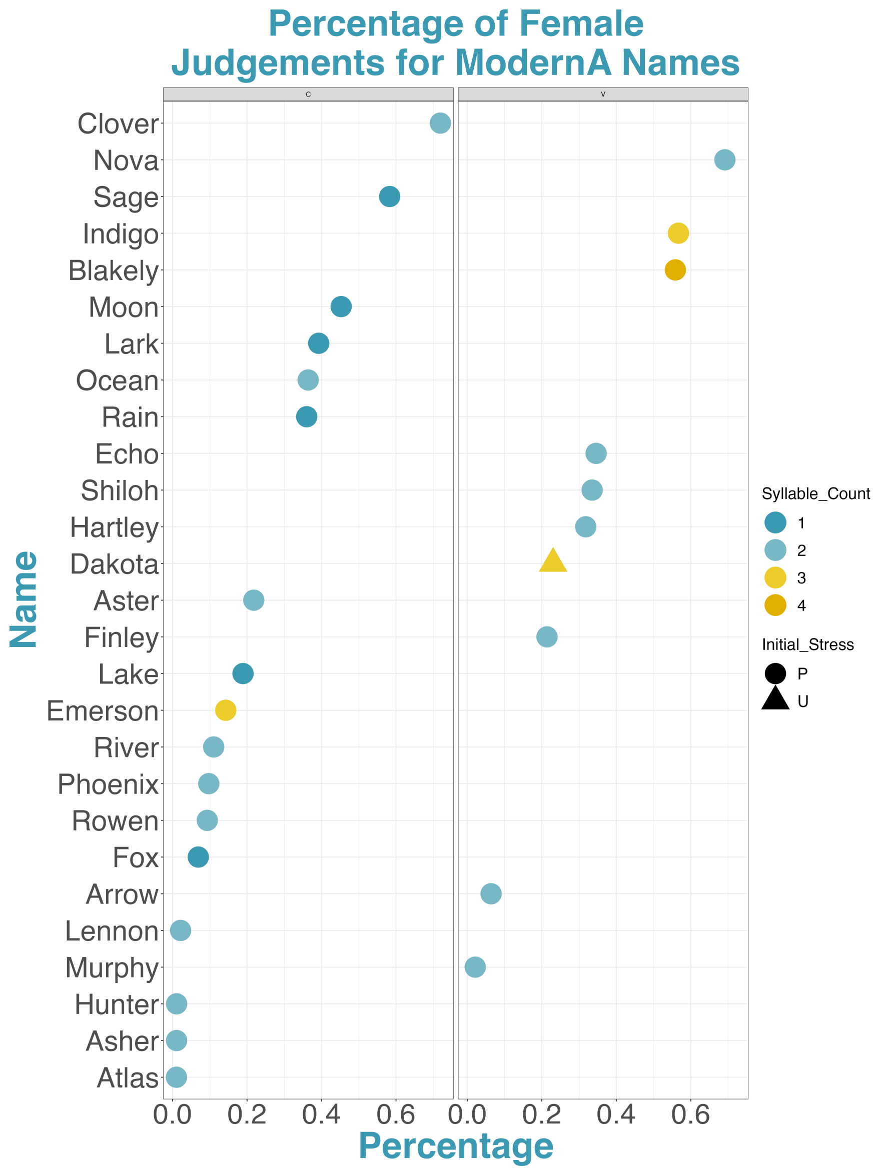

Figure 4.9: Percentage of female judgements for modern androgynous names

Figure 4.9 and Figure 4.10 compare the distribution of responses for modern androgynous names and traditionally androgynous names. Traditional androgynous names follow a very specific phonological pattern. All have either one or two syllables and primary stress, the majority of them end in a consonant, and on average a traditionally androgynous name was more likely to be assigned to a male category than a female category. On the other hand, modern androgynous names show much more phonological variability in terms of length and final phoneme, although all but two of the modern names still have initial primary stress.

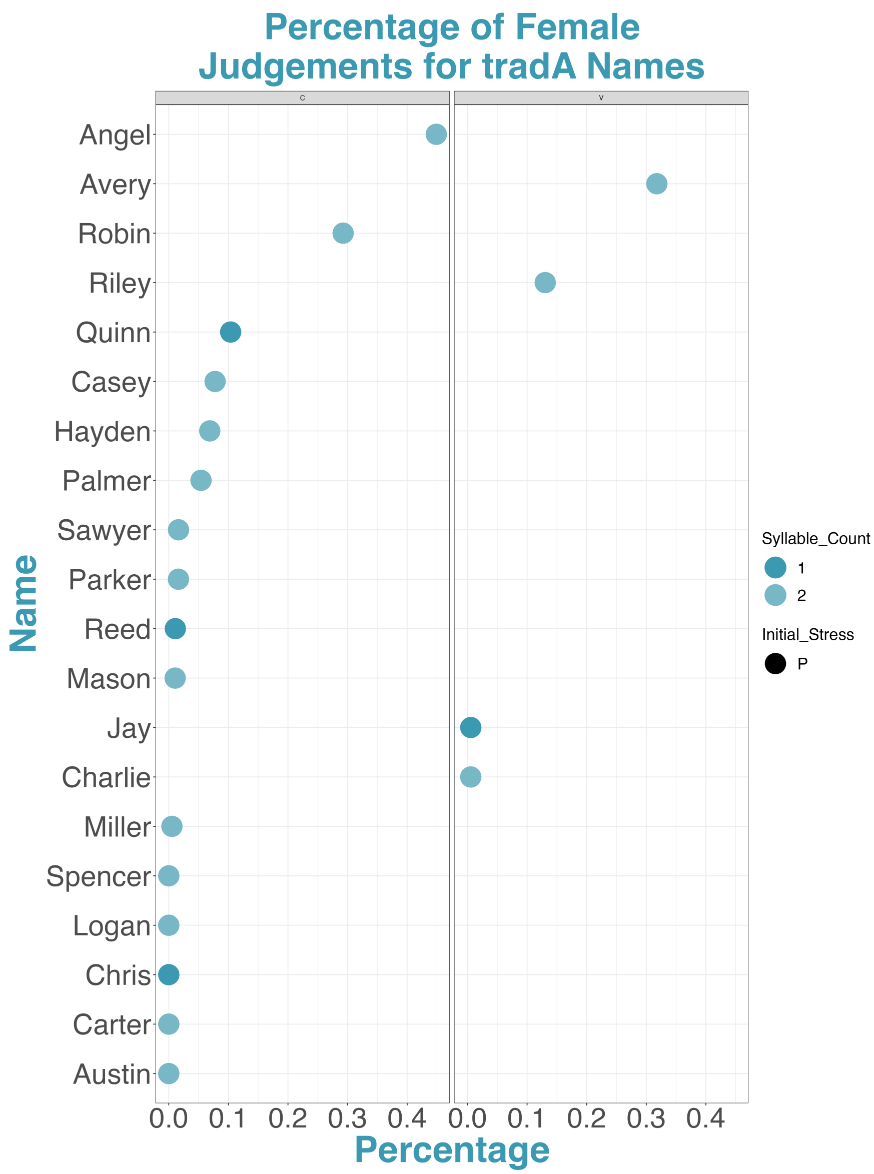

Figure 4.10: Percentage of female judgements for traditionally androgynous names

This final graph shows the response distribution for names that are traditionally understood to be male or female. Participants responded to these names as expected based on the understood sex of the name. The seven male names were never assigned to either of the female categories, and all seven of these names displayed the phonological characteristics previously found (Cutler et al 1986; Williams & Renwick 2022) to be associated with male names. The female names had a near 100% female assignment rate except for the name *Lindsey*, which was only assigned as female 60% of the time. The traditionally female names displayed a greater diversity of stress pattern, as well as a larger syllable count on average than the traditionally male names.

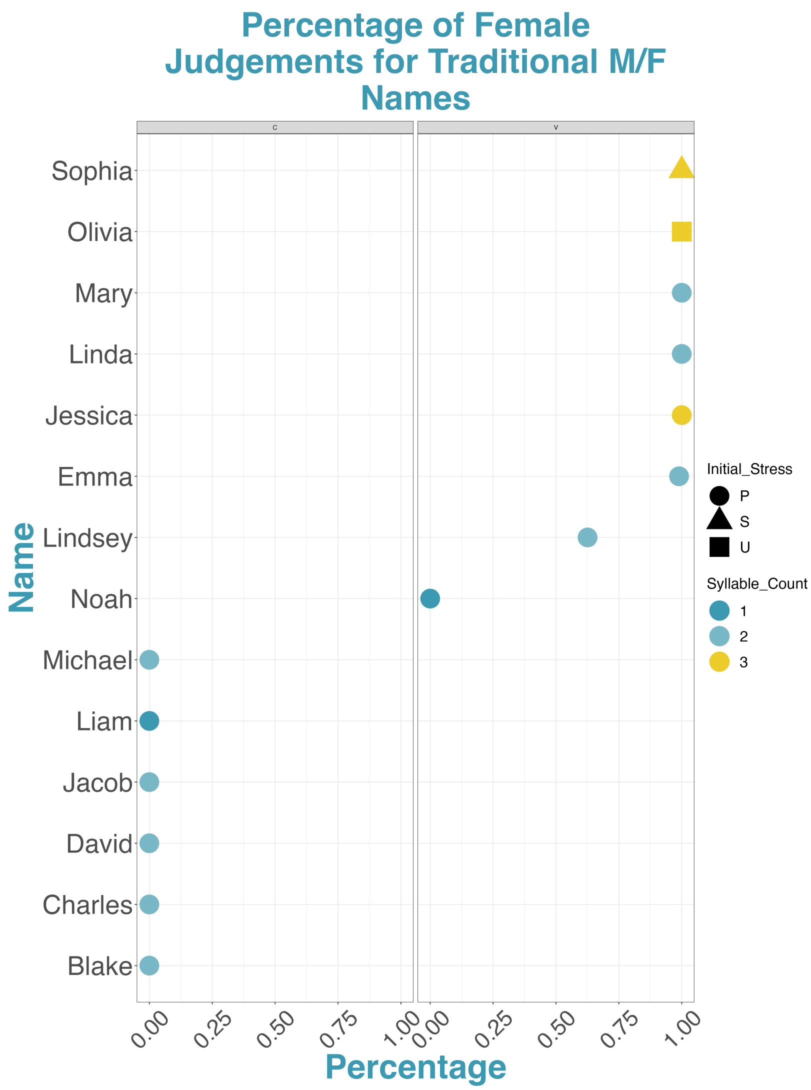

Figure 4.11: Percentage of female judgements for traditionally male and female names

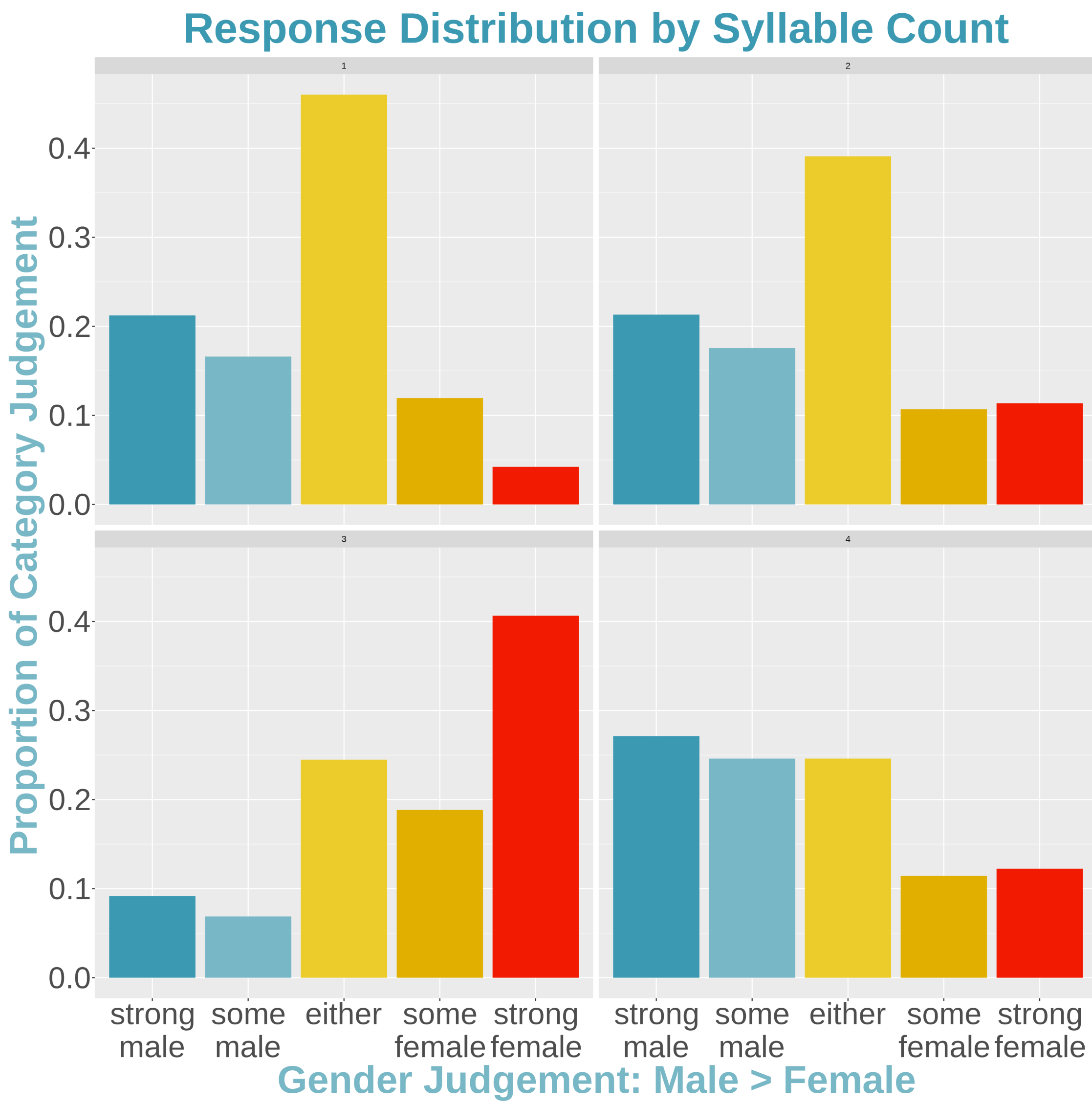

Figure 4.4: Distribution of participant responses by syllable count

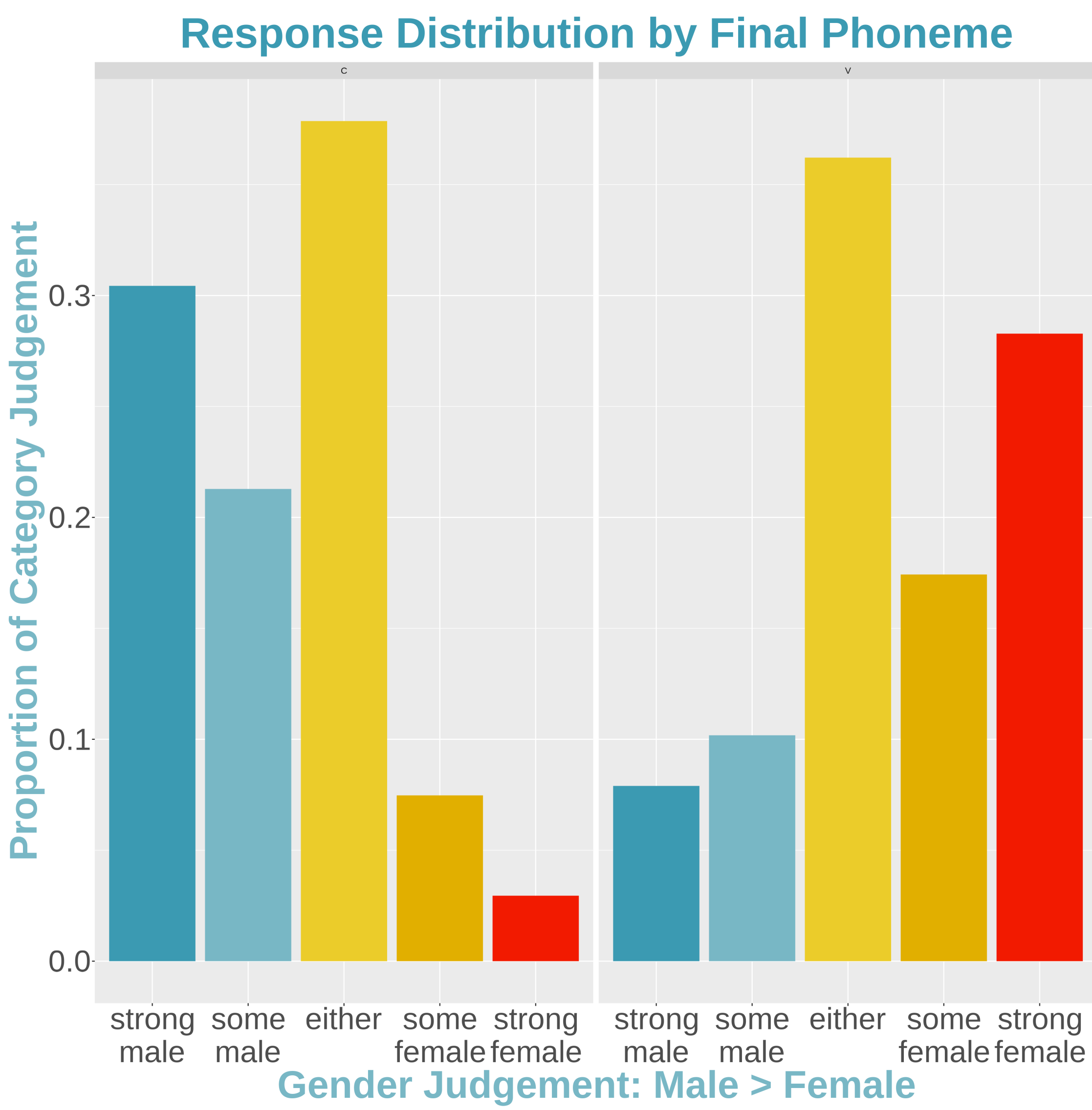

Figure 4.5: Distribution of participant responses by final phoneme

Chapter 5

Quantitative Analysis

The following section discusses and motivates the quantitative approach to the present dataset. First, I discuss the treatment of the response variable and present an argument in favor of preserving its ordered categories throughout the process of statistical analysis. I motivate this choice by contrasting key differences between metric and ordinal data, and explain why the response variable of the current study would benefit from an ordered approach. I then give a brief background of approaches to formally modeling ordinal data, with a focus on cumulative mixed models, document the progression of model creation and selection, and discuss the steps taken to extract useful information from the best fit model. The results of this analysis are discussed in Chapter 6.

5.1 Theoretical Approaches to Ordinal Data

When, as in the case of this study, the dependent variable of interest is the product of a Likert-style scaled judgment task, theoretical approaches to modeling such data diverge into two paths: preservation of the ordered categories vs. treating the categories as if they were unordered. Whether or not a dependent variable with ordered categories should be treated as an ordered categorical variable, an unordered categorical variable, or a continuous metric predictor has been the subject of much discussion and disagreement since Stevens' (1946) original argument in favor of a quantitative approach including ordered categories.

Distinctions between ordered categorical variables and other potential variable types are fairly easy to recognize given examples. A metric dependent variable in a linear regression could be the number of students graduating from a given class, where the independent variables are factors that could potentially impact graduation rates. If the dependent variable is condensed to a binary factor, where the two options are "did graduate" and "did not graduate", the distance between the categories is still clearly defined. Even if the number of categories is increased to "graduated early", "graduated on time", "graduated late", and "did not graduate", the boundaries of each category are still clear, and it is impossible for a student to occupy either more than one category, or a space in-between the categories.

Consider, however, a study that elicited people's *opinions* regarding whether or not a student would graduate based on some combination of presented variables. The response variable in this case would be an ordered categorical variable ranging from "most likely to graduate" to "least likely to graduate". Here the categories are not objectively defined, nor are they inherently universally understood the same way by each respondent. Though it is common to treat an ordinal variable as either an unordered categorical variable or a continuous linear predictor, the inherent difference in the function of these variable types indicates the need for a quantitative approach that will preserve the relationship between ordered categories without assuming that (a) they are equidistant from each other, or (b) the category boundaries are understood identically for every respondent.

In a regression analysis where the Likert scale data is treated as metric, it can be argued that the differences between participants' approach to the categories can be accounted for by including a random effect for participant. Though this may account for differences between participants' treatment of the scale data, what it does not allow for is analysis of the potential differences between the categories themselves. It is not that a random effect for participant is wrong to include, but rather that just including a random effect for participant does not allow for the exploration of potentially significant differences between category treatments.

The primary argument motivating an individualized approach to ordered categories is that ordered categories do not contain real metric information. Though numbers can be assigned to category levels, addition of, say, category 1 and category 2 does not necessarily produce category 3. While these numbers do correctly indicate the order of the categories, they do not give us any information about the intervals between the levels themselves, meaning that there could theoretically be a 1 unit increase between categories 1 and 2, and a 6 unit increase between category 3 and 4. Because of this, traditional statistical methods fail to capture the "psychological distance" (Burkner & Vuorre 2019) between categories, leading to models that can incorrectly estimate the extent of predictor impact and inflate Type I error rates (Liddell & Kruscke 2017).

In addition to prompting incorrect assumptions about the data, removing the "order" from ordered categories removes the potentially significant (or at the very least, interesting) findings regarding the boundaries that individual participants draw between the categories. A significant benefit of not assuming an equidistant relationship between categories is that it allows for the exploration and analysis of a psychological state that is not usually accessible for quantitative analysis. Often we are presented with data that is the result of a collection of beliefs or assumptions from a particular population, but it is less common that we are afforded the opportunity to investigate how a population defines its categories of belief. The assumption of equidistance between categories glosses over the potentially informative variation both between participants and between categories.

5.2 Application to the Current Dataset

Previous research discussing the types of data that would benefit from an ordinal approach has identified several specific characteristics that indicate the necessity of an ordinal approach, namely that (1) the categories of the data in question cannot be precisely measured (Baayen & Divjak 2017; Liddel, Burkner & Vuorre, Verrissimo 2021), (2) it cannot be assumed that the distance between each category is equal (Liddel, Burkner & Vuorrel 2019; Scot et al 1998; Verrissimo 2021), and (3), it cannot be assumed that there is equal category treatment by participants (Verrissimo 2021). In this section I justify the treatment of the current dataset as ordinal by highlighting the specific characteristics of the dataset that necessitate an ordinal approach, following the characteristics identified above.

It is easiest to understand and motivate the necessity of an ordinal approach when contrasting the characteristics of ordinal data to those of metric data. Based on the characteristics of ordinal data identified above, if we were going to treat the data collected in this study as metric, it would need to have the following characteristics:

1. We would need to be able to assume that the categories in question can be precisely measured. In this case, treating the response variable as a metric variable would mean relying on the underlying assumption that it is possible to measure and assign distinct, universally perceived boundaries to an abstract category such as "strongly male".

2. We would need to be able to assume that the distance between each category is equal, i.e. that there is an n unit increase between each level of the variable.

3. We would need to be able to assume identical category treatment by participants. That is, we must be able to assume that participant A's definition of the category "strongly male" is the same as that of participant B.

Regarding (1), a strong indicator that ordinal regression is the right approach for this dataset is that the categories within the scale do not have a definitive real-world referent. The categories "strongly male" and "strongly female" not only do not have an agreed upon value or range of values, but it is also not possible for me as the researcher to assign values to those categories. The values of the categories and the distances between them must therefore be constructed by each individual participant, and cannot be assigned a single universal value.

This leads to (2): if (1) is false, and we cannot precisely measure quantities in question, then it is logical to assume that we are also unable to assume that the distance between each category is equal. The decision to use an ordinal model in this case rests in the possibility of interesting results from intra-category variation. For example, there is a chance that the things that might influence a person's decision to rate a name as "strongly male" or "somewhat male" might be similar, whereas the factors that influence a person's decision between "strongly female" and "somewhat female" might differ, demonstrating greater variability in "female" traits versus "male" traits.

(3) requires equal treatment of categories across participants, which, given the discussion of (1) and (2), cannot be assumed for this dataset, as the lack of a definitive measure for both category boundaries and intervals results in an inability to guarantee that each participant's understanding of the categories is identical. There is also the possibility that participants

lean significantly more towards one category or side of the scale than the other, and the ability to incorporate this preference into the model could result in a better model fit.

Given the above discussion, we can conclude that, as we are unable to assume prior conception of category boundaries, equal intervals between categories, or consistent treatment across participants, preserving the ordered nature of the dependent variable (in addition to random effects per participant) will, in this case, not only ensure a more accurate model fit, but additionally provide valuable insights regarding the relationship between the categories themselves.

5.3 Ordinal Regression

Though several approaches to ordinal regression exist, the one best suited for the current dataset is a cumulative link model (Agresti 2002; Hensher & Greene 2009), which assumes that the dependent variable is the result of a latent continuous variable that explains the categorical division of the dependent variable. In this case, the latent variable would be the participants' beliefs regarding the response categories. This means that we are assuming that for each participant there is an underlying belief or assumption about the categories that influences their responses, and that we can use this underlying belief to divide the dependent variable into flexible categories.

5.3.1 Cumulative Link (Mixed) Models

In the case of this study we are interested in the probability that a value of `Gender_Judgement` (Y) falls into one of the established categories. Obtaining this probability requires assuming that there exists a latent continuous variable $\tilde{Y}$ that explains the categorization of Y. In this case $\tilde{Y}$ is the latent opinion about gender categorization. The formal modeling process begins by assuming that there are K thresholds, with an individual threshold represented as τ_k, and these thresholds divide $\tilde{Y}$ into *(K + 1)* observable categories of Y. The number of thresholds will be 1 less than the total number of categories in Y, so in this case we will have *(K + 1 = 5)* response categories and *(K = 4)* thresholds. Assuming the distribution of $\tilde{Y}$ is normal, a possible visualization of the modeling so far would look like 5.1:

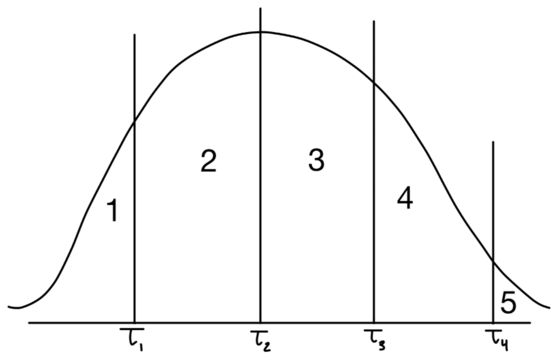

Figure 5.1: Visualization of possible model categories

Here, Y is divided into *(K + 1 = 5)* categories by K thresholds. We are interested in the probability that the value of Y falls within the boundaries of one of the categories. To calculate this, we will assume that $\tilde{Y}$ has a standard normal distribution and make use of the cumulative normal distribution function ϕ. We end up with the following equation for the probability that $Y = k$:

$$Pr(Y = k) = \phi(\tau_k) - \phi(\tau_{k-1}) \tag{5.1}$$

The above equation finds the probability that judgment Y is equal to a certain category of the dependent variable by subtracting the probability that Y is equal to one category from the probability that Y is equal to the category above it. Baayen & Divak (2017) succinctly describe this equation as calculating the size of a "rating bucket", where the probability of

Y being equal to k is assessed by calculating the distance between τ_k and τ_{k-1}. This is most easily illustrated with a toy problem, where we assume that $k = 2$, the first threshold is at –1 and the second is at 1.5. The resulting equation would be as follows:

$$Pr(Y = 2) = \phi(1.5) - \phi(-1) = 0.93 - 0.16 = 0.77 \tag{5.2}$$

Given the threshold values in the toy problem, the probability that Y is in the 2 rating "bucket" is 0.77. This equation, however, does not entirely fit our needs. We don't yet have a way to predict the probability that $Y = k$ given a set of predictor variables. For this, we will need to construct a linear regression for $\tilde{Y}$ using both a predictor term and an error term, where the predictor term will describe the variation in $\tilde{Y}$ that is explained by the predictors, and the error term will describe the variation in $\tilde{Y}$ that is not explained by the predictors. We can define the predictor term η as $\eta = b_1x_1 + b_2x_2...$ and the error term as ε. $\tilde{Y}$, then, is equal to $\eta + \varepsilon$. So, the probability that $Y = k$ given the predictor η is as follows:

$$Pr(Y = k|\eta) = \phi(\tau_k - \eta) - \phi((\tau_{k-1}) - \eta) \tag{5.3}$$

With the derivation up to this point established, we can now begin to relate these formulas to the data of interest in this study. Assume we want to predict the latent gender judgment opinion $\tilde{Y}$ given the phonological independent variable of `Final_Phoneme`. Here, x_n represents a value of the independent variable, and b_n the contrast between two values of the independent variable. For example, for the independent variable `Final_Phoneme` where the reference level is C, x_1 is equal to the effect that `Final_Phoneme` = V on the data with all other categories at their reference levels, and b_1 is the difference between the effects of `Final_Phoneme` = V and `Final_Phoneme` = C. If $\tilde{Y}$ is calculated using the following equation,

$$\tilde{Y}_k = \eta + \varepsilon = b_1 x_1 + \varepsilon \tag{5.4}$$

then the probability that $Y = k$ given the predictor variable is calculated as follows:

$$Pr(Y = k) = \phi(\tau_k - b_1 x_1) - \phi(\tau_{k-1} - b_1 x_1) \tag{5.5}$$

Understanding the derivation and application of the cumulative link model makes its resulting implementation in the R statistical software easier to both understand and to justify. Though the models discussed in the following section are more complicated than the equation above due to the inclusion of additional independent variables, the underlying formula is the same.

5.4 Model Creation

I employ a cumulative link mixed model to analyze the impact of phonological and sociological factors on gender judgement. A mixed model, one that includes random effects alongside fixed effects, is necessary here as there are multiple data points for each name and each respondent, and the inclusion of random effects for those two variables accounts for the non-independence.

Models were fit in R using the package `ordinal` (Christensen 2016). Of the available variables, `Choice_Order`, `Syllable_Count`, `Final_Phoneme`, `Initial_Stress`, `Respondent_Gender`, and `Respondent_Age` were chosen for inclusion in the model alongside the random effects of `Name` and `Participant`. Models containing the effects of `Respondent_Gender` and `Respondent_Age` failed to converge, and were therefore excluded from quantitative analysis. Variable selection was conducted using a stepwise method. An interaction between syllable count and stress pattern was attempted and found to be insignificant.

5.5 Extracting Predicted Probabilities

Unlike other packages for modeling in R, `ordinal` does not have built in functions for either extracting predictors or plotting them, nor does it integrate well with other packages intended for those purposes. Subsequently, in order to extract the necessary information from the model, I modified existing functions from Lankoski (2013) to fit my dataset. Namely, the original functions did not allow for the preservation of labels for the data, meaning that in any resulting graphs it was difficult or impossible to decipher one variable's effect from another, and this was particularly compounded by the amount of predictors I wanted to extract. In order to obtain predicted probabilities from the model given a combination of variables, the following lines of code extract the fixed effects regression parameters (`(clmm.9$beta)` and label them according to their corresponding variable.

```
varNames <- as.data.frame(rowSums(expand.grid(clmm.9$beta)))
rownames(varNames) <-
    c("syll2","syll3","syll4","stressS",
    "stressU","phonV","catG","catMA","catTA",
    "catTF","catTM","choice2")
eta <- rowSums(varNames)
```

Listing 5.1: Fixed effects extraction

Fixed effects, random effects and thresholds are extracted and stored as `theta`. An integer span of 1 to the length of `theta + 1` is stored in `cat`), and the inverse of the link function plogis is stored as inv.link; this will be used to transform the predictor values to the original scale of the response variable.

```
theta <- clmm.9$Theta
cat <- 1:(length(theta) + 1)
inv.link = plogis
Theta <- c(-1000, theta, 1000)
```

Listing 5.2: Fixed effects storage and transformation

Finally, for each value in Theta where Theta[j] is the lower threshold of a category and Theta[j + 1] is the higher threshold of a category, the difference between the predictor values for the high and low thresholds and the current data point are calculated. The resulting value for the low threshold is subtracted from the resulting value from the high threshold to obtain the predictor value.

```
preds <- sapply(cat, function(j) inv.link(Theta[j + 1] - eta) - inv.
link(Theta[j] - eta))
```

Listing 5.3: Calculation of thresholds

The resulting object can be transformed into a dataframe and used within traditional plotting packages such as ggplot; graphical exploration of predicted values is further discussed in Chapter 6.

Chapter 6

Results

The model summary for the final model is seen in Table 6.1 (fixed effects) and Table 6.2 (random effects) below. The only significant predictor in the final model was `Final_Phoneme` `(z = 4.414, p < 0.001)`. Other than the obviously significant predictor, the results of the model are difficult to interpret directly from the model summary. To remedy this, the following section discusses graphs made with the code outlined in §5.

Table 6.1: Model summary for fixed effects of final cumulative link mixed model

Variable	Estimate	Standard Error	z value	Pr(> $\|z\|$)
`Syllable_Count2`	-0.80159	0.54274	-1.477	0.140
`Syllable_Count3`	0.37544	0.84528	0.444	0.657
`Syllable_Count4`	-0.38638	1.40948	-0.274	0.784
`Initial_StressS`	-0.03695	0.85741	-0.043	0.966
`Initial_StressU`	0.07235	0.80045	0.090	0.928
`Final_PhonemeV`	1.98892	0.45057	4.414	1.01e-05 ***
`Choice_Order2`	0.06313	0.04719	1.338	0.181

Table 6.2: Model summary for random effects of final cumulative link mixed model

Groups	Variance	Standard Deviation
Response_ID	0.1075	0.3279
Name	3.2245	1.7957

6.1 Visualization of Predictor Impacts

Figure 6.1 shows the predicted impacts of each variable on the five possible response categories. As is consistent with the descriptive visualizations, the “either” category is predicted to have the highest number of responses. The top left panel of Figure 6.1 shows that, with all other variables at their reference level, a name having two syllables is more likely to be assigned as having “strong male” characteristics than a name with three or four syllables. A name ending in a vowel is almost never predicted to be assigned to the “strong male” category, consistent both with the original hypothesis that a word-final vowel is a characteristic usually attributed to female names, and to the descriptive graphs showing that participants were most likely to assign names ending in vowels to female categories.

The higher percentages on the y axis of the top left panel of Figure 6.1 shows that a name having two syllables had a bigger impact on a name being assigned as “somewhat male” than it did for a name being assigned “strongly male”. In fact, it seems that all predictors had a more significant effect on the “somewhat male” category than the “strong male” category, indicating that names in the “somewhat male” category had greater phonological variability than names in the “strongly male” category, perhaps compounded by the presence of the control names which showed higher concentrations of expected male characteristics. Note that the reference levels for the variables are a name having one syllable, ending in a consonant, and primary stress – these are traits hypothesized to be highly associated with male names.

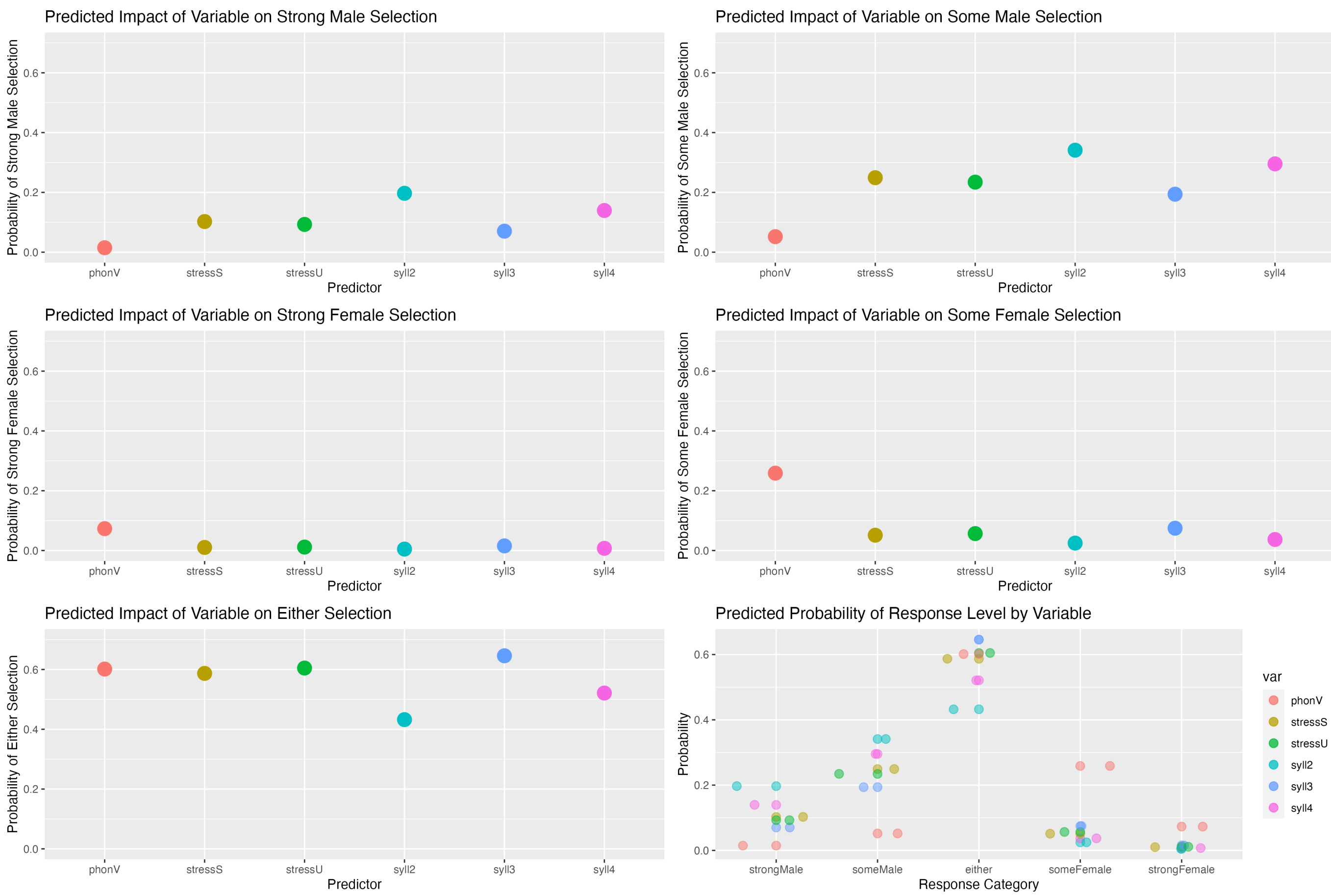

Figure 6.1: Predictor plots by category selection

Taken in conjunction with the top two panels of Figure 6.1, this shows that names in the "strongly male" category were likely to have the characteristics present in the reference level.

The second row of Figure 6.1 shows the predicted probabilities for the female response categories. The only variable to have more than a minimal predicted impact on response category here is if the name ends in a vowel, which marginally increases the chances of a name being assigned to the "strong female" category, and increases the probability of a name being assigned to the "somewhat female" category by a larger margin. It is important to note here that the probabilities of category assignment for both categories is not very high. This can be attributed to the fact that the reference levels for all variables are characteristics we expect to see in male names. These low probabilities then provide further support for the argument in favor of gendered phonology, as names having what are considered to be the strongest predictors of a male name have low predicted probabilities of being assigned female.

The bottom left panel of Figure 6.1 shows the impact of the variables on the "either" response category, which is selected more frequently than any of the other four categories. Most of the variables have an approximately equal impact on category assignment, pointing towards the status of the "either" category as a neutral category without any specific identifying phonological factors. The variable with the most significant impact on the "either" category is a name having two syllables, which, interestingly, is the strongest predictor of a name being assigned to a male category. This indicates that while "either" might not have strong positive category predictors, there are variables that make it less likely for a name to be assigned to the "either" category.

The final panel at the bottom right of Figure 6.1 shows the prediction of the probability of a response category being chosen as well as the variables that impact that probability.

Looking at the combined effects of the variables, the predictor `Final_Phoneme = V` visually distinguishes itself from the other predictors as having the strongest impact on a name's category assignment, where names ending in a vowel are less likely to be assigned to male categories, and more likely to be assigned to female categories, especially the "somewhat female" category. This is consistent with what we saw in the model summary, where `Final_Phoneme = V` was the only significant predictor.

Chapter 7

Discussion & Conclusion

7.1 Discussion of Hypotheses

Returning to the presentation of hypotheses in §2.4, we find that Hypothesis 1a was partially supported by the data. While participants did use phonological cues from previous corpus-based name studies to differentiate between male and female names, only one of the previously identified cues was found to be significant. Whether or not a name ended in a vowel or a consonant was the primary method of distinguishing between male and female names. The effects of stress and name length were not found to be significant predictors. This finding differs significantly from previous results of corpus-based studies (Cutler et al 1986; Whissel 2001; Williams & Renwick 2022), which found that stress and name length were predictive characteristics of male and female names in corpus data.

Two different directions could be taken to explain these results: either the results of this study highlight a gap between observed trends in popular names and the functionally salient characteristics of name sex, or there has been a change in linguistic conception in the last decade of naming practices. Previous studies did not collect perceptual data, meaning that it is possible that not all the trends observed in corpus data are actually salient or productive in real-world language use. Additionally, previous corpus studies only looked at naming trends up to the year 2010, meaning that in the last ten years a shift could have taken place where female and male names became less phonologically distinct. This is supported by results from Williams & Renwick (2022), which found that female and male names grew increasingly less distinctive after about 1960.

It is highly plausible that the increase in immigration and general diversification of the United States population plays a role in changing linguistic conventions of name sex. Choice of name is closely tied to the cultural and social conventions of communities, and as the population of those communities change, so do their conventions. This is supported by findings from Williams & Renwick (2022), who found a significant shift in the phonological distribution of both male and female names in the 1960s, which is also the decade in which the 1965 Act to Eliminate Race Discrimination in Immigration was passed and resulted in a diversification of the population of the United States.

Taken in sum, whether the results of this study are due to a gap in functional salience, a cultural and linguistic shift, or some combination of the two, it is abundantly clear that naming practices reflect a symbolic association between name sex and phonological distribution, and that this association is not static, leading to the conclusion that it is the result of widely applicable cultural and linguistic norms. This is proven further when looking at the results of Hypothesis 1b, which was found to be supported by the data: unfamiliar names patterned with known names in terms of category assignment. Participants were able to draw upon intuitions regarding name gender and use this to predict the gender of unfamiliar names, as is seen in the results for the generated nonce names as well as for names with a low familiarity score. This shows that gender association for names is not merely a function of previous knowledge of someone bearing that name, but rather a result of a set of phonological associations.

This discussion of the social and linguistic factors that impact changes in name phonology is also interesting when looked at in conjunction with the results of the hypotheses regarding demographic characteristics. Neither of the hypotheses regarding demographic characteristics were proven to be correct. The response patterns of participants did not vary significantly between demographic groups, meaning that the association of word-final vowels with female names was not impacted by any specific demographic factors. The fact that this hypothesis was shown to be false introduces interesting questions regarding the bounds of the broader demographic categories in which this same symbolic association would be found. Future research would benefit from a similar study with a focus on obtaining a larger gender and region diversity from participants. The majority of the participants in this study were women between the ages of 40 and 60 from the southeastern United States; demographic angles of similar future studies would benefit from an expanded focus on obtaining a diversity of responses for the categories of gender, age and region.

7.2 Conclusion

The results of this study show that there are discernible phonological characteristics in English first names that impact whether a name is categorized as male or female, and that this is not restricted to a particular name type or category. Additionally, participants were able to use phonological associations to determine the gender assignment of unfamiliar names, and their gender assignment was consistent with the expected characteristics of male and female names. These results suggest that there are learned patterns of sound and gender mapping that can be at least partially attributed to gendered phonology in English.

This result is particularly interesting when taken in the context of the introductory discussion of cross-linguistic sound symbolism. Though English is not a language that is traditionally understood to have sound-symbolic patterns, studies on specific subsets of English, including the present study, have shown that there are observable symbolic patterns that are individual to a particular area of English. Targeted approaches to sound-symbolic research have consistently found symbolic associations between sounds and meaning, especially when the act of naming is involved. This points to the conclusion that, while language as a whole might be arbitrary, users of language can form widely-held symbolic associations as a result of analogizing a particular sound to a sociological or cultural idea. This directs the lines of inquiry into sound symbolism away from investigating language as a whole as symbolic, which is nearly universally understood to be false, and towards the use of sound-symbolic data as potential lens through which to view the extra-linguistic information speakers' use to inform their perception.

Appendix A

Stimulus Names by Category

TA	MA	TM	TF	GG	TB
Angel	Arrow	Charles	Emma	Foh	BonQuavius
Austin	Asher	David	Jessica	Gerghi	Jada
Avery	Aster	Jacob	Linda	Howjisadow	Jamar
Blake	Atlas	Liam	Lindsey	Jachake	Jamari
Carter	Blakely	Michael	Mary	Jonsheeka	Janiya
Casey	Clover	Noah	Olivia	Khirey	Javon
Charlie	Dakota		Sophia	Sownikee	Jayla
Chris	Echo			Vashaethee	Journey
Hayden	Emerson			Vowklerg	Kameron
Hunter	Finley			Wai	Latoya
Jay	Fox			Wailock	Lawanda
Logan	Hartley			Wayzeekee	Quantavia
Mason	Indigo			Wee	Quantavius
Palmer	Lake			Whimsow	Quinta
Parker	Lark			Yaimchay	Quintin
Reed	Lennon			Zir	Reginald
Riley	Miller				Tamika
Robin	Moon				Tyree
Quinn	Murphy				Tyrell
Sawyer	Nova				
Spencer	Ocean				
	Phoenix				
	Rain				
	River				
	Rowen				
	Sage				
	Shiloh				

Appendix B

Generated Names

One Syllable	Two Syllables	Three Syllables	Four Syllables
Foh	Wailock	Wayzeekee	Howjisadow
Way	Vowklerg	Vashaethee	Gaysthatibley
Day	Jachake	Payowdiy	Zaypayjihtee
Tin	Yaimchay	Badlerzter	Midazlerda
Ned	Chunshah	Tivtheyshow	Hoovliygita
Wai	Taydig	Wechaythgays	Tanjervsheetag
Wee	Khirey	Sownikee	Pashaytaepdaow
Sow	Wayow	Dazgamrew	Moojidowmow
Zir	Whimsow	Jonsheeka	Dalergberday
Jih	Gerghi	Doetabay	Faylugaver

Appendix C

Average Familiarity Rating

Name	**Average Familiarity Rating**
David	3.83
Michael	3.8
Chris	3.78
Mary	3.76
Jessica	3.71
Emma	3.61
Charles	3.58
Olivia	3.57
Sophia	3.56
Linda	3.52
Lindsey	3.52
Jacob	3.51
Charlie	3.42
Casey	3.35

Jay	3.35	
Riley	3.33	
Avery	3.32	
Noah	3.29	
Robin	3.27	
Austin	3.24	
Mason	3.19	
Liam	3.18	
Blake	3.16	
Parker	3.16	
Logan	3.14	
Hunter	3.1	
Carter	3.06	
Spencer	2.97	
Angel	2.92	
Hayden	2.88	
Reed	2.87	
Quinn	2.76	
Kameron	2.72	
Dakota	2.63	
Sawyer	2.6	
Emerson	2.59	
Tamika	2.54	
Asher	2.5	

Echo	2.5	
Quintin	2.46	
Jada	2.43	
River	2.38	
Sage	2.38	
Reginald	2.36	
Palmer	2.35	
Rowen	2.34	
Miller	2.32	
Finley	2.29	
Murphy	2.29	
Blakely	2.22	
Latoya	2.22	
Jamar	2.18	
Shiloh	2.17	
Tyrell	2.17	
Phoenix	2.15	
Rain	2.12	
Jayla	2.05	
Tyree	2.02	
Javon	1.99	
Lennon	1.91	
Lawanda	1.85	
Jamari	1.82	

Lake	1.82	
Atlas	1.79	
Indigo	1.78	
Journey	1.73	
Hartley	1.71	
Aster	1.69	
Nova	1.68	
Quantavius	1.67	
Clover	1.61	
Fox	1.58	
Lark	1.58	
Moon	1.52	
Ocean	1.52	
Janiya	1.47	
Quantavia	1.36	
Arrow	1.29	
Quinta	1.22	
Wee	1.2	
Khirey	1.18	
Wai	1.17	
BonQuavius	1.13	
Vashaethee	1.11	
Foh	1.09	
Zir	1.07	

Wailock	1.05	
Gerghi	1.04	
Jonsheeka	1.03	
Vowklerg	1.02	
Wayzeekee	1.02	
Whimsow	1.02	
Yaimchay	1.02	
Howjisadow	1.01	
Jachake	1.01	
Sownikee	1.01	

Bibliography

Agresti, A. (1989). Tutorial on modeling ordered categorical response data [Place: US Publisher: American Psychological Association]. *Psychological Bulletin*, *105*(2), 290–301. https://doi.org/10.1037/0033-2909.105.2.290

Agresti, A., & Tarantola, C. (2018). Simple ways to interpret effects in modeling ordinal categorical data. *Statistica Neerlandica*, *72*(3), 210–223. https://doi.org/10.1111/stan.12130

Amidei, J., Piwek, P., & Willis, A. (2019). The use of rating and Likert scales in Natural Language Generation human evaluation tasks: A review and some recommendations. *Proceedings of the 12th International Conference on Natural Language Generation*, 397–402. https://doi.org/10.18653/v1/W19-8648

Ananth, C. V., & Kleinbaum, D. G. (1997). Regression models for ordinal responses: A review of methods and applications. *International Journal of Epidemiology*, *26*(6), 1323–1333. https://doi.org/10.1093/ije/26.6.1323

Auracher, J. (2017). Sound iconicity of abstract concepts: Place of articulation is implicitly associated with abstract concepts of size and social dominance [Publisher: Public Library of Science]. *PLOS ONE*, *12*(11), e0187196. https://doi.org/10.1371/journal.pone.0187196

Auracher, J., Albers, S., Zhai, Y., Gareeva, G., & Stavniychuk, T. (2010). P Is for Happiness, N Is for Sadness: Universals in Sound Iconicity to Detect Emotions in Poetry. *Discourse Processes*, *48*(1), 1–25. https://doi.org/10.1080/01638531003674894

Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism on object recognition. *Memory & Cognition*, *40*(1), 83–92. https://doi.org/10.3758/s13421-011-0139-3

Baayen, R. H., & Divjak, D. (2017). Ordinal GAMMs: A new window on human ratings.

Barry, H., & Harper, A. S. (2003). Final Letter Compared with Final Phoneme in Male and Female Names. *Names*, *51*(1). https://doi.org/10.1179/nam.2003.51.1.13

Barry III, H., & Harper, A. S. (2010). Racial and Gender Differences in Diversity of First Names [Publisher: Routledge]. *Names*, *58*(1), 47–54. https://doi.org/10.1179/175622710X12590782368107

Barton, D. (2016). *Exploring a social "bouba-kiki" effect and the affective consequences of having a congruent versus incongruent name* (Thesis) [Accepted: 2016-06-27T04:38:13Z]. University of Otago. Retrieved September 14, 2022, from https://ourarchive.otago.ac.nz/handle/10523/6647

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English [Publisher: Cambridge University Press]. *Phonology*, *3*, 255–309. https://doi.org/10.1017/S095267570000066X

Bergen, B. K. (2004). The Psychological Reality of Phonaesthemes [Publisher: Linguistic Society of America]. *Language*, *80*(2), 290–311. Retrieved August 27, 2020, from https://www.jstor.org/stable/4489664

Black, S., Wilcox, B., & Platt, B. (2014). Nicknames in Prison: Meaning and Manipulation in Inmate Monikers. *Names*, *62*(3), 127–136. https://doi.org/10.1179/0027773814Z.00000000077

Burkner, P.-C., & Vuorre, M. (n.d.). Ordinal Regression Models in Psychology: A Tutorial. https://doi.org/10.1177/2515245918823199

Cassidy, K. W., Kelly, M. H., & Sharoni, L. J. (1999). Inferring gender from name phonology [Place: US Publisher: American Psychological Association]. *Journal of Experimental Psychology: General*, *128*(3), 362–381. https://doi.org/10.1037/0096-3445.128.3.362

Chen, F., & Kenstowicz, M. (2022). Phonotactics of Gender in Mandarin Given Names: Patterns and Constraints [Number: 0]. *Proceedings of the Annual Meetings on Phonology*, *9*(0). https://doi.org/10.3765/amp.v9i0.5170

Christensen, R. H. B. (2015). Analysis of ordinal data with cumulative link models — estimation with the R-package ordinal.

Christensen, R. H. B. (2019). Cumulative Link Models for Ordinal Regression with the R Package ordinal.

Christensen, R. H. B. (2022). A Tutorial on fitting Cumulative Link Mixed Models with clmm2 from the ordinal Package.

Christensen, R. H. B., & Brockhoff, P. B. (2013). Analysis of sensory ratings data with cumulative link models. *Journal de la société française de statistique*, *154*(3), 58–79. Retrieved October 19, 2023, from http://www.numdam.org/item/JSFS_2013__154_3_58_0/

Cutler, A. (2015). Lexical Stress in English Pronunciation [Section: 6]. In *The Handbook of English Pronunciation* (pp. 106–124). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118346952.ch6

Cutler, A., McQueen, J., & Robinson, K. (1990). Elizabeth and John: Sound Patterns of Men's and Women's Names [Publisher: Cambridge University Press]. *Journal of Linguistics*, *26*(2), 471–482. https://www.jstor.org/stable/4176070

D'Anselmo, A., Prete, G., Zdybek, P., Tommasi, L., & Brancucci, A. (2019). Guessing Meaning From Word Sounds of Unfamiliar Languages: A Cross-Cultural Sound Symbolism Study [Publisher: Frontiers]. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00593

Darden, D. K., & Robinson, I. E. (1976). Multidimensional Scaling of Men's First Names: A Sociolinguistic Approach [Publisher: [American Sociological Association, Sage Publications, Inc.]]. *Sociometry*, *39*(4), 422–431. https://doi.org/10.2307/3033507

De Carolis, L. (2019). *Experimental assessment of sound symbolism and evolutionary considerations* (Doctoral dissertation).

De Carolis, L., Marsico, E., & Coupé, C. (2017). Evolutionary roots of sound symbolism. Association tasks of animal properties with phonetic features. *Language & Communication*, *54*, 21–35. https://doi.org/10.1016/j.langcom.2016.10.003

de Klerk, V., & Bosch, B. (1997). The sound patterns of English nicknames. *Language Sciences*, *19*(4), 289–301. https://doi.org/10.1016/S0388-0001(96)00070-8

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages [Publisher: Linguistic Society of America]. *Language*, *92*(2), e117–e133. https://doi.org/10.1353/lan.2016.0034

D'Onofrio, A. (2013). Phonetic Detail and Dimensionality in Sound-shape Correspondences: Refining the Bouba-Kiki Paradigm: [Publisher: SAGE PublicationsSage UK: London, England]. *Language and Speech*. https://doi.org/10.1177/0023830913507694

D'Onofrio, A., & Eckert, P. (2021). Affect and iconicity in phonological variation [Publisher: Cambridge University Press]. *Language in Society*, *50*(1), 29–51. https://doi.org/10.1017/S0047404520000871

Flego, S., & Forrest, J. (2021). Leveraging the temporal dynamics of anticipatory vowel-to-vowel coarticulation in linguistic prediction: A statistical modeling approach. *Journal of Phonetics*, *88*, 101093. https://doi.org/10.1016/j.wocn.2021.101093

Fredrickson, A. (2007). Phonological Cues to Gender in Sex-Typed and Unisex Names [Accepted: 2013-01-22T19:16:37Z]. Retrieved September 13, 2022, from https://scholarship.tricolib.brynmawr.edu/handle/10066/10189

Greene, W., & Hensher, D. (2009). *Modeling Ordered Choices: A Primer* [Journal Abbreviation: Modeling Ordered Choices: A Primer Publication Title: Modeling Ordered Choices: A Primer]. https://doi.org/10.1017/CBO9780511845062

Haiman, J. (2012). *Ideophone and the Evolution of Language.*

Hayes, B. P. (1980). *A metrical theory of stress rules* (Thesis) [Accepted: 2009-01-23T14:36:59Z]. Massachusetts Institute of Technology. Retrieved October 19, 2023, from https://dspace.mit.edu/handle/1721.1/16066

Heyman, T., Maerten, A.-S., Vankrunkelsven, H., Voorspoels, W., & Moors, P. (2019). Sound-Symbolism Effects in the Absence of Awareness: A Replication Study. *Psychological Science (0956-7976)*, *30*(11), 1638–1647. https://doi.org/10.1177/0956797619875482

Hinton, L., Nichols, J., & Ohala, J. J. (1994). *Sound symbolism.* Cambridge University Press.

Hopper, P., & Thompson, S. (1985). The iconicity of the universal categories "noun" and "verb". https://doi.org/10.1075/tsl.6.08hop

Hough, C. (2000). Towards an Explanation of Phonetic Differentiation in Masculine and Feminine Personal Names [Publisher: Cambridge University Press]. *Journal of Linguistics*, *36*(1), 1–11. Retrieved October 7, 2020, from https://www.jstor.org/stable/4176570

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution [Publisher: Royal Society]. *Philosophical Transactions: Biological Sciences*, *369*(1651), 1–13. Retrieved August 27, 2020, from https://www.jstor.org/stable/24501124

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*(1), 54–65. https://doi.org/10.1016/j.cognition.2008.07.015

Kasof, J. (1993). Sex bias in the naming of stimulus persons [Place: US Publisher: American Psychological Association]. *Psychological Bulletin*, *113*, 140–163. https://doi.org/10.1037/0033-2909.113.1.140

Kawahara, S. (2020). A wug-shaped curve in sound symbolism: The case of Japanese Pokémon names. *Phonology*, *37*(3), 383–418. https://doi.org/10.1017/S0952675720000202

Kawahara, S., Godoy, M. C., & Kumagai, G. (2020). Do Sibilants Fly? Evidence from a Sound Symbolic Pattern in Pokémon Names [Publisher: De Gruyter Section: Open Linguistics]. *Open Linguistics*, *6*(1), 386–400. https://doi.org/10.1515/opli-2020-0027

Kawahara, S., Noto, A., & Kumagai, G. (2018). Sound Symbolic Patterns in Pokémon Names [Publisher: Karger Publishers]. *Phonetica*, *75*(3), 219–244. https://doi.org/10.1159/000484938

Klink, R. R. (2001). Creating Meaningful New Brand Names: A Study of Semantics and Sound Symbolism [Publisher: Routledge]. *Journal of Marketing Theory and Practice*, *9*(2), 27–34. https://doi.org/10.1080/10696679.2001.11501889

Klink, R. R. (2003). Creating Meaningful Brands: The Relationship between Brand Name and Brand Mark [Publisher: Springer]. *Marketing Letters*, *14*(3), 143–157. Retrieved August 27, 2020, from https://www.jstor.org/stable/40216495

Knoeferle, K., Li, J., Maggioni, E., & Spence, C. (2017). What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings [Number: 1 Publisher: Nature Publishing Group]. *Scientific Reports*, *7*(1), 5562. https://doi.org/10.1038/s41598-017-05965-y

Köhler, W. (1929). *Gestalt psychology* [Pages: xi, 403]. Liveright.

Lankoski, S. (2013). R / Ordinal Scrips. https://petrilankoski.com/2013/01/05/r-ordinal-scrips-part-2/

Lapolla, R. (1994). An experimental investigation into phonetic symbolism as it relates to Mandarin Chinese. In *Sound Symbolism* (pp. 130–147). Cambridge University Press.

Lasserre, B. (2018). *Words that go ping: The ridiculously wonderful world of onomatopoeia / Barbara Lasserre ; with illustrations by Daniel Wlodarski*. Allen & Unwin.

Lawson, E. (1984). Personal names: 100 years of social science contributions. *32*, 45–73. https://doi.org/10.1179/nam.1984.32.1.102

Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, *79*, 328–348. https://doi.org/10.1016/j.jesp.2018.08.009

Maglio, S. J., & Feder, M. A. (2017). The Making of Social Experience from the Sounds in Names [Publisher: Guilford Publications Inc.]. *Social Cognition*, *35*(6), 663–674. https://doi.org/10.1521/soco.2017.35.6.663

Mangiafico, S. (2016). R Handbook: Two-way Repeated Ordinal ANOVA with CLMM. https://rcompanion.org/handbook/G_12.html

Margiotoudi, K., & Pulvermüller, F. (2020). Action sound–shape congruencies explain sound symbolism. *Scientific Reports*, *10*(1), 1–13. https://doi.org/10.1038/s41598-020-69528-4

Mattys, S. L. (2000). The perception of primary and secondary stress in English. *Perception & Psychophysics*, *62*(2), 253–265. https://doi.org/10.3758/BF03205547

Miall, D. S. (2001). Sounds of contrast: An empirical approach to phonemic iconicity. *Poetics*, *29*(1), 55–70. https://doi.org/10.1016/S0304-422X(00)00025-5

Miller, R. M. (2014). *The emotional weight of poetic sound: An exploration of phonemic iconicity in the haiku of Bashō* (Master's thesis) [ISBN: 9781321263909]. Southern Illinois University at Carbondale. United States – Illinois. Retrieved August 26, 2020, from https://search.proquest.com/docview/1614428392/abstract/2B1C8CAB00DE4A19PQ/1

Motschenbacher, H. (2020a). Corpus Linguistic Onomastics: A Plea for a Corpus-Based Investigation of Names [Publisher: Routledge]. *Names*, *68*(2), 88–103. https://doi.org/10.1080/00277738.2020.1731240

Motschenbacher, H. (2020b). Corpus Linguistic Onomastics: A Plea for a Corpus-Based Investigation of Names [Number: 2]. *Names*, *68*(2), 88–103. https://doi.org/10.1080/00277738.2020.1731240

Mutsukawa, M. (2014). Phonological and Semantic Gender Differences in English and Japanese Given Names. *Els noms en la vida quotidiana. Actes del XXIV Congrés Internacional d'ICOS sobre Ciències Onomàstiques*, 370–377. https://doi.org/10.2436/15.8040.01.41

Nazar, R., Renau, I., Acosta, N., Robledo, H., Soliman, M., & Zamora, S. (2021). Corpus-Based Methods for Recognizing the Gender of Anthroponyms [Number: 3]. *Names*, *69*(3), 16–27. https://doi.org/10.5195/names.2021.2238

Ning, Y., Ho, P. J., Støer, N. C., Lim, K. K., Wee, H.-L., Hartman, M., Reilly, M., & Tan, C. S. (2021). A New Procedure to Assess When Estimates from the Cumulative Link Model Can Be Interpreted as Differences for Ordinal Scales in Quality of Life Studies. *Clinical Epidemiology*, *13*, 53–65. https://doi.org/10.2147/CLEP.S288801

Nuckolls, J. B. (1999). The Case for Sound Symbolism [Publisher: Annual Reviews]. *Annual Review of Anthropology*, *28*, 225–252. Retrieved August 27, 2020, from https://www.jstor.org/stable/223394

Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, *114*(2), 173–186. https://doi.org/10.1016/j.jecp.2012.05.004

Parada, M. (2016). Ethnolinguistic and Gender Aspects of Latino Naming in Chicago: Exploring Regional Variation [Publisher: Routledge]. *Names*, *64*(1), 19–35. https://doi.org/10.1080/00277738.2016.1118858

Parault, S. J. (2002). *Sound symbolism : A possible piece in the puzzle of word learning.*

Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a General Property of Language: Evidence from Spoken and Signed Languages [Publisher: Frontiers]. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00227

Phonology of gender in English and French given names. (2018). https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/papers/ICPhS_2173.pdf

Pitcher, B. J., Mesoudi, A., & McElligott, A. G. (2013a). Sex-Biased Sound Symbolism in English-Language First Names [Publisher: Public Library of Science]. *PLOS ONE*, *8*(6), e64825. https://doi.org/10.1371/journal.pone.0064825

Pitcher, B. J., Mesoudi, A., & McElligott, A. G. (2013b). Sex-Biased Sound Symbolism in English-Language First Names [Publisher: Public Library of Science]. *PLOS ONE*, *8*(6), e64825. https://doi.org/10.1371/journal.pone.0064825

Sabet, P. G. P., & Zhang, G. (2020). First names in social and ethnic contexts: A socio-onomastic approach. *Language & Communication*, *70*, 1–12. https://doi.org/10.1016/j.langcom.2019.09.004

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*(3), 225–239. https://doi.org/10.1037/h0070931

Saussure, F. D. (1916). *Course in General Linguistics*.

Schmidtke, D., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity [Publisher: Frontiers]. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00080

Seguin, C., Julien, C., & Zhang, Y. (2021). The stability of androgynous names: Dynamics of gendered naming practices in the United States 1880–2016. *Poetics*, *85*, 101501. https://doi.org/10.1016/j.poetic.2020.101501

Sereno, J. A. (1994). Phonosyntactics. In *Sound Symbolism* (pp. 263–275). Cambridge University Press.

Shih, S., Ackerman, J., Hermalin, N., Inkelas, S., Jang, H., & Johnson, J. (2019). Cross-linguistic and language-specific sound symbolism: Pokémonastics. Retrieved October 7, 2020, from https://ling.auf.net/lingbuzz/004725

Shih, S., & Rudin, D. (2019). On sound symbolism in baseball player names. Retrieved December 2, 2020, from https://ling.auf.net/lingbuzz/004689

Shih, S. S. (2012). Linguistic determinants of English personal name choice, 22.

Shih, S. S. (2017). Phonological Influences in Syntactic Alternations. In V. Gribanova & S. S. Shih (Eds.), *The Morphosyntax-Phonology Connection* (pp. 223–252). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190210304.003.0009

Shih, S. S. (2020). Gradient categories in lexically-conditioned phonology: An example from sound symbolism [Number: 0]. *Proceedings of the Annual Meetings on Phonology*, *8*(0). https://doi.org/10.3765/amp.v8i0.4689

Shih, S. S., Ackerman, J., Hermalin, N., Inkelas, S., & Kavitskaya, D. (2018). Pokémonikers: A study of sound symbolism and Pokémon names [Number: 1]. *Proceedings of the Linguistic Society of America*, *3*(uno), 42–1–6. https://doi.org/10.3765/plsa.v3i1.4335

Sidhu, D. M., & Pexman, P. M. (2015). What's in a Name? Sound Symbolism and Gender in First Names [Publisher: Public Library of Science]. *PLOS ONE*, *10*(5), e0126809. https://doi.org/10.1371/journal.pone.0126809

Sidhu, D. M., & Pexman, P. M. (2019). The Sound Symbolism of Names [Publisher: SAGE Publications Inc]. *Current Directions in Psychological Science*, *28*(4), 398–402. https://doi.org/10.1177/0963721419850134

Sidhu, D. M., Pexman, P. M., & Saint-Aubin, J. (2019). Is un stylo sharper than une épée? Investigating the interaction of sound symbolism and grammatical gender in English and French speakers [Publisher: Public Library of Science]. *PLoS ONE*, *14*(12), 1–19. https://doi.org/10.1371/journal.pone.0225623

Slater, A. S., & Feinman, S. (1985a). Gender and the phonology of north American first names. *Sex Roles*, *13*(7-8), 429–440. https://doi.org/10.1007/BF00287953

Slater, A. S., & Feinman, S. (1985b). Gender and the phonology of north American first names. *Sex Roles*, *13*(7), 429–440. https://doi.org/10.1007/BF00287953

Starr, R. L. (2018). Sound symbolic effects in Mandarin and Cantonese personal names and Pokemon names. *Keio University*, 82.

Sullivan, L. (n.d.). Universal and cross-linguistic gender-based sound symbolism in Korean given names, 77.

Taylor, J. E., Rousselet, G. A., Scheepers, C., & Sereno, S. C. (2023). Rating norms should be calculated from cumulative link mixed effects models. *Behavior Research Methods*, *55*(5), 2175–2196. https://doi.org/10.3758/s13428-022-01814-7

Tutz, G. (2022). Ordinal regression: A review and a taxonomy of models. *WIREs Computational Statistics*, *14*(2), e1545. https://doi.org/10.1002/wics.1545

Uno, R. (2020a). What's in a villain's name? [Publisher: John Benjamins Publishing Company]. Retrieved October 6, 2020, from https://benjamins.com/catalog/rcl.00066.uno

Uno, R. (2020b). What's in a villain's name? [Publisher: John Benjamins Publishing Company]. Retrieved October 6, 2020, from https://benjamins.com/catalog/rcl.00066.uno

Vargas, V. M., Gutiérrez, P. A., & Hervás-Martínez, C. (2020). Cumulative link models for deep ordinal classification. *Neurocomputing*, *401*, 48–58. https://doi.org/10.1016/j.neucom.2020.03.034

Veríssimo, J. (2021). BLC mini-series: New statistical approaches and research practices for bilingualism research [Publisher: Cambridge University Press]. *Bilingualism: Language and Cognition*, *24*(5), 799–800. https://doi.org/10.1017/S1366728921000365

Whissell, C. (1999). Phonosymbolism and the Emotional Nature of Sounds: Evidence of the Preferential Use of Particular Phonemes in Texts of Differing Emotional Tone

[Publisher: SAGE Publications Inc]. *Perceptual and Motor Skills*, *89*(1), 19–48. https://doi.org/10.2466/pms.1999.89.1.19

Whissell, C. (2001a). Sound and Emotion in Given Names. *Names*, *49*(2), 97–120. https://doi.org/10.1179/nam.2001.49.2.97

Whissell, C. (2001b). Sound and Emotion in Given Names: Names: Vol 49, No 2. https://www.tandfonline.com/doi/abs/10.1179/nam.2001.49.2.97

Williams, R. (2016). Understanding and interpreting generalized ordered logit models. *The Journal of Mathematical Sociology*, *40*(1), 7–20. https://doi.org/10.1080/0022250X.2015.1112384

Williams, S. J., & Renwick, M. (2022). Modeling language change in English first names [Number: 1]. *Proceedings of the Linguistic Society of America*, *7*(1), 5243–5243. https://doi.org/10.3765/plsa.v7i1.5243

Wright, S. K. (2006). Phonological Cues Influence Sex Decisions about Novel Names [Publisher: SAGE Publications Inc]. *Psychological Reports*, *99*(2), 315–321. https://doi.org/10.2466/pr0.99.2.315-321