Phonology of Gender in English and French Given Names

By

Lisa Sullivan

Supervisor: Yoonjung Kang
Second Reader: Nathan Sanders

A forum paper submitted in conformity with the requirements

for the degree of Master of Arts

Graduate Department of Linguistics

University of Toronto

Abstract

Sound symbolism, the idea that meaning can sometimes be conveyed at a sub-morphemic level, challenges the notion that the form-meaning relation is completely arbitrary. One possible type of meaning that may be conveyed at the phonological level is gender. Corpus analyses of English given names (Cutler et al., 1990; MacAuley et al., in prep; Sidhu & Pexman, 2015; Slater & Feinman, 1985; Wright et al., 2005) have shown that there are phonological patterns which are correlated with gender. If these patterns are sound symbolic, it would be expected that they would be cross-linguistic and that speakers would be able to make use of them subconsciously.

The current study explores the possibility of sound symbolic meaning for gender through a corpus analysis of French and English given names and a name gendering task. The corpus analysis shows similarities in the phonological factor-gender correlation between French and English, although the significance of individual factors varies between the languages. This allows for the possibility that these correlations may be cross-linguistic.

The name gendering task shows that people rate phonologically female-biased names as more female than phonologically male-biased names. Furthermore, participants are even able to do this for names presented in a language other than their native one. This suggests that people have internalized and actively make use of these patterns, and that they have done so at an abstract level which allows them to apply them to other languages. It is not, however, clear how these patterns are acquired.

Future research should extend both the corpus analysis and the experiment to other, genetically unrelated, languages. This would provide more insight into the possibility that the patterns are cross-linguistic and address the possibility that there are sound symbolic meanings associated with gender.

Acknowledgements

I'd like to thank Yoonjung Kang, my supervisor, for her support, encouragement, guidance and feedback. I'd also like to thank Nathan Sanders for being my second reader. I'd like to thank Sali Tagliamonte, the graduate coordinator, and my forum cohort (Koorosh Ariyaee, Tim Gadanidis, Isabelle Ladouceur-Séguin, Mia Misic, Megan Parker, Ryan Pidhayny, Matthew Riopelle, Lisa Schlegl, Rachel Soo and Connie Ting) for providing feedback and support, both inside and outside forum meetings. I'd also like to thank my professors (including Lev Blumenfeld, Marie-Odile Junker, Beth MacLeod, Dan Siddiqi, Raj Singh and Ida Toivonen) from my undergraduate degree at Carleton University for fostering my interest in linguistics and encouraging me to apply to and pursue the M.A. program at the University of Toronto.

Data for my corpus analysis came from publicly available databases provided by the governments of Ontario and Québec. I'd like to thank Isabelle Ladouceur-Séguin for providing a recording of native speaker pronunciations of French names for transcription for the corpus analysis.

Funding for the experiment was provided by a SIG grant. The experiment would not have been possible without the help of numerous people. I'd like to thank Isabelle Ladouceur-Séguin for editing my written French materials and recording the French stimuli for my experiment, Megan Parker for recording my English stimuli and Connie Ting for helping me get set up in the phonetics lab to do my recordings for my experiment. I'd like to thank Hyoung Seok Kwon and Na-Young Ryu for the use of their online scripts and for helping me get started with them. I'd also like to thank Colleen Dulle, Anthony Fredette, Brian Sullivan and Simon Whedbee for helping me test various aspects of my experiment as well as the various friends, family and acquaintances who helped me recruit participants for the study.

Finally, I'd like to thank my family (especially my parents, Brian and Cheryl Sullivan and my siblings, Brianna, Alison, Patrick and Philip) and friends for their support and encouragement.

Table of Contents

1. Introduction

The view that the relationship between form and meaning in language is arbitrary, often associated with Saussure (1916), is widely held in linguistics. This is not, however, always the case. Onomatopoeic words (e.g. animal sounds meow, moo, oink etc.), for example, sound like what they are describing while phonosthemes are sound clusters which are associated with a particular meaning (e.g. /fl-/ and /gl-/ are associated with light in English) (Svantesson, 2017). Furthermore, certain sounds and natural classes of sounds have been claimed to symbolize different sizes and shapes (Svantesson, 2017). Such sound symbolism challenges the notion that the relationship between sound and meaning is always completely arbitrary. If certain sounds are associated with particular meanings, it suggests that sounds are not meaningless units.

One type of meaning sounds may symbolize is gender. Corpus analyses of English given names have found phonological patterns correlated with gender in these names (Cutler et al., 1990; MacAuley et al., in prep; Sidhu & Pexman, 2015; Slater & Feinman, 1985; Wright et al., 2005). It is possible that these patterns may be sound symbolic. Consider, for example, the types of vowels found in male and female names. The corpus analyses indicate that female names are more likely to have high and/or front vowels than male names (Wright et al., 2005). Size symbolism literature (Ohala, 1984; 2004) has argued that higher and more front vowels are associated with smaller sizes whereas lower and more back vowels are associated with larger sizes. Klink (2014) has argued that the high-front vowels are associated with femininity due to their higher acoustic frequency. Interestingly, English female given names are more likely to have high-front vowels than male ones (Cutler et al., 1990; Wright et al., 2005), suggesting that this pattern may in fact be sound symbolic. If this is the case, the phonology of given names may give clues to sound symbolic meanings for gender.

If the phonological patterns found in given names are, in fact, sound symbolic, people might be expected to have internalized these patterns and to actively make use of them when they encounter unfamiliar names. Experimental work (Cassidy et al., 1999; MacAuley et al., in prep; Sidhu & Pexman, 2015) which had participants link nonce and real given names with gender found that this appears to be the case, suggesting that these patterns are meaningful and not just arbitrary.

However, one shortcoming of previous research is that it only focused on English names, thus it is not clear if these patterns are English-specific or cross-linguistic. My research proposes to investigate the patterns found in English and French given names with three goals: (1) to look for cross-linguistic patterns by looking for phonological patterns in names from two languages, (2) to verify the trends found in previous studies in a new set of English names, and

(3) to test if people internalize and actively make use of these patterns. I begin with a discussion of research on sound symbolism (Section 2) and on the relationship between the phonology and gender of given names (Section 3). Next, Section 4 discusses a corpus analysis investigating the phonological patterns of a new set of English and French names. Section 5 describes an experiment based on the results of the corpus analysis in which participants were asked to rate how male or female they thought a set of nonce given names were. Finally, I conclude with a summary of the results, limitations of the study and directions for future research in Section 6.

2. Sound Symbolism

Sound symbolism occurs when individual sounds or sub-morphemic clusters of sounds convey meaning (Svantesson, 2017). This suggest that meaning can sometimes be found at the level of phonology and challenges the notion that morphemes are necessarily the smallest meaningful units in language. Two types of sound symbolism which focus on individual sounds are size symbolism and shape symbolism.

Size symbolism refers to the apparent relationship between vowel quality and size (Ohala 1984; 2004; Svantesson, 2017; Tanz, 1971). Lower and more back vowels are said to be associated with larger sizes whereas higher and more front vowels with smaller sizes. For example, Tanz (1971) did a cross-linguistic survey of words for proximal distance (‘here’ and ‘there’). She found that, across languages, where words for ‘here’ (smaller proximal distance) where similar to those for ‘there’ (larger proximal distance), the vowels tended to change such that a high front vowel became low and/or back or a low and/or back vowel was added on its own or as part of an additional syllable.

Shape symbolism refers to the apparent relationship between vowel or consonant quality and shape (Bremner et al., 2013; Mauer et al., 2006; Nielsen & Rendall, 2011; Svantesson, 2017; Westbury, 2005). Certain sounds are thought to be associated with round shapes while others are associated with sharp shapes. While the sounds proposed vary from study to study, round sounds generally include /b/, /l/, /m/, /n/, /a/ and /u/ while sharp sounds generally include /p/, /t/, /k/ and /i/. These patterns have been found to occur in both English speaking adults (Mauer et al., 2006; Nielsen & Rendall, 2011; Westbury, 2005) and children (Mauer et al., 2006), as well as in other languages (Bremner et al., 2013). While both vowels and consonants have been proposed to represent shape symbolism, Nielsen and Rendall (2011) evaluated the relative effects of the two types of sounds on shape symbolism and found that people associated the consonants, but not the vowels with the expected shapes. This suggests that the consonants are shape symbolic, but that the vowels may not be.

These kinds of sound symbolism have been applied in research focusing on the phonology of brand names (Klink, 2014; Shrum et al., 2012), chemotherapy drug names (Abel & Glinert, 2008), Pokémon names (Kawahara et al., 2015) and English given names (Sidhu & Pexman, 2015). Abel and Glinert (2008) conducted a corpus analysis of chemotherapy drug names which suggested that chemotherapy drugs contain voiceless stops, which they claim, following Newman (1933), are associated with lightness and fastness, at a rate significantly higher than general English vocabulary. Furthermore, while not significant, the chemotherapy drug names contain voiced consonants, which Newman (1933) associates with slowness and heaviness, at a lower rate than general English vocabulary. Kawahara et al.'s (2015) corpus analysis of Japanese Pokémon names found correlations between the phonology of the names and the Pokémon characters' size, weight, evolution levels and strength parameters. These studies suggest that people are able to make use of sound symbolic meanings in the creation of new words.

Klink (2014) and Shrum et al. (2012) conducted studies in which participants were presented with two nonce brand names which varied only in one segment type and were asked to choose which one they felt best described a particular attribute or product. Klink's (2014) study asked English participants which of two nonce names better described a particular attribute which was said to be sound symbolic. The results of the study demonstrated that participants generally selected words which would be expected to better describe the attributes according to sound symbolism. Shrum et al. (2012) asked English, Spanish, French and Chinese speakers which of two nonce names, which differed sound symbolically, better described a particular product. They found that participants displayed sound symbolic preferences, and that these preferences persisted across languages, suggesting that the sound symbolism may be universal. These studies suggest that people use sound symbolism when processing unfamiliar words.

Sidhu & Pexman (2015) tested the effects of sound symbolism with regards to sharp and round sounds using real and nonce English given names. They conducted a series of experiments in which participants paired names with shapes with various round or jagged contours. Names were categorized based on the presence of "sharp" and "round" sounding phonemes. Round phonemes included /b/, /l/, /m/, /n/, /u/, /o/ and /ɒ/. Sharp phonemes included /k/, /p/, /t/, /i/, /e/, /ɛ/, and /ʌ/. The results indicated that there may be a sound symbolic association between the names and the shape. Furthermore, participants were more likely to assign a female name to a round shape. An analysis of the names showed that more round phonemes were found in female names than male names. This suggests that people are able to use sound symbolism in processing both familiar and unfamiliar words.

Sidhu and Pexman's (2015) results showing a correlation between gender and shape, as well as a higher presence of round phonemes in female names, suggest that there may be sound symbolic correlations between phonological properties and gender. Klink (2014) also proposed a relationship between gender and sound whereby high-front vowels are associated with femininity because of their higher acoustic frequency. He found that, participants more strongly associated words with high-front vowels with femininity than words with other vowels. It seems possible, therefore, that there are sound symbolic associations with gender.

3. Name Gender Phonology

Research on the relationship between the phonology of names and their gender suggests both that there exist phonological patterns differentiating gender in names and that people are actively able to make use of these patterns. Several corpus analyses of English given names have been conducted which suggest the presence of contrastive phonological patterns in male and female names. These patterns are related to length, suprasegmental features, such as stress and syllable structure, type of initial sound, type of final sound, vowel type and consonant type.

In terms of length, Cutler et al. (1990), Slater & Feinman (1985) and Wright et al. (2005) found that female names contain both more syllables and more segments than male names. For example, *Jack* /dʒæk/ has one syllable but *Julia* /dʒu.li.æ/ has three. Furthermore, Wright et al. (2005) note that most names in their corpus were mono or bisyllabic, however, when comparing male and female names, the female ones were more likely to be multisyllabic.

In terms of suprasegmental features, stress and syllable structure patterns have been examined in English names. Examining stress patterns, Cutler et al. (1990) and Slater & Feinman (1985) found that, both male and female names are like nouns in that they are more likely to have initial stress than non-initial stress (e.g. *Brandon* /ˈbɹæn.dən/ and *Cara* /ˈkɛ.ɹə/). However, female names were more likely than male names to have non-initial stress (e.g. *Nathan* /ˈne.θən/ has initial stress but *Samantha* /sə. ˈmæn.θə/ has non-initial stress). Looking at syllable structure, Slater & Feinman (1985) found that female names had a higher ratio of open syllables than male names. For example, *Spencer* /spɛn.səɹ/ has two closed syllables but *Emma* /ɛ.mə/ has two open syllables.

Wright et al. (2005) examined the types of sounds that began English names. Comparing consonants and vowels, they found that female names were more likely to start with a vowel than male names (e.g. *Bruce* /bɹus/ vs. *Abigail* /æbəgel/). Examining different types of consonants, they found that male names were more likely to begin with an oral stop than female names (e.g. *Thomas* /tɑməs/ vs *Mary* /mɛri/). Male names were also more likely to begin with a voiced obstruent (stop, fricative or affricate) than female names (e.g. *Brian* /bɹiən/ vs *Theresa* /təɹisə/).

Slater & Feinman (1985) and Wright et al. (2005) examined the types of sounds that end English names. Comparing consonants and vowels, both studies found that female names are more likely to end with a vowel than male names (e.g. *Simon* /sajmən/ vs *Brianna* /bɹiænə/). Of the vowel segments, female names were more likely to end with a /ə/ compared to male names (e.g. *Andrew* /ændʒɹu/ vs *Maria* /mæɹiə/). Of the consonant segments, female names are more likely to end with a sonorant than male names (e.g. *Nicholas* /nɪkələs/ vs *Evelyn* /ɛvələn/). Female names are also more likely to end in a voiced consonant than male names (e.g. *Marcus* /mæɹkəs/ vs *Maeve* /mev/).

Looking at vowel segments, Cutler et al. (1990) found that female names were more likely to contain [i] and less likely to contain [ʊ], [ʌ] or [ɔ] than male names (e.g. *John* /dʒɑn/ vs *Eve* /iv/). Wright et al. (2005) similarly captured this when they found that female names were more likely than male names to contain high and/or front vowels. Note that [i] is both high and front while [ʊ], [ʌ] and [ɔ] are all back vowels, some of which are low. Wright et al. (2005) also found that monosyllabic female names were more likely to contain a long vowel or diphthong than monosyllabic male names.

In their exploration of the relationship between gender and shape symbolism, Sidhu and Pexman (2015) compared the occurrence of round and sharp consonants in a corpus of English names. Following Mauer et al. (2006), they selected /l/, /m/ and /n/ as round consonants and /p/, /t/ and /k/ as sharp consonants. Their results indicated that female names were more likely to contain round consonants than male names. For example, *Belinda* /bəlɪndə/ contains two round consonants ([l] and [n]) whereas *Cody* /kodi/ contains none. However, sharp consonants were not found to occur at a significantly higher rate in male or female names, although they were trending in the direction of occurring more in male names. MacAuley et al. (in prep) found that male names are more likely to end in a syllabic nasal than female names.

One shortcoming of these corpus analyses is that they only look at English names. Thus, it is not clear whether the observed trends are cross-linguistic or language-specific. Cutler et al. (1990) have, however, suggested similar trends in French: that most monosyllabic names are male and that male names are more likely to begin with a closed syllable than female names. The first corresponds with the observation that female names are longer than male names. The second corresponds with the higher open syllable ratio in female names and, less directly with the observation that female names are more likely to have non-initial stress. That is, male names are more likely to have more weight on the initial syllable than female names, either by virtue of stress or of having two moras. Extending the corpus analyses to include multiple languages would address this issue.

While these corpus analyses show that there exist phonological patterns in given names which are contrastive in terms of gender, they do not reveal if these are patterns that people have internalized and actively make use of. Sutton (2016) conducted a corpus analysis of super hero names which showed that these constructed names conform to the phonological patterns of American English names in terms of gender. This suggests that these patterns may be psychologically real and that people are able to use them. Furthermore, Cassidy et al. (1999) and MacAuley et al. (in prep) have conducted experiments to test these patterns and found that speakers appear to be able to make use of them to assign gender to names.

MacAuley et al. (in prep) examines a trend people have noticed of an increasing number of boys' names ending in syllabic nasal in North America. They conducted a corpus analysis of the top 200 boys' and girls' names in Ontario and in the United States across time. The results from both countries show a trend in boys' names but not girls' names. They followed this up with a forced choice lexical selection task in which participants were presented with two nonce words and asked which one was a boys' name. The results indicated that participants preferred nonce words ending in syllabic nasals for boys' names, consistent with what they found in the corpus analysis.

Cassidy et al. (1999) conducted a series of experiments to examine if English speakers internalize phonological tendencies in male and female names using tendencies observed in previous studies. Two of these experiments involved nonce names. The first had adults and children hear a nonce name and decide if it was the name of a male or a female doll. Stimuli were designed such that they consisted of pairs of which one member was a female variant and the other male. These pairs varied in terms of one of three phonological patterns: disyllabic stress pattern (female – iambic or male – trochaic), number of syllables (male – one or female – three) and final sound type (male – consonant or female – vowel). Both children and adults were found to assign names to the doll of the correct gender at an above chance rate.

Cassidy et al.'s (1999) second nonce-name experiment had adults complete sentences with a nonce name using a pronoun in an attempt to get at people's intuitions without specifically asking them about gender. Participants were presented with a sentence fragment containing a nonce name and were asked to complete the sentence. Sentences were designed in such a way that completion required the use of a male or female pronoun (e.g. "After Cora went to bed…" (Cassidy et al, 1999, p. 366)). Nonce names used the same phonological patterns as in the first study. Similar to the first study, participants assigned the correct pronoun to each nonce name at an above chance rate. Furthermore, adults appear to make more use of stress and final sound type more than syllable structure in assigning gender to names.

Cassidy et al. (1999) ran a third experiment which had adults classify real names as male or female. Participants saw a real name and were asked to indicate whether it was male or female. Stimuli consisted of real names which were either typical or atypical for their gender based on a connectionist model Cassidy et al. (1999) built of people's use of naming patterns in gendering names. Participants were able to assign the correct gender to names more often when they were phonologically typical than when they were phonologically atypical.

Cassidy et al. (1999) propose that their results reflect learned phonological naming patterns. However, it is not clear how these patterns are acquired. As Cassidy et al. (1999) only tested English patterns on English speakers, it is not possible to discern if these patterns are learned from English or by some other means. Their first experiment, which looked at both children and adults, shows that children and adults behave similarly in assigning names to gender. This suggests that children have acquired this knowledge at a young age, and allows for the possibility that acquisition was not strictly a result of their native language. Testing how speakers acquire these patterns would require looking at non-native name-gender patterns.

In order to address the issue of whether these patterns are cross-linguistic or language-specific I will conduct a corpus analysis of both French and English given names. This will allow me to compare the phonological patterns found in the two languages. If the sounds found in male and female names display distinct patterns, and these patterns are found across multiple languages, it could be suggestive of sound symbolism, challenging the notion that the form-meaning relationship is completely arbitrary (de Saussure, 1916). In my corpus analysis, I also hope to identify patterns that are language-specific in the sense that they can be found in French and not English or English and not French due to the phonology of those languages. For example, stress placement can't be used in French because French doesn't have lexical stress. I will then be able to use these patterns in a name gendering task to test if speakers of both languages are able to actively make use of them to assign gender to nonce names. This would address the issue of the method of acquisition of these patterns because, if French speakers are able to make use of English patterns, or English speakers of French patterns, this would suggest the patterns are acquired by some other means than exposure to the language, because speakers of the other language would not have been able to learn the patterns by this method.

4. Corpus Analysis

The corpus analysis investigates phonological patterns of French and English names with two goals: (1) to look for cross-linguistic patterns by looking for phonological patterns in names from two languages, and (2) to verify the trends found in previous studies in a new set of English names. Similarities between the English results of this study and previous results may indicate language-specific phonological patterns in male and female names. Similarities between the English and French results of this study may suggest cross-linguistic phonological tendencies. Such results would add to the body of knowledge on sound symbolism (Kawahara et al., 2015; Kawahara et al., 2017; Khöler, 1929, Maurer et al., 2006; Ramachandran & Hubbard, 2001; Svantesson, 2017).

4.1. Data Collection

Names for the corpus analysis were obtained from the databases of baby names compiled by ServiceOntario (2016a,b) and Retraite Québec (2017). These databases provide a record of all the first names registered in Ontario or Quebec in a given year along with the number of babies given that name in that year. The Ontario database provides records from 1917 through 2013. It excluded names with less than five records for privacy reasons. These names were used for the English names in the analysis. The Quebec database provides records from 2012 to 2017. These names were used as stand-ins for French names for the purposes of this analysis. The records from the year 2013 were selected for this study as they were the most recent year which was available in both databases. Names are recorded orthographically and, so, do not include pronunciations.

The 138 most frequent names for each gender from the Ontario database and 200 most frequent names for each gender from the Quebec database were phonetically transcribed. The goal for each language was to transcribe 100 names for each gender in each language. Additional names were transcribed in each language to account for the removal of alternate spellings of the same name and of non-English or non-French names. For the Ontario database, names were transcribed using the author's native speaker intuitions. For the Quebec names, a native female speaker of Quebec French in her twenties was recorded saying the names and her pronunciations were phonetically transcribed. English transcriptions included segmental information, syllable boundaries and stress placement. French transcriptions included segmental information, vowel nasalization and syllable boundaries. If multiple spellings of the same pronunciation occurred, the additional spellings were excluded from the analysis.

Each name was also coded for language of origin using the Behind the Name website (Campbell, 2013). A name was coded as English or French if it, or a spelling variant, was listed as being of English or French origin, otherwise it was coded as one of its languages of origin. Names on the Ontario list which were not English, as well as those on the Quebec list which were not French, were excluded from the analysis. This left 116 English girls' names, 122 English boys' names, 107 French girls' names and 92 French boys' names for analysis. A full list of names and transcriptions can be found in Appendix A.

4.2. Coding Schema

The names were coded for eleven factors in each language, ten of which were the same across both languages and one of which was different for English and French. The ten factors that were coded for both languages were: number of syllables, open syllable proportion, initial sound type, initial syllable type, final syllable type, high vowel proportion, low vowel proportion, back vowel proportion, round consonant proportion and sharp consonant proportion. The English-specific factor was stress placement. These eleven factors are meant to capture the trends observed in previous studies (Cutler et al., 1990; Slater & Feinman, 1985; Sidhu & Pexman, 2015; Wright et al., 2005). The French-specific factor was nasal vowel proportion. A summary of the factors, with their criteria and levels, can be found in Table 1.

Number of syllables is a continuous variable which is the number of syllables in the name. Open syllable proportion, high vowel proportion, low vowel proportion, back vowel proportion, round consonant proportion and sharp consonant proportion are continuous variables with a range of 0 to 1. Each is calculated by dividing the number of target syllables, vowels or consonants in the word by the total number of syllables, vowels or consonants in the name. For example, high vowel percentage is calculated by dividing the number of high vowels in a name by the total number of vowels in that name. These variables were coded as percentages rather than the counts used in previous research to avoid confounds with name length. Initial sound type and final sound type are categorical variables which are coded as "C" if the initial/final sound is a consonant and "V" if it is a vowel. Initial syllable type and final syllable type are categorical variables which are coded as "O" if the initial/final syllable is open and "C" if it is closed. Final sound type was not coded for as this is reflected in the coding for final syllable type since open syllables would end in a vowel and closed syllables, in a consonant. For French, syllables containing nasal vowels were coded as closed, following Paradis and Prunet (2000) who argue that these sequences are biphonemic and are underlyingly a vowel and nasal consonant sequence in which the nasal consonant is not realized, but its [+nasal] features spreads to be realized on the vowel. Stress placement is a categorical variable which is coded as 1, 2 or 3 depending on which syllable, from left to right, primary stress is on in the name. Nasal vowel percentage is a continuous variable calculated by dividing the number of nasal vowels in a name by the total number of vowels in the same name.

Table 1

Factors coded for in corpus analysis

Factor	**Criteria**	**Levels**	**e.g. Emma /ˈɛ.mə/**
Number of Syllables	Number of syllables in the name	Continuous (integers)	2
Open Syllable Proportion	Number of open syllables divided by total number of syllables	Continuous (range 0-1)	1
Initial Sound Type	Is the initial sound a consonant or a vowel?	C (consonant), V (vowel)	V
Initial Syllable Type	Is the initial syllable open or closed?	O (Open), C (Closed)	O
Final Syllable Type	Is the final syllable open or closed?	O (Open), C (Closed)	O
High Vowel Proportion	Number of high vowels divided by number of vowels	Continuous (range: 0-1)	0
Low Vowel Proportion	Number of low vowels divided by number of vowels	Continuous (range: 0-1)	0
Back Vowel Proportion	Number of back vowels divided by number of vowels	Continuous (range: 0-1)	0
Round Consonant Proportion	Number of /l/, /m/, /n/ in the name divided by total number of consonants	Continuous (range: 0-1)	1
Sharp Consonant Proportion	Number of /p/, /t/, /k/ divided by total number of consonants	Continuous (range: 0-1)	0
Stress Placement	Primary stress location from left to right	1, 2, 3	1
Nasal Vowel Proportion	Number of nasal vowels divided by total number of nasal vowels	Continuous (range: 0-1)	0

4.3. Hypotheses

In verifying that the trends found in previous studies exist in English, I expect the results of my English corpus analysis to be in line with those of previous studies (Cutler et al., 1990; Slater & Feinman, 1985; Sidhu & Pexman, 2015; Wright et al., 2005). Furthermore, if the observed patterns are cross-linguistic and not just langauge-specific, I expect that my French results should match my English ones. More specifically, regarding factors for which trends have been found in previous research, I expect to find that female names have more syllables (Cutler et al., 1990; Slater & Feinman, 1985; Wright et al., 2005), a higher proportion of open syllables (Slater & Feinman, 1985), a higher proportion of high vowels (Cutler et al., 1990; Wright et al., 2005), a lower proportion of low vowels (Cutler et al., 1990), a lower proportion of back vowels (Cutler et al., 1990; Wright et al., 2005), and a higher proportion of round consonants (Sidhu & Pexman, 2015) compared to male names, as well as being more likely to begin with a vowel (Wright et al., 1985) and to have non-initial stress (Cutler et al., 1990; Slater & Feinman, 1985).

Predictions can also be made regarding those factors which have not been studied or for which no trends have been found. Initial syllable type was not studied in previous research, however, Cutler et al. (1990) suggested that, in French, male names began with closed syllables more often than female names. If this is true, I expect to find that more male names begin with closed syllables than female names in my corpus. Likewise, final syllable type was not evaluated in previous studies, but it is essentially identical to final sound type (consonant or vowel) for English. Slater & Feinman (1985) and Wright et al. (2005) found that female names are more likely to end with a vowel than male names. Thus, I expect to find that female names are more likely than male names to end with an open syllable, which is synonymous to ending with a vowel in final sound type. Sidhu & Pexman (2015) found no significant difference between the presence of sharp consonants in male and female names in their corpus, although the trend for sharp consonants was that there were more sharp consonants in male, compared to female names. Based on this, I expect to find that male names have a higher proportion of sharp consonants than female names, although this trend may be non-significant. As nasal vowels are not found in English, their presence in names has not been looked at in previous studies. However, since nasal vowels often contrast with vowel + nasal consonant combinations in French in such a way that the nasal vowels are indicative of male nouns (e.g. *un gardien* /gaʀdiɛ̃/ 'guardian (male)' vs *une gardienne* /gaʀdiɛn/ 'guardian (female)') or adjectives modifying male nouns (e.g. *canadien* /kanadiɛ̃/ 'Canadian (male)' vs *canadienne* /kanadien/ 'Canadian (female)'), it might be expected that nasal vowels would be more common in male names than female names. As this is a property of French phonology, the trend, if found may be language specific, not universal.

4.4. Results

Each factor was described individually by gender and language. Univariate and multivariate analyses were conducted for each factor comparing the distribution of that factor in male and female names in each language. The results of both types of analysis were compared across languages.

Number of Syllables: Female names have more syllables on average than male names in both English (Table 2, Figure 1) and French (Table 3, Figure 2).

Table 2
Mean and standard deviation of number of syllables in English male and female names

	M	***SD***
Female	2.42	0.75
Male	2.10	0.57

Table 3
Mean and standard deviation of number of syllables in French male and female names

	M	***SD***
Female	2.60	0.66
Male	2.43	0.68

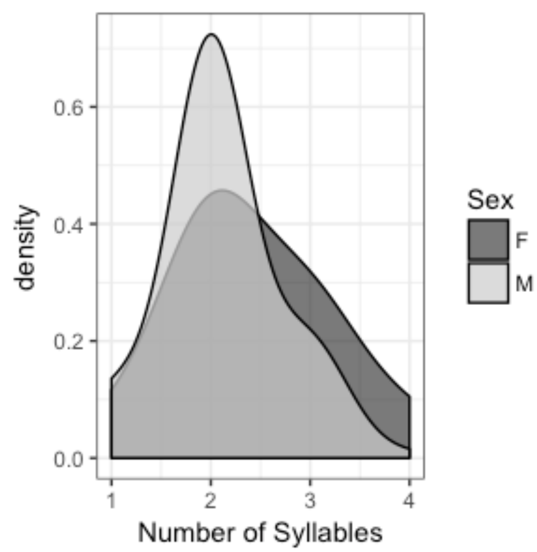

Figure 1 Density plot of number of syllables in English male and female names

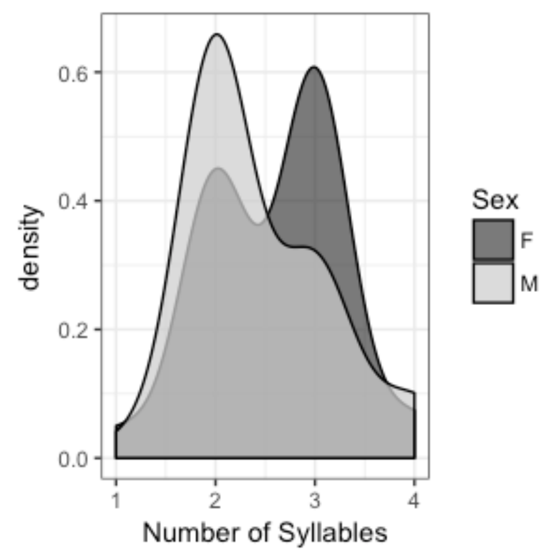

Figure 2 Density plot of number of syllables in French male and female names

Open Syllable Proportion: Female names have a higher proportion of open syllables than male names in both English (Table 4, Figure 3) and French (Table 5, Figure 4).

Table 4

Mean and standard deviation for open syllable proportion in English male and female names

	M	*SD*
Female	0.75	0.33
Male	0.44	0.30

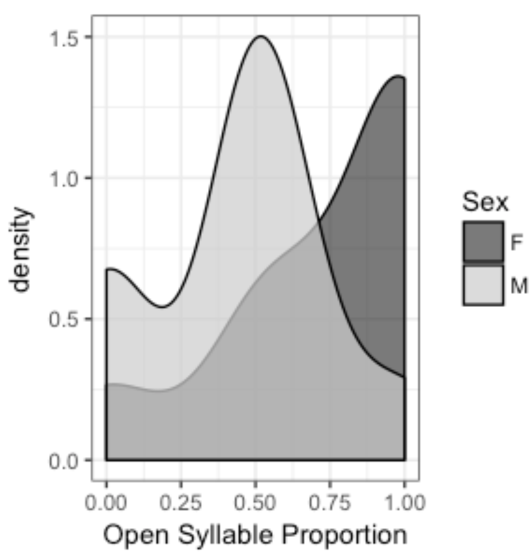

Figure 3 Density plot of open syllable proportion in English male and female names

Table 5

Mean and standard deviation for open syllable proportion in French male and female names

	M	*SD*
Female	0.66	0.30
Male	0.56	0.30

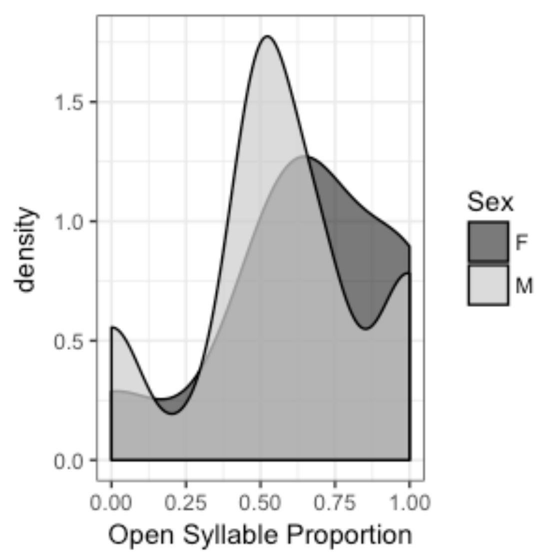

Figure 4 Density plot of open syllable proportion in French male and female names

Initial Sound Type: Female names began with a vowel more often than male name in both English (Table 6, Figure 5) and French (Table 7, Figure 6).

Table 6

Distribution of initial consonants and vowels in English male and female names

	Female		Male		Total	
	N	%	N	%	N	%
Consonant	83	0.72	97	0.80	180	0.76
Vowel	33	0.28	25	0.20	58	0.24
Total	116		122		238	

Table 7

Distribution of initial consonants and vowels in French male and female names

	Female		Male		Total	
	N	%	N	%	N	%
Consonant	67	0.63	72	0.78	139	0.70
Vowel	40	0.37	20	0.22	60	0.30
Total	107		92		199	

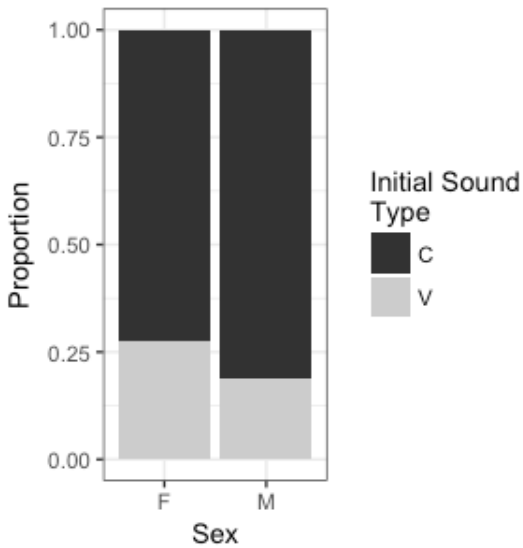

Figure 5 Bar chart of the distribution of initial consonants and vowels in English male and female names

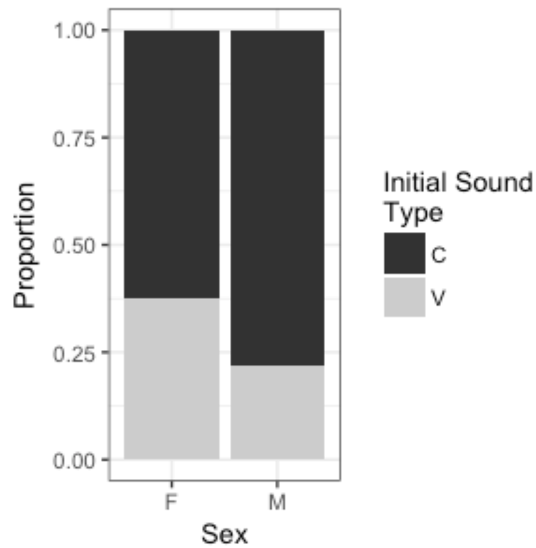

Figure 6 Bar chart of the distribution of initial consonants and vowels in French male and female names

Initial Syllable Type: Female names began with an open syllable more often than male names in both English (Table 8, Figure 7) and French (Table 9, Figure 8).

Table 8

Distribution of initial open and closed syllables in English male and female names

	Female		**Male**		**Total**	
	N	%	N	%	N	%
Closed	88	0.76	118	0.97	206	0.87
Open	28	0.24	4	0.03	32	0.13
Total	116		122		238	

Table 9

Distribution of initial open and closed syllables in French male and female names

	Female		**Male**		**Total**	
	N	%	N	%	N	%
Closed	15	0.14	19	0.21	34	0.17
Open	92	0.86	73	0.79	165	0.83
Total	107		92		199	

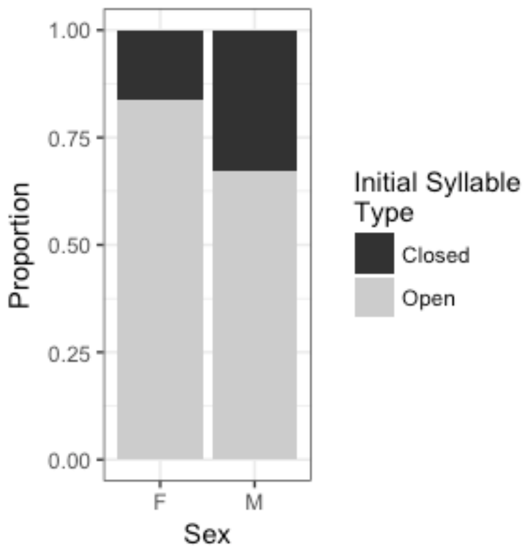

Figure 7 Bar chart of the distribution of initial open and closed syllables in English male and female names

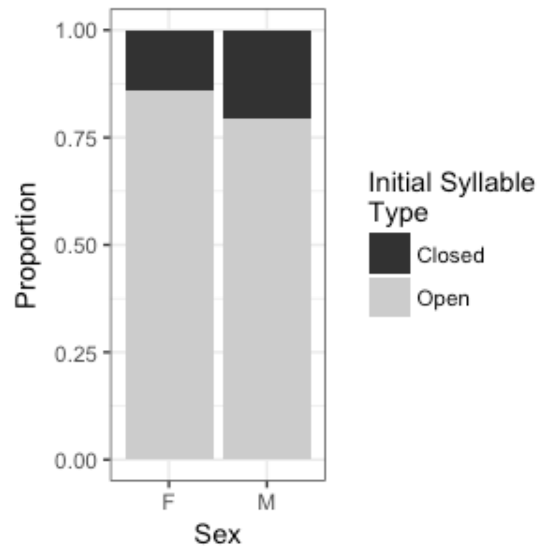

Figure 8 Bar chart of the distribution of initial open and closed syllables in French male and female names

Final Syllable Type: Female names end in an open syllable more often than male names in both English (Table 10, Figure 9) and French (Table 11, Figure 10).

Table 10

Distribution of final open and closed syllables in English male and female names

	Female		**Male**		**Total**	
	N	%	N	%	N	%
Closed	43	0.37	106	0.87	149	0.63
Open	73	0.63	16	0.13	89	0.37
Total	116		122		238	

Table 11

Distribution of final open and closed syllables in French male and female names

	Female		**Male**		**Total**	
	N	%	N	%	N	%
Closed	70	0.65	68	0.74	138	0.69
Open	37	0.35	24	0.26	61	0.31
Total	107		92		199	

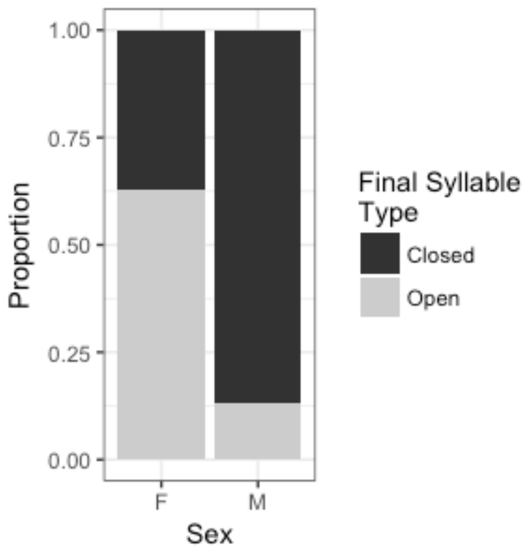

Figure 9 Bar chart of the distribution of final open and closed syllables in English male and female names

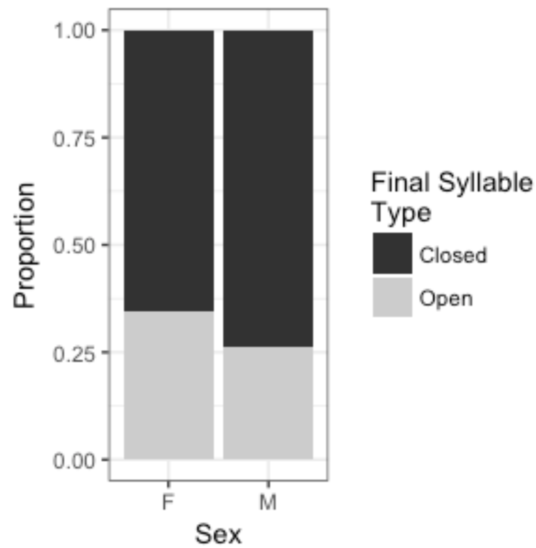

Figure 10 Bar chart of the distribution of final open and closed syllables in English male and female names

High Vowel Proportion: Female names had a higher proportion of high vowels than male names in English (Table 12, Figure 11), but not in French, where these numbers were very similar for both genders (Table 13, Figure 12).

Table 12

Mean and standard deviation for high vowel proportion in English male and female names

	M	***SD***
Female	0.25	0.29
Male	0.16	0.24

Table 13

Mean and standard deviation for high vowel proportion in French male and female names

	M	***SD***
Female	0.26	0.24
Male	0.28	0.29

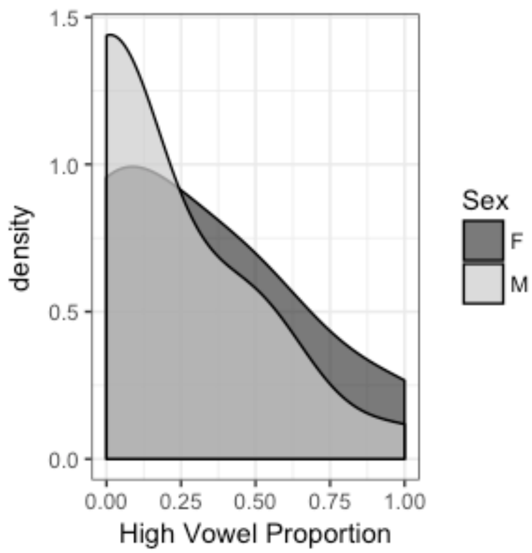

Figure 11 Density plot of high vowel proportion in English male and female names

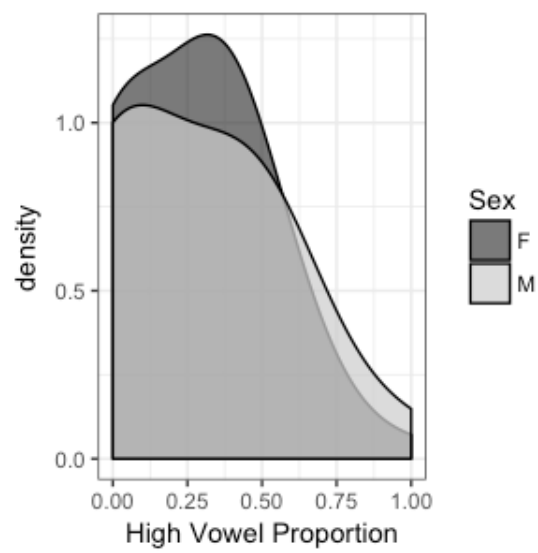

Figure 12 Density plot of high vowel proportion in French male and female names

Low Vowel Proportion: Female names had a lower proportion of low vowels than male names in English (Table 14, Figure 13), but not in French (Table 15, Figure 14), where these numbers were very similar for both genders.

Table 14

Mean and standard deviation for low vowel proportion in English male and female names

	M	***SD***
Female	0.15	0.21
Male	0.23	0.27

Table 15

Mean and standard deviation for low vowel proportion in French male and female names

	M	***SD***
Female	0.35	0.28
Male	0.34	0.30

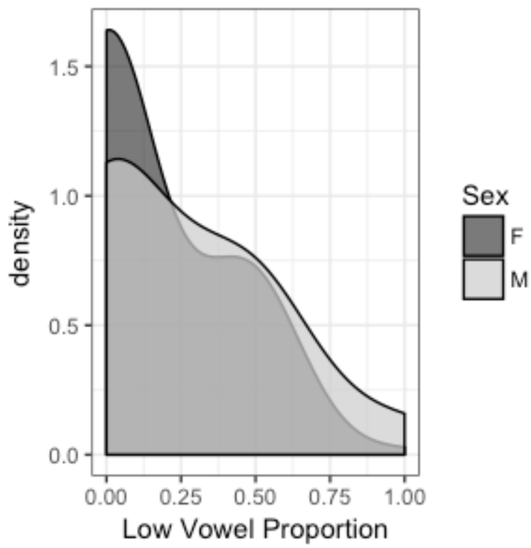

Figure 13 Density plot of low vowel proportion in English male and female names

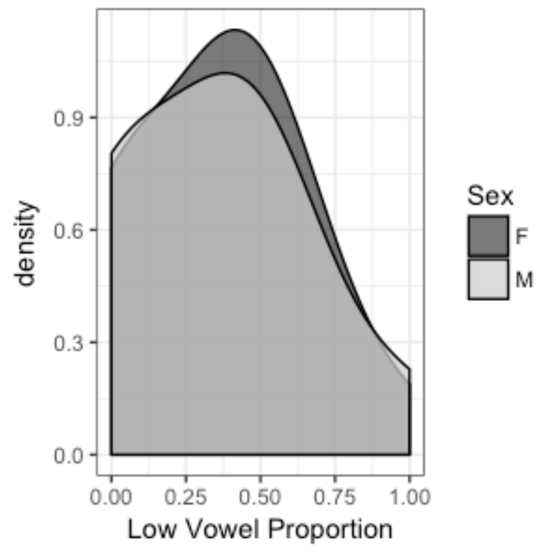

Figure 14 Density plot of low vowel proportion in French male and female names

Back Vowel Proportion: Female names had a lower proportion of back vowels than male names in both English (Table 16, Figure 15) and French (Table 17, Figure 16).

Table 16
Mean and standard deviation for back vowel proportion in English male and female names

	M	***SD***
Female	0.12	0.22
Male	0.17	0.27

Table 17
Mean and standard deviation for back vowel proportion in French male and female names

	M	***SD***
Female	0.17	0.25
Male	0.29	0.28

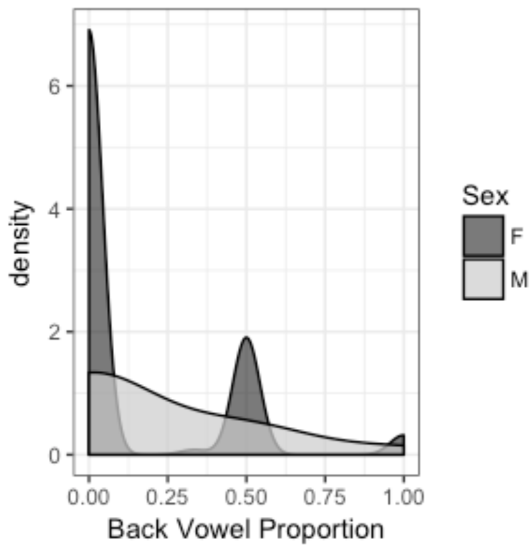

Figure 15 Density plot of back vowel proportion in English male and female names

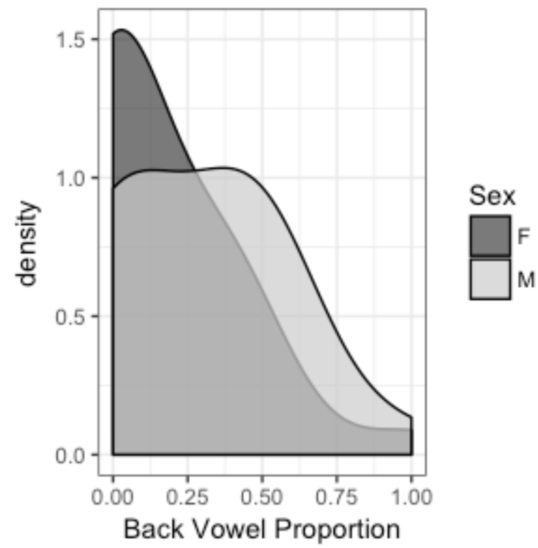

Figure 16 Density plot of back vowel proportion in French male and female names

Round Consonant Proportion: Female names had a higher proportion of round consonants than male names in both English (Table 18, Figure 17) and French (Table 19, Figure 18).

Table 18

Mean and standard deviation for round consonant proportion in English male and female names

	M	***SD***
Female	0.42	0.37
Male	0.34	0.26

Table 19

Mean and standard deviation for round consonant proportion in French male and female names

	M	***SD***
Female	0.52	0.31
Male	0.39	0.31

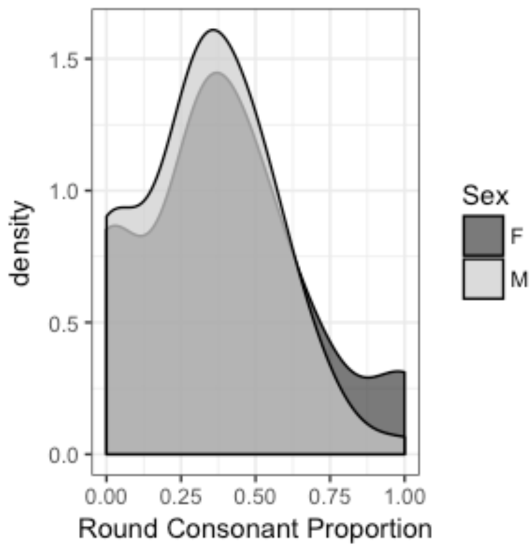

Figure 17 Density plot of round consonant proportion in English male and female names

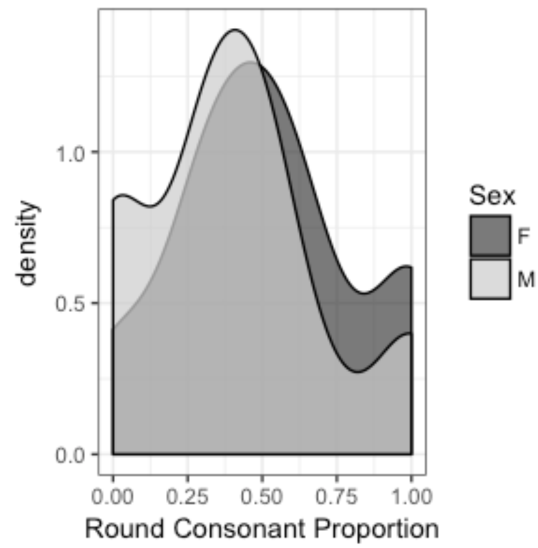

Figure 18 Density plot of round consonant proportion in French male and female names

Sharp Consonant Proportion: Female names had a lower proportion of sharp consonants than male names in both English (Table 20, Figure 19) and French (Table 21, Figure 20).

Table 20

Mean and standard deviation for sharp consonant proportion in English male and female names

	M	***SD***
Female	0.13	0.21
Male	0.17	0.20

Table 21

Mean and standard deviation for sharp consonant proportion in French male and female names

	M	***SD***
Female	0.08	0.15
Male	0.17	0.21

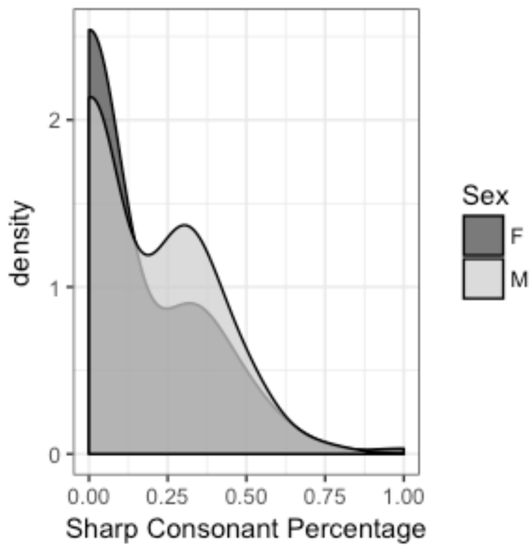

Figure 19 Density plot of sharp consonant proportion in English male and female names

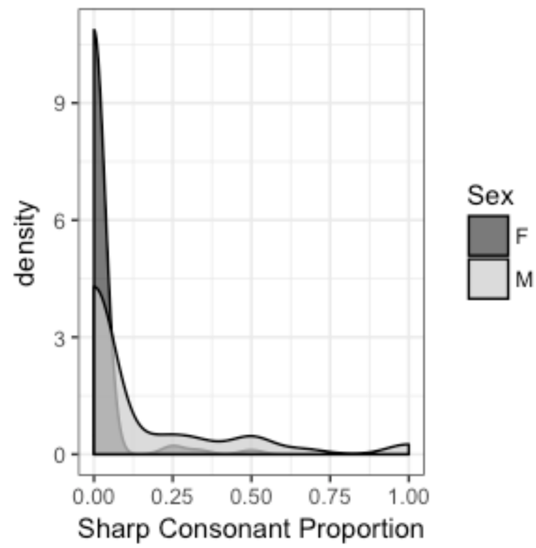

Figure 20 Density plot of sharp consonant proportion in English male and female names

Stress Placement (English Only): The English names in my corpus had stress on one of the first 3 syllables. Stress on the second and third syllable was more common in female names than male names, as can be seen in Table 21 and Figure 22. (Stress was not coded for in French, because the language lacks lexical stress.) In order to compare initial and non-initial stress, as was done in previous studies (Cutler et al., 1990; Slater & Feinman, 1985), and because of the low number of names with stress on the third syllable, stress on the second and third syllable was collapsed to form two categories: initial and non-initial. This also reduces the confound between stress placement and number of syllables, whereby the number of syllables in a name affects the possible locations of primary stress. A much higher proportion of female names (41%) have three or more syllables than male names (20%), meaning there are more female names which could have stress on the third syllable. Collapsing stress on the second and third syllable to non-initial stress eliminates the possible effect of this disparity. There is still a possible confound in that monosyllabic names must have primary stress, however such names only account for 9% of the overall data and 7% and 11% of female and male names, respectively, compared to 30% for names will three or more syllables. Thus, monosyllabic names will have less effect on the data overall. Stress occurred more often on non-initial syllables in female names compared to male names (Table 24).

Table 23

Distribution of primary stress placement in English male and female names

	Female		Male		Total	
	N	%	N	%	N	%
1	85	0.73	115	0.94	200	0.84
2	26	0.22	6	0.05	32	0.13
3	5	0.04	1	0.01	6	0.03
Total	116		122		238	

Table 24

Distribution of initial and non-initial primary stress placement in English male and female names

	Female		Male		Total	
	N	%	N	%	N	%
1	85	0.73	115	0.94	200	0.84
2+	31	0.27	7	0.06	38	0.16
Total	116		122		238	

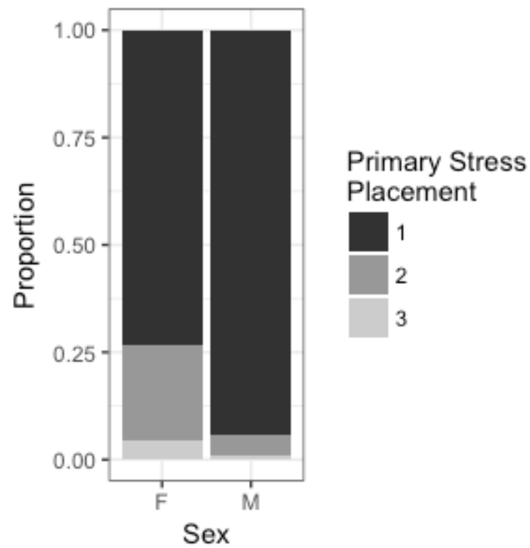

Figure 21 Bar chart of the distribution of primary stress placement in English male and female names

Nasal Vowel Percentage (French Only): In my French corpus, the mean proportion of nasal vowels (Table 25, Figure 22) was lower in female names ($M = 0.02$, $SD = 0.09$) than in male names ($M = 0.13$, $SD = 0.23$).

Table 25

Mean and standard deviation for nasal vowel proportion in French male and female names

	M	***SD***
Female	0.02	0.09
Male	0.13	0.23

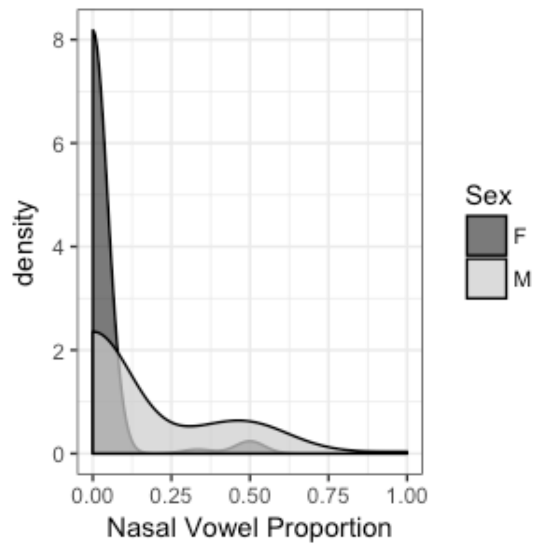

Figure 22 Density plot of nasal vowel proportion in French male and female names

4.5. Univariate Analyses

A univariate logistic regression model was constructed for each factor in each language to test if the factors were significant predictors of name gender on their own. The results, which

Table 26

Summary of the results of the corpus analysis

Factor	English		French	
	z-score	*p-value*	*z-score*	*p-value*
Number of Syllables	-3.565	< 0.001***	-1.700	< 0.1.
Open Syllable Proportion	-6.283	< 0.001***	-2.207	< 0.05*
Initial Sound Type	-1.424	0.154	-2.375	< 0.05*
Initial Syllable Type	-4.051	< 0.001***	-1.233	0.218
Final Syllable Type	-7.334	< 0.001***	-1.292	0.196
High Vowel Proportion	-2.613	< 0.01**	0.369	0.712
Low Vowel Proportion	2.420	< 0.05*	-0.303	0.762
Back Vowel Proportion	1.653	< 0.1.	3.074	< 0.01**
Round Consonant Proportion	-1.937	< 0.1.	-2.776	< 0.01**
Sharp Consonant Proportion	1.410	0.158	3.458	< 0.001***
Stress Placement	-4.048	< 0.001***		
Nasal Vowel Proportion			3.24	< 0.001***

are summarized in Table 26, indicate that, while not all factors were significant predictors of gender in English, most were trending in the direction expected based on the results of previous analyses. Most of the factors were also trending in the same direction in French as in English. Only high vowel proportion ($z = 0.369$, $p = 0.712$) and low vowel proportion ($z = -0.303$, $p = 0.762$) were not trending in the same direction as English; however, these factors display no discernable pattern relative to gender in French. Differences in significance between French and English suggest that, while many of the same factors are predictive of gender in both languages, which factors are most predictive varies across the languages.

4.6. Multivariate Analysis

Multivariate logistic regression models were constructed separately for the French and English names in the corpus. Each model included gender as the target factor and the ten common factors as well as the language-specific factor for each language as the predictor variables (1).

(1) glm(formula = Gender ~ Initial Sound Type + Number of Syllables + Open Syllable Percentage + Initial Syllable Type + Final Syllable Type + High Vowel Percentage + Low Vowel Percentage + Back Vowel Percentage + Round Consonant Percentage + Sharp Consonant Percentage + Stress Placement (English) OR Nasal Vowel Percentage (French), family = "binomial", data = ont)

The categorical predictor variables (initial sound type, initial syllable type, final syllable type and stress placement) were coded with the levels described in Table 1 (Consonant and Vowel, Open and Closed or Initial and Non-initial, depending on the variable). Sum contrast coding was used for these variables (Consonant, Closed and Initial = -0.5, Vowel, Open and Non-initial = 0.5).

The large number of predictor factors in the full models leads to the possibility that there are interactions or collinearities between the variables. In order to address this concern, stepwise logistic regression models (2) were constructed for each language using the step() function in R (R Core Team, 2018) to trim the models.

(2) step(glm(formula = Gender ~ Initial Sound Type + Number of Syllables + Open Syllable Percentage + Initial Syllable Type + Final Syllable Type + High Vowel Percentage + Low Vowel Percentage + Back Vowel Percentage + Round Consonant Percentage + Sharp Consonant Percentage + Stress Placement (English) OR Nasal Vowel Percentage (French), family = "binomial", data = ont))

These models were used to determine which predictor factors contribute to the model which best fits the data in each language. A best fit logistic regression model, like the one in (1), was constructed for each language based on the results of the stepwise model for that language.

4.6.1. English

The results of the full logistic regression model for English are summarized in Table 27. Only final syllable type ($z = -2.2014$, $p < 0.01$) and stress placement ($z = -1.2353$, $p < 0.05$) came out as significant. The negative coefficients suggest that female names are less likely to end in a final closed syllable or to have initial stress than male names.

Table 27

Summary of coefficients in the English logistic regression model for gender with number of syllables, open syllable percentage, initial sound type, initial syllable type, final syllable type, high vowel percentage, low vowel percentage, back vowel percentage, round consonant percentage, sharp consonant percentage and stress placement as predictor variables

	Estimate	Standard Error	z-value	Pr(>|z|)
(intercept)	-0.9142	1.1153	-0.820	0.412
Number of Syllables	0.2630	0.3446	0.763	0.445
Open Syllable %	-0.2618	1.6150	-0.162	0.871
Initial Sound Type	-0.3146	0.3940	-0.798	0.425
Initial Syllable Type	-0.5149	0.8303	-0.620	0.535
Final Syllable Type	-2.2014	0.8003	-2.751	<0.01**
High Vowel %	-0.1216	0.5577	-0.218	0.827
Low Vowel %	0.8558	0.6196	1.381	0.167
Back Vowel %	0.6801	0.5952	1.143	0.253
Round Consonant %	-0.8713	0.6635	-1.313	0.189
Sharp Consonant %	-0.9950	0.8916	-1.116	0.264
Stress Placement	-1.2353	0.5909	-2.091	<0.05*

The stepwise logistic regression model for English shows that the best fit model (AIC = 262.2) includes initial syllable type, final syllable type, low vowel percentage and stress placement as predictor variables. The results of this best fit logistic regression model are summarized in Table 28. Both final syllable type ($z = -6.311$, $p < 0.001$) and stress placement ($z = -2.141$, $p < 0.05$) came out as significant while initial syllable type ($z = -1.634$, $p =$

0.102) and low vowel proportion (z = 1.477, p = 0.140) approached marginal significance. These results indicate that female names are more likely to have non-initial stress and to end with an open syllable compared to male names.

Table 28

Summary of coefficients in the English logistic regression model for gender with initial syllable type, final syllable type, low vowel percentage and stress placement as predictor variables

	Estimate	Standard Error	z-value	Pr(>|z|)
(intercept)	-0.6717	0.2876	-2.335	<0.05*
Initial Syllable Type	-0.5999	0.3670	-1.634	0.102
Final Syllable Type	-2.1534	0.3412	-6.311	<0.001***
Low Vowel %	0.8684	0.5881	1.477	0.140
Stress Placement	-1.0609	0.4955	-2.141	<0.05*

4.6.2. French

The results of the full logistic regression model for French are summarized in Table 29.

Table 29

Summary of coefficients in the French logistic regression model for gender with number of syllables, open syllable percentage, initial sound type, initial syllable type, final syllable type, high vowel percentage, low vowel percentage, back vowel percentage, round consonant percentage, sharp consonant percentage and nasal vowel percentage as predictor variables

	Estimate	Standard Error	z-value	Pr(>|z|)
(intercept)	1.39494	1.00116	1.393	0.164
Number of Syllables	0.02105	0.28665	0.073	0.942
Open Syllable %	-2.97239	1.83561	-1.619	0.105
Initial Sound Type	-0.6606	0.38111	-1.732	<0.1.
Initial Syllable Type	1.35960	0.97878	1.389	0.165
Final Syllable Type	1.16254	0.79165	1.468	0.142
High Vowel %	0.47755	0.67098	0.712	0.477
Low Vowel %	-0.76479	0.69240	-1.105	0.269
Back Vowel %	0.48452	0.69890	0.693	0.488
Round Consonant %	-0.77177	0.56463	-1.367	0.172
Sharp Consonant %	5.12582	4.07453	1.258	0.208
Nasal Vowel %	-0.10798	3.47741	-0.031	0.975

No factors came out as significant, however initial sound type ($z = -0.6606$, $p < 0.1$) came out as marginally significant. The negative coefficient suggests that female names may be more likely to begin with a vowel than male names.

The stepwise logistic regression model for French shows that the best fit model (AIC = 252.9) includes open syllable percent, initial sound type, initial syllable type, final syllable type, low vowel percentage and sharp consonant percentage as predictor variables. The results of this best fit logistic regression model are summarized in Table 30. Open syllable percentage ($z = -2.092$, $p < 0.05$), initial sound type ($z = -2.279$, $p < 0.05$) and sharp consonant percentage ($z = 3.301$, $p < 0.001$) came out as significant while final syllable type ($z = 1.647$, $p < 0.1$) and low vowel percentage ($z = -1.877$, $p < 0.1$) came out as marginally significant. Initial syllable type ($z = 1.441$, $p = 0.150$) approached marginal significance. These results indicate that female names are more likely to have a higher proportion of open syllables and to begin with a vowel while male names are more likely to contain a higher proportion of sharp consonants.

Table 30

Summary of coefficients in the French logistic regression model for gender with open syllable percentage, initial sound type, initial syllable type, final syllable type, low vowel percentage and sharp consonant percentage as predictor variables

	Estimate	Standard Error	z-value	Pr(>|z|)
(intercept)	1.6077	0.8977	1.791	<0.1.
Open Syllable %	-3.2901	1.5725	-2.092	<0.05*
Initial Sound Type	-0.8010	0.3515	-2.279	<0.05*
Initial Syllable Type	1.3470	0.9347	1.441	0.150
Final Syllable Type	1.1865	0.7202	1.647	<0.1.
Low Vowel %	-1.1255	0.5997	-1.877	<0.1.
Sharp Consonant %	5.5302	1.6754	3.301	<0.001***

4.7. Discussion

The by-factor analysis showed that, for English, all phonological factors were trending in the expected direction. For French, all factors except high vowel proportion and low vowel proportion were trending in the expected direction. There was no discernable pattern for high vowel proportion or low vowel proportion. Open syllable proportion was found to be significant in both languages. Number of syllables was significant in English and marginally significant in French. Back vowel proportion and round consonant proportion were significant in French and marginally significant in English. Initial syllable type, final syllable type, high vowel proportion and low vowel proportion were significant in English, but not French. Initial sound type and sharp consonant percentage were significant in French, but not English. Stress placement and nasal vowel proportion were significant in English and French, respectively.

The multivariate stepwise logistic regression run for each language assessed the relative effects of the factors when all factors were considered. The best fit model for English included initial syllable type, final syllable type, low vowel proportion and stress placement as predictor variables, with final syllable type and stress placement coming out as significant. The best fit model for French included open syllable proportion, initial sound type, initial syllable type, final syllable type, low vowel proportion and sharp consonant proportion as predictor variables with open syllable proportion, initial sound type and sharp consonant proportion coming out as significant and final syllable type and low vowel percentage coming out as marginally significant.

Table 31

Summary of results of the descriptive and inferential analyses of the French and English corpus data

	English			**French**		
Factor	*Trend*	*Univariate*	*Multivariate*	*Trend*	*Univariate*	*Multivariate*
Number of Syllables	✓	✓		✓	•	
Open Syllable %	✓	✓		✓	✓	✓
Initial Sound Type	✓			✓	✓	✓
Initial Syllable Type	✓	✓		✓		
Final Syllable Type	✓	✓	✓	✓		•
High Vowel %	✓	✓				
Low Vowel %	✓	✓				•
Back Vowel %	✓	•		✓	✓	
Round Consonant %	✓	•		✓	✓	
Sharp Consonant %	✓			✓	✓	✓
Stress Placement	✓	✓	✓			
Nasal Vowel %				✓	✓	

Note: ✓ = significant ($p < 0.05$), • = marginally significant ($p < 0.1$)

The results, which are summarized in Table 31, suggest that there may be cross-linguistic and/or language-specific phonological patterns in gender in given names. First, the analysis of a new set of English given names shows that they display the same patterns observed in previous studies (Cutler et al., 1990; Sidhu & Pexman, 2015; Slater & Feinman, 1985; Wright et al., 2005). This confirms the existence of these patterns in English. As with the previous studies, this study looked at popular English names in North America. It is possible that an analysis of English names from another part of the world, such as England or Australia, would yield different results. Second, the analysis of French names shows that they display many of the same patterns as English names. This suggests that, at least some of the patterns may be cross-linguistic while others may be language-specific. The results show that which factors are most important varies between the two languages, suggesting that, if the patterns are cross-linguistic, the relative importance of each factor may be determined by each language. To show that these patterns are cross-linguistic, the corpus analysis needs to be extended to other, genetically unrelated languages.

5. Name Gendering Experiment

While the corpus analysis shows that there are correlations between the gender of given names and their phonology, it does not reveal whether these patterns are psychologically available to people or if they can actively make use of them. An online experiment was, therefore, conducted to explore if people internalize and actively make use of the cross-linguistic and language-specific phonological naming patterns found in the corpus analysis. The experiment consisted of three parts: a language background questionnaire, a name gendering task and an AX discrimination task. The language background questionnaire was used to determine if participants met the language background criteria for participation in the experiment. The main task was the name gendering task which tested if people assign nonce names, which were created to be biased to one gender, to that gender at an above-chance rate. The nonce names in this task were controlled for four factors (number of syllables, final syllable type, presence of a back vowel and presence of a round consonant) present in both French and English and one language-specific factor for each language. It is expected that, if participants have internalized and actively make use of phonological naming patterns, they should categorize nonce names with more phonological properties of one gender than the other as being of that gender at an above-chance rate. Finally, the AX discrimination task tested if participants could perceive the language-specific factors tested in the name gendering task.

5.1. Participants

27 monolingual North American English speakers participated in the online experiment. The participants were 18 years of age or older, had normal speech, hearing and vision, and had limited exposure to languages other than their native language. They were required to have access to headphones and a computer with internet access and a functional audio system. Participants were recruited online through email and social media and received a $5 Amazon.ca gift card as compensation. Participants provided informed consent using the form in Appendix B prior to completing the experiment. Nine participants were excluded from the analysis due to technical errors (2), failing to complete the experiment properly (1), failure to meet the language requirements (2) and failure to discriminate nasal and non-nasal vowels (4). Monolingual North American French speakers were also recruited to complete the experiment, however, their results aren't reported here due to low numbers of participants.

5.2. Language Background Questionnaire

A language background questionnaire was conducted to assess speakers' level of proficiency in and exposure to French, English and any other language(s) that they speak. The questionnaire was based on Chambers & Lapierre's (2011) Language Use Index (LUI) and Marian, Blumenfeld & Kaushanskaya's (2007) Language Experience and Proficiency Quetionnaire (LEAP-Q).

For French and English, the questionnaire had three parts: current exposure, historical exposure and education level. The current exposure section, based on the LUI (Chambers & Lapierre, 2011), asked participants to rate how often they used French or English in five areas: at home, at work or school, with family, with friends and in the media. Participants rated their use in each area as always, often, seldom or never. The historical exposure section asked participants to list each time they had lived in a Francophone or Anglophone community for at least three months, including the country they were living in, their age when they lived there and how long they lived there. The education level section asked participants to indicate the amount of education they received at the primary, secondary and post-secondary levels by indicating how many years they were taught the language at each level, as well as their degree of immersion (native, immersion, core).

For languages other than English, participants were asked to list all the other languages they knew as well as their level of proficiency in each language (beginner, intermediate, advanced or native). A copy of the language background questionnaire can be found in Appendix C.

5.3. Stimuli

5.3.1. Name Gendering Task

The target stimuli consisted of minimal pairs of nonce names whose members differed from each other in terms of one of six target factors: number of syllables, final syllable type, vowel backness, presence of a round consonant, stress placement and presence of a nasal vowel. These factors, which were selected to cover a range of types of phonological phenomena including word length, weight, vowel quality and consonant type, are summarized in Table 32. The language specific factors (stress placement, presence of a nasal vowel) were selected because they were not phonologically possible in the other language. Female names tend to have more syllables than male names, are more likely to end in an open syllable, contain a sonorant consonant or have non-initial stress placement, and are less likely to contain a back vowel or a nasal vowel.

Table 32

Target factors selected for the "Name Gendering" task

Factor	**Male**	**Female**
Length Number of syllables	Less Syllables	More Syllables
Weight Final syllable type	More likely to end in a closed syllable	More likely to end in an open syllable
Vowel Quality Vowel backness	More likely to have a back vowel	Less likely to have a back vowel
Consonant Type Presence of round consonant	Less likely to have a sonorant consonant	More likely to have a sonorant consonant
Language Specific – English Stress placement	Less likely to have non-initial stress	More likely to have non-initial stress (compared to male names)
Language Specific – French Presence of a nasal vowel	More likely to contain a nasal vowel	Less likely to contain a nasal vowel

In order to generate these pairs, a list of sequences of two open syllables (CVCV sequences), using 13 consonants and 7 vowels found in both English and French was generated using R (R Core Team, 2018). The consonants used were /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /s/, /z/, /m/, /n/ and /l/. The vowels used were /i/, /e/, /ɛ/, /ə/, /u/, /o/ and /ɑ/. The list generated contained 8281 CVCV sequences. This list was reduced by eliminating real words, sequences with more than one instance of the same segment and those that were biased, according to the factors established in the corpus analysis, towards being more male or female.

To eliminate real words, the CVCV sequences were matched against pronunciations from the CMU Pronouncing Dictionary (Carnegie Mellon University, 2014) for English words and Lexique (New et al., 2001) for French words. Any sequence which matched a pronunciation from either of these dictionaries was removed. This reduced the number of available sequences to 7050. CVCV sequences in which either both consonants or both vowels were the same were then eliminated, reducing the number of sequences to 5619.

Each of the remaining sequences was coded for factors which were found to vary according to gender in the corpus analysis. These factors were number of back vowels, number of low vowels, number of high vowels, presence of schwa, number of sonorant consonants and number of voiceless stops. The sequences were coded such that they received a score of +1 for every more female characteristic and -1 for every more male characteristic. For example, number of back vowels was coded such that a sequence received a score of -1 for every back vowel in it and +1 for every front vowel in it, with /ə/ being excluded as neither back nor front. A summary of the coding criteria can be found in Table 33.

Table 33

Criteria for coding CVCV sequences for gender bias

Factor	Criteria	Example (/vupə/)
Number of Back Vowels	-1 for every back vowel; +1 for every front vowel; /ə/ not included in calculation	-1
Number of Low Vowels	-1 for every low vowel	0
Number of High Vowels	+1 for every high vowel	+1
Presence of Schwa	1 if a schwa is present; otherwise 0	+1
Number of Round Consonants	+1 for every round consonant	0
Number of Sharp Consonants	-1 for every sharp consonant	-1

These factors were used to calculate three gender bias scores for each word: vowel gender bias, consonant gender bias and overall gender bias. Vowel gender bias was calculated by summing number of low vowels, number of back vowels and number of high vowels. Consonant gender bias was calculated by summing number of sharp consonants and number of round consonants. Overall gender bias was calculated by summing the vowel and gender bias scores. All sequences with a bias value in any of the three categories that was not 0, as well as any sequences containing a schwa, were eliminated to avoid sequences which may be biased to being more male or female. This reduced the number of sequences available to 334.

Twenty of the remaining sequences were pseudo-randomly selected to be gender-neutral target nonce names. Each of these names was then modified to create a second target nonce name which varied from the first by being either more male or more female than the first based on the target factors listed in Table 3. Four were modified for each of number of syllables, final syllable type, presence of a back vowel and presence of a round vowel. Two were modified for each of stress placement and presence of a nasal vowel. This created twenty minimal pairs consisting of a total of forty target names.

For some of the factors (number of syllables, consonant type and stress placement), neutral names were modified to make them more female-leaning, while for other factors (final syllable type, presence of a back vowel and presence of a nasal vowel), the modification made the names more male-leaning. For English names, the default stress placement was initial, keeping with the preference for initial stress in names of both genders found in the corpus analysis. For number of syllables, an additional syllable (/və/) was added between the two existing syllables (e.g. /leto/ becomes /levəto/). For presence of a round consonant, a consonant that was not a voiceless stop was changed to a sonorant (e.g. /boze/ becomes /bole/). For stress placement, stress was shifted from the initial syllable to the final syllable (e.g. /ˈbevo/ becomes /beˈvo/). For syllable type, a /v/ was added to the final syllable so that it became closed rather than open (e.g. /vɑdi/ becomes /vɑdiv/). For vowel backness, the first vowel in the word, which was a front vowel, was changed to a back vowel with similar vowel height (e.g. /sifɑ/ becomes /sufɑ/). For the presence of a nasal vowel, the first vowel in the word was nasalized (e.g. /dɛzo/ becomes /dɛ̃zo/). The sounds /v/ was chosen for consonant insertion because it was neither a round, nor a sharp consonant. The sound /ə/ was chosen for vowel insertion because it is central, and therefore neutral, in terms of both backness and height. Where possible, sounds that were modified were those that were neutral in terms of the target factors.

The target stimuli were divided into two groups, one to represent English names and the other, French names. For the factors common to both languages, two pairs were assigned to each language group. The stress placement pairs were assigned to the English group and the presence of a nasal vowel pairs to the French group. Four additional CVCV sequences were selected pseudo-randomly to serve as training items. These were divided into two groups, such that there were four training items for each language. A complete list of target and training stimuli can be found in Table 34.

Table 34

Stimuli for the "Name Gendering" Task

	English		French	
	Neutral	*Biased*	*Neutral*	*Biased*
Number of syllables	lɛto	lɛvəto	gefo	gevəfo
	dɑfi	dɑvəfi	tilɑ	tivəlɑ
Final syllable type	vɑdi	vɑdiv	fobɛ	fobɛv
	nope	nopev	likɑ	likɑv
Vowel backness	sifɑ	sufɑ	nikɑ	nukɑ
	nɛpo	nɑpo	kɛlo	kɑlo
Presence of round consonant	boze	bole	gofɛ	golɛ
	bisɑ	bilɑ	gibɑ	gilɑ
Stress placement (English Only)	ˈnɑpi	nɑˈpi		
	ˈbevo	beˈvo		
Presence of a nasal vowel (French Only)			dɛzo	dɛ̃zo
			gɑsi	gɑ̃si
Training items	kəmu, vofe, zɛgo, nəku		temo, vode, dəfu, fego	

5.3.2. AX Discrimination Task

The stimuli for the AX discrimination task consisted of the stimulus pairs from the "Name Gendering" task for the language specific factors – nasal vowels and stress placement. Two additional pairs were generated for each factor by selecting four additional CVCV sequences from the list generated for the name gendering task. These items served as fillers and were not included in the analysis. Each of these sequences was then modified in the same way as was done for the name gendering task to create a minimal pair that varied only in terms of the target factor. Two were modified for stress placement and two for nasal vowels. A complete list of stimuli for this task can be found in Table 35.

For each stimulus pair, each item was paired with itself, and the two different items were paired with each other in two different orders, giving four AX discrimination pairs per stimulus pair. This means there are four pairs, two in which the items are the same (/gɑsi/-/gɑsi/ and /gɑ̃si/-/gɑ̃si/) and two in which they are different (/gɑsi/-/gɑ̃si/ and /gɑ̃si/-/gɑsi/).

Table 35
Stimuli for the AX discrimination task

	Stress Placement		**Nasal Vowels**	
Target Items	ˈgisɑ	giˈsɑ	dɛzo	dɛ̃zo
	ˈbevo	beˈvo	gɑsi	gɑ̃si
Fillers	ˈzɑde	zɑˈde	pɛke	pɛ̃ke
	ˈmɛzu	mɛˈzu	lɑtə	lɑ̃tə

5.3.3. Recording and preparation of audio stimuli

The stimuli were recorded by a native speaker of Canadian English and a native speaker of Canadian French. Both speakers were phonetically trained linguists. Stimuli were recording using a Sound Devices 722 digital audio recorder and a DPA 4011 unidirectional cardioid microphone. The speakers were presented with the stimuli corresponding to their native language in IPA. They were asked to read each stimulus item clearly three times into the microphone.

The stimuli were prepared using Praat (Boersma & Weenink, 2016). Two of the three recordings were extracted from the audio file for each word. Generally, this was the second and third recording; however, if there was a problem with one of these recordings, the first recording was selected instead of the problematic one. For one word (/dɛ̃zo/), only the first recording was correct, so only one recording was used of that word. For each extracted word, approximately 500ms of silence was added before and after the word and the intensity of the word was adjusted to 70db. Each of these sounds was then saved as a wav file and converted to an mp3 file. The better of the two recordings was used for the name gendering task. Both recordings were used for the AX discrimination task, with the better recording being used when each word was paired with a different word, and the worse recording only being used when the word was being paired with itself. The exception is /dɛ̃zo/ for which there was only one usable recording. That recording was paired with itself and the /dɛ̃zo/-/dɛ̃zo/ pair was excluded from the analysis.

A pure tone was also created in Praat (Boersma & Weenink, 2016) so that participants could test their sound level prior to beginning the experiment. The intensity of this sound was adjusted to 80db to make it comparable to the intensity of the stimuli.

5.4. Procedure

In order to participate in the experiment, participants sent an email to the researchers. Participants were assigned to one of two conditions: French first or English first. The purpose of this was to counterbalance the order of the two blocks of the experiment. Participants were then emailed a participation code and a link to the version of the experiment they were to complete. They completed the experimental tasks in a web browser on their computer. The script used to present the experiment to the participants was developed using jsPsych (de Leeuw, 2015) and jsPsych scripts developed by Hyoung Seok Kwon and Na-Young Ryu. Upon arriving at the website, participants saw a welcome screen and were asked to click “Next” to begin the experiment. They were then asked to enter their participation code. This was followed by the informed consent form (Appendix B), which asked participants to indicate whether or not they agreed to participate in the experiment. If participants consented, they proceeded with the experiment. Otherwise, they received a message thanking them for their time.

The first experimental task participants were presented with the language background questionnaire, which is described in section 5.2. English-speaking participants were presented with this part of the questionnaire for English first and French second. French-speaking participants were presented with the French section first and the English second. All participants were then presented with the other languages section.

Following the completion of the language background questionnaire, participants were reminded that they need headphones for the experiment and will be asked to indicate which type of headphones they are using. This information was not used by the experimenters; rather, it served as a reminder for participants to use headphones. Following this, participants were asked to check their sound using a button which played a pure tone with similar volume to the experimental stimuli and to adjust their volume so they could hear the tone clearly at a comfortable sound level.

Next, the stimuli were presented to the participants in two blocks, one for each language. Each block consisted of a name gendering task followed by an AX discrimination task. The order of the two blocks was counterbalanced across participants. Each name gendering task began with instructions, which were followed by the four training items. The target stimuli were presented after the training items.

For the name gendering task (Figure 23), participants were given the prompt: *A new family from Ottawa has moved in next door. They have two kids: a boy and a girl. The kids have unusual names and you are trying to figure out what each kid’s name is.* They were then instructed that they would hear a name and be asked to rate how male or female they think it is. They were also reminded to make sure their headphones are in and their volume is on. Each stimulus item was presented aurally to the participant. The participant first heard the target name while an audio symbol was displayed on the screen. Once the sound was finished playing, they were presented with the question: *Is this name male or female?* and a 6-point

Likert-type scale ranging from definitely female to definitely male. Participants rated how male or female they thought the name was and clicked “Next” to proceed to the next trial. The first four items the participant heard were training items which were randomized by participant. Following this, the participant heard the target items, which were also randomized by participant.

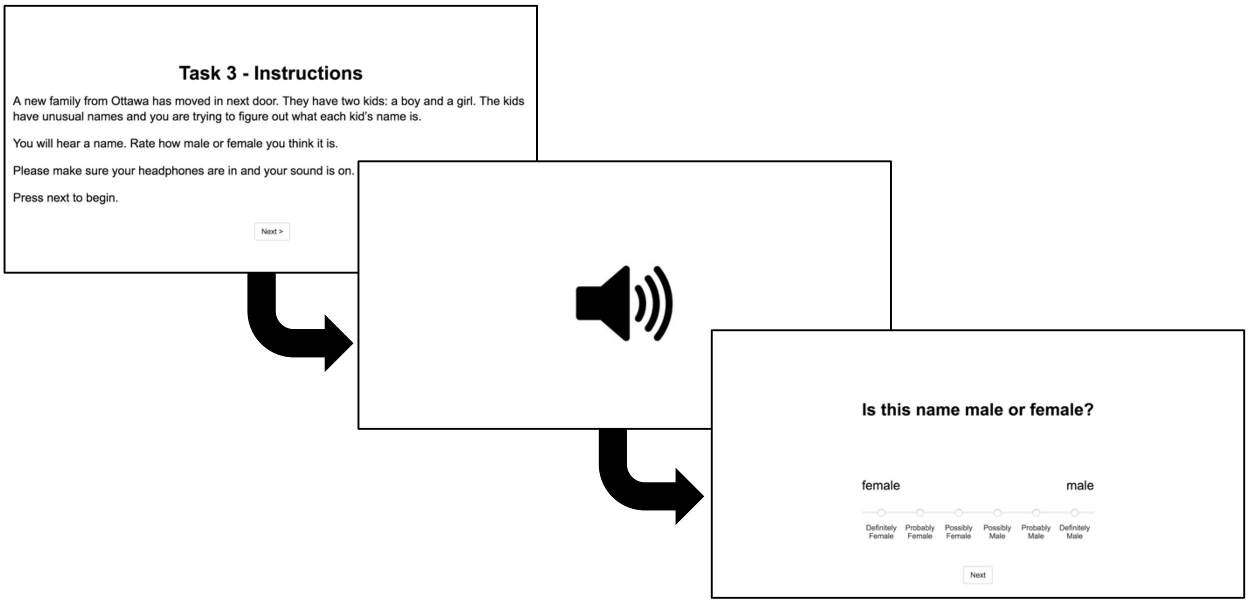

Figure 23 Name Gendering Task Procedure

The AX discrimination task (Figure 24) following the English name gendering task was for stress placement whereas the one following the French name gendering task was for nasal vowels. For this task, participants were told that they would hear two words and would have to decide if the words were the same or different. They were again reminded to make sure their headphones were in and their volume was on. Participants heard the sound pair while two audio symbols were presented on the screen. The interstimulus interval (ISI) was approximately 1 second. They were then prompted with the question *Are these two words the same*? and had to choose “Yes” or “No”. Once they made their choice, the next trial started. Participants first heard four training items selected from the eight discrimination pairs generated from the fillers. These items were randomized by participant. Next, the participant heard the discrimination pairs generated from the target stimuli interspersed with the remaining filler items. These items were also randomized by participant.

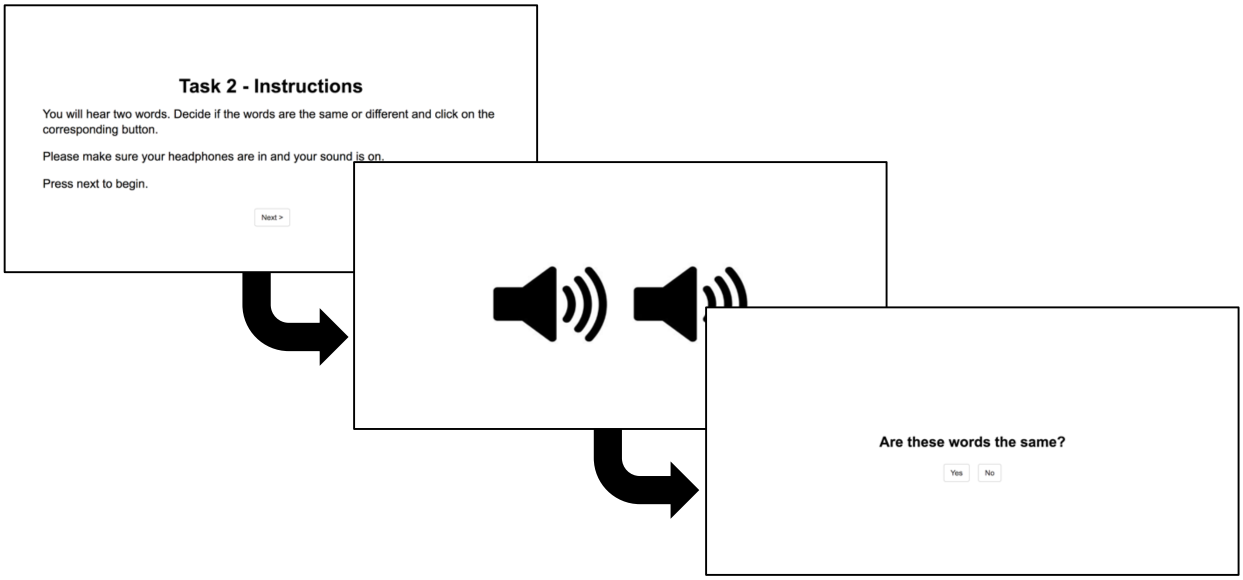

Figure 24 AX Discrimination Task Procedure

After the completion of the first block, either French or English, participants completed the second block. The procedure and instructions for this block were identical to those of the first block, except that the stimuli corresponded to the other language. Once participants completed this block, they were thanked for their time and informed that they would receive compensation within one week of completing the experiment.

5.5. Exclusion Criteria

In addition to exclusions due to technical errors and failure to complete the tasks properly, participants were also excluded from the analysis for failure to meet the language background criteria, based on the results of the language background questionnaire and failure to discriminate between nasal and non-nasal vowels based on the results of the AX discrimination task.

The language background questionnaire was coded for 3 measures of language proficiency and exposure: current exposure, historical exposure and education level as well as proficiency in languages other than French or English. Participants who exceed thresholds in any of the categories were excluded from the analysis. A current exposure score was calculated for French by assigning a score based on the participant's responses to each of the factors in the first question (Always – 3, Often – 2, Seldom – 1, Never – 0). These scores were summed to get a total current exposure score. Participants with scores greater than 5 were excluded from the analysis. A historical exposure score was tabulated for French by summing the total number of years and months of exposure to French. Participants with more than two years total spent living in French-speaking communities were excluded from the analysis. For education level, participants with more than two years of French immersion were excluded from the analysis. Participants who were native speakers of a language other than English were also excluded from the analysis. This excluded two participants.

For each target item in the AX discrimination task, participants' responses were coded as "1" if they were correct and "0" otherwise. These scores were summed separately for stress and nasal vowel discrimination for each participant. Excluding the participants removed from the analysis due to incomplete data or failure to meet the language requirements, participants had mean scores of 0.90 ($N = 22$, $SD = 0.19$) for stress discrimination and 0.82 ($N = 22$, $SD = 0.21$) for nasal vowel discrimination. Participants who responded correctly to less than 70% of the nasal vowel target trials were excluded from the analysis. This excluded four participants.

5.6. Results

Participants, on average, rated female biased names ($M = 3.34$, $SD = 1.17$) as more female than male biased names ($M = 3.54$, $SD = 1.21$). Ratings were given on a 6-point scale where lower numbers (1-3) represented more female names and higher numbers (4-6) represented more male names. When divided by factor and gender (Table 36), there is some variability in the ratings, with final syllable type in French and presence of a round consonant in English having female-biased names rated as more male than male-biased names and stress placement and number of syllables in English being rated as equally male for both male and female-biased names. However, in most cases, more male-biased names were rated more male than female biased names. Combining all the factors and looking at the languages separately shows that French names (female: $M=3.45$, $SD=1.12$; male: $M=3.72$, $SD=1.14$) were rated higher than English names (female: $M=3.23$, $SD=1.21$; male: $M=3.36$, $SD=1.25$) and that female-biased names were rated as more female than male-biased names. Combining both languages and looking at each factor shows that, for every factor except stress, in which both male- and female-biased were rated as equally male, male-biased names were rated as more male than female-biased names (Figure 25). Additional tables and graphs can be found in Appendix D.

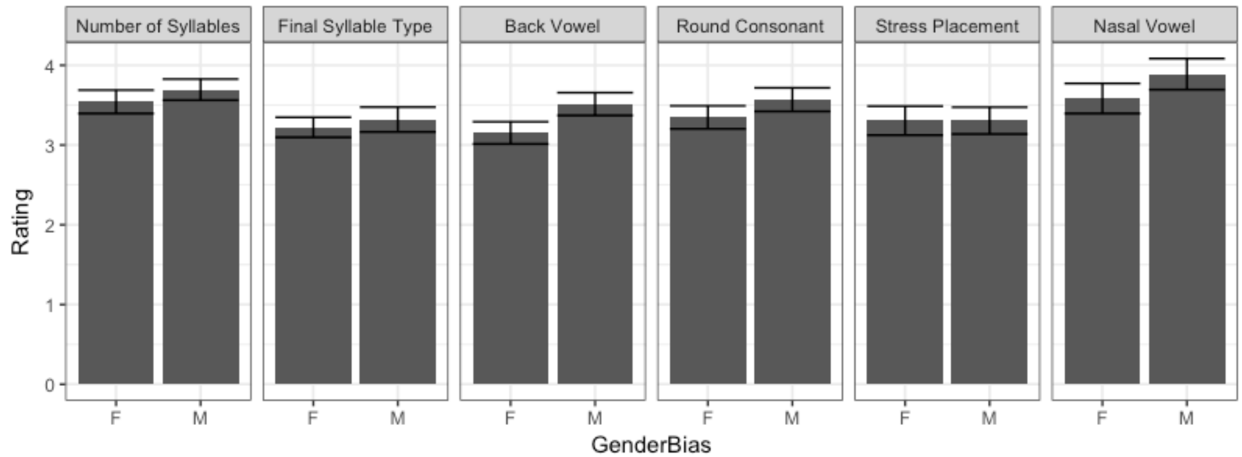

Figure 25 Mean ratings of male- and female- biased names by factor

Table 36

Mean ratings and standard deviations for male-biased and female-biased names by language and factor

	English				French			
	Female		Male		Female		Male	
Factor	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Number of Syllables	3.50	1.32	3.50	1.28	3.58	1.18	3.89	0.92
Final Syllable Type	3.31	1.12	3.67	1.39	3.14	1.02	2.97	1.16
Presence of a Back Vowel	3.00	1.17	3.31	1.26	3.31	1.19	3.72	1.14
Presence of a Round Consonant	3.06	1.35	3.03	1.25	3.64	1.02	4.11	1.01
Stress Placement	3.31	1.09	3.31	1.01				
Presence of a Nasal Vowel					3.58	1.13	3.89	1.17

A mixed effects linear regression model was built to assess the effects of gender bias, language and factor on participants' ratings of nonce names (3). The model also assesses the possibility of interactions between gender bias and factor and between gender bias and language. This assesses if ratings for gender bias are equal across all factors or speaker languages. Stimulus pair and participants were coded as random factors with gender bias as a random intercept for participants. This controls for individual differences in ratings across participants and stimulus pairs.

(3) lmer(Rating~GenderBias*Factor + GenderBias*Language + (1|Pair) + (1 + GenderBias|Participant), names)

The results indicate that there are no significant effects, however, gender bias is trending in the expected direction in the model ($t = 1.609$, $p = 0.108$). An ANOVA of the model indicates a significant effect of gender bias ($F_{(1, 573.56)} = 5.3605$, $p < 0.05$) but not of factor ($F_{(5,13)} = 0.2045$, $p = 0.955$), language ($F_{(1,13)} = 0.7509$, $p = 0.402$), the interaction between gender bias and factor ($F_{(5,676)} = 0$, $p = 0.894$), or the interaction between gender bias and language ($F_{(1,676)} = 0.3132$, $p = 0.576$). A table of results for this model can be found in Appendix D.

As gender bias was the only significant predictor of participants' ratings in the ANOVA, a trimmed down mixed effects linear regression model was constructed in which the random effects and target variable were the same, but only gender bias served as a predictor variable (4).

(4) lmer(Rating~GenderBias + (1|Pair) + (1 + GenderBias|Participant), names)

The results of this model, which are summarized in Table 37, indicate that gender bias is a significant predictor of rating, with higher ratings being associated with male-biased names and lower ratings being associated with female-biased names ($t = 2.542$, $p < 0.05$).

Table 37

Summary of coefficients in the full mixed effects linear regression model for gender with gender bias as the predictor variable and stimulus pair and participant as random effects

	Estimate	Standard Error	t-value	Pr(>|t|)
(intercept)	3.34167	0.14630	33.841	<0.001***
Gender Bias (M)	0.19722	0.0.07759	2.542	<0.05

5.7. Discussion

The results indicate that gender bias, but not phonological factor or speaker language has a significant effect on participant rating of the gender of nonce names. These results pattern in the expected direction, with female-biased names receiving more female ratings than male-biased names. This is consistent with the results of previous studies looking at peoples' use of phonology in assigning gender to nonce names (Cassidy et al., 1999; MacAuley, in prep). The results also suggest that there are no significant differences between the phonological factors or between the speaker languages. That is, the results for both languages pattern in the same way, as do the results for the individual phonological factors.

Furthermore, there is no significant effect of either the interaction between gender bias and phonological factor or that between gender bias and speaker language. This means that there was no significant difference in how ratings pattern in terms of gender bias either across speaker languages, or across factors. It is possible, however, that the lack of interaction between gender bias and phonological factor is the result of lack of power, rather than the fact that there are no differences in rating patterns across factors. Each factor had only two to four pairs of stimulus items, meaning only two to four items which were either female-biased or male-biased. Such low numbers are difficult or impossible to generalize from, especially when the ratings are close, as was the case in this experiment. This is especially problematic for the stress placement and presence of a nasal vowel factors, which only had two stimulus pairs because they were language specific. This may explain the results for stress placement, in particular, which showed no difference in ratings between male-biased and female-biased names. These two pairs displayed opposing patterns and, when the means for male- and female-biased names were calculated, they averaged out. In order to get a better idea of how each factor behaves, it would be necessary to have more stimulus pairs for each phonological factor.

Overall, these results suggest that speakers are able to internalize and make use of phonological patterns to assign gender to names. This suggests that these patterns are psychologically available. Furthermore, the fact that the English and French language results pattern similarly, including the French-only presence of a nasal vowel factor, suggests that speakers have internalized these at an abstract level and are able to apply them to names from another language. It is not clear, however, how speakers acquired these patterns. The fact that the French results pattern with the English ones allows for the possibility that the patterns are learned by a method other than exposure to a specific language, however, as most of the factors are the same across languages, there is not sufficient evidence to support this hypothesis over the hypothesis that the patterns are learned from exposure to English or French.

The behavior of participants on the presence of a nasal vowel factor potentially provides the best indication of method of acquisition because it is French-specific. It is possible that participants were able to associate this more with male names, not because they associate nasal vowels with this, but because of its effect on syllable structure. In the stimuli used in the experiment, changing an oral vowel to a nasal vowel resulted in creating what some abstract analyses of French phonology (such as Paradis & Prunet, 2000) consider an initial closed syllable, which decreased the proportion of open syllables in the name. The change, thus, made the name more male in ways which are associated with factors found in English, in addition to adjusting the type of vowel. Using default stimuli with initial syllables which end in a consonantal nasal would help isolate the nasal vowel factor. Even if this factor were not conflated with non-language specific other factors, it is possible that it does behave differently from the other factors. This is difficult to determine due to the lack of power in the experiment, as discussed above. Increasing the number of stimulus items for this factor would increase the power of the study and give a better indication of how the factor behaves.

6. Conclusion

A corpus analysis and name gendering experiment were conducted to determine if there were correlations between phonology and gender in French and English given names. The results of the corpus analysis indicate that these correlations exist. For English, the patterns match those found in previous corpus analyses (Cutler et al., 1990; MacAuley et al., in prep; Sidhu & Pexman, 2015; Slater & Feinman, 1985; Wright et al., 2005). For French, most of the patterns are the same, or similar to, those found in English, suggesting that the patterns may be cross-linguistic and not just language specific. A multivariate analysis showed that the patterns which were most significant in each language differed, suggesting that, while the patterns may be universal, there may be a degree of specialization which goes on within each language.

One shortcoming of the corpus analysis is that it only looked at two genetically related languages, which are in close contact. While the results suggest the patterns found may be cross-linguistic, it would be necessary to extend the corpus analysis to more languages, including those genetically unrelated to French or English. Similarities in phonology-gender correlations in given names across a wide variety of languages would provide strong evidence for the notion that these correlations are cross-linguistic and not just language-specific.

Additionally, as names were selected for the corpus analysis based on popularity and were from one time point, it is possible that some of the patterns observed are social rather than phonological. To address this, future research could extend the analysis to look at historical data, analyzing the most popular names at several time points to get an idea of what, if any patterns change across time. Such changes may indicate that these patterns are social rather than phonological. It is possible that some of the patterns found in the current data could be social trends in both languages due to the contact between French and English and the fact that the names were collected from databases from two adjacent Canadian provinces.

The results of the name gendering task show that speakers have internalized and actively make use of the patterns, as was previously found by Cassidy et al. (1999) and MacAuley et al. (in prep). Speakers rated more female-biased names as more female than male biased names, regardless of phonological factor or speaker language. This suggests that participants have externalized these patterns in an abstract way and are able to apply them to novel situations and languages. However, due to the low number of items for individual factors, the study lacked power to determine if individual phonological factors behaved the same or differently. Additionally, the experiment failed to address the method of acquisition question because the presence of a nasal vowel factor is conflated with syllable structure factors which are present in English and could account for participants' behavior on this factor.

In order to better address the issue of method of acquisition, future research should focus on increasing the amount of non-native data assessed. One way to do this would be to increase the number of stimulus items for the presence of a nasal vowel factor to get a better idea of how it behaves. Another would be to extend the research to include monolingual French speakers. The current experiment has stress placement as a factor, which is not present in French due to its lack of lexical stress. Thus, if French speakers were able to use stress placement to assign gender to nonce names in the expected way, this would lend additional support to the notion that the patterns are acquired by a means other than exposure to the native language.

A third way to increase the amount of non-native data would be to include phonological factors from other languages which correlate with gender but are not present in English (or French) phonology. English (and French) speakers using these in the expected way would be indicative of an alternative method of acquisition. This would also address the issue of the genetic and areal proximity of French and English, which could account for English speakers' assessment of the presence of a nasal vowel feature. Ideally, having extended the corpus analysis to include additional languages, it would be possible to do all three of these things, as well as including multiple speakers of languages other than English or French in the analysis. Speakers of multiple languages behaving similarly on factors from a variety of languages, which may or may not be phonologically possible in their language, provides strong support for the idea that these patterns are learned by a method other than exposure to the source language.

References

Abel, G. A., & Glinert, L. H. (2008). Chemotherapy as language: Sound symbolism in cancer medication names. *Social Science & Medicine*, *66*(8), 1863-1869.

Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.11, retrieved from http://www.praat.org/

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). "Bouba" and "Kiki" in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. *Cognition*, *126*(2), 165-172.

Campbell, M.D. (2013). Behind the name: the etymology and history of first names. Retrieved from https://www.behindthename.com/

Carnegie Mellon University (2014). The Carnegie mellon university pronouncing dictionary [cmudict 0.7b]. Retrieved from http://www.speech.cs.cmu.edu/cgi-bin/cmudict

Cassidy, K. W., Kelly, M. H., & Sharoni, L. J. (1999). Inferring Gender From Name Phonology. *Journal of Experimental Psychology*, *128*(3), 362–381.

Chambers, J. K., & Lapierre, A. (2011). Dialect variants in the bilingual belt. *Le français en contact. Hommage à Raymond Mougeon, Québec, Les Presses de l'Unversité Laval*, 35-85.

Cutler, A., Mcqueen, J., & Robinson, K. (1990). Elizabeth and John, Sound Patterns of Men's and Women's Names. *Journal of Linguistics*, *26*(2), 471–482.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1-12. doi:10.3758/s13428-014-0458-y.

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

Kawahara, S., Noto, A., & Kumagai, G. (2017). Sound symbolic patterns in Pokémon names.

Kawahara, S., Shinohara, K., & Grady, J. (2015). Iconic inferences about personality: From sounds and shapes. *Iconicity: East Meets West*, 57–69.

Klink, R. R., & Wu, L. (2014). The role of position, type, and combination of sound symbolism imbeds in brand names. *Marketing Letters*, *25*(1), 13-24.

Köhler, W. (1929). *Gestalt Psychology*.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940-967.

MacAuley, L., Siddiqi, D., & Toivonen, I. (in prep). The Age of Aidans: Cognitive underpinnings of a new trend in English boys' names.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: sound – shape correspondences in toddlers and adults. *Developmental Science*, 316–322.

Nielsen, A., & Rendall, D. (2011). The sound of round: evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *65*(2), 115.

New B., Pallier C., Ferrand L., Matos R. (2001) Une base de données lexicales du français contemporain sur internet: LEXIQUE, L'Année Psychologique, 101, 447-462. http://www.lexique.org

Newman, S. S. (1933). Further experiments in phonetic symbolism. *The American Journal of Psychology*, *45*(1), 53-75.

Ohala, J.J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. Phonetica, 41 1, 1-16.

Ohala, J.J. (2004). Sound Symbolism.

Paradis, C., & Prunet, J. F. (2000). Nasal vowels as two segments: Evidence from borrowings. *Language*, 324-357.

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Ramachandran, V. S., & Hubbard, E. M. (2001). Perception, Thought and Language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Retraite Québec. (2017). *List of baby names* [Data file]. Retrieved from http://www.rrq.gouv.qc.ca/en/enfants/banque_prenoms/Pages/banque_prenoms.aspx

ServiceOntario. (2016a). *Ontario top baby names (female)* [Data file]. Retrieved from https://www.ontario.ca/data/ontario-top-baby-names-female

ServiceOntario. (2016b). *Ontario top baby names (male)* [Data file]. Retrieved from https://www.ontario.ca/data/ontario-top-baby-names-male

Shrum, L. J., Lowrey, T. M., Luna, D., Lerman, D. B., & Liu, M. (2012). Sound symbolism effects across languages: Implications for global brand names. *International Journal of Research in Marketing*, *29*(3), 275-279.

Sidhu, D. M., & Pexman, P. M. (2015). What's in a Name? Sound Symbolism and Gender in First Names. *PLoS ONE*, *10*(5), 1–22. https://doi.org/10.1371/journal.pone.0126809

Slater, A. S., & Feinman, S. (1985). Gender and the Phonology of North American First Names. *Sex Roles*, *13*, 429–440.

Sutton, L. (2016). Aliens Are Just Like Us: Personal Names in The Legion of Super-Heroes. *Names*, *64*(2), 109-119.

Svantesson, J. (2017). Sound symbolism: the role of word sound in meaning. *WIREs Cogn Sci 2017*, *8* (October). https://doi.org/10.1002/wcs.1441

Tanz, C. (1971). Sound symbolism in words relating to proximity and distance. *Language and speech*, *14*(3), 266-276.

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and language*, *93*(1), 10-19.

Wright, S. K., Hay, J., & Bent, T. (2005). Ladies first? Phonology, frequency, and the naming conspiracy. *Linguistics*, *43*(3), 531–561. https://doi.org/10.1515/ling.2005.43.3.531

<u>Appendix A</u>

<u>Corpus of male and female French and English given Names</u>

Sources

- English names: ServiceOntario (2016a, b) baby name database – 2013
- French names: Retraite Québec (2017) baby name database – 2013

English Female Names

Olivia	ə.ˈlɪ.vi.ə	Julia	ˈdʒu.li.ə	Alice	ˈæ.ləs
Emma	ˈɛ.mə	Evelyn	ˈɛ.və.lən	Jasmine	ˈdʒæz.mən
Sophia	so.ˈfi.ə	Brooklyn	ˈbɹʊ.klən	Harper	ˈhɑɹ.pəɹ
Ava	ˈe.və	Alyssa	ə.ˈlɪ.sə	Rachel	ˈɹe.tʃəl
Charlotte	ˈʃɑɹ.lət	Anna	ˈæ.nə	Naomi	naj.ˈo.mi
Emily	ˈɛ.mə.li	Lauren	ˈlo.ɹən	Sydney	ˈsɪd.ni
Isabella	ɪ.zə.ˈbɛ.lə	Aria	ˈɑ.ri.ə	Clara	ˈklɛ.rə
Chloe	ˈklo.i	Zoey	ˈzo.i	Faith	ˈfeθ
Abigail	ˈæ.bə.gel	Mackenzie	mə.ˈkɛn.zi	Savannah	sə.ˈvæ.nə
Avery	ˈe.və.ri	Addison	ˈæ.də.sən	Georgia	ˈdʒoɹ.dʒə
Ella	ˈɛ.lə	Samantha	sə.ˈmæn.θə	Peyton	ˈpe.tən
Mia	ˈmi.ə	Violet	ˈvaj.lət	Nevaeh	nə.ˈve.ə
Hannah	ˈhæ.nə	Scarlett	ˈskɑɹ.lət	Nora	ˈno.ɹə
Victoria	vɪk.ˈto.ɹi.ə	Sadie	ˈse.di	Brooke	ˈbrʊk
Lily	ˈlɪ.li	Eva	ˈi.və	Paige	ˈpedʒ
Maya	ˈmaj.ə	Isabelle	ˈɪ.zə.bɛl	Jessica	ˈdʒɛ.sə.kə
Amelia	ə.ˈmi.li.ə	Alexis	ə.ˈlɛk.səs	Ivy	ˈaj.vi
Grace	ˈgɹes	Arianna	ə.ɹi.ˈæ.nə	Ellie	ˈɛ.li
Sophie	ˈso.fi	Lillian	ˈlɪ.li.ən	Katherine	ˈkæ.θɹən
Madison	ˈmæ.də.sən	Layla	ˈle.lə	Keira	ˈki.ɹə
Leah	ˈli	Stella	ˈstɛ.lə	Ruby	ˈɹu.bi
Hailey	ˈhe.li	Alexandra	æ.lək.ˈsæn.dɹə	Madeline	ˈmæ.də.lən
Sarah	ˈsɛ.rə	Gabriella	gæ.bɹi.ˈɛ.lə	Annabelle	ˈæ.nə.bɛl
Elizabeth	ə.ˈlɪ.zə.bəθ	Aubrey	ˈɑ.bɹi	Maria	mə.ˈɹi.ə
Audrey	ˈɑ.dɹi	Natalie	ˈnæ.ɾə.li	Lucy	ˈlu.si
Claire	ˈkleɹ	Quinn	ˈkwɪn	Chelsea	ˈtʃɛl.si

Aaliyah	æ.ˈli.ə	Kate	ˈket	Madeleine	ˈmæ.də.lajn
Nicole	nə.ˈkol	Kayla	ˈke.lə	Summer	ˈsʌ.məɹ
Riley	ˈɹaj.li	Rebecca	ɹə.ˈbɛ.kə	Tessa	ˈtɛ.sə
Sienna	si.ˈɛ.nə	Gabrielle	ˈgæ.bɹi.ɛl	Piper	ˈpaj.pəɹ
Taylor	ˈte.ləɹ	Alexa	ə.ˈlɛk.sə	Skylar	ˈskaj.ləɹ
Kylie	ˈkaj.li	Fiona	fi.ˈo.nə	Brielle	bɹi.ˈɛl
Serena	sə.ˈɹi.nə	Angelina	æn.gə.ˈli.nə	Kennedy	ˈkɛ.nə.di
Willow	ˈwɪ.lo	Eleanor	ˈɛ.lə.nəɹ	Kaitlyn	ˈket.lən
Autumn	ˈɑ.ɾəm	Molly	ˈmɑ.li	Mikayla	mə.ˈke.lə
Vanessa	və.ˈnɛ.sə	Brianna	bɹi.ˈæ.nə	Zara	ˈzɑ.ɾə
Kaylee	ˈke.li	Penelope	pə.ˈnɛ.lə.pi	Khloe	ˈklo.i
Lyla	ˈlaj.lə	Eden	ˈi.dən	Talia	ˈtæ.li.ə
Morgan	ˈmoɹ.gən	Elise	ə.ˈliz		

English Male Names

Liam	ˈli.əm	Oliver	ˈɑ.lə.vəɹ	Zachary	ˈzæ.kə.ɹi
Ethan	ˈi.θən	Samuel	ˈsæ.mu.əl	Luke	ˈluk
Lucas	ˈlu.kəs	Michael	ˈmaj.kəl	Cameron	ˈkæm.ɹən
Noah	ˈno.ə	Evan	ˈɛ.vən	Austin	ˈɑ.stən
Benjamin	ˈbɛn.dʒə.mən	Hunter	ˈhʌn.təɹ	Connor	ˈkɑ.nəɹ
Jacob	ˈdʒe.kəb	Gabriel	ˈge.bɹi.əl	Nolan	ˈno.lən
William	ˈwɪ.li.əm	Adam	ˈæ.dəm	Sebastian	sə.ˈbæs.tʃən
Nathan	ˈne.θən	Jayden	ˈdʒe.dən	Blake	ˈblek
Mason	ˈme.sən	Aiden	ˈe.dən	Tyler	ˈtaj.ləɹ
Alexander	æ.lək.ˈsæn.dəɹ	Thomas	ˈtɑ.məs	Gavin	ˈgæ.vən
Logan	ˈlo.gən	Dylan	ˈdɪ.lən	Xavier	ɛk.ˈze.vjəɹ
Daniel	ˈdæ.njəl	Nicholas	ˈnɪ.kə.ləs	Anthony	ˈæn.θə.ni
Owen	ˈo.wən	Andrew	ˈæn.dɹu	Elijah	ə.ˈlaj.ʒə
Jack	ˈdʒæk	Isaac	ˈaj.zək	John	ˈdʒɑn
James	ˈdʒemz	David	ˈde.vəd	Grayson	ˈgɹe.sən
Jackson	ˈdʒæk.sən	Henry	ˈhɛn.ɹi	Tristan	ˈtɹɪ.stən
Carter	ˈkɑɹ.təɹ	Joseph	ˈdʒo.səf	Chase	ˈtʃes
Joshua	ˈdʒɑ.ʃə.wə	Hudson	ˈhʌd.sən	Christian	ˈkɹɪs.tʃən
Matthew	ˈmæ.θju	Colton	ˈkol.tən	Lincoln	ˈlɪn.kən
Ryan	ˈɹaj.ən	Caleb	ˈke.ləb	Brayden	ˈbɹe.dən

Parker	ˈpɑɹ.kəɹ	Jace	ˈdʒes	Alex	ˈæ.ləks
Dominic	ˈdɑ.mə.nɪk	Hayden	ˈhe.dən	Myles	ˈmaj.əlz
Cole	ˈkol	Jason	ˈdʒe.sən	Simon	ˈsaj.mən
Charles	ˈtʃɑ.əɹəlz	Adrian	ˈe.dɹi.ən	Sawyer	ˈsɔj.əɹ
Marcus	ˈmɑɹ.kəs	Cooper	ˈku.pəɹ	George	ˈdʒoɹdʒ
Wyatt	ˈwaj.ət	Eli	ˈi.laj	Peter	ˈpi.ɾəɹ
Aaron	ˈɛ.ɹən	Carson	ˈkɑɹ.sən	Jeremy	ˈjɛ.ɹə.mi
Jordan	ˈdʒoɹ.dən	Easton	ˈi.stən	Reid	ˈrid
Nathaniel	nə.ˈθæ.njəl	Robert	ˈɹɑ.bəɹt	Vincent	ˈvɪn.sənt
Christopher	ˈkɹɪ.stə.fəɹ	Spencer	ˈspɛn.səɹ	Harrison	ˈhɛ.ɹə.sən
Jonathan	ˈdʒɑ.nə.θən	Jake	ˈdʒek	Patrick	ˈpæ.tɹək
Leo	ˈli.o	Charlie	ˈtʃɑɹ.li	Mark	ˈmæɹk
Landon	ˈlæn.dən	Justin	ˈdʒʌ.stən	Elliot	ˈɛ.li.ət
Emmett	ˈɛ.mət	Eric	ˈɛ.ɹɪk	Felix	ˈfi.ləks
Bentley	ˈbɛnt.li	Riley	ˈɹaj.li	Bennett	ˈbɛ.nət
Levi	ˈli.vaj	Kyle	ˈkaj.əl	Asher	ˈæ.ʃəɹ
Julian	ˈdʒu.li.ən	Tyson	ˈtaj.sən	Ashton	ˈæʃ.tən
Max	ˈmæks	Wesley	ˈwɛs.li	Cohen	ˈko.hən
Maxwell	ˈmæks.wɛl	Brandon	ˈbɹæn.dən	Omar	ˈo.mɑɹ
Ryder	ˈɹaj.dəɹ	Brody	ˈbɹo.di	Ryker	ˈɹaj.kəɹ
Isaiah	aj.ˈze.jə	Theodore	ˈθi.ə.doɹ		

French Female Names

Léa	le.a	Béatrice	be.a.tʀis	Élodie	e.lo.di
Emma	ɛ.ma	Maeva	ma.e.va	Raphaëlle	ra.fa.ɛl
Florence	flɔ.ʀɑ̃s	Laurence	lɔ.ʀɑ̃s	Mégane	me.gan
Alice	a.lis	Maëlie	ma.ɛ.li	Aurélie	ɔ.ʀe.li
Zoé	zo.e	Sarah	sa.ʀa	Sophie	so.fi
Rosalie	ʀo.za.li	Noémie	no.e.mi	Alexia	a.lɛk.si.a
Juliette	ʒu.li.ɛt	Elizabeth	e.li.za.bɛt	Emy	ɛ.mi
Chloé	klo.e	Rose	ʀoz	Léonie	le.o.ni
Charlotte	ʃaʀ.lɔt	Coralie	co.ʀa.li	Jasmine	ʒas.min
Jade	ʒad	Annabelle	a.na.bɛl	Maïka	ma.i.ka
Camille	ka.mij	Ève	ɛv	Océane	o.se.an
Anaïs	a.na.is	Gabrielle	ga.bʀi.ɛl	Émilie	ɛ.mi.li

Marianne	ma.ʀi.an	Amélie	a.me.li	Roxanne	ʀok.san
Justine	ʒu.stin	Évelyne	ɛ.və.lin	Anne-Sophie	an.so.fi
Ariane	a.ʀi.an	Julianne	ʒu.li.an	Isabelle	i.za.bɛl
Flavie	fla.vi	Simone	si.mon	Clémence	kle.mɛns
Jeanne	ʒan	Alyson	a.lɪ.sən	Maryam	ma.ʀi.am
Mélodie	me.lo.di	Laurianne	lo.ʀi.an	Mélina	me.li.na
Adèle	a.dɛl	Amy	e.mi	Violette	vi.o.lɛt
Éliane	e.li.an	Maëlle	ma.ɛ.li	Constance	kɔ̃.stɑ̃s
Delphine	dɛl.fin	Lauralie	lo.ra.li	Anaeve	a.na.ɛv
Magalie	ma.ga.li	Catherine	ka.tʀin	Élianne	ɛ.li.an
Daphnée	daf.ne	Estelle	ɛ.stɛl	Madeleine	ma.də.lɛn
Mathilde	ma.tild	Marguerite	maʀ.ga.ʀit	Sandrine	sɑ̃.drin
Marilou	ma.ri.lu	Ariel	a.ʀi.ɛl	Lya	li.a
Frédérique	fʀɛ.dɛ.ʀik	Eléonore	ɛ.le.o.noʀ	Alexane	a.lɛk.san
Lily	li.li	Pénélope	pɛ.nɛ.lop	Axelle	ak.sɛl
Arielle	a.ʀi.ɛl	Agathe	a.gat	Lara	la.ʀa
Éloïse	ɛ.lo.iz	Inès	i.nɛs	Liliane	li.li.an
Léa-Rose	le.a.ʀoz	Myriam	mi.ʀi.am	Lili-Rose	li.li.ʀoz
Ophélie	o.fe.li	Rachel	ʀa.ʃɛl	Sarah-Maude	sa.ʀa.mod
Naomie	na.o.mi	Marion	ma.ʀi.ɔ̃	Anaëlle	a.na.ɛl
Alexandra	a.lɛk.san.dʀa	Abigaëlle	a.bi.ga.ɛl	Emy-Rose	ɛ.mi.ʀoz
Léane	le.an	Elie	ɛ.li	Esther	ɛ.stɛʀ
Maxim	mak.sim	Emmanuelle	ɛ.ma.nu.ɛl	Morgane	moʀ.gan
Rébecca	ʀe.bɛ.ka	Romane	ʀo.man		

French Male Names

Nathan	na.tɑ̃	Noah	no.a	Zachary	za.k.ʀi
Samuel	sa.mu.ɛl	Xavier	gza.vi.e	Alexandre	a.lɛk.sɑ̃dʀ
Alexis	a.lɛk.si	Benjamin	bɛ̃.ʒa.mɛ̃	Lucas	lu.ka
Olivier	o.li.vi.je	Adam	a.dɑ̃	Théo	te.o
Thomas	to.ma	Émile	e.mil	Victor	vik.tɑr
Félix	fe.liks	Charles	ʃaʀ.lə	Loïc	lo.ik
Gabriel	ga.bʀi.ɛl	Léo	le.o	Justin	ʒu.stɛ̃
Antoine	ɑ̃.twan	Mathis	ma.tis	Nicolas	ni.ko.la
Raphaël	ʀa.fa.jɛl	Édouard	ed.waʀ	Louis	lu.i

Tristan	tʀi.stɑ̃	Mathias	ma.ti.as	Bastien	ba.stjɛ̃
Jeremy	ʒɛ.ʀi.mi	Mathieu	ma.ti.œ	Jason	dʒe.sən
Vincent	vɛ̃.sɑ̃	Louka	lu.ka	Maxence	mak.sɑ̃s
Étienne	e.tjɛn	Jordan	ʒoʀ.dɑ̃	Samy	sa.mi
Ethan	i.θən	Émrick	ɛ.mʀik	Charles-Olivier	ʃaʀl.o.li.vje
Arthur	aʀ.tuʀ	Jules	ʒul	Clément	kle.mɑ̃
Philippe	fi.lip	Cédric	sɛ.dʀik	Léon	le.ɔ̃
David	da.vid	Élie	e.li	Frédéric	fʀɛ.dɛ.ʀik
Alex	a.lɛks	Léonard	le.o.naʀ	Sebastian	se.ba.sti.an
Arnaud	aʀ.no	Guillaume	gi.jom	Lohan	lo.an
Maxime	mak.sim	Jonathan	jɑ̃.na.tɑ̃	Adrien	a.dʀi.ɛ̃
Laurent	lɔ.rɑ̃	Joseph	ʒo.zef	Paul	pɑl
Simon	si.mɔ̃	Rémi	ʀe.mi	Albert	al.bɛʀ
Henri	ɑ̃.ʀi	Théodore	te.o.dɔʀ	Louis-Félix	lu.i.fe.liks
Damien	da.mi.ɛ̃	Nathaniël	na.ta.ni.el	Louis-Philippe	lu.i.fi.lip
Mathéo	ma.te.o	Charles-Antoine	ʃaʀl.ɑ̃.twan	Romain	ʀo.mɛ̃
Ludovic	lu.do.vik	Christophe	kʀi.stɔf	Christian	kʀi.sti.ɑ̃
Michaël	mi.ka.el	Emmanuel	e.ma.nu.ɛl	Marc-Antoine	maʀk.ɑ̃.twan
Hubert	u.bɛʀ	Jérôme	ʒe.ʀom	Yohan	jo.an
Éloi	ɛ.lo.a	Lyam	li.am	Renaud	ʀœ.no
Julien	ʒu.li.ɛ̃	Maël	ma.ɛl	Gaël	ga.ɛl
Daniel	da.njel	Félix-Antoine	fe.liks.ɑ̃.twan		

Appendix B

Informed Consent Form

LINGUISTIC EXPERIMENT

Consent for Participation in a Research Project

Invitation to Participate and Description of Project

You are invited to participate in a linguistic experiment. You are free to decline to participate if you wish. In order to decide whether or not you wish to be a part of this research study you should know enough about its general purpose, risks and benefits to make an informed judgment. This consent form gives you detailed information about the procedures of the experiment, as well as any risks of the procedures and any possible benefits of these studies. Once you understand the procedures, you will be asked to indicate if you wish to participate by selecting a radio button next to a statement indicating whether or not you consent to participate.

Purpose of Research

The principal researcher is Lisa Sullivan, an M.A. student in the Department of Linguistics at the University of Toronto. She is supervised by Professor Yoonjung Kang, a faculty member of the Centre for French and Linguistics at UTSC and of the Department of Linguistics at the University of Toronto. The general purpose of the research is to understand the relationship between language and mind.

Description of the Procedure

The experiment will consist of three tasks. The first task will be a language use questionnaire. During the second task, you will hear names and be asked to indicate whether you think they are male or female names and how confident you are in that regard. During the third task, you will hear two sounds and be asked to indicate if the second sound was the same as the first one. You will need headphones to complete the experiment. The experiment is expected to take 30 minutes to complete.

Risks and Inconveniences

The procedure involves no known risk or discomfort.

Benefits

This study will provide no benefit to you but will provide us with information that may lead to the future benefit of others.

Compensation

You will receive a $5 Amazon gift card as compensation for 30 minutes of your time. Furthermore, you should know that if you fail to complete the full experimental procedure, you will, nonetheless, receive your compensation in proportion to the tasks you completed.

Confidentiality

Your name will not be recorded and, therefore, will not appear in subsequent scholarly presentations or publications. Other personal information gathered in this study will not be disclosed to any persons other than the investigators and their collaborators. Furthermore, effort will be made to keep this information secure and to destroy it once the results of the study have been published. Following the standard practice in linguistics research, recorded linguistic data, however, will not be disposed (unless by specific request of the participant), as its preservation is necessary for further studies involving wider samples of speakers or for studies examining linguistic change over time.

The research study you are participating in may be reviewed for quality assurance to make sure that the required laws and guidelines are followed. If chosen, (a) representative(s) of the Human Research Ethics Program (HREP) may access study-related data and/or consent materials as part of the review. All information accessed by the HREP will be upheld to the same level of confidentiality that has been stated by the research team.

Voluntary Participation

You are free to choose not to participate, and, if you do become a participant, you are free to withdraw from the experiment at any time during its course by closing the browser window. You can expect to receive compensation, relative to the portion of the experimental tasks you completed, within one week of withdrawing. If you wish to withdraw following the completion of the experiment, you may do so by emailing the investigators with your intent to withdraw and your participation code. Should you terminate your participation, you may do so without jeopardizing any of the following that may apply to you, namely, any opportunities to serve as a participant in future experiments or your standing as a student if you are a student.

If you have further questions about this project, please contact the principal investigator, Lisa Sullivan, lisa.sullivan@mail.utoronto.ca, or the research supervisor, Dr. Yoonjung Kang, 416-287-7172 or kang@utsc.utoronto.ca. If you have any questions about your rights as a research participant, please contact the Office of Research Ethics, 416-946-3273 or ethics.review@utoronto.ca.

Authorization

- I have read this form and decided that I will participate in the experiment described above. Its general purpose, procedure, possible risks and benefits have been explained to my satisfaction.
- I have read this form and decided that I will not participate in the experiment described above.

Appendix C

Language Background Questionnaire

The questionnaire is based on Chambers & Lapierre's (2011) Language Use Index (LUI) and Marian, Blumenfeld & Kaushanskaya's (2007) Language Experience and Proficiency Quetionnaire (LEAP-Q).

English

1. How often do you use English in the following ways?

At home	Always	Often	Seldom	Never
At work or school	Always	Often	Seldom	Never
With your friends	Always	Often	Seldom	Never
With your relatives	Always	Often	Seldom	Never
In the media	Always	Often	Seldom	Never

2. For any time you have lived in an English-speaking community for more than three months, please list the country, how old you were and how long you lived there.

3. What is your level of education in English? Please indicate the number of years you studied English at the primary, secondary and post-secondary levels and the degree of immersion (native, immersion, language classes only)

French

1. How often do you use French in the following ways?

At home	Always	Often	Seldom	Never
At work or school	Always	Often	Seldom	Never
With your friends	Always	Often	Seldom	Never
With your relatives	Always	Often	Seldom	Never
In the media	Always	Often	Seldom	Never

2. For any time you have lived in a French-speaking community for more than three months, please list the country, how old you were and how long you lived there.

3. What is your level of education in French? Please indicate the number of years you studied English at the primary, secondary and post-secondary levels and the degree of immersion (native, immersion, language classes only)

Other Languages

Please list any other languages you speak and your level of proficiency (beginner, intermediate, advanced, native)

Appendix D

Additional Tables and Graphs of Experimental Results

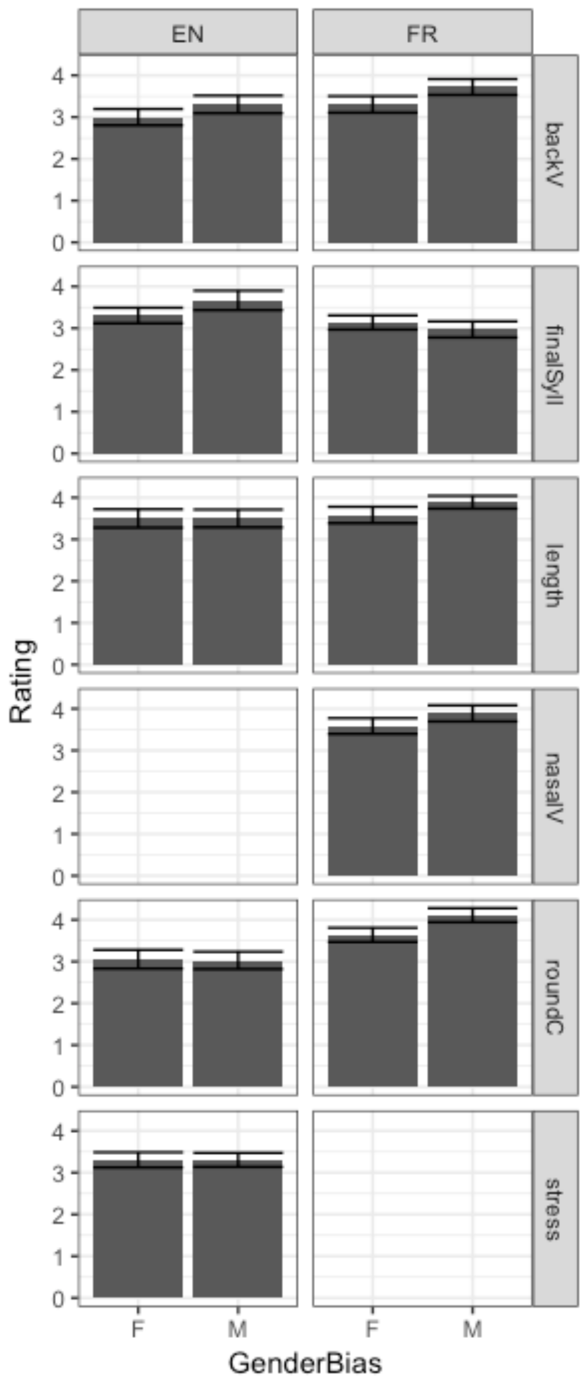

Figure 1 Mean ratings of male- and female-biased names by language and factor

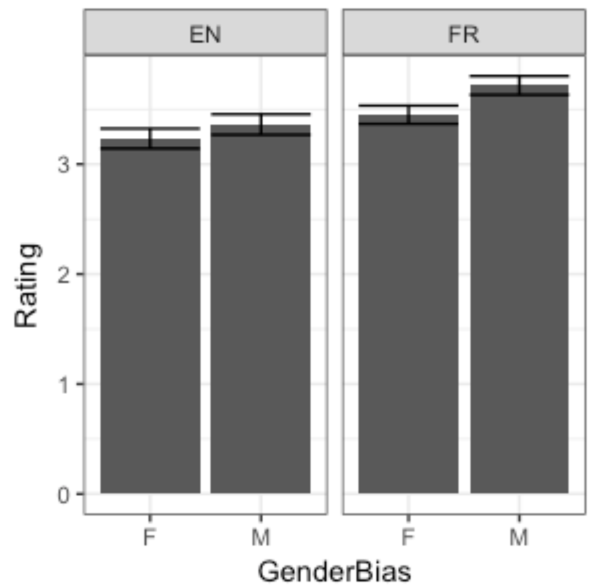

Figure 2 Mean ratings of male- and female-biased names by language

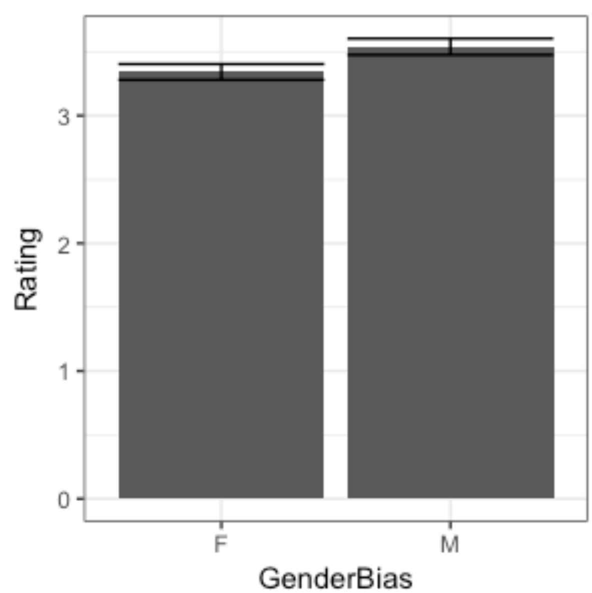

Figure 3 Mean ratings of male- and female-biased names

Table 1

Mean ratings and standard deviations for male-biased and female-biased names by language

	Female		Male	
Language	*M*	*SD*	*M*	*SD*
English	3.23	1.21	3.36	1.25
French	3.45	1.12	3.72	1.14

Table 2

Mean ratings and standard deviations for male-biased and female-biased names by factor

	Female		Male	
Factor	*M*	*SD*	*M*	*SD*
Presence of a Back Vowel	3.15	1.18	3.51	1.21
Final Syllable Type	3.22	1.06	3.32	1.32
Number of Syllables	3.54	1.24	3.69	1.12
Presence of a Nasal Vowel	3.58	1.13	3.89	1.17
Presence of a Round Consonant	3.47	1.22	3.57	1.25
Stress Placement	3.31	1.09	3.31	1.01

Appendix E

Results of Full Linear Regression Model for Experiment

Table 1

Summary of coefficients in the full mixed effects linear regression model for gender with gender bias, factor, language, the interaction between gender bias and factor and the interaction between gender bias and language as predictor variables and stimulus pair and participant (random slope = gender bias) as random effects

	Estimate	Standard Error	t-value	Pr(>|t|)
(intercept)	3.05208	0.34641	8.811	<0.001***
Gender Bias (M)	0.31250	0.19426	1.609	0.108
Factor – Final Syll	0.06944	0.42608	0.163	0.873
Factor – Length	0.38889	0.42608	0.913	0.375
Factor – Nasal V	0.32986	0.54315	0.607	0.552
Factor – Round C	0.19444	0.42608	0.456	0.654
Factor – Stress	0.25347	0.54315	0.467	0.647
Language (FR)	0.20139	0.30129	0.668	0.514
Gender (M)*Factor (Final Syll)	-0.26389	0.24567	-1.074	0.283
Gender (M)*Factor (Length)	-0.20833	0.24567	-0.848	0.397
Gender (M)*Factor (Nasal V)	-0.10417	0.31316	-0.333	0.740
Gender (M)*Factor (Round C)	-0.13889	0. 24567	-0.565	0.572
Gender (M)*Factor (Stress)	-0.31250	0.31316	-0.998	0.319
Gender (M)*Language (FR)	0.09722	0.17371	0.560	0.576