Underexplored Queer Voices: Constructing Transfeminine Identities in Toronto English

Morgan Terry

Supervisors: Ruth King Ph.D, Chandan Narayan Ph.D

Submitted: July 30th, 2026

Introduction

This paper explores the phenomenon of the "gay voice", a topic that has been discussed in sociolinguistics for decades and is also a relatively well known linguistic phenomenon in everyday life. The gay man with a lisp or a "queeny" voice is a well known stereotype, but most people do not know what makes a gay man's voice sound "gay", only that they can distinguish between it and the average straight man's voice. However, while there has been a lot of research on the voices of gay men and what linguistic variables are being used to make their voices sound more queer, there has been significantly less research on the voices of queer women. Sociolinguistic research from the 2000s to the 2020s has determined that the most common variable that gay men use to produce this type of voice is /s/ center of gravity, but there is still plenty of debate on whether there is even a "gay voice" for queer women.

There has been some research on the voices of queer women, like Moonwomon-Baird (1997) or Bucholtz and Hall (2004), but this topic has received a lot less attention than the voice of gay men. In addition to this lack of research on the speech of lesbians, there has been even less research on the differences between the speech of different types of queer women. With the field of sociolinguistics advancing and a lot of these old studies on lesbian speech not being up to date, it seems necessary to fill this gap, while also bringing in another group in the queer community that has received very scarce attention: trans women and transfeminine people.

In the past decade and a half, research on trans voices has increased significantly, as trans people complicate the ideas that people have had for a long time about the connections between gender and the voice. One area of research that has become popular is examining the relationship between trans people's voices and their gender presentations, since gender identity and gender presentation are very salient for trans people. One of the most well known studies in this area is

Zimman (2017a), who investigated trans masc individuals and examined how their voices changed over time when on Hormone Replacement Therapy. However, just like how the voices of gay men have received much more attention in research compared to lesbian women, there has been a similar problem with trans men and transmasculine people receiving a lot more attention than trans women and transfeminine people.

This gap in both the research in the voices of gay and bisexual women as well as the voices of trans women and transfeminine people inspired the idea behind this research project. If there is a gay voice for gay men and a lesbian voice for lesbians, what would this look like for queer and straight trans women and transfeminine people? Would there be any similar type of voice based on sexuality, when countless previous research has found the same for gay men, and to a lesser extent lesbians. However, rather than just looking at the sexuality of the speakers, examining the gender presentation might be productive as well. Gender identity and gender presentation are very salient for trans speakers, so examining the intersection of gender presentation and sexuality in their voice might illuminate how these individual variables have an effect.

This research project will investigate whether there is phonetic variation within the transfeminine communities based on the sexuality and gender presentation of the speaker. Here, gender presentation refers to an individual's expression of their gender identity, which often manifests through clothes, hairstyles, makeup, but the voice can also be used as a resource to express one's gender presentation. This expression is often either an expression of femininity or masculinity; however, some people also try to express more of an androgynous or neutral gender presentation. To investigate these issues, there will be an examination of specific phonetic variables. The phonetic variation being examined in the present study will be the center of

gravity for /s/, as well as the F1 and F2 formants for five vowels, which are /ɑ/, /æ/, /ɛ/, /i/, and /o/. First, an examination of the phonetic variation based on the speaker's membership in broader social categories, like sexuality, will be examined.

This investigation into the macro level interactions and whether there is phonetic variation to be seen solely based on the sexuality of the speakers will be determined. Next, a more micro level investigation will be conducted, which will examine if there is phonetic variation based on the gender presentation of the speaker. While not as broad a social category as sexuality, gender presentation is still an important part of one's identity. After this, the paper will look at how these two variables, sexuality and gender presentation, interact and whether this interaction leads to any phonetic variation among the speakers.

Finally, a transcription analysis will be undertaken, and will be used to see the most micro level interactions that the speakers engage in and whether they use these phonetic variables as resources to index their sexuality and gender presentation. As has been shown in previous research (Campbell-Kibler, 2007; Podesva and Van Hofwegen, 2016; Zimman 2017a), /s/ center of gravity is associated with a wide variety of meanings, including sexuality and gender presentation. Munson et. al (2005) also used F1 and F2 formants to examine sexuality. However, it does not seem like any previous research has examined both /s/ center of gravity and F1/F2 formants as indexing both sexuality and gender presentation directly. By examining whether the speakers use these variables to index their sexuality and gender presentation, it will help reveal the process of identity construction that the speakers are engaging in. Previous work has largely not examined the intersection of sexuality and gender presentation among trans women and transfeminine people, so the aim is to hopefully shed light on a topic that has gone neglected.

Literature Review

Style and Stylistic Bricolage

Key to this investigation is the concept of style and stylistic practice. Eckert (2003) explains that “style (like language) is not a thing but a practice. It is the activity in which people create social meaning, as style is the visible manifestation of social meaning. And inasmuch as social meaning is not static, neither are styles" (p. 43). Previous studies, like Zimman (2017a), have examined stylistic practices in queer communities and argue that focus should be placed on utilizing concepts like sociolinguistic style and stylistic bricolage to explain differences in the voice based on gender. This process of stylistic bricolage, a term that Eckert (2003) uses to describe how "stylistic practice involves a process of bricolage, by which people combine a range of existing resources to construct new meanings or new twists on old meanings" (p. 43), is what leads to the reanalysis of social meanings on linguistic resources.

In this project, the main goal is examining whether speakers are engaging in this process by taking the existing meanings that are placed on the linguistic variables, which are center of gravity and F1/F2, reconstructing these meanings, and then using these phonetic variables to index these new meanings. However, one important distinction to make is that the use of stylistic bricolage in analysis does not mean that speakers are aware of their combination of linguistic resources into styles. Zimman (2017a) makes the argument that this process is not always something that speakers are aware of. He says, “the theory itself does not require intentionality on the part of the speaker… The core function of the concept of bricolage is to help us understand how meaning is constructed, not to make claims about the degree to which speakers are conscious of the process” (Zimman, 2017a, p. 366). This distinction is vital to my analysis, as my claim is not that all of the speakers are actively and knowingly engaging in this process of

indexing their gender presentation and sexuality. This investigation examines whether there is phonetic variation among these speakers, regardless of whether they are intentionally using these linguistic variables or not.

Another crucial concept is how such meanings vary among speakers. Eckert (2008) proposes that the meanings of variables are not fixed and immutable but instead are related by an indexical field, which she defines as a "constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable" (p. 453). While a linguistic variable, such as /t/ release, may convey a specific meaning for one group, it can hold different meanings for other groups or for specific individuals. If one of the speakers in this investigation chooses to use one of the vowels to index their sexuality and gender presentation, this does not mean that these meanings are the same for every speaker.

There is not a static set of meanings that each variable holds and are only activated when chosen by the speaker. Instead, each speaker engages in the process of stylistic bricolage by taking the pre-existing meanings attached to these variables and reconstructing them for their own purposes. However, one important question is where these pre-existing meanings attached to these variables originate. To answer that question, the concept of sound symbolism needs to be introduced.

Sound Symbolism

The most important concept for this investigation is the concept of sound symbolism, which is the idea that there is a link between speech sounds and the meanings of the words that these sounds convey. Levon (2017) discusses this when talking about the meanings attached to /s/, and talks about how the assumption in modern linguistics is “that the link between a signified and its signifier is arbitrary” (p. 984). Sound symbolism is in opposition to this belief, as it

argues that there is a connection between the signifier (the speech sounds used) and the signified (the meaning of a word), or in other words, a connection between the phonetic form and the meaning conveyed by this form.

Hinton et. al (1994) describes four types of sound symbolism, which are corporeal sound symbolism, imitative sound symbolism, conventional sound symbolism, and synesthetic sound symbolism. Levon (2017) says that corporeal sound symbolism, "involves both segmental and intonational patterns that express the physical and emotional states of speakers, including so-called "symptomatic" sounds, like *achoo*, and expressive elements of prosody and voice quality" (p. 984). Imitative sound symbolism is often referred to as onomatopoeia, which are sounds that attempt to mimic other sounds, such as sounds from animals, like *meow*, or inanimate sounds, like *splash*. Conventional sound symbolism does not refer to the phonetic form, but it instead refers to how certain phonetic forms have come to be associated with certain meanings. An example from Hinton et. al (1994) comes from English words that begin with [gl], which have become associated with words referring to brightness, such as glow or glimmer. Hinton et. al (1994) argued that these clusters of sounds, such as [gl], became associated with specific meanings and then spread out to words with similar semantic domains.

The final type of sound symbolism, and the type that is most relevant to this paper, is synesthetic sound symbolism. Levon (2017) states that synesthetic sound symbolism is "the association of sounds with properties in the world" (p. 984) and that "the most common form of synesthetic sound symbolism is magnitude symbolism, that is the perception that certain sounds are naturally expressive of size distinctions (i.e., large versus small)" (p. 984). Ohala (1984; 1994) says that the primary phonetic correlates of magnitude symbolism are formant frequencies for vowels (F1/F2) and the fundamental frequency of pitch (F0) (p. 985). He also states that there

is a general tendency for phonetic tokens with higher frequency (such as front vowels) to carry the meaning of "smallness" while phonetic tokens with (such as back vowels) to carry the meaning of "bigness" (Ohala, 1984; 1994).

Another important concept related to sound symbolism is the idea of informational and affective meanings, which was proposed by Gussenhoven (2002). He proposed that informational meanings convey the message being communicated by the phonetic form (such as emphasis or when an utterance is ending) while affective meanings communicate perceived stances of the speaker (such as friendliness). Silverstein (1994; 2003)'s approach uses these affective meanings and describes how they can vary cross-linguistically. He describes how this sound symbolic system works in Wasco, which is an indigenous language of the Americas, where the augmentative/diminutive system does index meanings such as "bigness" and "smallness", but also indexes further indexical relationships, such as "impersonal" versus "personal" (Silverstein, 1994). In addition, in Wasco, these diminutive and augmentative systems seemed to correlate with the vowel formant frequency variation described by Ohala (1984; 1994). In Wasco, fronting vowels was associated with meaning diminutive, and backing was associated with meaning augmentative.

Silverstein (2003) expands upon this idea of the complex indexical system by introducing the idea of the indexical order, which demonstrates the process that linguistic signs go through to acquire meanings. This order of indexicality is organized in degrees, including first-order indexicality or second-order indexicality. When looking at sound symbolism in Wasco, this order of indexicality can be applied to the layers of indexical meanings that are attached to linguistic signs. At the level of *n*, which is the first order, the meaning attached to these front vowels is "smallness" while the meaning attached to the back vowels is "bigness". At the level of *n+1*,

which is the second order, these same linguistic signs can index further meanings, such as "impersonal" or "personal" (Silverstein, 1994).

There are further orders of indexicality, such as *n+1+1*, but the main point here is that these meanings that come from sound symbolism, like "smallness" and "bigness" being attached to vowel formant frequency, seems to most often occupy this *n* level of meaning and seems to exist across languages and cultures, as described by Ohala's Frequency Code (Ohala, 1984; 1994). However, the meanings that occupy this *n+1* level differ cross-linguistically and cross-culturally, as shown by the Wasco example where the same linguistic signs that index the *n* level meanings of "smallness" versus "bigness" for front and back vowels also index the *n+1* meanings of "personal" versus "impersonal".

There are further examples of linguistic meanings that have their origin in sound symbolism across cultures. Joseph (1994) found that coronal affricates in Modern Greek were found in words that related to small things, such as little girl, as well as things that sting, such as nettle. He argued that these meanings exist as part of a "relatedness network", where the meanings are in relation to each other and that the sound symbolism meaning for affrication is how the *n* level of meaning was first created (Joseph, 1994). Similarly, Hamano (1994) found that the palatalization of alveolar consonants in Japanese has the meaning of "childishness" and then further meanings such as "instability" and "unreliability".

This "relatedness network" that Joseph (1994) commented on is very similar to the concept of the indexical field that Eckert (2008) discussed, which was built off of Ochs (1992)'s work on direct and indirect indexicality. As previously stated, Eckert (2008)'s idea of the indexical field is one of a "constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable" (p. 453). These concepts of the indexical field are

informed by Silverstein (2003)'s idea of the indexical order. All of these meanings that exist in first-order indexicality (*n*), second-order indexicality (*n+1*), and so on, all belong to this indexical field, which can be activated by a speaker at any time. The connection between sound symbolism, indexical order, and the indexical field is that these meanings come from sound symbolism as "smallness" or "bigness" are associated with front and back vowels, and exist on the level of *n*. However, more complex meanings that stem from this dichotomy exist within these other orders of indexicality, and all of these meanings exist at the same time but have to be activated by a speaker first.

An important paper that investigates phonetic variation in speech and the influence that sound symbolism has on this variation is Eckert (2010)'s study of indexicality in front and back vowels. This study was based on the ethnographic work that Eckert did in two schools in San Jose, California, which followed groups of kids from fifth to eighth grade. Here, Eckert finds that there was systemic variation of the vowels /o/ and /ay/ based on the type of personae that the adolescent girls in these groups wanted to embody. In particular, when the girls wanted to present themselves as "nice", "friendly", and "positive", their productions of the /o/ and /ay/ vowels were significantly fronted and lowered (Eckert 2010). When the girls wanted to present themselves as "cynical", "negative", or "mature", the vowels were backed and raised (Eckert, 2010). As she states, "the meanings of fronting and backing of these low vowels constitute a range, or field, of affect that is both coherent and culturally specific" (Eckert 2010, p.78).

As this section has outlined, there is a broad range of meanings associated with speech sounds that have their origins in sound symbolism. In the present study, the vowel frequency formants of the speakers will be measured, by examining whether speakers are manipulating these variables or not. However, fronting or backing these vowels and /s/ productions does not

mean that speakers are trying to index the meanings at the *n* level or even the *n+1* level or further beyond that. Instead, through the process of stylistic bricolage, these speakers will instead reconstruct such meanings attached to these forms through these previous orders of indexicality and attaching their own meanings. Even if there is no phonetic variation based on categories like sexuality, these speakers may instead be choosing to do this with their vowels. However, if these speakers are reconstructing the meanings attached to these forms and creating new ones, what are they choosing to index with these new meanings? My hypothesis is that these speakers are indexing locally constructed categories that relate to their gender presentation or sexuality, and that these locally constructed categories do not solely originate from their local communities of practice, but instead from the "imagined trans community" these speakers all belong to.

Imagined Communities

The concept of the imagined community originated from Anderson (1983). In Anderson (1983), the term *imagined community* is first used to explain that nations are imagined communities, "because the members of even the smallest nation will never know most of their fellow-members, meet them, or even hear of them, yet in the minds of each lives the image of their communion" (p. 6). This characteristic that most of the members of this community will have never met, yet still feel this shared sense of identity and belonging to the community, is the core argument behind the term imagined community.

While this term originated in the field of political science, it has been used by sociolinguistics as well. In Zimman (2017a), his participants were not in a single community of practice, however he argues that they can instead be visualized as an imagined community, building on Anderson's concept. Zimman argues that this notion can be extended beyond the conceptualization of the nation and that, "Anderson (1983)'s notion of the imagined community

is one way to conceptualize the sense of community that can be shared by people who have never met" (p. 348). He also cites Valentine (2007)'s imagining of "the transgender community" to envision his participants as belonging to one imagined transgender community, which they positioned themselves within.

Kanno & Norton (2012) explain imagined communities as "groups of people, not immediately tangible and accessible, with whom we connect through the power of imagination" (p. 242). Following Anderson (1983)'s definition of the nation as an imagined community, Kanno & Norton (2012) argue that "the notion of imagined communities provides a theoretical framework for the exploration of creativity, hope, and desire in identity construction" (p. 248). While imagined communities are not necessarily communities of practice, as they are not defined by engagement in shared common social practices, this does not prevent analysis of the stylistic practices that members of an imagined community engage in, as Zimman (2017a) shows. Following Zimman (2017a)'s use of these frameworks, this paper will also be analyzing the participants as part of a shared imagined trans community.

As all of the participants in this study are part of this shared imagined trans community, the knowledge of locally constructed categories that relate to their gender presentation and sexuality originate there. Within this community, common shared knowledge about certain categories, like butch, femme, or "doll", is shared through in-person social networks and social media. While the social practices of individuals in this community may vary significantly, since it does not have the shared social practices that define a typical community of practice, the meanings that the members of this community have access to are shared. For example, there was a locally constructed category that every speaker in this research project had knowledge of. This category is called "the dolls", who are straight trans women that have a hyperfeminine gender

presentation. Whether it was through in-person social networks or social media, they were aware that such a group existed.

Even though none of these speakers had ever met each other, they all knew about this locally constructed category due to their membership in this imagined trans community and the common knowledge in this community. While each individual speaker may have belonged to their own community of practice in Toronto, they all still identified with belonging to "the trans community" and they all had access to this imagined community. The primary claim of this paper is that these speakers are indexing locally constructed categories that relate to their gender presentation or sexuality, and that the process of stylistic bricolage is taking place, where the speakers are changing the old meanings associated with these variables by attaching new ones.

The phonetic variables

The acoustic correlates being analyzed for this study include /s/ center of gravity and the F1 and F2 of /ɑ/, /æ/, /ɛ/, /i/, and /o/. The /s/ center of gravity was selected as a variable due to its long history as the primary phonetic variable associated with gay men's speech. However, it has also been associated with a lot of other meanings, such as class and age in Glaswegian English (Stuart-Smith, 2007), place and region (Campbell-Kibler, 2007; Podesva and Von Hofwegen, 2016), and sexuality (Munson, 2005; 2007). In addition, Zimman (2013; 2017) studied /s/ variation among trans men in San Francisco and found that it is a high frequency /s/ and more fronted /s/ that are linked with a "gay sounding" voice in trans men. A higher frequency and fronted /s/ means that the center of gravity is higher. As reported by Flipsen et. al (1999), the average center of gravity range for cisgender English-speaking men is 4000 - 7000 Hz, and the range for cisgender women is 6500 Hz - 8100 Hz. In Zimman (2017), the speakers' center of

gravity ranged from the lower end for men to the higher end for women, with speakers who had an average /s/ center of gravity at the higher end of the spectrum often identifying as gay and having a more feminine gender presentation. Overall, there is a long history of a higher frequency /s/ center of gravity being associated with gay sounding voices for men, which is why this acoustic correlate is being used in this study.

For the other acoustic correlate, the vowel formant frequencies for /ɑ/, /æ/, /ɛ/, /i/, and /o/, the main inspiration behind that is Munson et. al (2005)'s paper. While /s/ skewness is examined here, the formant frequency for these five vowels is also examined. This seems to be inspired by Pierrehumbert et. al (2004), who used /i/, /ɛ/, /æ/, /ɑ/, and /u/ as stimuli and measured duration, F1, and F2. In Munson et. al (2005), a larger variety of vowels were chosen, so to simplify the stimuli for this study, a mix of Pierrehumbert et. al (2004) and Munson et. al (2005)'s stimuli was used, which came out to be ɑ/, /æ/, /ɛ/, /i/, and /o/ to represent the entirety of the vowel space. Another reason why vowels were used in this study is the findings from Ohala (1984; 1994)'s Frequency Code, which showed that vowel formant frequencies (F1 and F2) seemed to be the primary acoustic correlate for sound symbolism, and Eckert (2010), who found that these front and back vowels were indexing a wide range of meanings. These previous studies have found that these vowels seem to already hold some meanings at the *n* level from sound symbolism and are used further at the *n+1* levels across many different languages, as well as in different social situations. That means that these speakers would already have familiarity using these vowels to index the *n* level meanings and beyond, so it is more likely that they would choose these vowels and reanalyze these meanings to use them for other purposes, like indexing their gender presentation and sexuality.

In this study, the primary argument is that the *n* level of meaning attached to these vowels originates from sound symbolism, but that these meanings are reconstructed and the speakers index their own meanings using these linguistic forms. As Eckert (2010) shows, the primary variable being manipulated in these vowels involved the fronted/lowered versus backed/raised dichotomy. To measure how fronted or backed these vowels are, F1 and F2 need to be measured, which is why they are the other acoustic correlates.

Hypothesis & Research Questions

Research Questions

The primary goal of this study is to investigate macro level phonetic variation, to examine if any phonetic variation is closely aligned with the speakers' gender presentation and sexuality, and to see if micro level interactions based on if these speakers use these phonetic variables as linguistic resources as part of their stylistic practices. In addition, since these speakers do not constitute a community of practice and instead belong to an imagined community, my second goal is to understand how these speakers reconstruct the meanings attached to these variables to position themselves within these imagined communities based on locally constructed categories.

My research questions are laid out below:

1. Is there phonetic variation in the trans-feminine communities in Toronto based on sexuality, similar to some of the variation found in cisgender communities? Do these variables vary among the lesbian, bisexual, and heterosexual participants?

2. Is there also phonetic variation based on the gender presentation of the speakers? Is there phonetic variation based on the interaction of the sexuality and gender presentation of the speaker?
3. What goals do these speakers seek to achieve by using these variables?
4. What meanings are attached to these variables? How do these meanings vary among the speakers of this imagined community?

Hypotheses

Based on previous research, such as Zimman (2017) and Munson et. al (2005), my hypotheses are laid out below. Zimman (2017) found that /s/ center of gravity varied based on the gender presentation of the speaker, with more feminine presenting trans men having a higher center of gravity and the masculine presenting trans men having a lower center of gravity. Munson et. al (2005) found that higher values of F1 and F2 (fronted vowels) were loosely being associated with gay men and lower values (backed vowels) were being associated with bisexual and lesbian participants.

1. The F1 and F2 of the front and back vowels of the speakers will vary based on their sexual orientation, as the bisexual speakers will have more fronted/lowered vowels while the lesbian speakers will use more backed/raised vowels. This same pattern will also continue in the /s/ center of gravity for the speakers based on their sexual orientation, as the bisexual speakers will have a higher center of gravity with a more fronted production and the lesbian speakers will have a lower center of gravity with a more backed production.

2. Speakers use these variables as a linguistic resource to create meaning, such as indexing sexual orientation, constructing their gender presentation and gender identity, and positioning themselves within locally constructed categories.

My goal is not to make broad statements such as bisexual speakers have higher values for center of gravity and F1/F2, and vice versa for lesbians, because of stereotypes like bisexual women being more associated with femininity while lesbians are more associated with masculinity. These types of statements are reductive and don't account for the wide variety of gender identities, presentations and styles within these communities. Instead, my goal is to explore how meanings are reconstructed by these speakers using these phonetic variables as resources, and how these speakers index locally constructed categories. By the end of this paper, there should be an understanding of how these meanings are used by these speakers in the process of constructing their gender identities, gender presentations, and sexualities to position themselves within locally constructed categories in these imagined communities.

Participants

Participants were primarily recruited via posters that were placed around the campus of York University and six of the eight participants were recruited that way. Two of the participants were recruited via word-of-mouth and do not attend York University. While I attempted to recruit lesbian, bisexual, and heterosexual participants, there were no heterosexual participants. Five of the participants identified as lesbian and three of the participants identified as bisexual. When the participants signed up, they were asked to fill out a form on Calendly and to sign up for a date and time slot. On the Calendly form, the participants were asked to read and fill out the consent form. The two participants who were recruited via word-of-mouth knew each other, and did not

know any of the other participants. The other six participants did not know any of the other participants. All of the participants were compensated for their time in the form of $15 Amazon gift cards.

Four of the eight participants identified as white, with the other four participants identifying as trans people of color. The age of the participants ranged from 18 to 29. During the interviews, information about the gender identity, gender presentation, and sexualities of the participants were obtained. While the demographic information of the participants was not obtained on the Calendly form, as there was too much to collect, their information was collected through a form sent to the participants. The participants were asked to self-identify basic demographic information, such as age, ethnicity, class and birthplace (Table 1).

The participants were also asked on the Calendly form if they would like to use pseudonyms for the duration of their participation in the project, and only two participants chose to do so. These speakers chose to go by the names of Nyx and Starra for the duration of the study, while the rest of the speakers were assigned numbers (such as speaker 1, speaker 2, etc). Not all of the participants responded to the form sent out, so any information missing is described as not applicable. In addition, other demographic information, such as gender identity, gender presentation, and sexual orientation, was collected during the interview.

All of the gender identities of the participants were the same, as they all identified as women. Participants were asked to self-identify their gender presentations and sexual orientation, which are listed in Table 1. For the bisexual interviewees, Table 1 displays what gender identities they desire. These data were gathered during the interview. The purpose of presenting what gender identities the bisexual speakers desire is that the speakers differed in terms of which groups they were interested in dating. While speakers 1 and 4 did not mention having a

preference for one gender identity over the other and seemed to equally desire being in relationships with both men and women, speaker 2 talked about having a strong desire for men over women. Since this may potentially have an effect on how these speakers use these linguistic resources, it seems relevant to have this information in Table 1.

Methodology

Tasks

There were two tasks used for this project, with the first task being a sociolinguistic interview and the second task being a reading passage task. In the sociolinguistic interview, the eight participants were asked a number of questions about their identities, with each interview ranging from 20-40 minutes. In the reading passage task, participants read the Rainbow Passage (Fairbanks 1960) out loud twice. The Rainbow Passage is a reading passage that has been used in previous studies on this topic, such as Zimman (2017a), which is why it will be used for this methodology as well. The goal of the different tasks was to measure whether the phonetic variables being measured varied significantly among the participants across different speech contexts. The sociolinguistic interview was used to measure casual speech while the reading passage was used to measure read speech. As Zimman (2017b) notes, the reading task allows for participants to perform their gender identity. By examining how participants use these phonetic variables in these performances, it should allow for deeper insight into why these variables are being used in these practices and what strategies they employ to construct their identities.

Both tasks were recorded at the York University Phonetics Studio, were sampled at 16000 Hz, and were conducted by the author. Participants were allowed to look at the interview questions before the interview started as well as before each question was asked. Participants

were given the Rainbow Passage (Fairbanks 1960) before the task and were asked to practice reading it out loud. Each participant read the task out loud twice and the reading with the fewest speech errors (such as pauses or stutters) was chosen. For most participants, their second reading was chosen. However, one participant (Starra)'s first reading was chosen.

Montreal Forced Aligner, WhisperAI, & ELAN

Following the conclusion of the interviews, they were transcribed using WhisperAI. After the completion of the automatic transcription, the transcription was checked to ensure there were no errors in spelling or grammar. The transcription software ELAN was used afterwards to manually align the transcriptions with the audio file and converted into TextGrids. The audio from the interviewer was not transcribed so only the interviewee's audio would be used for the transcriptions. The TextGrids were then checked in Praat and the pitch settings were set to the default settings at 50 Hz and 800 Hz. Since F0 was not being examined, and the maximum and minimum F1 and F2 were not a concern, the settings were kept at their default values.

After confirming the TextGrids were aligned, the Montreal Forced Aligner (McAuliffe et al., 2017) was used to align the graphemes from the transcriptions with their phonemic representations using the software Anaconda. During this process, the dictionary used to convert these grapheme representations was trained using a grapheme-to-phoneme model to include any foreign words, such as slang words and community specific words, that were not already present in the default English US ARPA G2P model v2.0.0. Once this process of aligning the graphemes with the phonemes was complete and the corresponding TextGrid was created, these TextGrids were then used to extract the measurements needed for this analysis, mainly the /s/ center of gravity and F1/F2 of the measured vowels.

FAVE & Spectral Moments Analysis

FAVE Analysis

Using the TextGrids generated by the MFA, they were then analyzed using FAVE (Forced Aligner and Vowel Extraction), which is a tool for automating and optimizing vowel formant extraction (Fruehwald, 2026). FAVE automatically collected a wide range of measurements, such as F0/F1/F2/F3, vowel intensity, duration, and the word each vowel occurred in. However, for the purposes of this research, only the F1/F2, the word, and the context (such as internal, initial, or final) were considered. The context in which the vowels occurred within the word was not controlled, as trying to select only word-initial vowels or word-final vowels resulted in a significantly lower number of tokens and a discrepancy between the total number of tokens per speaker. This means that the contexts the vowels occurred in were word-initial, word-internal, and word-final. Using FAVE, six vowels, /o/, /i/, /ae/, /ɛ/, and /ɑ/ were collected. However, following the collection of the vowels, /u/ was ultimately not used due to the small number of tokens that was found in the speech of the participants. While previous research, such as Zimman (2017), examined variables by dividing the interviews into intonational phrases, these interviews were divided based on the question context that each variable occurred in. For example, if a participant was asked question 10 and their speech contained one of the five vowels, the occurrence of that vowel would be labeled as occurring during question 10. This allowed for the speech to be divided into different speech contexts. This also meant that if some participants varied in terms of how long they responded to each question, as many of them did, then some of the questions would have more tokens then the others. Ultimately, I decided that being able to examine the different speech contexts and how that could affect the measurements being collected was worth this downside.

Spectral Moments Analysis

To collect the /s/ center of gravity, DiCanio (2013)'s spectral moments Praat script was used. This spectral moments analysis collected the duration, intensity, center of gravity, standard deviation, skew, kurtosis and the word the token occurred in. For the purposes of this research, only the duration, center of gravity and word were considered. Similar to the vowels, the /s/ tokens were divided based on the question context that each variable occurred in. Following the collection of these measurements, the tokens were filtered so that only words with initial /s/ were considered, which follows Zimman (2017a)'s methodology. This ensures that the phonological context remains consistent, while the speech context varies. In addition, each /s/ token was divided based on whether a consonant followed the word-initial /s/ or whether a vowel followed. For the models, word-initial /s/ tokens followed by vowels were exclusively used.

Models

All of the data collected using FAVE and the spectral moments analyses were modeled with mixed-effects models. The data modeled includes the main measurements collected (/s/ center of gravity, F1, and F2) for all eight participants. In addition, the models include the measurements per question type for each speaker, with the two question types being the interview questions and the reading passage. As discussed in the literature review, reading passage tasks can lead to a performance of one's gender presentation that is often not seen in semi-structured interview tasks. These models of question types should be able to show whether these measurements differ significantly based on the question type or not. These question type models were also calculated using all contexts for the /s/ tokens, rather than only word-initial /s/ tokens like in the previous models. This was done due to the small amount of tokens in each

reading passage task compared to the significantly larger amount of tokens in the interview questions.

Linear Mixed-Effects Models

Linear Mixed-Effects Models (LMMs) were fitted in R Studio to model the response variables, fixed effects, and random effects. The reason Linear Mixed-Effect Models were chosen to model this data is that LMMs maintain variation across multiple data points. As Winter (2020) states, "by-participant variability in the ratings is not retained in the final analysis if you are working with by-word averages" (p. 234). He also says that if averaged data points are the only data a model uses, it is going to underestimate the amount of variation in the data (Winter, 2020). This is why LMMs are used, as they do not rely on averages. Instead, LMMs introduce "random effects", which is non-independent data, consisting of non-independent intercepts and slopes (Winter 2020). This is the main difference between LMMs and other models, as they mix fixed effects with random effects.

Center of gravity, F1, & F2 were the dependent variables for these models. One way of thinking about these relationships between the response/dependent variable and the fixed effect is that these models are predicting how the fixed effects *Gender Presentation* and *Sexuality* are influencing the dependent variables *Center of Gravity, F1,* and *F2*. The result of a LMM is that it shows the intercept based on the fixed effects and random effects.

In a normal linear regression model, the intercept is the value of the dependent variable *Y* when the independent variable *X* is equal to zero. The intercept can also be referred to as the reference for each model. The intercept in an LMM is the mean of the response variable, in this case either center of gravity or F1/F2, when all the predictor variables, which are the fixed

effects, are equal to zero. The intercept is automatically chosen based on the levels of the data frame, which is the table where the data is stored in R studio. This can be changed, so the intercept was changed for this analysis so that the gender category would be *feminine* presentation and the sexuality would be *lesbian*. This was done so that the fixed effects for the gender category would be compared to the intercept *feminine*, which is the category that the majority of speakers identify with. The goal was to compare the other gender presentations to the majority to examine whether there were differences in the intercepts. The reason the sexuality intercept was selected to be *lesbian* was because the majority of speakers identified as lesbians, so the reasoning is the same as the previous intercept, to compare the other intercepts to the majority.

The fixed effects depend on whether the model is examining intraspeaker variation or interspeaker variation. The first relationship modeled for the interspeaker variation examines the relationship between /s/, F1, and F2 with the sexuality and gender presentation of the speaker, so the fixed effects will be the sexuality and gender presentation. When the model is examining the interspeaker variation, such as whether /s/, F1, and F2 pattern based on question type, the fixed effect will be the question type. The random effect for both will be the speaker. This is due to the fact that each speaker produces multiple tokens, so having the speaker be the random effect will measure the effect that the speaker will have on these measurements. As can be seen in Table 1A, each speaker was assigned a gender category based on their stated gender presentation and their interview responses. The main analysis will be examining the correlation of the fixed effects to the intercept. Six of the participants are in the “feminine” gender category, and two are in the “fem & masc” category. As previously stated, five of the participants identified as lesbian and three of the speakers identified as bisexual. For the sexuality gender effect, its intercept was

the lesbian category, since I wanted to compare the bisexual speakers to the intercept, since there are more lesbian speakers.

Vowel Normalization

After collecting all of the F1 and F2 data from FAVE, vowel normalization procedures were used. Based on Adank et al. (2004), vowel normalization is a process that preserves phonemic information, preserves sociolinguistic information for each speaker, and minimizes the variation that the physiology of the speaker can cause. This means that all remaining variation seen in the data is caused by the sociolinguistic factors, rather than any physiological factors. The vowel normalization method utilized in this research is the Lobanov normalization, which is also known as the z-score. Lobanov normalization is an example of a vowel-extrinsic procedure, which take information across multiple vowels, rather than a vowel-intrinsic procedure, which only uses a single vowel token for the normalization process (Adank et al., 2004). In addition, according to Adank et al. (2004), vowel-extrinsic normalization procedures, like the Lobanov procedure, performed better than the vowel-intrinsic procedures.

The Lobanov normalization was carried out in R Studio and followed Fruehwald (2022)'s tutorial. The Z-score for each vowel (/o/, /i/, /ae/, /ɛ/, and /ɑ/) were calculated and a sample of the results can be seen in Table 1B. While these Z-scores aren't useful on their own, they are useful for plotting and modeling F1 and F2 values while ensuring that any potential variation is only due to the sociolinguistic factors, rather than physiological factors. However, only the raw values for F1 and F2 will be used for the models and analysis. While it is useful to have the vowels be normalized, the z-scores and normalized vowels were not used when comparing speaker's formant frequencies in previous research, such as Munson et. al (2005).

Results

Models and Plots

These next sections will examine the results of the Linear Mixed-Effects models as well as the plots. These sections are divided into the individual fixed effects, the additive fixed effects, and the iteration models. During the iteration models section, the plots will be discussed alongside the models. As previously stated, all of the participants identified as women, so their gender identities are not factors in their pronunciations of these variables. This means that the gender presentation and sexualities will be the main factors used for comparison.

First off, the Linear Mixed-Effects Model for all of the speakers based on their center of gravity can be seen in Table 4A. Here, each fixed effect was tested on its own so that the intercept, that being *feminine* for the gender category fixed effect and *lesbian* for the sexuality fixed effect, could be isolated on its own. The first equation for the center of gravity model is *cog ~ sexuality + (1|speaker)*, with center of gravity as the response variable, sexuality as the fixed effect and speaker as the random effect. The second equation for the model is *cog ~ gender.category + (1|speaker)*.

These equations result in a variety of different measurements, but the only values that will be recorded in this are the fixed effects coefficients and the significance relationships, which are based on the p-values per model. The coefficients represent how much the fixed effect intercepts vary from the primary intercept. The significance relationships represent how statistically significant a result is as well as if there is sufficient evidence against the null hypothesis of model equivalence (Winter 2022). The null hypothesis argues that there is no relationship between variables, so if there is not sufficient evidence against it (p-value > .05),

then the null hypothesis is supported and there is most likely no relationship between the variables.

Individual Effects - Center of Gravity

The first two models examined are in Table 4A. The first model is the gender presentation model, with the equation *cog ~ gender.category+ (1|speaker)*. In these equations, the first variable is the response variable, which in this case is center of gravity. The second variable is the fixed effect, which is gender presentation, but there can be multiple fixed effects at once. The final variable in the parentheses is the random effect, which is the speaker. The results of this model can be seen here:

Measurement	Gender Presentation Intercept (Feminine) (Hz)	Gender Presentation Fem & Masc (Hz)
CoG	5471	357, NS

Table 4A - Center of Gravity Individual Effects Models by Gender Presentation

The gender presentation intercept, which as previously discussed is the *feminine* presentation, has a value of 5471 Hz for the intercept. In R studio, p-values will have significance codes placed after numbers that meet certain thresholds. For values below .05, they are given one asterisk *, for values below .01, they are given two asterisks **, and for values below .001, they are given three asterisks ***. For the *fem & masc* intercept, the coefficient is 357 Hz. In Table 4A, since the *fem & masc* intercept does not have a significant value, it was labeled as NS for not significant. From these results, it shows that the *fem & masc* intercept is 357 Hz above the *feminine* intercept, but that this difference between the intercepts is not statistically significant, which suggests that there is no significant effect of gender presentation on center of gravity for the individual effects models.

The second model being examined in Table 4A is the sexuality model, with the equation *cog ~ sexuality + (1|speaker)*. The results of this model can be seen here:

Measurement	Sexuality Intercept (Lesbian) (Hz)	Sexuality Bisexual (Hz)
CoG	5360	464, NS

Table 4A - Center of Gravity Individual Effects Models by Sexuality

The fixed effect in this model is sexuality. The sexuality intercept, which as previously discussed is the lesbian speakers, has a 5360 Hz intercept. For the bisexual intercept, it has a 464 Hz intercept. This means that the bisexual intercept is 464 Hz above the lesbian intercept, and it is closer to the significance threshold than the *fem & masc* intercept, but it is still not below that threshold. Since both of these intercepts were not below the significance threshold, it suggests that there was not a significant effect of sexuality on the center of gravity models.

The final models examined for the center of gravity models are the vowel fixed effects. The tokens examined for the center of gravity models only contained words with a word initial /s/. However, for a word with a word-initial consonant-vowel cluster, this means that the initial /s/ center of gravity can be affected by the following vowel. These models use the following vowel as a fixed effect to see if the following vowel has an effect on the /s/ center of gravity. These following vowels are limited to the vowels being examined in the F1 and F2 measurements (/ɑ/, /æ/, /ɛ/, /i/, and /o/). The equation for these models was *cog ~ following.vowel + (1|speaker)*, with the following vowel being the fixed effect. The intercept in these models was /æ/, which was chosen since it had the fewest tokens. The results of this models can be seen here:

Measurement	Following Vowels Intercept (æ)	Following Vowels ɑ	Following Vowels ɛ	Following Vowels i	Following Vowels o

CoG	6607	-1266, NS	-845, NS	-652, NS	-1262, NS

Table 4B - Center of Gravity Individual Effects Models by Following Vowel

The intercept coefficient was 6607 Hz intercept, with an /ɑ/ coefficient of -1266 Hz and an /ɛ/ coefficient of -845 Hz. The final coefficient was /o/, with a -1262 Hz. Overall, none of these individual fixed effects have a significant relationship with center of gravity. Although some of these coefficients seem to be drastically different from the intercept coefficient, they are not statistically significant. This suggests that the following vowel variable did not have a significant effect on the center of gravity model.

These differences between the intercepts are not in any specific direction, as none of the variables had a significant effect on the center of gravity models. The coefficients in these models do not support the conjecture that the bisexual speakers would have a more fronted production of /s/. These coefficients also do not support the claim that the speakers differ based on gender presentation. Overall, the variance in these center of gravity models are not accounted for by any of the gender presentation, sexuality, or following vowel variables.

F1 Models

The F1 models will now be examined, which are almost identical to the previous center of gravity models. However, now the response variable/dependent variable is F1 instead of center of gravity. In addition, there will be no following vowel fixed effect, since that is not relevant to the examination of the vowels. Each model will be split among the five vowels, meaning that there are five models, one for each vowel. However, there will be fewer intercepts for these five models due to there only being the two fixed effects.

The first model is the gender presentation model, which has this equation *F1 ~ gender.category+ (1|speaker)*. For these models, only the secondary intercepts with significant

relationships to the formant frequencies will be reported for the sake of brevity. The results of this model can be seen here:

Measurement	Gender Presentation Intercept (Feminine) (Hz)	Gender Presentation Fem & Masc (Hz)
F1 ɑ	511	44 **
F1 æ	701	77 **
F1 ɛ	595	55 **
F1 i	366	17, NS
F1 o	628	38, NS

Table 4C - F1 Individual Effects Models by Gender Presentation

The /ɑ/ intercept for this model has a 511 Hz intercept coefficient. The /ɑ/ *fem & masc* intercept coefficient is 44 Hz **. The /æ/ main intercept is 701 Hz, while the *fem & masc* intercept coefficient is 77 Hz **. The /ɛ/ main intercept is 595 Hz, and the *fem & masc* intercept coefficient is 55 Hz**. Overall, the three vowels that had significant relationships for their *fem & masc* intercepts were /ɑ/, /æ/, and /ɛ/, with neither /i/ nor /o/ having a significant relationship. This implies that there is a significant effect of the gender presentation variable on F1.

One important topic to discuss before moving on is whether changing the main intercept will change the values that each model gives. What does the gender presentation model look like if the intercept changes from *feminine* to *fem & masc*? This was not relevant for the center of gravity models, since none of the secondary intercepts had significant relationships to center of gravity, which is different in these F1 models as many of the secondary intercepts have significant relationships. For example, for the /ɑ/ intercept, the coefficient was 511 Hz when the main intercept was *feminine* and the *fem & masc* intercept was 44 Hz with a p-value of 0.00297 **.When the main intercept is *fem & masc*, the coefficient is 556 Hz. When the secondary

intercept is *feminine*, the coefficient is -44 with a p-value of 0.00297 **. When considering what the coefficients represent, this makes sense. The *feminine* coefficient is 511 Hz, which is -44 Hz below the *fem & masc* coefficient of 556 Hz, which is 44 Hz above. In addition, while the main intercept p-values are both different, the second p-values are the same at 0.00297 **. This is because the intercept coefficients will always be the opposite values, and since the significance relationship is based on the coefficient, the significance relationship will stay the same regardless of which intercept is used as the main intercept. If the model supports a correlation for the coefficients in one direction, such as the *fem & masc* variable having higher values, it has to support a correlation in the opposite direction, meaning that the *feminine* variable has to have lower values. Due to this relationship, the intercept for every model will stay as lesbian for the sexuality models and *feminine* for the gender presentation models, since it is unnecessary to change the intercepts in the models.

For the sexuality model, which has this equation *F1 ~ sexuality + (1|speaker)*. All of the findings can be found in the tables (Tables 4 - 8), but only the intercepts with significant relationships are relevant to this project. The results of this model can be seen here:

Measurement	Sexuality Intercept (Lesbian) (Hz)	Sexuality Bisexual (Hz)
F1 ɑ	515	13, NS
F1 æ	702	48, NS
F1 ɛ	598	25, NS
F1 i	373	-6, NS
F1 o	618	49 *

Table 4C - F1 Individual Effects Models by Sexuality

In the sexuality model, the only intercept with a significant relationship was the /o/ bisexual intercept. The main intercept had a coefficient of 618 Hz, and the bisexual intercept had a coefficient of 49 Hz *. Looking at both the gender presentation and sexuality models, the gender presentation models had three vowels with significant relationships, /ɑ/, /æ/, and /ɛ/, while the sexuality models only had one vowel with a significant relationship, /o/. This suggests there is a significant effect of sexuality on the F1 models.

The differences in these models are in the expected direction, at least for the sexuality models. The trends from these gender presentation models seem to suggest that the effect of the *fem & masc* variable on F1 leads to higher values than the *feminine* variable, as the *fem & masc* coefficients were higher than the *feminine* coefficients. Since it was not predicted in the hypothesis how the speakers would differ based on gender presentation, it cannot be concluded whether this direction for gender presentation is expected. For the sexuality models, these trends are in the expected directions, as the bisexual coefficients were higher than the lesbian coefficients for four of the vowels. The coefficients from these F1 models seem to support the idea that bisexual speakers will have higher formant frequencies than the lesbian speakers. Overall, it seems that a significant proportion of the variance for the F1 individual effects models can be accounted for by gender presentation, while only the variance in the /o/ model can be accounted for by sexuality. It remains to be seen whether the trends from the individual effects models continue into the future models.

F2 Models

The F2 models will now be examined, which are the same as the F1 models. The equations used for the F2 models were *F2 ~ sexuality + (1|speaker)* for the sexuality models and

F2 ~ gender.category + (1|speaker) for the gender presentation models. The gender presentation models will be examined first. The results of this model can be seen here:

Measurement	Gender Presentation Intercept (Feminine) (Hz)	Gender Presentation Fem & Masc (Hz)
F2 ɑ	1557	97 **
F2 æ	1702	29, NS
F2 ɛ	1681	121 **
F2 i	2167	201 *
F2 o	1251	60, NS

Table 4D - F2 Individual Effects Models by Gender Presentation

The first intercept is the /ɑ/ model, which has a main intercept coefficient of 1557 Hz, while the /ɑ/ *fem & masc* intercept is 97 Hz **. The main intercept for the /ɛ/ model has a coefficient of 1681 Hz, while the /ɛ/ *fem & masc* intercept is 121 Hz **. The main intercept for the /i/ model has a coefficient of 2167 Hz, while the /i/ *fem & masc* intercept is 201 Hz *. It seems like there is a significant effect of gender presentation on the F2 models. This continues the trend seen in the F1 models. However, unlike the F1 models, gender presentation has a significant effect on the /i/ model, instead of the /æ/ model.

For the F2 sexuality models, the results can be seen here:

Measurement	Sexuality Intercept (Lesbian) (Hz)	Sexuality Bisexual (Hz)
F2 ɑ	1568	31, NS

F2 æ	1707	5, NS
F2 ɛ	1704	19, NS
F2 i	2203	37, NS
F2 o	1222	110 **

Table 4E - F1 Individual Effects Models by Sexuality

For the sexuality model, the /o/ main intercept has a coefficient of 1222 Hz, while the /o/ bisexual intercept has a coefficient of 110 Hz **. Other than that, the F2 sexuality models remain the same as the F1 models, meaning that the only difference between the F1 models and the F2 models is in the gender presentation intercepts. For the F2 models, the /i/ intercept has a significant gender presentation intercept instead of the /æ/ intercept. The differences in these models were in the expected direction based on the results of the F1 models, with the *fem & masc* coefficients and bisexual coefficients being higher. The F2 models also suggest a significant effect of both gender presentation and sexuality on F2. Similar to the F1 models, the coefficients from these models support the idea that the bisexual speakers will have higher formant frequencies than the lesbian speakers. These coefficients also support the trends seen in the F1 models for the gender presentation variables, as the *fem & masc* coefficients were higher than the *feminine* coefficients. Overall, these models seem to suggest again that a significant proportion of the variance for the F2 individual effects models can be accounted for by gender presentation, while only the variance in the /o/ model can be accounted for by sexuality.

Additive Models

Now that the individual fixed effects models have been examined, it is time to look at how the relationship between the categorical variables (sexuality & gender presentation) and the

measurements are models using the LMMs. The two types of models that can be used for this are additive models and interaction models. The main difference between the models is that additive models examine how each independent variable (sexuality & gender presentation) affect the dependent variables (measurements), without assuming that each independent variable affects the other.

For example, the additive models look at how sexuality affects the measurements, and how gender presentation affects the measurements, but it does not assume that a speaker's gender presentation and sexuality could interact and affect the measurements. So if speaker X was bisexual with a feminine presentation and speaker Y was lesbian with a fem & masc presentation, the additive model would assume that sexuality is the only variable having an effect and ignore the effect that the gender presentation combined with the sexuality would have on the measurement. The additive models answer research questions #1 & #2, which is whether there is phonetic variation based on sexuality and gender presentation.

In contrast, the interaction models assume that the independent variables do have an effect on each other and can influence the measurements. If speaker X was bisexual with a feminine presentation and speaker Y was lesbian with a feminine presentation, the interaction model would assume that these speakers would have different outcomes for their measurements, since it assumes that these variables are interacting with each other. Even though both speakers have the same gender presentation, the interaction models assume that their sexualities have an effect on their gender presentations and will lead to different outcomes. The interaction model also answers question #2, which is how these independent variables interact and if their interactions lead to different outcomes for the speakers.

This is why both the additive and interaction models will be used. As previously discussed, due to the volume of values that these models have resulted in, only the intercepts that lead to significant relationships will be mentioned. The additive models are still following the rules of the previous models, so only intercepts with significant relationships will be reported as well as the secondary intercepts. The Center of Gravity additive model results can be seen here:

Measurement	Gender Presentation + Sexuality + Following Vowels: (Intercept:Lesbian + Feminine + æ) (Hz)	Gender Presentation + Sexuality + Following Vowels: Bisexual (Hz)	Gender Presentation + Sexuality + Following Vowels: Fem & Masc (Hz)
CoG	6398	428, NS	192, NS

Table 5A

Measurement	Following Vowels ɑ (Hz)	Following Vowels ɛ (Hz)	Following Vowels i (Hz)	Following Vowels o (Hz)
CoG	-1272, NS	-865, NS	-671, NS	-1280, NS

Table 5B - Center of Gravity Additive Models

The differences between the intercepts were in the expected direction, as the *fem & masc* coefficients and the bisexual coefficients were higher. However, none of the intercepts from the center of gravity models have significant relationships, which implies there is not a significant effect of any of the variables on center of gravity. These coefficients do not support the idea that the bisexual speakers will have more fronted productions of /s/, since it does not seem like there is a significant effect of sexuality on center of gravity. However, these coefficients do support the trend from the previous models, as the *fem & masc* and bisexual coefficients are higher, even if there is no significant relationship between both gender presentation and sexuality on center of gravity. Overall, the variance in these center of gravity additive models cannot be accounted for by either gender presentation or sexuality.

F1 Additive Models

The F1 additive models use this equation *F1 ~ sexuality + gender:category + (1|speaker)*. The results of the model can be seen here:

Measurement	Gender Presentation + Sexuality (Intercept:Lesbian + Feminine) (Hz)	Gender Presentation + Sexuality: Bisexual (Hz)	Gender Presentation + Sexuality: Fem & Masc (Hz)
F1 ɑ	508	6, NS	43***
F1 æ	687	38 *	70 **
F1 ɛ	588	18, NS	50**
F1 i	369	-9, NS	19, NS
F1 o	612	44 *	32, NS

Table 5C - F1 Additive Models

For the intercepts, the /ɑ/ main intercept coefficient is 508.812 Hz, with a *fem & masc* intercept of 43 Hz ***. The /æ/ main intercept coefficient is 687 Hz, with an /æ/ *fem & masc* intercept coefficient of 70 Hz **, and an /æ/ bisexual intercept coefficient of 38 Hz *. The /ɛ/ main intercept coefficient is 588 Hz, while the /ɛ/ *fem & masc* intercept coefficient is 50 Hz **. The /o/ main intercept coefficient is 612 Hz and a bisexual intercept coefficient of 44 Hz *. Overall, the F1 additive models have almost the exact same results as the individual effects models. For the gender presentation variable, the three vowels that were previously significant, /ɑ/, /æ/, and /ɛ/, are significant again. However, for the sexuality variable, the two vowels that are significant are /æ/ and /o/, while previously only /o/ had a significant relationship.

The differences in these models are in the expected direction, with the *fem & masc* and bisexual coefficients being higher. There seems to be a significant effect of both gender

presentation and sexuality on F1. The coefficients from these models seem to support the conjecture that the bisexual speakers will have higher values. In addition, the coefficients support the trends from the previous models for the gender presentation variable in regards to the *fem & masc* speakers having higher values. These results imply that the trends from the individual effects models continue in these additive models. A significant proportion of the variance in these F1 additive models can be accounted for by gender presentation, with a smaller amount of the variance being accounted for by sexuality, although it is more than the F1 individual models as /æ/ also has a significant relationship in addition to /o/.

F2 Additive Models

Next are the F2 additive models, which use this equation *F2 ~ sexuality + gender.category + (1|speaker)*. The results of the model can be seen here:

Measurement	Gender Presentation + Sexuality (Intercept:Lesbian + Feminine) (Hz)	Gender Presentation + Sexuality: Bisexual (Hz)	Gender Presentation + Sexuality: Fem & Masc (Hz)
F2 ɑ	1550	16, NS	94 **
F2 æ	1702	0.8, NS	29, NS
F2 ɛ	1680	1.8, NS	120 **
F2 i	2163	10, NS	199 *
F2 o	1217	103 ***	34, NS

Table 5D - F2 Additive Models

For the intercepts, the ɑ main intercept coefficient is 1550 Hz, with a *fem & masc* coefficient of 94 **. The /ɛ/ main intercept coefficient is 1680 Hz, with a *fem & masc* coefficient of 120 **. The /i/ main intercept coefficient is 2163 Hz, with a *fem & masc* coefficient of 199 *. The /o/ main intercept coefficient is 1217 Hz, with a bisexual coefficient of 103 Hz ***. For the

gender presentation variable, the three vowels with significant p-vowels are different, which are /ɑ/, /ɛ/, and /i/. For the sexuality variable, there is only one significant vowel again, which is /o/. In this case, there does not seem to be any relationship between the three vowels. Overall, across all the models, the one consistent vowel that seems to have significant relationship for sexuality is /o/, and the only two vowels with significant relationship are /ɑ/ and /ɛ/. All of the other vowels have had significant relationships for gender presentation, but /o/ has never had a significant relationship for gender presentation and the only other vowel with a significant relationship for sexuality is /æ/.

All of the trends from the F2 individual effects models and both the F1 models are supported here. The direction for the effects is the same here, as the bisexual and *fem & masc* coefficients are higher. The conjectures from the previous models are also supported here. Overall, there are not any findings in the results from this additive model that go against the findings of the previous models.

Iteration models and plots

One final set of models that can be used to model the relationships between these categorical variables is interaction models. Previous models examined each fixed effect individually to see their effect on the measurements. However, by changing the equation to have both fixed effects added and adding an asterisk between the fixed effects, that changes the model to an interaction model. As previously discussed, these interaction models are essential for answering question 2. However, while the results of the Linear Mixed-Effects models are useful when comparing the significance relationships across the models, there are better options. That is why, instead of interaction models themselves being used, significance tests need to be used.

Following Winter (2020), likelihood ratio comparisons (‘deviance tests’) will be used as significance tests.

While you can use likelihood ratio tests to calculate the significance relationships for every combination of the measurement and fixed effects, this can result in errors that may cause issues in the data. When there are models with many fixed effects, Winter (2020) recommends using iteration models that combine multiple likelihood ratio tests. Instead of modeling each fixed effect individually, alongside the additive models and the interaction models, they end up nested within a full model. This compares the results of all of the models across every model type. While you can specify whether the iteration models should be additive or interaction, only the interaction models will be examined. The additive model results can be seen in Table 7, however they do not differ from the other additive model results so they will not be discussed here. During this section, the results of the plots will be discussed alongside the models.

One way of showing the significance between the categorical variables, which are gender presentation, sexuality, and question type, and the continuous variables, the center of gravity, F1, and F2 are box plots. Box plots are very useful for plotting the distribution of continuous variables across categorical variables. There are multiple sections of box plots, with the main box of the plot representing all of the data between the first quartile and third quartile, which is 25% and 75% of the data, also known as the interquartile range. Within this interquartile range, the median is also present. The other sections of box plots are called the whiskers, which represent the minimum and maximum of the plot, calculated by first quartile - 1.5 times the interquartile range and third quartile + 1.5 times the interquartile range for the minimum and maximum respectively. There are also outliers, which lay outside of the whiskers. Another reason why box plots are going to be used to plot the relationships between the variables is that notches can be

added to box plots to make them notched box plots. Notched box plots are useful because they add a 95% confidence interval around the median, which makes the center of the box notched. Notched box plots are essential when comparing the medians of multiple groups, since if the notches from different boxes do not overlap, it means that there is a statistically significant difference between the medians. However, if the sample size is not large enough, this can lead to the confidence interval extending beyond the quartiles and becoming distorted.

This is why data from all of the speakers is being plotted, as plotting data from the individual speakers can lead to a small sample size and lead to distortions. All of the notched box plots are available to view in plots 1A to 8C. They represent the data from the center of gravity from every speaker, split by every speaker, by gender presentation, and by sexuality. This continues for the F1 and F2 plots as well, which are divided per vowel for all speakers per F1 and F2, and then divided by categorical variables. Plots 8A through 8C plot the relationship between the question type and the measurements. Although these plots do not provide all of the evidence for whether the relationships between the categorical variables and the measurements are statistically significant, they are another useful tool when combined with the models to examine these relationships between variables and measurements. The plots only show the differences between the medians, but can show some of the trends of the effects that the variables have on the measurements. When combined with the models and the significance relationships, both the models and the plots will be useful in displaying the general trends that are present in the relationships between the variables and the measurements.

Center of Gravity Iteration Models

For the Center of Gravity models, the equation used is *cog ~ sexuality * gender.category + (1 | speaker)*, there are a total of six models, and the random effects variance is 103086. The results of the model can be seen here:

Measurement	Intercept: Lesbian * Feminine * æ (Hz)	Sexuality: Bisexual (Hz)	Gender Presentation: Fem & Masc (Hz)
CoG	5561	-420, NS	279, NS

Table 8A - Center of Gravity Iteration Models

Measurement	Following Vowel: ɑ (Hz)	Following Vowel: ɛ (Hz)	Following Vowel: i (Hz)	Following Vowel: o (Hz)
CoG	-872, NS	276, NS	327, NS	263, NS

Table 8B - Center of Gravity Iteration Models

Measurement	Fem & Masc * Bisexual (Hz)	Bisexual * ɑ (Hz)	Bisexual * ɛ (Hz)	Bisexual * i (Hz)	Fem & Masc * ɑ (Hz)
CoG	121, NS	-906, NS	287, NS	167, NS	1671, NS

Table 8C - Center of Gravity Iteration Models

Measurement	Fem & Masc * ɛ (Hz)	Fem & Masc * i (Hz)	Bisexual * Fem & Masc * ɑ (Hz)	Bisexual * Fem & Masc * ɛ (Hz)	Bisexual * Fem & Masc * i (Hz)
CoG	-857, NS	-263, NS	993, NS	-223, NS	-402, NS

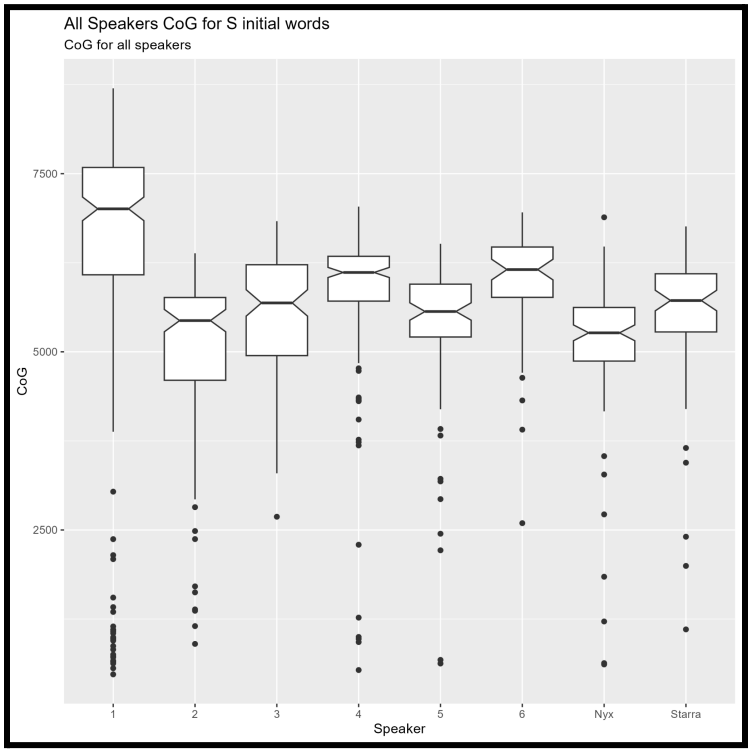
Table 8D - Center of Gravity Iteration Models

Overall, none of the relationships between the intercepts are significant. The differences among the models are unexpected, as the bisexual coefficients are lower than the lesbian

coefficients. However, the *fem & masc* coefficients are still higher than the *feminine* coefficients. The trends for gender presentation still continue here from the previous center of gravity models, but the trends for sexuality do not continue. These iteration models do not support the conjecture that bisexual speakers will have more fronted productions of /s/. However, the conjecture that the *fem & masc* speakers will have more fronted productions seems to be supported, even if there is not a significant effect of gender presentation on center of gravity.

Center of Gravity Plots

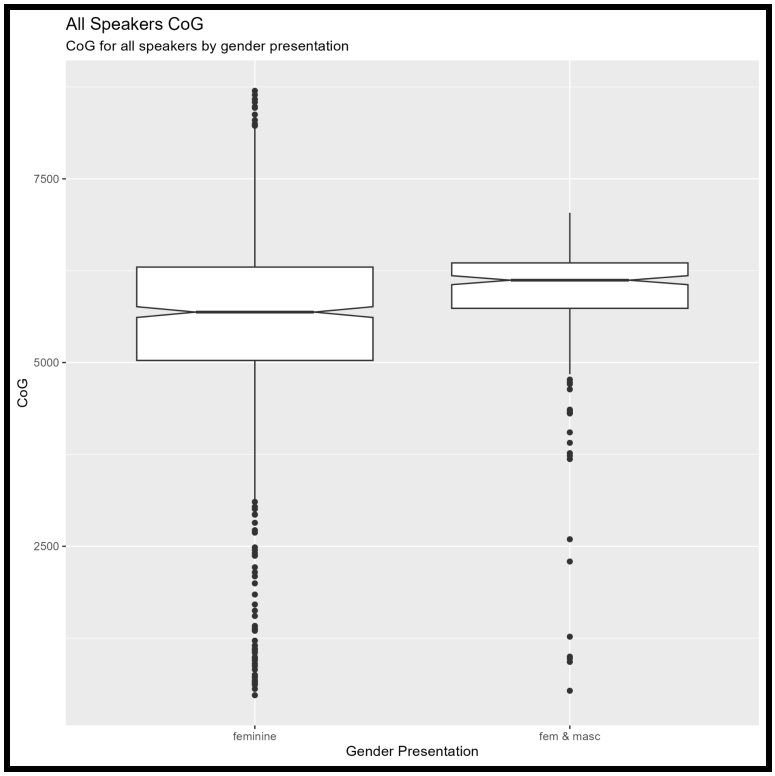
The first plots being examined will be the center of gravity plots, which can be viewed in Plots 1A through 1C. These plots show the distribution of every center of gravity token per speaker (Plot 1A), the distribution of the tokens per the gender presentation variable based on the two subcategories *feminine* and *fem & masc* (Plot 1B), and the distribution per sexuality, based on the two subcategories bisexual and lesbian (Plot 1C). Plot 1A can be seen here:



Plot 1A - Center of Gravity for all speakers per /s/ initial words

When examining Plot 1A, it is important to establish the meaning of the three parts of the box plot, which are the box, the whiskers, and the notched median. The first observation from Plot 1A is that the majority of speakers seem to fall within the same range for their interquartile range, with only speakers 4 and 6 being slightly elevated. However, speaker 1 has a large difference in their interquartile range compared to the rest of the speakers, with the bottom of their first quartile being above the IQRs of three speakers. In addition, their notched median clearly does not overlap with any other speaker, showing that their median is statistically significantly different. For the other speakers, it seems like speaker 2 has the lowest minimum and first quartile, while speaker 4 has the highest maximum and speaker 6 has the highest third quartile. Overall, speaker 1 has clearly the highest median center of gravity at just below 7500 Hz, which is just above the range for the average cis man, which is 7100 Hz, and falls within the range of the average cis woman, which is between 6500 and 8100 Hz. The rest of the speakers seem to pattern together and have overlapping medians, except speakers 4 and 6, whose medians are higher and do not overlap with anyone else's but each other. It seems like the general trends from the models have continued and are in the expected directions, as two of the bisexual speakers (1 and 4) have higher values than the rest of the speakers.

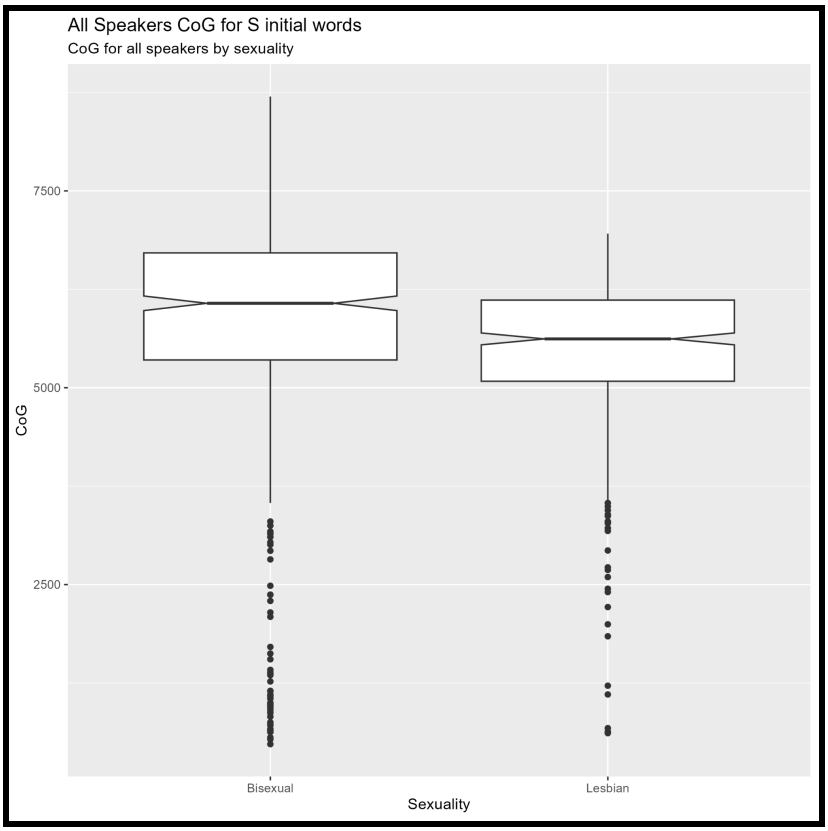
The next plot being examined is the center of gravity by gender presentation plot (Plot 1B), which can be seen here:



Plot 1B - Center of Gravity for all speakers by gender presentation

For this plot, the interquartile range of the two boxes do overlap, with the *fem & masc* IQR being slightly higher than the *feminine* plot, while the *feminine* plot's maximum whisker is significantly higher than the *fem & masc* plot. However, the *fem & masc* plot's median is clearly statistically significantly different from the *feminine* plot, as the notched medians do not overlap at all. As with the previous plot, this does not confirm anything itself. However, it does show that the median center of gravity for the *fem & masc* speakers is significantly different from the *feminine* speakers, which may be useful information when comparing the findings from the plots and models. These plots also continue the general trends and are in the expected direction, as the *fem & masc* speakers have higher medians than the *feminine* speakers.

The final plot being examined for the center of gravity plots is the sexuality plot (Plot 1C), which can be seen here:



Plot 1C - Center of Gravity for all speakers by sexuality

For this plot, the interquartile range does have some overlap, albeit with the bisexual plot's third quartile being noticeably higher than the lesbian plot's third quartile. The bisexual plot's maximum is also significantly higher than the lesbian plot's maximum. Finally, the bisexual plot's notched median does not overlap with the lesbian plot's notched median. This means that the medians are statistically significantly different, and that the bisexual speakers had a higher median than the lesbian speakers. These plots also continue the general trends and are in the expected direction, as the bisexual speakers have higher medians than the lesbian speakers.

When examining the results of these three plots, it is important to start with Plot 1A since it provides a summary of each speaker's center of gravity measurements. It is also important to identify some of the gender presentations and sexualities observed for the speakers in 1A. To start off, speaker 1 had clearly the highest median center of gravity, and they were a bisexual speaker with a *feminine* gender presentation. From plot 1B, the bisexual speakers had a lower

median center of gravity, while the bisexual speakers had a higher median center of gravity in plot 1C.

When compared to the other bisexual speakers, speaker 1 clearly has a significantly higher median. However, the difference is stark among the three bisexual speakers. Speaker 2, who is also bisexual, has the lowest minimum for this plot of all the speakers, and has an overlapping median with the other lower median speakers. In contrast, speaker 4 is also bisexual, and they have a median that is significantly different from speaker 2 but is also significantly lower than speaker 1. One big difference between speaker 2 and speaker 4 is that speaker 2 has a *feminine* presentation while speaker 4 has a *fem & masc* presentation.

For the lesbian speakers, speakers 3, 5, Nyx, and Starra seem to have significant overlap in their box plots, with all speakers having overlapping medians. All four of these speakers also have *feminine* presentations. The final lesbian speaker, number 6, does not have overlapping medians with any of the other speakers. Interestingly, speaker 6 also has a *fem & masc* presentation. Two of the three speakers with the highest median center of gravity have *fem & masc* presentations. While there are two bisexual speakers in the top three highest medians, the lowest speaker overall is bisexual.

It seems like the differences in these plots were in the expected directions, as the bisexual speakers (1 and 4) and *fem & masc* speakers (4 and 6) had higher medians. For the next plots, it will be interesting to see if general trends continue. If the bisexual speakers have such significant differences and the lesbian speakers pattern differently across their gender presentations for the F1 and F2 plots, then it would seem that the general trends will have continued.

Center of Gravity Iteration Models and Plots Discussion

The iteration models for the center of gravity models do not show any significant relationships between center of gravity and any of the variables. However, both the models and center of gravity plots do show some general trends for the variables, namely that speakers 1, 4, and 6 have higher than average center of gravity medians, and the speakers with a *fem & masc* gender presentation tend to have higher medians and coefficients. The effects of sexuality on center of gravity differ in terms of direction for the iteration model and plots, as the bisexual speakers tend to have a higher median but have lower coefficients. While there is not a significant effect of either gender presentation or sexuality on center of gravity, the general trends across the individual effects, additive and iteration models, as well as the medians from the plots, seem to be consistent.

The coefficients from these different model types generally support the conjecture that the bisexual speakers have more fronted production of /s/, even if there does not seem to be a significant effect of sexuality on center of gravity from these models. This lack of a significant effect could be due to the low number of speakers and tokens, so future research on these effects is necessary. Overall, across all of these models, it does not seem like the significant proportion of variance in center of gravity is accounted for by gender presentation, sexuality, or the following vowel, but the models and plots do show a general trend. The next models and plots being examined are for F1 and F2, so it remains to be seen if these trends continue.

F1 Iteration Tests

For the F1 model, the equation used is *F1 ~ sexuality * gender.category + (1 | speaker),* and there are six models for each vowel. The results of the model can be seen here:

Measurement	Intercept: Gender Presentation + Sexuality + Following Vowels (Feminine * Lesbian) (Hz)	Sexuality: Bisexual (Hz)	Gender Presentation: Fem & Masc (Hz)	Sexuality * Gender Presentation (Bisexual * Fem & Masc) (Hz)
F1 ɑ: Variance = 0	534 SINGULAR FIT	-0.8, NS SINGULAR FIT	-22*** SINGULAR FIT	-5, NS SINGULAR FIT
F1 æ Variance = 362.9	741	-20 *	-35 **	2, NS
F1 ɛ Variance = 80.91	625	-4, NS	-25 **	-10, NS
F1 i Variance = 150	375	7, NS	-9, NS	-6, NS
F1 o Variance = 142.4	648	-31 ***	-14 **	18 **

Table 8E - F1 Iteration Models

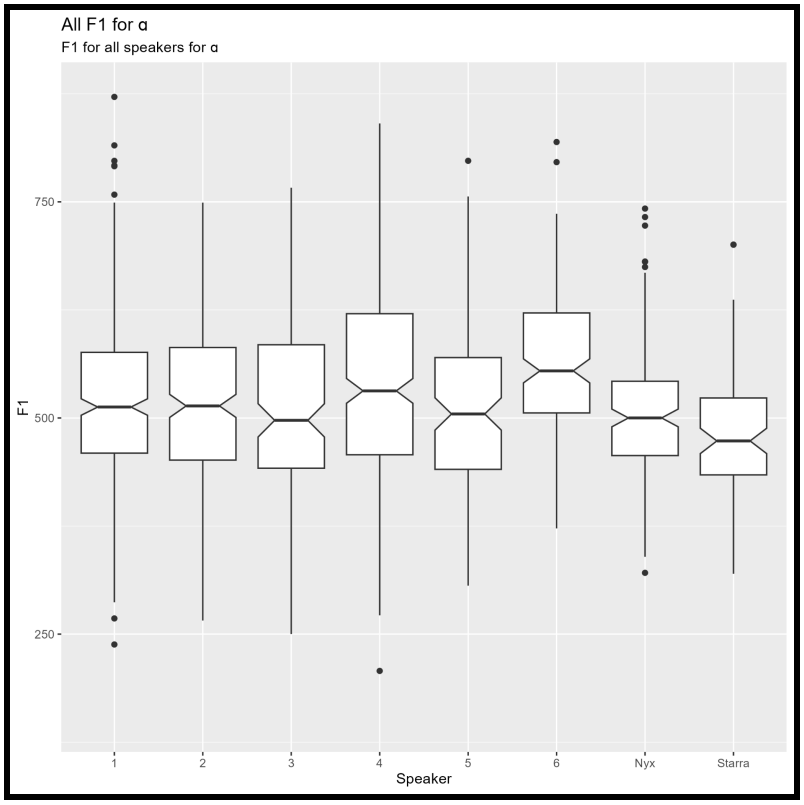
The /ɑ/ model main intercept has a coefficient of 534 Hz, and the *fem & masc* coefficient is 22 Hz ***. However, while this model was being run, a message occurred in R studio stating that this model had a singular fit. As previously stated, the random effects variance being zero means that there is a singular fit. While this does not mean that this model should not be used, it does mean that some caution should be used when interpreting these results. The /æ/ main intercept has a coefficient of 741 Hz, and the bisexual coefficient is 20 Hz *, and an /æ/ *fem & masc* coefficient of -35 Hz **. The /ɛ/ main intercept coefficient is 625 Hz, and the *fem & masc* coefficient is 25 Hz **. The /o/ main intercept has a coefficient of 648 Hz, and the bisexual coefficient is -31 Hz ***, and an /o/ *fem & masc* coefficient of -14 Hz **. Finally, the /o/ bisexual * *fem & masc* coefficient is 18 Hz **.

These iteration interaction models have some quite different results from the previous tables, with all vowels except for /i/ having a significant relationship with the gender presentation intercept. Previously, only /ɑ/, /æ/, /ɛ/ had significant values for gender presentation, but now /o/ is significant. In addition, /o/ is typically the only vowel that has a significant value for sexuality, but /æ/ has a significant value as well. Finally, /o/ is the only vowel with a significant relationship with the sexuality and gender presentation interaction.

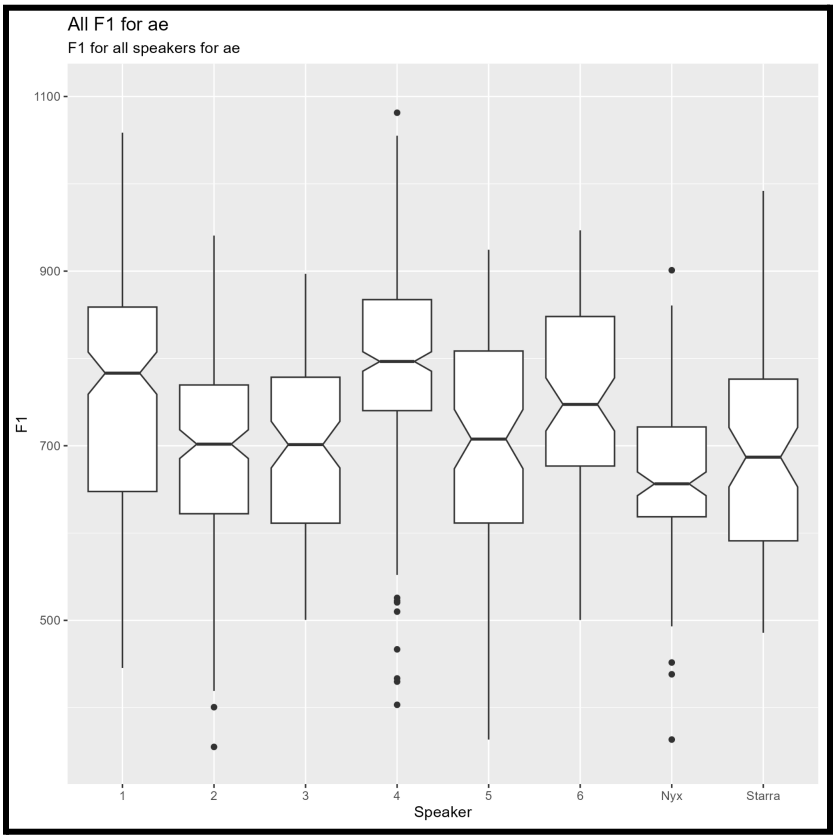
The differences in the F1 iteration models are in an unexpected direction. Even though gender presentation and sexuality seem to have a significant effect on F1, the bisexual and *fem & masc* coefficients are lower than the main intercept. All of the models that have significant relationships to the variables have lower coefficients except for one, which is the /o/ model for the bisexual * *fem & masc* and it has a coefficient higher than the main intercept. This means that the general trends from the previous models are not supported, as the bisexual and *fem & masc* coefficients are lower, but gender presentation and sexuality still seem to have a significant effect on F1. These coefficients do not support the conjecture that bisexual and *fem & masc* speakers will have a higher formant frequencies, even though it does seem that a significant proportion of the variance in these models is accounted for by gender presentation and sexuality. However, it does seem to support the hypothesis that there is an interaction between gender presentation and sexuality, based on the results of the /o/ coefficients for the interaction intercept.

F1 All speakers' Plots

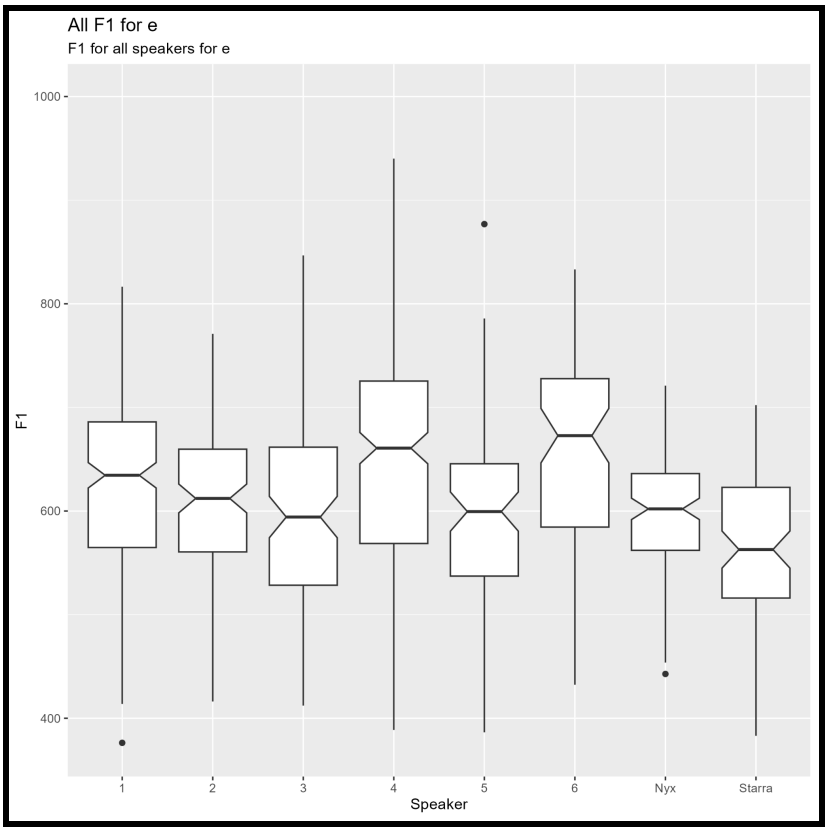
The F1 and F2 plots will be split into three different plots for each measurement. First, the F1 per speaker, per gender presentation, and per sexuality plots will be examined. The next plots being examined are the F1 plots (plots 2A through 2E) for each vowel per speaker. These plots can be seen here:



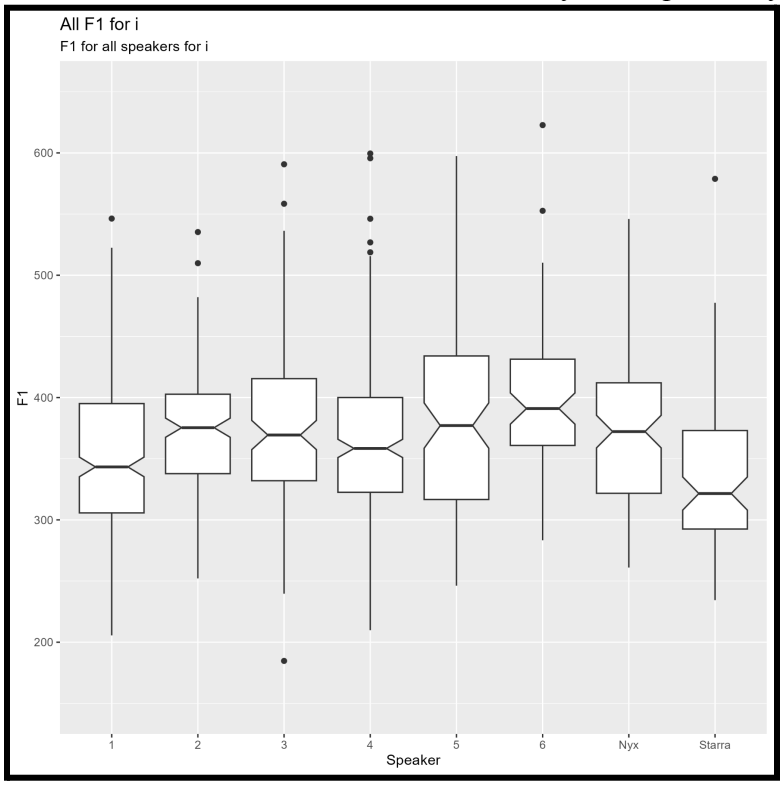
Plot 2A - F1 for all speakers for /a/



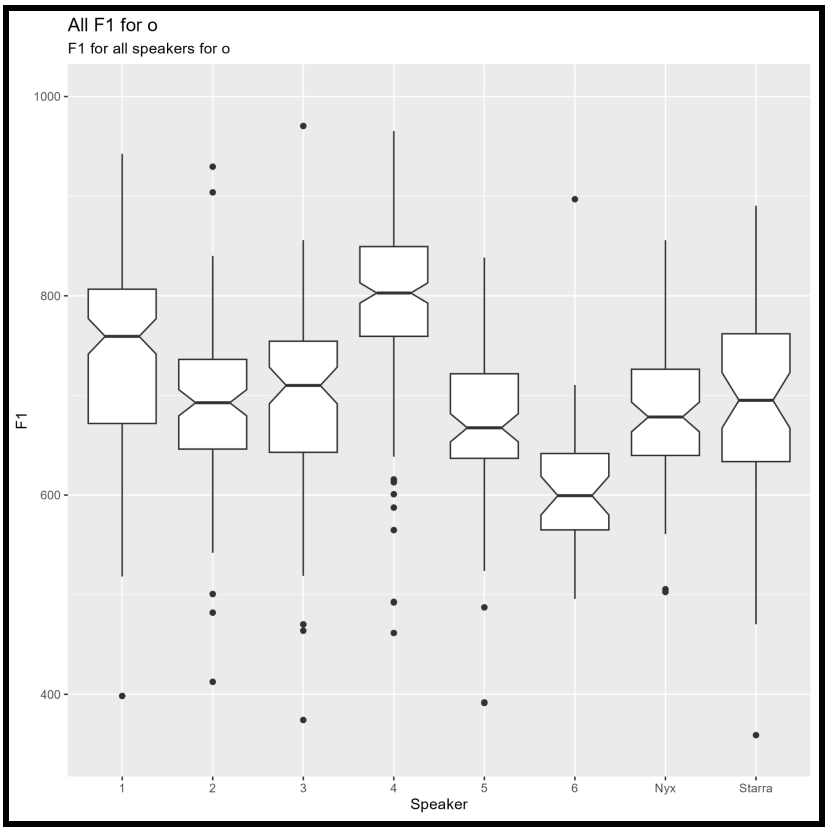
Plot 2B - F1 for all speakers for /æ/



Plot 2C - F1 for all speakers for /ɛ/



Plot 2D - F1 for all speakers for /i/



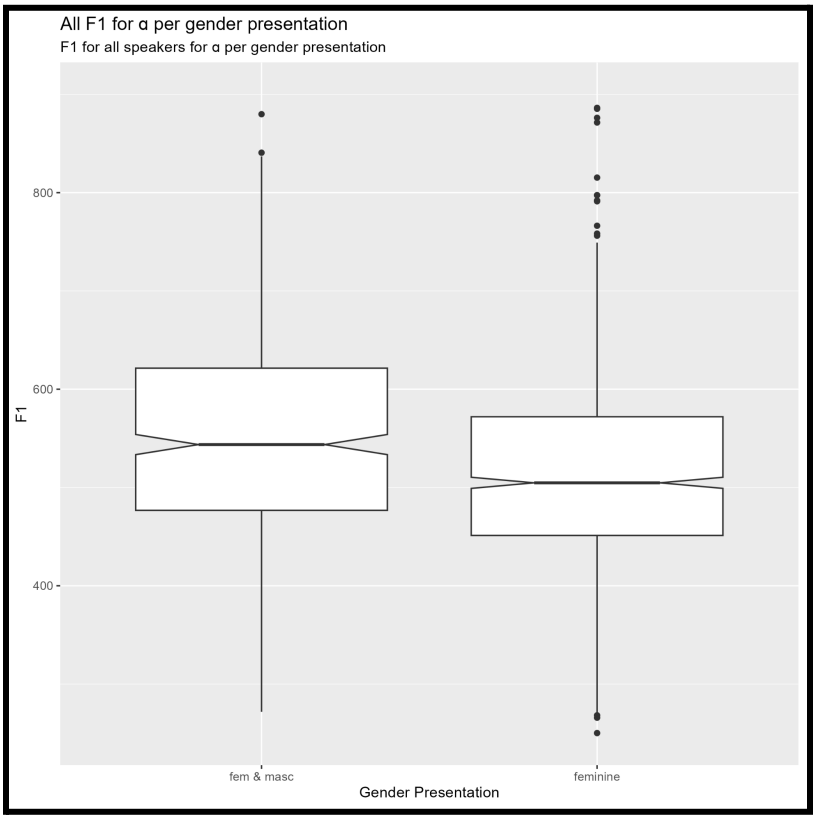
Plot 2E - F1 for all speakers for /o/

For these plots, there seems to be a lot of overlap for the medians. The only speakers who are noteworthy are speaker 6 for the /ɑ/ plot, speakers 1 and 4 for the /æ/ plot, speakers 4 and 6 for the /ɛ/ plot, and speakers 1, 4 and 6 for the /o/ plot. Only speaker 6 for the /o/ plot is notable for having a significantly lower value than the rest, every other speaker has a significantly higher F1 with almost no overlap for their medians. Beyond that, the other speakers do not seem to vary significantly from each other, and all of their medians seem to share overlap in the majority of these charts. These results do not seem to continue the trends from the previous models and plots in either direction, as speakers 1, 4, and 6 varied considerably across the vowels.

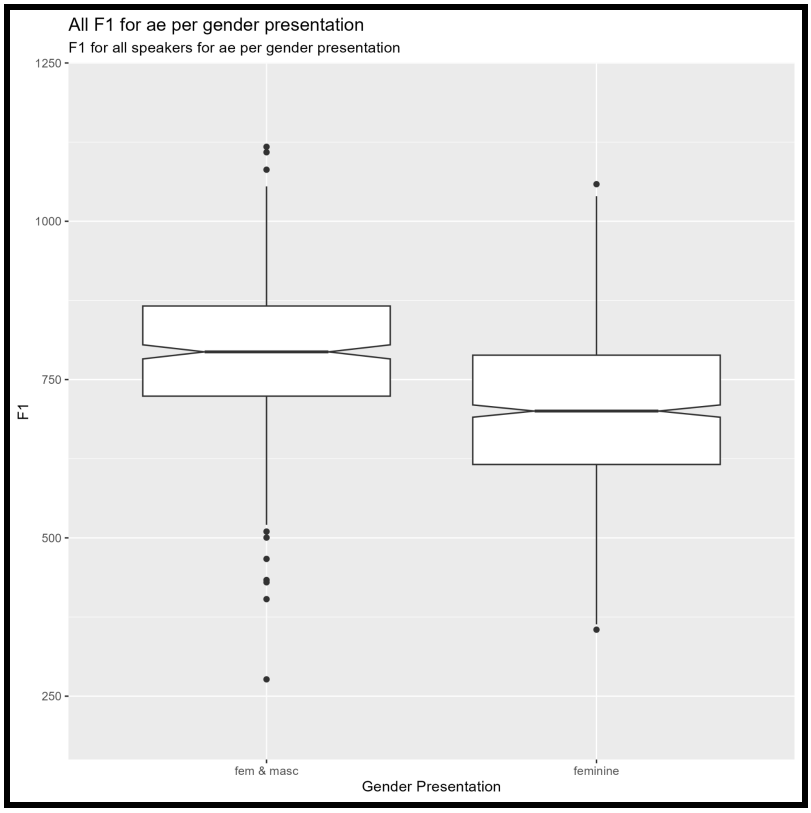
F1 Gender presentation plots

Now that the per speaker plots for F1 have been presented and discussed, the focus will shift to the gender presentation per vowel plots for F1 (plots 4A through 4E). Similar to the

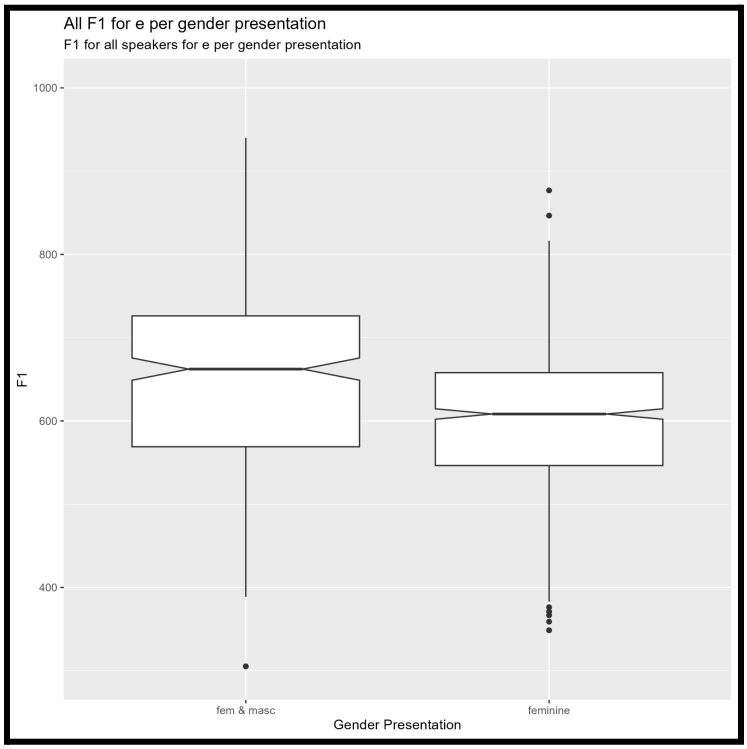
previous plots, these plots are divided by vowel, except for now they will only be examining the gender presentation of the speaker, which is only limited to two variables, *feminine* presentation and *fem & masc* presentation. The F1 plots can be seen here:



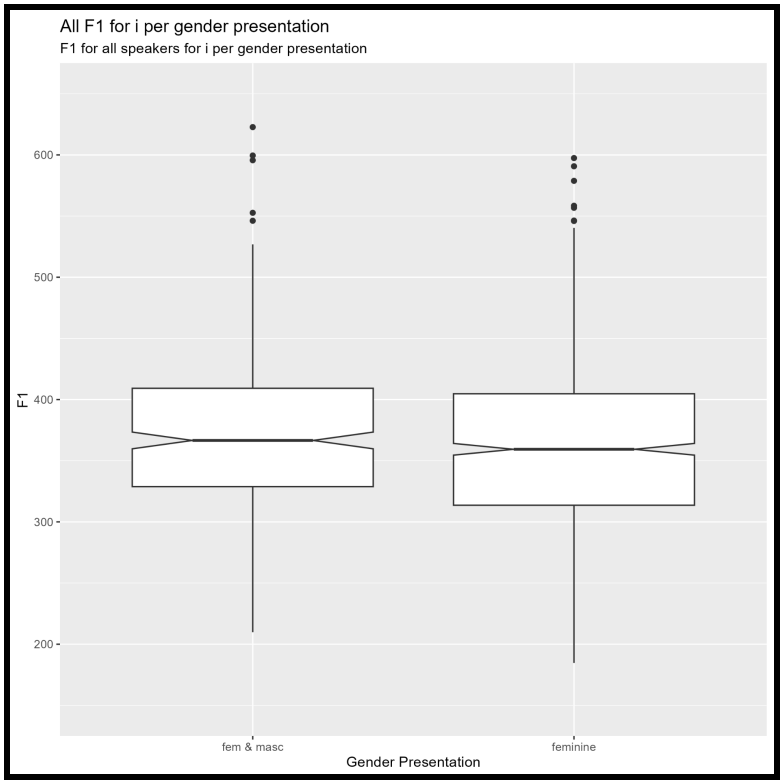
Plot 4A - F1 for all speakers for /a/ by gender presentation



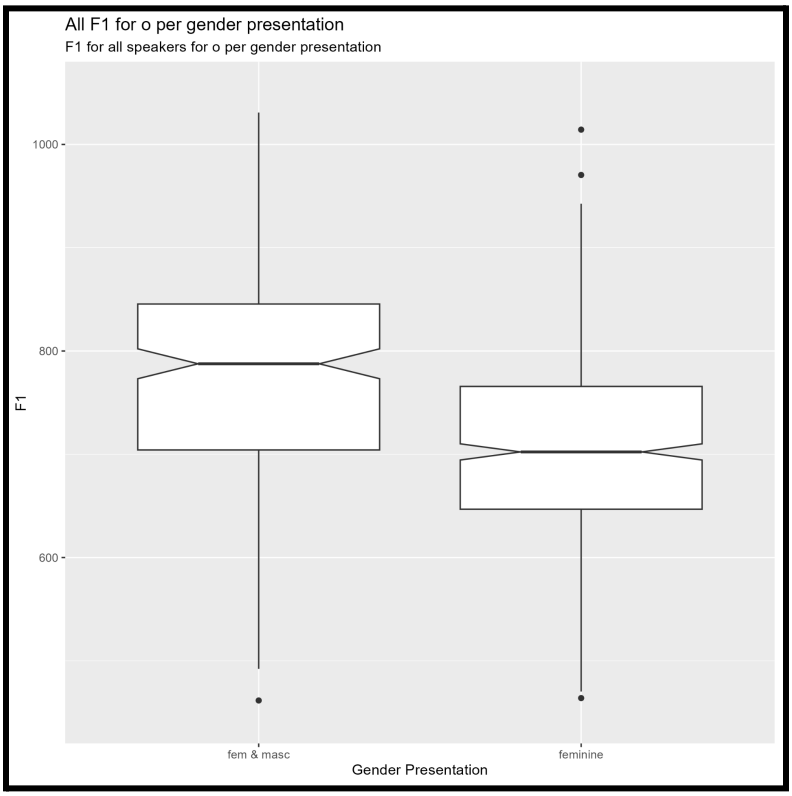
Plot 4B - F1 for all speakers for /æ/ by gender presentation



Plot 4C - F1 for all speakers for /ɛ/ by gender presentation



Plot 4D - F1 for all speakers for /i/ by gender presentation

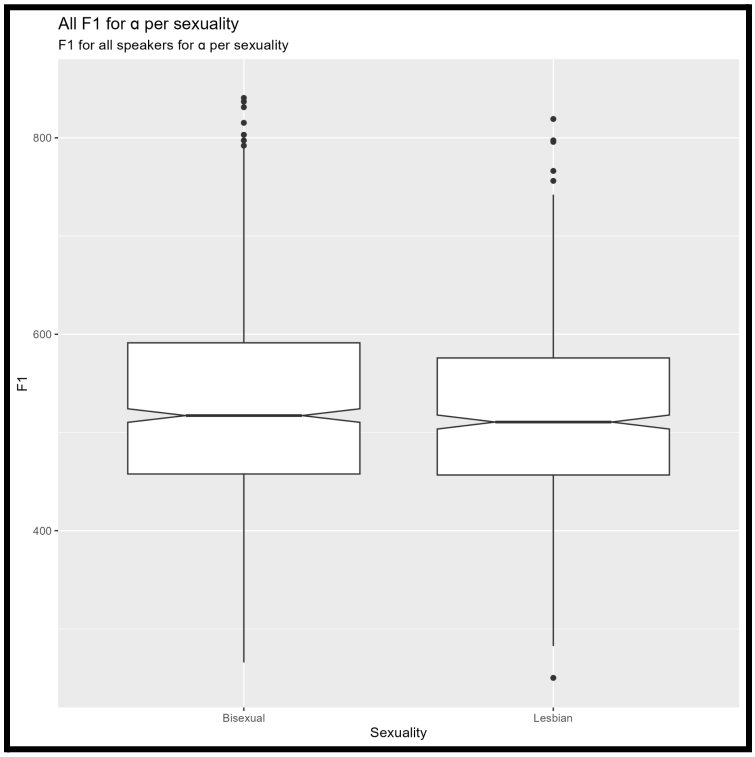


Plot 4E - F1 for all speakers for /o/ by gender presentation

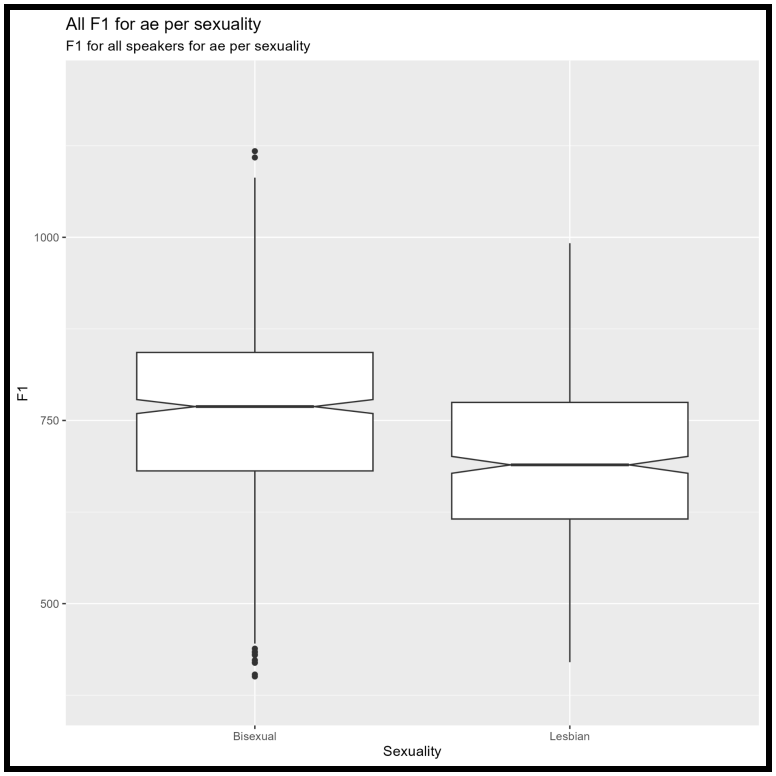
The first observation is that all but one of the plots had *fem & masc* presentation as the variable with the highest value and did not have the medians overlap. The only plot that had the medians overlap was the /i/ plot (plot 4D). In addition, the one plot that seemed to have medians with a significant difference between the two was the /o/ plot (plot 4E). The trends from these plots are in the expected direction, as the *fem & masc* plots had higher medians for the majority of the plots.

F1 Sexuality plots

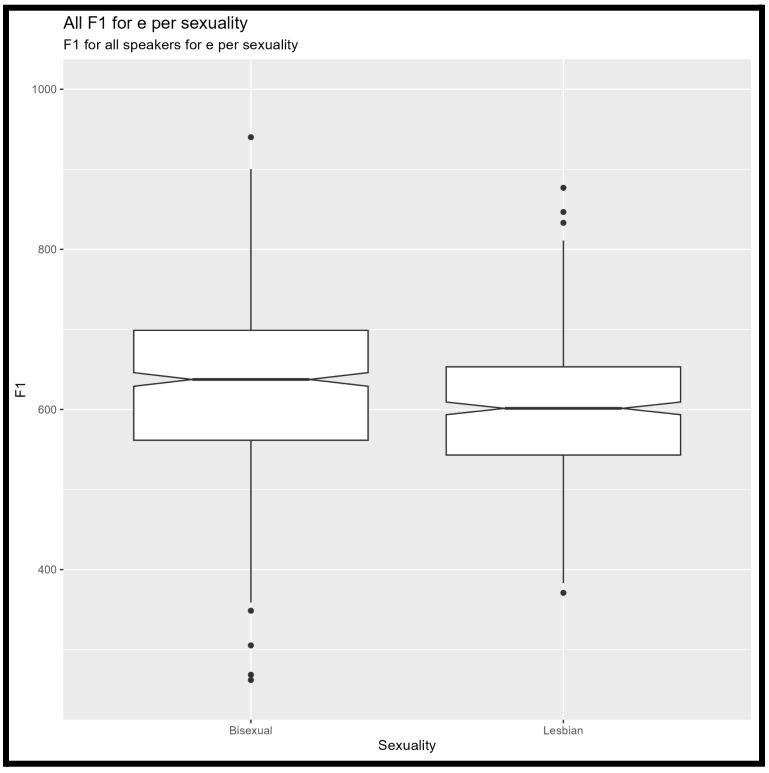
The final plots to be examined are the F1/F2 sexuality plots, which are plots 6A through 7E. These plots are divided by vowel and sexuality, which is limited to bisexual and lesbian. The F1 plots can be seen here:



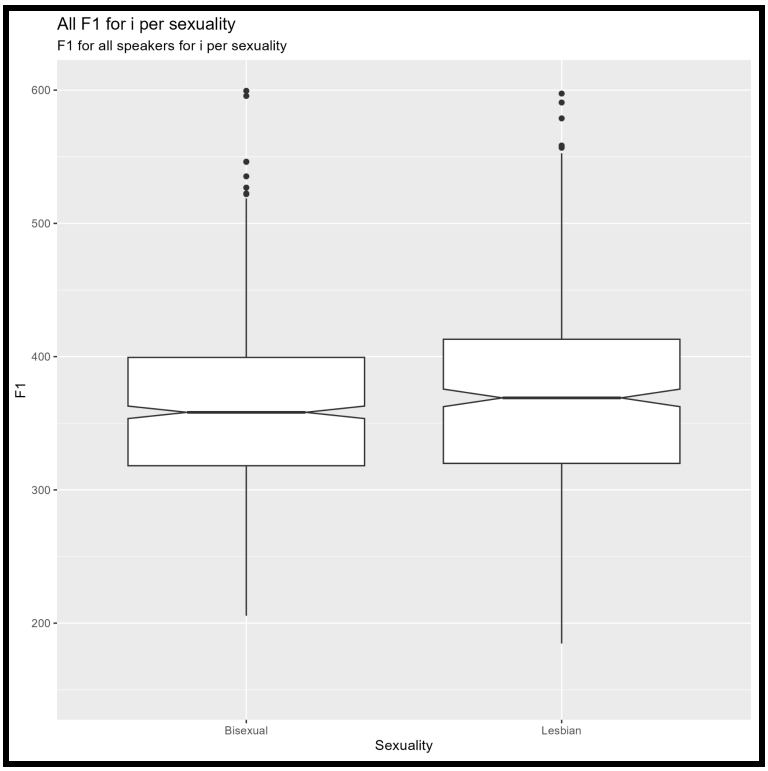
Plot 6A - F1 for all speakers for /a/ by sexuality



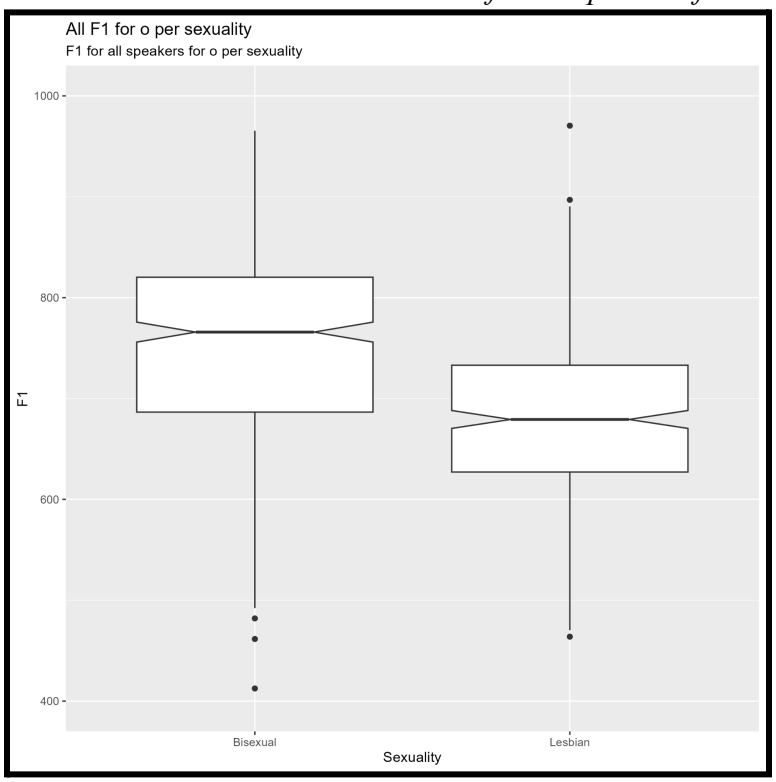
Plot 6B - F1 for all speakers for /æ/ by sexuality



Plot 6C - F1 for all speakers for /ɛ/ by sexuality



Plot 6D - F1 for all speakers for /i/ by sexuality



Plot 6E - F1 for all speakers for /o/ by sexuality

The main observation is that these F1 plots are a lot more evenly divided, as there are two plots with overlapping medians and three plots with the bisexual median being significantly higher than the lesbian median. The only plots with overlapping medians are the /ɑ/ and /i/ plots. In addition, the plot with the largest gap between the medians is the /o/ plot. The trends from the previous models and plots are continued here and are in the expected direction, as the bisexual plots have higher medians. However, these trends are not as strongly supported as previous plots, as only three of the plots have higher bisexual plots.

F1 Discussion

For the F1 plots, the per speaker models were very similar to the center of gravity plots Again, speakers 1, 4, and 6 had higher medians than the rest of the speakers for most of the plots. The gender presentation plots also continued the general trends of the center of gravity plots, with four of the five plots having the *fem & masc* variable median as higher than the *feminine* variable median. However, the F1 plots have some different results for the sexuality plots. Only three of the five plots had the bisexual median as higher than the lesbian median, which means that the trend from the center of gravity sexuality plot is not as strong in the F1 sexuality plots.

These models and plots seem to suggest that there is a significant effect of gender presentation and sexuality on F1. However, the trends from the previous models are not as strongly supported here, as the direction of the effects in the models and plots is unexpected. Most of the F1 models had the bisexual and *fem & masc* coefficients as lower, while the sexuality plots were more evenly split. This implies that the conjecture that bisexual and *fem & masc* speakers have higher formant frequencies is not supported based on the coefficients, even if these variables still seem to have a significant effect. Overall, a significant proportion of the

variance in F1 seems to be accounted for by gender presentation and sexuality, even if the trends from the previous models and plots did not continue as strongly here.

F2 Iteration Tests

For the F2 model, the equation is *F2 ~ sexuality * gender.category + (1 | speaker),* and there are six models for each vowel. The results of this model can be seen here:

Measurement	Intercept: Gender Presentation + Sexuality + Following Vowels (Feminine * Lesbian) (Hz)	Sexuality: Bisexual (Hz)	Gender Presentation: Fem & Masc (Hz)	Sexuality * Gender Presentation (Bisexual * Fem & Masc) (Hz)
F2 ɑ Variance = 405	1607	-4, NS	-47 **	-8, NS
F2 æ Variance = 939	1715	-5, NS	-14, NS	10, NS
F2 ɛ Variance = 796.4	1739	-9, NS	-60 **	17, NS
F2 i Variance = 2707	2277	28, NS	-95 **	-72 *
F2 o Variance = 879.3	1289	-45 **	-19, NS	-10, NS

Table 8F - F2 Iteration Models

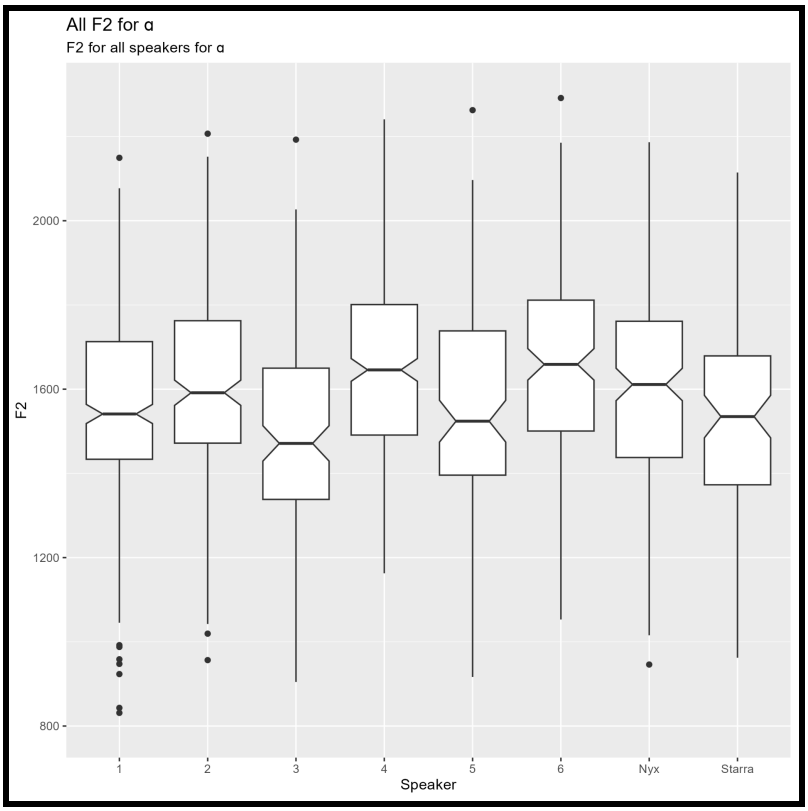
The /ɑ/ model main intercept has a coefficient of 1607 Hz, and the *fem & masc* coefficient has a 47 Hz **. The /ɛ/ main intercept has a coefficient of 1739 Hz, and the *fem & masc* coefficient has a 60 Hz **. The /i/ main intercept has a 2277 Hz, and the *fem & masc* coefficient is -95 Hz **, and also a bisexual * *fem & masc* coefficient of -72 Hz *. The /o/ main intercept has a coefficient of 1289 Hz, and the bisexual coefficient is -45 Hz **.

Unlike the F1 iteration models, there are only three vowels with significant values for the gender presentation variable and only /o/ has a significant value for the sexuality variable. /ɑ/, /æ/, /ɛ/ are typically the three vowels that have significant values for gender presentation across the previous models, but /æ/ is not significant for F2 gender presentation intercepts. Instead, /i/ has a significant value for the gender presentation intercepts. In addition, the one vowel which has a significant relationship with the bisexual * *fem & masc* intercept is now /i/, when it was /o/ for the F1 models.

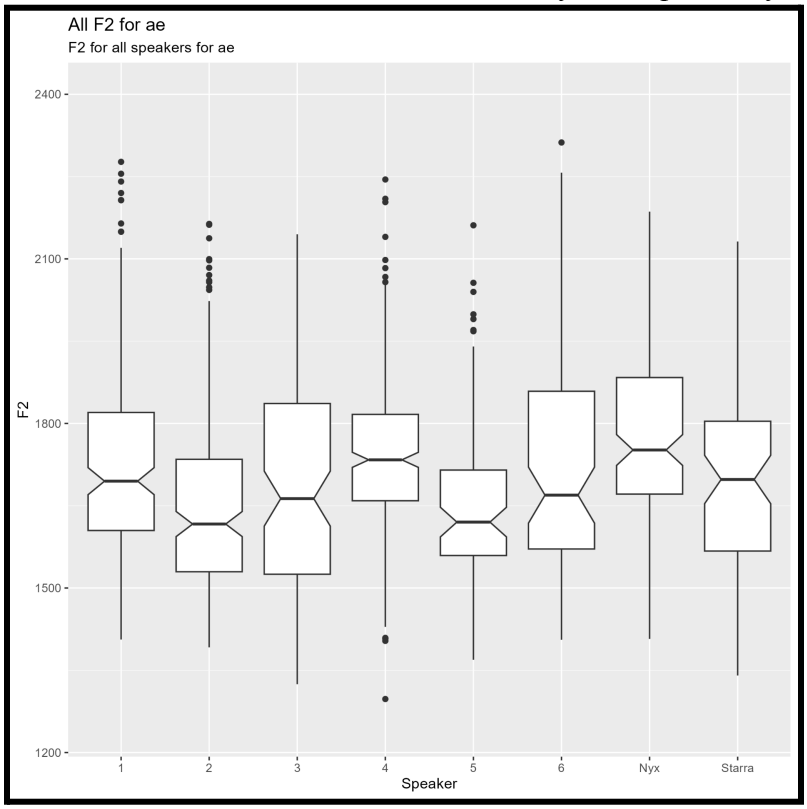
These F2 iteration models do not support the trends and the direction of the effects are unexpected in comparison to the previous models and plots, as the bisexual and *fem & masc* coefficients are lower, similar to the coefficients for the F1 models. However, it still seems like there is a significant effect of gender presentation and sexuality on F2, which is supported by the significant relationships for the gender presentation and the bisexual * *fem & masc* intercepts. For sexuality, this does not seem as significant as the F1 models, since only one vowel has a significant relationship instead of two.

F2 All speakers' Plots

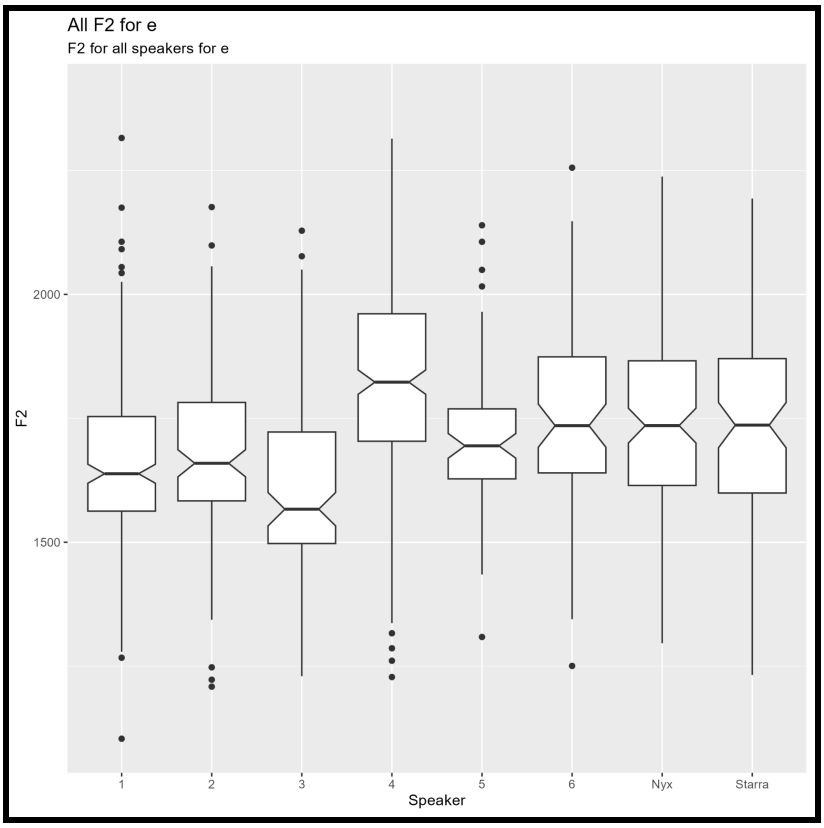
Next, the F2 plots per speaker for each vowel (3A through 3E) will be examined. These plots can be seen here:



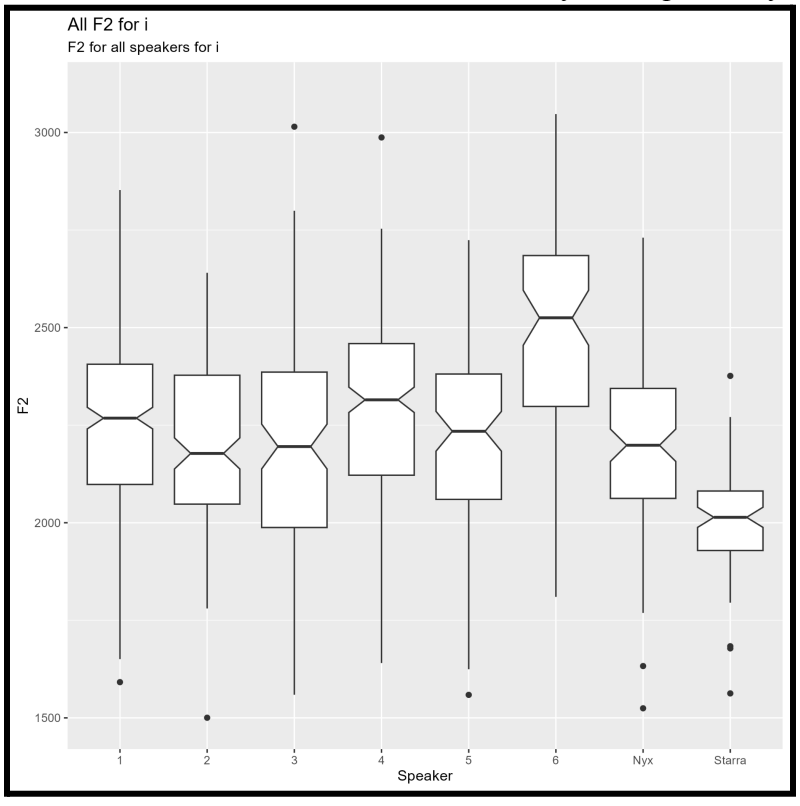
Plot 3A - F2 for all speakers for /a/



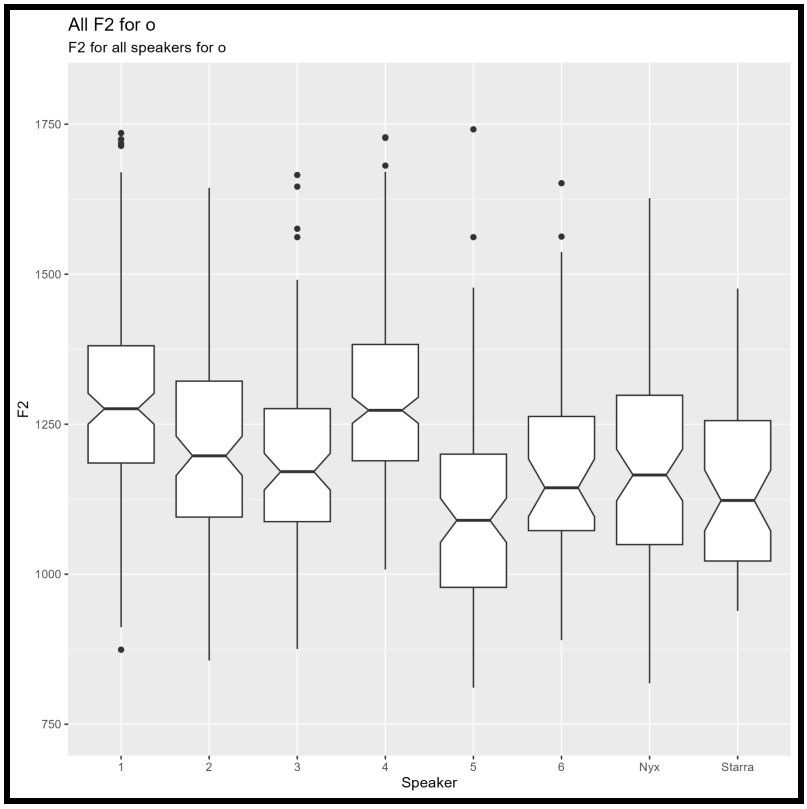
Plot 3B - F2 for all speakers for /æ/



Plot 3C - F2 for all speakers for /ɛ/



Plot 3D - F2 for all speakers for /i/

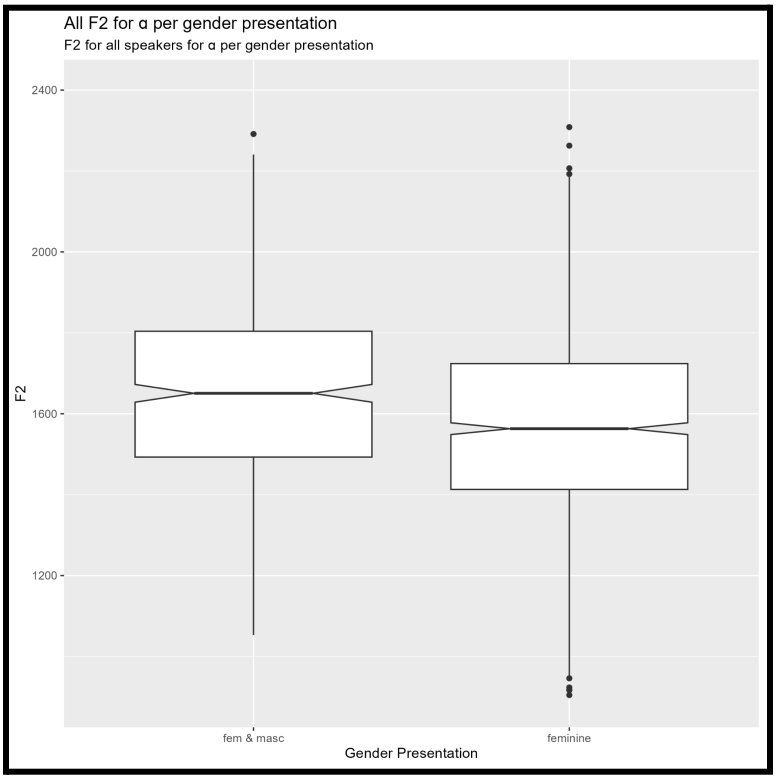


Plot 3E - F2 for all speakers for /o/

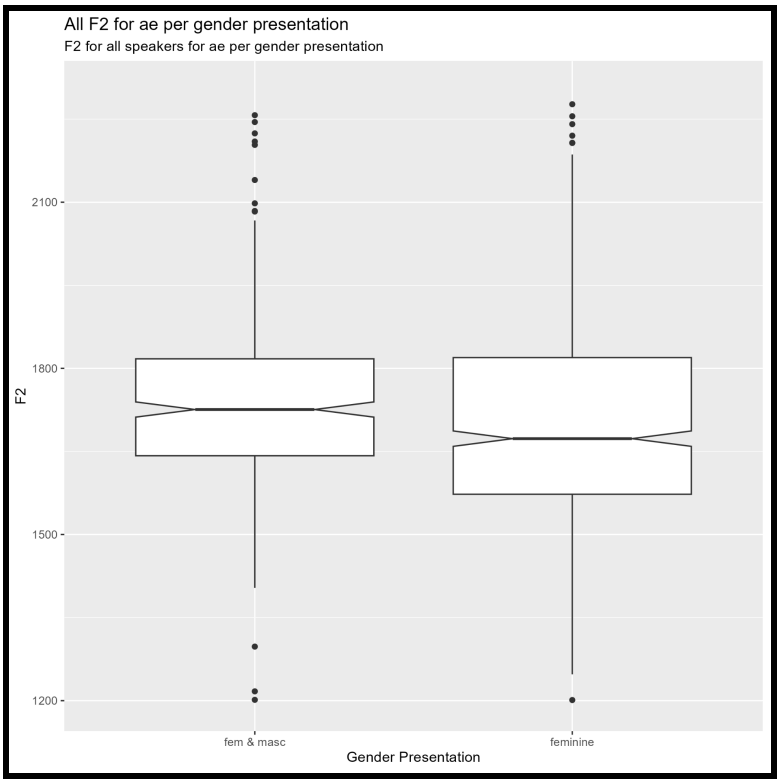
The only noteworthy speakers is speaker 4 for the /ɛ/ plot, speaker 6 and Starra for the /i/ plot, and speakers 1 and 4 for the /o/ plot. The only speakers who had a significantly lower value than the rest across the plots was Starra for the /i/ plot and speaker 3 for the /ɛ/ plot. Across these F1 and F2 plots, only speakers 1, 4, 6 and Starra seemed to have medians that differed significantly, with speakers 6 and Starra having significantly lower values. Interestingly, speaker 6 is the only speaker with significantly lower and higher values than the rest of the group across these plots. These results are similar to the findings from the center of gravity plot, which also found that speaker 1, 4 and 6 had significantly higher values. It is notable that the same speakers also had some significantly higher values across these plots, even though the significance of this meaning is not known yet. Overall, these plots seem to continue the trends from the previous plots and in the expected direction, as speakers 1, 4, and 6 are the only speakers with higher values across multiple plots.

F2 Gender presentation plots

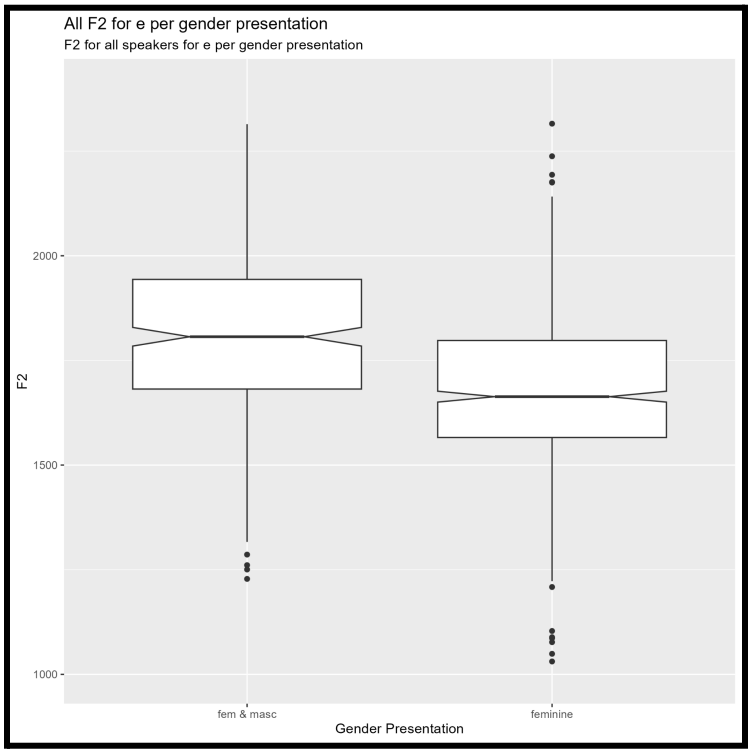
In contrast to the F1 plots, none of the F2 gender presentation plots had overlapping medians for the two plots. The results of these plots can be seen here:



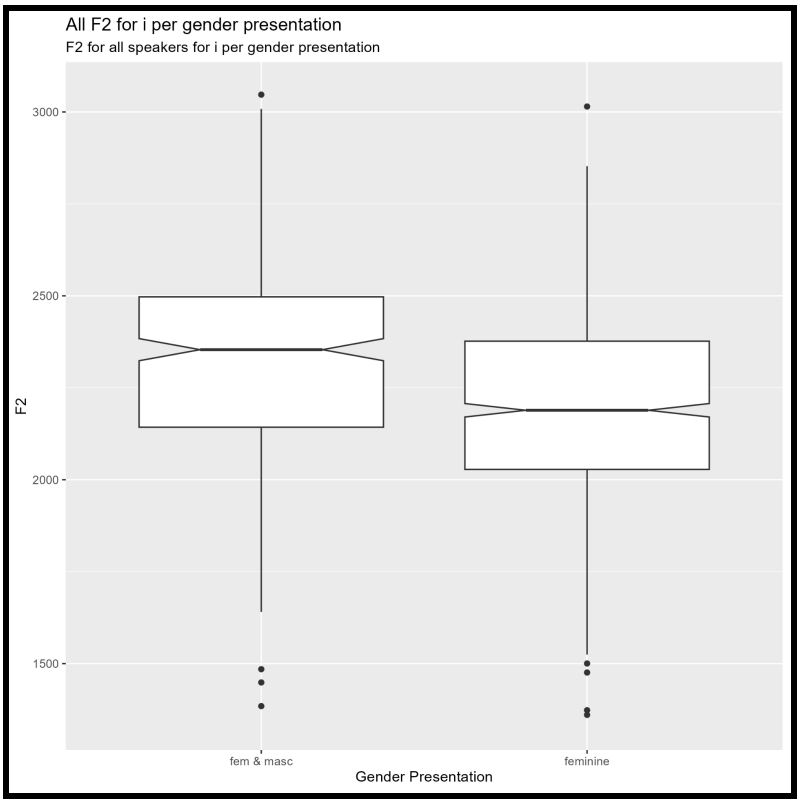
Plot 5A - F2 for all speakers for /a/ by gender presentation



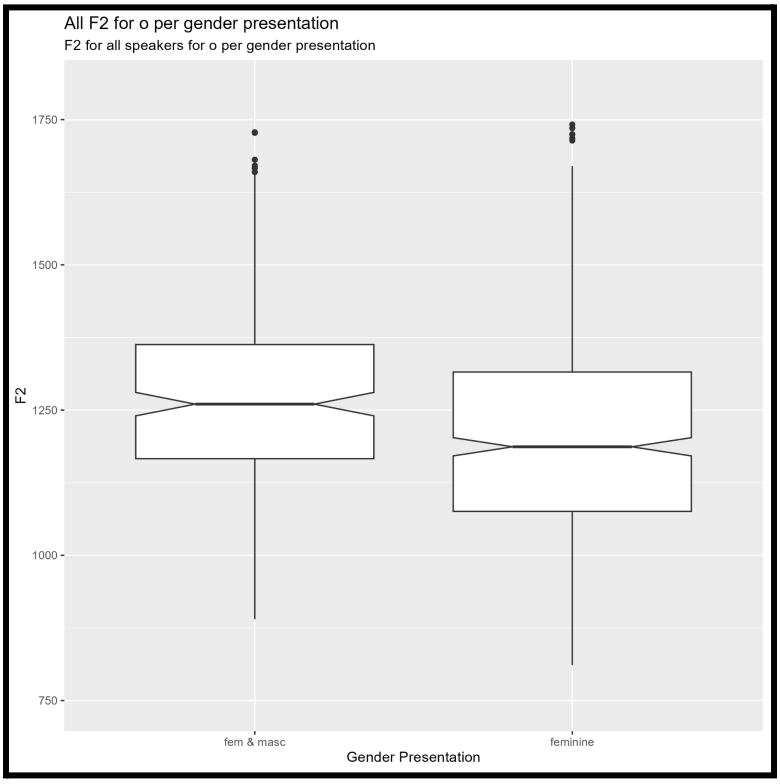
Plot 5B - F2 for all speakers for /æ/ by gender presentation



Plot 5C - F2 for all speakers for /ɛ/ by gender presentation



Plot 5D - F2 for all speakers for /i/ by gender presentation

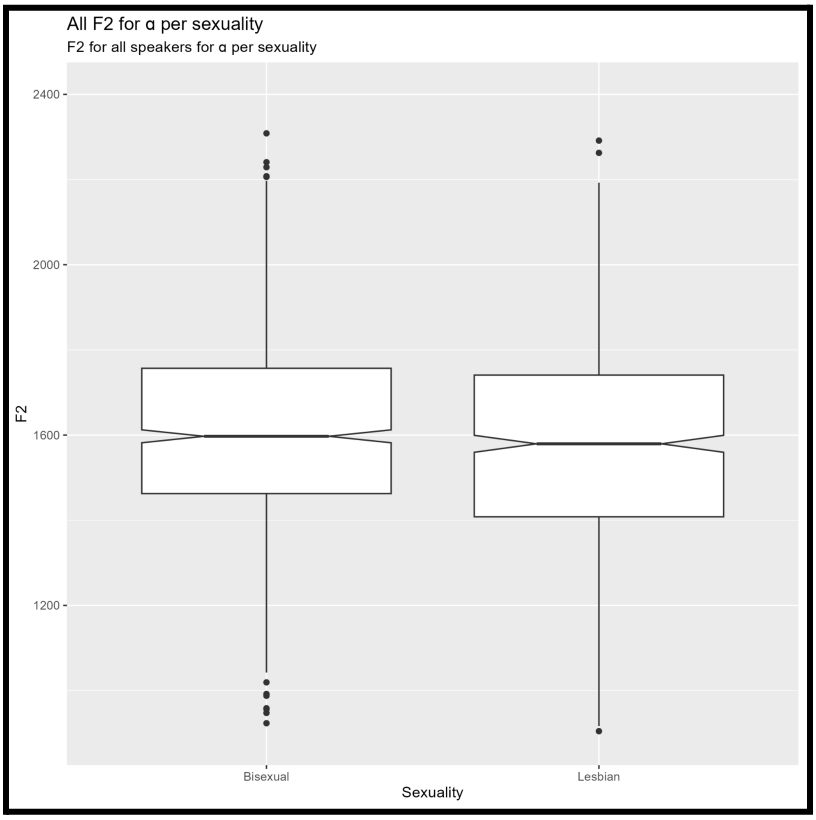


Plot 5E - F2 for all speakers for /o/ by gender presentation

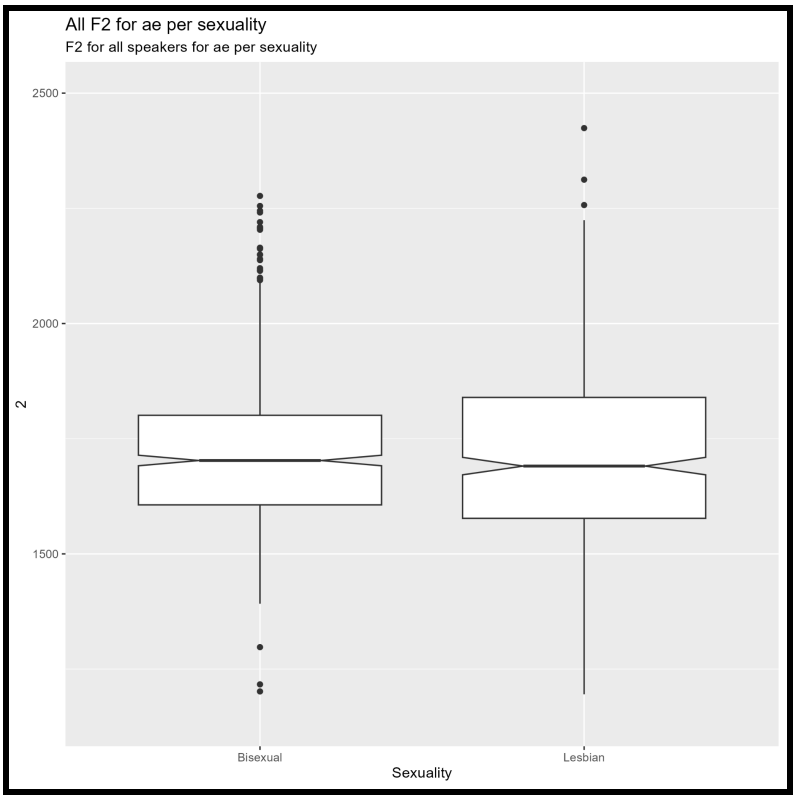
Even the /i/ F2 plot, which did have overlapping medians for the F1 plot, did not have an overlapping median. The two plots that had the closest medians was the /ɑ/ and /æ/ plots. Overall, nine of the ten gender presentation plots had the *fem & masc* plot have a significantly higher median than the *feminine* plot, to the point where their medians did not overlap. The trends from the previous plots are continued here and in the expected direction, as the *fem & masc* plot has higher medians in the majority of the plots.

F2 Sexuality plots

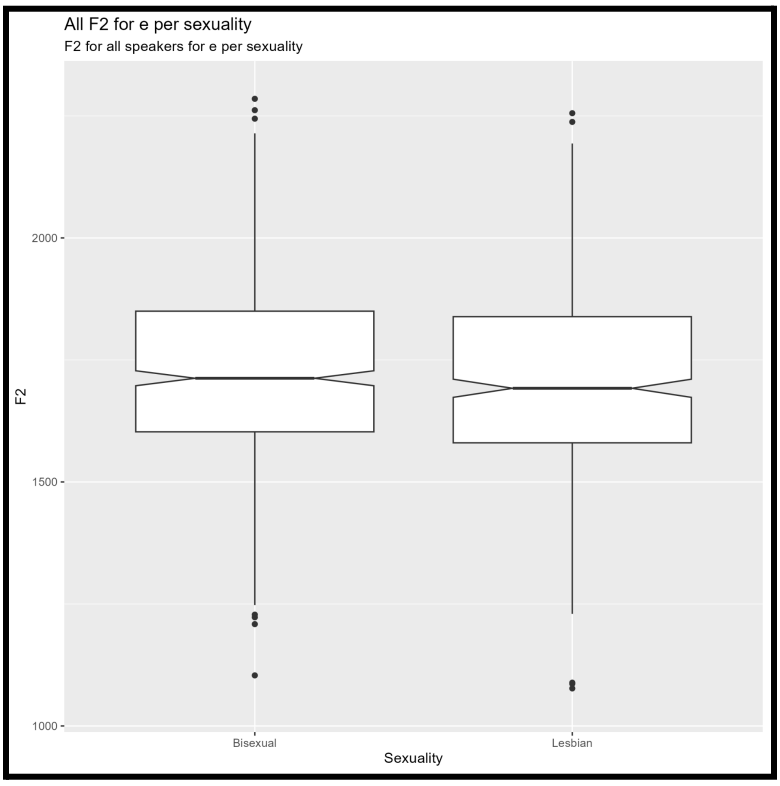
The F2 sexuality plots are also different from the F1 plots, and they can be seen here:



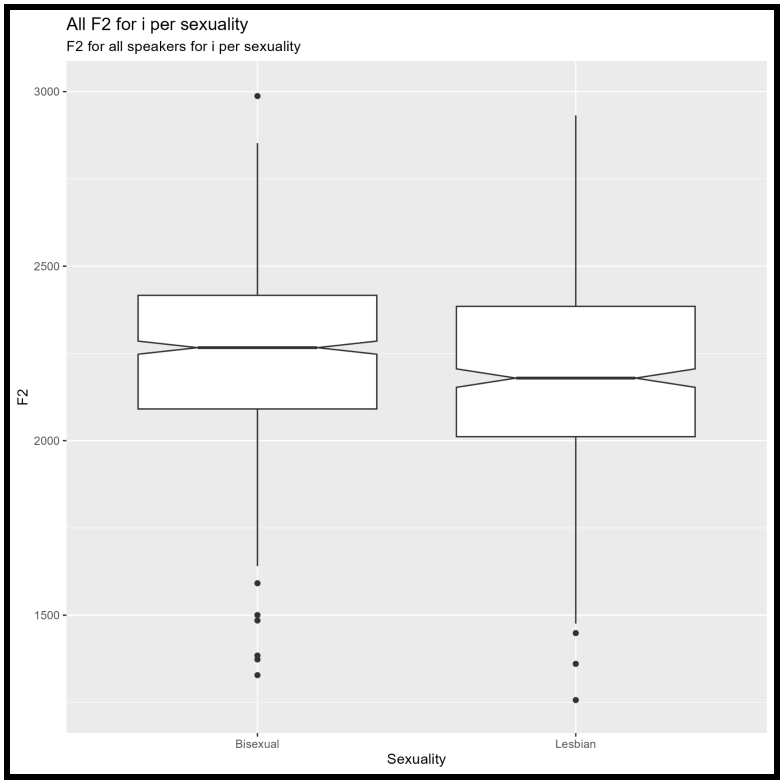
Plot 7A - F2 for all speakers for /a/ by sexuality



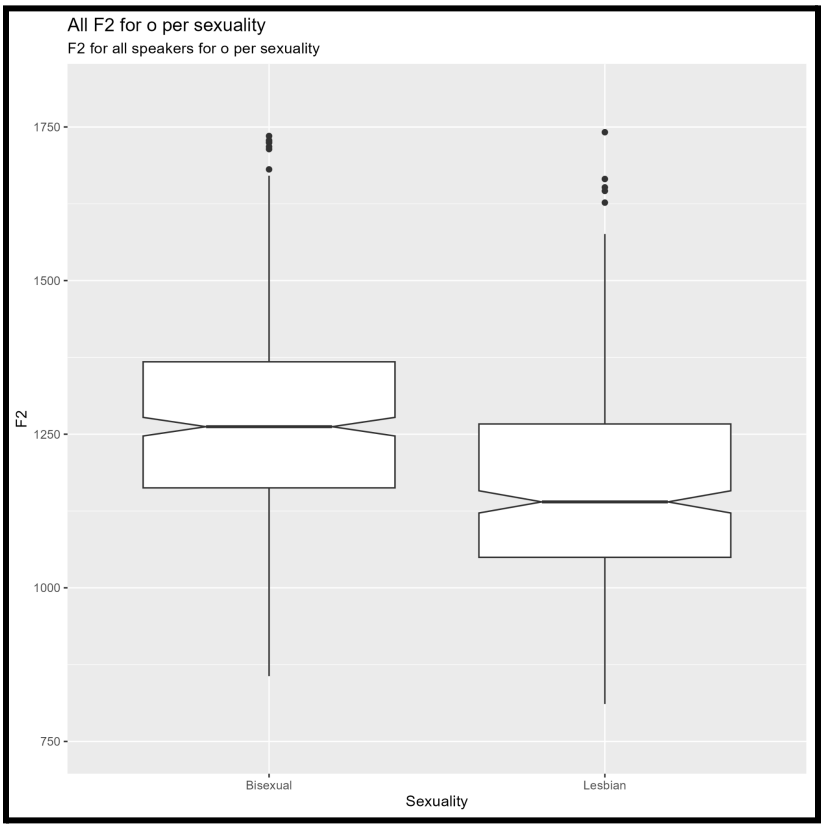
Plot 7B - F2 for all speakers for /æ/ by sexuality



Plot 7C - F2 for all speakers for /ɛ/ by sexuality



Plot 7D - F2 for all speakers for /i/ by sexuality



Plot 7E - F2 for all speakers for /o/ by sexuality

The F2 plots differ from the F1 plots again, with three F2 plots with overlapping medians and two plots without overlapping medians. The three plots with overlapping medians are the /ɑ/, /æ/, and /ɛ/ plots, and the plots without the overlapping medians are the /i/ and /o/ plots. The only plots across both F1 and F2 that are the same are the /ɑ/ and /o/ plots. Finally, the plot with the largest gap between the medians again is the /o/ plot. The trends from the previous F2 plots are continued here and in the expected direction, although the direction is not as significant.

Iteration Tests and Plots Results

For the iteration tests models, the center of gravity models once again have no significant relationships with any of the intercepts, even though the general trend for the variables is present here and the direction of the effects is expected. Overall, there are three vowels that seem to consistently have significant relationships with their models, which are /ɑ/ and /ɛ/ for the gender presentation variable, and o for sexuality across both F1 and F2. Most interestingly, /o/ is the most consistent vowel across all the vowels, in that it always is significant for the sexuality variable across all models, it is one of two vowels that is significant for the gender presentation and sexuality interaction, and the only vowel that has had a significant value for all three intercepts.

The one vowel that consistently had a significant relationship with the sexuality variable, /o/, is present again in the F1 models. The biggest difference between the F1 and F2 models again is that /i/ has significant relationship with the bisexual and *fem & masc* interaction for F1, and it is /o/ that has significant values for the interaction for F2. Occasionally, /æ/ and /i/ have significant relationships with some of the intercepts. However, these are not as consistent as /ɑ/, /ɛ/, and /o/. This also means that the only vowels in this model who have significant relationship

with the sexuality and gender presentation interaction, /o/ and /i/, are the high back and high front vowels, respectively.

The F1 and F2 models largely follow the patterns of the previous models, as both gender presentation and sexuality have a significant effect on F1 and F2. However, the general trend is not continued in the F1 and F2 models and direction of the effects for the coefficients is unexpected. The majority of the coefficients for bisexual and *fem & masc* are lower than the main coefficients. These results were not present in the previous individual effects and additive models, as the direction of the effects was more consistent with the bisexual and *fem & masc* coefficients being higher than the main coefficients. For these F1 and F2 iteration models, the coefficients do not support the conjecture that the bisexual speakers and *fem & masc* speakers will have higher values, despite the fact that gender presentation and sexuality seem to have significant effects on F1 and F2. Overall, it seems that a significant proportion of the variance in F1 and F2 can be accounted for by gender presentation and sexuality.

Overall Results

Across the individual, additive and iteration tests models, the general trends seem to be fairly stable for the center of gravity models and the F1 models. Across all of the center of gravity models, gender presentation and sexuality do not have a significant effect on center of gravity. In contrast, across all of the F1 and F2 models, it seems like gender presentation and sexuality have a significant effect on F1 and F2. For these models, the general trend is that the same three vowels have significant values for the gender presentation variable and the same one vowel has significant vowels for the sexuality variable. The only major differences seen in these models are the F2 models. /ɑ/ and /o/ are present in every model, but /æ/, /i/, and /ɛ/ are not

consistent across the models. However, /i/ is the only other variable for F2 besides /o/ that has any significant relationship with the sexuality variable.

For all of the vowel models across both F1 and F2 though, /ɑ/ is the only vowel that did not have a significant relationship for the sexuality intercept at any time. Looking at these outliers, /ɑ/ only ever had a significant relationship with the gender presentation intercept, and /o/ only had a significant relationship with the sexuality intercept. The rest of the vowels across both F1 and F2 had significant relationships for both variables at least one time across all of the models. The meaning of this variation is not known at this point, such as why it is specifically these back vowels that are consistent in what variables they are attached to, but it should still be noted.

Across all of the plots, there do seem to be some general trends. For the per speaker plots, it seems like speaker 1, 4 and 6 are the only speakers who have had medians that are significantly higher than the rest of the speakers, who often have overlapping medians. While speakers 1 and 6 would occasionally have lower values depending on the vowel, speaker 4 seemed to have consistently higher values across center of gravity, F1 and F2. For the gender presentation plots, there was very little variation in the results, with the *fem & masc* plots having significantly higher medians for ten of the eleven plots and overlapping medians for one of the plots.

The sexuality plots had more even results, as there were five plots with overlapping medians and five plots with the bisexual plot having a significantly higher median. However, there were no plots that had the lesbian plot higher than the bisexual plot. The most notable observation from the sexuality plots is that the /o/ plot for F1 and F2 had significantly higher medians than the rest of the vowels. This continues the pattern seen from the LMMs, where the

/o/ models had significant relationships with the sexuality intercepts. Since the results of the plots and models are finished being discussed, the mean for each measurement per speaker can be examined.

The general trends across all the models and plots suggest that there are significant effects of the two variables, and that the direction of the effects is expected. However, there are individual instances within the models where the direction of the effects is unexpected. The coefficients across these models vary, with the individual and additive models supporting the conjecture that bisexual and *fem & masc* speakers have higher formant frequencies, while the iteration models do not support it. In conclusion, for the models and plots, a significant proportion of variance in the F1 and F2 models can be accounted for by gender presentation and sexuality. This seems to suggest that the strongest predictor of F1 and F2 was these two variables.

Question Types Models and Plots

As discussed in the methodology, there were two tasks for this project, which were the interview task and the reading task. Up until now, these tasks have not been separated in the data, since the type of question each data point came from was not relevant to the categorical variables being investigated. However, the question of whether the measurements vary based on the type of task used is still relevant. Zimman (2017b) stated that reading tasks allow for participants to put on a "performance" of their gender identity, which is directly tied to their gender presentation. By doing so, the speakers may then use this performance during the process of constructing their gender presentation. Due to the low number of tokens, all of the speakers will be analyzed at once, rather than examining each speaker's results individually. Like the previous

models, each model's results will be shown and their relationships will be analyzed to examine the influence of the intercept on the measurement.

The previous models examined the effect of the fixed effects, gender presentation and sexuality, on the measurement and whether the fixed effects had a significant relationship, such as whether they were affecting the outcome of the measurement. The intercepts and relationships gathered from each model conveyed the relationship between the intercept, like *fem & masc* intercept or bisexual intercept, and the measurement. In these models, however, they show the effect of the question type on the measurement. In addition, plots will be used to convey the relationship between the variable and the measurements. These plots can be seen in plots 8A through 8C. Plot 8A shows the distribution of the tokens based on the type of task, which is the interview questions or the reading passage task, and represents every speaker. Plot 8B shows the distribution based on the gender presentation, and plot 8C shows the distribution based on sexuality. If an intercept has a significant relationship, that means that it is likely that the question type had an effect on the measurement. If an intercept does not have a significant relationship, that will mean that the question type did not have an effect.

CoG

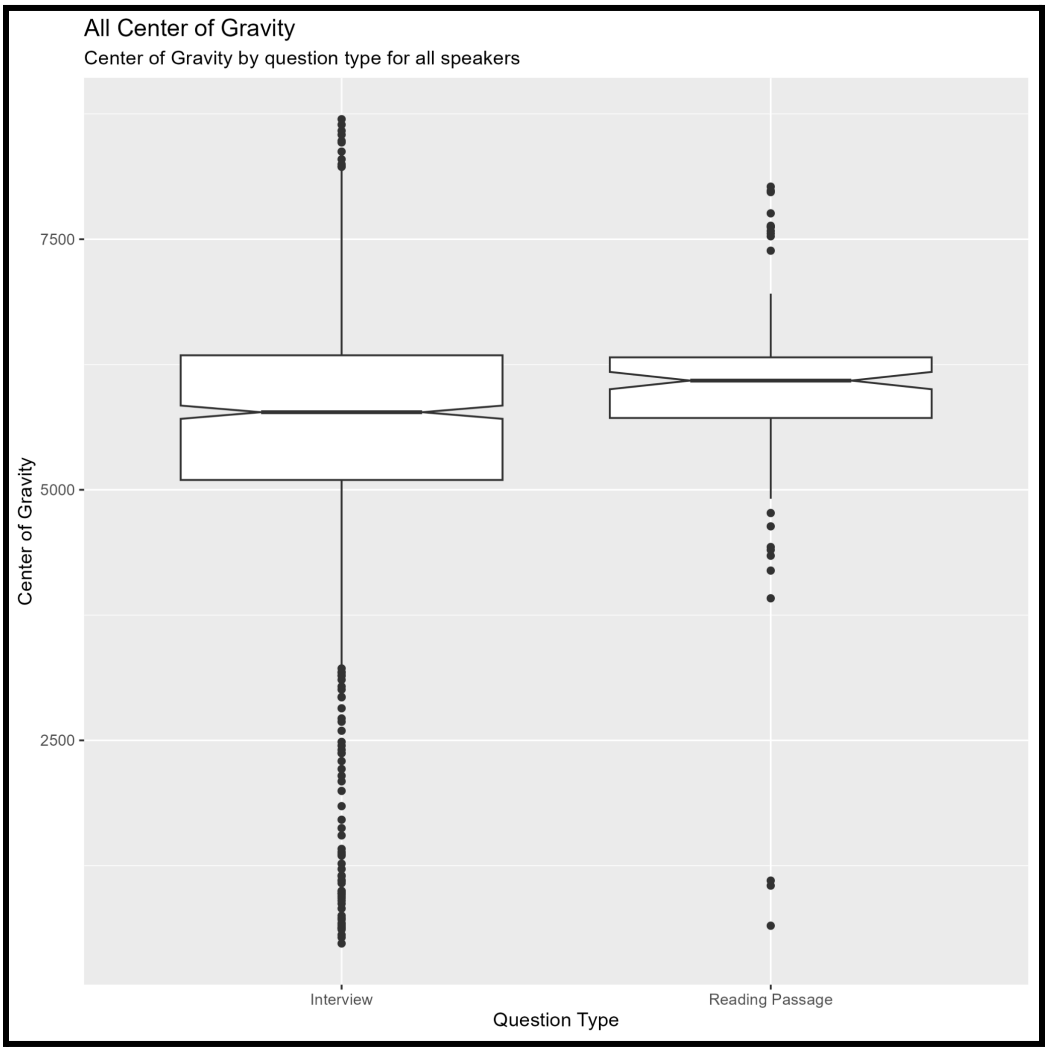
The first model shown will be the center of gravity model. The equation used for this model is *cog ~ question.type + (1 | speaker)*. Like previous models, the response variable for this equation is the measurement, in this case the center of gravity. There is only one fixed effect this time, which is the type of question, meaning whether the task was the interview task or the reading passage task. Also, the random effects variable is still the speaker, since all of the data is being examined at once, rather than the data per speaker. The results for this model can be seen here:

Measurement	Question Type - (Intercept) (Hz)	Question Type - Reading Passage (Hz)
Center of Gravity (Hz)	5505	445 ***

Table 12A - Center of Gravity Individual Effects Models by Question Type

For the center of gravity model, the random variance is 153386. For this model, the main intercept (which is the interview task) coefficient is 5505 Hz. For the reading passage task, the intercept is 445 Hz ***. There does seem to be some effect of the type of task on center of gravity, and the differences across these coefficients are in the expected direction.

Plot 8A, which is the center of gravity plot, can be seen here:



Plot 8A - Center of Gravity for all speakers by question type

The initial observation from plot 8A shows a clear difference of the medians of the two plots. The reading passage task median and the interview task median clearly do not overlap, and there is a significant difference between the two. The reading passage IQR, minimum and maximum is smaller than those of the interview task, but this may be due to the smaller amount of tokens from the reading passage task.

F1

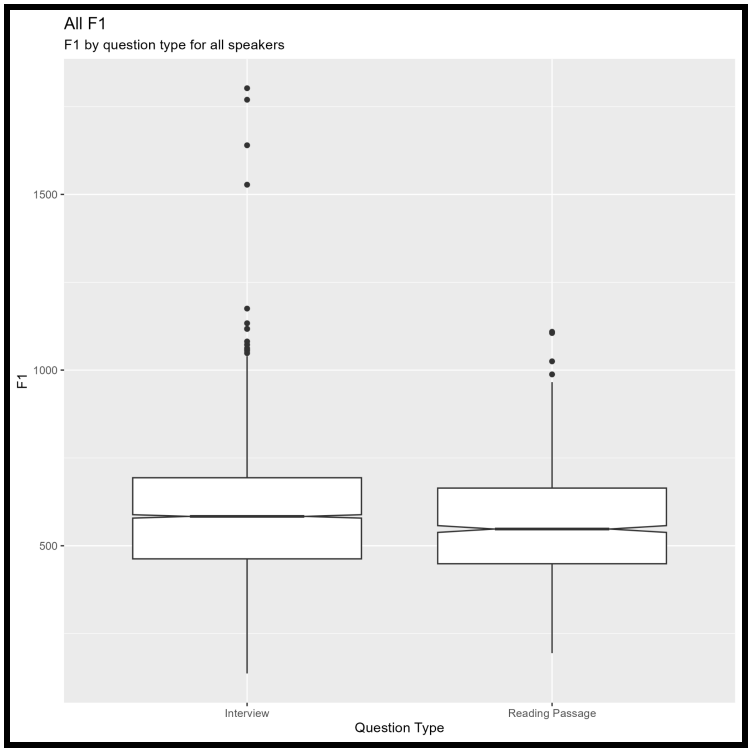
The next model shown will be the F1 model. The equation used for this model is *F1 ~ question.type + (1 | speaker)*. The response variable this time is F1, and the random effects variable is the same as the center of gravity model, and the random variance is 659.2. The results of this model can be seen here:

Measurement	Question Type - (Intercept) (Hz)	Question Type - Reading Passage (Hz)
F1 (Hz)	574	-21 ***

Table 12B - F1 Individual Effects Models by Question Type

For this model, the main intercept coefficient is 574 Hz. For the reading passage task, the intercept is -21 Hz ***. The trend from the center of gravity model continues here, and the effects were in the expected direction again.

Plot 8B shows a different result from the previous models and plots, and can be seen here:



Plot 8B - F1 for all speakers by question type

The medians of both tasks do not overlap again, but that the reading passage task has a lower median this time. This time, they had very similar IQRs as well as similar minimums and maximums. The differences in Plot 8B were in the unexpected direction, as the interview plot had a higher median. The general trends from the previous models and plots are not supported here.

F2

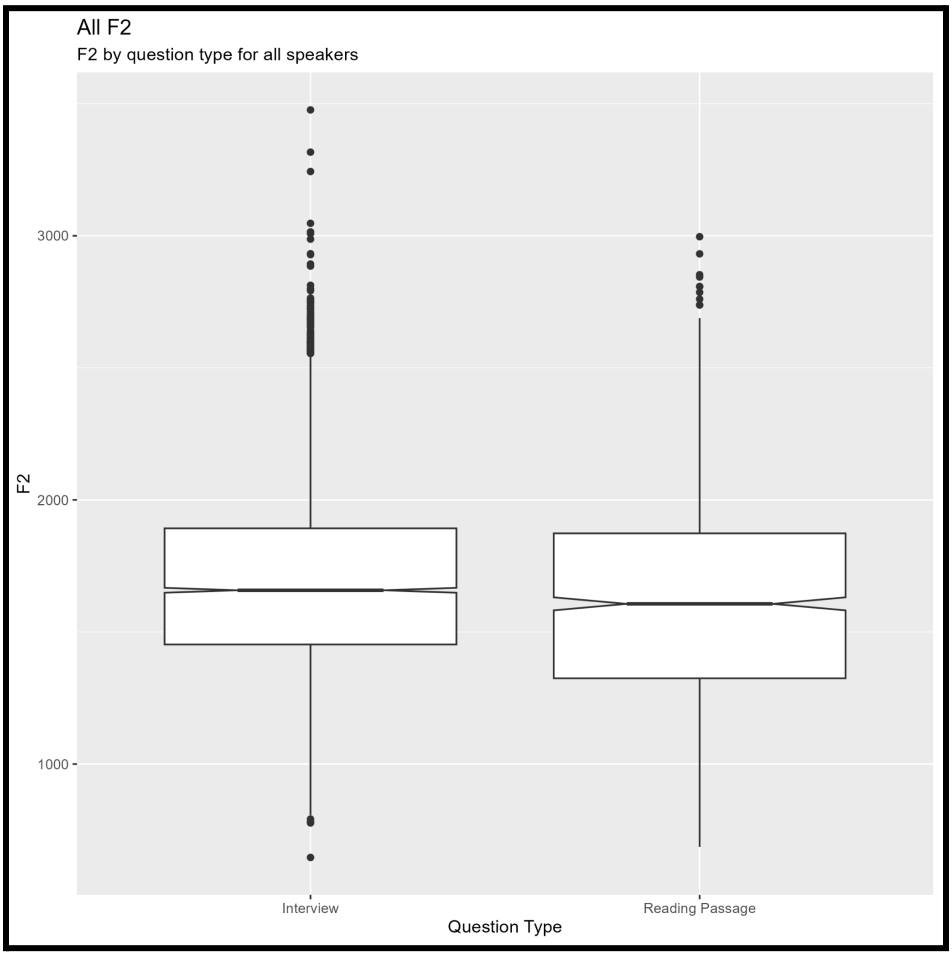
The final model shown will be the F2 model. The equation used for this model is *F2 ~ question.type + (1 | speaker)*. The response variable is F2, and the random variance is 2451. The results for this model can be seen here:

Measurement	Question Type - (Intercept) (Hz)	Question Type - Reading Passage (Hz)
F2 (Hz)	1676	-50 ***

Table 12C - F2 Individual Effects Models by Question Type

For this model, the main intercept is still the same, and the coefficient is 1676 Hz. For the reading passage task, the intercept is -50 Hz ***. This final model continues the general trend of these two different tasks having a significant effect on the F2 values.

Plot 8C shows the final plot, and can be seen here:



Plot 8C - F2 for all speakers by question type

It seems like there is a very slight overlap between the reading passage task median and the interview median in Plot 8C. However, the reading passage task median would be below the interview task median if they are not overlapping. The rest of the box plot components also seem to be very similar, just like the F1 plots. Similar to the F1 plot, the differences across the plots were in the unexpected direction, as the interview task had either a higher median or overlapping

medians. The general trends from the models and the center of gravity plot are not supported here.

Question Type Results

Overall, these results show that the reading passage task only had a larger median for center of gravity, but consistently lower results for F1/F2. This seems to suggest that the type of question did not have a significant effect across all of the models. While the relationships between the coefficients seemed to be significant for the type of task, these results do not follow the same pattern for the plots. The direction of the effects were not consistent across the models and plots, and the general trends were not supported. Due to the inconclusive evidence, it does not seem like it can be argued that a significant proportion of the variance for the measurements can be accounted for by the question type. It does not seem like the strongest predictor of the measurements was the type of question. More investigation into this would be required to make an accurate conclusion.

Means

For the analysis of the mean measurement for every speaker, each mean for center of gravity, F1 and F2 has been calculated and can be seen in Tables 9 through 11. These tables show each speaker's mean for each measurement, and then ranks them compared to the other speakers. For example, a speaker with (3) next to their value in the table has the third highest mean for that measurement, such as the third highest center of gravity mean. A speaker with an (8) next to their value has the eighth lowest mean for the measurement, which would equal the lowest center of gravity mean among all of the speakers. Examining the average ranking per speaker could

potentially reveal more about the relationship between the measurements and the categorical variables.

All Rankings

The average of all of the rankings can be found in Table 11. For the center of gravity rankings, the means were selected per speaker based on the words with initial /s/ followed by a vowel. The data collected for /s/ center of gravity did include both words with initial /s/ followed by both vowels and consonants, however the /s/-consonant cluster words will not be used to maintain consistency since only the /s/-vowel words were used in the linear mixed-effects models. For the F1 and F2 rankings, they were separated by vowel, which can be seen in tables 9 and 10, and then the average was found by getting the average of the five vowels. The initial observation for the center of gravity rankings (Table 11) is that speakers 1, 4, and 6, who had consistently the highest medians from the plots, have the three highest center of gravity means. The bottom three speakers are Nyx, speaker 5, and speaker 2.

Before getting into the F1 and F2 average rankings, it is necessary to check how the means vary per vowel for each speaker. The F1 and F2 rankings can be seen here in Table 9 and 10:

Speaker	ɑ F1	æ F1	ɛ F1	i F1	o F1
1	520.1 (3)	759.0 (2)	629.1 (3)	355.9 (7)	658.2 (2)
2	516.0 (4)	690.2 (7)	597.6 (4)	373.2 (5)	630.8 (3)
3	514.2 (5)	696.4 (6)	593.0 (6)	381.3 (2)	625.1 (5)
4	552.7 (2)	800.9 (1)	645.3 (2)	370.9 (6)	713.1 (1)
5	510.9 (6)	702.1 (4)	586.9 (7)	378.3 (3)	609.6 (8)
6	562.9 (1)	753.4 (3)	657.2 (1)	399.3 (1)	613.2 (7)

Nyx	507.1 (7)	662.8 (8)	596.1 (5)	374.0 (4)	629.5 (4)
Starra	480.6 (8)	698.1 (5)	561.1 (8)	333.4 (8)	614.9 (6)

Table 9 - F1 Average Rankings

Speaker	ɑ F2	æ F2	ɛ F2	i F2	o F2
1	1549.8 (5)	1731 (3)	1661 (7)	2244 (3)	1327.9 (2)
2	1598.6 (4)	1661 (7)	1682 (6)	2206 (4)	1325.7 (3)
3	1507.1 (8)	1696.1 (6)	1608.2 (8)	2147 (7)	1249.1 (5)
4	1651 (2)	1747 (2)	1827 (1)	2272 (2)	1344.7 (1)
5	1541.0 (6)	1656 (8)	1705 (5)	2202 (5)	1176.5 (8)
6	1659 (1)	1714 (4)	1772 (2)	2473 (1)	1274.4 (4)
Nyx	1599 (3)	1770 (1)	1727 (3)	2198 (6)	1238.1 (6)
Starra	1529.7 (7)	1697 (5)	1712 (4)	1998 (8)	1180.2 (7)

Table 10 - F2 Average Rankings

In Table 9, speakers 1, 4 and 6 have some interesting patterns with their vowels. Speakers 1 and 4 have the top 3 highest vowel means for /ɑ/, /æ/, /ɛ/, and /o/. However, they drop to the 7th lowest and 6th lowest mean respectively for /i/. Speaker 6 is the same, as they are in the top 3 highest vowel means for /ɑ/, /æ/, /ɛ/, and /i/, but suddenly drop to the 7th lowest mean for /o/. For the other speakers, there aren't any consistent trends. Speaker 2 drops to the 7th lowest mean for /æ/ but stays between 3rd and 5th place for the other vowels. Speaker 3 stays around 5th and 6th for almost all of the vowels, but then has the second highest mean for /i/. Speaker 5 is in the bottom half of the means for /ɑ/, /ɛ/, and /o/, but then has the 4th mean for /æ/ and 3rd for /i/. Nyx goes through 4th and 8th for every vowel, and Starra often was in the lower half. However, these trends are not as significant as the patterns that speakers 1, 4 and 6 have with their vowels.

For the F2 vowel means per vowel, the results are different. Speaker 1 does not have a consistent trend for their vowel means anymore, while speakers 4 and 6 still are very consistent in their rankings. However, that sudden drop that speaker 4 had for /i/ and speaker 6 had for /o/ are not present in the F2 rankings. In addition, Nyx, who had the majority of their means be in the bottom 4 for F1, has the third highest mean for /ɑ/, the highest mean for /æ/, the third highest mean for /ɛ/, and then the 6th lowest means for /i/ and /o/. Beyond that, none of the other speakers have notable trends. Overall, the means for speaker 4 and 6 are consistently in the top three across both F1 and F2. The only other notable trend across both rankings is the pattern that speakers 1 and 4 share in F1 and the drop that speaker 6 experienced for /o/ in F1.

All of the rankings across these three measurements were averaged up and put into the all rankings average. Looking at Table 11, the trends spotted in the rankings for each measurement are consistent in the averages rankings. Speakers 1, 4, and 6 have significantly higher rankings than the other five speakers, with speakers 2 and 3 sharing an average ranking of 5.33, which is the 4th/5th highest ranking. The speaker with the lowest average ranking overall is Starra, with an average ranking of 6.33 across the three measurements. Despite Starra having the 4th highest center of gravity mean, they had the lowest F1 and F2 means.

Rankings Discussion

The average rankings per measurement can be seen in Table 11:

Speaker	CoG Rankings	F1 Rankings (all vowels)	F2 Rankings (all vowels)	All Rankings Average
1	6283.7 (1)	584.46 (3)	1702.74 (4)	2.6 (3)
2	5239 (6)	561.56 (5)	1694.66 (5)	5.33 (4/5)
3	5551 (5)	562.0 (4)	1641.5 (7)	5.33 (4/5)

4	5776 (3)	616.58 (1)	1768.34 (2)	2 (2)
5	5198.1 (7)	557.56 (6)	1656.1 (6)	6.33 (7)
6	5919 (2)	597.2 (2)	1778.48 (1)	1.66 (1)
Nyx	4785 (8)	553.9 (7)	1706.42 (3)	6 (6)
Starra	5681 (4)	537.62 (8)	1623.38 (8)	6.66 (8)

Table 11 - Average Rankings per Measurements

Examining the means based on gender presentation, the two speakers with the consistently highest medians and means, speakers 4 and 6, also identified with the *fem & masc* presentation. Since every other speaker identified with the *feminine* presentation, it is not worth it to examine them on the basis of gender presentation. The speaker with the third highest average ranking is speaker 1, who is bisexual with a *feminine* presentation, and then speakers 2 and 3, who share the 4th/5th highest average ranking, are bisexual with a *feminine* presentation and lesbian with a *feminine* presentation respectively. The bottom three speakers all identify as lesbian with a *feminine* presentation. It is notable that speakers 4 and 6 both have a *fem & masc* presentation but differ in their sexualities. The only other lesbian speaker in the top 4 average rankings is speaker 3, with the rest of the top speakers being bisexual.

These rankings do seem to show some interaction between sexuality and gender presentation from the means, since there is a gap of three to four rankings between speaker 6 and speaker 3 and the only difference between the two speakers is their differences in gender presentation. There are two bisexual speakers in the top 3 rankings, and three lesbian speakers in the bottom 3 rankings, with the lesbian speaker in the top 3 rankings being the only lesbian speaker with a *fem & masc* presentation. Overall, these rankings seem to continue the trend that the plots showed, which is that speakers 1, 4, and 6 have significantly higher values than the rest of the speakers. The mean rankings also potentially show some significant relationship between

the gender presentation of the speaker and the measurements, as the top two speakers in the averages rankings are the only speakers with a *fem & masc* presentation.

Results Conclusion

To conclude this section, the results from the means analysis seems to indicate that there are patterns in regards to how the categorical variables, gender presentation and sexuality, have an effect on center of gravity, F1 and F2. In addition, the data from the linear mixed-effects models as well as the plots support these trends. There seems to be a stronger relationship between gender presentation and the measurements compared to sexuality and the measurements, as speakers 4 and 6 seem to have higher values for the measurements due to their gender presentations. The results of the medians and means indicate that speakers 1, 4 and 6 consistently have higher values for the measurements, although speakers 4 and 6 seem to be more consistent than speaker 1. In addition, the results seem to show that while sexuality may be somewhat significant in predicting whether a speaker has a higher or lower value for the measurements, it seems like gender presentation has a significant relationship with higher values for the measurements.

Overall, the differences in the means rankings were in the expected direction, as speakers 1 and 4 are both bisexual and in the top 3 average rankings, and speaker 6 has the highest average while presenting *fem & masc*. The trends seen in the means rankings continue from the models and plots, and supports the conjecture that the bisexual speakers will have higher values, as three of the four speakers with the highest average rankings are bisexual. In addition, it supports the conjecture that *fem & masc* speakers will have higher values, as the top two speakers in terms of average rankings have a *fem & masc* presentation. All of the results from the models, plots, and means seem to suggest that there is a significant effect of gender presentation

and sexuality on F1 and F2. There is contradictory evidence as to whether there is a significant effect of gender presentation and sexuality on center of gravity, as the models do not support this conclusion. For the F1 and F2 models, plots, and means, the results suggest that the strongest predictor of F1 and F2 is gender presentation and sexuality and that a significant proportion of the variance for F1 and F2 can be accounted for by gender presentation and sexuality.

Analysis

The relationship between center of gravity & F1/F2

As the results have shown from the means analysis, there seems to be some mixed results in terms of how significant a relationship there is between each categorical variable and the measurements. However, as previously discussed, the scope of this research is not to only investigate the macro level interactions between these categorical variables and the measurements. Following the examination of the significance of the effect of these variables on the measurements, an investigation into the micro level interactions by using transcription analysis is necessary as well. By examining how these measurements are used in specific instances, it will show whether speakers are using these linguistic resources for their stylistic practices, or if there does not seem to be a significant relationship between the measurements and the micro level interactions.

Sexuality and Gender Presentation

The investigation of the effect of sexuality on primarily center of gravity, as well as F1/F2, was the primary goal of this research. A lot of previous research, such as Campbell-Kibler (2007) and Podesva and Van Hofwegen (2016), have shown that /s/ indexes a range of meanings,

and in particular it has been used by gay men to index their status as gay men. The means analysis shows that the speakers that consistently have the highest values for center of gravity, such as speakers 1, 4 and 6, also have the highest values for the formant frequencies. These findings are also supported by the findings from the plots, as speakers 1, 4, and 6 also had the highest medians. However, the models did not show a significant relationship between center of gravity and either of the variables. This could be due to the lack of data, which could have had an effect on the models. As the means analysis and the plots support this while the models do not, it seems like there is not a strong association between center of gravity and either of the variables.

In contrast with /s/ center of gravity, there seemed to be a stronger association between formant frequencies and the means, which is supported by the trends seen in the plots and models. For F1 /ɑ/, /æ/, and /ɛ/, there were significant relationships between the vowels and the gender presentation variable. In addition, for F1 /o/, there was a significant relationship with the sexuality variable. These exact significance relationships do not continue for F2. However, over both the F1 and F2 models, there were four vowels that had a significant relationship for the gender presentation intercept, and the same vowel, /o/, had a significant relationship for both F1 and F2.

This tracks with the medians from the plots, as ten out of the eleven plots had the *fem & masc* median higher than the *feminine* median and not overlapping. Since four of the vowels have significant relationships with the gender presentation variable, it's not a surprise that the medians were consistently not overlapping. In addition, the fact that only the /o/ vowel had a significant relationship with the sexuality variable makes sense, as the sexuality plots had six plots with the bisexual plot having a higher median than the lesbian plot, and five plots with the medians overlapping. In addition, the impact of the relationship between /o/ and the sexuality

intercept can be seen in plots 6E and 7E. These plots show the significant difference between the medians of the two plots, which is larger than every other vowel.

These trends that are seen in the models and plots are confirmed by the analysis of the speaker's means. In terms of the average ranking (Table 11), the top five speakers by average ranking consists of three bisexuals and two lesbians. However, when looking at the average rankings, the average ranking between speaker 6 and speaker 3, who are the two lesbian speakers, is 3.67 rankings, which is the largest gap in rankings between speakers of the same sexuality. There also seems to be a large gap between the bisexual speakers in regards to their rankings. Speakers 1 and 4 have the highest center of gravity mean (6283 Hz) and the third highest mean (5776 Hz), respectively. However, speaker 2 has the 6th lowest mean center of gravity (5239 Hz), which is 1000 Hz lower than speaker 1 and 500 Hz lower than speaker 3. This difference in rankings between bisexual speakers still exists in the F1 rankings, even though it is not as pronounced, with speaker 1 having the 2nd highest F1 mean, speaker 4 has the highest F1 mean, and speaker 2 has the 5th highest mean. In the F2 rankings, speakers 1 and 2 have rankings that are much more similar, while speaker 2 still has the higher ranking.

While all of these speakers are bisexual, speakers 1 and 2 share the same gender presentation, while speaker 4 does not. This may be a potential explanation for the differences in the average ranking, but it does not explain the differences among the other bisexual speakers. As speakers 1 and 2, despite sharing sexualities and gender presentations, had significantly different center of gravity means, as well as different F1 rankings and almost the exact same F2 means rankings. This implies that there may be another explanation beyond the effect of these categorical variables on these measurements.

While there seems to be differences among the bisexual speakers, there also seems to be differences among the lesbian speakers, as there is only one lesbian speaker in the top 3, with the rest of the lesbians being in the bottom 4/5 rankings. This one lesbian is speaker 6, who had the second highest average ranking. This could potentially be a reason why the mixed-effects models had insignificant relationships sexuality intercepts so often, as this speaker was consistently above every other lesbian speaker, while the rest of the speakers sat in the bottom 4 rankings almost every single time. There were speakers who individually had top 4 rankings, such as Starra having the fourth highest center of gravity mean, speaker 3 having the fourth highest F1 mean, and Nyx having the third highest F2 mean. However, as a whole, the four lesbian speakers with *feminine* presentations all were in the bottom half of the average rankings.

In a group of eight speakers, with five lesbians and three bisexual speakers, there was inevitably going to be one lesbian speaker in the top 4 average rankings. However, there seems to be a big difference between speaker 6 and the rest of the lesbian speakers that may explain why this speaker consistently placed in the top half of the rankings while every other lesbian speaker placed in the bottom half. This is because speaker 6 has a *fem & masc* gender presentation, while every other lesbian speaker has a *feminine* presentation. Instead of it being completely random that speaker 6 is the one lesbian speaker with consistently high values for these measurements, it seems like the gender presentation of the speaker may have influenced this placement.

This seems to be even more significant than the difference between the bisexual speakers who had the *feminine* presentations and speaker 4, who had the *fem & masc* presentation. In the average ranking, speaker 4 had an average ranking of 2, while the next closest bisexual speaker, speaker 1, had an average ranking of 2.6. In contrast, speaker 6 had an average ranking of 1.6,

while the next closest lesbian speaker, speaker 3, had an average ranking of 5.33. While the gender presentation may have had an impact on the rankings of the measurements for speaker 4, it is hard to definitively say since speaker 1 also has a high ranking. However, it seems like the gender presentation did have a significant impact on the rankings of the measurements for speaker 6.

Interaction between sexuality and gender presentation

When discussing the interactions between sexuality and gender presentation in the means rankings, speakers 4 and 6 are the main examples. However, since only two speakers identified as *fem & masc* and the rest identified as *feminine*, it is hard to discuss some of the other trends seen in the rankings. However, it seems to be undeniable that speaker 6 has the second highest ranking across all measurements due to their *fem & masc* presentation, despite the fact that it seems like being a lesbian speaker means they should be below speakers 1 and 2 in the rankings.

For some of the other trends seen in the rankings, there does not seem to be a large difference between the rankings of the measurements for speakers 1 and 4. They are the two bisexual speakers who are the most comparable, but there is not a large difference in their measurements rankings. Since the results of the center of gravity models are in conflict with the medians and means, that means that only the F1 and F2 vowels are examined.

For the F1 and F2 vowels, there has already been discussion about which vowels are associated with which variables. For the additive F1 models, /ɑ/, /æ/, and /ɛ/ had significant relationships with gender presentation and for F2, ɑ, ɛ, and i had significant relationships. For both F1 and F2, only /o/ had significant relationships for sexuality. While these models give insight as to which vowels are associated with which variables, these are only the additive

models. Since it is important to examine how these two variables are interacting, the interaction models can now be used.

Looking at the center of gravity models, there are no intercepts with significant relationships. When looking at the F1 models, they are very similar to the additive models. All three of /ɑ/, /æ/, and /ɛ/ have significant relationships for the gender presentation intercept. Similar to the additive models, /o/ has a significant relationship for the sexuality intercept. However, these interaction models show some new significant relationships, which are that /æ/ has a significant relationship for the sexuality intercept, and that /o/ has a significant relationship for the bisexual * *fem & masc* intercept, which is the intercept that models the interaction between the gender presentation intercept and the sexuality intercept.

For the F2 models, they are a bit different than the additive models, in that only /ɑ/ and /ɛ/ had significant relationship for the gender presentation intercept, and not /i/. The same trend across every model for the sexuality intercept continues here, as /o/ is the only sexuality intercept. However, there is one big difference in these F2 models, which is that /i/ has a significant relationship for the interaction between the gender presentation and the sexuality intercept.

To summarize these findings, /ɑ/, /æ/, and /ɛ/ have significant relationships to gender presentation for F1 and /ɑ/ and /ɛ/ for F2. /o/ has a significant relationship to sexuality for both F1 and F2, while /æ/ only has a significant relationship for F1. In addition, /o/ has a significant relationship to the interaction between the gender presentation intercept and the sexuality intercept for F1, while /i/ has a significant relationship for F2. The vowels and which intercepts they have significant relationships with are summarized in Table 13. As previously discussed, the coefficients and relationship for these intercepts are the same across both the bisexual and

lesbian intercepts and the *feminine* and *fem & masc* intercepts. This implies that all of the speakers may be using these vowels as linguistic resources, to index both sexualities and both gender presentations.

Vowel Analysis

Now that we have established that specific vowels have significant relationships with certain variables, examining the speaker's means for these vowels is critical. To start off with the F1 table (Table 9), looking at speakers 4 and 6 for /ɑ/, /æ/, and /ɛ/, the values and means rankings appear to be very similar. The largest difference for these speakers is that speaker 4 has the highest mean for /o/ and speaker 6 has the lowest mean for /o/. Speaker 6 having the 7th lowest mean for /o/ is very inconsistent with their rankings for the other vowels, which have the highest rank for three of the vowels and third highest rank for the remaining vowel. To go from the highest mean for three of the four vowels to the 7th lowest mean is strange. From Table 13, /o/ is one of the vowels that has a significant relationship with sexuality, and the only variable difference between speakers 4 and 6 is their sexuality. This does not necessarily mean that speaker 6 is using /o/ to index their sexuality, as other lesbian speakers have a high ranking for /o/, such as speaker 3 who has the third highest /o/ mean.

For speakers 1 and 4, there is also a weird trend of every vowel but /i/ having high values. Both speakers randomly have the 7th and 6th lowest means for /i/, despite having high values for every other vowel. It is strange that /i/ is the only one with the low value, but looking at Table 13, /i/ is the only vowel that does not have a significant relationship with any of the intercepts. This is not to say that every trend in this F1 table is associated with these variables, as the rankings numbers are not as significantly different as speaker 6. However, it is noteworthy that

there seems to be a clustering of high or low values based on these vowels and their associations, especially for speakers 1, 4 and 6.

For the F2 table (Table 10), there seems to be one significant trend, which is that speaker 3 had means in the bottom 3 rankings for every vowel, except they have the third highest mean for /o/. In Table 13, /o/ is associated with sexuality, which makes it seem like this gap between the /o/ ranking and the rest of the rankings seem significant. Beyond that, the only other notable trend is that speaker 6 again has the 5th lowest ranking for /o/, when the other vowels are much higher. This is not as large of a gap, since speaker 6 has the highest, the 4th highest, the 2nd highest, and the highest rankings for the previous vowels, so the fact that one of the other previous vowels is also not as high seems to indicate this relationship is not as significant as in Table 9.

Moving beyond the measurements

I do not think it is useful to extrapolate beyond these single instances to make statements such as a more fronted F1 and F2 for these specific vowels indexing the *fem & masc* presentation, while a more backed F1 and F2 indexes *feminine*, or a similar relationship for sexuality. A more appropriate explanation is that these speakers have these measurements available as resources and a spectrum of available options, from lower values of F1/F2 to higher values. Alongside this spectrum of measurements is a set of meanings attached, with these meanings coming from a variety of places. As seen in previous research, a high center of gravity for /s/ broadly carries the meaning of a more “gay” sounding voice in many communities of practice. However, a high center of gravity /s/ does not mean “gay” sounding in every community of practice. In the context of North American English, it could be argued that the meaning of a high center of gravity /s/ carries the meaning of “gay”, but is then reanalyzed based

on more local categories, as in Campbell-Kibler (2007) or Stuart-Smith (2007), within these communities of practice and given new meanings.

However, due to the fact that this analysis is looking at all of these macro level interactions, it is difficult to say exactly how these speakers are manipulating these linguistic variables. The best we can do is make assumptions based on trends. Another way of examining how these speakers use these linguistic resources is using a transcription analysis to examine the micro level interactions and how these speakers use this spectrum of high and low values to reconstruct meanings.

Other factors

Beyond gender presentation and sexuality as the most prominent factors, this section will briefly examine some of the other factors that could have an effect on the measurements. There are three factors that were measured and collected from the survey sent out to the participants. These factors include age, ethnicity, and place of origin. However, three speakers did not respond to these surveys, which were speakers 1, 5, and 6. This information can be seen in Table 1A. Some information could still be filled out from information collected during the interviews, but any information that could not be collected is labeled as NA in the table.

The argument this paper is making is not that speakers with lower values are choosing to speak like cis men or are failing to move away from sounding like a cis man, nor that the speakers with higher values are successful in moving away from sounding like cis men. The point of discussing these benchmarks and where these speakers fall on this spectrum is that analyzing how these trans speakers compare to the values for the average cis speaker reveals and what it reveals about the choices these trans speakers are making. In comparison to cis speakers, who, I would argue, typically do not think about what makes their voice sound feminine or

masculine, it is a common experience for trans speakers to actively reflect on their voice and whether they want to change it or not when socially and medically transitioning.

This is especially true for trans women and transfeminine people, as their voices do not change at all as a result of estrogen therapy. While trans men and transmasculine people experience their voices deepening as a result of testosterone therapy, this is not the case for trans women and transfeminine people, as estrogen does not have any significant effects on the voice once someone who is assigned male at birth goes through puberty. This means that, for trans women and transfeminine people, it is a larger part of their daily lives to think about their voices and make choices about how they want to sound to others as part of their gender presentation.

That is why I think that the results from these speakers reveal some of the choices they make everyday regarding their voice. Every speaker brought up their experiences with voice training during their interviews, with some talking about how long they had been consistently voice training to reach a certain goal and some talking about how it just did not fit them. These discussions show that every speaker is aware of the way they speak and knows how to manipulate their voice to sound a certain way. While some may be more proficient at this than others, all of the speakers are choosing to sound the way they currently do. A speaker with higher or lower values for these measurements is not accidentally speaking this way, it is a conscious choice they have made and these choices intersect with their sexuality and gender presentation.

There may be factors beyond sexuality and gender presentation that have a greater influence on their voices, such as whether the speaker is socially/medically transitioning, how long they have been transitioning for, or their attitudes towards voice training, but these factors will be discussed later. The main argument of this paper is that these factors are not isolated

effects and can be separated from the speaker's sexuality or gender presentation. These factors all intersect and ultimately inform each other. A speaker's attitude towards voice training and whether they're socially/medically transitioning or not informs how they want to present their gender which then affects the choices they make regarding how they want their voice to sound. All of these factors have to be engaged with when discussing all of the results brought up earlier, such as the relationships between the variables and means, and is part of why this analysis cannot be completed by just looking at sexuality and gender presentation.

Age

As Zimman (2017) states, age is an important factor for trans people, as younger trans people often have an easier time dealing with their changing gender identity while older trans people may have a harder time. From Table 1, there is not that large of a difference between the ages of all the speakers, with the youngest speaker being 20 and the oldest speaker being 29. This can lead to older trans people having difficulty using these linguistic resources to align their speech productions with more normative values for their cisgender counterparts. However, this does not seem to be a major influence on the results. The bisexual speakers were older on average, while the lesbian speakers were younger, typically around 20. However, there is no indication that this is anything other than random chance. In addition, of the lesbian speakers whose age is known (speaker 3, Nyx, and Starra), they were recruited for this project from posters placed around York University's campus. Two of the bisexual speakers, speakers 1 and 2, were recruited from off campus, which seems to be the explanation for why they are older, since they aren't from the university community. Looking at the average ranking in Table 1H, the speaker with the number one average ranking is speaker 4, who is 29 years old, and speaker 6, who did not give an age but was assumed to be around 20 due to conversations from the

interview. In addition, the speaker with the lowest average ranking, Starra, is 21, so age does not seem to be a significant factor in the measurements.

One of the only factors that age may have an effect on are factors like length of time spent voice training or how long a speaker has been transitioning. Although this has been affected by outside factors like the availability of supplies needed for transitioning, which has changed significantly over time, most trans people do not begin medically transitioning before the age of 18. This is due to the fact that it requires a certain amount of resources to begin transitioning, like the money to afford the estrogen or testosterone therapy, or a safe place to begin transitioning if one's parents are hostile to the idea, which trans youth often do not have. In addition to having the resources to transition, most trans people do not even realize they are trans until they have the opportunity to explore their gender identity, which is often not available until they are adults. These factors mean that the speakers who are young, like speaker 3, speaker 6, Nyx, and Starra, often have not been voice training for that long and have either not been socially/medically transitioning for long or started transitioning at all.

Ethnicity

The second factor measured was ethnicity. Three of the speakers identified as white, one speaker identified as white/mixed, two identified as Vietnamese-Canadian, and one identified as Italian-Canadian. Overall, this means there were three trans people of color. However, this does not seem to have a massive effect on the measurements. Looking at the average rankings, the top 3 speakers include speaker 4, who is Vietnamese-Canadian, speaker 1, who is white, and speaker 2, who is mixed. The speaker with the lowest average ranking is Starra, who is Vietnamese-Canadian. With the top speaker and the bottom speaker both being

Vietnamese-Canadian, this suggests that there is no effect of being a trans person of color on these measurements.

Place of Origin

Similar to the effect of ethnicity on these measurements, there seems to be no significant effect of place of origin/dialect on these measurements. This stems from the fact that almost every speaker was from Toronto, except for speaker 2, who was from Boston. The lack of variation in dialect means that it is hard to associate any of the variation with dialect specifically.

Voice Training

One of the most important factors that was not measured in these surveys but was talked about extensively in the interviews was voice training. During the interview, questions were specifically asked about whether the speaker had done voice training, but many participants also brought up this topic on their own. There was considerable discussion on this topic, with people holding a wide range of opinions on what voice training means to them, why they chose to do voice training as well as why they chose to stop, and their opinions on what the meaning of voice training is within the community. Since this seemed to be such an important topic due to its connections to gender presentation, sexuality, and identity construction more broadly, I decided to include these attitudes the speakers held.

Transitioning Status

Similar to voice training, this is another factor that was not brought up during the surveys but was extremely important to every speaker. For the speakers, there was a wide range of stages they were in their transitions. Some speakers had been transitioning and voice training for several years and talked about how their gender presentation or sexuality changed over time.

Other speakers were only socially transitioning for certain groups and were not out in their everyday lives, which led to them not wanting to present as feminine despite feeling like that aligned with their gender presentation. This is significantly linked to their gender presentation and their voices, as if they did not feel safe presenting feminine in their everyday lives, they would definitely not change their voice.

However, this does not mean that transitioning status directly corresponds to whether someone changes their voice. For example, speaker 5 had talked about how they had been transitioning for some time and had done some voice training before, but ultimately decided that they did not want to continue that and that they could be a woman without doing voice training. Speaker 6 also talked about having only transitioned for less than a year but doing a significant amount of voice training during that time before ultimately not wanting to do it because they did not want to sound “hyper feminine”. While these factors could play a role, a deeper investigation into one of these factors specifically would be needed to determine how significant its relationship is to these measurements.

Indexical Order

Now that the analysis of the macro level interactions has been completed, the micro level analysis can begin. To do this, this section will use the transcription analysis in Calder (2019). In this paper, Calder shows the link between linguistic forms, in this case /s/ center of gravity, and specific meanings that his participants construct using fronted /s/ as a stylistic resource. He argues that individuals produce /s/ tokens that are more fronted when constructing their identity, particularly by aligning themselves with femininity or constructing an opposition to femininity (Calder, 2019). In his examples, /s/ tokens that are higher than the mean center of gravity for the

speaker are claimed as indexing such stances, which then are part of the ongoing process of identity construction that the speaker is engaging in.

For this analysis, the goal is to examine whether these speakers are engaged in the same process of identity construction by employing the previously examined vowels. While both center of gravity and F1/F2 were measured, the previous sections show that only the F1 and F2 LMMs had significant relationships with the gender presentation and sexuality intercepts. That is why only the vowels will be examined during this part. This section will be answering the hypothesis, as the hypothesis claims that speakers will be using these vowels to construct their identity, by manipulating the location of the vowel.

In the previous sections, the speakers were shown to be on a spectrum of values for their vowels. Some speakers had vowels with higher formants, while others had much lower formants. By manipulating these formants, the speakers can control the height and frontness of the vowels. If a speaker produces /o/ with an F1 of 700 and another speaker produces it with an F1 of 600, the first speaker will have a much lower production of /o/ than the other speaker. Similarly, if the first speaker produces /o/ with a higher F2, that /o/ will be much more fronted. By having higher formants for the vowels, speakers with higher formants will be producing much more fronted and lower vowels, while speakers with lower formants will be producing more backed and higher vowels.

The argument here is that the meanings that these vowels are associated with first originated from this sound symbology, but then these meanings are reanalyzed by the individual speakers. This process of reconstructing the meanings takes place as the speakers use identity practices to define their gender presentations and sexualities. This idea of identity practices relies on the framework established in Bucholtz (1999). In this study, Bucholotz presents the

framework of identity practices, which are positive and negative practices that help define what someone's identity is. They can mark belonging to a group by defining which members of a group belong, while also creating distance from individuals who do not belong to the group. Positive identity practices are practices that people use to construct their chosen identity and define what practitioners are. Negative identity practices define what their users are not and are used to create distance from a rejected identity. These positive and negative identity practices can be seen as indexing a positive or negative affective meaning, as previously discussed in the sound symbolism section.

By associating these identity practices with these vowels as linguistic variables, the speakers attach meanings to these variables. This process of stylistic bricolage is used by the speakers, where the meanings attached to these fronted/lowered and backed/raised vowels that originated from sound symbolism (level *n*) and cultural meanings attached to these vowels (level *n+1* and beyond) are removed and new meanings are attached, with the positive/negative identity practices being used as a tool in this process. While these variables are used in defining the gender presentation and sexuality of the speaker, the meanings are ultimately unique to the speaker themselves. Even though a fronted and lowered vowel may have originally indexed meanings like "femininity" or "friendliness" in earlier levels of the indexical order, it does not hold that original meaning after this process of reconstruction.

Transcription Analysis Process

To start off, the vowel /o/ is first examined since it was the only vowel that had a significant relationship for every intercept (Table 8). As Calder (2019) did his analysis by looking at values that were above the mean, this analysis will be conducted by looking at instances of /o/ where both the F1 and F2 were above the mean for that speaker. However, since

it is likely that not every speaker wanted to make their vowels more fronted and low, there will be an examination of the instances where both F1 and F2 were below the mean. This is because previous sections showed that some speakers may be manipulating the phonetic variables by having lower values rather than higher. For example, speaker 6 having the 7th lowest mean /o/, despite having much higher means for the other vowels. Whether the F1 and F2 above the mean or below the mean will be examined is based on the speaker's ranking for that vowel. If a speaker has a higher ranking for the vowel, it is likely that it would make sense to examine the higher values, and vice versa for low ranking. This way, both instances of speakers raising their vowels significantly and lowering them significantly will be captured.

Once tokens have been identified, the instance in their transcription will be found and the location of the token will be identified. If these locations seem to be relevant to the speaker's gender presentation or sexuality, then it seems like there is a link between the vowel and the manipulation of the formant. If it seems irrelevant, then it is likely that there was no active manipulation and it may just be due to random chance.

Transcription Analysis

This analysis will focus on three speakers that have previously been identified as having trends across the median and means of the measurements, as well as the measurements rankings in Table 9 through 11. These speakers are speakers 1, 4 and 6, and the analysis will begin with them.

In the mean rankings analysis, one of the most notable trends was that speaker 6 had significantly higher measurements than the rest of the lesbian speakers. This seemed to be motivated by their gender presentation, which is *fem & masc*, while the rest of the lesbian speakers had a *feminine* presentation. In the following passage, speaker 6 talks about how their

gender presentation is not necessarily aligned with what they call “hyper feminine” trans women and not following what trans women think of as passing:

(1) **Speaker 6** F1 /o/ mean: 613.2 & F2 /o/ mean: 1274.4

I feel like a lot of trans women have a really toxic mindset behind transitioning. It's like, oh, like you're not a woman if you don't, um, sound this way…And it's like, /nah/ [896.95; 1651.68], like I've always really known that like, oh, transitioning is just like about your gender expression. Like the goal for me was, it was to pass, but it wasn't to be like hyper feminine. Like the goal for people isn't, like it doesn't /always/ [1000.78; 1922.63] have to be to pass. /And/ [771.18; 1972.32] that's the thing I knew, but that's not a thing I internalized really.

In this section, speaker 6’s production of /o/ has significantly higher formants than their mean productions. This is notable because this speaker had the 7th lowest and 4th lowest F1 and F2 means, so it is rare that they had vowel formants this fronted and this low in comparison to their average productions. Here, the speaker discusses how they view transitioning as part of their gender presentation/expression, but that the popular ideal of what a passing voice sounds like is not something they align with. They want to pass, but they don’t align themselves with the standard hyper femininity that is expected of trans women. Instead, they align with more of an alternative femininity that does not try to meet these standards.

In this second section, speaker 6 expands on their views of voice training and what having a passing voice sounds like:

(2) **Speaker 6** F1 /o/ mean: 613.2 & F2 /o/ mean: 1274.4

If I were to be like a straight person, I would probably do voice training like more. I would probably like try harder to pass. Whereas now I'm kind of like, I don't really need

to pass that much to be seen as a woman. I just need to— it's very simple to pass… Like, /there's/ [759.12; 2424.23] more masculine /women/ [778.82; 1731.28] than I am who are cisgender. And I am also lesbian and I also did voice training to— I think I did it more for myself and for like my projection into the world than for the male gaze.

Speaker 6 continues to set themselves apart from standard femininity and the type of femininity that men expect. In contrast with straight trans women, who speaker 6 talks about doing voice training for men's expectations, speaker 6 wanted to do voice training for their own self-presentation. They do not feel the need to try harder to pass and attain this more "standard" heterosexual femininity, and the type of voice that is associated with that gender presentation. Instead, they create a contrast between their own reasons for doing voice training as a lesbian and the reasons that straight trans women do voice training, which speaker 6 sees as for male expectations.

In both sections, speaker 6 constructs their gender presentation and their alternative femininity to standard femininity that many trans women seek to use as their gender presentation. The gender presentation of speaker 6 is certainly constructed in opposition to this standard femininity, which could be argued that speaker 6 is engaging in negative identity practices. In these examples, speaker 6 defines their *fem & masc* gender presentation by comparing it to what it is not, mainly the practices that are embodied by a standard feminine presentation. Their gender presentation is also linked to their sexuality in this process, as they use negative identity practices to define what their identity, as a lesbian with an alternative feminine presentation, is in opposition to, which is a standard femininity practiced by more straight/heterosexual leaning trans women. In this process of constructing their gender

presentation, they use these high vowel formants for /o/ and link it to specific meanings, mainly that of an alternative queer femininity that is juxtaposed with a standard heterosexual femininity.

The next speaker to be analyzed is speaker 1, who also had very high formants for both F1 and F2, but is much different than speaker 6. Speaker 1 is bisexual with a *feminine* presentation, so it will be interesting to see if the way they use /o/ is similar to speaker 6. In this next section, speaker 1 talks about their journey with their gender presentation and how it has changed over time:

1. **Speaker 1** F1 mean: 658.2 F2 mean: 1327.9

   Definitely like switched up a little bit and /gotten/ [716.94; 1463.64] more comfortable with my femininity as times have /gone/ [710.36; 1400.71] by. There was a time where I was like sort of trying to walk a little bit of a line because I was presenting masculine previously and then so getting into a little bit more sort of an /androgynous/ [827.68; 1385.89] area for a little while, but as I've /gotten/ [690.13; 1533.70] more comfortable with myself, I've been comfortable just being pretty stereotypically feminine, I guess

As speaker 1 narrates their journey with their gender presentation changing over time, their /o/ F1 and F2 formants are consistently higher than their mean. This seems to provide a contrast between what their previous presentation was, which was sometimes leaning masculine-androgynous, with what they consider to represent their current feminine presentation. Speaker 1 seems to identify strongly with a stereotypical femininity, which makes sense as to why they have fronted and lowered their vowels. They do rely on some negative identity practices by contrasting their previous presentations with their current gender presentation. It also seems like they are using positive identity practices by mentioning multiple times that

they’re much more comfortable with their current stereotypical feminine presentation. Examining the words they have these higher formants on, “gotten” and “gone” are part of their discussion of how they’re now more comfortable with femininity. It also seems like the high formant on “androgynous” is used to contrast how their presentation, including their voice, is now much more feminine and is not androgynous anymore.

In the next section, speaker 1 manipulates vowel formants to construct their gender presentation and sexuality are in this case by using both higher and lower formants:

(2) **Speaker 1** F1 mean: 658.2 F2 mean: 1327.9

{**In your ideal scenario, how would you approach a cis man in a bar?**}

So they would come up and they'd start talking to me and I would be a little bit higher, a little bit breathier, and I would sort of, you know, just, hi, how are you? What's going /on?/ [761.93; 1404.11] What are you drinking? Oh yeah, you here with your friends? Yeah. How long have you guys come here /often?/ [906.83; 1347.46] A little bit like that. I do sometimes worry that I sound like it's associating towards closer, like a valley girl almost kind of a vibe. And I do worry that sometimes makes me seem less intelligent than I am, and especially around men. I have my guard up less around women, but around men sometimes I do feel like I have to /also/ [647.27; 936.30] make sure that I sound smart.

Speaker 1 uses the higher F1 and F2 formants during this section when they imagine talking to a man they’re attracted to at a bar. Similar to the first section, the higher formants are used to index a standard femininity, which speaker 1 wants to perform and seems to know how to

do so, as they bring up making their voice higher and breathier. During their interview, speaker 1 mentioned how when they're with men, they like to go with a "femme fatale kind of look". Their practices in this section seem to be positive identity actions that align themselves with this icon that they want to perform.

However, something interesting happens at the end of the section, when they discuss their worries about seeming less intelligent by performing this type of standard femininity that the valley girl accent is associated with. It seems like they use a negative identity practice to distance themselves from this valley girl type of femininity. In doing so, they construct what their own femininity they want to perform is supposed to sound like. When they are in the process of opposing this valley girl femininity, their /o/ formants drop sharply. While their F1 drops about 100-200 Hz, their F2 drops about 400 points. This places the /o/ significantly further back and more raised, compared to the more fronted position it occupied before. By manipulating the /o/ formants, they construct their gender presentation of the "femme fatale", while also distancing themselves from more negative forms of femininity, like the valley girl.

The final speaker of these three that needs to be looked at is speaker 4. Speaker 4 is also bisexual, like speaker 1, but they have a *fem & masc* presentation, like speaker 6. In this section, speaker 4 talks about how they view their gender presentation as being influenced by their sexuality:

(1) **Speaker 4** F1 mean: 713.1 & F2 mean: 1344.7

{And then do you think your gender presentation is influenced by your sexuality?}

That's not [845.79; 1477.20] a question I really think about, but like if I do think about it, then it definitely does. I would have to think about more and how it does. When I was like kind of— it kind of fluctuates my sexuality, but when I'm feeling a lot more like leaning towards femme people and like being attracted to more femmes, I'll dress more like androgynous [838.77; 1348.43] or futch and everything. And then when I'm like unfortunately being attracted to more like masc folks, it's just like, well, I guess I feel like dressing more like a femme now, you know.

Speaker 4 explicitly discusses how their gender presentation is influenced by their sexuality, with their presentation changing based on who they are interested in. Their gender presentation is defined as being more androgynous or "futch", which is a blend of femme and butch, when they are talking to queer women. Futch is a locally constructed category that is known in butchfemme lesbian communities. These lesbian communities are a type of imagined community, where the participants gain membership by identifying as lesbians, and it is made up of many communities of practice. However, commonly shared knowledge about certain categories, like butch or femme, is shared across all of these communities of practice and in the imagined butchfemme lesbian community, as well as in the broader queer imagined community. This knowledge is spread through in-person social networks and social media. In this section, speaker 4 is using /o/ as a resource to index their gender presentation, which is based on the locally constructed category of futch.

When speaker 4 talks about their attraction to men, however, they use more negative identity practices to define their gender presentation. Their femme gender presentation is created

in contrast to their androgynous presentation, just like their attraction to queer women is in contrast to them dating men, as their attraction to men is defined as “unfortunate”. Speaker 4’s gender presentation is defined by their sexuality, as their sexual attraction to queer women is accompanied by a more queer femininity, while their attraction to men results in a more standard heterosexual-leaning femininity. In this example, speaker 4 uses the higher vowel formants for both positive identity practices and negative identity practices. They define their queer femininity through their attraction to queer women and a more masculine presentation, but their more standard heterosexual femininity is defined in contrast to their queer femininity.

In the next section, speaker 4 continues to discuss their gender presentation and uses more positive identity practices in defining their queer femininity.

(2) **Speaker 4** F1 mean: 713.1 & F2 mean: 1344.7

> *In terms of like, uh, dressing to present for like more femme folks and everything, um, I feel like a lot more free, a lot more leeway. I wanna [658.85; 1091.34] present, um, more, I guess, like I want to embody both masculinity and femininity to show kind of like, I guess this is kind of who I am really. Um, I definitely have to think about that a part a bit more, you know. Um, but I definitely feel a lot more comfortable dressing more androgynous [762.27; 1487.66] futch around femme-presenting people.*

Speaker 4 uses positive identity practices more to define their androgynous futch *fem & masc* gender presentation, but also defines their gender presentation in opposition to another style. Similar to speaker 1 when talking about their gender presentation, speaker 4 talks about how they are a lot more comfortable with their queer feminine futch presentation. For their

gender presentation, they state that “both masculinity and femininity” are “who i am really”, and define what their gender presentation is. However, the first production of /o/ occurs when they are contrasting dressing for femmes and queer women in comparison to their standard femininity that they discussed in section 1.

They also pronounce /o/ with vowel formants that are above and below their means. When discussing how they want their presentation to be both masculine and feminine, they use both a much more fronted and lowered /o/ as well as a much more backed /o/ and raised. Rather than being random or coincidental, this seems like a much more deliberate choice in this speech context. Speaker 4 uses the much more backed production to define what their gender presentation is in opposition to, which is a standard femininity that is not as free and does not give them as much leeway. They then use the much more fronted /o/ when defining what their gender presentation is, which is androgynous, more futch, and more comfortable for them.

When looking at these three speakers overall, it seems like they used the more fronted/lowered and more backed/raised productions of /o/ in their positive and negative identity practices. In contrast with speaker 6, who used this fronted /o/ to construct their gender presentation as an alternative feminine style, speaker 1 uses the same variable to construct their presentation as standard femininity. This division also reveals that not all of the speakers manipulated the formants in the same way. While speakers 1 and 4 clearly used more fronted and more backed productions to divide between positive and negative identity practices, speaker 6 used more fronted productions for all of their identity practices. This reveals that these linguistic resources are not all used in the same way.

Transcription Analysis Conclusions

One of the conclusions of the analyses of these transcripts is that /o/ is used to index a wide variety of meanings for all of these speakers. Speaker 6 uses it to define their gender presentation in opposition to heterosexual femininity, with their identity as a lesbian being a key factor in the decisions they make regarding their gender presentation. Speaker 1 uses it to construct their ideal type of femininity that includes heterosexual relations, while also distancing themselves from certain negative types of femininity. Speaker 4 also employs it to mark their preference for a more androgynous futch *fem & masc* presentation around queer women and femmes, which is in opposition to their more begrudging standard femininity and attraction to men. The interaction between the speaker's sexuality and gender presentation is clear, as all of the speakers clearly kept in mind who they were attracted to and what type of femininity (and masculinity) they wanted to utilize in their presentations.

Now that the /o/ analysis is completed, all of the analyses are finished. There remains potential for the investigation into the micro level interactions that could occur in regards to each of the other four vowels in future research. However, it is not relevant to this paper, since it has already been shown that at least one of these vowels is being manipulated by these speakers as part of their stylistic practices. For the sake of brevity, only /o/ will be investigated. In addition, while the other speakers had some interesting transcriptions that could be discussed, the majority followed the same patterns laid out in the previous sections in regards to how they used /o/ as a linguistic variable in their stylistic practices. It does not seem fruitful to investigate these instances when the main point has already been made clear through the examples of speakers 1, 4 and 6.

Discussion

Through the findings of the LMMs, the medians and means, and the transcription analysis, it can be concluded that one linguistic variable, /o/, is used by these speakers as part of their stylistic practices. Through the process of stylistic bricolage, they attached new meanings to this variable, which they use to construct their gender presentations and sexualities. This variable does not hold only one meaning, but the meaning seems like it can be shifted by manipulating the formants of the vowel to create new meanings. Speakers 1, 4 and 6 often used a more fronted and lowered vowel to index their positive identity practices, through which they defined what their gender presentations and sexualities would be. However, speaker 6 also used these fronted vowels to index their negative identity practices, while speakers 1 and 4 used more backed productions of /o/ to define what their gender presentations and sexualities were not.

Imagined communities and meanings

What is the role of the imagined community in all of this discussion about meanings and reconstruction? As has been shown in the transcription analysis, the speakers used these reconstructed meanings to construct their identity and index these locally constructed categories, such as futch, as a part of their stylistic practices. Since these speakers do not share a single community of practice, it is argued that these locally constructed categories originate from the imagined communities that all of these speakers are a part of. While the participants in my study were not all part of the same community of practice in Toronto and not every participant knew each other, they are still positioned within this imagined community, with access to this community centered around being transgender.

Knowledge of locally constructed categories in this imagined trans community is spread through in-person social networks and social media. In addition, knowledge of common beliefs around passing and gender presentation is circulated in this imagined trans community. In section 1 of speaker 6's analysis, they stated "*I feel like a lot of trans women have a really toxic mindset behind transitioning. It's like, oh, like you're not a woman if you don't, um, sound this way…And it's like, nah"*. This sentiment about what it sounds like to pass when you're a trans woman is knowledge that is spread through this imagined community, as well as other beliefs.

As in any community, there is a wide variety of beliefs and values, with intercommunity groups and divisions. The first type of imagined community that Anderson (1983) explored was the nation state, which still have their common cultures, values and beliefs when examined on a macro scale, despite internal divisions. The imagined trans community, just like any other imagined community, also has widespread beliefs, values, and icons that individuals compare themselves to. As speaker 6 discussed in their interview, they were comparing their gender presentation, with which their voice is one part, to the commonly held beliefs about what a standard femininity for a trans woman in this community is supposed to look like and sound like. Speaker 4, who knows about the butchfemme lesbian community, uses one of the locally constructed categories, futch, to define their gender presentation. Speaker 1 seems to be relying on these commonly held beliefs to embody a type of stereotypical trans femininity, that is still in opposition to other types of femininity like the valley girl.

All of these speakers belong to the same imagined trans community and are using these categories to construct their gender presentations and sexualities. By using these categories to position themselves around, either in opposition against or in embrace of, they reconstructed the meanings attached to the fronted/lowered or backed/raised dichotomy for /o/ and constructed

their own meanings. Ultimately, the speakers had the same variable, /o/, and manipulated the fronting and backing to reach different goals. This process of stylistic bricolage that the speakers engaged in allowed the speakers to index these locally constructed categories and place themselves in relation to them. Constructing their voice was just one part of the speaker's stylistic practices, and the positive and negative identity practices were only one tool used in this voice construction.

Future Directions

It cannot be determined in this paper that these speakers manipulate the fronting and backing of /o/. However, there seem to be some significant relationships between these variables and these meanings, which makes it worth it to investigate further to see if this is an actual phenomenon. The results from the means suggest there is a relationship between the measurements and the formant frequencies, and the models, plots and means support this trend. Even if the results can't be extrapolated since only three speakers were discussed, the transcription analysis does seem to show that at least something is happening with these variables in regards to these formants being manipulated.

There are additional factors that were not able to be fully explored in this study due to scope limitations. For the effect of question type on the measurements, the results were very inconclusive and need further investigation. In addition, the effects of voice training and transitioning status would also be fruitful avenues for future investigation. All trans people are different in regards to their transitioning status, and trans women and trans feminine people especially vary with regards to how much vocal training they have done. Examining these factors, both with regards to sexuality and gender presentation as well as independently of those variables, seems like a useful direction for future research.

Conclusion

In summary, it does not appear that there is phonetic variation in the trans-feminine communities in Toronto based solely on sexuality. While there was a loose association between the sexuality of the speaker and their means per measurement, it was not sufficiently significant to argue that it was meaningful. In a larger study with more participants, it seems very likely that this association would not be maintained. However, it does seem that a significant proportion of the variation in F1 and F2 can be accounted for by gender presentation, with a smaller proportion, like the /o/ models, being accounted for by sexuality. There also seemed to be an interaction between the sexuality and gender presentation of the speaker, with speakers 4 and 6 having the highest formant frequencies, most likely from their gender presentations.

By examining the transcription analysis, this variation emerged as speakers were using fronted/lowered and backed/raised vowels as resources in their stylistic practices. The meanings attached to these variables varied based on the sexuality and gender presentation of the speaker, but these meanings were related to an imagined trans community that all of the speakers are members of. These speakers used these variables to position themselves in relation to locally constructed categories that are known in this imagined community.

No variation based on broader social categories, like sexuality, was found. However, it seems that this phonetic variation occurred on the micro level, within more local categories like gender presentation and in the interactions between sexuality and gender presentation. Going forward, it seems best to move on from trying to find “gay voice” by comparing groups of speakers based on larger categories like sexuality. Phonetic variation tied to identities within the queer community, such as gender presentation, seems to be the most fruitful avenue going

forward, and any future investigation must take into consideration the locally constructed categories that speakers belong to.

References

Adank, P., Smits, R., & Van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, *116*(5), 3099–3107.

Anderson, B. (1983). Imagined communities: Reflections on the origin and spread of nationalism. In *The new social theory reader* (pp. 282–288). Routledge.

Bucholtz, M. (1999). "Why be normal?": Language and identity practices in a community of nerd girls. *Language in Society*, *28*(2), 203–223.

Bucholtz, M., & Hall, K. (2004). Theorizing identity in language and sexuality research. *Language in Society*, *33*(04). https://doi.org/10.1017/S0047404504334020

Calder, J. (2019). From *Sissy* to *Sickening:* The Indexical Landscape of /s/ in SoMa, San Francisco. *Journal of Linguistic Anthropology*, *29*(3), 332–358. https://doi.org/10.1111/jola.12218

Campbell-Kibler, K. (2007). ACCENT, (ING), AND THE SOCIAL LOGIC OF LISTENER PERCEPTIONS. *American Speech*, *82*(1), 32–64. https://doi.org/10.1215/00031283-2007-002

DiCanio, C. (2013). *Spectral moments of fricative spectra* [Praat script] https://www.acsu.buffalo.edu/~cdicanio/scripts/Time_averaging_for_fricatives_4.0.praat

Eckert, P. (2003). *The meaning of style*. na.

Eckert, P. (2008). Variation and the indexical field[1]. *Journal of Sociolinguistics*, *12*(4), 453–476. https://doi.org/10.1111/j.1467-9841.2008.00374.x

Eckert, P. (2010). *Affect, sound symbolism, and variation*.

Eckert, P., & Rickford, J. R. (2001). *Style and sociolinguistic variation*. Cambridge University Press.

Fairbanks, G. (1960). Voice and articulation drillbook. *(No Title)*.

Flipsen, P., Shriberg, L., Weismer, G., Karlsson, H., & McSweeny, J. (1999). Acoustic characteristics of/s/in adolescents. *Journal of Speech, Language, and Hearing Research*, *42*(3), 663–677.

Fruehwald, J. (2022, October 20). *How to normalize your vowels using the tidyverse*. *Linguistics Methods Hub*. Zenodo. https://doi.org/10.5281/zenodo.7232341

Fruehwald, J., & Maier, W. (2026). *Forced-Alignment-and-Vowel-Extraction/new-fave* (Version 1.3.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.18613417

Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and phonology. *Speech Prosody 2002*, 47–57. https://doi.org/10.21437/SpeechProsody.2002-7

Hamano, S. (1994). Palatalization in Japanese sound symbolism. *Sound Symbolism*, 148–157.

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

Joseph, B. D. (1994). Modern Greek ts: Beyond sound symbolism[1]. *Sound Symbolism*, 222.

Kanno, Y., & Norton, B. (Eds.). (2012). *Imagined Communities and Educational Possibilities: A Special Issue of the journal of Language, Identity, and Education*. Routledge. https://doi.org/10.4324/9780203063316

Levon, E., & Mendes, R. B. (Eds.). (2016). *Language, sexuality, and power: Studies in intersectional sociolinguistics*. Oxford University Press.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner, and Morgan Sonderegger (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In

Proceedings of the 18th Conference of the International Speech Communication Association.

Moonwomon-Baird, B. (1997). Toward a Study of Lesbian Speech. In A. Livia & K. Hall (Eds.), *Queerly Phrased* (pp. 202–213). Oxford University PressNew York, NY. https://doi.org/10.1093/oso/9780195104707.003.0012

Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, *34*(2), 202–240. https://doi.org/10.1016/j.wocn.2005.05.003

Ochs, E. (1992). Indexing Gender. *Rethinking Context: Language as an Interactive Phenomenon*, *11*(11), 335.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F_0 of voice. *Phonetica*, *41*(1), 1–16.

Ohala, J. J. (1995). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound Symbolism* (1st ed., pp. 325–347). Cambridge University Press. https://doi.org/10.1017/CBO9780511751806.022

Pierrehumbert, J. B., Bent, T., Munson, B., Bradlow, A. R., & Bailey, J. M. (2004). The influence of sexual orientation on vowel production. *The Journal of the Acoustical Society of America*, *116*(4), 1905–1908.

Podesva, R. J., & Hofwegen, J. V. (2015). /S/exuality in Smalltown California. In E. Levon & R. B. Mendes (Eds.), *Language, Sexuality, and Power* (pp. 168–188). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190210366.003.0009

Silverstein, M. (1994). Relative motivation in denotational and indexical sound symbolism of Wasco-Wishram Chinookan. *Sound Symbolism*, 40–60.

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, *23*(3–4), 193–229. https://doi.org/10.1016/S0271-5309(03)00013-2

Stuart-Smith, J. (2007). *Empirical evidence for gendered speech production:/s/in Glaswegian*.

Sulkin, E. (2025). Analyzing acoustic correlates of gender presentation in the lesbian community. *Journal of Language and Sexuality*, *14*(2), 228–261. https://doi.org/10.1075/jls.24023.sul

Valentine, D. (2020). *Imagining transgender: An ethnography of a category*. Duke University Press.

Winter, B. (2020). *Statistics for linguists: An introduction using R*. Routledge.

Zimman, L. (2013). Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men. *Journal of Language and Sexuality*, *2*(1), 1–39. https://doi.org/10.1075/jls.2.1.01zim

Zimman, L. (2017a). Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/. *Language in Society*, *46*(3), 339–370. https://doi.org/10.1017/S0047404517000070

Zimman, L. (2017b). Variability in /s/ among transgender speakers: Evidence for a socially grounded account of gender and sibilants. *Linguistics*, *55*(5). https://doi.org/10.1515/ling-2017-0018

Zimman, L. (2018). Transgender voices: Insights on identity, embodiment, and the gender of the voice. *Language and Linguistics Compass*, *12*(8), e12284. https://doi.org/10.1111/lnc3.12284

Appendices

Tables

Table 1 - Demographics

Speaker	Age	Ethnicity	Origin/Dialect	Gender Identity	Gender Presentation	Gender Category	Sexuality
1	27	White	NA	Woman	“Stereotypical feminine”	Feminine	Bisexual - no preference
2	26	White / Mixed	U.S.	Woman	“Standard Woman”	Feminine	Bisexual - mostly men
3	20	Italian-Canadian	Toronto	Woman	“Femme”	Feminine	Lesbian
4	29	Vietnamese-Canadian	Toronto	Woman	“Futch”	Fem & Masc	Bisexual - no preference
5	NA	White	NA	Woman	“Basic Femme”	Feminine	Lesbian
6	NA	NA	NA	Woman	“Mostly Feminine / Mix"	Fem & Masc	Lesbian
Nyx	21	White	Toronto	Woman	“Twink”	Masc Androgynous	Lesbian
Starra	21	Vietnamese-Canadian	Toronto	Woman	“Transfem”	Feminine	Lesbian

Table 2 - Gender Presentation

Speaker	Gender Presentation
Speaker 1	Feminine
Speaker 2	Feminine
Speaker 3	Feminine
Speaker 4	Fem & Masc
Speaker 5	Feminine

Speaker 6	Fem & Masc
Nyx	Masc Androgynous
Starra	Feminine

Table 3 - Z-scores example

speaker	F1	F2	vowel	F1z-score	F2z-score
1	395.2254296	1910.419227	i	-0.9987856561	0.5457125464
1	353.3467126	2524.24083	i	-1.247318886	2.175850502
1	358.0267227	2258.164733	i	-1.219544921	1.469227042
1	278.8847359	2357.27673	i	-1.689220569	1.732440698
1	312.3333995	1751.956793	i	-1.490716293	0.1248807822
1	300.6345369	2491.002012	i	-1.560144306	2.087577527
1	319.0057083	2337.022177	i	-1.451118842	1.678650288
1	304.002552	2356.35988	i	-1.5401565	1.730005801
4	753.2671326	2044.916134	e	0.6547230888	0.8603002066
4	611.0361028	1815.416822	e	-0.0915522775	0.2317832219
4	470.6648413	1899.335808	e	-0.8280695813	0.4616074753
4	714.8241287	1752.157199	e	0.4530155791	0.0585376154
4	639.9111828	1881.626893	e	0.05995306514	0.4131090561
4	545.9517075	1878.267026	e	-0.4330451361	0.4039075754
4	702.4563444	1894.982394	e	0.3881227581	0.4496850229
4	531.813029	1666.422346	e	-0.5072297039	-0.1762596518
4	511.3349916	1912.842144	e	-0.6146764044	0.4985965256
4	471.1313168	2166.273063	e	-0.82562202	1.19265367
4	751.5466028	1793.650979	e	0.6456956001	0.1721743199

Table 4 - Fixed Effects Coefficients: Individual Fixed Effects

A.

Measurement	Gender Presentation Intercept (Feminine)	Gender Presentation Fem & Masc	Sexuality Intercept (Lesbian)	Sexuality Bisexual

CoG	5471.712:1.77e-10 ***:	357.312:0.371	5360.530:5.9e-10 ***:	464.289:0.15:

B.

Measurement	Following Vowels Intercept (æ)	Following Vowels ɑ	Following Vowels ɛ	Following Vowels i	Following Vowels o
CoG	6607.3:3.22e-06 ***:	-1266.0:0.364	-845.8:0.544	-652.8:0.641	-1262.8:0.363

C.

Measurement	Gender Presentation Intercept (Feminine):	Gender Presentation Fem & Masc	Sexuality Intercept (Lesbian)	Sexuality Bisexual
F1 ɑ	511.872:1.3e-09 ***	44.580:0.00297 **	515.923:1.33e-11 ***	13.787:0.433
F1 æ	701.759:1.3e-12 ***	77.425:0.00994 **	702.046:1.95e-11 ***	48.337:0.108
F1 ɛ	595.293:1.36e-13 ***	55.150:0.00535 **	598.499:4.62e-11 ***	25.937:0.223
F1 i	366.141:3.99e-11 ***	17.698:0.204	373.238:6.38e-11 ***	-6.657:0.602
F1 o	628.396:1.68e-11 ***	38.206:0.138	618.630: 8.88e-14 ***	49.546: 0.0169 *

D.

Measurement	Gender Presentation Intercept (Feminine)	Gender Presentation Fem & Masc	Sexuality Intercept (Lesbian)	Sexuality Bisexual
F2 ɑ	1557.078:9.49e-14 ***	97.326:0.00446 **	1568.578: 3.74e-13 ***	31.062: 0.42
F2 æ	1702.733:1.16e-14 ***	29.584:0.338	1707.959:2.27e-15 ***	5.366:0.849

F2 ɛ	1681.674:8.97e-14 ***	121.085:0.00408 **	1704.425:1.53e-12 ***	19.139:0.686
F2 i	2167.30:7.75e-12 ***	201.52:0.02 *	2203.182:1.95e-10 ***	37.878:0.678
F2 o	1251.060:1.22e-11 ***	60.689:0.219	1222.987:8.12e-13 ***	110.303:0.00174 **

Table 5 - Fixed Effects Coefficients: Additive Fixed Effects

a.

Measurement	Gender Presentation + Sexuality + Following Vowels: (Intercept:Lesbian + Feminine + æ)	Gender Presentation + Sexuality + Following Vowels: Bisexual	Gender Presentation + Sexuality + Following Vowels: Fem & Masc
CoG	6398.678: 7e-06 ***	428.718: 0.195	192.594: 0.598

b.

Measurement	Following Vowels ɑ	Following Vowels ɛ	Following Vowels i	Following Vowels o
CoG	-1272.957: 0.361	-865.940: 0.535	-671.923: 0.632	-1280.596: 0.357

c.

Measurement	Gender Presentation + Sexuality (Intercept:Lesbian + Feminine)	Gender Presentation + Sexuality: Bisexual	Gender Presentation + Sexuality: Fem & Masc
F1 ɑ	508.812: < 2e-16 ***	6.611: 0.226	43.078: 2.95e-13 ***
F1 æ	687.663: 2.05e-14 ***	38.768: 0.04042 *	70.740: 0.00396 **
F1 ɛ	588.841: 3.46e-14 ***	18.832: 0.13357	50.128: 0.00454 **

F1 i	369.576: 4.2e-11 ***	-9.572:0.403	19.380: 0.155
F1 o	612.886:5.85e-13 ***	44.466:0.0113 *	32.566:0.0651 .

D.

Measurement	Gender Presentation + Sexuality (Intercept:Lesbian + Feminine)	Gender Presentation + Sexuality: Bisexual	Gender Presentation + Sexuality: Fem & Masc
F2 ɑ	1550.463: 2.55e-16 ***	16.517: 0.43264	94.744: 0.00422 **
F2 æ	1702.4528: 6.03e-15***	0.7733: 0.977	29.4330: 0.346
F2 ɛ	1680.998: 5.41e-14 ***	1.886: 0.94551	120.697: 0.00432 **
F2 i	2163.640: 1.66e-11 ***	10.684: 0.8680	199.607: 0.0218 *
F2 o	1217.246:3.95e-13 ***	103.196:0.000918 ***	34.831:0.120472

Table 6 - Fixed Effects Coefficients: Interaction Fixed Effects

A.

Measurement	Gender Presentation + Sexuality + Following Vowels: (Intercept:Lesbian * Feminine * æ)	Gender Presentation + Sexuality + Following Vowels: Bisexual	Gender Presentation + Sexuality + Following Vowels: Fem & Masc
CoG	6422.39: 6.69e-06 ***	428.83: 0.329	1032.40: 0.166

B.

Measurement	Following	Following	Following	Following

	Vowels ɑ	Vowels ɛ	Vowels i	Vowels o
CoG	-1398.18: 0.336	-1050.58: 0.456	-616.54: 0.667	-1341.98: 0.331

C.

Measurement	Bisexual * Fem & Masc	Bisexual * ɑ	Bisexual * ɛ	Bisexual * i	Fem & Masc * ɑ	Fem & Masc * ɛ	Fem & Masc * i
CoG	-982.90: 0.289	40.32: 0.944	637.62: 0.208	168.21: 0.794	327.85: 0.788	-503.66: 0.598	-1834.69: 0.128

D.

Measurement	Bisexual * Fem & Masc * ɑ	Bisexual * Fem & Masc * ɛ	Bisexual * Fem & Masc * i
CoG	576.84: 0.720	-139.29: 0.901	1469.99: 0.286

E.

Measurement	Gender Presentation + Sexuality (Intercept:Lesbian * Feminine)	Gender Presentation + Sexuality: Bisexual	Gender Presentation + Sexuality: Fem & Masc	Gender Presentation + Sexuality: Bisexual * Fem & Masc
F1 ɑ	505.190: < 2e-16 ***SINGULAR FIT	13.341: 0.0389 * SINGULAR FIT	57.693: 1.62e-09 *** SINGULAR FIT	-23.503: 0.0515 SINGULAR FIT
F1 æ	688.830: 1.72e-14 ***	35.597: 0.0909	64.579: 0.0296 *	11.897: 0.7492
F1 ɛ	585.078: 1.96e-15 ***	30.179: 0.023529 *	72.166: 0.000494 ***	-42.138: 0.075637
F1 i	367.244: 3.17e-11 ***	-2.990: 0.8040	32.016: 0.0783	-25.324: 0.2990
F1 o	620.172:< 2e-16 ***	25.989:0.01282 *	-6.975:0.62105	73.865:0.00105 **

F.

Measurement	Gender Presentation + Sexuality (Intercept:Lesbian * Feminine)	Gender Presentation + Sexuality: Bisexual	Gender Presentation + Sexuality: Fem & Masc	Gender Presentation + Sexuality: Bisexual * Fem & Masc
F2 ɑ	1546.741: <2e-16 ***	25.673: 0.2926	111.923: 0.0056 **	-33.396: 0.4600
F2 ae	1706.794: 5.03e-15 ***	-10.966: 0.718	7.514: 0.856	43.318: 0.466
F2 ɛ	1687.731: 2.98e-14 ***	-17.166: 0.5595	84.231: 0.0585	71.714: 0.2275
F2 i	2137.686: 3.1e-12 ***	87.997: 0.111136	335.266: 0.000767 ***	-288.451: 0.017035 *
F2 o	1213.359:7.07e-13 ***	113.528:0.00141 **	61.011:0.04821 *	-43.218:0.25827

Table 7 - Iteration Models / Likelihood Ratio Tests Additive

A.

Measurement	Intercept: Gender Presentation + Sexuality + Following Vowels / GP + Presentation (Feminine + Lesbian + æ) / Feminine + Lesbian	Sexuality: Bisexual	Gender Presentation: Fem & Masc
CoG	5891.050: <2e-16 ***	214.359: 0.195	96.297: 0.598

B.

Measurement	Following Vowels ɑ	Following Vowels ɛ	Following Vowels i	Following Vowels o
CoG	-454.674: 0.168	818.283: 0.461	-47.657: 0.880	146.360: 0.661

C.

Measurement	Intercept: Gender Presentation + Sexuality + Following Vowels / GP + Presentation (Feminine + Lesbian + æ) / Feminine + Lesbian	Sexuality: Bisexual	Gender Presentation: Fem & Masc
F1 ɑ Variance = 0	533.657: < 2e-16 *** SINGULAR FIT	3.306: 0.226 SINGULAR FIT	21.539: 2.95e-13 *** SINGULAR FIT
F1 æ Speaker Variance = 368	742.417: 2.65e-13 ***	-19.384: 0.04042 *	-35.370: 0.00396 **
F1 ɛ Variance = 162.7	623.321: 6.67e-13 ***	-9.416: 0.13357	-25.064: 0.00454 **
F1 i Variance = 180.6	374.481:7.25e-11 ***	4.786:0.403	0.403:0.155
F1 o Variance = 1073	709.972: 3.94e-11 ***	35.766: 0.0227 *	2.816: 0.8486

D.

Measurement	Intercept: Gender Presentation + Sexuality + Following Vowels / GP + Presentation (Feminine + Lesbian + æ) / Feminine + Lesbian	Sexuality: Bisexual	Gender Presentation: Fem & Masc
F2 ɑ Variance = 451.6	1606.093: 1.19e-12 ***	-8.259: 0.43264	-47.372: 0.00422 **
F2 æ Variance = 1032	1717.5559: 7.7e-14 ***	-0.3866: 0.977	-14.7165: 0.346
F2 ɛ Variance = 1041	1742.2891: 2.75e-13 ***	-0.9432: 0.94551	-60.3484: 0.00432 **

F2 i Variance = 6677	2268.786:7.17e-12 ***	-5.342:0.8680	-99.803:0.0218 *
F2 o Variance = 911.9	1286.260:3.63e-10 ***	-51.598:0.000918 ***	-17.416:0.120472

Table 8 - Iteration Models / Likelihood Ratio Tests Interaction

A.

Measurement	Intercept	Sexuality: Bisexual	Gender Presentation
CoG	5561.11:<2e-16 ***	-420.29: 0.1973	279.37: 0.3900

B.

Measurement	Following Vowel: ɑ	Following Vowel: ɛ	Following Vowel: i	Following Vowel: o
CoG	-872.13:0.6328	276.29:0.6360	327.76:0.5155	263.89:0.5684

C.

Measurement	Fem & Masc * Bisexual	Bisexual * ɑ	Bisexual * ɛ	Bisexual * i	Fem & Masc * ɑ
CoG	121.77:0.7074	-906.45:0.3603	287.23:0.5364	167.61:0.6435	1671.02:0.0923

D.

Measurement	Fem & Masc * ɛ	Fem & Masc * i	Bisexual * Fem & Masc * ɑ	Bisexual * Fem & Masc * ɛ	Bisexual * Fem & Masc * i
CoG	-857.98:0.0655	-263.19:0.4676	993.11:0.3163	-223.29:0.6308	-402.32:0.2671

E.

Measurement	Intercept: Gender Presentation + Sexuality + Following Vowels (Feminine * Lesbian)	Sexuality: Bisexual	Gender Presentation: Fem & Masc	Sexuality * Gender Presentation (Bisexual * Fem & Masc)
F1 ɑ: Variance = 0	534.8309: < 2e-16*** SINGULAR FIT	-0.7948: 0.7921 SINGULAR FIT	-22.9706: 4.31e-14*** SINGULAR FIT	-5.8759: 0.0515 SINGULAR FIT
F1 æ Variance = 362.9	741.892: 5.76e-14 ***	-20.773: 0.04744 *	-35.264: 0.00375 **	2.974: 0.74925
F1 ɛ Variance = 80.91	625.716: 1.66e-14 ***	-4.555: 0.40399	-25.549: 0.00109 **	-10.534: 0.07564
F1 i Variance = 150	375.426:2.86e-11 ***	7.826:0.208	-9.677:0.130	-6.331:0.299
F1 o Variance = 142.4	648.145: < 2e-16 ***	-31.461:1.11e-05 ***	-14.9790.00440 **	18.4660.00105 **

F.

Measurement	Intercept: Gender Presentation + Sexuality + Following Vowels (Feminine * Lesbian)	Sexuality: Bisexual	Gender Presentation: Fem & Masc	Sexuality * Gender Presentation (Bisexual * Fem & Masc)
F2 ɑ Variance = 405	1607.190: 3.87e-13***	-4.487: 0.6868	- 47.612: 0.0032**	-8.349: 0.4600
F2 æ Variance = 939	1715.897: 2.93e-14***	-5.346: 0.716	-14.586: 0.333	10.830: 0.466

F2 ɛ Variance = 796.4	1739.192: 6.14e-14***	-9.346: 0.5141	-60.044: 0.0026**	17.928: 0.2275
F2 i Variance = 2707	2277.205:1.74e-12***	28.114:0.26671	-95.520:0.00437**	-72.113:0.01704*
F2 o Variance = 879.3	1289.824:2.07e-12***	-45.960:0.00159**	-19.701:0.06195.	-10.805:0.25827

Table 9 - F1 Rankings

Speaker	ɑ	æ	ɛ	i	o
1	520.1 (3)	759.0 (2)	629.1 (3)	355.9 (7)	658.2 (2)
2	516.0 (4)	690.2 (7)	597.6 (4)	373.2 (5)	630.8 (3)
3	514.2 (5)	696.4 (6)	593.0 (6)	381.3 (2)	625.1 (5)
4	552.7 (2)	800.9 (1)	645.3 (2)	370.9 (6)	713.1 (1)
5	510.9 (6)	702.1 (4)	586.9 (7)	378.3 (3)	609.6 (8)
6	562.9 (1)	753.4 (3)	657.2 (1)	399.3 (1)	613.2 (7)
Nyx	507.1 (7)	662.8 (8)	596.1 (5)	374.0 (4)	629.5 (4)
Starra	480.6 (8)	698.1 (5)	561.1 (8)	333.4 (8)	614.9 (6)

Table 10 - F2 Rankings

Speaker	ɑ	æ	ɛ	i	o
1	1549.8 (5)	1731 (3)	1661 (7)	2244 (3)	1327.9 (2)
2	1598.6 (4)	1661 (7)	1682 (6)	2206 (4)	1325.7 (3)
3	1507.1 (8)	1696.1 (6)	1608.2 (8)	2147 (7)	1249.1 (5)
4	1651 (2)	1747 (2)	1827 (1)	2272 (2)	1344.7 (1)
5	1541.0 (6)	1656 (8)	1705 (5)	2202 (5)	1176.5 (8)

6	1659 (1)	1714 (4)	1772 (2)	2473 (1)	1274.4 (4)
Nyx	1599 (3)	1770 (1)	1727 (3)	2198 (6)	1238.1 (6)
Starra	1529.7 (7)	1697 (5)	1712 (4)	1998 (8)	1180.2 (7)

Table 11 - Average Rankings

Speaker	CoG Rankings	F1 Rankings (all vowels)	F2 Rankings (all vowels)	All Rankings Average
1	6283.7 (1)	584.46 (3)	1702.74 (4)	2.6 (3)
2	5239 (6)	561.56 (5)	1694.66 (5)	5.33 (4/5)
3	5551 (5)	562.0 (4)	1641.5 (7)	5.33 (4/5)
4	5776 (3)	616.58 (1)	1768.34 (2)	2 (2)
5	5198.1 (7)	557.56 (6)	1656.1 (6)	6.33 (7)
6	5919 (2)	597.2 (2)	1778.48 (1)	1.66 (1)
Nyx	4785 (8)	553.9 (7)	1706.42 (3)	6 (6)
Starra	5681 (4)	537.62 (8)	1623.38 (8)	6.66 (8)

Table 12 - Question Type Models

Measurement	Question Type - (Intercept)	Question Type - Reading Passage
Center of Gravity (Hz)	5505.7: 8.83e-11***	445.3:0.000667***
F1 (Hz)	574.67556: 2.08e-12***	-21.81709: 2.28e-05***
F2 (Hz)	1676.23702: 1.51e-13***	-50.32991: 3.28e-05***

Table 13 - Vowel Significance - Based on Iteration Interaction Models

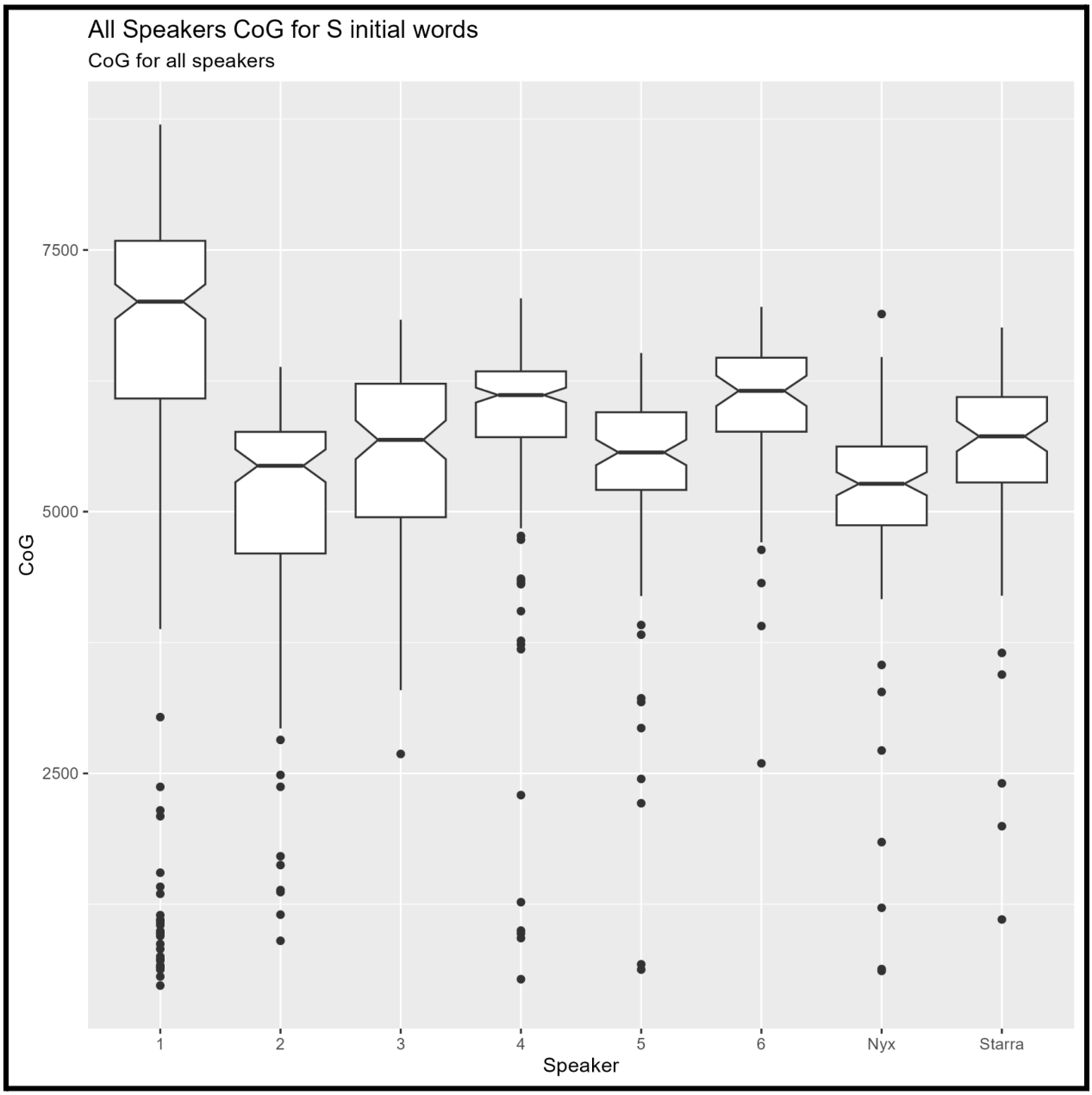
Measurement	Sexuality Significance	Gender Presentation Significance	Sexuality * Gender Presentation Significance
F1	**/o/** and /æ/	/ɑ/, /æ/, /ɛ/, and **/o/**	**/o/**
F2	/o/	/ɑ/, /ɛ/, and /i/	/i/

Models

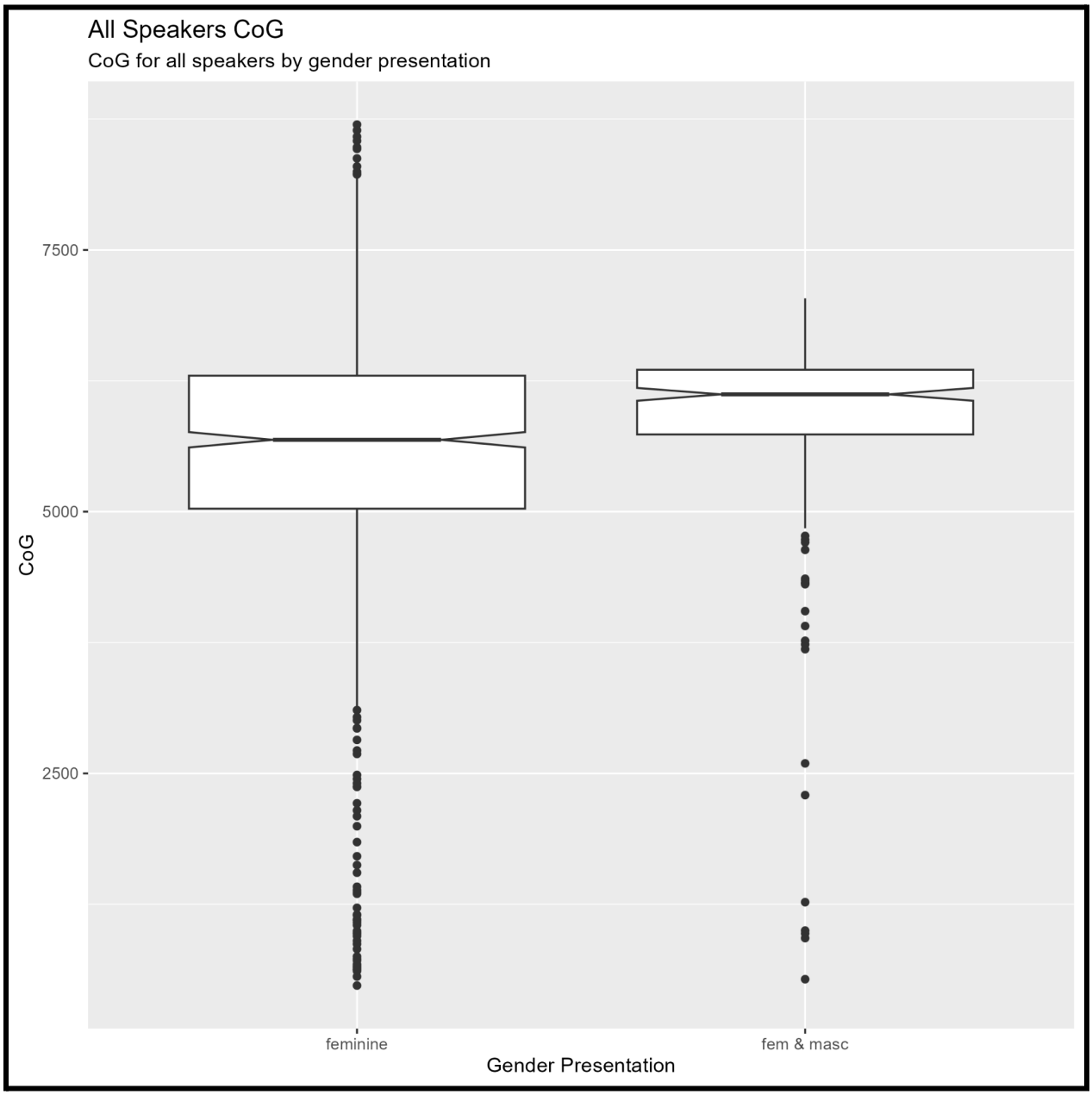
Plots

F1z per vowel?
F2z per vowel?

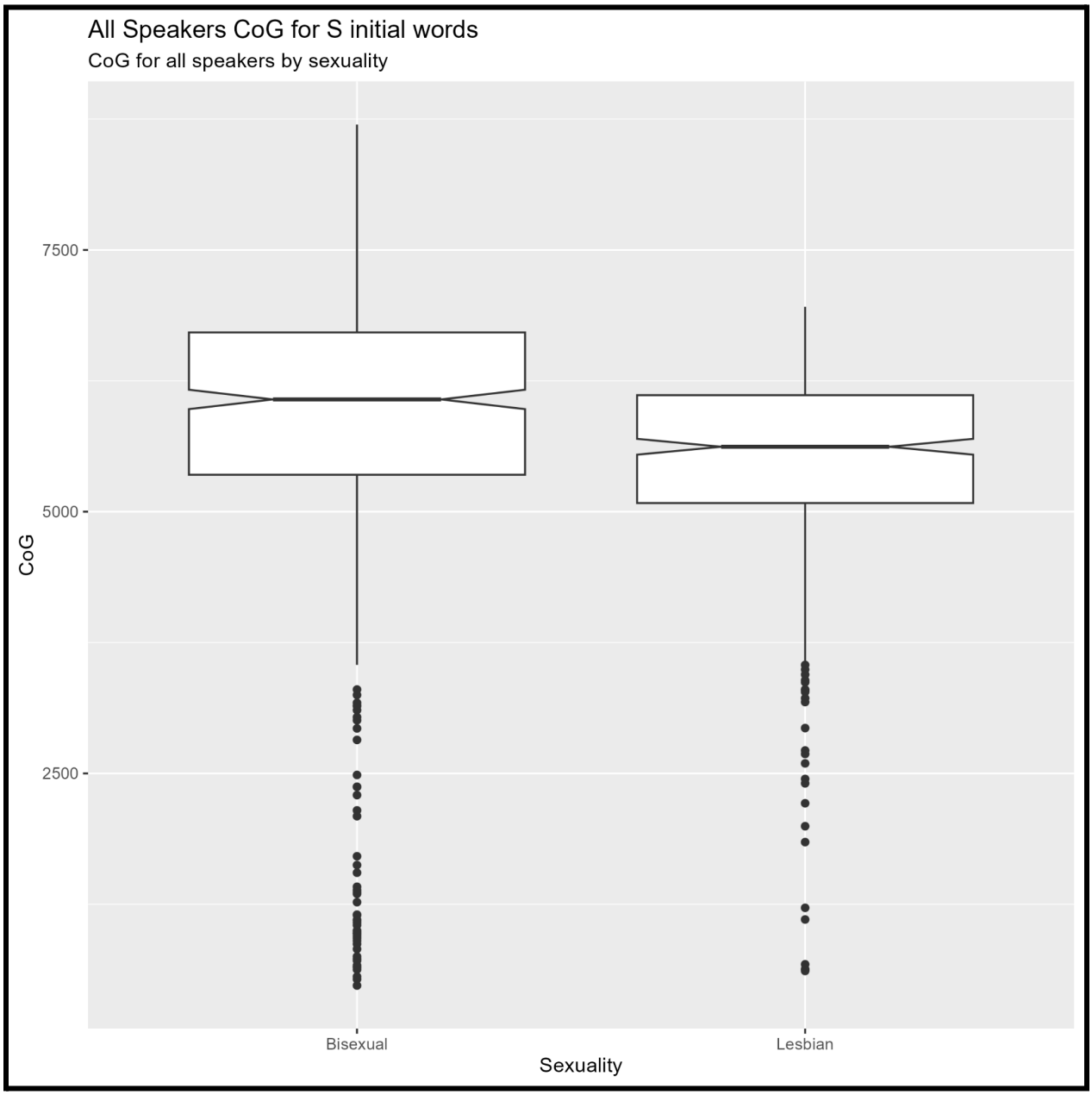
Plot 1A


All Speakers CoG for S initial words
CoG for all speakers
CoG
7500
5000
2500
1
2
3
4
5
6
Nyx
Starra
Speaker

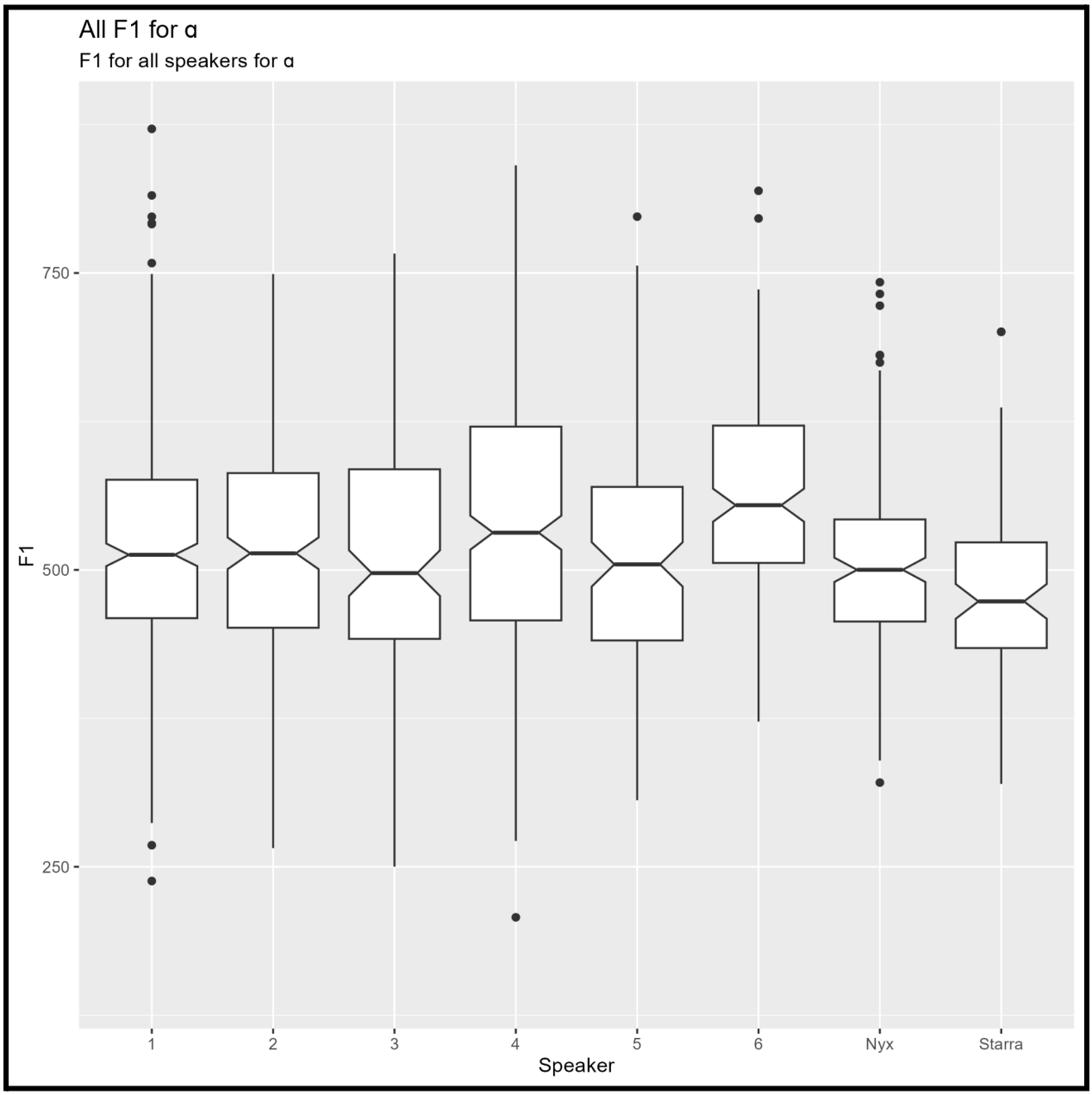
Plot 1B


All Speakers CoG
CoG for all speakers by gender presentation
7500
5000
2500
CoG
feminine
fem & masc
Gender Presentation

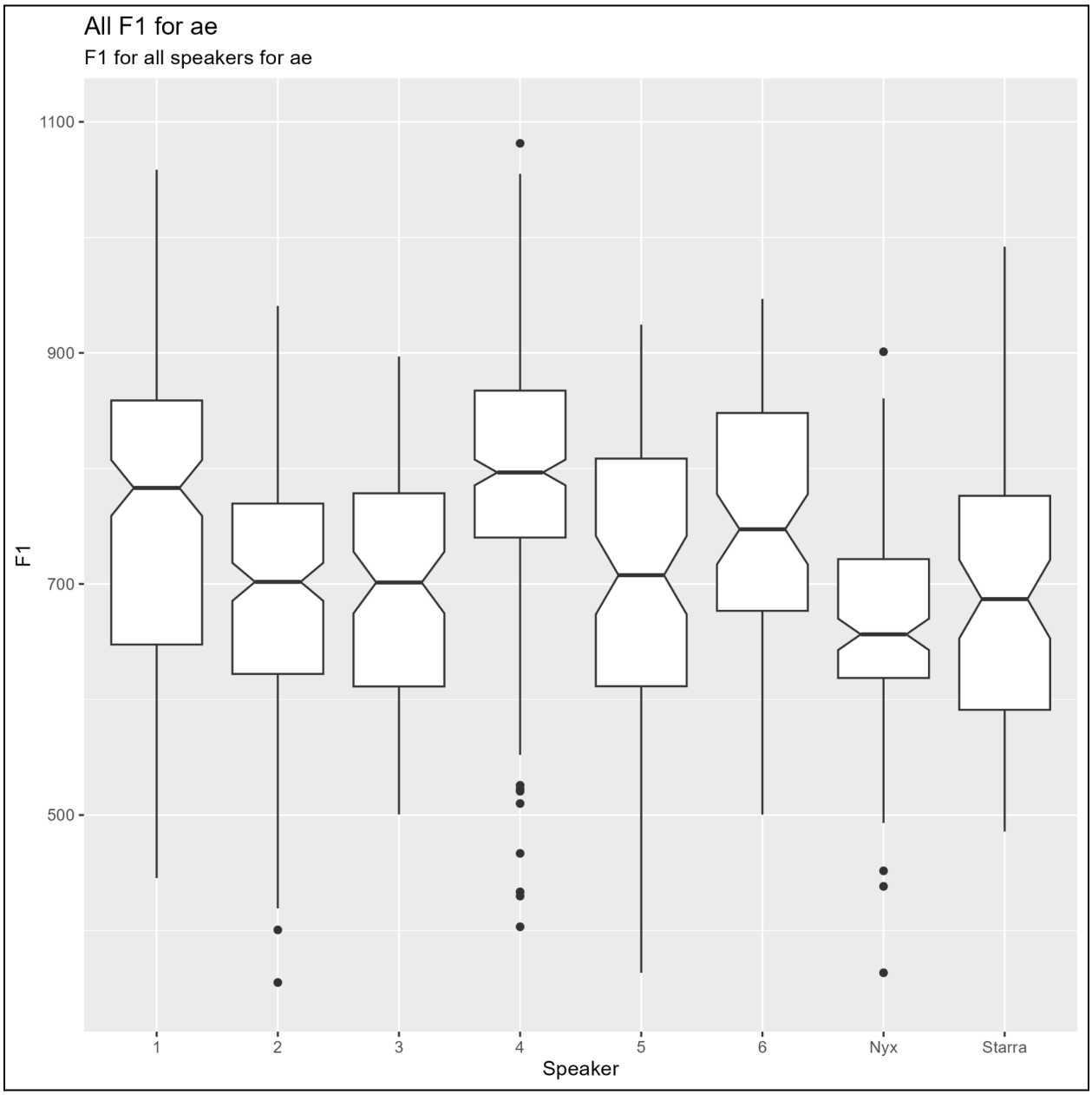
Plot 1C


All Speakers CoG for S initial words
CoG for all speakers by sexuality
7500
5000
2500
CoG
Bisexual
Lesbian
Sexuality

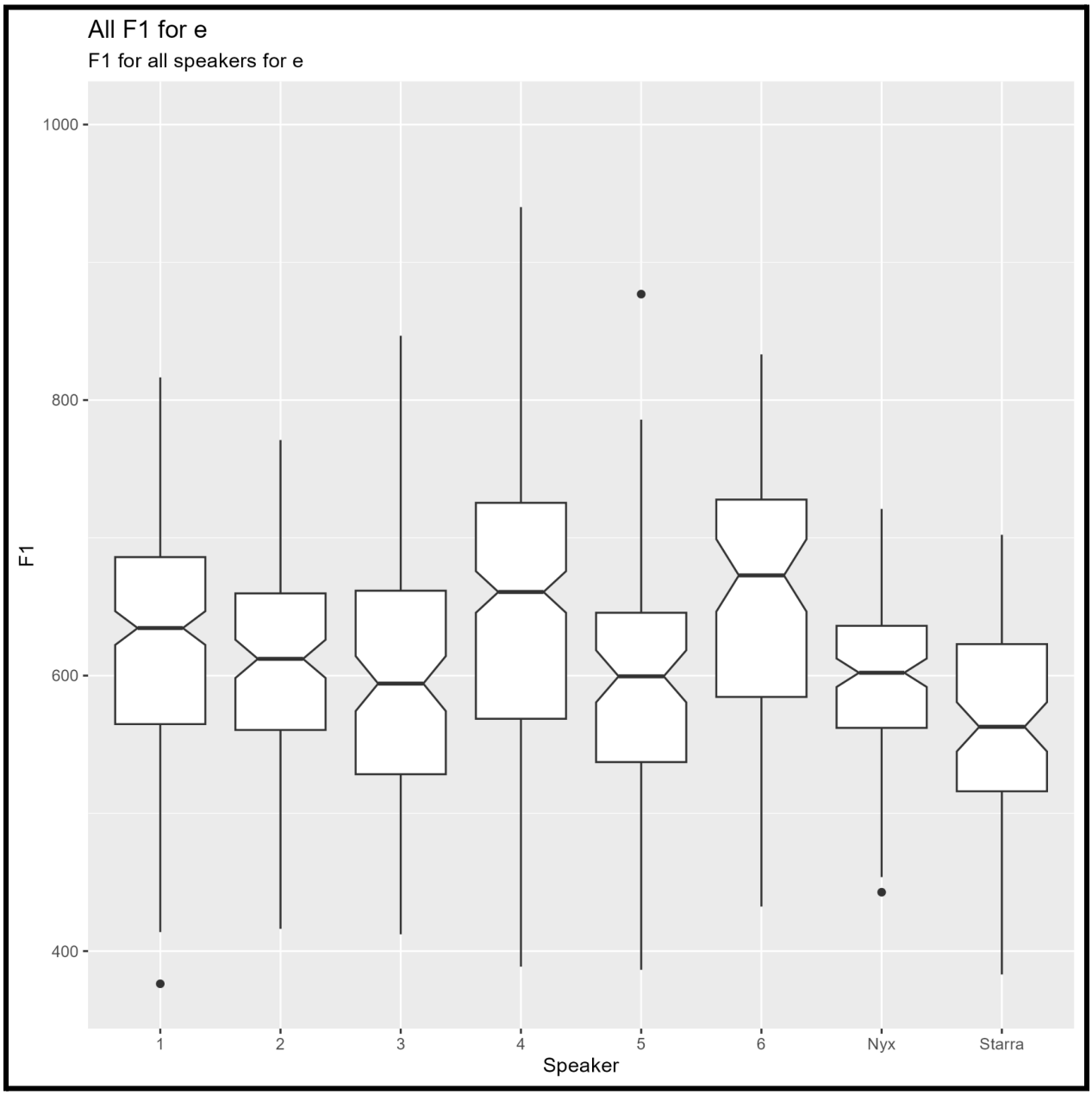
Plot 2A


All F1 for ɑ
F1 for all speakers for ɑ
F1
750
500
250
1
2
3
4
5
6
Nyx
Starra
Speaker

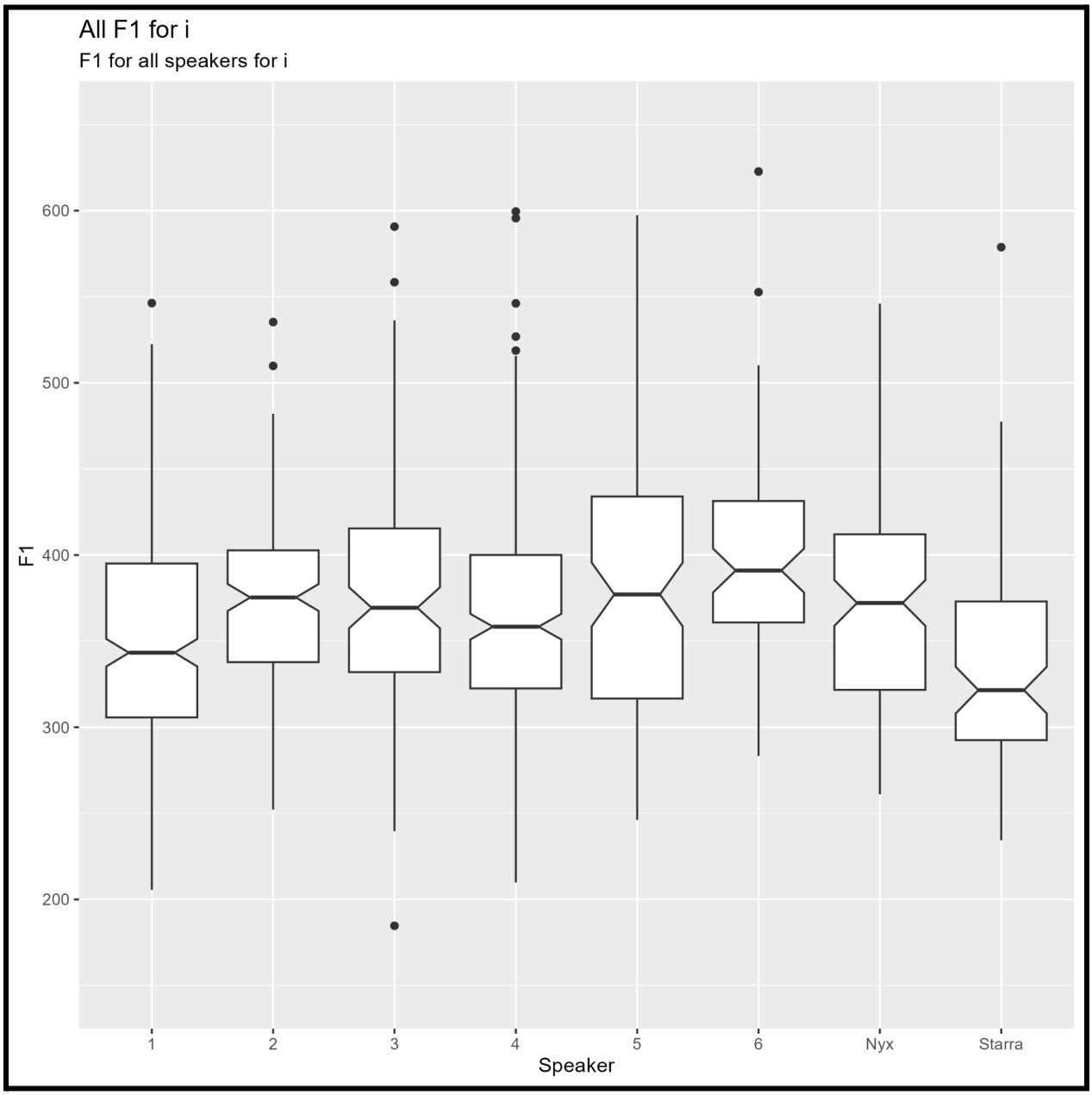
Plot 2B


All F1 for ae
F1 for all speakers for ae
1100
900
700
500
F1
1
2
3
4
5
6
Nyx
Starra
Speaker

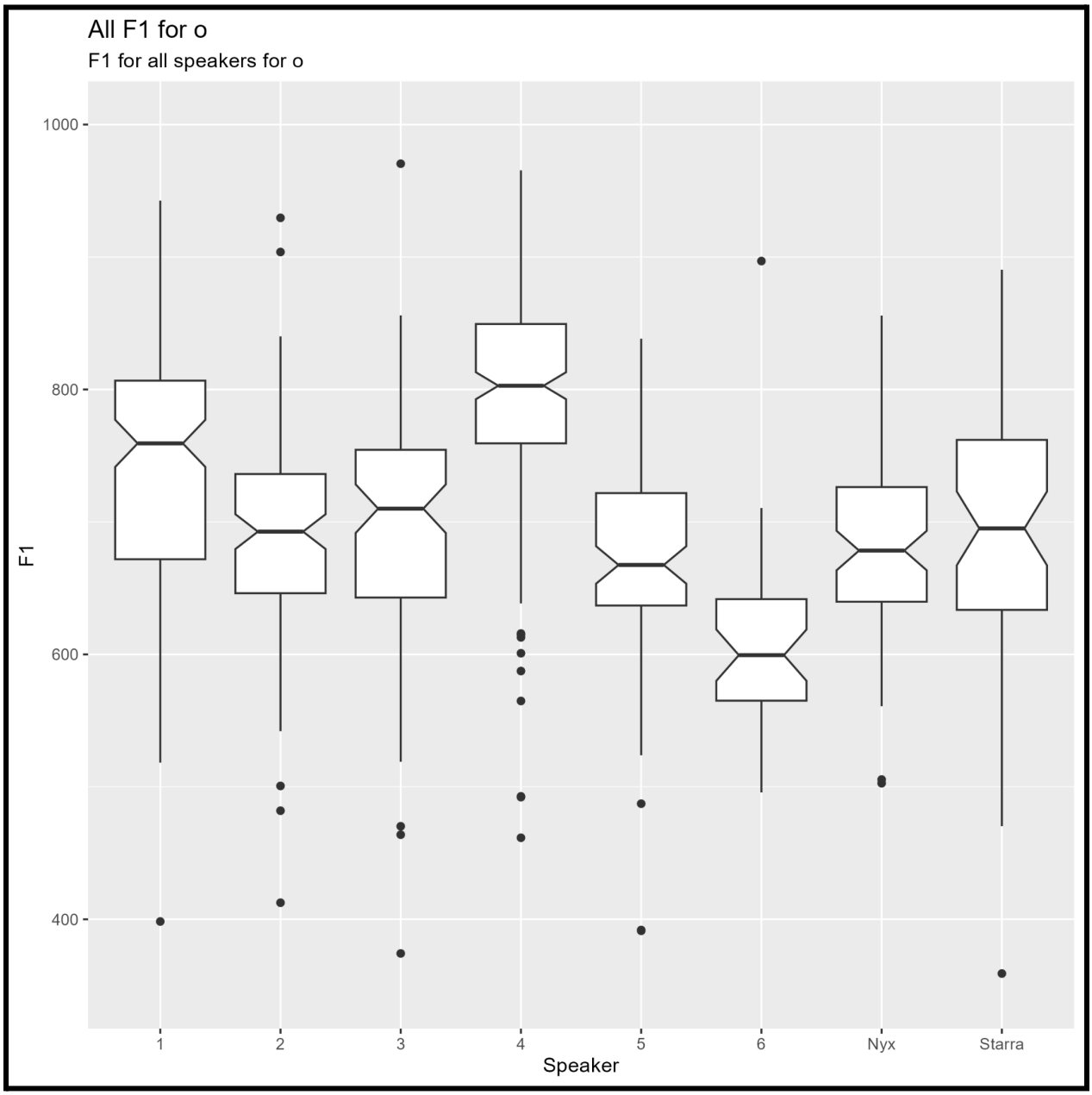
Plot 2C


All F1 for e
F1 for all speakers for e
1000
800
600
400
F1
1
2
3
4
5
6
Nyx
Starra
Speaker

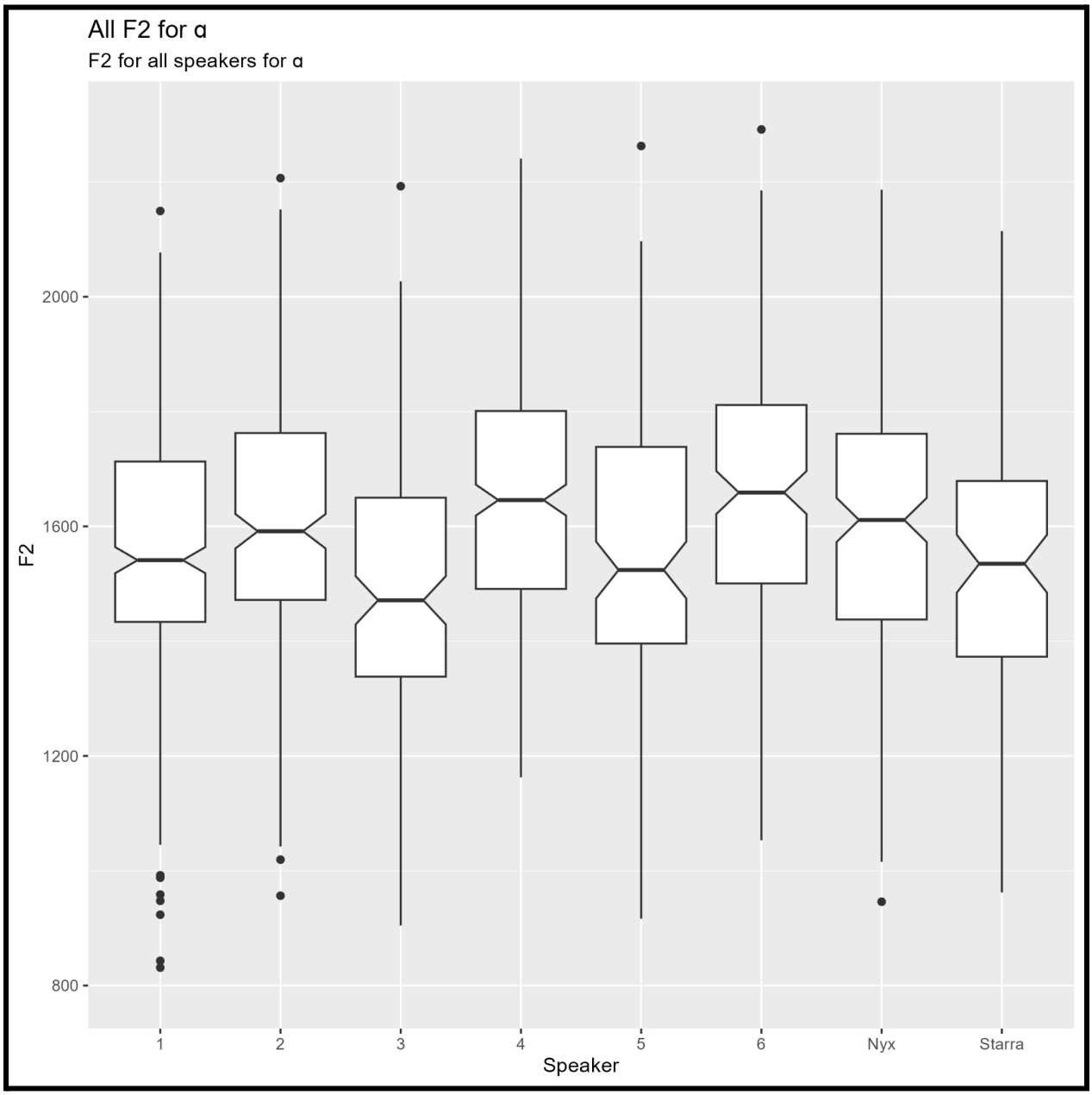
Plot 2D


All F1 for i
F1 for all speakers for i
F1
600
500
400
300
200
1
2
3
4
5
6
Nyx
Starra
Speaker

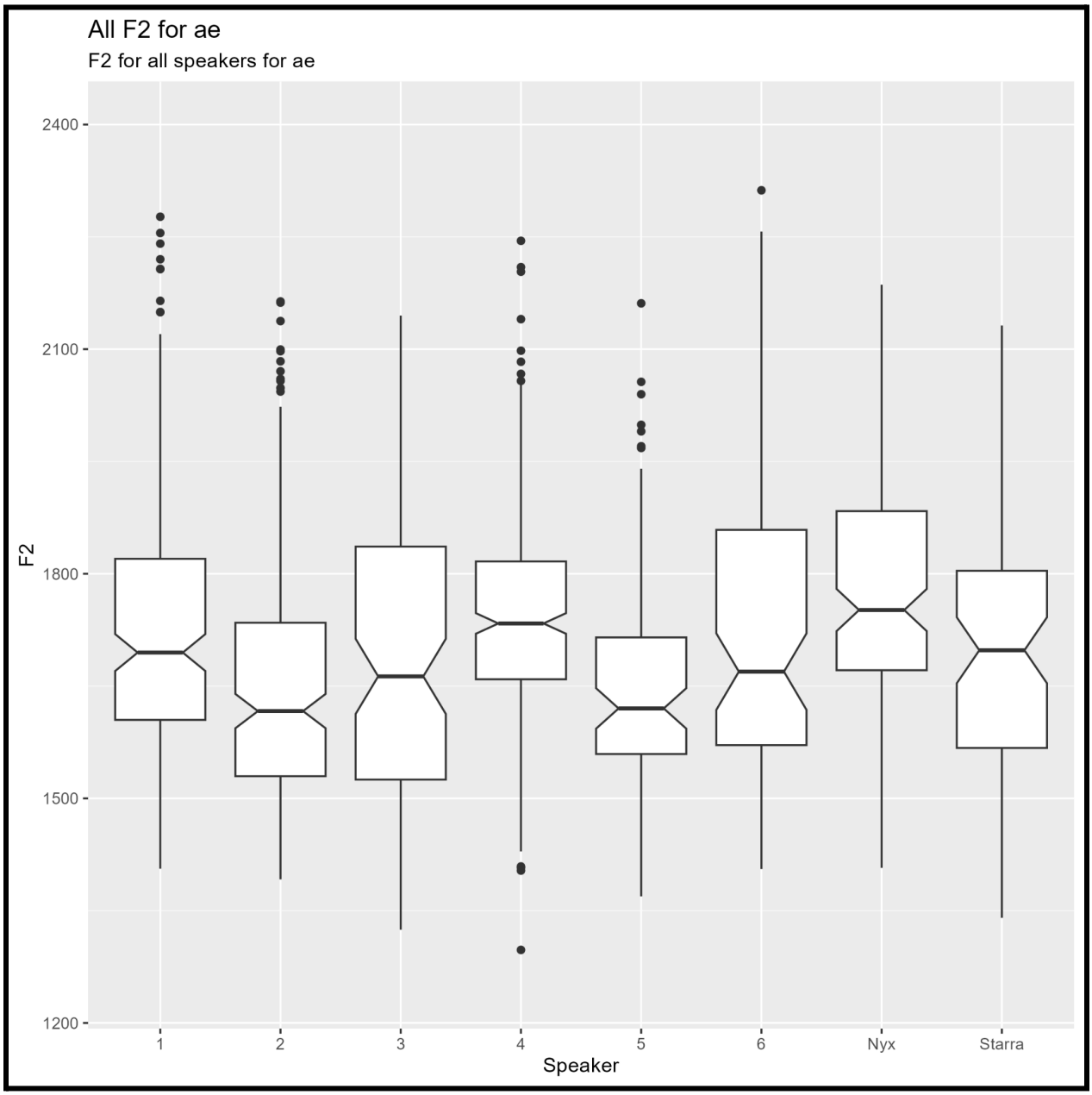
Plot 2E


All F1 for o
F1 for all speakers for o
1000
800
600
400
F1
1
2
3
4
5
6
Nyx
Starra
Speaker

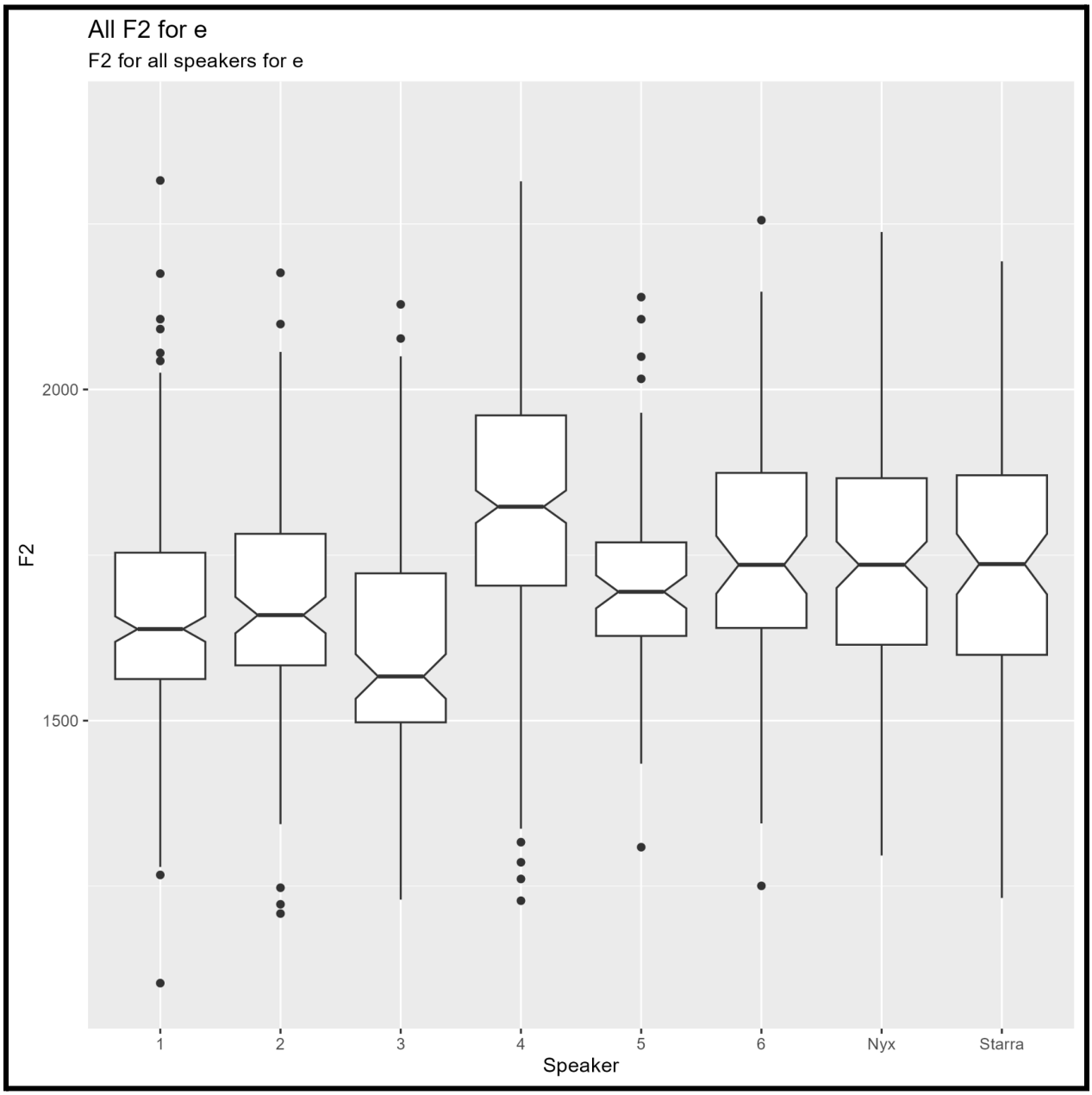
Plot 3A


All F2 for ɑ
F2 for all speakers for ɑ
F2
2000
1600
1200
800
1
2
3
4
5
6
Nyx
Starra
Speaker

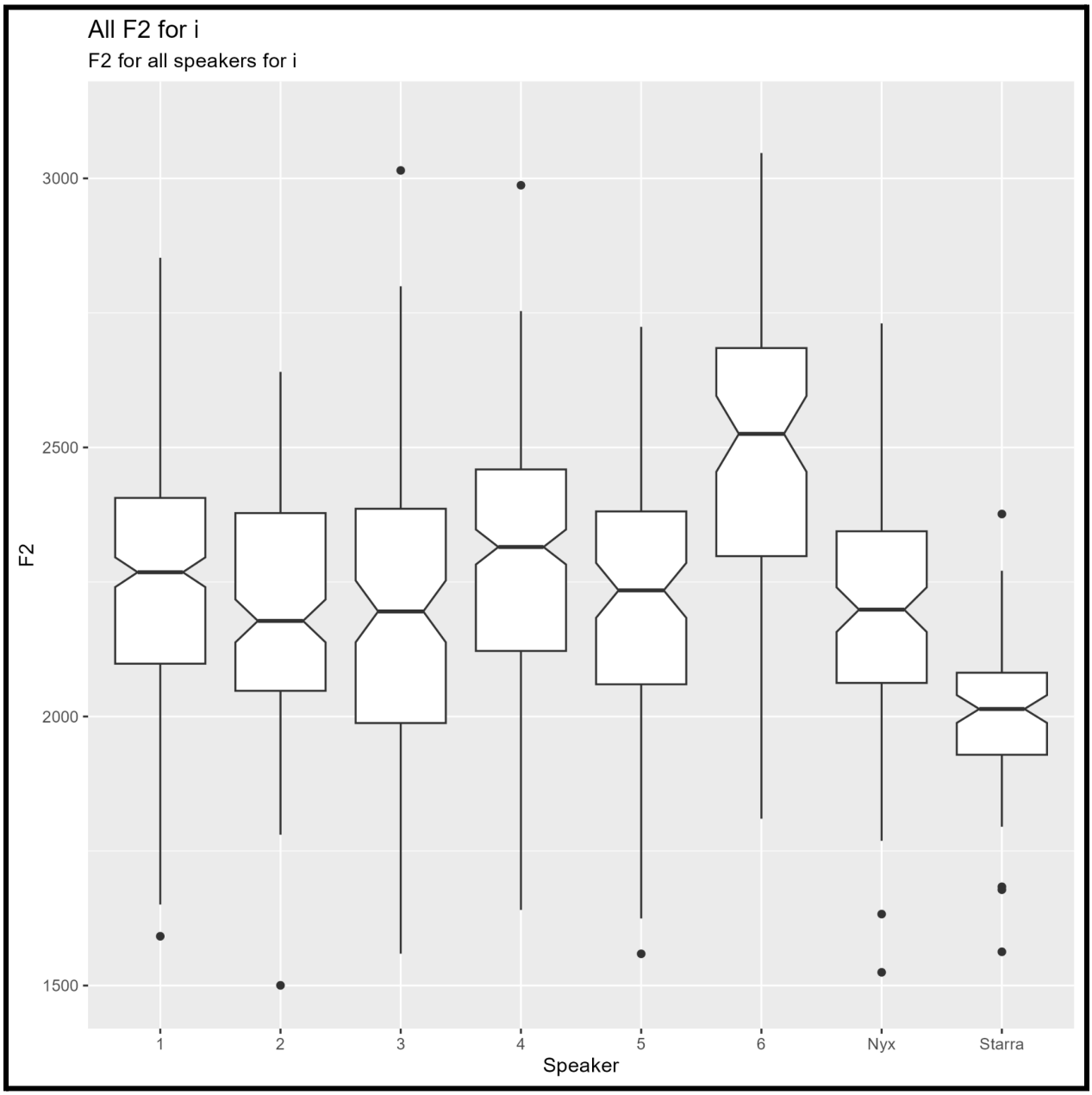
Plot 3B


All F2 for ae
F2 for all speakers for ae
2400
2100
1800
1500
1200
F2
1
2
3
4
5
6
Nyx
Starra
Speaker

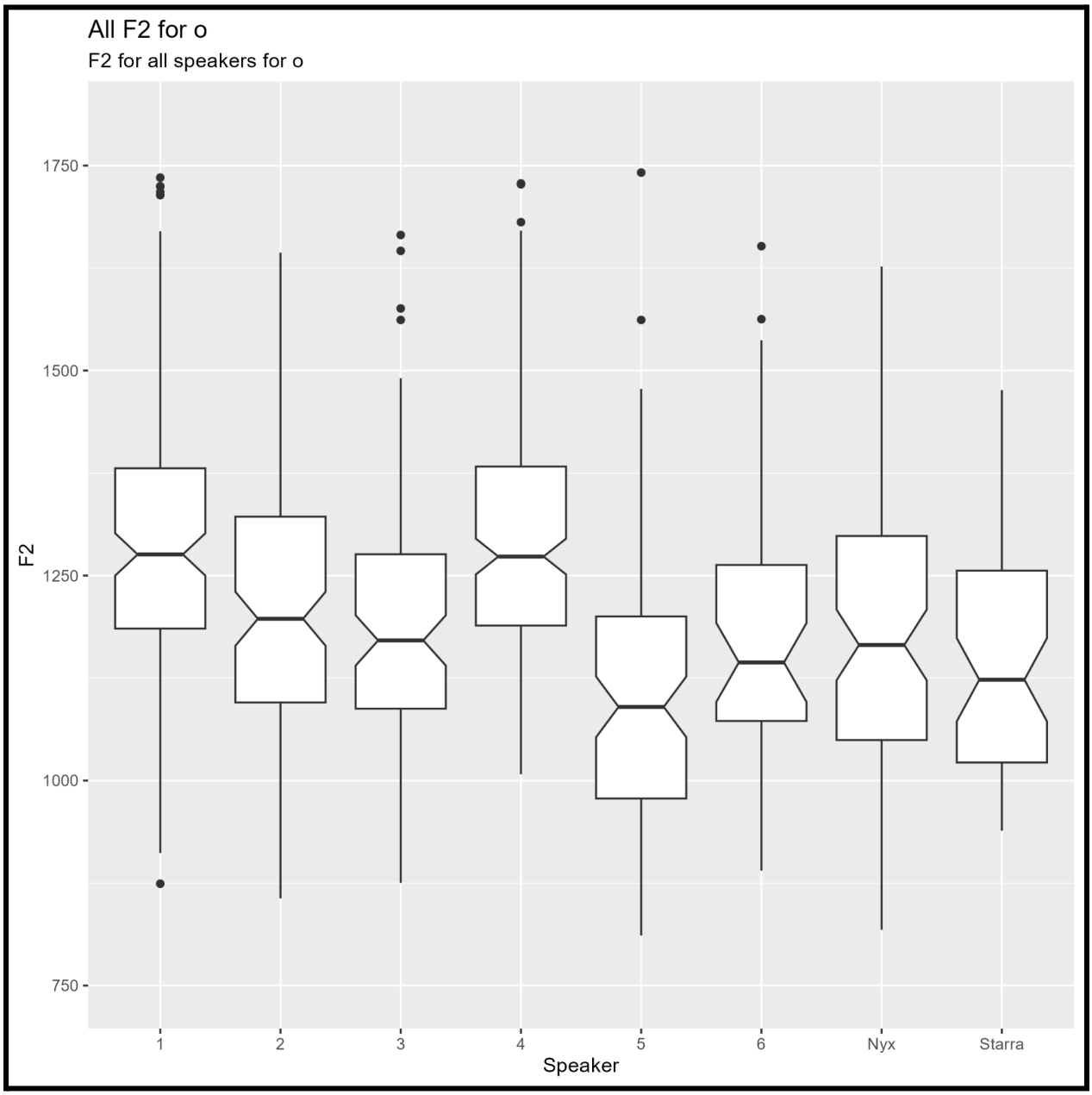
Plot 3C


All F2 for e
F2 for all speakers for e
F2
2000
1500
1
2
3
4
5
6
Nyx
Starra
Speaker

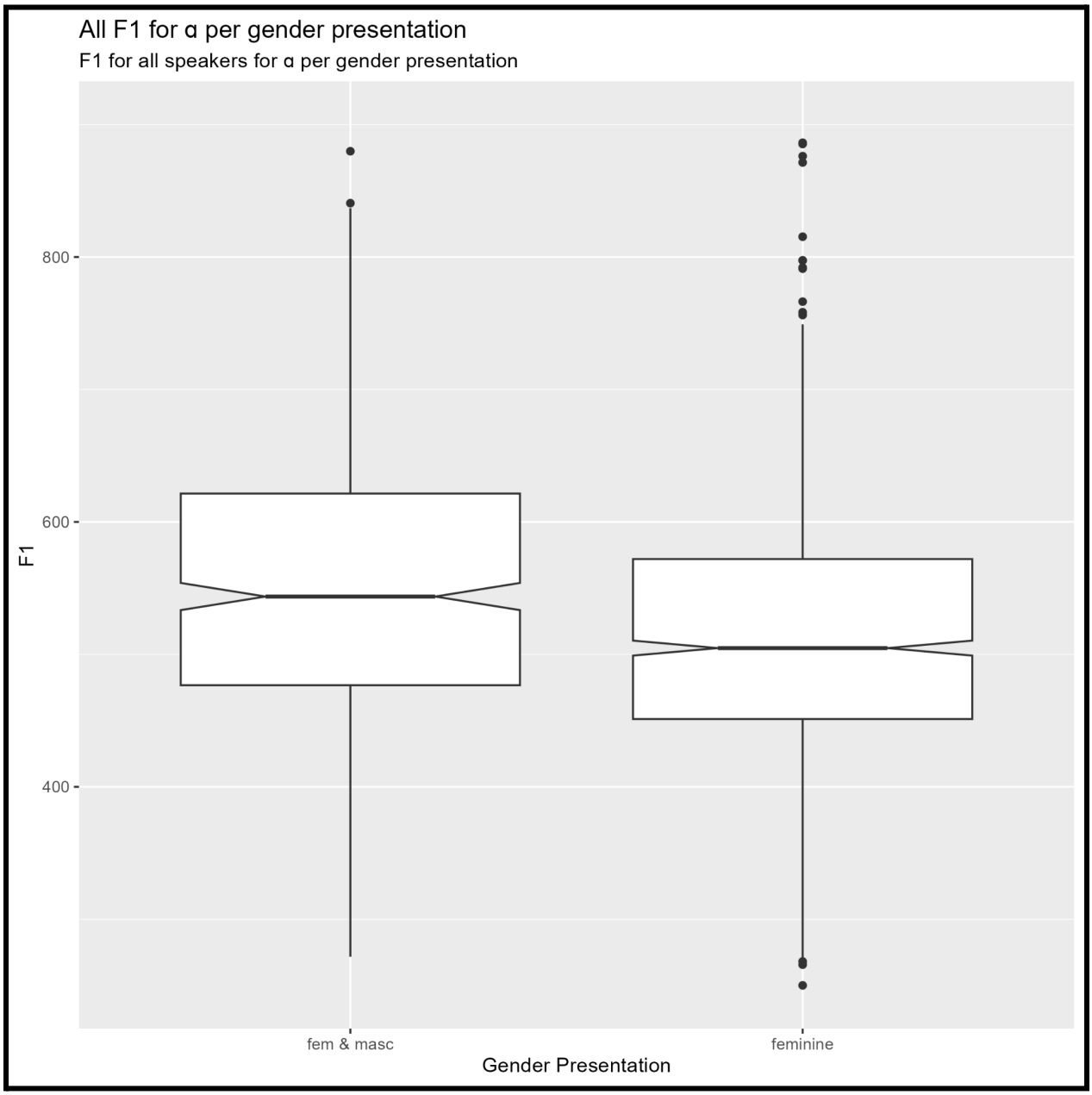
Plot 3D


All F2 for i
F2 for all speakers for i
3000
2500
2000
1500
F2
1
2
3
4
5
6
Nyx
Starra
Speaker

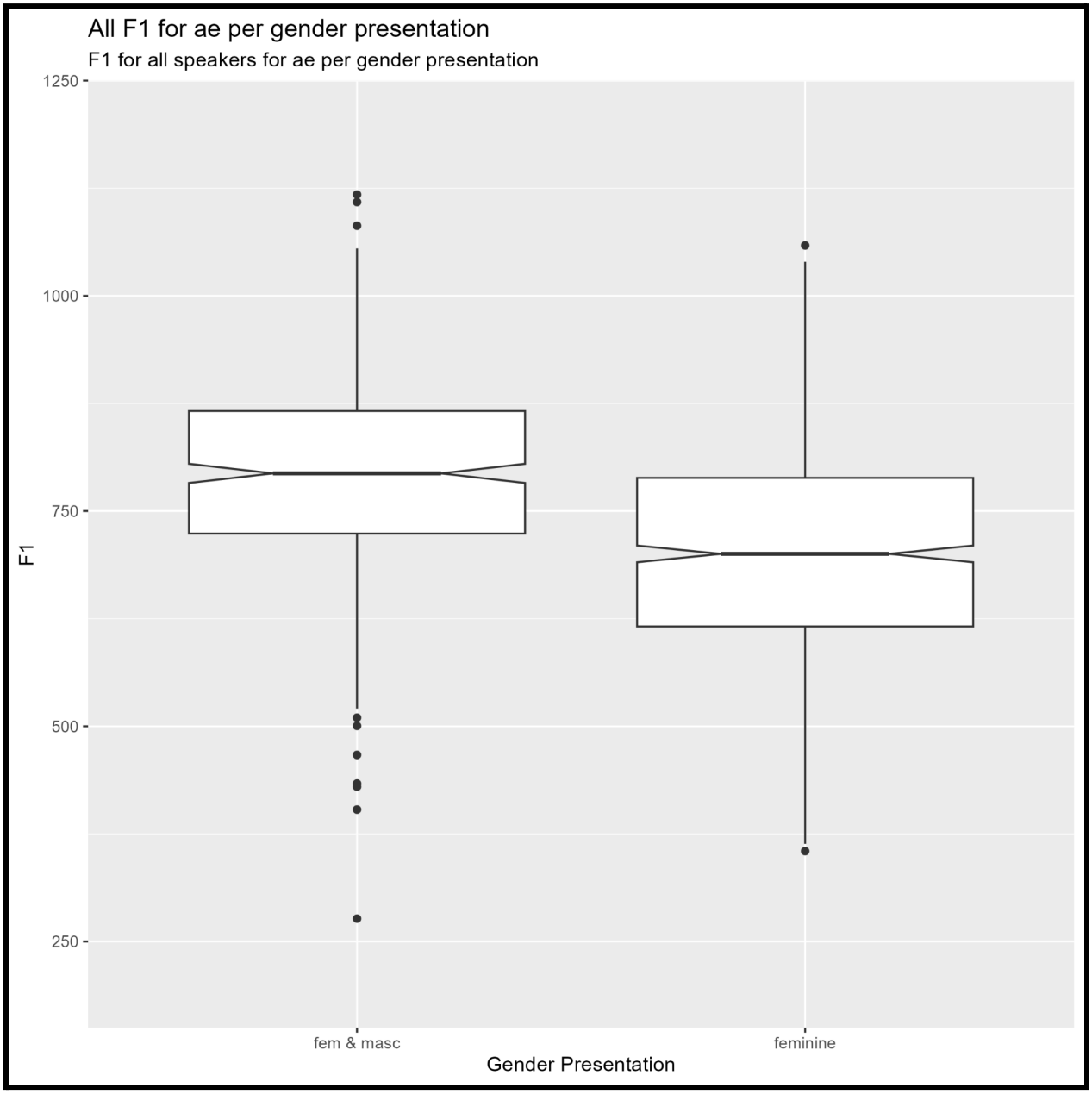
Plot 3E


All F2 for o
F2 for all speakers for o
F2
1750
1500
1250
1000
750
1
2
3
4
5
6
Nyx
Starra
Speaker

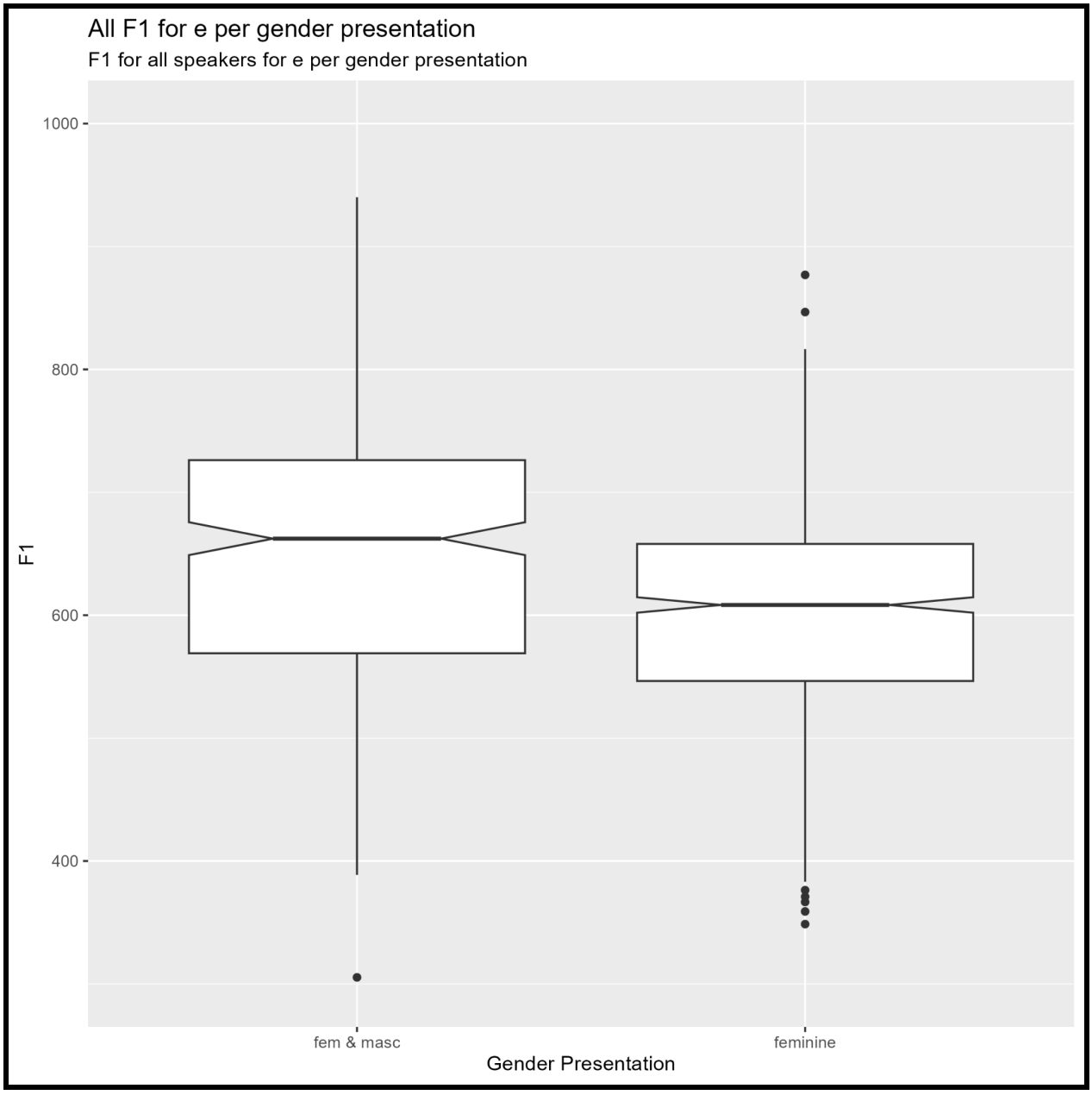
Plot 4A


All F1 for ɑ per gender presentation
F1 for all speakers for ɑ per gender presentation
800
600
400
F1
fem & masc
feminine
Gender Presentation

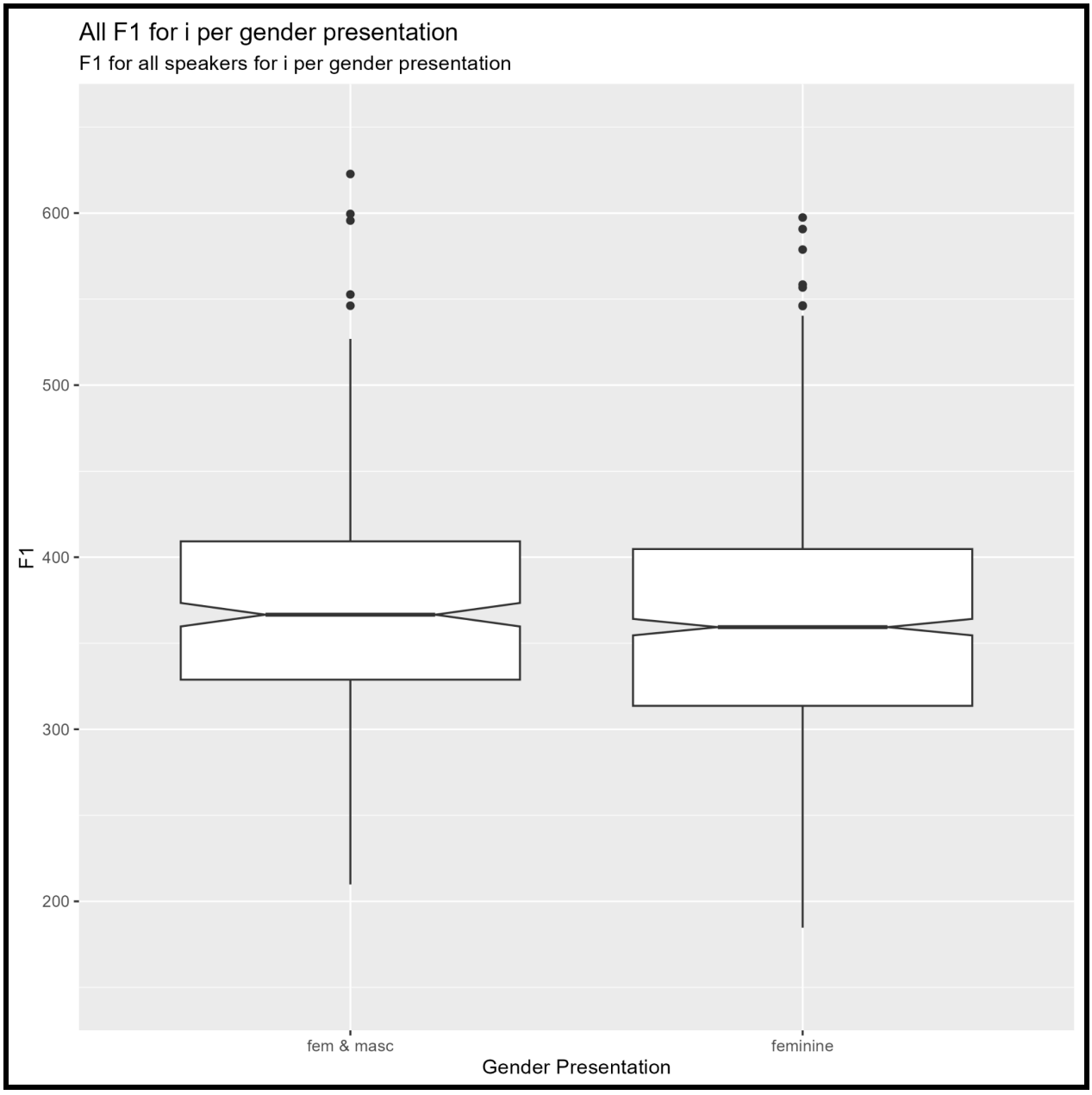
Plot 4B



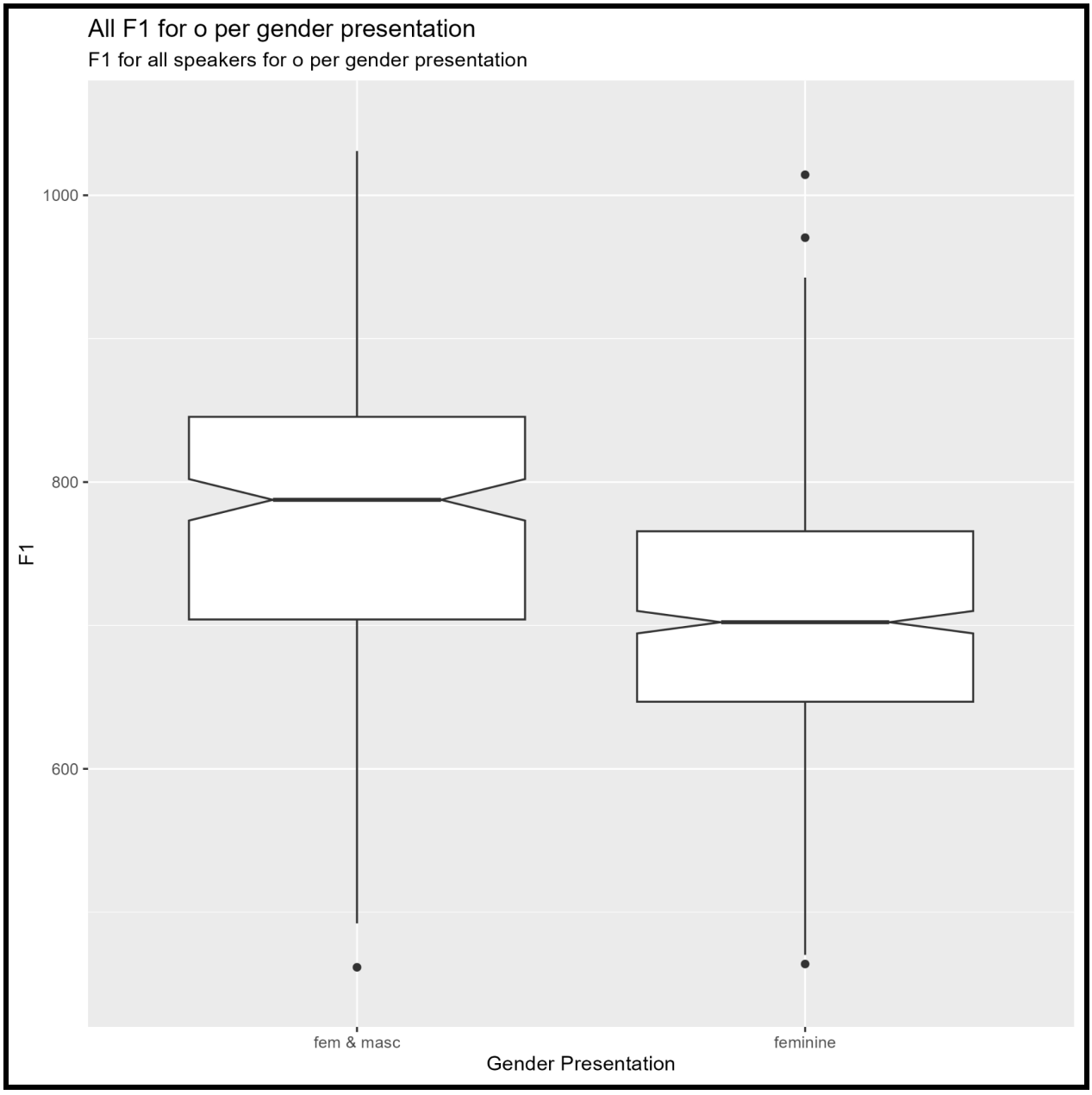
Plot 4C


All F1 for e per gender presentation
F1 for all speakers for e per gender presentation
1000
800
600
400
F1
fem & masc
feminine
Gender Presentation

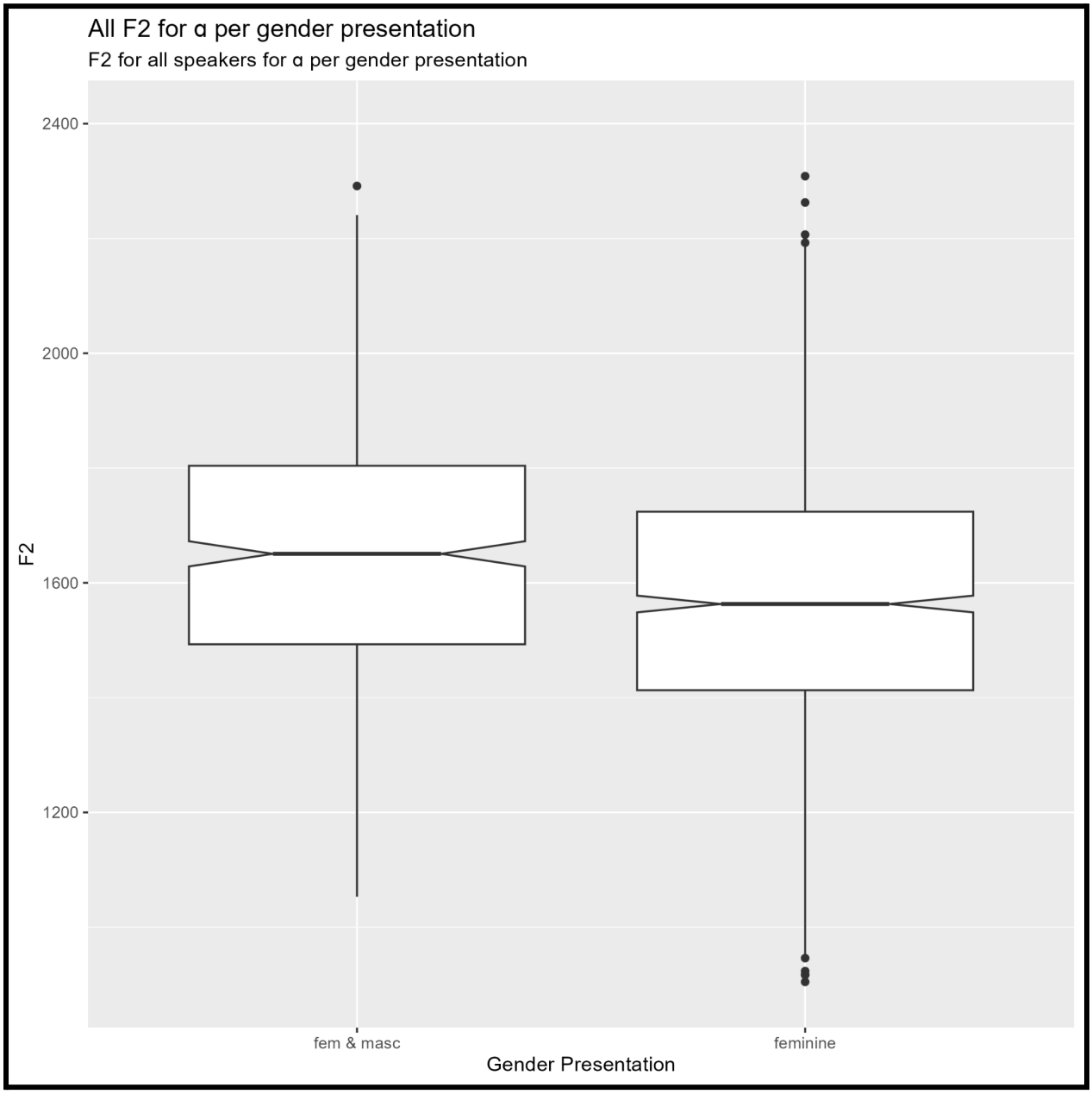
Plot 4D


All F1 for i per gender presentation
F1 for all speakers for i per gender presentation
600
500
400
300
200
F1
fem & masc
feminine
Gender Presentation

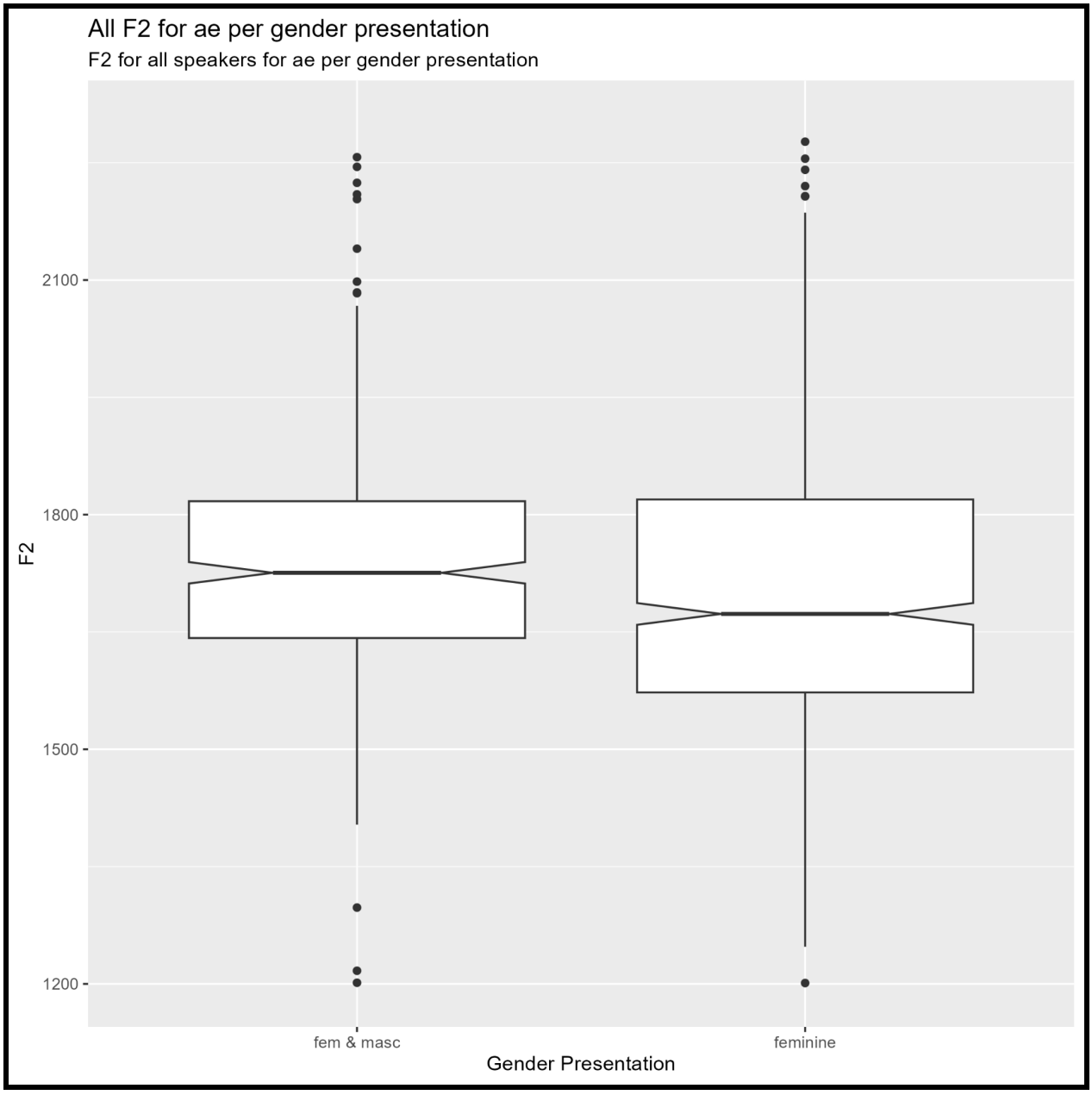
Plot 4E


All F1 for o per gender presentation
F1 for all speakers for o per gender presentation
1000
800
600
F1
fem & masc
feminine
Gender Presentation

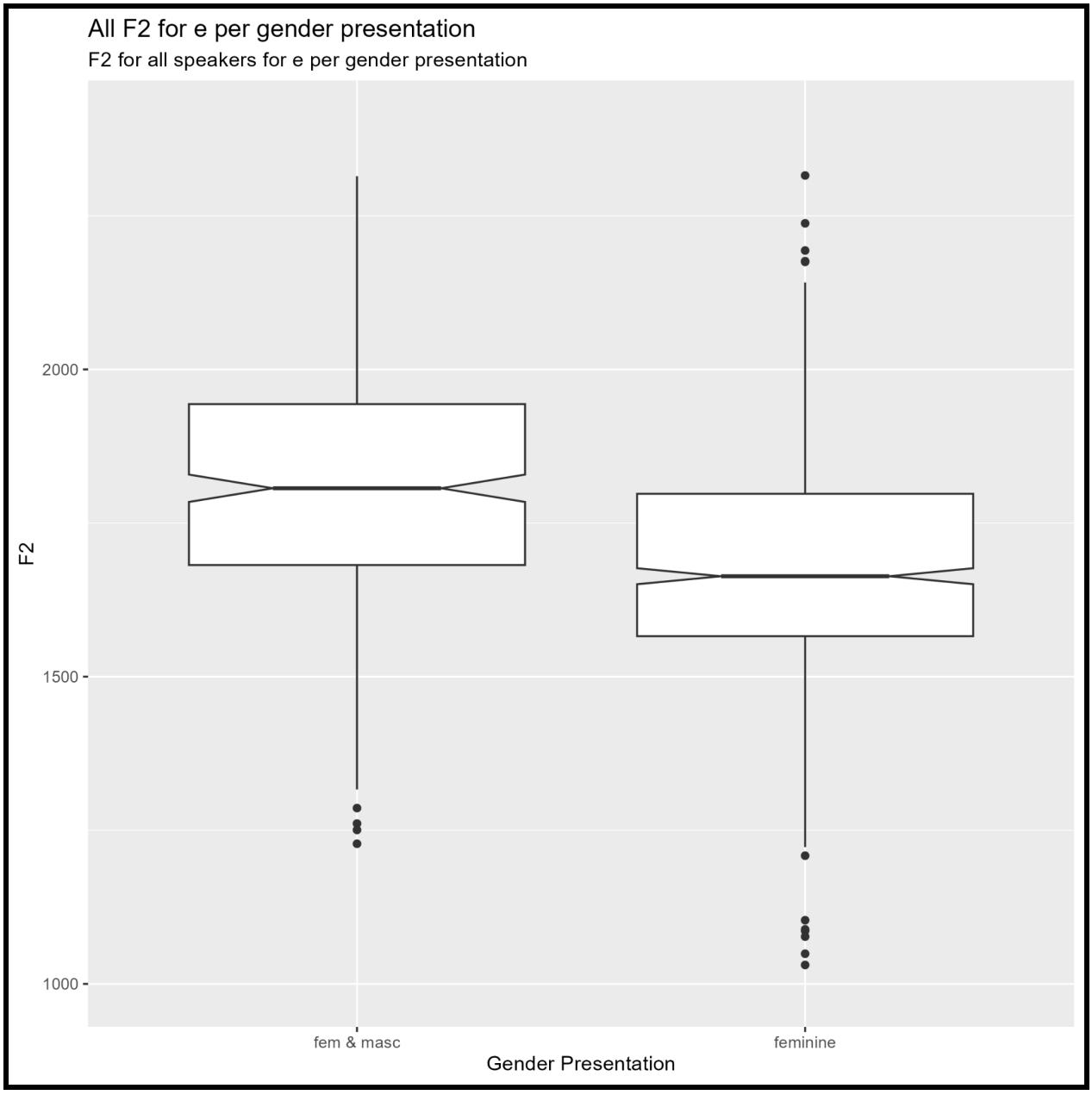
Plot 5A



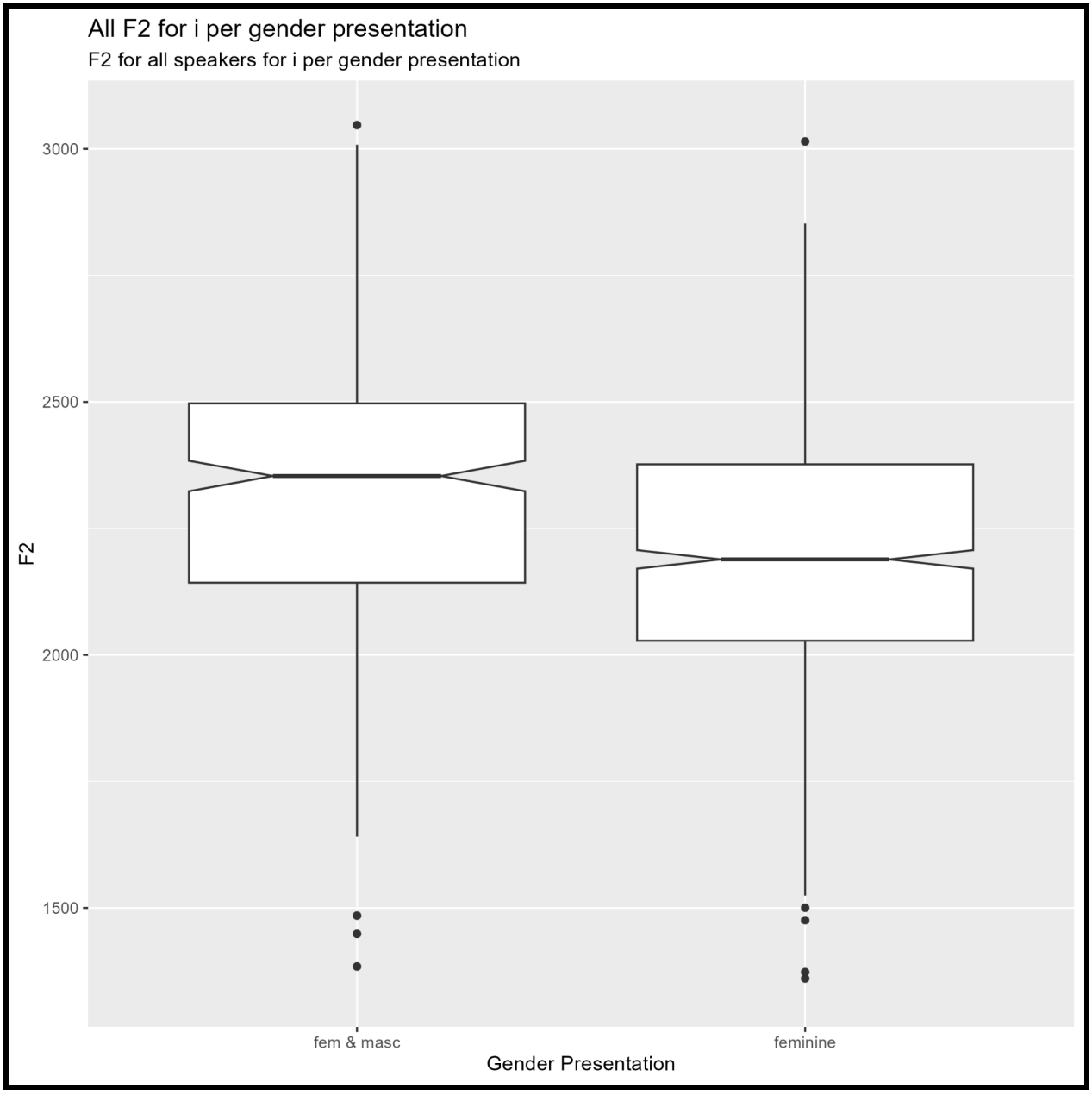
Plot 5B


All F2 for ae per gender presentation
F2 for all speakers for ae per gender presentation
2100
1800
1500
1200
F2
fem & masc
feminine
Gender Presentation

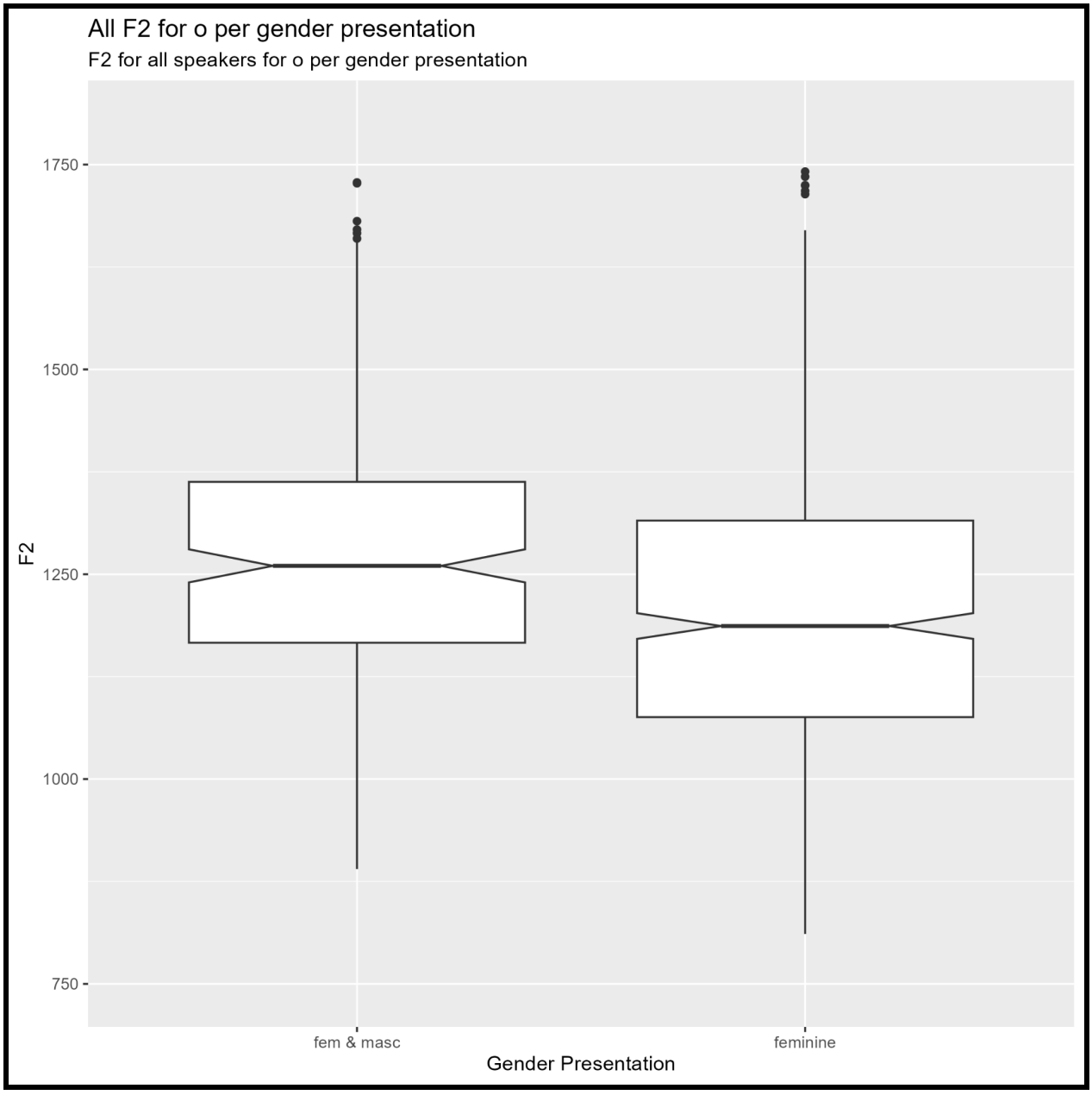
Plot 5C


All F2 for e per gender presentation
F2 for all speakers for e per gender presentation
F2
2000
1500
1000
fem & masc
feminine
Gender Presentation

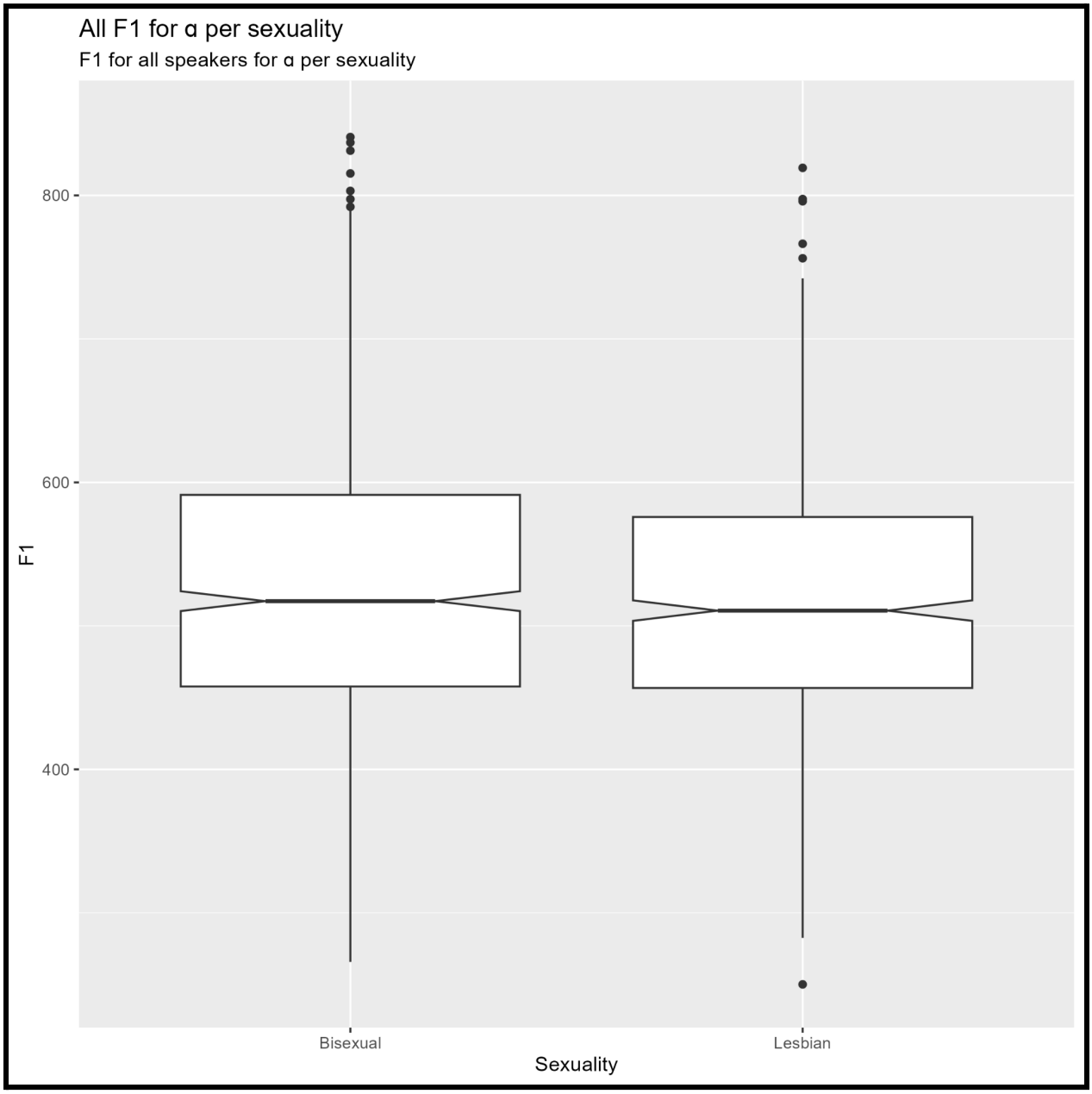
Plot 5D


All F2 for i per gender presentation
F2 for all speakers for i per gender presentation
3000
2500
2000
1500
F2
fem & masc
feminine
Gender Presentation

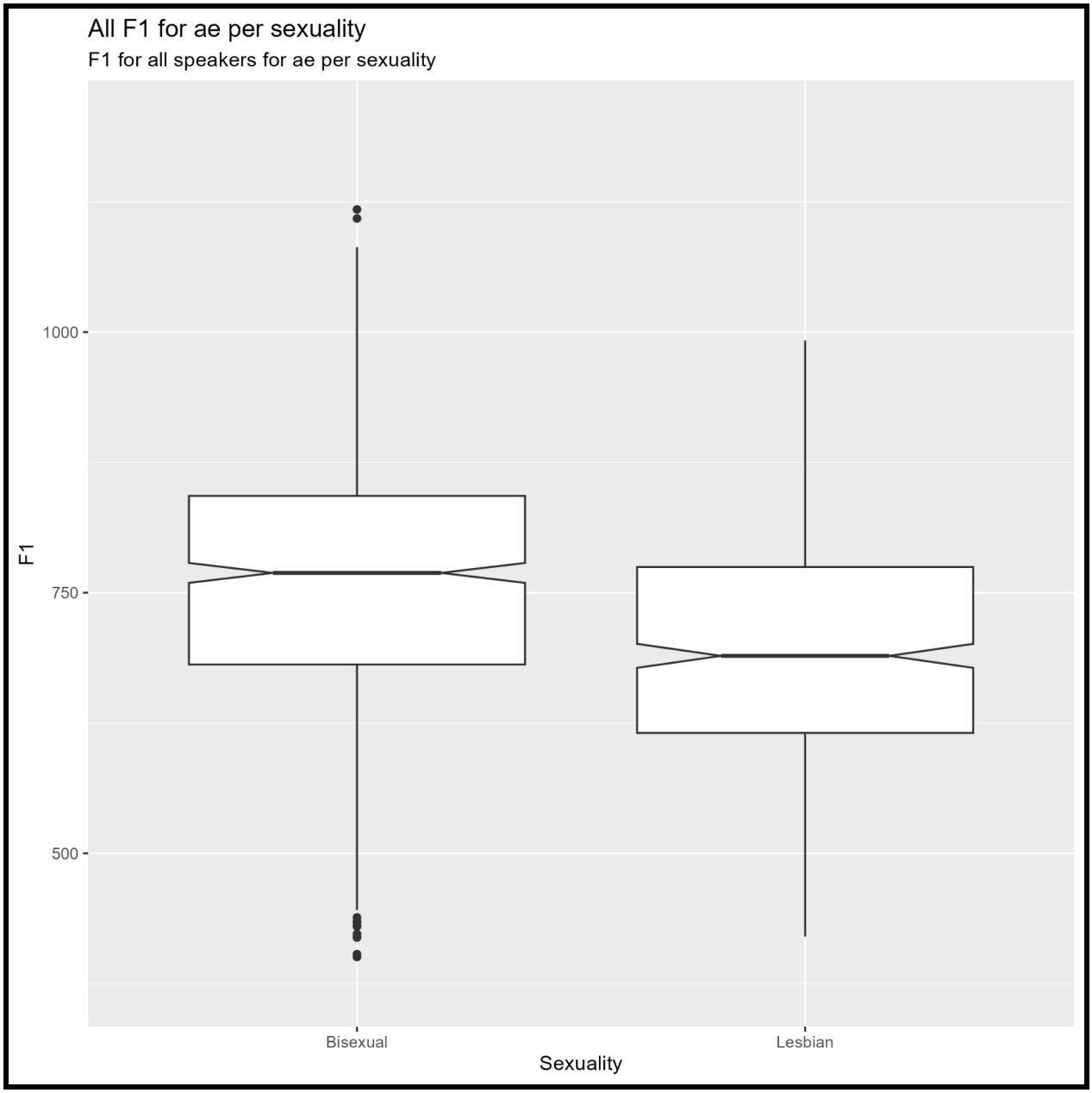
Plot 5E


All F2 for o per gender presentation
F2 for all speakers for o per gender presentation
1750
1500
1250
1000
750
F2
fem & masc
feminine
Gender Presentation

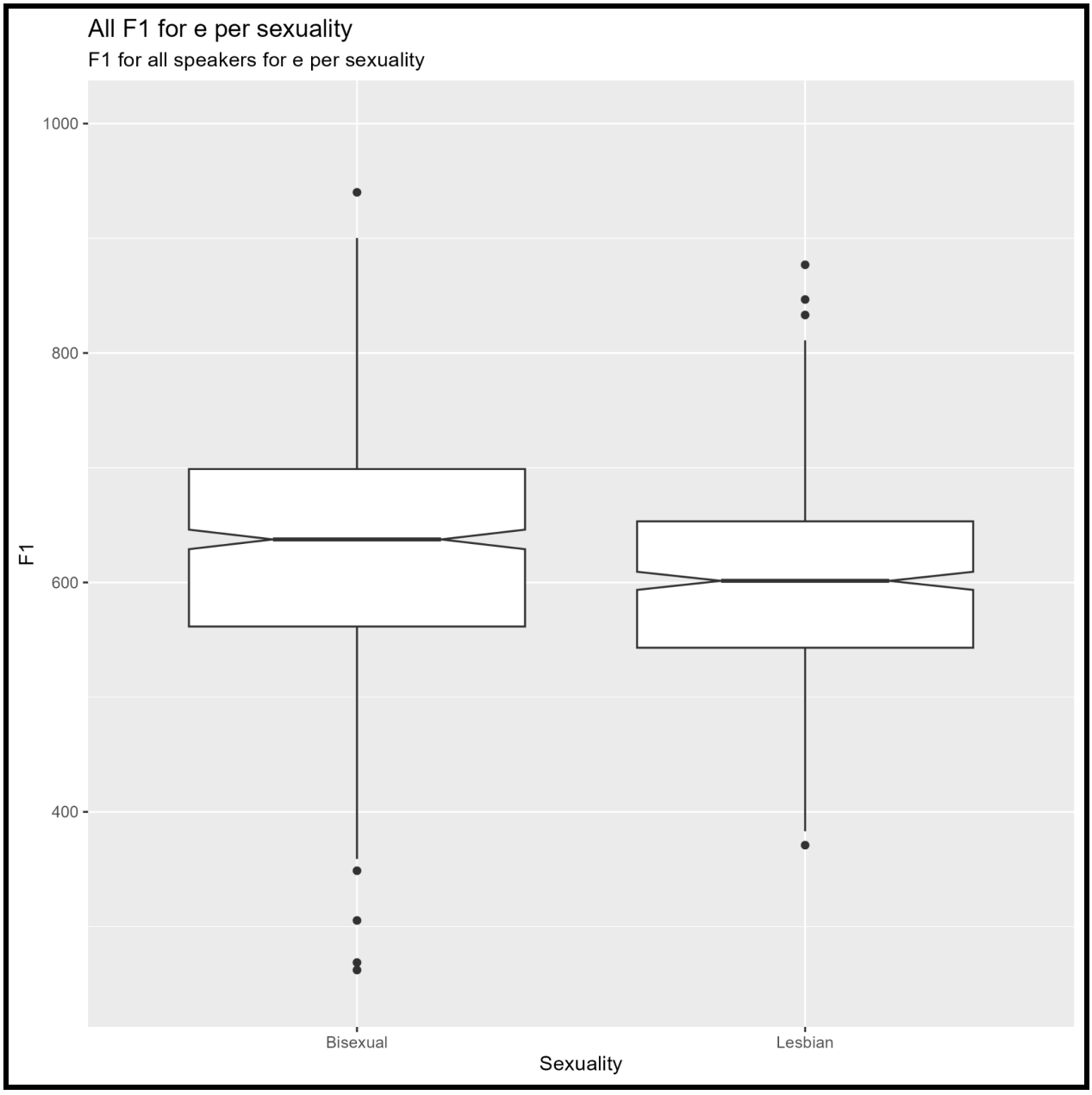
Plot 6A



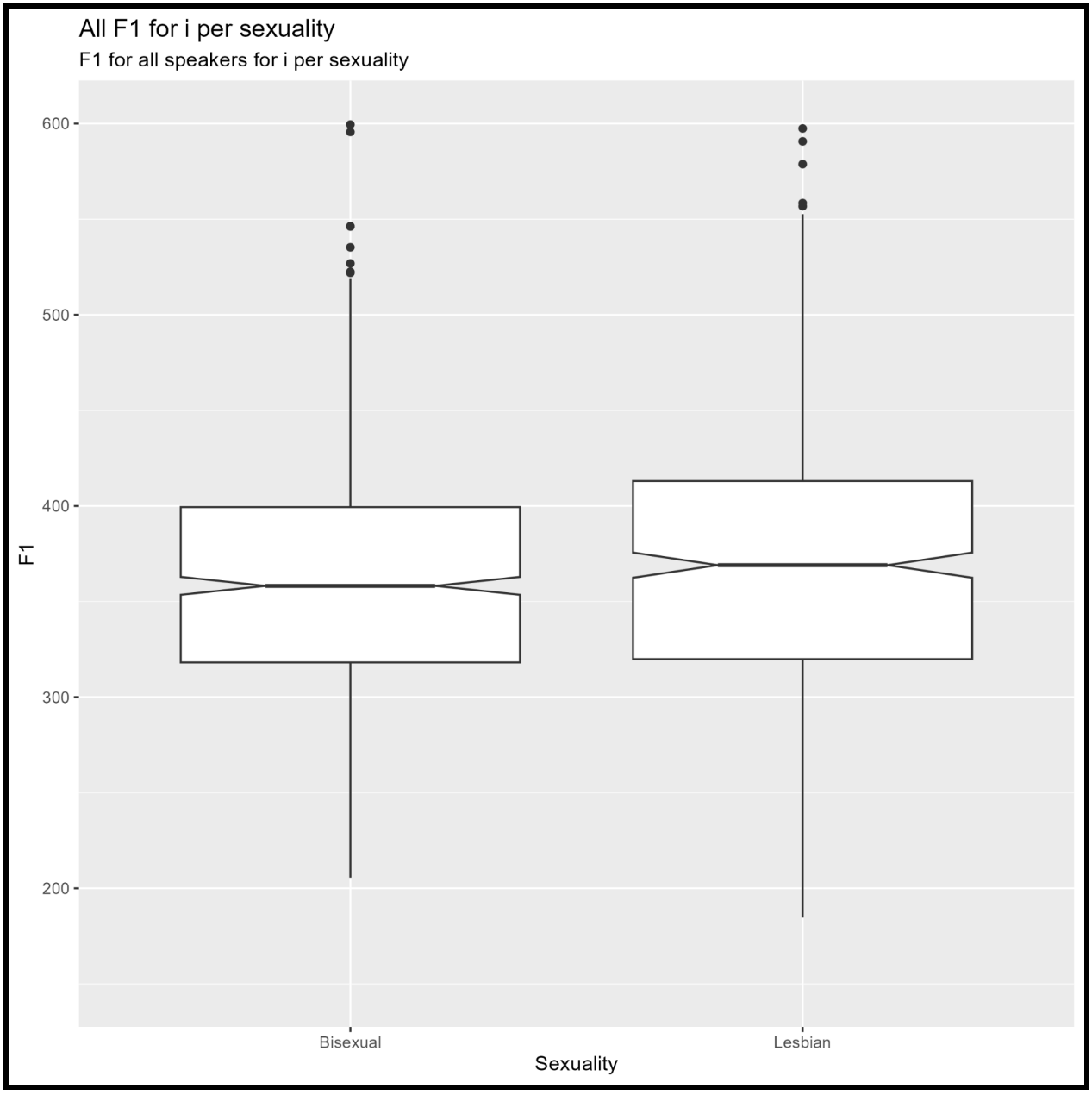
Plot 6B


All F1 for ae per sexuality
F1 for all speakers for ae per sexuality
F1
1000
750
500
Bisexual
Lesbian
Sexuality

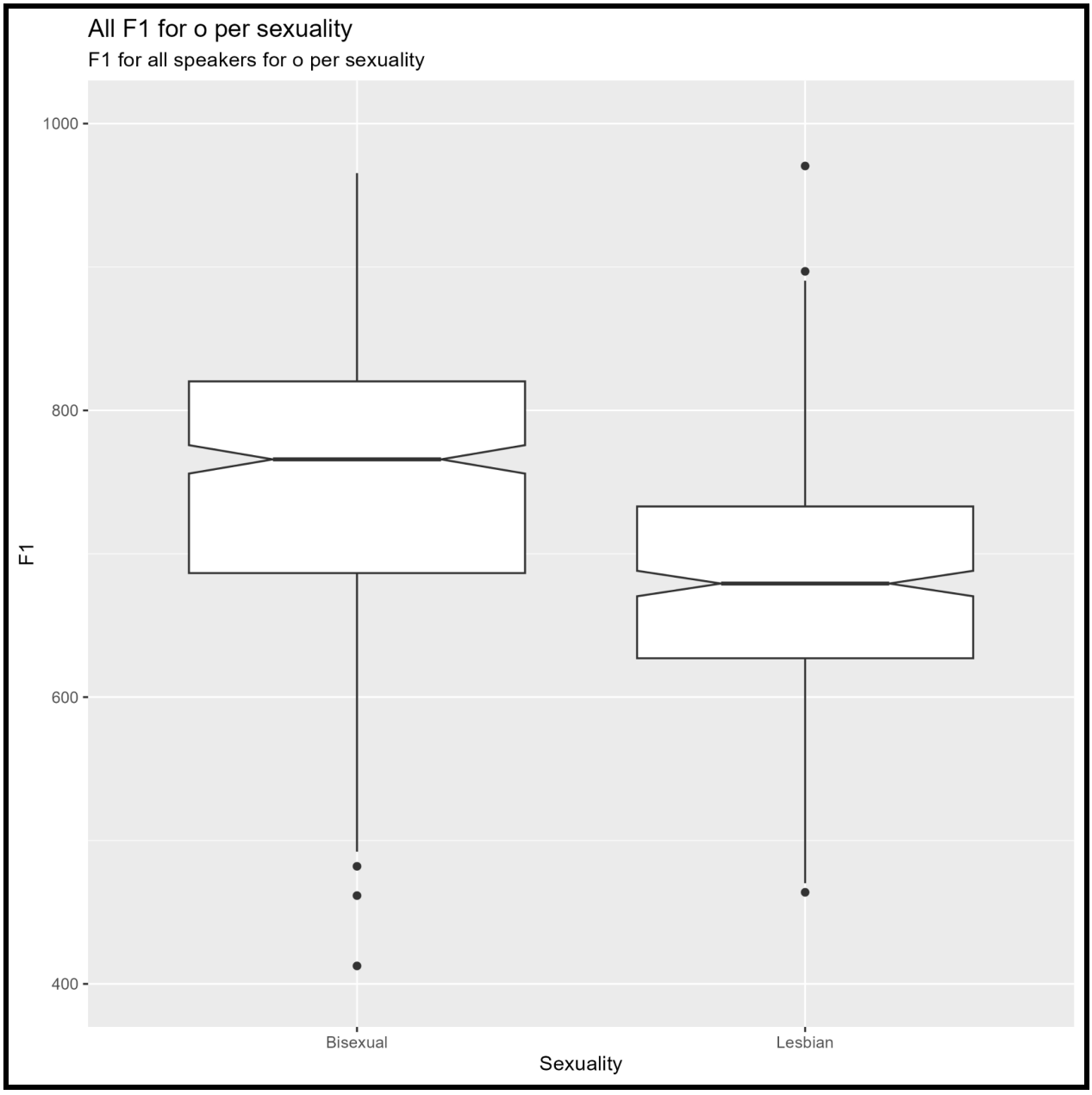
Plot 6C


All F1 for e per sexuality
F1 for all speakers for e per sexuality
1000
800
600
400
F1
Bisexual
Lesbian
Sexuality

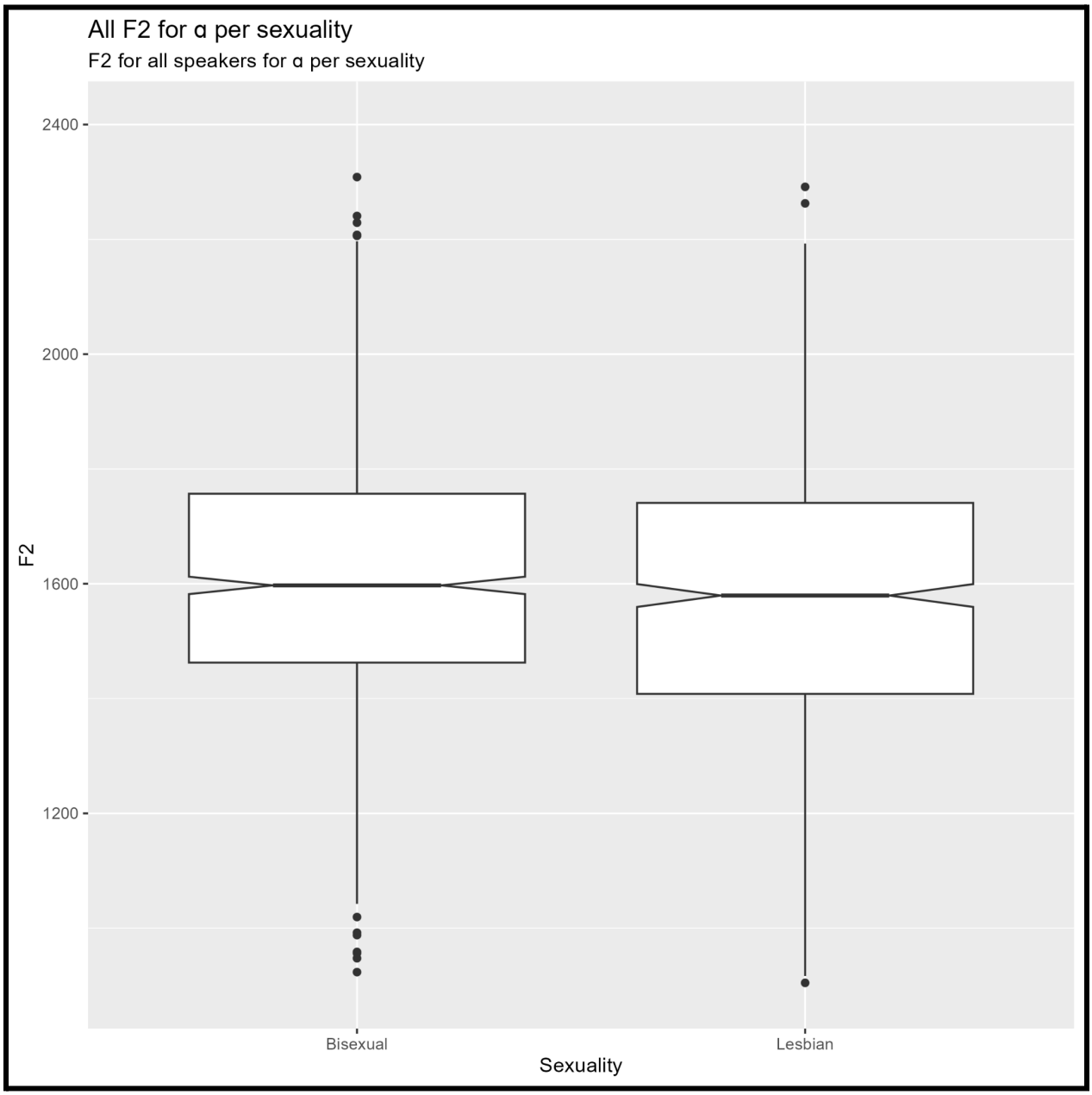
Plot 6D


All F1 for i per sexuality
F1 for all speakers for i per sexuality
600
500
400
300
200
F1
Bisexual
Lesbian
Sexuality

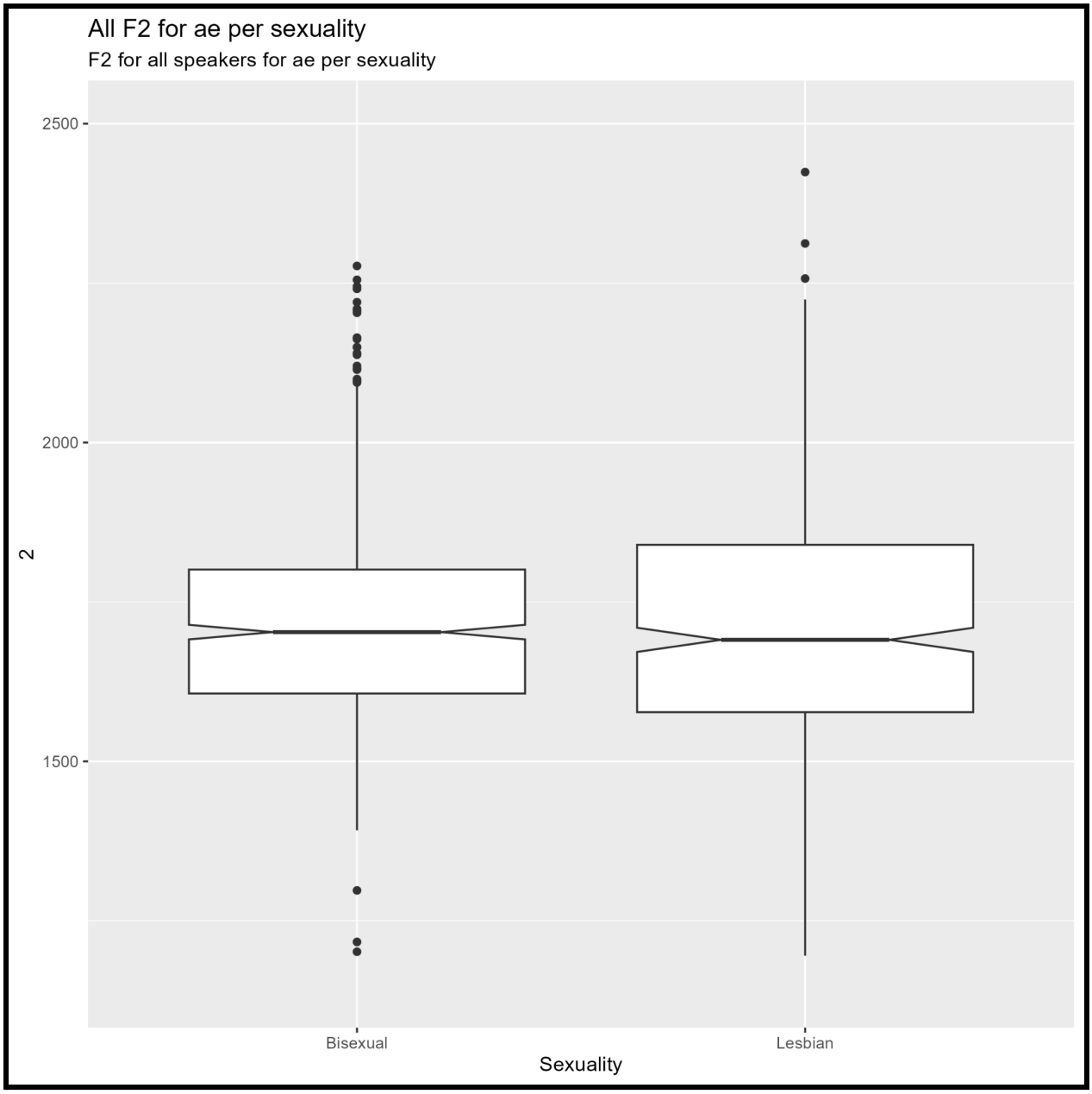
Plot 6E


All F1 for o per sexuality
F1 for all speakers for o per sexuality
1000
800
600
400
F1
Bisexual
Lesbian
Sexuality

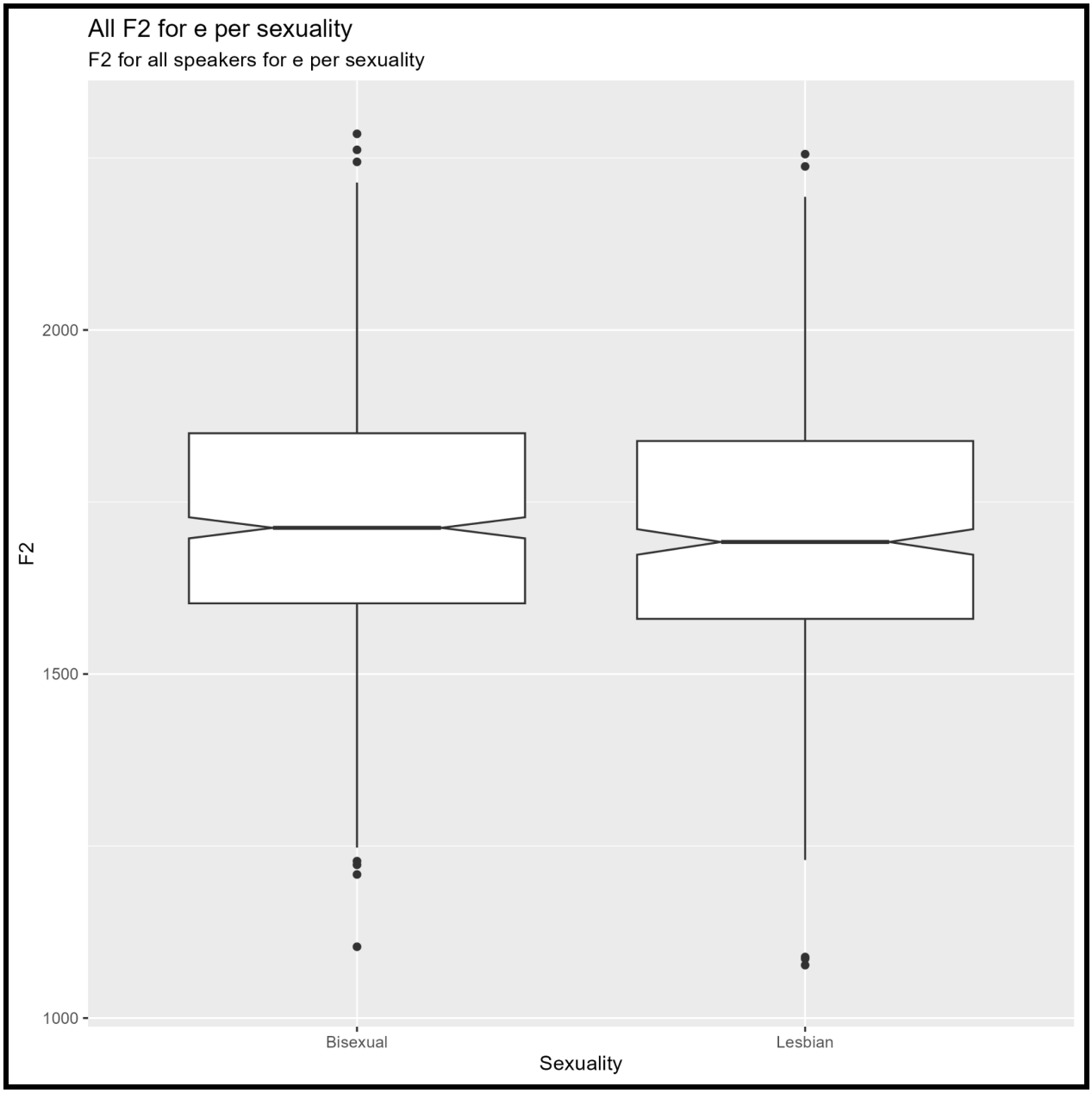
Plot 7A


All F2 for ɑ per sexuality
F2 for all speakers for ɑ per sexuality
2400
2000
1600
1200
F2
Bisexual
Lesbian
Sexuality

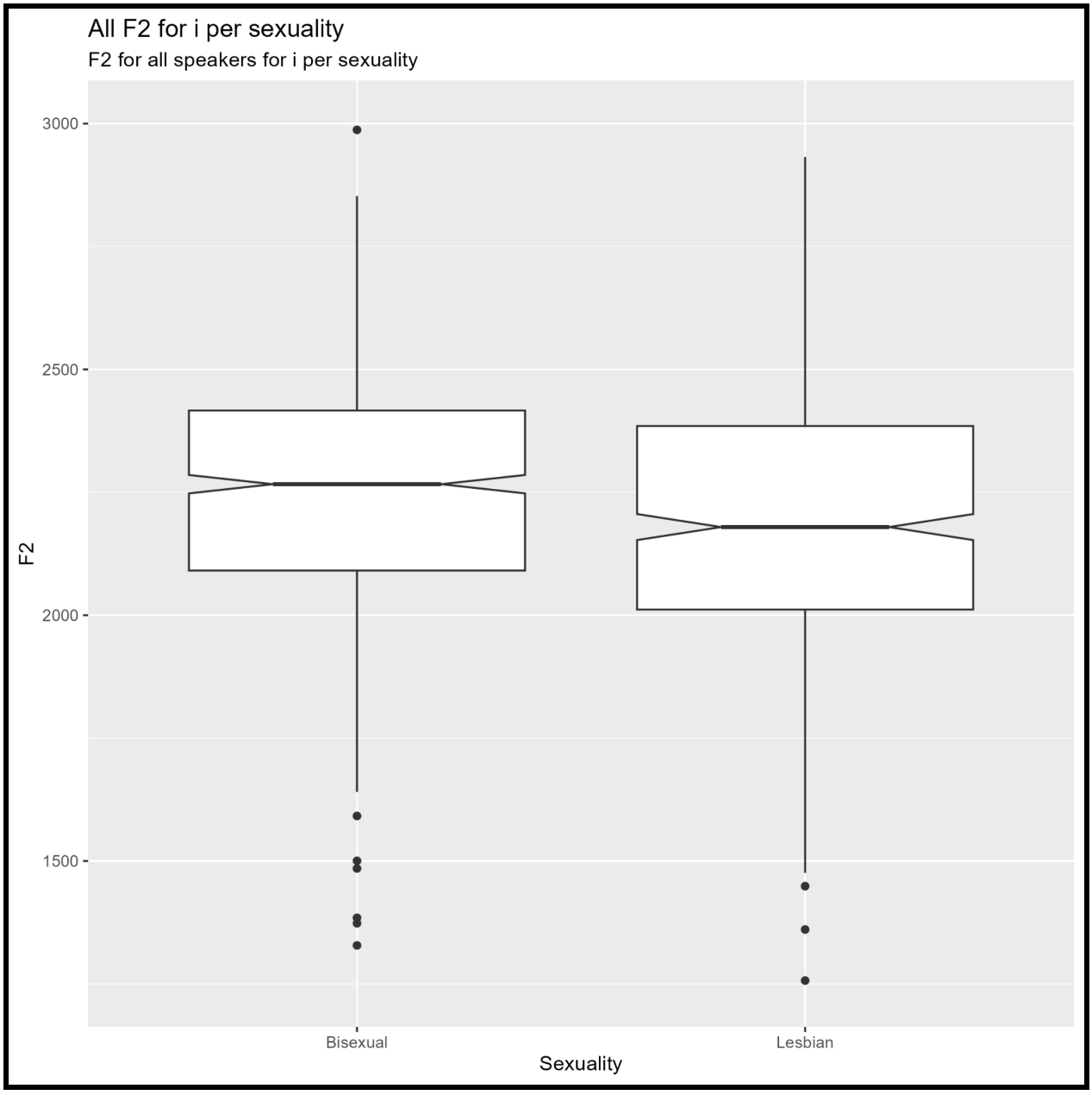
Plot 7B



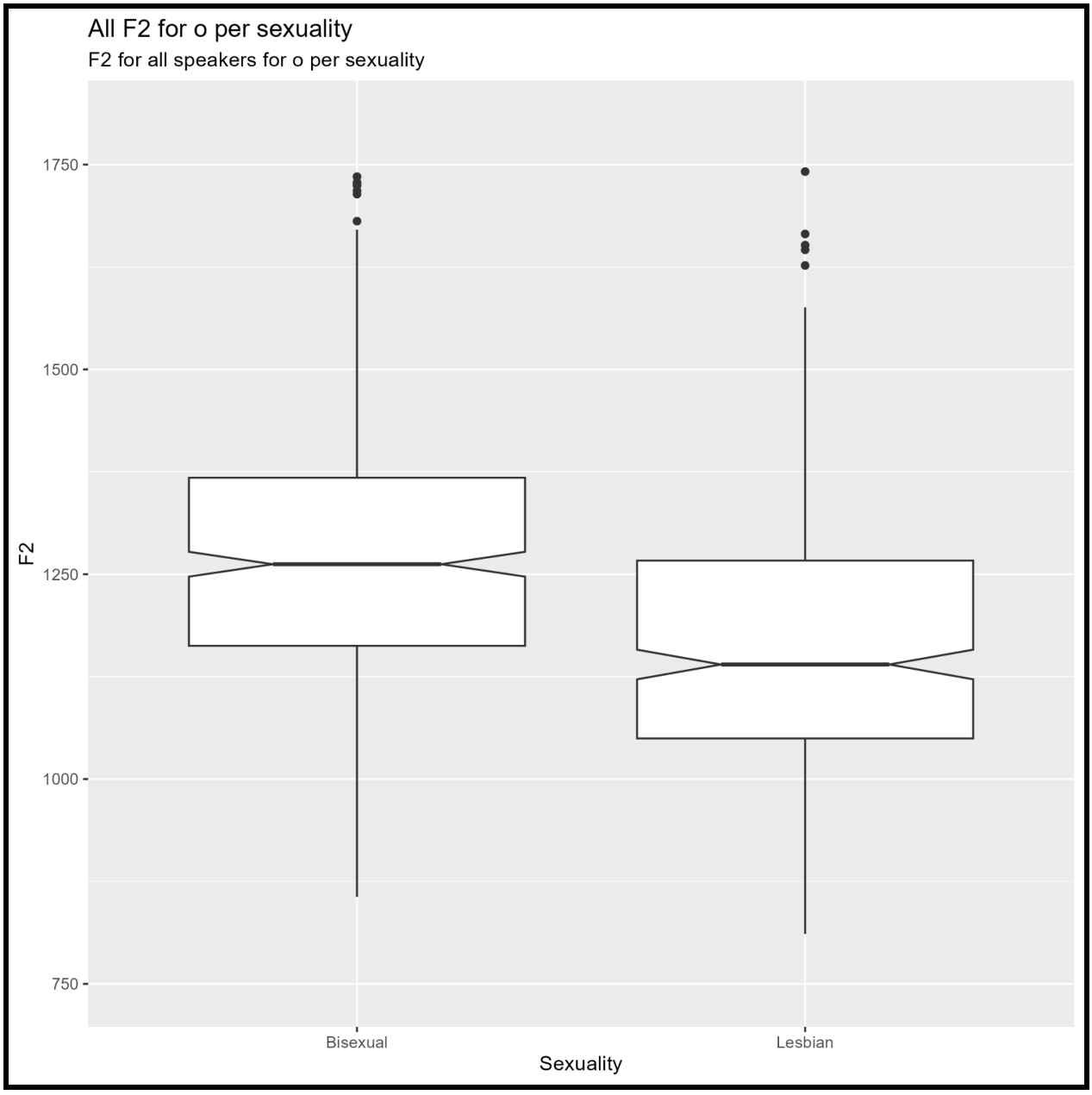
Plot 7C


All F2 for e per sexuality
F2 for all speakers for e per sexuality
2000
1500
1000
F2
Bisexual
Lesbian
Sexuality

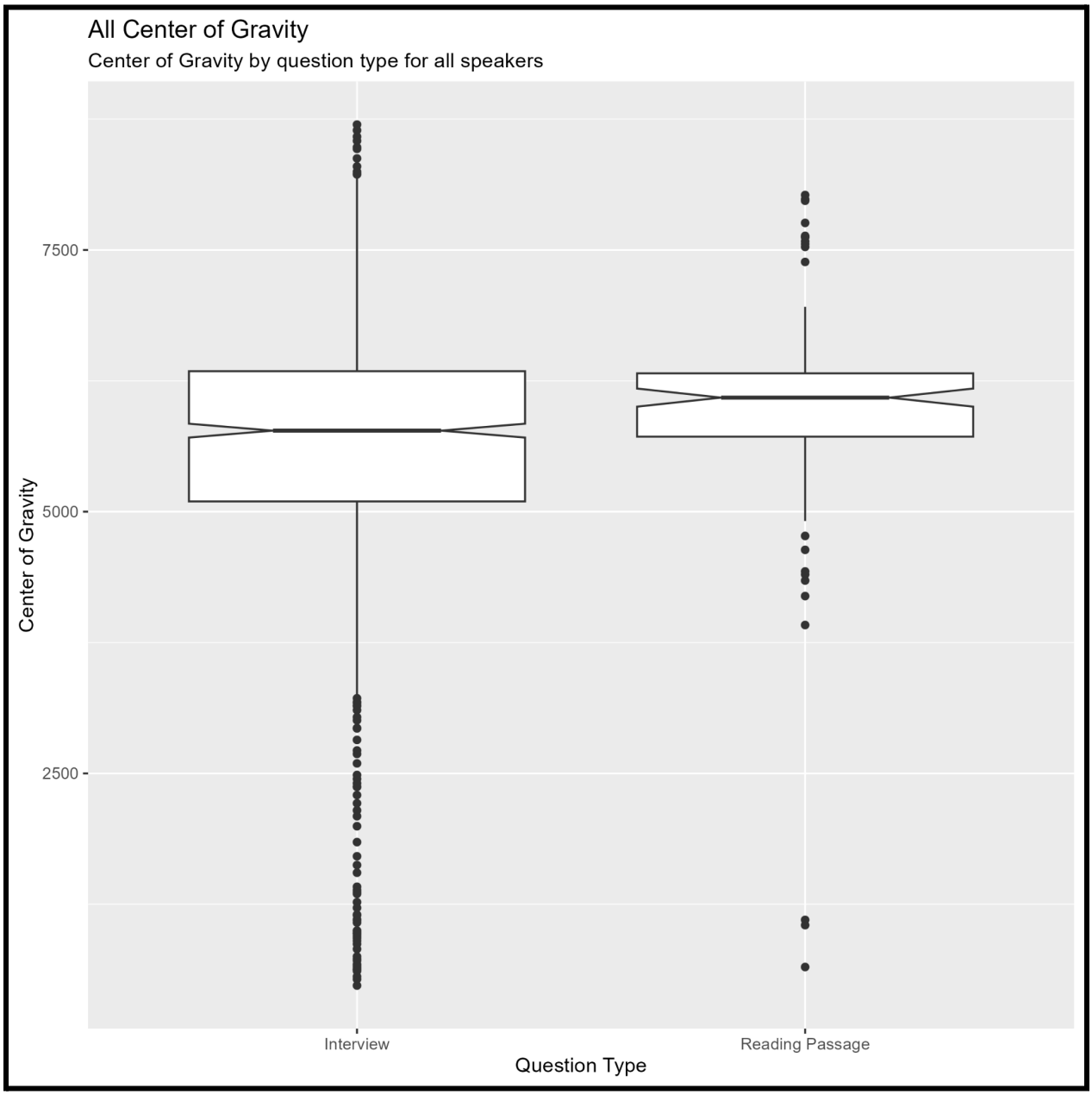
Plot 7D


All F2 for i per sexuality
F2 for all speakers for i per sexuality
3000
2500
2000
1500
F2
Bisexual
Lesbian
Sexuality

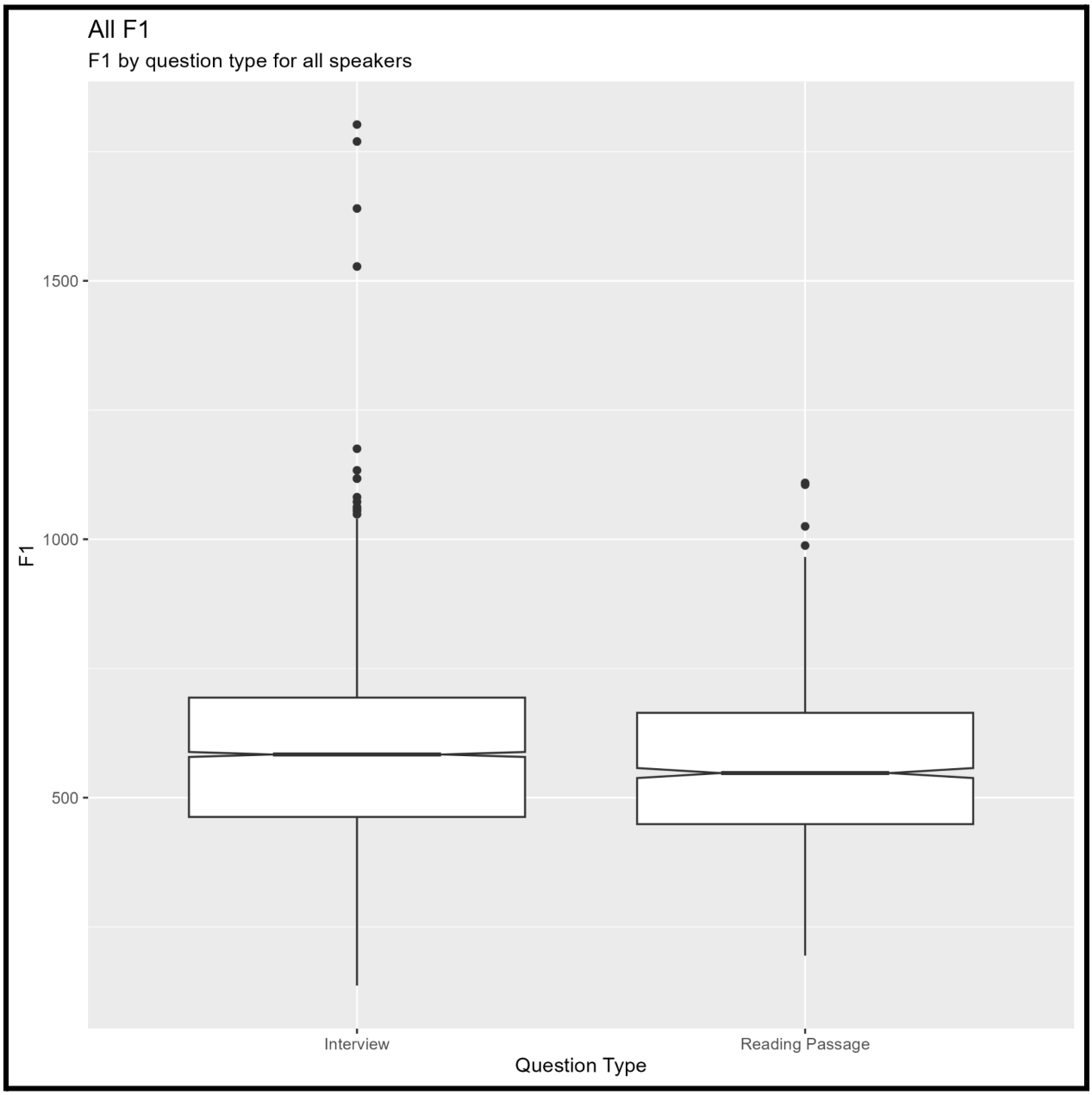
Plot 7E


All F2 for o per sexuality
F2 for all speakers for o per sexuality
1750
1500
1250
1000
750
F2
Bisexual
Lesbian
Sexuality

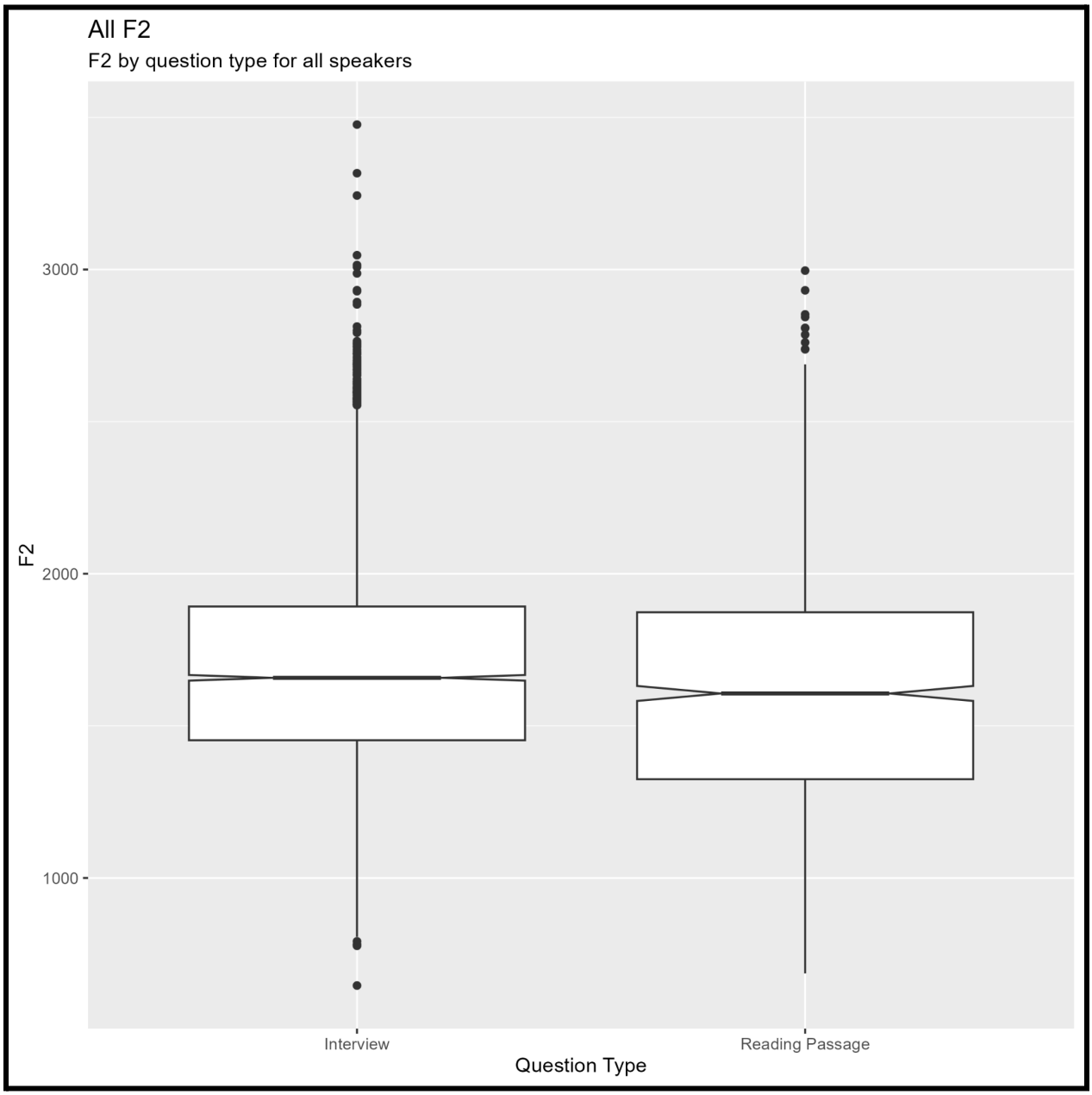
Plot 8A


All Center of Gravity
Center of Gravity by question type for all speakers
7500
5000
2500
Center of Gravity
Interview
Reading Passage
Question Type

Plot 8B


All F1
F1 by question type for all speakers
1500
1000
500
F1
Interview
Reading Passage
Question Type

Plot 8C


All F2
F2 by question type for all speakers
F2
3000
2000
1000
Interview
Reading Passage
Question Type