PHONETIC CONVERGENCE OF GENDER NONCONFORMING INDIVIDUALS*

Anissa Baird and Jessamyn Schertz
University of Toronto and University of Toronto Mississauga

1. Introduction

1.1 Background

Language is a crucial component to the construction of social identities. For the construction of one's gender expression or gender performance (Butler 1990) specifically, one's voice is "an enormously important aspect [...], particularly for those who are transitioning from one gender role, identity, or presentation to another" (Zimman 2018: 2). For this reason, one area of particular interest for those in queer linguistics is a form of style-shifting (Bell 1984) known as phonetic convergence, defined as "an increase in the similarity of acoustic-phonetic form between talkers" (Pardo 2013: 559). This linguistic phenomenon has a rich body of work on a wide spectrum of acoustic-phonetic variables such as F0 (e.g., Babel and Bulatov 2012), VOT (e.g., Nielsen 2011), and vowel space (e.g., Pardo et al. 2017), to name only a few. While there is a plentiful amount of phonetic convergence studies regarding binary gender's affect on this linguistic phenomenon (e.g., Namy et al. 2002, Pardo 2006, Pardo et al. 2013), they have had mixed and/or inconclusive findings (see Coles-Harris 2017 for a review).

However, the same cannot be said with regard to gender nonconformity. For example, Parnell-Mooney (2019) looks at how gender nonconforming (GNC) individuals use phonetic convergence of /s/-peak frequency (the loudest frequency of an /s/-sound) to establish their GNC identities. For cisgender individuals, women typically have higher /s/-peak frequencies than men (e.g., Levon and Holmes-Elliott 2013). However, working with six GNC individuals in the Central Belt of Scotland via sociolinguistic interviews, Parnell-Mooney (2019) makes two key findings: (i) the general trend that assigned female at birth (AFAB) transmasculine individuals have an average /s/-peak frequency that is even lower than that of cisgender male speakers and (ii) that these individuals, when speaking to an individual of their assigned sex at birth (ASAB), change their /s/-peak frequency to further themselves from the "perceived standard" of their ASAB and avoid misgendering. For example, an AFAB GNC individual, when speaking to a cisgender woman, would phonetically diverge and lower their /s/-peak frequency, away from higher frequencies associated with cisgender women and towards lower frequencies associated with cisgender men.

This trend is not only replicated in other studies on phonetic convergence of GNC individuals (e.g., Rechsteiner 2023, Rechsteiner and Sneller 2023), but it is also shown in

*This research was supported in part by NSERC grant #RGPIN 2020 05841 to Jessamyn Schertz.

Actes du congrès annuel de l'Association canadienne de linguistique 2023.
Proceedings of the 2023 annual conference of the Canadian Linguistic Association.

other types of style-shifting as well. As an example, Gratton (2016) looks at *-in/-ing* variation (e.g., *singin'* vs. *singing*) via sociolinguistic interviews with two nonbinary individuals (one of each binary ASAB) and finds, when in public, not strictly 2SLGBTQIA+-friendly spaces, there is significantly more frequent use of the variation less associated with one's ASAB.[1] For instance, an assigned male at birth (AMAB) nonbinary speaker will start using more of the *-ing* variant than *-in*, which is what is expected of cisgender female individuals (e.g., Trudgill 1974).

In all of the aforementioned studies regarding gender nonconformity, it is hypothesized that these style-shifts are a subconscious way in which GNC individuals establish their GNC identities in an attempt to avoid being misgendered. However, most of these studies include a modest number of participants [N=6 for Parnell-Mooney 2019, N=15 for Rechsteiner and Sneller 2023, and N=2 for Gratton 2016; Rechsteiner's (2023) study, which was published after we had completed data collection, has a larger total of 45 participants in one experiment, albeit split evenly across three experiment conditions into smaller subgroups]. As such, the primary motivation of our study is to look at phonetic convergence (or, in this case, divergence) with a larger group of GNC individuals in order to establish this trend on a grander scale.

1.2 Targets of phonetic convergence

Selecting which acoustic-phonetic variables would be used is based on two key criteria. First, the variables selected should be phonetic cues of gender nonconformity already established in the literature. Second, these phonetic cues should have previous evidence of being targets of phonetic convergence in cisgender groups using "asocial" paradigms. An example of such a paradigm is a word shadowing task, where one repeats words that they hear, which involves little-to-no "social engagement" with another individual. The reasoning behind this second criterion is due to it showing a history of the selected phonetic variables being sensitive to phonetic convergence, even in not-so-social situations. This is important to keep in mind, as convergence can surface differently, depending on the level of social engagement of the paradigm used (Pardo et al. 2018).

Following these criteria, the first variable chosen is /s/-Center of Gravity (/s/-COG), defined as the average of frequencies of an /s/-sound, weighted by intensity. With regard to gender nonconformity, Hazenberg (2012) looks at individuals in the Ottawa area and finds the following: (i) the general trend that women, as a whole, have a higher /s/-COG than that of men and (ii) that cisgender women and cisgender men define the high and low extremes of /s/-COG, respectively, with all women groups (including trans women) having a higher /s/-COG than all men groups (including trans men). Furthermore, studies exist that show evidence of this variable being prone to phonetic convergence in cisgender populations using asocial paradigms (e.g., Long 2018).

The second and final variable chosen is F0. This is a heavily studied variable, both for binary gender (e.g., Whiteside 2001, Simpson 2009) and nonconforming gender (e.g., Zimman 2017, Schmid and Bradley 2019, Brown and Pillot-Loiseau 2022). Many studies

[1] Two-spirit, Lesbian, Gay, Bisexual, Transgender, Queer/Questioning, Intersex, Aromatic/Asexual, etc.

that focus on F0 and binary gender, including those mentioned above, attribute gender differences to biological sex-based differences, with cisgender women having higher F0 and cisgender men having lower F0. However, regardless of any biologically-based factors at play, F0 is shown to be able to be significantly altered by trans individuals to the point that they are perceived as a cisgender individual of their gender identity over the phone (Adler et al. 2006). Most importantly for the purposes of our study, this can be done without the use of surgery or hormone therapy. As such, no matter any biological factors that can affect F0, this variable can be further manipulated to significant degrees for socially-based purposes. Further still, as with /s/-COG, F0 also shows evidence of phonetic convergence in cisgender groups using asocial paradigms (e.g., Babel and Bulatov 2012).

1.3 Research questions and hypotheses

Our research questions are as follows: (1) *Do GNC individuals, on a larger scale, show phonetic convergence of /s/-COG and/or F0?* (2) *Do GNC individuals, on a larger scale, show phonetic divergence of /s/-COG and/or F0, dependant on the perceived gender identity of the individual they are speaking with?* Based on the above literature, our hypotheses are as shown below:

1. There will be at least some evidence of phonetic convergence of /s/-COG and/or F0.

2. There will be at least some phonetic divergence of one or both variables, when listening to an individual who is (perceivably) of one's same ASAB.

2. Method

2.1 Participants

In total, 26 AFAB adults from North America who both identified as GNC and considered English to be their first language were included in the final analysis of this study. Originally, participants were gathered by word of mouth and given a CAN$15.00 Indigo gift card as compensation (N=7). However, participant gathering this way proved to be slower than desired, resulting in the remainder of participants being gathered via Prolific (www.prolific.co) and compensated CAN$15.00 through Prolific's system (N=19). While an additional 19 participants completed the study, 15 were excluded (although still compensated) due to poor audio quality, leaving a final total of 30 participants (26 AFAB, 3 AMAB, and 1 intersex). Due to the low numbers remaining for AMAB and intersex participants, and due to the second hypothesis relying on participant ASAB, only data from the aforementioned 26 AFAB participants were analyzed.

2.2 Materials

The study consisted of a reading task, a shadowing task, and a questionnaire. For the reading task, a total of 50 single-syllable words were selected for participants to read (see Table 1). The *critical* stimuli consisted of 38 of these words, which were selected such that there existed two /s/-initial and two /s/-final words for each North American English vowel possible (with the exception of /ʊ/, for which only one word of each type was selected). Two of the remaining 12 words were *practice* stimuli: words without an /s/-sound that were always the first two words for the participant to read as practice while they acclimated to the task's format. The other 10 words were *filler* stimuli with no /s/-sound, meant to distract participants from the importance of /s/ for this experiment.

Table 1. Full list of word stimuli, by type.

Stimuli Type	Word Stimuli				
Critical	suit	sit	set	sock	suck
	soup	sip	sec	sop	sup
	seek	sake	sack	soap	soot
	seat	safe	sat	soak	
	moose	miss	mess	moss	muss
	goose	kiss	guess	boss	fuss
	geese	mace	mass	gross	wuss
	niece	case	gas	dose	
Practice	moon	cat			
Filler	sheep	shoot	goat	pear	dime
	shop	shape	coat	bear	time

The shadowing task used the same list of words as the reading task. In this case, the words were given via voice stimuli, divided into two blocks of two voices each: AFAB and AMAB. To create these voices, the word list was first recorded by one cisgender AFAB speaker and one cisgender AMAB speaker. Both speakers were (at the time of recording) 62-year-old native English speakers who had lived in the Greater Toronto Area their entire lives. The AFAB recording had all sibilant frequencies scaled up, while the AMAB recording had all sibilant frequencies scaled down. The motivation behind this decision was to create a more extreme difference for /s/-COG measurements between the AFAB and AMAB stimuli, in order to make the difference between these two voices more perceptible and maximize the possibility of finding convergence and/or divergence. Finally, both sets of recordings then had second copies created by having their non-/s/-segments pitch-shifted in opposite directions (down for the AFAB copy and up for the AMAB copy). For this decision, the intention was to create more perceptively "gender ambiguous" voices, to see if it might affect the degree of convergence and/or divergence. Figure 1 below shows a visualization of these manipulations.

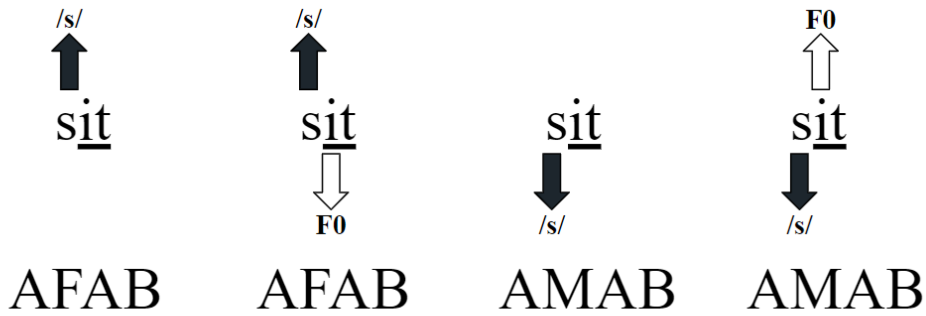

Figure 1. Visual demonstration of the manipulation of stimuli voices, including sibilant scaling (the black arrows) and non-/s/-segment pitch-shifting (the white arrows).

Finally, a questionnaire was created for all participants to complete. The first portion of this questionnaire collected some basic, relevant demographic information such as gender identity, ASAB, month and year of birth, birthplace, current location (only as specific as the city/town level), understood languages, and racial/ethnic identity. Additionally, inspired by the Genderbread Person, participants were asked to judge their gender expression, based on a set of two sliders (see Figure 2).[2] This was done in an attempt to get a quantifiable measure of gender expression for quantitative analysis while still allowing for individuality and nuance. Finally, participants were asked to answer a series of Likert scale questions about their experiences as a GNC individual. These questions ranged from asking about if participants have other family that are GNC to asking about how connected to the 2SLGBTQIA+ community participants feel as a whole. Questions were phrased such that "Strongly Agree" meant a more positive experience, meaning higher total scores also meant a more positive experience. This was done based on Gratton's (2016) work, the implication being that GNC individuals might be more prone to divergence if they have had more negative experiences from others due to their being GNC.

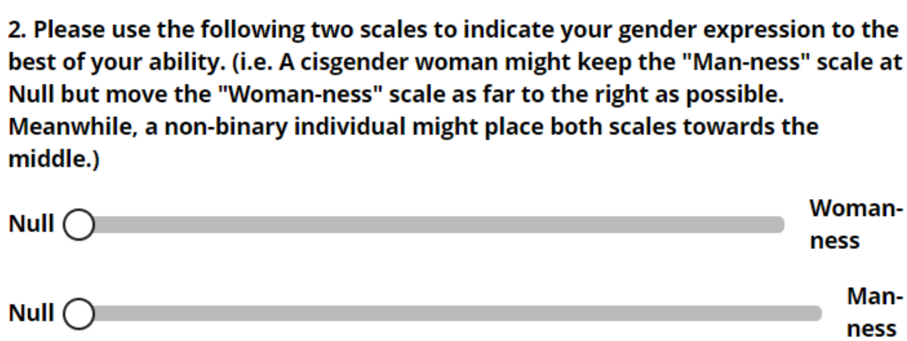

Figure 2. Snapshot of a question from the questionnaire given to participants asking them to judge their personal gender expression via a pair of sliders.

[2] The author, Sam Killermann, has uncopyrighted this material. It can be found here: https://www.itspronouncedmetrosexual.com/genderbread-person

2.3 Procedure

The experiment itself was hosted on the Gorilla Experiment Builder (www.gorilla.sc) and conducted online. The reasoning behind the decision to run the study online was to get a larger group of total participants while still respecting the COVID-19 restrictions in place at the time. When beginning the experiment, participants were first led to a consent form and equipment test. The equipment test ensured that participants had a chance to double-check that both their audio input and output was working correctly, before the experiment began proper, in order to help limit the number of exclusions due to recording/audio quality issues.

Following this, participants underwent the reading task to attain their baseline measurements. Participants were given the words in text form on their screens and were explicitly given three seconds to read each word, with a visible timer given, to avoid cut-offs (see Figure 3 for an example). The words were semi-randomized, per participant, such that the two *practice* words were always shown first (in either order) before all other words, with *critical* and *filler* words being randomized together.

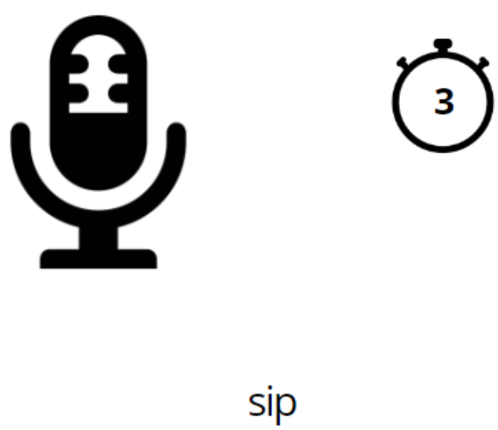

Figure 3. Example of a participant's screen during the reading task, including the word to be read and a visual timer.

After this was the shadowing task. Participants heard audio stimuli of the same list of words from the reading task and were asked to repeat each word. These stimuli were presented in two blocks: one for the AFAB stimuli and one for the AMAB stimuli. As previously mentioned, both blocks of stimuli had two "voices" differing in F0, meaning participants repeated the original list of words a total of four times, two times in each block, for this task. At the beginning of each block, participants were introduced to the voices (described as "[AFAB/AMAB] speakers"), who were visually displayed as either a purple or orange silhouette, the latter always being indicative of the pitch-shifted voice. Participants heard both voices greet them, before the task began proper. For each word, participants were shown the coloured silhouette of the stimuli speaker and were now given five seconds to repeat each word, with a visible timer given, in order to account for any possible auditory perception delays (see Figure 4). Blocks were randomized per participant such that either the AFAB block or the AMAB block could appear first. The stimuli within both blocks were also randomized, per participant, in the same way. First, the two voices in each block were completely randomized together. Second, as with the reading task, words were semi randomized such that all starter words (from both voices in each block) appeared first before all other words (also from both voices in each block).

Figure 4. Example of a participant's screen during the shadowing task, including a visual timer and the silhouette of the speaker "giving" the current word to be shadowed (which would be either purple or orange, depending on the stimuli voice being heard).

Following this, participants were asked to fill out the questionnaire before being given a full debriefing as to the study and its purpose. While not a task of the experiment, the explicit mention of a debriefing is of particular importance, in this case. The participant pool, by nature of the study, is from a marginalized group. As such, it is felt necessary to emphasize that participants were not only told the purpose of the study after their completion of it, but they were also reminded that they could choose to withdraw from the study at any time, even after completing the experiment.

2.4 Acoustic data measurements

To measure /s/-COG and F0, boundaries were created by hand around the /s/-sound and vowel of each word via TextGrids in Praat (Boersma 2001). The /s/-sounds were selected such that no voicing from the following or previous vowel was included in the selection, based on the spectrogram of the word. The vowel was selected such that its initial/final boundary (depending on whether said vowel was part of an /s/-initial or /s/-final sound) was shared with the /s/-sound's final/initial boundary. The remaining boundary for each vowel sound was made such that it did not overlap with the following or previous non-/s/-sound, if one existed. All boundaries were placed at zero-crossings.

The measurements for /s/-COG were completed via a Praat script. The /s/-segment was Hann-band-filtered to remove frequencies from 0-500 Hz to avoid capturing any residual voicing frequency before calculating COG across the full /s/-segment. F0 was measured automatically at 20 ms after vowel onset, using speaker-specific pitch ranges that had been chosen via visual inspection prior to measurement. After this, F0 was checked and manually corrected, if necessary, for all tokens. Tokens that were not analyzable due to poor recording quality, noise, or undefined F0 (usually due to creaky voice) were omitted: 1.44% of /s/-COG shadowing task measurements and 10.35% of F0 shadowing task measurements.

3. Results

Phonetic convergence of /s/-COG is first analyzed by raw measurement (see Figure 5). Specifically, each token included is a single, raw measurement from a GNC AFAB speaker producing an /s/-sound of some word stimuli in the shadowing task. Overall, the distribution of values looks remarkably similar across all voice stimuli "types." This hints that the stimuli voice being shadowed does not play a meaningful role in any potential convergence and/or divergence exhibited by these speakers in the shadowing task.

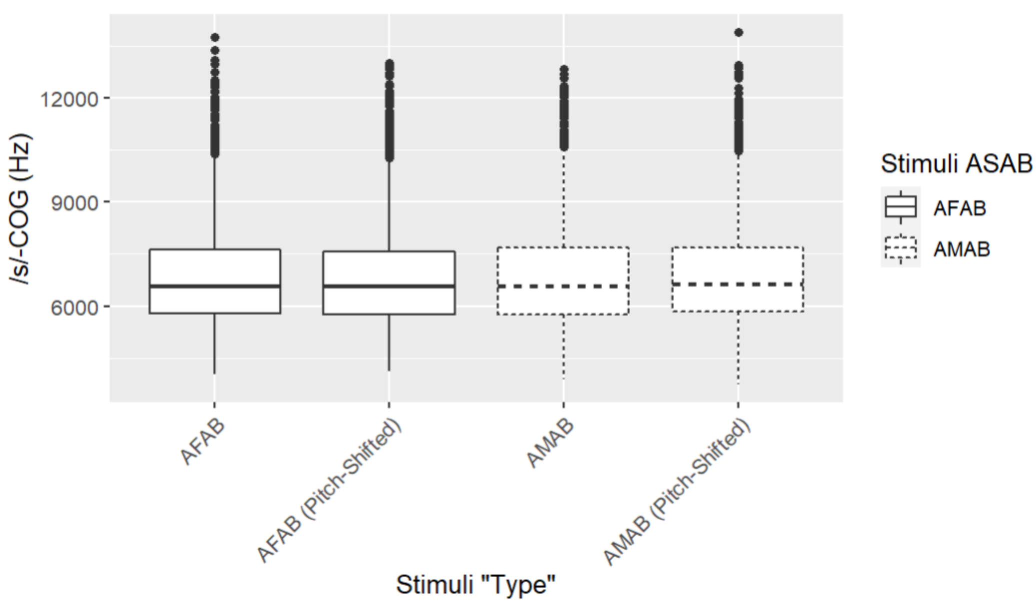

Figure 5. Raw /s/-COG measurements (N=3895) from the word shadowing task, split by stimuli voice being shadowed.

To attain a clearer picture, /s/-COG is also analyzed using Difference in Distance (DID) measurements (Pardo et al. 2017). DID is a measurement that can be used to measure phonetic convergence, by comparing measurements from the reading task to the shadowing task, in terms of similarity to the stimuli measurements. To calculate DID, the following steps are made: (i) determine the difference between a participant reading task measurement and the corresponding stimuli voice measurement (reading task measurement - stimuli measurement), (ii) determine the difference between the corresponding participant shadowing task measurement and the corresponding stimuli voice measurement (shadowing task measurement - stimuli measurement), and (iii) calculate the difference between the absolute values of the two previous calculations (|i| - |ii|). Simply put, the more positive a DID measurement, the more convergence is present. The more negative a DID measurement, the more divergence is present. A DID measurement of 0 means absolutely no change between the reading task and shadowing task measurements for a given /s/-sound in a given word spoken by a given speaker.

To investigate phonetic convergence and/or divergence using DID measurements, each raw /s/-COG shadowing task measurement used requires a corresponding reading task measurement from the same participant. As such, using DID measurements results in a drop of 1.21% of tokens used in the raw measurement analysis. When split across the four stimuli voices, as shown in Figure 6, the DID measurements for /s/-COG present a similar picture as the raw measurements for /s/-COG: there is an overall lack of change across all voices being shadowed. For all stimuli voices, however, the medians of these data sets are also crucially very close to a measurement of 0. Thus, the DID measurements for /s/-COG not only confirm a lack of change, across stimuli voice, but also indicate a lack of convergence or divergence from this group of participants.

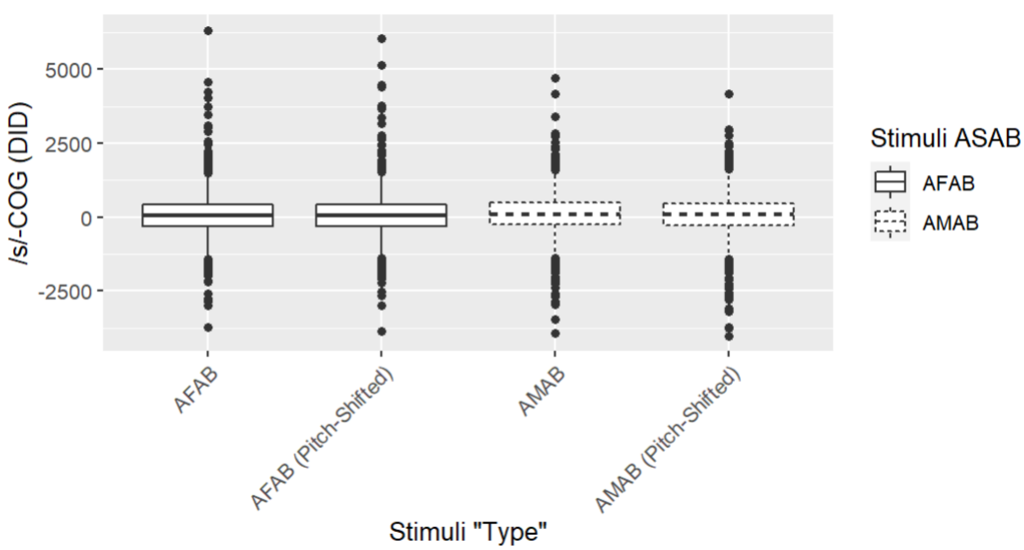

Figure 6. /s/-COG DID measurements (N=3848) from the word shadowing task, split by stimuli voice being shadowed.

As with /s/-COG, F0 is also first analyzed by raw measurement (see Figure 7). For this variable, each token is a vowel measurement in a given word spoken by a given speaker in the shadowing task. When split across stimuli voice, the F0 raw measurements show a pattern of no change similar to that of /s/-COG. Further still, DID measurements for F0 (which involve a 7.11% decrease in tokens analyzed, from the number used in the raw measurement analysis) also show the same trend as the /s/-COG DID measurements: there is not only no visual difference across stimuli voice, but the medians of these measurement groups all hover around 0, indicating no evidence of convergence or divergence across stimuli voice (see Figure 8).

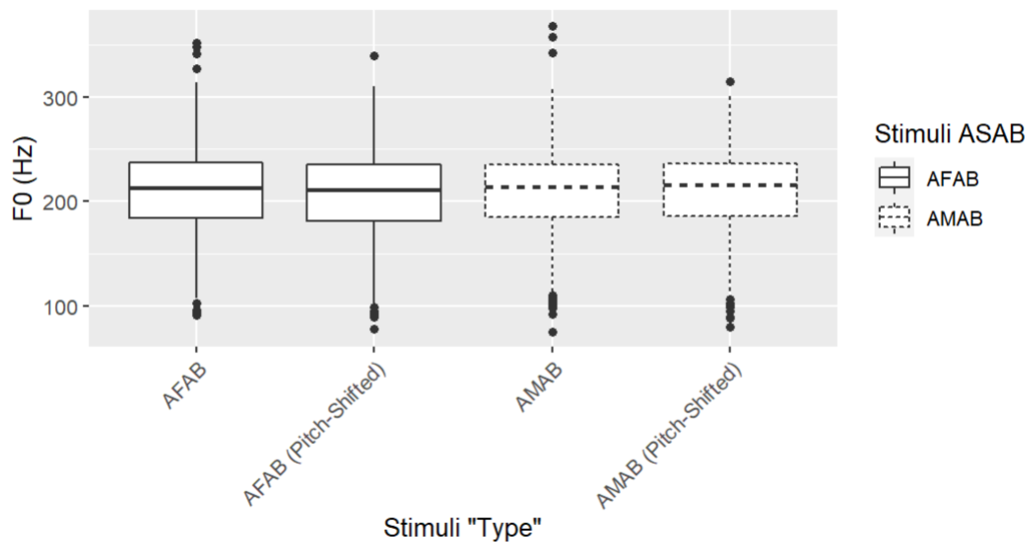

Figure 7. Raw F0 measurements (N=3543) from the word shadowing task, split by stimuli voice being shadowed.

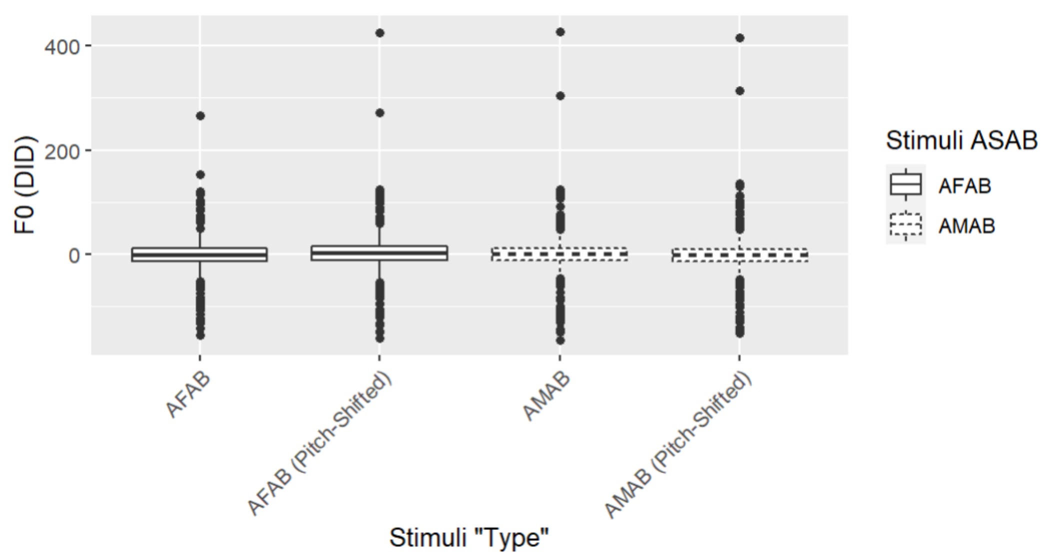

Figure 8. F0 DID measurements (N=3291) from the word shadowing task, split by stimuli voice being shadowed.

4. Discussion

The results above strongly suggest that participants in this study do not show either convergence or divergence in either /s/-COG or F0 when shadowing different voices. This supports neither of the hypotheses of this study, which also goes against most of the literature on this subject. There is no evidence of phonetic divergence for AFAB speakers, when shadowing either AFAB voice, despite numerous studies showing a trend of GNC individuals phonetically diverging from cisgender individuals of their same ASAB (e.g., Parnell-Mooney 2019, Rechsteiner 2023, Rechsteiner and Sneller 2023). Furthermore, the lack of phonetic convergence and/or divergence as a whole contradicts the findings of other studies looking at these phonetic variables with cisgender populations (e.g., Long 2018 for /s/-COG and Babel and Bulatov 2012 for F0).

We posit three potential reasonings to explain these null effects. First, it is possible these effects only seem null on a group level. That is to say, there could be individual differences, with some participants converging while others are diverging when shadowing the same voice(s), causing the graphing of data at a group level to average out the same, across each stimuli voice. However, investigation into the data still shows these same null effects at an individual level, making this possibility highly unlikely.

Second, it is possible that neurodivergence masks the trend of phonetic divergence expected of GNC individuals. Recent studies find that individuals on the autism spectrum are less likely to converge (Hogstrom et al. 2022). It is also known that individuals on the autism spectrum are more likely to be GNC (Warrier et al. 2020). However, none of the mentioned previous studies on phonetic convergence of GNC individuals (this study included) attain any information about participant neurodivergence. As such, it is possible (however improbable) that the present study ended up with a group of mostly neurodivergent GNC individuals, while studies such as Parnell-Mooney (2019) did not. Regrettably, without such information for the bulk of participants in this study, the influence that neurodivergence may have had on these results will never be known. However, it does bring up an interesting question for future investigation.

Finally, the third possibility, and the most likely, is simply that the "social threshold" for activating phonetic convergence and/or divergence was simply not met via the tasks implemented. There are a few considerations that make this reasoning the most plausible: (i) although Babel and Bulatov (2012) find convergence of F0 with a shadowing task, the effect is only slight, (ii) all but one study involving GNC phonetic convergence/divergence mentioned (Rechsteiner and Sneller 2023) includes interviews as a source of data collection in some capacity, and (iii) to reiterate, Pardo et al. (2018) finds that the "social engagement" level of a task influences the degree of phonetic convergence. Therefore, it is possible that a more "social" task, such as a sociolinguistic interview, may have provided different, non-null results, with this group of speakers.

With these potential reasonings in mind, there are a number of ways that future iterations of this study can expand. One way is to add questions about neurodivergent identity to the questionnaire in order to investigate how the intersection of neurodivergence and gender nonconformity (or lack thereof, for some GNC individuals) may affect phonetic convergence and/or divergence. Further still, research on phonetic convergence in GNC individuals could also benefit from looking at other aspects of queer intersectionality as well, including those involving sexuality, ethnicity/race, class, and more (see Gray and Cooke 2018 for further discussion). Another way to expand on the present study is to include a more social task, such as a sociolinguistic interview or a map task (Pardo et al. 2018), in addition to a word shadowing task, in order to determine the effect that the "social engagement" of a task has, if any. Additionally, no individual differences were found, even when considering the quantified measurement for gender expression. While self-labels are still the most important identifier to consider for gender identity, the scales used to attempt to quantify gender expression in this study could benefit from additional nuance, in order to fully encompass gender expression, gender identity, and the importance of gender as a concept for an individual as a whole. Finally, given that the known trend of GNC individuals using phonetic divergence is affected by ASAB (in that GNC individuals diverge based on the ASAB of the cisgender individual they are speaking with), running a study with more participants, especially more AMAB and intersex GNC participants, as well as with a cisgender control group, would provide a more complete picture on phonetic convergence in GNC individuals.

5. Conclusion

The present study sought to replicate previous work showing phonetic divergence in GNC individuals using a larger scale, focusing on /s/-COG and F0 from 26 AFAB GNC individuals in reading and word shadowing tasks. Participants repeated stimuli produced by model speakers of different genders (cisgender female and cisgender male) with natural and manipulated F0 in order to test whether productions shifted either toward or away from the model speakers. In contrast to previous results, there was no evidence of either convergence or divergence. It is posited that the most likely reason for these results is the lack of social interaction of the paradigm used. Although further investigation is needed to confirm this hypothesis, if found to be true, it highlights the importance of this linguistic phenomenon's connection to social context and gender performance, particularly for GNC individuals.

References

Adler, Richard K., and John van Borsel. 2006. Female-to-male considerations. In *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide*, ed. Richard K. Adler, Sandy Hirsch, and Michelle Mordaunt, 139–167. San Diego, CA: Plural Publishing.

Babel, Molly, and Dasha Bulatov. 2012. The Role of Fundamental Frequency in Phonetic Accommodation. *Language and Speech* 55(2): 231–248.

Bell, Allan. 1984. Language style as audience design. *Language in Society* 13(2): 145–204.

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot. Int.* 5(9): 341-345.

Brown, LeAnn, & Claire Pillot-Loiseau. 2022. Bright Voice Quality and Fundamental Frequency Variation in Non-binary Speakers. *Journal of voice: official journal of the Voice Foundation*.

Butler, Judith. 1990. *Gender trouble : feminism and the subversion of identity.* London: Routledge.

Coles-Harris, Evan H. 2017. Perspectives on the motivations for phonetic convergence. *Language and Linguistics Compass* 11(12): e12268.

Gratton, Chantal. 2016. Resisting the gender binary: The use of (ING) in the construction of non-binary transgender identities. *University of Pennsylvania Working Papers in Linguistics* 22(2): 51–60.

Gray , John, and Melanie Cooke. 2018. Intersectionality, language and queer lives. *Gender and Language* 12(4): 401–415.

Hazenberg, Evan N. L. 2012. Language and identity practice: a sociolinguistic study of gender in Ottawa, Ontario. Master's thesis, Memorial University of Newfoundland.

Hogstrom, Anders, Rachel Theodore, Allison Canfield, Brian Castelluccio, Joshua Green, Christina Irvine, and Inge-Marie Eigsti. 2022. Social and sensory influences on linguistic alignment: Phonetic convergence in autism spectrum disorder. *Evolutionary Linguistic Theory* 4(1): 102–128.

Levon, Erez, and Sophie Holmes-Elliott. 2013. East end boys and west end girls:/s/-fronting in Southeast England. *University of Pennsylvania Working Papers in Linguistics* 19(2): 111–120.

Long, Yanyu. 2018. The perception-production link in sibilant convergence. *The Journal of the Acoustical Society of America* 144(3): 1722–1722.

Namy, Laura L., Lynne C. Nygaard, and Denise Sauerteig. 2002. Gender Differences in Vocal Accommodation: *The Role of Perception. Journal of Language and Social Psychology* 21(4): 422–432.

Nielsen, Kuniko. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39(2): 132–142.

Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119(4): 2382–2393.

Pardo, Jennifer S. 2013. Measuring phonetic convergence in speech production. *Frontiers in Psychology* 4: 559–559.

Pardo, Jennifer S., Kelly Jordan, Rolliene Mallari, Caitlin Scanlon, and Eva Lewandowski. 2013. Phonetic convergence in shadowed speech: The relation between acoustic and perceptual measures. *Journal of Memory and Language* 69(3): 183–195.

Pardo, Jennifer S., Adelya Urmanche, Sherilyn Wilman, and Jaclyn Wiener. 2017. Phonetic convergence across multiple measures and model talkers. *Attention, Perception & Psychophysics* 79(2): 637–659.

Pardo, Jennifer S., Adelya Urmanche, Sherilyn Wilman, Jaclyn Wiener, Nicholas Mason, Keagan Francis, and Melanie Ward. 2018. A comparison of phonetic convergence in conversational interaction and speech shadowing. *Journal of Phonetics* 69: 1–11.

Parnell-Mooney, James. 2019. Tran/s/gender: assessing the effects of the social construction of gender on speech: a focus on transgender /s/ realisations. Master's thesis, University of Glasgow.

Rechsteiner, Jack. 2023. Sociophonetic Variation and Imitation in Nonbinary Speakers. Master's thesis, Michigan State University.

Rechsteiner, Jack, and Betsy Sneller. 2023. The impact of social information on VOT shadowing by nonbinary speakers. In *Proceedings of the 20th International Congress of Phonetic Sciences* 2023, ed. Radek Skarnitzl and Jan Volín, 3720–3723. Guarant International.

Schmid, Maxwell, and Evan Bradley. 2019. Vocal pitch and intonation characteristics of those who are gender non-binary. In *Proceedings of the 19th International Conference of Phonetic Sciences* 2019, ed. Sasha Calhoun, Paola Escudero, Marija Tabain and Paul Warren, 2685–2689. Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Simpson, Adrian P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass* 3(2): 621–40.

Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich.* Cambridge: Cambridge University Press.

Warrier, Varun, David M. Greenberg, Elizabeth Weir, Clara Buckingham, Paula Smith, Meng-Chuan Lai, Carrie Allison, and Simon Baron-Cohen. 2020. Elevated rates of autism, other neurodevelopmental and psychiatric diagnoses, and autistic traits in transgender and gender-diverse individuals. *Nature Communications* 11(1): 3959–3959.

Whiteside, Sandra P. (2001). Sex-specific fundamental and formant frequency patterns in a cross-sectional study. *Journal of the Acoustical Society of America* 110(1): 464–78.

Zimman, Lal. 2017. Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/. *Language in Society* 46(3): 339–370.

Zimman, Lal. 2018. Transgender voices: Insights on identity, embodiment, and the gender of the voice: Transgender voices. *Language and Linguistics Compass* 12(8): e12284.