

Crosslinguistic perceptions of /s/ among English, French, and German listeners

ZAC BOYD

The University of Edinburgh

JOSEF FRUEHWALD

The University of Kentucky

LAUREN HALL-LEW

The University of Edinburgh

ABSTRACT

This study reports the results of a crosslinguistic matched guise test examining /s/ and pitch variation in judgments of sexual orientation and nonnormative masculinity among English, French, and German listeners. Listeners responded to /s/ and pitch manipulations in native and other language stimuli (English, French, German, and Estonian). All listener groups rate higher pitch guises as more gay- and effeminate-sounding than lower pitch guises. However, only English listeners hear [s+] guises as more gay- and effeminate-sounding than [s] or [s−] guises for all stimuli languages. French and German listeners do not hear [s+] guises as more gay- or effeminate-sounding in any stimulus language, despite this feature's presence in native speech production. English listener results show evidence of *indexical transfer*, when indexical knowledge is applied to the perception of unknown languages. French and German listener results show how the enregistered status of /s/ variation affects perception, despite crosslinguistic similarities in production.

It is now widely accepted within sociolinguistics that phonetic variation can convey social meaning. Phonetic variation can often index things such as speaker gender and sexual orientation, and these social meanings are indexed regardless of the speaker's actual identity (some straight men 'sound gay,' etc.). Interestingly, some of these cues appear to be crosslinguistic (e.g., sibilant variation, especially within /s/).

Theoretically, the social meaning of a given variable is specifically tied to phonetic properties of *language* and not an abstraction of the acoustic sounds. [Figure 1](#) illustrates this point. The acoustics of /s/ are comparable to that of white noise, like the hiss of a tire. Variation between different acoustic features of a tire hiss might indicate differences in the size of the leak; while some leaks might acoustically resemble a threatening hiss (see Eckert, 2017:1198), others

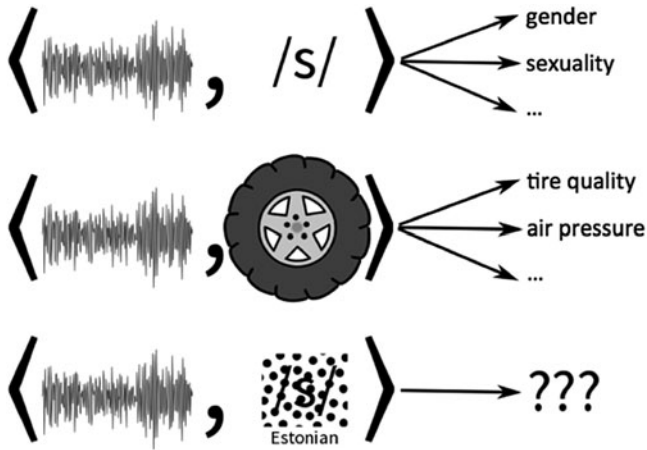


FIGURE 1. The same acoustic signal in three different contexts.

might more acoustically approximate variation in /s/. However, in a context where the /s/-like sound corresponds to the phoneme /s/, variation may be associated with cues to a speaker's gender identity (Levon & Holmes-Elliott, 2013) or sexual orientation (Munson & Babel, 2007), or both (Zimman, 2013). Indexicality is situated in and not an inherent property of the cue itself (Pharao, Maegaard, Møller, & Kristiansen, 2014; Stuart-Smith, 2007), and there is some suggestion that nonspeech sounds can accrue indexicality if contextualized as speech (Rogers & Smyth, 2003; Szakay, 2006). This line of reasoning extends, for example, to the line between indexicality and sound symbolism (e.g., Eckert, 2012, 2017; Levon, Maegaard, & Pharao, 2017), which remains an active area of inquiry. One approach is to test for shared social meanings for the "same" sound across different languages.

In this paper we examine two things. First is the extent to which acoustic variation in the voiceless sibilant /s/ (specifically, fronted versus non-fronted articulations) indexes the same things across different languages (e.g., Bekker & Levon, 2017; Pharao et al., 2014; Zimman, 2013). Second is the extent to which variation in /s/ carries social meaning when embedded within speech signals that are plausibly not parsable to a listener (i.e., languages unknown or unfamiliar to listeners; Figure 1). Based on the results of a matched guise experiment, we obtain language attitude data on /s/ variation as perceived by English, French, and German listeners. We look at attitudes towards /s/ in a listener's native language and expand this paradigm to include crosslinguistic perceptions of /s/ where English, French, and German native language listeners rate speech samples in English, French, or German (whichever is not their native language), as well as Estonian.

Our results reveal two findings. First, we see indexes apparent in production that are not recognized in perception. Despite previous findings showing /s/ to vary

according to sexual orientation in French and German speech production (Boyd, 2018a, 2018b), French and German listeners here show no difference in their rating of fronted /s/ versus non-fronted /s/, in their native language or in any other language, with regard to judgments of perceived sexual orientation. These results add to the growing evidence for patterned mismatches between production and perception of sociophonetic variation, which we analyze with respect to enregisterment (Agha, 2003, 2005) and “meaning potential” (Eckert, 2016). Second, we observe other listeners applying their indexical knowledge even to unknown languages. For English listeners, fronted /s/ stimuli are rated as more gay- and effeminate-sounding, not only in English, but across all language stimuli regardless of the listener’s knowledge of the others. Here we propose a model of *indexical transfer*. Both findings point toward a need for a cognitive model of indexical representation.

BACKGROUND

Enregisterment

The process of enregisterment occurs when “distinct forms of speech come to be socially recognized as indexical of speaker attributes by a population of language users... [which are] *reflexive models of language* that are disseminated along identifiable trajectories in social space through communicative processes” (Agha, 2005:38, original emphasis; see also 2003). Enregisterment is often discussed in terms of regional dialect or social class variation (e.g., RP [Agha, 2003]; “Pittsburghese” [Johnstone, Andrus, & Danielson, 2006]; “Sheffield” or “northern” [Beal, 2009]). Here, we suggest that “sounding gay” (Gaudio, 1994) is also an enregistered style of English, albeit one whose meanings are variable across different social contexts. In this way, “sounding gay” is enregistered in the way that “netspeak” or “chatspeak” are enregistered, being not necessarily “geographically bounded” (Squires, 2010:461). While “chatspeak” is enregistered as a result of “standard language ideology and deterministic views of technology” (Squires, 2010:457), “sounding gay” is enregistered as a result of hegemonic masculinity (Zimman, 2013), and is a style we might call gender bounded, being unavailable in the same way to speakers constrained by hegemonic femininity. For men, heterosexuality and “sounding straight” rely on hegemonic masculinity and its distinction from “subordinate masculinities” (Talbot, 2010:169), for example, gay masculinities, because it “must negate them” (Talbot, 2010:169). A classic example of how this may manifest in practice is seen in Cameron’s (1997) study of five fraternity brothers, all of whom subscribe to aspects of hegemonic masculinity, where being “gay” has little to do with sexual desire but is instead regarded as being “insufficiently” masculine.

A large body of research shows that sibilant variation has become enregistered with a gay male speaking style in multiple languages. Not only is its indexicality evidenced in patterns of interspeaker variation,¹ but its ability to index “gay” has

reached a level of metadiscourse, at least in English. The concept of a “gay /s/” features prevalently in English language pop culture, often dubbed colloquially (albeit incorrectly) as “the gay lisp.” Examples of this can be seen, for example, in the documentary *Do I Sound Gay?* (Thorpe, 2014).

Cross-Variety Indexicality

Our enregisterment analysis of English and languages other than English is based on a crosslinguistic comparison of indexicality in perception: French listeners of French compared to German listeners of German, etc. At the same time, we also present a crosslinguistic comparison of how listeners perceive variation in languages with which they are not proficient: French listeners of Estonian, etc. There are two relevant areas of research for understanding this latter case: work on attributing social meaning to unfamiliar languages, and work on attributing social meaning to unfamiliar but mutually intelligible varieties.

Previous work has considered the judgments listeners can make when listening to languages they do not know. For example, Eisenstein (1982) found that English learners at different levels of proficiency had the same relative status ranking of English varieties as native English listeners (“Standard English” as higher than both “black English” and “New York English”), but that more proficient listeners more closely resembled native listeners’ ratings (cf., Major, 2007; Vaughn & Bradlow, 2017). As these studies did not control for the specific linguistic variables driving the responses, it might be the case that listeners of nonfamiliar languages use different strategies than native speakers despite similar judgments of social meaning. For example, Clopper and Bradlow (2009) found that non-native listeners of US English, in contrast to native listeners, do not rely on monophthongization in /ɑj/ as a variable distinguishing Southern and non-Southern regional varieties.

However, segment-specific variation is relevant when there is a potential to map segmental variation from one language onto another. Brown and Lambert (1976) found that monolingual English listeners were able to accurately identify the socioeconomic status of Canadian French speakers speaking French and suggested that English respondents base their judgments on features that appear in both languages and happen to correlate with status in similar ways. Moreau, Thiam, Harmegnies, and Huet (2014) showed that “European” listeners with no prior knowledge of Wolof were only slightly less accurate in identifying Wolof speakers’ social status than Senegalese students and suggest that this is due to (unnamed) features of Wolof borrowed from French that carry similar indexical values. Clopper and Bradlow (2009) showed that Mandarin Chinese listeners classifying US English regional dialects attend to variables that are also variables in Mandarin Chinese (fricative voicing and post-vocalic r-lessness). These studies speak to a process of what we describe here as *indexical transfer*: when listeners’ socioindexical knowledge about familiar languages is applied to unfamiliar languages.

Previous work has also examined the perception of social meaning across unfamiliar varieties of the same language. These mostly evidence the differential role of sociolinguistic knowledge. Foulkes, Docherty, Kattab, and Yaeger-Dror (2010) found that gender indexes of /t/ realization in Tyneside English were identified by Tyneside listeners but not by other English listeners. Montgomery and Moore (2018) found that listeners from the Isles of Scilly were better at differentiating Scillonian personae than other English listeners. What we do not know is if Tynesiders and Scillonians classify non-Tynesiders and non-Scillonians into the same social categories, using those variables; are there examples of intralanguage indexical transfer? One thing we do know is that, if a linguistic variable is not regionally specific, then there appears to be no consistent relationship between listener and speaker regional background in the evaluation of a variant's indexical meaning, (e.g., [ING] [Campbell-Kibler, 2007, 2011]). /s/ and pitch are similar variables to (ING), being widespread across English varieties and indexing globally similar but locally distinctive social meanings. Indexical transfer seems likely when an English listener hears this variation in a different English variety, but we also investigate what happens when listening to a different language.

Sexual Orientation, Pitch, and Sibillance

In this study, we present the results of a perception experiment using stimuli that manipulate two phonetic variables that appear to be enregistered with gay male speaking styles across multiple languages: *pitch* and *sibillance*. Our analysis focuses on the crosslinguistic perception of indexicality and /s/ variation, with the use of pitch variation as a comparison variable.

Pitch. Research has shown that women have, on average, higher pitch and utilize a wider pitch range than men (e.g., Titze, 1989; Whiteside, 2001), although the size of the difference across men and women differs crossculturally (e.g., Van Bezooijen, 1995). The relationship between pitch and sexual orientation is less straightforward. The earliest research on pitch, sexuality, and masculinity seemed to suggest that there was no correlation in either production or perception. Gaudio (1994) examined the pitch differences between four gay and four straight men from the San Francisco area and found no significant differences in the pitch properties (neither fundamental frequency nor f0 range) between the two groups. In perception, listeners were consistently able to identify which of the speakers were gay, but these ratings were unrelated to the speakers' f0 values (see Avery & Liss [1996] for a similar study).

Smyth, Jacobs, and Rogers (2003) examined the relationship between the pitch properties of the speech of twenty-five men from Toronto and the perception of their voices as masculine/feminine- and gay/straight-sounding. They found a correlation between pitch and listener judgments of masculinity and femininity, though not sexual orientation. Voices with low mean f0 were rated as "sounding gay," but were unlikely to be rated as "feminine." Voices with higher mean f0 were rated as both feminine-sounding and gay-sounding. These findings indicate

that “sounding gay” and “sounding feminine” may be related but are ultimately distinct concepts (Smyth, Jacobs, & Rogers, 2003:342). Rogers and Smyth (2003) further showed that perceived pitch and intonational variability correlate with perceptions (but not productions) of gayness if, and only if, all segmental information is removed from a speech signal. This may suggest that, while no reliable phonetic differences exist in pitch production between gay and straight men, listeners may associate high pitch and high pitch variability in male voices with femininity, and in some contexts this may be further associated with gayness, but not directly.

While pitch generally does not correlate with sexual orientation, it can be used stylistically to index a gay speech style. In his study of an individual gay speaker, Podesva (2007) showed the utilization of falsetto as a stylistic marker to index a gay identity, specifically what he refers to as this speaker’s (Heath’s) “diva persona.” Zimman’s (2013:1) analysis of trans men’s pitch range and variability furthers the case that these features are not indexing femininity or gayness directly but rather represent “deviation from the hegemonic norm” that give way to these readings.

Sibilance. Like pitch, sibilance is consistently shown to differ according to (binary) gender identification, with women consistently being found to produce /s/ with a higher center of gravity (CoG) than male speakers, and that the difference is not wholly attributable to physiological differences (e.g. Fuchs & Toda, 2010; Stuart-Smith, 2007). The past two decades have seen a wealth of research establishing /s/ variation as an index of a gay male speech style and/or nonnormative masculinity, in English (Campbell-Kibler, 2011; Levon, 2006; Munson, McDonald, DeBoe, & White, 2006; Smyth et al., 2003;) but also other languages (*Danish*: Maegaard & Pharao, 2016; Pharao et al., 2014; *Spanish*: Mack, 2010; Walker, García, Cortés, & Campbell-Kibler, 2014; *Hungarian*: Rácz & Schepáz, 2013).

While most previous work has focused on variation within a language, Bekker and Levon (2017) looked at the perception of /s/-fronting in both Afrikaans and White South African English. A total of 214 native Afrikaans listeners, all with at least a moderate degree of English proficiency, participated in a matched guise test listening to /s/ variation in both Afrikaans and English. For the male guises, the fronted-/s/ stimuli were rated as less masculine and more gay-sounding than the nonfronted variants, regardless of the language heard.

The present study builds on this work by designing a matched guise experiment that draws on the production results from studies on French and German. Hobart (2013) examined /s/ variation in the speech of bilingual French men from Aix-en-Provence, finding that for both French and English, the gay speakers of the study produced /s/ with a higher CoG than the straight speakers. However, as his follow-up study contradicts these findings (Hobart, 2014), Hobart suggested speakers in that study may not accurately represent the gay bilingual population (2014), and/or the sample may overly reflect the fact that not all gay men produce the features of a gay speech style. Russell (2017) examined overtly performative speech of six individuals based in Paris, finding higher /s/ CoG

values and longer sibilant durations when speakers were tasked with “sounding gay” than when tasked with “sounding straight.”

As to German, Guzik (2006) looked at the pitch (and vowel space peripherality) of two speakers, showing that the “less masculine sounding” speaker produced average and maximum fundamental frequency values at a much higher range than the “more masculine sounding speaker,” suggesting pitch as a potential resource for nongender-conforming speech acts in German men. Fuchs and Toda (2010) showed that, though German speakers showed more similarities in palate length than English-speaking counterparts, female speakers of German produced /s/ with a fronter articulation than German men beyond what may be attributed to physiological difference. Kachel, Simpson, and Steffens (2018), based on speech from fifty-four German speakers from Jena, showed gay men to produce higher /s/ Center of Gravity (CoG) than the straight men of the study; however, this result was not significant.

Boyd (2018a) demonstrated that both French and German gay and straight men reliably differ in /s/ production. Speakers were asked if they can tell if a French or German person is gay by how they speak. Only one participant responded “no” (Table 1). When asked what aspects of speech signal gayness, the only consensus was that /s/ (or as mentioned above, the “gay lisp”) is *not* part of a French or German gay speech style. As one speaker put it: “Oh, I’ve heard of [the “gay lisp”] in English, but we definitely don’t have it.” All the other speakers flatly stated that they had never heard of it in either English or their native language. The question we ask here is if the indexical association between fronted /s/ and a gay male speech style might be present in the general French- and German-speaking populations in a more implicit matched guise experiment.

METHODS

The experimental design of the present study draws its inspiration from Levon (2006, 2007) and Pharao et al. (2014), employing a matched-guise technique (Lambert, Hodgson, Gardener, & Fillenbaum, 1960). The audio used in testing comes from read speech of four cis-gender male speakers: one English speaker from Essex (England), one French speaker from Lyon, one German speaker from Düsseldorf, and one Estonian speaker from Püünsi (a village 17km from Tallinn). A sample of each speaker’s English read speech was pretested on scales of Straight/Gay and Masculine/Effeminate (cf., Levon, 2006:61) rated by fifteen lifelong English listeners (Table 2). Speakers were chosen for the pilot study because of their relative similarity in pretesting ratings as compared to all other speakers in the speaker sample (Boyd, 2018a), and their having all been rated overall as relatively Straight and Masculine as compared to those other speakers. Subsequent rating of these speakers’ guises as more Gay or Effeminate can therefore be attributed to the manipulations in the guises rather than strong differences in their unmanipulated speech samples.

TABLE 1. Responses to “Can you tell if someone is gay by how they speak?” by both French L1 and German L1 participants (from Boyd, 2018a)

‘Something’ in Speech	Prosodic Cues	/s/ in English	/s/ in L1
18/19 (95%)	13/19 (68%)	1/19 (5%)	0/19 (0%)

Following the pretest, two audio segments (average 4.5 seconds) were taken from each of the four speakers’ readings of a fairy tale in their native language: *Snow White* (English), *Le Petit Chaperon Rouge* (French), *Rotkäppchen* (German), and *Venevere Muinasjutt* (Estonian). One segment contained sibilants while the other did not. From these segments we created two sets of guises, one set for /s/ stimuli and one for pitch.²

We tested pitch and /s/ in isolation as the first step in determining whether /s/ variability holds the same indexical values in French and German as has been shown in English. Furthermore, testing the variables in isolation ensured that the survey averaged under thirty minutes for all participants.

For /s/ guises, speech segments were selected from the readings that contained at least four instances of /s/ and no other sibilants (/z/, /ʃ/). Due to phonotactic differences between the languages, the instances of /s/ were not controlled for syllable position or phonological environment. The guises were created by splicing into these recordings tokens of /s/ produced in isolation by the first author (see Campbell-Kibler, 2011; Mack & Munson, 2012) created under similar recording conditions as the original interviews. Approximately thirty /s/ tokens were produced and analyzed for CoG and skewness. The two stable tokens of /s/ that most closely matched (on measures of CoG and skewness) the two speakers with the lowest and highest average /s/ productions of CoG in Boyd (2018a) were selected as the stimuli to be spliced in for the [s−] and [s+] respectively. A middle token was selected that is comparable to that of the overall speaker production average of the same study. All naturally occurring tokens of /s/ were spliced out of the original speech and replaced with the stimulus /s/ tokens in Audacity (Audacity Team, 2016). The inserted /s/ tokens were matched for both intensity and duration of the original speech. Intensity was matched auditorily as slight liberty was taken with this to make the inserted stimuli sound as natural as possible. Though several previous studies have altered durational aspects of the sibilant (Levon, 2006; Linville, 1998; Rogers & Smyth, 2003), we felt that altering the sibilant durations made the speech sound highly unnatural, and instead we chose to match the stimuli with original duration produced by the speaker. The resulting stimuli consist of three versions of each sentence with identical [s−], [s], or [s+] tokens across all four languages. These three specific /s/ tokens were selected based on production data from Boyd (2018a) where [s+] and [s−] are representative of the highest and lowest average /s/ CoG values produced by the two most extreme speakers

TABLE 2. *Pretest Results of Each Speaker for Manipulation on a 7-point scale (1 is Straight/Masculine and 7 is Gay/Effeminate)*

	Straight/Gay	Masc./Effem.
English	1.733	2
French	2.866	2.333
German	2.333	1.866
Estonian	2.333	2

of that study, and [s] is representative of the average CoG of that study's overall speaker average. Table 3 gives the acoustic measurements of each guise.

For the pitch guises, different instances of speech from the same reading passages were selected. These clips, which were approximately as long as the /s/ clips, contained no sibilants at all. For the "mid-pitch" stimuli, each baseline stimulus was manually adjusted with very minor manipulations to average the pitch across all speakers (within $\pm 5\text{Hz}$) by altering the pitch points via Praat's "Manipulate" function. This "mid-pitch" can be considered representative of the speakers' natural pitch. For the "high" and "low" stimuli, the "mid-pitch" was adjusted by $\pm 25\text{Hz}$ across the entire utterance utilizing Praat's "Shift pitch frequencies..." function. The decision to adjust the pitch by $\pm 25\text{Hz}$ was based on the need to have the low/mid/high categories maximally distinct while maintaining stimuli pitches that occurred within the natural pitch range of the participants as seen across the full interviews of Boyd (2018a, 2018b). $\pm 25\text{Hz}$ seemed auditorily distinct enough to elicit a listener response while ensuring that the speech did not sound unnaturally high or low. For Estonian, there were no instances in the reading passage of a sentence without any sibilants, and instead a sentence containing only two instances of sibilance was selected. These sibilants were then spliced out of the recording and pitch manipulation continued as per the other languages. We deemed this acceptable as Estonian listeners are not part of the current experiment and a *post hoc* examination of the data shows none of the participants having any prior knowledge of Estonian, so the lack of individual phonemes would not have been noticed.

With three levels of manipulation ("low," "mid," and "high") on both /s/ and pitch guises across four languages the experiment resulted in a total of 24 guises, each approximately five seconds long. The pitch and /s/ guises were presented separately with the pitch of the /s/ guises being analogous to the "mid" pitch. The order of the experiment began with a short practice phase showing the format of the test. For each phase (practice and testing), the respondents were always first given the stimuli set corresponding to their native language. The order in which they responded to the remaining three (nonnative) languages was then randomized, and within each language all stimuli for that respective language was also randomized, but each language was presented separately from

TABLE 3. Centre of Gravity and Skewness values of the /s/ variants spliced into all four language stimuli

Variant	CoG	Skewness
[s-]	5208	1.1502
[s]	6436	0.033
[s+]	7988	-1.0795

the other three languages. Each respondent heard all twenty-four guises. Demographic information was collected at the end of the survey.

The survey was distributed across multiple social media platforms (i.e., Facebook and Twitter), via the authors' personal social networks, with separate recruitment text used to recruit French participants in French and German participants in German. The survey was completed online via Qualtrics (Qualtrics, 2018) by native speakers of English, French, and German, regardless of geographical location. Participants were asked to rate each guise on six semantic differential scales (Educated/Uneducated [French: *Cultivé/Peu instruit*; German: *Gebildet/Ungebildet*]; Straight/Gay [*Hétérosexuel/Homosexuel; Heterosexuell/Homosexuell*]; Lazy/Hardworking [*Paresseux/Travailleur; Faul/Fleißig*]; Friendly/Mean [*Amical/Inamical; Freundlich/Unfreundlich*]; Masculine/Effeminate [*Masculin/Efféminé; Maskulin/Feminin*]; Natural/Synthetic [*Naturel/Synthétique; Natürlich/Synthetisch*]), on a 100-point continuous scale (no numerical values of the rating were shown to the listener). These scales are largely analogous to those used in Levon (2006, 2007) with the addition of "Natural/Synthetic" as a fail-safe of sorts to ensure that all language stimuli appear natural to the respondents.

One challenge we encountered was the lack of a direct translation for "Masculine/Effeminate" in German. The pilot study respondents raised two possible translations: *maskulin* and *männlich*. Though it is possible to say someone has a masculine voice, "*maskulin Stimme*," a voice might also be described as "*So männlich*." Following conversations with multiple native German-speaking linguists, it was suggested that, though *maskulin* is not unambiguous (referring also to grammatical gender), it is unlikely that respondents outside of linguistics would be confused by the alternative meaning. Under their advice, we decided on the German pair, "*Maskulin/Feminin*."

RESULTS

Participants were excluded if they did not identify English, French, or German as their native language in the respective survey (Table 4). These remaining participants vary widely with respect to regional background and country of residence. English listeners were raised in Australia ($n = 1$), New Zealand ($n = 1$), various parts of the United Kingdom ($n = 9$), and the United States

TABLE 4. Total Number of Respondents by Language and Respondent Foreign Language Study

Survey Language	Total Participants	(of which) Studied English	(of which) Studied German	(of which) Studied French	(of which) Studied Estonian
English	27	--	9	16	0
German	23	25	--	19	0
French	32	34	14	--	0

($n = 16$). French listeners were from Belgium ($n = 1$), Canada ($n = 4$), France ($n = 26$), and Switzerland ($n = 1$). German listeners were from Austria ($n = 13$), Germany ($n = 11$), Italy ($n = 1$), Switzerland ($n = 1$), or unknown ($n = 1$). This regional variation was not possible to fully model quantitatively, but analysis suggests that regional dialect of the listener did not affect the results (in the interest of space these results are not reported). Table 4 also summarizes the number of survey respondents who reported having studied any of the stimuli languages; none of the participants were natively bilingual in the stimuli languages. A post-hoc examination reported below showed that crosslinguistic proficiency had no effect on ratings for any listener or listener groups. Summary statistics for each respondent language across all measures is included in Appendix 1.

One mean and standard deviation was estimated for each participant by pooling their responses on all rating scales for all guises, and these were used to perform z-score conversions on each participant's ratings. For each guise frame, a participant would rate it exactly three times (the high, mid, and low guises). As such, guise ratings can be treated as *paired*, or characterized in terms of the difference between two guise levels. In doing so, we can simplify our statistical analyses. For example, rather than conducting two-sample tests to compare the high and mid guises, a single-sample test can be used on the difference. Or, instead of estimating main effects of guise level, stimulus language, and an interaction between the two, we can simply fit a main effect of the difference between guise levels. Difference scores between the high and mid guise, and between the mid and low guise, were estimated within each participant, within each stimulus language, within each manipulation (/s/ and pitch), and within each rating scale. Table 5 presents a representative example from a participant in the English language survey for the Effeminate scale. A positive value indicates that the stimulus on the left (*high* or *mid*, respectively) was rated as more effeminate than the stimuli on the right (*mid* or *low*, respectively). Because these difference scores are calculated on subjects' z-scored ratings, these differences can be thought of as the magnitude of the difference relative to the range of the scale subjects used. If a given participant only used a narrow range of the scale, they could still have a large difference score, if they utilized opposite ends of their own range for these two ratings.

TABLE 5. Subset of results for one participant for one scale ('Effeminate')

Participant	Manipulation	Stimulus Language	Scale	High - Mid	Mid - Low
1	pitch	English	Effeminate	0.68	1.21
1	s	English	Effeminate	0.98	-0.15
1	pitch	Estonian	Effeminate	2.65	-3.11
1	s	Estonian	Effeminate	0.68	0.76
1	pitch	French	Effeminate	0.68	0.91
1	s	French	Effeminate	0.30	-0.76
1	pitch	German	Effeminate	-1.36	2.12
1	s	German	Effeminate	0.30	-1.44

English Survey Results

The difference in ratings that listeners gave to each guise between the high and mid and mid and low manipulations serve as our dependent variables (e.g., Table 6). The difference scores for a given label and a manipulation guise are symmetrical, without clear left or right skewness, as seen in Figure 2, making the use of single sample nonparametric tests appropriate.

As a first pass, we estimated pseudomedians and confidence intervals for each manipulated linguistic feature, stimulus language, rating scale, and dependent variable using the Hodges-Lehman estimator (Hollander & Wolfe, 1999).³ It is also possible to estimate *p*-values for these estimates using a one-sample Mann-Whitney U test, but with so many tests it is necessary to correct for multiple comparisons. We did so using the Holm-Bonferroni method, whereby the smallest *p*-value is multiplied by the number of tests *n*, the second smallest *p*-value is multiplied by *n*-1, etc. (Holm, 1979). Despite the adjustments, these *p*-values should still be treated with some caution, because of the number of tests conducted. The pseudomedians and confidence intervals, on the other hand, would remain unchanged regardless of the number of tests carried out.

In the bottom panel of Figure 3, we can see that the difference between mid and low manipulations of /s/ and pitch had effectively no reliable effect on listeners' ratings on any scale. But the top panel shows several reliable effects between the high and mid manipulations. For all stimulus languages, English listeners have rated guises with higher pitch as more Effeminate than mid pitches by about one standard deviation. There is an effect of similar magnitude on the Gay rating scale for the /s/ manipulations. However, it is not clear if the magnitude of difference is the same across stimulus languages; it certainly appears to be less for German, for example. In a cross-stimulus analysis, it is possible to fit a mixed effects model, and we did so with the high-mid difference score as the outcome variable, stimulus language as the predictor, and a random intercept by participant (R Core Team, 2020).⁴ The model estimates along with 95% bootstrap confidence intervals⁵ are presented in Table 6. This was the only model specification that was fit to the data, as we were solely focused on the effect of language stimulus on the "gay/straight" rating scale.

TABLE 6. *Fixed effects estimates for English listener's Gay rating differences between high and mid manipulations of /s/ (by-speaker random intercept sd = 0.54, residual deviance = 1.03)*

parameter	Estimate	(CI)	Std. Error	t-value
intercept	1.13	(0.69, 1.57)	0.22	5.03
stim = Estonian	-0.01	(-0.57, 0.54)	0.28	-0.05
stim = French	-0.02	(-0.57, 0.53)	0.28	-0.07
stim = German	-0.47	(-1.01, 0.08)	0.28	-1.67

In Table 6, the intercept corresponds to the estimated difference score on the English guise, which replicates the effect displayed in Figure 3 of front /s/ being rated about one z-score more Gay than mid /s/. The remaining stimulus language effects describe the difference between the difference score on English and these languages. The only stimulus language to have a large estimated difference from English is German. The direction of this effect would mean that English listeners do not rate front /s/ in the German stimulus as Gayer than they do in the English stimulus, but this effect is not statistically reliable (the bootstrap confidence interval includes zero, and the t-value is less than two). It's also the case that, in Figure 3, the difference score between mid and front /s/ was not reliably different from zero in the German guise. This is an apparently equivocal result: English listeners do not treat the German stimulus significantly different from English, but also do not rate it as significantly Gayer, either. There is too much statistical uncertainty to conclude whether or not the /s/ manipulation had an effect on English listeners' Gay rating in the German guise. What is clear, however, is that, for French and Estonian, front /s/ was rated as Gayer by English listeners to a degree indistinguishable from their ratings of English.

English listeners' Gayness rating differences between front and mid /s/ were most similar between the English and French guise, but otherwise varied in the magnitude of their sensitivity across the other guises. Despite this variability across stimulus languages, within each stimulus language, listeners tended to evaluate the front /s/ as Gayer than the mid /s/. Furthermore, our results show that this pattern is not limited to a small group of individuals being highly attuned to this variation. In other words, each respondent response varied greatly between the stimulus languages but the trend to rate front /s/ as Gayer than mid /s/ is consistent regardless of the individual variation in ratings (see Appendix 2 for more detail).

Data was collected on the participants' familiarity with the languages included in the experiment. Two sample Mann-Whitney U tests did not find a significant effect of having studied French on the French stimuli results ($U = 125$, $n_1 = 14$, $n_2 = 13$, $HL\Delta = 0.94$, $\rho = 0.67$, $p = 0.1$), nor of having studied German on the German stimuli results ($U = 94$, $n_1 = 19$, $n_2 = 8$, $HL\Delta = 0.46$, $\rho = 0.61$, $p = 0.5$).

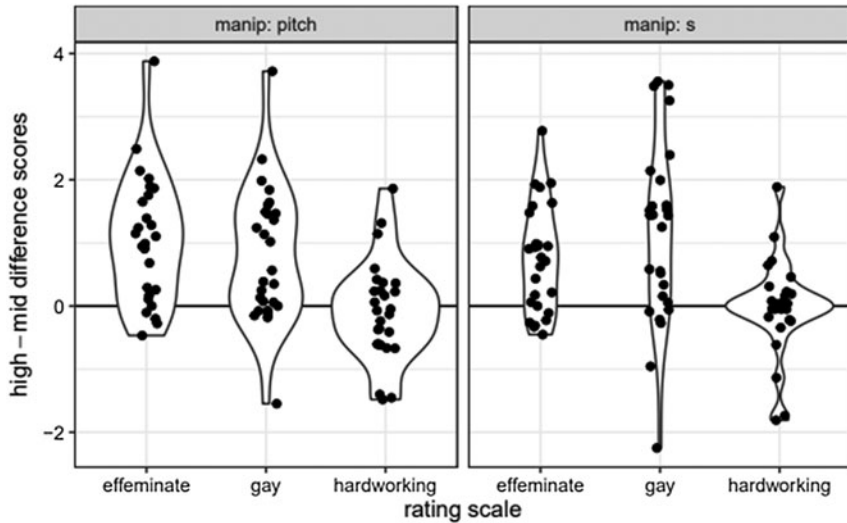


FIGURE 2. Distributions of high-mid difference scores for English listeners listening to English audio in the pitch (left) and /s/ (right) guises.

Finally, we examined how highly correlated participants' front versus mid /s/ difference scores for 'Effeminate' and 'Gay' were. Interestingly, participants' difference scores for these scales were moderately correlated for Estonian, French, and German but more weakly so for English, as illustrated in Figure 4.

French and German Survey Results

We followed the same initial procedure for the results from the French and German language surveys. Figure 5 displays the pseudomedians and confidence intervals for French respondents' high versus mid and mid versus low ratings. Again, p -values were estimated using a one sample Mann-Whitney U test and adjusted using the Holm-Bonferroni method.

The French listeners show no reliable differences between the mid and low guises for either linguistic variable for any language or rating scale. French listeners rate the higher pitch guises more Effeminate than the mid pitch guises for all languages. For the French guise, they have also rated the higher pitch guise as less Educated.

Their results for /s/ are different. While English listeners reliably rated front /s/ as more Effeminate for all language guises, and Gayer for all guises except for perhaps German, French listeners only reliably rate front /s/ as more Effeminate in English, and front /s/ appears to have no effect on their Gay ratings for any language stimulus. The effect size is also smaller for their rating of the English guise, with the front /s/ being rated approximately 0.5 standard deviations more Effeminate than the mid /s/, while the English listeners had an effect size of about one standard deviation more Effeminate. A two-sample Mann-Whitney U

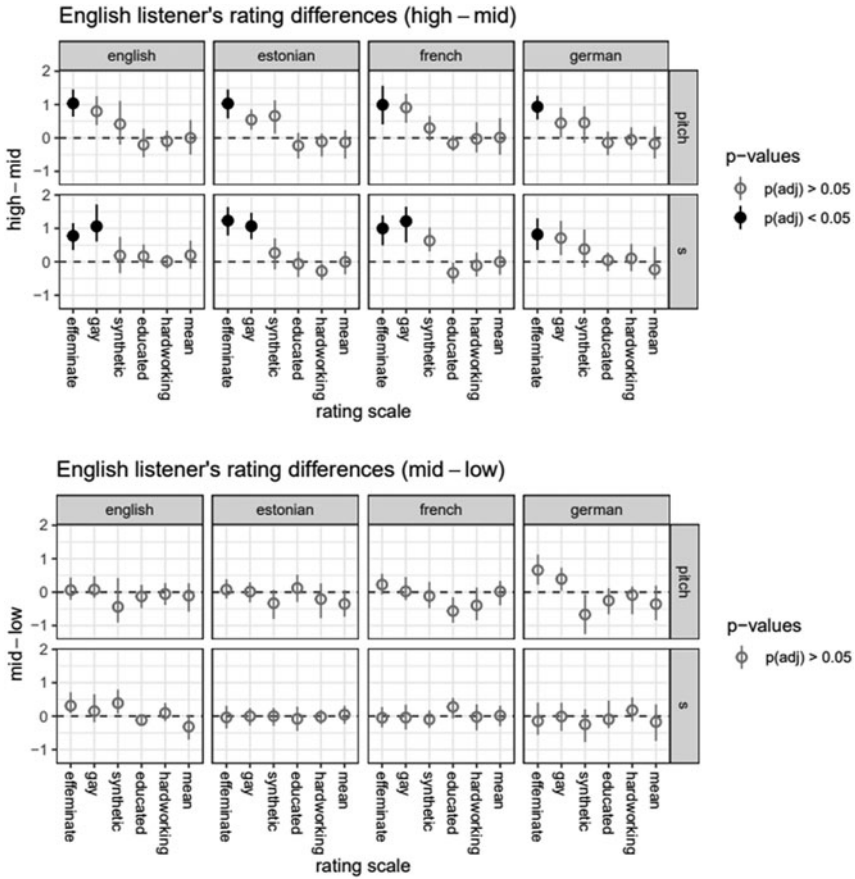


FIGURE 3. English listeners' responses to high versus mid (top) and mid versus low (bottom) manipulation of /s/ and pitch in four languages.

test found that there was no significant effect of English language ability on these Effeminate difference scores ($U = 103, n_1 = 10, n_2 = 22, HL\Delta -4 \times 10^{-5}, \rho = 0.47, p = 0.8$).

We fit a mixed effects model for effect of stimulus language on Gay rating scale differences for the front versus mid /s/ manipulation, with participant as a random intercept (Table 7). None of the parameters are reliably different from zero, meaning that French listeners were not consistently rating front /s/ as Gayer than mid /s/ for any language guise. Like English, French respondents show no listener or subset of listeners who consistently rated fronted /s/ as gayer than mid /s/ in any language guise (see Appendix 3).

There are even fewer effects of our guise manipulations for the German listeners. Again, the difference between mid and low manipulations had no effect on any rating scale for either manipulation. Moreover, there does not appear to be any

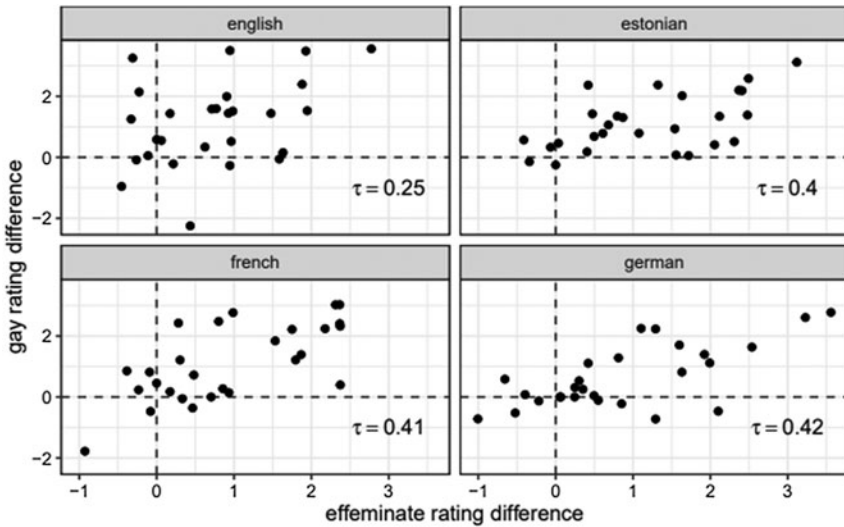


FIGURE 4. Correlation of English listeners' Gay rating difference and Effeminate rating differences between high and mid /s/ by language, with Kendall's τ .

difference between the high and mid manipulations, either, except perhaps for pitch on the Effeminate scale for the English stimuli. We did not carry out any further analysis of the German data, as the noneffect of our guise manipulation seems to be clear enough from Figure 6, though the pitch effect requires replication and closer evaluation.

DISCUSSION

Our results show two main findings of interest:

1. English listeners associate fronted /s/ and Gayness when listening to English, other languages they know, and languages they do not know.
2. French and German listeners do *not* associate fronted /s/ and Gayness, for any language, despite the presence of this feature in speech production by gay men in their respective languages.

The first finding we refer to as *indexical transfer* (see Bekker & Levon, 2017). The second finding calls forth a discussion of the “meaning potential” (Eckert, 2016) of /s/ in French and German.

Indexical Transfer

To account for the results from the English listeners, we propose a model of *indexical transfer*, drawing on Silverstein's (2003) theory of the indexical order, and Eckert's (2017) analysis of the indexicality of /s/. Whether or not the

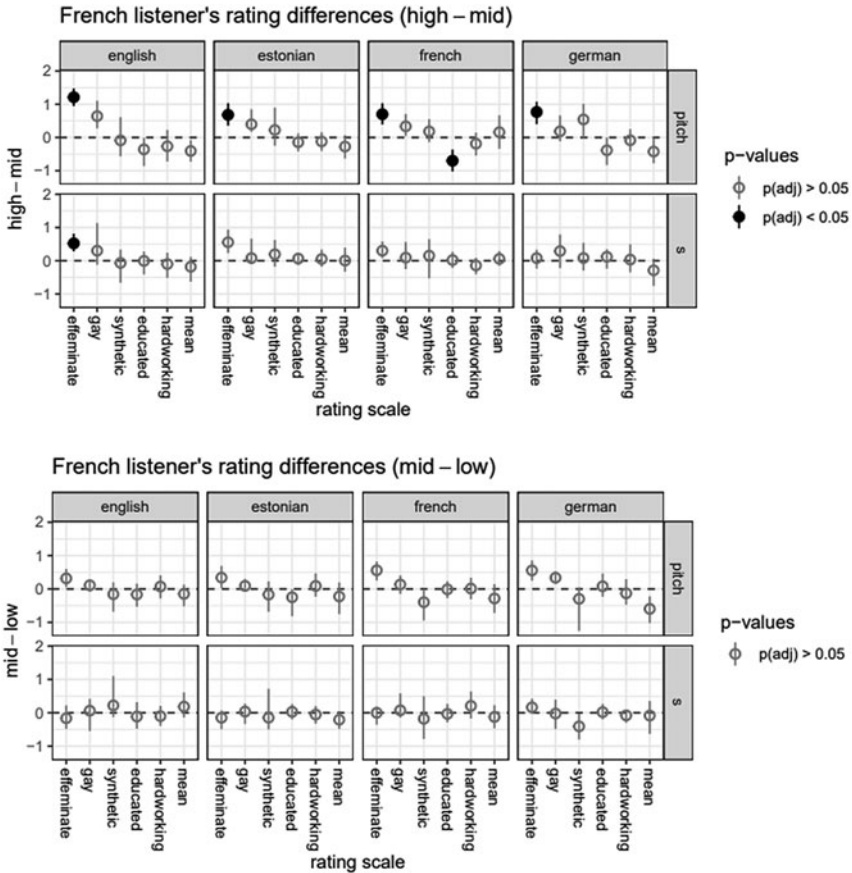


FIGURE 5. French listeners’ responses to high versus mid (top) and mid versus low (bottom) manipulation of /s/ and pitch in four languages.

TABLE 7. Fixed effects estimates for French listener’s Gay rating differences between high and mid manipulations of /s/ (by-speaker random intercept $sd = 0.32$, residual deviance = 0.82)

parameter	Estimate	(CI)	Std. Error	t-value
intercept	0.27	(-0.04, 0.57)	0.16	1.71
stim = estonian	-0.05	(-0.45, 0.35)	0.21	-0.25
stim = french	-0.18	(-0.59, 0.22)	0.21	-0.90
stim = german	-0.06	(-0.47, 0.34)	0.21	-0.29

process we are proposing is expressly one of a “transfer” (such as L1 transfer) is something that needs further exploration, but we use this term here because it reflects our interpretation of the empirical finding that English listeners make the

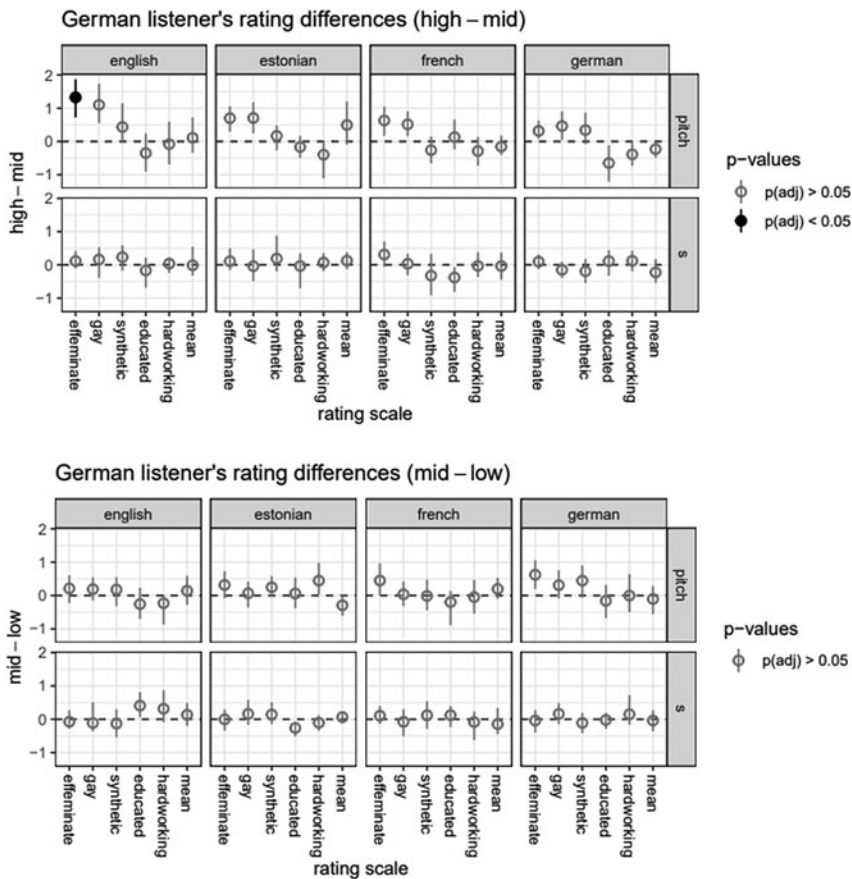


FIGURE 6. German listeners' responses to high versus mid (top) and mid versus low (bottom) manipulation of /s/ and pitch in four languages.

same indexical judgments about /s/ variation in English as they do about /s/ variation in non-English languages. We propose conceptualizing this as *transfer*, or extension, of their sociolinguistic knowledge about English to the other languages. This implies a temporal process whereby indexical associations are formed first in a native language and then later applied to other languages. Note that we are not suggesting that all social meaning associated with nonnative or unknown languages necessarily derives from native language knowledge. Nor are we suggesting that this transfer process will always happen when listeners encounter all possible linguistic variables in all nonnative languages. Indeed, it has been noted (Eckert, 2017:1198) that /s/ is “one sign that fits the bill” for Peirce’s (1934:448) “most perfect of signs” in terms of having symbolic, indexical, and iconic qualities. It might be the case that variation that has such a rich semiotic scope is also more available for crosslinguistic indexical transfer.

Modeling the indexicality of /s/ variation will draw on our results, those of Boyd (2018a, 2018b), and those in the 2017 special issue of *Linguistics* on “The Sociophonetics of /s/” (Levon, Maegaard, & Pharaoh, 2017), and the papers therein. First, recall that any attributions of /s/ variation to differences in vocal tract anatomy cannot account for all observed social differences in /s/ production (Fuchs & Toda, 2010). Indeed, if indexical interpretation were a direct result of learned correlates with physiological effects (see, e.g., Barreda, 2017), then we should see no differences in the current study according to listener L1.

What, then, is indexical transfer? Recalling our example of the tire hiss (Figure 1), we predict that listeners would not give the same attitudinal responses to a nonlinguistic production of an acoustic signal that corresponds to /s/. Just as a snake’s hiss is interpreted as a sign of danger (Eckert, 2017), a tire’s hiss is interpreted as a sign of a leak in the tire. But furthermore, an actual tire’s hiss is highly unlikely to actually match the acoustic signal that corresponds to /s/, given the differences in articulation (so to speak), and this also makes it an entirely different sign. As Eckert (2017:1198) noted, a human hiss (e.g., by an evil villain), is produced with a different articulatory configuration than the variability in /s/ that is the focus of the current study; it is therefore “the phonetic *process*, not just the individual segment, that constitutes the sociolinguistic variable” (emphasis original). This provides a framework for interpreting how English listeners are parsing the sociophonetic variation in non-English speech stimuli, a process which first relies on the ability for the listener to recognize a segment as a segment with an acoustic signal capable of indicating (social) meaning. First, we must establish that listeners are recognizing non-English speech as speech, before making indexical inferences about that speech. Second, English listeners appear to be recognizing non-English /s/ segments as speech segments, comparable to English /s/. In order for indexical transfer to occur, listeners must link the phonetic segments of the input language to those corresponding within the languages that they are more familiar with. Third, English listeners appear to be extracting (social) meaning from non-English speech. It is perhaps not surprising that, in the absence of the ability to extract referential meaning from speech, listeners attempt to extract indexical information. Since they lack full or even any linguistic knowledge of that speech,⁶ they rely on the same processes that they would rely on when making indexical inferences. In other words, the indexical order of /s/ that an English listener orients to in any given moment is the same whether listening to English or a non-English language. There are no additional *n*+1st meanings (Silverstein, 2003) that arise from this process, and indeed it is unlikely that the indexical order will be updated or changed from the process (of hearing speech stimuli in a laboratory setting) because a listener of an unfamiliar language will presumably lack the social knowledge (of, for example, relevant persona in that linguistic community) to update that order (other than to add the meaning “speaker of another language/language X”). The quantitative results are, therefore, identical across all languages, rather than being, for example, stronger for English than the non-English languages.

In other words, when a listener hears a language they have little to no knowledge of,⁷ they apply whatever interpretive resources they have available to them. Lacking indexical knowledge or sociolinguistic competence in an unfamiliar language, the listener might apply an indexical interpretation as an attempt to extract meaning where lexical and grammatical meaning fails. For an English listener, the indexical field (Eckert, 2008) of /s/ may contain indexes of social class, gender, sexual orientation, level of education, and so on, but indexes of gender and sexual orientation, gayness as well as effeminacy, hold very strong metadiscursive value and may likely be the set of indexes that are activated when there is little else to signal meaning. Just as language learners filter their L2 phonology through their L1 phonology (Flege, 1987), listeners may filter the indexical interpretation of any speech stream through their first language indexical order. This parallelism suggests a cognitive embedding such that “learned acoustic patterns are mapped simultaneously to linguistic representations and to social representations” (Sumner, Kim, King, & McGowan, 2014). Future work on multilingual speakers and learners of different proficiency levels, and individual differences within all groups, would give us a fuller picture of these representations and processes, given what we know about how bilinguals shift their perceptual boundaries (e.g., Elman, Diehl, & Buchwald, 1977).

What are the alternatives to an analysis of indexical transfer? Levon, Maegaard, and Pharaoh (2017:984) and related papers have pointed out that “there are striking similarities in the perceived meanings of fine-grained phonetic variation in /s/ production across a range of linguistic and cultural contexts.” Perhaps our English listeners are, at some level, aware of this fact. Levon et al. (2017:984; see also Eckert, 2017) theorized the concept of “synesthetic sound symbolism,” specifically “magnitude symbolism” with respect to /s/ variation, noting the ways in which /s/ variation is linked to the perceived size of the speaker, which is then linked to gender, which is then linked to sexual orientation. However, they (2017:984, 986) expanded on Silverstein (2003) and others to show how that process (i.e., from n to $n + 1$ to $n + 2$) is necessarily “taken up and interpreted in language- and culture-specific ways,” that are what enable the emergence of indexes of gender and sexual orientation. Therefore, even if a similar indexical process is at work across the many different languages studied thus far, the process of interpreting social meaning is still necessarily tied to the language and sociolinguistic context in question.

Furthermore, the results for the English listeners here are orthogonal to the actual patterning of variation in those non-English languages. The fact that /s/ production does pattern with gender in German (Fuchs & Toda, 2010), or with sexual orientation in German (Boyd, 2018a, 2018b) has no bearing on how English listeners (who do not know German) hear /s/ in German. We therefore expect indexical transfer to apply even in cases unlike those described here, where the actual production patterns in a language are at odds with (or just do not correlate at all with) the indexicalities identified by nonnative listeners. This has interesting implications for a phonetic level of crosscultural misunderstanding.

One further item to note is our failure to replicate past results on the correlation between fronted /s/ variants and higher perceived level of education (Campbell-Kibler, 2011; Levon, 2014), specifically with regard to the English guises and respondents. This may be due in part to our use of an English speaker who has a noticeable Essex accent. Holmes-Elliott's and Levon's (2017) study of /s/ variation analyzed speakers from Essex as representing lower socioeconomic status, which might be connected to the speaker's perceived level of education (see, e.g., Cepeda, 1995). However, two-thirds of our English respondents are from outside of the UK, with likely limited ability in placing the accent and its associated array of social meanings. In terms of the French and German listeners, the association between /s/ variation and perceived level of education has, to our knowledge, not been previously tested.

Another possibility is that variation in /s/ is somehow made more salient in this speaker's voice, and that the results would not obtain for the same manipulation in another speaker's voice. For example, despite some evidence that /s/ variation plays a role in perceptions of nonhegemonic femininity (e.g., Bekker & Levon, 2017; Podesva & Van Hofwegen, 2014; Saigusa, 2016), the results would likely be muted for a speaker clearly heard as female. Another obvious follow-up study would be to replicate the study with a speaker of a US English variety, given the high proportion of US-based respondents. At the same time, the evidence for crosslinguistic indexical transfer itself suggests that the results are quite robust to variation between talkers. In other words, if an English listener is willing and able to respond to languages they are entirely unfamiliar with, they are probably also likely to do so for any male speaker of English, regardless of the regional variety.

Indexicality in Production but not Perception

For the French and German respondents, we see vastly different results than those seen for the English respondents. Where the English listeners appeared to be attuned to variation of /s/, regardless of the languages, the French and German listeners did not show this for any language, including their native language. These results and those seen in Boyd (2018a), where /s/ is robustly shown to vary according to sexual orientation in French and German men's speech production, suggest that the indexical meaning of /s/ for French and German speakers and listeners is not straightforward.

One descriptive framework for understanding any production/perception mismatch is Labov's (1972) taxonomic distinction of *indicators*, *markers*, and *stereotypes*. Variation in /s/ in the speech of native French and German speaking men can be seen to pattern like a *marker*: the variation patterns according to social group differences and exhibits topic-linked stylistic variation (Boyd, 2018b) but appears to be below the level of awareness. In contrast, *stereotypes* are variables subject to social evaluation, and the distinguishing factor between them lies in the level of social awareness (Eckert, 2008; Labov, 1972:314–15). While /s/ appears to be a stereotype in English, the same cannot be said for French and German /s/.

Relatedly, it seems that indexicality in production precedes indexicality in perception. Indexical orders rely on “recognition” (Agha, 2003) of signs as *being signs*, that is, as marking stylistic distinctiveness (Irvine, 2001). French and German [s+] may currently have “meaning potential” (Eckert, 2016), waiting for a “baptismal moment” (Silverstein, 2003) to be taken up as an index of gay identity in perception. Differences in production suggest that [s+] may become enregistered through continued socially differentiated use in interaction. Building on Eckert’s observation that “innovative personae are the more immediately accessible manifestations—indeed agents—of change” (2018:190), Boyd (2018b) showed that those gay French and German speakers who produce fronted [s+] variants are also those enacting and embodying specific types of *counter-hegemonic* gay personae. Iterated use in the construction of such personae increases the likelihood that [s+] will become associated with those personae.

Though we have shown that /s/, currently, does not clearly index sexual orientation for French listeners, Russell (2017) showed that French speakers who have been asked to perform “gay sounding speech” produce higher frequency /s/ CoG when performing a “gay” speech style than when asked to perform as “straight.” The differences seen between the straight and gay styles produced by Russell’s speakers is much less than those differences seen in Boyd (2018a), but it is nonetheless interesting that /s/ is an available resource to draw on in stereotyped “mock” speech. Why it is not heard as “gay” or “effeminate” in a perception task remains unanswered.

It may be that there is an important cultural difference in terms of how participants engage with ratings of gayness and masculinity on a matched-guise test. First, Boyd’s (2018a) speaker sample included a much higher proportion of queer-identified speakers than the current listener sample. Additionally, while English listeners are willing to rate voices on dimensions of gayness and masculinity, French and German listeners may be more reticent, for various reasons. The finding that the German listeners were willing to make a determination for one voice with respect to pitch might reflect more about the relative social acceptability of pitch as a cue to sexual orientation than other speech features, like /s/, which (like English) might be stigmatized.

However, the metalinguistic commentary described earlier and by Boyd (2018b) suggests that /s/ variation probably does not factor into judgments of “Gay” or “Effeminate” sounding by French and German listeners. Since it is possible that the null results presented here are limited by the fact that /s/ and pitch were tested in isolation, future work should investigate the covariation of these and any number of other potential phonetic variants (such as the “gay nasal” stereotype in German, [e.g., Kachel et al., 2018]), for a fuller picture (cf., Campbell-Kibler, 2011; Levon, 2007; Pharo & Maegaard, 2017).

Our findings also speak to our understanding of the mechanisms behind production/perception mismatches in the wider scope of phonetics. The results presented here are broadly akin to a phenomenon like near-mergers (e.g., Labov, Karen, & Miller, 1991), where speakers have a phonetic distinction between two

historically distinct phonemes even though they do not perceive any difference between those phonemes. The difference there is that near-merger is typically evidenced by a mismatch within the same speaker-listeners, whereas here we see a production difference across a set of speakers and a perception effect (or lack thereof) within a different group of listeners.

CONCLUSIONS

The results shown here demonstrate the potency of speakers' indexical knowledge, with English listeners making the same indexical inferences regardless of their level of familiarity with the language to which they are listening. English listeners know that the phoneme /s/ produced with a high frequency indexes a nonhegemonic masculinity and demonstrate this knowledge in perception. We now know that they will also extend this knowledge to unfamiliar linguistic contexts where the language is clearly not English, and our results suggest that this is only possible for English listeners because of the *enregistered* status of fronted /s/ in the language. The results also demonstrate the danger of imputing indexical associations in perception from production data alone, in that we see a mismatch between two of the three listener groups. Despite the fact that gay and straight French and German men appear to produce similar differences in /s/ production as seen in English, fronted /s/ is not an enregistered feature of gay speech in these languages (yet). Perhaps what we are seeing with the French and German /s/ is a potential index waiting for its "baptismal moment" to be taken up as an index of gay identity that just is not there yet.

NOTES

1. Under the theory of Indexical Orders (Silverstein, 2003), interspeaker variation such as this might constitute an n -th order indexicality, potentially emerging from a prior $n-1$ th order, and with the potential for new $n + 1$ th indexicalities to develop off of it.
2. While the reading of fairy tales opens up the possibility of a more "animated" or "clear-speech" style (e.g., Maniwa, Jongman, & Wade, 2009), Boyd (2018b) showed no difference between interview speech and L1 read speech for the French and German speakers.
3. The Hodges-Lehman estimator is a nonparametric estimator of the "location parameter," or the center of a distribution. It is estimated by taking the mean of all possible pairs of data points (including self-pairings) and then taking the median of the resulting distribution.
4. Here, a random slope of stimulus language by participant would not be identifiable, since each participant only contributed one data point per stimulus language.
5. Based on 10,000 parametric bootstraps using *bootMer* from the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015).
6. Indeed, Kabak and Maniwa (2007) showed that listeners demonstrate "difficulty in perceiving foreign-language contrasts," specifically with respect to fricative voicing and place of articulation differences between English and German. However, this should not have impacted results as our stimuli explicitly excluded all sibilants other than /s/. Furthermore, Boyd (2018a) reported French and German /s/ productions at frequencies comparable to English /s/ productions (cf., Zimman [2017] for discussion of English /s/). Thus, respondents would be likely to have similar perception of the variable in question across language boundaries.
7. It could be that knowledge of the language interrupts or mediates indexical transfer (see Phrao et al., 2014; Walker et al., 2014), but our results show no indication of this effect.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S0954394521000089>

ACKNOWLEDGMENTS

We would like to thank Mirjam Eiswirth and Michäel Gautier for their help in translating the texts for the German and French versions of the matched guise test. We would also like to thank Erez Levon and Claire Cowie, as well as the audiences of NVAV 44 in Vancouver and UKLVC 11 in Cardiff, and two anonymous reviewers, for their feedback on earlier versions of this project.

REFERENCES

- Agha, Asif. (2003). The social life of cultural value. *Language and Communication*, 23(3–4), 231–73.
- . (2005). Voice, footing, enregisterment. *Journal of Linguistic Anthropology*, 15(1), 38–59.
- Audacity Team. (2016). Audacity(R): Free Audio Editor and Recorder [Computer program]. Version 2.1.2. Retrieved from <http://www.audacityteam.org/>
- Avery, Jack D., & Liss, Julie M. (1996). Acoustic characteristics of less-masculine-sounding male speech. *The Journal of the Acoustical Society of America*, 99(6), 3738–48.
- Barreda, Santiago. (2017). Listeners respond to phoneme-specific spectral information when assessing speaker size from speech. *Journal of Phonetics*, 63, 1–18.
- Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve. (2015) “Fitting linear mixed-effects models using lme4.” *Journal of Statistical Software*, 67(1), 1–48.
- Beal, Joan C. (2009). “You’re not from New York city, You’re from Rotherham”: Dialect and identity in British indie music. *Journal of English Linguistics*, 37(3), 223–40.
- Bekker, Ian, & Levon, Erez. (2017). The embedded indexical value of /s/-fronting in Afrikaans and South African English. *Linguistics*, 55(5), 1109–39.
- Boyd, Zac. (2018a). *Crosslinguistic variation of /s/ as an index of non-normative sexual orientation and masculinity in French and German men*. Doctoral dissertation, The University of Edinburgh.
- . (2018b). A certain kind of gay identity: [s+] and contextually mediated variation in bilingual French and German men. Unpublished Manuscript.
- Brown, Bruce L., & Lambert, Wallace E. (1976). A cross-cultural study of social status markers in speech. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, 8(1), 39–55.
- Cameron, Deborah. (1997). Performing gender identity: Young men’s talk and the construction of heterosexual masculinity. In S. Johnson & U. H. Meinhof (Eds.), *Language and Masculinity* (pp. 47–64). Oxford: Blackwell.
- Campbell-Kibler, Kathryn. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech*, 82(1), 32–64.
- . (2011). The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change*, 22(3), 423–41.
- Cepeda, Gladys. (1995). Retention and deletion of word-final /s/ in Valdivian Spanish (Chile). *Hispanic Linguistics*, (6), 329–54.
- Clopper, Cynthia G., & Bradlow, Ann R. (2009). Free classification of American English dialects by native and non-native listeners. *Journal of Phonetics*, 37(4), 436–51.
- Eckert, Penelope. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4), 453–76.
- . (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, 41, 87–100.
- . (2016). Variation, meaning, and social change. In N. Coupland (Ed.), *Sociolinguistics: Theoretical Debates*. Cambridge University Press.
- . (2017). Comment: The most perfect of signs: Iconicity in variation. *Linguistics*. 55(5), 1197–1207.
- Eisenstein, Miriam. (1982). A study of social variation in adult second language acquisition. *Language Learning*, 32(2), 367–91.

- Elman, Jeffrey L., Diehl, Randy L., & Buchwald, Susan E. (1977). Perceptual switching in bilinguals. *The Journal of the Acoustical Society of America*, 62(4), 971–4.
- Flege, James E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15(1), 47–65.
- Foulkes, Paul, Docherty, Gerard J., Khattab, Ghada, & Yaeger-Dror, Malcah. (2010). In D. Preston & N. Niedzielski (Eds.), *A Reader in Sociophonetics*. Berlin: Mouton.
- Fuchs, Susanne, & Toda, Martine. (2010). Do differences in male versus female /s/ reflect biological or sociophonetic factors? In S. Fuchs, M. Toda, & M. Žygis (Eds.), *Turbulent sounds: An interdisciplinary guide* (Vol. 21). Walter de Gruyter.
- Gaudio, Rudolf P. (1994). Sounding gay: Pitch properties in the speech of gay and straight men. *American Speech*, 69(1), 30–57.
- Guzik, Karita M. (2006). Acoustic analysis of phonetic parameters of less masculine sounding German speech. *Arbeitsberichte Des Instituts Für Phonetik Und Digitale Sprachverarbeitung Universität Kiel*, 36, 15–29.
- Hobart, Michael. (2013). *Manifestation of sexual identity in the speech of French second language speakers of English: Evidence from the gay /s/*. Master’s thesis, The University of Edinburgh.
- . (2014). *One Social Meaning – Two Languages: Sociophonetic Markers Indexing Identity in Bilinguals*. Master’s thesis, The University of Edinburgh.
- Hollander, Myles, & Wolfe, Douglas A. (1999). *Nonparametric Statistical Methods, Second Edition*. Wiley-Blackwell.
- Holm, Sture. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.
- Holmes-Elliott, Sophie, & Levon, Erez. (2017). The substance of style: Gender, social class and interactional stance in /s/-fronting in southeast England. *Linguistics*, 55(5), 1045–72.
- Irvine, Judith. (2001). Style as distinctiveness: The culture and ideology of linguistic differentiation. In J. R. Rickford & P. Eckert (Eds.), *Style and Variation* (21–43). Cambridge: Cambridge University Press.
- Johnstone, Barbara, Andrus, Jennifer, & Danielson, Andrew E. (2006). Mobility, Indexicality, and the Enregisterment of “Pittsburghese.” *Journal of English Linguistics*, 34(2), 77–104.
- Kabak, Barış, & Maniwa, Kazumi. (2007) “L2 perception of English fricatives in clear and conversational speech: The role of phonemic, phonetic, and acoustic factors.” *Proceedings of ICPhS 2007*.
- Kachel, Sven, Simpson, Adrian P., & Steffens, Melanie C. (2018). “Do I sound straight?” - Acoustic correlates of actual and perceived sexual orientation and masculinity/femininity in men’s speech. *Journal of Speech, Language, and Hearing Research*, 61(7), 1560–78.
- Labov, William. (1972). *Sociolinguistic patterns* (No. 4). University of Pennsylvania Press.
- Labov, William, Karen, Mark, & Miller, Corey. (1991). Near-mergers and the suspension of phonemic contrast. *Language Variation and Change*, 3(1), 33–74.
- Lambert, Wallace E., Hodgson, Richard C., Gardner, Robert C., & Fillenbaum, Samuel. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology*, 60(1), 44–51.
- Levon, Erez. (2006). Hearing “Gay”: Prosody, interpretation, and the affective judgements of men’s speech. *American Speech*, 81(1), 56–78.
- . (2007). Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society*, 36, 533–54.
- . (2014). Categories, stereotypes, and the linguistic perception of sexuality. *Language in Society*, 43, 539–66.
- Levon, Erez, & Holmes-Elliott, Sophie. (2013). East end boys and west end girls: /s/-fronting in Southeast England. *University of Pennsylvania Working Papers in Linguistics*, 19(2), 111–20.
- Levon, Erez, Maegaard, Marie., & Phrao, Nicolai. (2017). Introduction: Tracing the origin of /s/ variation. *Linguistics*, 55(5), 979–92.
- Linville, Sue Ellen. (1998). Acoustic correlates of perceived versus actual sexual orientation in men’s speech. *Folia Phoniatrica et Logopaedica*, 50, 35–48.
- Mack, Sara. (2010). Perception and identity: Stereotypes of speech and sexual orientation in Puerto Rican Spanish. In *Selected Proceedings of the 12th Hispanic Linguistic Symposium* (136–47). Sommerville, MA: Cascadilla Proceedings.
- Mack, Sara, and Benjamin Munson. (2012). The influence of /s/ quality on ratings of men’s sexual orientation: Explicit and implicit measures of the ‘gay lisp’ stereotype. *Journal of Phonetics* 40(1): 198–212.

- Maegaard, Marie, & Pharao, Nicolai. (2016). /s/ variation and perceptions of male sexuality in Denmark. In E. Levon & R. B. Mendes (Eds.), *Language, Sexuality, and Power* (88–104). Oxford University Press.
- Major, Roy C. (2007). Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition*, 29(4), 539–56.
- Maniwa, Kazumi, Jongman, Allen, & Wade, Travis. (2009). Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America*, 125(6), 3962–73.
- Montgomery, Chris, & Moore, Emma. (2018). Evaluating S(c)illy voices: The effects of salience, stereotypes, and co-present language variables on real-time reactions to regional speech. *Language*, 94(3), 629–61.
- Moreau, Marie-Louise, Thiam, Ndiassé, Harmegnies, Bernard, & Huet, Kathy. (2014). Can listeners assess the sociocultural status of speakers who use a language they are unfamiliar with? A case study of Senegalese and European students listening to Wolof speakers. *Language in Society*, 43(3), 333–48.
- Munson, Benjamin, & Babel, Molly. (2007). Loose lips and silver tongues, or, projecting sexual orientation through speech. *Language and Linguistics Compass*, 1(5), 416–49.
- Munson, Benjamin, McDonald, Elizabeth C., DeBoe, Nancy L., & White, Aubrey R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, 34(2), 202–40.
- Peirce, Charles S. 1931–58. In Charles Hartshorne, Paul Weiss, and A. W. Burks (Eds.), *The Collected Papers of Charles S. Peirce*. 8 vols. Cambridge, MA: Harvard University Press.
- Pharao, Nicolai, & Maegaard, Marie. (2017). On the influence of coronal sibilants and stops on the perception of social meanings in Copenhagen Danish. *Linguistics*, 55(5), 1141–67.
- Pharao, Nicolai, Maegaard, Marie, Møller, Janus Spindler, & Kristiansen, Tore. (2014). Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society*, 43, 1–31.
- Podesva, Robert J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics*, 11(4), 478–504.
- Podesva, Robert J., & Van Hofwegen, Janneke. (2014). How Conservatism and Normative Gender Constrain Variation in Inland California: The Case of /s/. *University of Pennsylvania Working Papers in Linguistics*, 20(2), 129.
- Qualtrics. (2018). *Qualtrics*. Provo, Utah, USA.
- Rácz, Péter, & Schepács, András. (2013). The perception of high frequency sibilants in Hungarian male speech. *Acta Linguistica Hungarica*, 60(4), 457–68.
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rogers, Henry, & Smyth, Ron. (2003). Phonetic differences between gay- and straight-sounding male speakers of North American English. *Proceedings of the 15th International Congress of Phonetic Sciences*, 1855–8.
- Russell, Eric Louis. (2017). Style shifting and the phonetic performance of gay vs. straight: A case study from French. *Journal of Language and Sexuality*, 6(1), 128–76.
- Saigusa, Julie. (2016). Jane Lynch and /s/: The Effect of Addressee Sexuality on Fricative Realization. *Lifespans and Styles*, 2(1), 10.
- Silverstein, Michael. (2003). Indexical order and the dialectics of sociolinguistic life. *Language and Communication*, 23(3–4), 193–229.
- Smyth, R., Jacobs, G., & Rogers, H. (2003). Male voices and perceived sexual orientation: An experimental and theoretical approach. *Language in Society*, 32(3), 329–50.
- Squires, Lauren. (2010). Enregistering internet language. *Language in Society*, 39(4), 457–92.
- Stuart-Smith, Jane. (2007). Empirical evidence for gendered speech production: /s/ in Glaswegian. In *Change in Phonology: Papers in Laboratory Phonology 9* (Vol. 42, 65–86). Benjamins Stevens, MIT Press Strand.
- Sumner, Meghan, Kim, Seung Kyung, King, Ed, & McGowan, Kevin B. (2014). The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology*, 4, 1015.
- Szakay, Anita. (2006). Rhythm and pitch as markers of ethnicity in New Zealand English. In *Proceedings of the 11th Australasian International Conference on Speech Science & Technology*, University of Auckland, 421–6.
- Talbot, Mary. (2010). *Language and Gender* (2nd ed.). Cambridge, UK: Polity Press.
- Thorpe, David. (2014). *Do I Sound Gay?* (Film). United States: Sundance Selects.

- Titze, Ingo R. (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America*, 85(4), 1699–1707.
- Van Bezooijen, Renée. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38(3): 253–65.
- Vaughn, Charlotte R., & Bradlow, Ann R. 2017. Processing Relationships Between Language-Being-Spoken and Other Speech Dimensions in Monolingual and Bilingual Listeners.” *Language and Speech* 60(4): 530–61.
- Walker, Abby, García, Christina, Cortés, Yomi, & Campbell-Kibler, Kathryn. (2014). Comparing social meanings across listener and speaker groups: The indexical field of Spanish /s/. *Language Variation and Change*, 26(2), 169–89.
- Whiteside, Sandra P. (2001). Sex-specific fundamental and formant frequency patterns in a cross-sectional study. *Journal of the Acoustical Society of America*, 1110(1), 464–78.
- Zimman, Lal. (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.
- _____. (2017). Variability in /s/ among transgender speakers: Evidence for a socially grounded account of gender and sibilants. *Linguistics*, 55(5), 993–1019.