



Volume 23

Issue 2 *Selected Papers from New Ways of Analyzing
Variation (NWAY 45)*

Article 11

11-1-2017

Phonetic Variation and Self-Recorded Data

Lauren Hall-Lew
University of Edinburgh

Zac Boyd
University of Edinburgh

Phonetic Variation and Self-Recorded Data

Abstract

Self-recordings, when speakers record themselves without a researcher present, are attractive for potentially eliciting a wider range of styles than is obtained through interviews. To compare the stylistic differences between self-recorded speech and interview speech, we present an analysis of sibilant production among four speakers in both contexts. Our results show that the contrast between self-recordings and interviews can be a reliable predictor, with differences often surpassing those between interview speech and read speech. We suggest that self-recordings may be stylistically different enough from interviews to justify overcoming the practical challenges of their collection, integrating the self-recording into standard sociolinguistic methodologies, at least for studies of intraspeaker variation and the description of variable phenomena.

Phonetic Variation and Self-Recorded Data

Lauren Hall-Lew and Zac Boyd

1 Introduction

Self-recording, a type of “non-participant observation” (Levon 2013:210), refers to when participants record their speech without a researcher present. Self-recordings present an appealing method for data collection in sociolinguistics because of the presumption that they improve the ecological validity of the speech event and the implication that they may reduce the Observer’s Paradox (Labov 1972). In fact, they are appealing even if they only alter the Observer’s Paradox (cf., Wilson 1987), such as eliciting speech activities that are not typical to interviews (Schøning and Møller 2009), which may result in a broader production of stylistic variation. In this paper, we present phonetic data from eleven self-recordings from four speakers to argue that this is the case.

Studies based on self-recorded data are relatively rare, in part because such data presents a number of ethical, logistical, and analytical challenges. Ethically, the challenge is around obtaining informed consent from all individuals in all of the self-recordings. While those making the self-recordings can be instructed on how to obtain informed consent, it is not always feasible or desirable to obtain the consent of all those whose voices become recorded, particularly if the recording is made in a public place. Logistically, the researcher gives up knowledge and control over the context of data collection: they are unable to control the amount of ambient noise on the recording, for example, and are unable to elicit more speech in a given recording if the recording is too short. There are also practical issues similar to managing research assistants, since this is effectively what the participants themselves become, including scheduling, paperwork, recording equipment, and the like. Analytically, the downside to self-recordings is that researchers lack much more information about the social context of the recording than they would typically have in fieldwork contexts. What’s more, different self-recordings often differ stylistically (Meyerhoff, Schlee, & MacKenzie 2015:56; Podesva 2007, 2011a, 2011b). We return to this point later.

Self-recordings are more readily obtained now than in the past because people’s comfort levels have shifted with respect to being recorded and making their own recordings. Both are increasingly commonplace experiences, cross-culturally. While our focus in this paper is on self-recordings that are motivated by a linguist, the sheer feasibility of this method is based on the fact that the world is now full of self-recordings that are made for personal or other reasons, from casual, everyday recordings made on one’s phone, to video blogs created for consumption on YouTube, to family oral history interviews made for school projects.¹ On the whole, potential participants and their interlocutors are more likely to have recorded themselves or others prior to the research than they ever have been before.² Self-recordings are, for many, a well-practiced method.

In the present paper we compare sociolinguistic interview data with self-recording data acquired from speakers who are personally comfortable and experienced with portable recording technology. Despite differences in the directionality of the effect between speakers, overall we find that the inclusion of self-recordings results in capturing a fuller picture of a speaker’s range of phonetic variation than what would be captured from a sociolinguistic interview, alone.

2 Background

2.1 Self-Recordings

Self-recordings are not as yet typical in sociolinguistic speech elicitation and data collection. For example, there is no section in Mallinson, Childs, and Van Herk’s (2013) textbook on data collec-

¹<https://storycorps.org/discover/storycorpsu/teacher-resources/>

²This fact suggests that younger speakers may orient to even a standard interview context in a more familiar way than in previous decades. If so, Apparent Time inferences might reflect a change in style over time that could impact the interpretation of a change in progress. Testing this idea is beyond the scope of the present paper, as it would require a microanalysis of age-stratified longitudinal data.

tion in sociolinguistics on the topic of self-recordings (but see Levon 2013). While sociolinguists are known for their elicitation of a wide range of speech styles, the researcher is usually present for the recording of all of those styles. There are some early examples of pseudo-self-recordings, where a researcher is present for the recording but is acting as a conversational participant rather than an interviewer (Hindle 1979), or acting as a background observer rather than an interviewer (Coupland 1980). There are, however, some studies that focus only on self-recorded data. Podesva has analyzed the phonetic variation of one (Podesva 2007; 2011b) to three (Podesva 2011a) gay American men who each made three self-recordings in comparable settings: professional one-on-one, social one-on-one, and social group contexts. In each case, the social persona relevant to each context corresponded to reliable phonetic differences between contexts. While this data collection was researcher-led, amateur self-recordings made for posting on YouTube or similar sites are also examples of self-recorded data with potential for linguistic analysis. Lee (2017) accounts for vocalic variation in the speech of one popular YouTuber by comparing his speech in four different YouTube styles. Schneider's (2016) overview of the use of YouTube for World Englishes research includes both other-recorded and self-recorded speech, and the pros and cons he details are parallel to those for self-recordings: a wider stylistic range might be obtained, but at the expense of the researcher understanding the full social context in which the recording was made.

Researchers are increasingly collecting data from both self-recordings and sociolinguistic interviews. Sharma (2011) analyzed phonetic variation among residents of Southall, a majority South Asian London neighborhood, with respect to individual speaker variation by interlocutor. One of the interlocutors was always a sociolinguistic interviewer, and all the others were determined based on self-recordings participants made in the course of their daily lives. Sharma found that some speakers had broad linguistic repertoires, shifting nearly categorically according to audience, while others had more 'fused' repertoires and shifted little. For those who shifted the most, the self-recorded speech samples demonstrated a far wider range of variability than was obtained within the interview speech sample. Several recent studies have found a significant difference between interview speech and self-recorded speech with respect to a wide range of linguistic variables (Tseng 2014; Van Hofwegen 2016), although not always (Saisuwan 2016).

We are still in the early days of understanding if self-recorded data is sufficiently insightful to justify the costs. Levon (2013:211) observed that, "the use of self-recording with semi-informed collaborators does not necessarily lead to significant performative shifts, possibly because the exigencies of the actual interaction tend to dominate." The present study is motivated by the desire to better understand how any effects of self- versus researcher-recording interact with the effects of the interaction and the interlocutor. The current paper builds on an analysis of a single speaker whose vowels were analyzed with respect to both self-recordings and sociolinguistic interview recordings, as well as other tasks (Boyd et al. 2015). The speaker in that study is a speaker of California English participating in the California Vowel Shift (Eckert 2008). Those findings revealed multiple significant differences in relation to task type and the California Vowel Shift, wherein the self-recorded speech shows more advanced productions of many of the vowels that participate in the California Vowel Shift than those seen in laboratory tasks or the sociolinguistic interview. Here, we look at that same speaker's /s/ production, and consider three additional speakers. We take our results to confirm that the stylistic range captured by self-recorded data justifies the extra challenges that the method brings.

2.2 Sociophonetics and /s/ Variation

Sociophonetics research has witnessed a recent rise in the number of studies investigating variation in the production and perception of the sibilant /s/, especially in US English. Since variation in /s/ has been previously modeled with respect to social factors and style/situational factors, it presents a reasonable choice for our study. Sibilant variation is typically studied with respect to segment duration and the first four spectral moments: Center of Gravity (CoG), Standard Deviation, Skewness, and Kurtosis. Studies also often consider Peak Frequency. In the current paper, we use (non-normalized) CoG values for sake of comparison with earlier findings.

Previous work on /s/ in (socio)phonetics has found its realization to be significantly predicted by both speech style and speaker identity. Maniwa et al. (2009) compared 'clear speech' and 'conversational speech' elicited in laboratory conditions among the same speakers, finding a higher

CoG for all fricatives (/s, f, v, θ, z/) in clear speech than in conversational speech. Tucker et al. (2016) analyzed /s/ production in two corpora: TIMIT (read speech) and Buckeye (conversational speech) and found the same pattern across all fricatives: higher CoG in the read speech contexts.

Work on /s/ and speaker identity has largely focused on gender and/or sexual orientation, finding /s/ to be a reliable index both between and within male-identified and female-identified speakers. Many studies suggest that women produce, or are at least expected to produce, /s/ at higher frequencies than men, and that these differences are strongly influenced by social motivations beyond just anatomical differences (Schwartz 1968; Flipsen et al. 1999; Jongman et al. 2000; Stuart-Smith 2007; Fuchs and Toda 2010; Hazenberg 2012). While Flipsen et al. (1999) found women producing /s/ peak frequencies higher than men, other factors can mitigate this effect. For instance, other studies (largely outside of North American contexts) have shown /s/ variability to correlate with speaker age and social class (Stuart-Smith 2007; Levon and Holmes-Elliott 2013; Pharoa et al. 2014). Stuart-Smith (2007) shows younger working class Glaswegian women producing /s/ with a much lower peak frequency than the other women of the study, nearing production values seen in the male participants. She argues that this is not an act of the young working class women associating themselves with the males, but rather is an act of indexing an identity which distances themselves from the middle-class females.

Beyond gender, variation of English /s/ is widely studied in work on sexual orientation. Specifically, many studies show higher frequency /s/, both in production and perception, being strongly correlated with non-normative masculinity and a gay speech style (Linville 1998; Smyth et al. 2003; Munson et al. 2006; Levon 2007, 2014; Smyth and Rogers 2008; Mack and Munson 2012; Zimman 2015, 2017; Podesva & Hofwegan 2016). Furthermore, a higher frequency /s/ has been reported to correlate with listener perceptions of a speaker sounding more feminine (Strand 1999; Munson 2007; Campbell-Kibler 2011)

On the whole, there are fewer studies of /s/ in US English that include or focus on females, as we do here. Saigusa (2016) analyzed the speech of a lesbian celebrity in two televised interviews, and found that the speaker produced /s/ with a higher CoG when speaking with two straight women than she did when speaking with a lesbian interviewer. Podesva and Van Hofwegen (2016) compared /s/ production across a large sample of sociolinguistic interviews with speakers in Redding, California, and found a highly significant difference among straight women on the one hand, with a relatively higher CoG, and trans men, lesbians, trans women, and gay/bi men, on the other, with a relatively lower CoG. There is, to our knowledge, no production study specifically focused on intraspeaker variation in /s/ production among straight, cis-gendered women.

3 Methods

3.1 Speakers

	Born	R'd	Born	IV	Heritage	OtherL1	US Gen	Work
Kat	1986	2016	SF East Bay, CA	Edinburgh	Chinese	Taiwanese Mandarin	2	Asst Prof
Piper	1988	2016	Louisville, KY	Edinburgh	Greek	N/A	3	PhD student
Vicky	1985	2013	San Francisco, CA	San Francisco	Chinese	Shanghainese	2	Lawyer
Virginia	1990	2016	San Antonio, TX	Edinburgh	Mexican	N/A	3	PhD student

Table 1: Speakers analyzed and their social characteristics. ‘R’d’ = year recorded, ‘IV’ = place interviewed, ‘US Gen’ = immigration generation in the United States.

We analyzed four speakers (Table 1). Because all of the relevant comparisons we are making are at an intraspeaker level, the number and type of speakers is incidental. However, we recruited participants who were vaguely socially similar, in order to maintain focus on the effects of style, more

generally, and self-recordings, in particular. Building off of the analysis in Boyd et al. (2015), this meant that all four speakers analyzed here are heterosexual, cis-gender, highly-educated middle class American women born between 1985 and 1990. Due to practical constraints, they are regionally and ethnically diverse, and three of the four were interviewed in Edinburgh, Scotland. While cross-speaker comparisons are not our objective, these facts may become relevant if we consider that those speakers who had been living for several years as expatriates (Piper and Virginia) might have developed a broader sociolinguistic linguistic repertoire than those speakers who were just traveling (Kat) or recorded close to home (Vicky). The three-year time difference between recordings might also have an effect, especially in the sense that the fieldworker (the first author) had spent those three years as an American expatriate as well, possibly resulting in stylistic differences on the part of the interviewer for Vicky than for the other three. Those issues may be explored in future work but will not be seen to impact the results of the present study.

3.2 Tasks & Styles

The present study is part of a larger project considering a range of speech-elicitation methods. In the present paper we compare speakers according to only three of the different task types: interviews, read speech, and self-recordings. The interviews are comprised of standard sociolinguistic questions about participant biography and social life. The read speech is comprised of word lists and seven different reading passages, taken from previous studies (see Boyd et al. 2015). The self-recordings, described in Table 2, all fit the criteria of recordings made without the researchers acting in researcher roles. Most of the time this means that the researcher was not present, although there are some exceptions: in Kat’s *lunch with friends*, the first author was present as a non-researching co-participant, and in Virginia’s *hangout with friends* and *game night*, the second author was present as a non-researching co-participant. In both cases the speaker had been good friends with the respective researcher for several years prior to the recording. The nature of the relationships between the participants in each self-recordings is, of course, different for each speaker, but they are all interactions with friends or family members that they would have had regardless of their participation in this project. All are face-to-face interactions, with all participants in the same room, except for Kat and Virginia’s *Skype* calls. All interlocutors are native speakers of varieties of US English except for Piper’s *party* and Virginia’s *hangout* and *game night*, all of which include mixed groups of Americans and non-Americans of variation nationalities. All interlocutors are white except for Kat’s *lunch*, and Vicky’s *hangout* and *home*, which include one Chinese American interlocutor each, and Virginia’s *Skype*, which includes her Mexican American mother. The gender distribution of interlocutors is indicated in Table 2.

Speaker	SelfRec1	SelfRec2	SelfRec3
Kat	Hangout with friend (♂)	Lunch with friends (♀*,♀)	Skype with friend (♀)
Piper	Party with friends (♀,♀,♀)	[no recording; technical error]	Home with visiting mother (♀)
Vicky	Hangout with friends (♂,♂)	Lunch with friend (♂)	Home with sister (♀)
Virginia	Hangout with friends (♂*,♂)	Game night with friends (♂*,♂,♂,♂)	Skype with parents (♂,♀)

Table 2: Speakers analyzed and the contexts of their self-recordings including the gender of interlocutors present (* = researcher present as co-participant). Skype is an online video chat service; recordings were made with the same recording set-up as all other recordings, not over Skype.

3.3 Coding for /s/ Variation

All audio was transcribed in ELAN and underwent FAVE forced-alignment (Rosenfelder et al. 2014). Segments of speech were not transcribed if they contained any background noise; one recording was excluded entirely for this reason (Piper’s second self-recording; see De Decker 2016). Each instance of /s/ was extracted in Praat (Boersma and Weenink 2017) from the sibilant’s tem-

poral midpoint, getting measurements for CoG, as well as for Peak Frequency and Skewness (although these results are not analyzed here). To temper any slight background noise or influences of co-articulatory voicing from the surrounding segments in more rapid speech, a high-pass filter was implemented at 1,000 Hz (see Stuart-Smith 2007; Zimman 2015, 2017). Prior to analysis, we removed all tokens less than 30ms, all /STR/ clusters (as these are known to retract; Baker et al. 2011), and sibilants which obviously failed Peak Frequency tracking, and then adjusted the dataset within ± 2 standard deviations, resulting in 6726 tokens. Kat is represented by the most tokens (N=2202), followed by Virginia (N=1614), then Vicky (N=1465), and then Piper (N=1445).

3.4 Statistical Models

Best-fit models for /s/ CoG were obtained by speaker, rather than across the whole subsample, because the effect of task operates differently for each speaker (Figure 1), and because data were not normalized for between-speaker comparisons. Internal factors include /s/ DURATION (*ms*), /s/'s LOCATION (*onset, medial, final*), the FOLLOWING and PRECEDING phonological context (*apical, dorsal, non-lingual*), and the speech TASK (*interview, reading, selfrec1, selfrec2, selfrec3*). These all obtained significance in the final models except DURATION and LOCATION for Virginia. Final models were obtained through by-hand drop-one ANOVA comparisons of mixed-effect linear regression models built with lme4 (Bates et al. 2014). Models include a random intercept of word.

4 Results

Overall, we find that the contrast between self-recordings and interviews can be a reliable predictor of phonetic variation, and that the effect and its direction varies by speaker and by self-recording. Each speaker differs from the others in her production of /s/, at least with respect to non-normalized Center of Gravity values (Figure 1). Virginia's /s/ has the highest average CoG and Vicky's the lowest. While there may be interesting ways to explain these differences between speakers, we do not pursue this here given the size of the sample, other than to note that Kat and Vicky are the two Chinese Americans from the San Francisco Bay Area, while Virginia and Piper are the two expatriates in Scotland, and neither pair patterns together. Here, we focus on each speakers' differences across tasks, specifically the comparison of self-recording and interview speech. All speakers show a consistent and significant difference between their read speech and speech in interaction with the first author (Figure 1; Table 3). The difference goes in the expected direction: /s/ is produced with a higher center of gravity in the read contexts (Tucker et al. 2016).

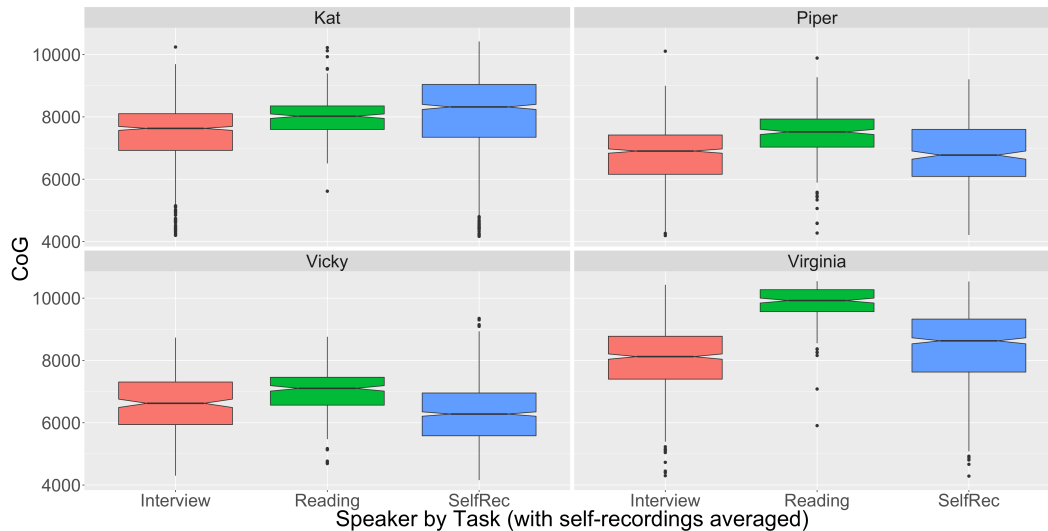


Figure 1: /s/ Center of Gravity by speaker and by task, averaging across self-recordings.

The differences and direction of differences for self-recorded speech are less consistent (Figure 1).

For Virginia and Kat, self-recordings appear to have elicited /s/ productions with higher CoGs than interview speech, while for Vicky they have elicited the opposite, and Piper appears to show little difference between her self-recordings and interview speech. Taking a closer look at each separate self-recording in Figure 2 and Table 3, we find that /s/ production in all but one of the self-recordings is significantly different from the interview data, and that the direction and strength of the effect varies between different self-recordings for some individuals.

For Kat, all three self-recordings resulted in high CoG values relative to her interview speech; two of these surpass the shift observed for her read speech. For Piper, one of her self-recordings resulted in her highest CoG values, surpassing the contrast between her interview speech and read speech, while her other self-recording resulted in her lowest CoG values of any task. For Vicky, all three of her self-recordings show lower CoG values than her interview (or read) speech. For Virginia, one of her self-recordings does not show significantly different CoG values from the interview speech, while the other two result in significantly higher CoG values, although neither effect is as strong as the effect of read speech.

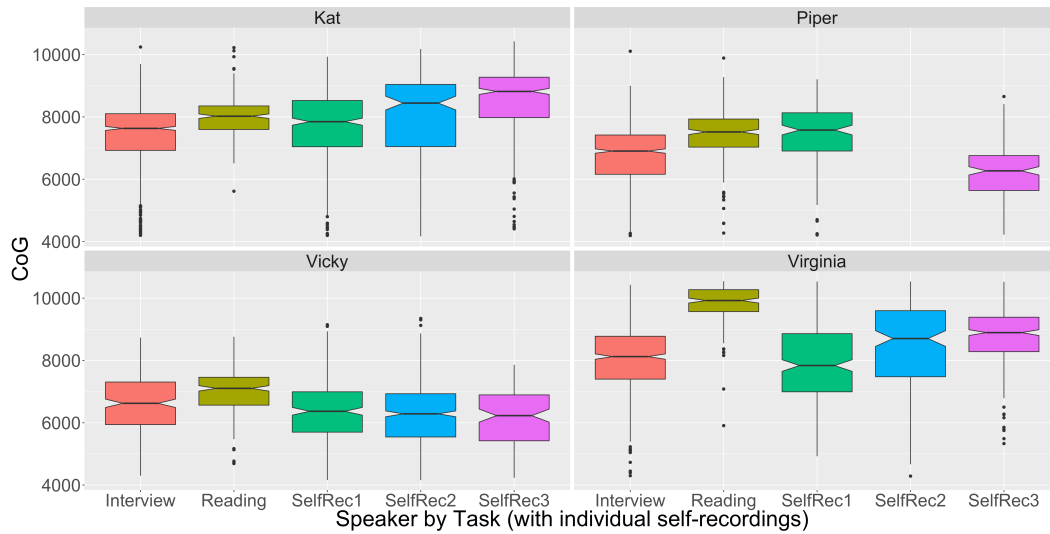


Figure 2: /s/ Center of Gravity by speaker and by task, comparing across self-recordings.

Speaker = KAT	Estimate	Std. Error	t-value	p-value
(Intercept)	7218.21735	87.26277	82.718177	< 0.001
Duration	3127.22323	421.04286	7.4273276	< 0.001*
Location-in-Word = Medial	12.41971	79.00359	0.1572044	0.875
Location-in-Word = Onset	261.17321	77.18588	3.3836916	0.001*
Following = dorsal	-163.07858	60.14621	-2.7113693	0.007*
Following = non-lingual	-247.85651	78.45236	-3.159325	0.002*
Preceding = dorsal	-134.19272	56.10189	-2.3919464	0.017*
Preceding = non-lingual	-183.86749	77.72885	-2.3654988	0.018*
Task = Reading	436.78552	76.12569	5.7376886	< 0.001*
Task = SelfRec1	281.60332	63.37886	4.443174	< 0.001*
Task = SelfRec2	575.35949	82.56067	6.9689294	< 0.001*
Task = SelfRec3	1040.53558	62.70408	16.5943827	< 0.001*
Speaker = PIPER	Estimate	Std. Error	t-value	p-value

(Intercept)	6494.7921	84.698	76.6817605	< 0.001
Duration	3823.0225	429.48334	8.9014454	< 0.001*
Location-in-Word = Medial	52.5059	76.32485	0.6879266	0.491
Location-in-Word = Onset	276.0906	71.4327	3.865045	< 0.001*
Following = Dorsal	-241.331	58.0677	-4.1560274	< 0.001*
Following = Non-lingual	-377.8901	76.01113	-4.9715106	< 0.001*
Preceding = Dorsal	-235.3765	55.87978	-4.2121944	< 0.001*
Preceding = Non-lingual	-144.5988	80.41183	-1.7982279	0.072
Task = Reading	638.2647	61.46587	10.3840513	< 0.001*
Task = SelfRec1	651.2551	77.23254	8.4323927	< 0.001*
Task = SelfRec3	-579.7256	73.60959	-7.8756806	< 0.001*
Speaker = VICKY	Estimate	Std. Error	t-value	p-value
(Intercept)	6339.82485	105.32786	60.1913408	< 0.001
Duration	4101.20415	575.78761	7.1227725	< 0.001*
Location-in-Word = Medial	-82.88436	81.6993	-1.0145051	0.310
Location-in-Word = Onset	244.21619	76.40984	3.1961353	0.001*
Following = Dorsal	-282.00802	63.54771	-4.4377373	< 0.001*
Following = Non-lingual	-71.84763	76.158	-0.9434024	0.345
Preceding = Dorsal	-70.77034	57.54367	-1.2298544	0.219
Preceding = Non-lingual	-217.22163	81.92689	-2.6514082	0.008*
Task = Reading	348.37296	81.40847	4.2793207	< 0.001*
Task = SelfRec1	-220.57123	78.34473	-2.8153932	0.005*
Task = SelfRec2	-329.3753	69.25914	-4.7556942	< 0.001*
Task = SelfRec3	-396.91716	99.23847	-3.9996299	< 0.001*
Speaker = VIRGINIA	Estimate	Std. Error	t-value	p-value
(Intercept)	8436.6769	66.87038	126.164623	< 0.001
Following = Dorsal	-281.4723	69.03406	-4.077296	< 0.001*
Following = Non-lingual	-529.9732	83.43567	-6.351878	< 0.001*
Preceding = Dorsal	-159.7462	65.49549	-2.439041	0.015*
Preceding = Non-lingual	-126.9851	83.92982	-1.512992	0.130
Task = Reading	1661.9705	87.07723	19.086167	< 0.001*
Task = SelfRec1	-152.0535	78.54809	-1.935801	0.053
Task = SelfRec2	241.2991	88.70445	2.72026	0.007*
Task = SelfRec3	654.8792	69.6696	9.399784	< 0.001*

Table 3: By-speaker best-fit mixed effect regression models of /s/ Center of Gravity, with WORD as random intercept. Asterisks indicate significance at 0.05 alpha.

5 Discussion

Based on data from four speakers, each recorded in a sociolinguistic interview with the first author

and two-to-three self-recordings with friends and family, we find that the contrast between interview speech and self-recorded speech is a strong predictor of sibilant variation, sometimes eliciting phonetic shifts just as large or larger than those found for the classic stylistic contrast between interview speech and read speech. We can also see that the direction and strength of the effect varies by individual and by self-recording context. This is unsurprising, given previous work on the stylistic diversity of self-recording contexts and their linguistic correlates (Podesva 2007; 2011a; 2011b; Sharma 2011). Here we see suggestive evidence to support Levon's (2013) observation that self-recordings do not necessarily lead to significant differences, and that the situation of recording is the ultimate predictor. For example, both Kat's and Virginia's third self-recording resulted in their highest Center of Gravity values for /s/, and both of these recordings were made while they were conducting Skype calls, which perhaps motivated clearer articulations than would be found for offline, face-to-face communication. The fact that /s/ variation in Virginia's first self-recording does not significantly differ from her interview data is the exception that proves the rule: it shows that while self-recordings do not necessarily result in style shifts, they are quite likely to do so. Indeed, it is for this reason that Meyerhoff et al. (2015) specifically caution researchers on the use of the self-recordings, since the different styles obtained by different speakers make cross-speaker comparisons tricky. We agree that self-recordings are most fit for studies of intraspeaker variation. Beyond this, we argue that self-recordings are ideal for any study that seeks to obtain a basic description of the range of variation available to a speaker, since they seem to be a vehicle for obtaining evidence of speaking patterns not easily obtainable by traditional sociolinguistic interview methods. This possibility of greater descriptive accuracy would seem to be a good motivation to justify taking on the ethical, logistical, and analytic hurdles inherent in obtaining self-recorded speech data.

The data here further suggest that the situation can have more of an impact on intraspeaker patterns of /s/ variation than the interlocutor. Although we did not analyze the effect of interlocutor in the models, we do have several self-recordings where one of the two researchers was present as a non-researching co-participant: the first author was present for Kat's second self-recording, and the second author was present for Virginia's first and second self-recordings. There does not seem to be any obvious effect of these facts in our data: for example, Kat's self-recording with the first author present is still significantly different from the interview speech, and to a greater extent than in her first self-recording. There also does not seem to be any clear effect in these data of interlocutor gender, age, ethnicity, or nationality, although these were not balanced across speakers and recordings. We have already noted that the four speakers themselves differ from one another with respect to their average /s/ CoG, but that this difference is not readily explained by any demographic characteristics distinguishing them from one another.

The results for /s/ Center of Gravity dovetail with previous findings for seven of twelve vocalic variables for the speaker we call Vicky (four of six F1 variables and three of six F2 variables; Boyd et al. 2015). And as one of few studies to demonstrate robust style-shifting effects for sibilants (see also Maniwa et al. 2009, Saigusa 2016), our results have implications for the growing body of work on sibilant variation in North American English (e.g., Linville 1998, Flipsen et al. 1999, Smyth et al. 2003, Levon 2007, Campbell-Kibler 2011, Mack and Munson 2012, Podesva and Van Hofwegen 2016, Zimman 2013). This is potentially important given the fact that it is at times difficult to compare across studies because of their stylistic differences.

6 Conclusion

Self-recordings have featured in several recent studies on style-shifting and phonetic variation (e.g., Podesva 2007; Sharma 2011). Such recordings are attractive for potentially eliciting a wider range of styles than can be recorded in typical sociolinguistic interviews. However, obtaining self-recordings also presents ethical, logistical, and analytic challenges inherent in training a participant to be a temporary fieldworker. Perhaps for this reason they are not a standard part of the variationist toolkit, and before now we have lacked systematic studies comparing the phonetic differences between self-recorded speech and speech obtained from sociolinguistic interviews. The results from our study suggest that the classic sociolinguistic interview may capture a relatively narrow range of potential phonetic variation. Speech in self-recordings may be different enough from speech in interviews to justify overcoming the practical challenges of its collection and integrating

it into standard sociolinguistic methodology.

References

- Baker, Adam, Diana Archangeli, and Jeff Mielke. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23:347-374.
- Bates, Douglas, Martin Maechler, Ben Bolker and Steven Walker. 2014. lme4: linear mixed-effects models using Eigen and S4. R package [Computer program]. Version 1.1-7, retrieved March 2015 from <http://CRAN.R-project.org/package=lme4>.
- Boersma, Paul & Weenink, David. 2017. Praat: doing phonetics by computer [Computer program]. Version 6.0.28, retrieved March 2017 from <http://www.praat.org/>.
- Boyd, Zac, Zuzana Elliott, Josef Fruehwald, Lauren Hall-Lew, and Daniel Lawrence. 2015. An Evaluation of Different Sociolinguistic Elicitation Methods. In *Proceedings of the 18th International Congress of Phonetic Sciences*, ed. by The Scottish Consortium for ICPhS 2015. The University of Glasgow.
- Campbell-Kibler, Kathryn. 2011. Intersecting variables and perceived sexual orientation in men. *American Speech* 86:52-68.
- Coupland, Nikolas. 1980. Style-shifting in a Cardiff work-setting. *Language in Society* 9:1-12.
- De Decker, Paul. 2016. An evaluation of noise on LPC-based vowel formant estimates: Implications for sociolinguistic data collection. *Linguistics Vanguard* 2:83-101.
- Eckert, Penelope. 2008. Where do ethnolects stop? *International Journal of Bilingualism* 12:25-42.
- Flipsen, Peter, Lawrence Shriberg, Gary Weismer, Heather Karlsson, and Jane McSweeney. 1999. Acoustic characteristics of /s/ in adolescents. *Journal of Speech, Language, and Hearing Research* 42:663-677.
- Fuchs, Susanne and Martine Toda. 2010. Do differences in male versus female /s/ reflect biological or socio-phonetic factors. In *Turbulent Sounds: An Interdisciplinary Guide*, ed. by S. Fuchs, M. Toda, and M. Zygis, 281-302. Walter de Gruyter.
- Hazenberg, Evan Nicholas Leo. 2012. Language and Identity Practice: A Sociolinguistic Study of Gender in Ottawa, Ontario. Masters thesis, Memorial University of Newfoundland.
- Hindle, Donald Morris. 1979. The Social and Situational Conditioning of Phonetic Variation. Doctoral dissertation, University of Pennsylvania.
- Jongman, Allard, Ratreay Wayland, and Serena Wong. 2000. Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America* 108:1252-1263.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania.
- Lee, Sarah. 2017. Style-Shifting in Vlogging: An Acoustic Analysis of “YouTube Voice”. *Lifespans & Styles: Undergraduate Working Papers on Intraspeaker Variation* 3:28-39.
- Levon, Erez. 2014. Categories, stereotypes, and the linguistic perception of sexuality. *Language in Society*, 43:539-566.
- Levon, Erez. 2013. Ethnography and recording interaction. In *Research Methods in Linguistics*, ed. by R.J. Podesva and D. Sharma, 195-215. Cambridge: Cambridge University Press.
- Levon, E. 2007. Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society* 36:533-554.
- 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: Selected Papers from NAWAV 41*, ed. by A. Freeman.
- Linville, Sue Ellen. 1998. Acoustic correlates of perceived versus actual sexual orientation in men’s speech. *Folia Phoniatrica et Logopaedica* 50:35-48.
- Mack, Sarah, & Munson, Benjamin. 2012. The association between /s/ quality and perceived sexual orientation of men’s voices: implicit and explicit measures. *Journal of Phonetics* 40:198-212.
- Mallinson, Christine, Becky Childs, and Gerard Van Herk. 2013. *Data collection in sociolinguistics: methods and applications*. Routledge.
- Maniwa, Kazumi, Allard Jongman, & Travis Wade. 2009. Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America* 125:3962-3973.
- Meyerhoff, M., Schlee, E., & MacKenzie, L. 2015. *Doing sociolinguistics: A practical guide to data collection and analysis*. Routledge.
- Munson, Benjamin. 2007. The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and Speech* 50:125-142.
- Munson, Benjamin, Elizabeth C. McDonald, Nancy L. DeBoe, and Aubrey R. White. 2006. The acoustic and perceptual bases of judgments of women and men’s sexual orientation from read speech. *Journal of Phonetics* 34:202-240.

- Pharao, Nicolai, Marie Maegaard, Janus Spindler Møller, and Tore Kristiansen. 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* 11:478–504.
- Podesva, Robert J. 2011a. Saliency and the Social Meaning of Declarative Contours: Three Case Studies of Gay Professionals. *Journal of English Linguistics* 39:233–264.
- Podesva, Robert J. 2011b. The California Vowel Shift and Gay Identity. *American Speech* 86:32–51.
- Podesva, Robert J., and Janneke Van Hofwegen. 2016. /s/sexuality in Smalltown California: Gender Normativity and the Acoustic Realization of /s/. In *Language, Sexuality, and Power*, ed. by E. Levon and R.B. Mendes, 168–188. Oxford: Oxford University Press.
- Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, Scott Seyfarth, Kyle Gorman, Hilary Prichard, and Jiahong Yuan. 2015. FAVE (Forced Alignment and Vowel Extraction) 1.1.3. ZENODO. doi:10.5281/zenodo.9846.
- Saigusa, Julie. 2016. Jane Lynch and /s/: The Effect of Addressee Sexuality on Fricative Realization. *Lifespans & Styles: Undergraduate Working Papers on Intraspeaker Variation* 2:10–16.
- Saisuwan, Pavadee. 2016. Male Femininity in Thai Among Men who Identify with Non-Normative Male Roles. Doctoral dissertation, Queen Mary University of London.
- Schneider, Edgar W. 2016. World Englishes on YouTube: Treasure trove or nightmare?. In *World Englishes: New theoretical and methodological considerations*, ed. by E. Seoane and C. Suárez-Gómez, 253–282. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Schøning, Signe, and Janus Spindler Møller. 2009. Self-recordings as a social activity. *Nordic Journal of Linguistics* 32:245–269.
- Schwartz, Martin F. 1968. Identification of speaker sex from isolated, voiceless fricatives. *The Journal of the Acoustical Society of America* 43:1178–1179.
- Sharma, Devyani. 2011. Style repertoire and social change in British Asian English. *Journal of Sociolinguistics* 15:464–492.
- Smyth, Ron, Greg Jacobs, and Henry Rogers. 2003. Male voices and perceived sexual orientation: An experimental and theoretical approach. *Language in Society* 32:329–350.
- Smyth, Ron, and Henry Rogers. 2008. Do gay-sounding men speak like women? *Toronto Working Papers in Linguistics* 27.
- Strand, Elizabeth A. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18:86–100.
- Stuart-Smith, Jane. 2007. Empirical evidence for gendered speech production: /s/ in Glaswegian. In *Laboratory phonology 9*, ed. by J. Cole and J.I. Hualde, 65–86. New York: Mouton de Gruyter.
- Tseng, Amelia. 2015. Vowel Variation, Style, and Identity Construction in the English of Latinos in Washington, DC. Doctoral dissertation, Georgetown University.
- Tucker, Benjamin, Viktor Kharlamov, and Daniel Brenner. 2016. What's the Zed? The Acoustics of Conversational Fricatives in Mid-Western American English. Paper presented at NorthWest Phonetics & Phonology Conference 2016. University of Oregon, Eugene, Oregon.
- Van Hofwegen, Janneke. 2016. A day in the life: What self-recordings reveal about “everyday” language. Paper presented at NWAV 45, Vancouver, BC.
- Wagner, Petra, Jürgen Trouvain, and Frank Zimmerer. 2015. In defense of stylistic diversity in speech research. *Journal of Phonetics* 48:1–12.
- Wilson, John. 1987. The Sociolinguistic Paradox: Data as a Methodological Product. *Language & Communication* 7:161–177.
- Zimman, Lal. 2017. Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/. *Language in Society* 1–32.
- Zimman, Lal. 2015. Transmasculinity and the voice: Gender assignment, identity, and presentation. In *Language and Masculinities: Performances, Intersections, Dislocations*, ed. by T. Milani, 197–219. Routledge, New York.
- Zimman, Lal. 2013. Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men. *Journal of Language and Sexuality* 2:1–39.

Linguistics & English Language
 University of Edinburgh
 3 Charles Street
 Edinburgh EH8 9AD
 United Kingdom
 Lauren.Hall-Lew@ed.ac.uk
 Z.C.Boyd@sms.ed.ac.uk