Evaluating S(c)illy voices: The effects of salience, stereotypes, and co-present language variables on real-time reactions to regional speech

Chris Montgomery, Emma Moore

EVALUATING S(C)ILLY VOICES: THE EFFECTS OF SALIENCE, STEREOTYPES, AND CO-PRESENT LANGUAGE VARIABLES ON REAL-TIME REACTIONS TO REGIONAL SPEECH

CHRIS MONTGOMERY   EMMA MOORE

University of Sheffield   *University of Sheffield*

This article explores the relationship between salience, stereotypes, and cooccurring language variables in the social perception of language. Following previous work, we argue that sociolinguistic perception is dependent upon the ability of listeners to map the linguistic cues contained in a speech signal to stereotypes. However, we contend that the understanding of which language features contribute to those stereotypes, and how they do so in the specific context of talk, has been limited because of the tendency to focus on preselected variables and to control for the context in which they occur. We advance an account of the role of stereotypes in the social perception of language by using a new tool for capturing, visualizing, and querying listeners' real-time reactions to voice samples. Our survey instrument collects reactions to two topically distinct guises from the same speaker (taken from the Scilly Voices corpus), both of which contained a similar number of regionally distinctive accent features. As our survey instrument includes a review function enabling listeners to provide information on why certain features were notable to them, we are able to interrogate listeners' ability to respond to unspecified linguistic features. Ultimately, this enables us to build a more nuanced account of the interaction between a range of linguistic features and their relationship to message content, and allows us to demonstrate that both do evaluative and perceptual work.

Our findings have important implications for those interested in understanding the situated meaning of linguistic features and, in particular, how researchers might continue to develop exemplar models of the ways in which social information is indexed to linguistic features. We argue that no experiment can be context-free and, as a result, researchers must consider ways of modeling the effects of co-present variants on a given exemplar, not just the social indices of specific exemplars themselves.*

Keywords: perception, real-time reactions, salience, language regard, social meaning, stereotypes

1. INTRODUCTION. Interest in nonspecialists' beliefs about and reactions to language use has been long-standing and widespread (see, for example, Lambert et al. 1960, Giles 1970, Giles & Powesland 1975, Ryan & Giles 1982, Preston 1989, Niedzielski & Preston 2003, Clopper & Pisoni 2004a,b, Garrett 2010). Early research on the perception of dialect features tended to elicit nonspecialists' beliefs and reactions to regional accents, with the principal findings demonstrating disparities between perceived 'standard' and 'nonstandard' varieties. This research points to a general, and well-established, pattern linking linguistic standardness to intelligence, and (some kinds of) linguistic nonstandardness to social attractiveness (e.g. Paltridge & Giles 1984:71). Such studies demonstrate that listeners 'can and do make a number of attitudinal judgments about a talker based on his or her speech' (Clopper & Pisoni 2002:273).

More recently, however, attention has turned to isolating particular accent features considered to have 'salience'. As noted by Nycz (2016) and Drager and Kirtley (2016: 12), in linguistics, the meaning of 'salience' is contested, although '[i]n sociolinguistics, salience is tied with both noticeability and awareness of sociolinguistic variables'. No-

ticeability, according to Nycz (2016:64), is the 'conscious awareness and subjective experience of a linguistic feature; it is a step beyond mere perception of a feature'. As the review by Drager and Kirtley (2016) outlines, noticeability has been explained by 'localizedness', linguistic prominence due to factors such as pitch and position, and a feature's unexpectedness or 'surprisal value' (Rácz 2013:51). All of these dimensions point toward the idea that 'salience is the degree to which something stands out relative to other, neighbouring items' (Drager & Kirtley 2016:12, citing Hogg & Vaughan 2008:61).

Research dealing with the salience of particular accent features has tended to be of two kinds. The first focuses on success at voice placement. For instance, Williams, Garrett, and Coupland (1999) investigated factors contributing to dialect recognition in Wales, speculating that some of their recordings might have contained 'more, or more salient, phonological cues than others did' (1999:353), which therefore enable better recognition rates. Although they do not generally discuss which features these might be, they mention the raised /aː/ shared by the Cardiff and north-west Wales speakers. Similarly, but more explicitly, Plichta and Preston (2005) examined the role played by the degree of /ay/ monophthongization in perceptions of north and south in the United States. Also in the US, Clopper and Pisoni (2004b:131) found some features to be better than others at predicting speaker location when respondents were asked to categorize six dialects of US English on the basis of eleven variables. Extending the analysis to native and nonnative speakers, Clopper and Bradlow (2009:444) found that it was categorical consonant features (i.e. the presence or absence of rhoticity) that were the 'significant predictors of performance' in relation to assigning speakers to a US region, irrespective of listener nativeness. Consonants were also found to be more important than vowels when children were asked to categorize speakers into groups in a study of contrasts between Home dialect (Midland American English), a regional variant (British English), and a second-language variant (Indian English) (Wagner, Clopper, & Pate 2014:1080).

The second kind of research on salience has focused on the effects specific features have on the social evaluation of speakers. For instance, the effect of (-ing) on a speaker's suitability for a job as a newsreader has been examined in the US (Labov et al. 2011) and the UK (Levon & Fox 2014). In both experiments, listeners were presented with a number of guises manipulated to increase or decrease tokens of /n/ and /ŋ/. Labov and colleagues (2011:440) found a significant effect of even a small increase in the percentage of /n/ in the guise, with 'professionalism' ratings decreasing dramatically with the increase of the nonstandard /n/ form. Levon and Fox were not able to replicate these findings in Great Britain. However, they also investigated responses to (th)-fronting (the fronting of /θ/ and /ð/ to [f] and [v], respectively, a feature said to be spreading through Great Britain by Kerswill (2003)) and found that it showed some effects for northern English respondents (Levon & Fox 2014:205). However, it should be noted that these findings did not replicate the logarithmic pattern found by Labov and colleagues (2011).

Other research on (-ing) has highlighted the interaction between a linguistic variant and the larger social style in which it occurs. For instance, Campbell-Kibler (2009, 2010) observed that listeners' perceptions of speakers' regional and class backgrounds strongly condition the correlation between high use of the alveolar variant and high intelligence/education levels. Similarly, Pharao and colleagues (2014) show that the indexical meanings of alveolar versus fronted (s) ([s] versus [s+]) differ dependent upon whether the features occur in 'modern Copenhagen speech' or 'street language', with [s+] more strongly indexing femininity and gayness in the former. While Pharao and colleagues (2014) focus on the effects of register, Podesva, Reynolds, Callier, and Bap-

tiste (2015) have demonstrated that the social meanings of a feature may also be affected by the perceptions of how a speaker 'usually' talks. They examine released /t/ in the talk of US politicians and show that, contrary to other findings, Barack Obama is rated as more intelligent in guises where he does not release /t/ because the (predominantly Democrat) listeners sampled consider this to more closely approximate how he usually talks.

Studies that prime listeners with social information have provided further support for the interaction of social factors and perception. For example, Hay, Nolan, and Drager (2006) have shown that listeners process speech signals differently based on what they believe about the speakers' age, gender, and class. Both Niedzielski (1999) and Hay and colleagues (2006) have shown that priming listeners with geographical labels also affects their perception of the vowel space. Hay and Drager (2010) have even suggested that the presence of soft toys that are readily associated with national identities may influence the perception of voice samples. More recently, D'Onofrio (2015) has demonstrated that what listeners are told about a speaker's social persona can impact the way they react to a Californian 'Valley girl' voice sample.

While this research points to the importance of the larger social style in which specific linguistic features occur, as our review of the literature suggests, the tendency has been for research to consider one linguistic item in isolation. This is problematic, given recent production studies that suggest the importance of 'clusters' of linguistic variants on the situated social meaning of language (for instance, Eckert 2008, Moore & Podesva 2009, Sharma 2011) and Preston's work on language regard (e.g. Preston 2015), which argues that linguistic features are noticed and classified differently in different situations. Consequently, recent work has attempted to show how multiple features interact to influence perception. For example, Campbell-Kibler (2011) considers how pitch, (s)-fronting or backing, and (ING) interact to affect perceptions of masculinity; Levon (2014) examines whether different combinations of variables compatible with stereotypes about gender (pitch), sexuality (sibilance), and class ((th)-fronting) result in different listeners' perceptions; and Pharao and colleagues (2014) explore how sibilance interacts with rhythm/ prosodic frame to affect the indexical meanings of fronted (s).

While these studies provide clear evidence of multiple features interacting to affect perception, the analysis provides only a partial account of the contextual frame in which those features occur. This is not to say that the features selected for analysis are not well justified—they usually are (see, for instance, Levon (2014:546), who provides a detailed rationale of why pitch, sibilance, and (th)-fronting are pertinent to his interest in gender, sexuality, and class, respectively).[1] Nonetheless, these studies do not account for the features that are present in the data samples but not selected for analysis. Of course, the presence or absence of unexamined features is 'controlled' across samples used in these studies. However, controlling for the presence or absence of a particular feature in a stretch of discourse does not account for the possibility that awareness of this feature may change when it occurs in the alternative guises. For instance, imagine two samples of talk constructed using the sentence *The north wind and the sun had a quarrel about which of them was stronger*: sample 1, which includes fronted (s) and has a relatively high average fundamental frequency, and sample 2, which includes fronted (s) and has a relatively low average fundamental frequency. We can draw conclusions

---

[1] Although also see Levon and Fox's (2014) study. This draws upon Labov et al. 2011, but ultimately demonstrates the lack of salience of (-ing) in a British context, suggesting that this was not a pertinent feature for analysis.

about how these forms interact, and their subsequent social meanings, on the basis of how listeners respond differently to the two samples. However, what we cannot know is whether fronted (s) and a relatively high average fundamental frequency in sample 1 made other aspects of the discourse more or less salient in this sample than they were in sample 2, and whether this had an effect on the overall evaluation of the sample. For instance, it could be that an instance of /h/-retention in a word that regularly features /h/-dropping (such as the word *had* in the sentence proposed in this hypothetical scenario) has differential salience on the basis of its occurrence in sample 1 or sample 2. Given that 'salience is the degree to which something stands out relative to other, neighbouring items' (Drager & Kirtley 2016:12, citing Hogg & Vaughan 2008:61), as we noted earlier, we might expect a change in one or two variables to influence the salience and subsequent effect of other adjacent forms.

Providing only a partial account of the wider linguistic style in which analyzed features occur has implications for our understanding of the salience of linguistic features and their relationship to stereotypes. We use the term 'stereotype' here, rather than something more generic like 'social meaning', because several researchers have suggested that the social perception of language is dependent upon the ability of speakers to map the linguistic cues contained in a speech signal to stereotypes (Hay et al. 2006, Campbell-Kibler 2008, 2016, Levon 2014, Drager & Kirtley 2016).[2] This research has been used to support exemplar-based models of speech production and perception. Drager and Kirtley (2016:2), who provide a detailed review of work relating to EXEMPLAR THEORY, note that it is 'a collection of cognitive models in which experiences are encoded in the mind as episodic memories, known as exemplars' (see also Brooks 1978, Johnson 1997, Pierrehumbert 2001). The theory is attractive to sociolinguists because it proposes that social information is indexed to linguistic exemplars, making sociolinguistic variation a natural outcome of the model. It also offers an explanation of how speech perception works, given that incoming speech is believed to activate associated exemplars and their related social indices. Of course, these indices may be multiple, given that the social meanings of linguistic features shift according to the contexts in which they occur. Consequently, research using exemplar-based models has suggested that contextual factors determine precisely which social meaning is activated at any one point in time (e.g. Hay & Drager 2010). They do so because contextual factors trigger stereotypical reactions to speech, which in turn 'frame the listeners' evaluations of fine-grained sociolinguistic variation' (Levon 2014:540). Put another way, we might assume that 'all research on perception is also research on metalinguistic characterization' (Campbell-Kibler 2016:130).

The cues that generate stereotypes can include the specific combination of linguistic features heard and message content, among other factors (see Campbell-Kibler 2016: 130). Crucially, then, to understand sociolinguistic perception, we need to understand at least three key components: (i) which stereotypes are generated by a stretch of discourse, (ii) which language features contribute to those stereotypes, and (iii) how they do so in the context of the specific content of the talk. Accessing (i) is relatively straightforward, given that it has, in various ways, long been a goal of sociolinguistic research. However, we contend that the understanding of (ii), and its interaction with (iii), has been limited by experimental methodologies that make assumptions about

---

[2] We follow Levon (2014:544) in defining stereotypes as 'a form of associative social knowledge that serve to encode popular ideologies about social groups by linking the perceptual activation of a group concept [e.g. *man*] with the activation of relevant trait attributes [e.g. *athletic*, *domineering*] and roles [e.g. *father*]'.

which linguistic features are relevant to perception. As Levon (2014:540) has noted, production studies are increasingly providing more nuanced accounts of the social meaning of language variation which reflect the 'indexical mutability' of language (Eckert 2012:94). If we want our perception studies to do the same, it will be necessary to capture more information about the relationship between variables, and the ways in which their ideological loads shift in given contexts of talk.

1.1. CAPTURING LISTENER RESPONSES IN REAL TIME. It is necessary to find ways to record how listeners react to linguistic variants as they hear them if we are to demonstrate that the contribution a variant makes to social meaning changes according to its precise placement and the context in which it occurs. Capturing listeners responding in real time makes it possible to gather responses to multiple linguistic features, and to be sure which feature is having an effect on perception (and—importantly—when it does or does not have that effect). Previous research has been limited in its ability to identify the exact point at which listeners make an evaluation of what they hear. Nonetheless, there have been some attempts to capture real-time reactions to voice samples. Labov and colleagues (2011) conclude their article by mentioning the ability of one of their experiments to capture real-time reactions, although no in-depth analysis of this is presented. However, Watson and Clark (2013) focus exclusively on real-time reactions. They investigated reactions to the NURSE-SQUARE merger in north-west England using an online interface. This enabled listeners to move a slider on the screen when they believed the sample to sound more or less 'posh'. Watson and Clark found 'a fairly high proportion of identified change points cluster at or around the same time as an occurrence of NURSE or SQUARE' (2013:315).[3] Although they do not provide an in-depth analysis of the contexts in which their tokens occur, they observe that their findings suggest 'a particular linguistic feature can be more or less salient depending on both the local social context and the micro-linguistic context in which it appears' (Watson & Clark 2013:321).

As noted above, much previous work uses manipulated and spliced linguistic tokens to control for the effect of certain phonetic variables. Building on their 2013 article, however, Watson and Clark (2015) examine real-time reactions to unmanipulated regional speech. British voice samples from Cambridge, Cardiff, Dublin, Liverpool, and Newcastle upon Tyne were presented to listeners, who were asked to react to what they heard using two sliders, indicating how 'posh' and how 'friendly' each speaker was (Watson & Clark 2015:45). As with their previous study, however, they experienced difficulties in assessing whether 'change points correlate with particular linguistic features [because] … these features do not occur in isolation of others' (2015:55). They note that nonstandard linguistic features often cluster together in talk and that unpicking the perceptual prominence of individual features is not a straightforward task.

In this article, we attempt to address the limitations we have identified in previous studies by presenting a new technique for capturing real-time reactions to vocal stimuli. In particular, we aim to demonstrate that it is possible to measure the effects of multiple linguistic variants in real time. Furthermore, in examining a number of features at once, we also suggest that it is possible to better capture the effects of multiple language features (as opposed to observing the effects of linguistic features that happen to be of in-

---

[3] 'Change points' relate to a statistical technique used by Watson and Clark in their 2013 and 2015 articles. It is 'used to identify the points at which the statistical properties of a time-ordered data set change' (Watson & Clark 2013:311). Change-point analysis is not used in this article due to the difference in the type of data gathered by our survey instrument.

terest to the linguist), and to record the effects of individual features within a broader style. Finally, drawing upon research into the social meaning of linguistic features, we also consider how extralinguistic information affects perception by examining the ways that topic and discourse content interact with the occurrence of particular linguistic features to produce different evaluative responses from listeners.

Our analysis uses data from the Scilly Voices project. This project has been examining language variation and change on the Isles of Scilly, a group of islands off the southwest coast of England, using data from the Isles of Scilly Museum's Oral History Archive. The archive dates from the 1970s onward and is a series of recordings made by local people interviewing other local people (the archive can be searched online at https://www.dhi.ac.uk/scillyvoices/). We are particularly motivated to explore sociolinguistic perceptions of this variety given a mismatch between historical metalinguistic commentary on the variety and recent production analyses. Historical texts across time emphasize the standard nature of the Scillonian dialect, as can be seen from the following quotations.[4]

> (1) … the Language of Scilly refines upon what is spoken in many Parts of Cornwall; probably from the more frequent Intercourse of the Inhabitants, some more than others, with those who speak the Standard English best … (Heath 1750:436)
> (2) The Islanders are remarkable for speaking good English—far preferable, at least, to what is generally heard amongst the humbler classes of any county, at some distance from the metropolis … (Woodley 1822:105)
> (3) … in the Isles, all dialect has been educated out. (Ellis 1890:41)
> (4) The English spoken today (1979) by natives of the Isles of Scilly … is scarcely removed from Standard (southern) English, using a slightly modified 'received pronunciation' (R.P.) as of educated persons. (Thomas 1979:109)

While these quotations suggest that the Scillonian dialect differs substantially from adjacent mainland varieties, Moore and Carter (2015, 2018) demonstrate that the variety clearly shares vernacular features with dialects in East and West Cornwall. Consequently, we explore if context affects whether listeners are able to perceive vernacular features in Scillonian speech; which features, or groups of features, generate listener responses; and what factors affect the way this specific variety is evaluated.

To summarize, our article has two main aims: (i) to better understand the interaction between a wider range of linguistic features and their relationship to message content, in order to advance our understanding of the relationship between salience, stereotypes, and cooccurring language variables, and (ii) to expand and test advances in perception-testing methodologies, using a variety of British English as a test case. To this end, in the next section, we begin by presenting our methodology. This includes a discussion of a calibration test which provides evidence that listeners are able to respond to linguistic data in real time. The calibration test also prepared listeners for the test proper, and we discuss their reactions to Scillonian English in our results section. We conclude by reflecting upon the social meaning of linguistic features and considering what real-time perception data contribute to this knowledge.

2. METHODOLOGY. We created a survey instrument that allowed respondents to listen to voice samples remotely via an internet connection and react to them in real time. Un-

---

[4] For a more detailed discussion of the history of the Isles of Scilly, please see Moore & Carter 2015 and Moore & Montgomery 2018.

like other similar online survey tools (Watson & Clark 2013, 2015), which only gather the points in time at which a respondent reacts to something, we required that our respondents be able to both react in real time and go back and review their reactions. This allowed us to be more confident about the features listeners were responding to. Below, we describe the survey instrument, the guises respondents were asked to listen to, and the outcome of our calibration test, before turning to our results in §3.

2.1. THE SURVEY INSTRUMENT. In addition to capturing real-time reactions to multiple linguistic features, the survey instrument also collected attitude data for each voice sample, along with an array of biographical and location data from respondents. After an introductory screen containing information about the project and obtaining ethics consent from participants, respondents were invited to provide information about their biographies (age, gender), their location (based on postcode), their travel experience, and the places where they had lived (see Clopper & Pisoni 2006 for some of the effects of residential history on perception).[5] They then navigated to the next screen, where they were asked to complete a calibration test. The calibration test sample (see Figure 1) was constructed from the speech of a sixty-five-year-old male from the East End of London[6] and contained twelve instances of (th)-fronting in various linguistic environments (three in final position, two in medial position, and the remaining seven in word-initial position). We selected this feature because it has already been shown by Levon and Fox (2014) to be salient to British speakers, as discussed in §1. As with the samples used in the test proper, all tokens were natural and unmanipulated. Following Watson and Clark's approach (2013:306), a series of 'pips' were added to the start of every sample used in the survey in order to allow respondents a 'foreperiod' (Niemi & Näätänen 1981) in which to focus on the task. As shown in Figure 2, listeners were presented with instructions above the waveform, with a large green 'Click' button below.[7] They were instructed: 'When you hear him [the speaker] use an "f" sound in place of a "th" sound, please click the green button below the sound wave straightaway'. This part of the survey permitted us to test listeners' abilities to respond to a linguistic feature in real time and to examine the speed of reactions to a known feature. The results of this part of the experiment are reported in §2.3.

Test sample

[six seconds of 'pips'…] I was out of work for <u>three</u> <u>months</u>, I <u>think</u> and they said, 'Oh, we can't do <u>nothing</u> for you for three <u>months</u>, you still come and sign on.' I went, 'Oh no,' I said. 'I ain't gonna come and sign on if I ain't gonna get <u>nothing</u>, if I've got <u>three</u> <u>months</u> lieu of notice, you don't want to see me for <u>three</u> <u>month</u>.' I said, 'Er, I got better things to do with my life than come here.' But, as luck had it, there was an advert in the local paper for a job in the library, and I- I- I applied for a job and I was lucky enough to get it I <u>think</u> the <u>three</u> of us all started at the same time.

FIGURE 1. Calibration test sample. Instances of (th)-fronting are underlined.

[5] We make only limited use of these data in this article, but there is, of course, much scope to develop the analysis further.

[6] Thanks to Sue Fox for supplying this sample.

[7] The decision to use a click button as opposed to a slider was due to the need to gather binary data relating to features of regional speech, and not gradient evaluation data as in Watson & Clark 2013, 2015. Although sliders provide the facility to look at the direction of evaluation, this is not how they have tended to be used in past research. In any case, evaluation in a slider task is predominantly in one direction (Watson & Clark 2015:47), meaning that analysts can only look at single points in time. Slider-based methods also do not permit researchers to generalize across variants (only how the trajectory moves from one variant to another variant in a particular segment of talk). This is another reason why they are not suitable for the type of research we outline here.

The Voices Project

Test Sample

In one of the tasks that follow, we will ask you to listen to a voice sample and click a button whenever you notice anything in the way the person sounds which makes you wonder where he is from (or confirms where you already think he is from). To allow us to link your click to a point in the sound file, we need to get an idea of how fast your reaction times are.

To do this, we need you to listen to the extract below. The person in this extract says lots of words where he pronounces words that begin 'th' as 'f'. For instance, he has a tendency to pronounce words like 'thanks' and 'Thursday' as "fanks" and "Fursday". When you hear him use an 'f' sound in place of a 'th' sound, please click the green button below the sound wave straightaway. You can do this by hovering your mouse over the green button and clicking. This will record the exact point in time where you reacted to the sound clip.

Press play when you are ready. You will only hear the voice sample once, but don't worry if you don't catch all of the 'f' pronunciations.

Click

Next

FIGURE 2. Click interface and instructions for the test sample.

After completing the calibration test, respondents then moved onto the survey proper. Here, they listened to four male voice samples, the second and fourth of which were samples from the Scilly Voices corpus that varied by content but not by speaker, in order to test the effects of topic and discourse on perception (see §2.2 for a detailed account of this speaker). The first and third samples were taken from corpora of recordings from two other British varieties of English, Stoke-on-Trent (a city in Staffordshire in the North-West Midlands region) and Barnsley (a town in Yorkshire, in the north of England), respectively. These served as distractor samples.[8]

The listeners completed the same series of tasks for each test-proper sample in turn (completing all tasks before moving on to the next sample). For each guise, listeners were first asked to listen to the voice sample and complete five-point semantic differential scales, designed to gain reaction data along a number of attitudinal dimensions. Principal component analysis (PCA) identified three main factors from the ten dimensions (cf. Kristiansen, Garrett, & Coupland 2005:16), and maximum-likelihood factor analysis identified which dimensions should be grouped within factors. We restrict our discussion to two of these factors: Status ('educated ~ uneducated', 'ambitious ~ unambitious', 'articulate ~ unarticulate', 'confident ~ shy') and Solidarity ('friendly ~ unfriendly', 'reliable ~ unreliable', 'talking to best friend ~ talking to stranger', 'laid back ~ uptight').[9] After supplying the regard data, listeners were asked to provide additional information about the region they believed the speaker to be from, and in what kind of

---

[8] Thanks to Hannah Leach and Kate Burland for supplying these samples.

[9] The third factor contained only one dimension: 'fast talker ~ slow talker'. While the effect of perceived speech rate is, in itself, an interesting question, in line with previous research, the following discussion focuses on the perceived personality traits of our speaker.

locale the speaker might live (for instance, 'on the coast' or 'in a city'), following, for example, Campbell-Kibler (2006:95–96). Respondents then heard the sample again and were asked to 'listen out for anything in the way this person sounds which makes you wonder where he is from (or confirms where you already think he is from) … When you hear something that sounds distinctive, please click the button below the sound wave straightaway', as shown in Figure 3. This part of the test was intended to monitor the salience of specific language features to enable us to observe the effects of individual features and groups of features.

The Voices Project

Voice sample 1

For the next task, listen to the voice sample again and listen out for anything in the way this person sounds which makes you wonder where he is from (or confirms where you already think he is from). This might include the way he pronounces certain words or phrases, the quality of his voice (its speed, loudness or pitch) or the words he uses. When you hear something that sounds distinctive, please click the button below the sound wave straightaway. You can do this by hovering your mouse over the green button and left clicking. This will record the exact point in time where you reacted to the sound clip. You may hear the same distinctive feature of language multiple times and you can click as many times as you like.

You will only hear the voice sample once, so don't think too much about what it is that made you feel this way. Just click straightaway whenever you hear something which triggers these feelings. Don't worry if you miss something or click accidentally. You can correct any mistakes in the next section.
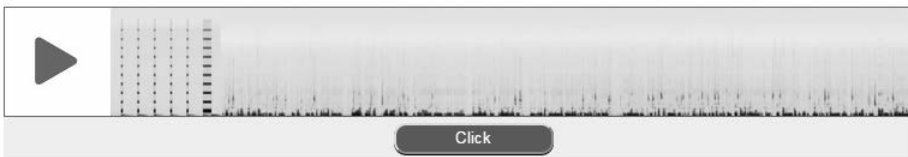
Click

Next

FIGURE 3. Click interface and instructions for the survey instrument proper.

After the completion of the click test, respondents were presented with a screen that showed where they had clicked along with the transcript of the voice sample ±3 seconds from the point of their click, as shown in Figure 4. On this screen, they were invited to state why they had clicked where they had, to indicate that they were unsure about the reason for their click, or to remove the click from the database by selecting 'I made a mistake and didn't mean to click here'. Consequently, the review function of the survey instrument minimized the possibility of not being able to discern why respondents had clicked where they had when listening to the samples—providing us with the data to measure the effects of variants and groups of variants in real time.[10]

Note that respondents were asked to listen to each sample twice, and that the ordering of the tasks ensured that they reflected on their evaluation of the speakers before

[10] Note that this was the only point at which it was possible for listeners to indicate that they had made a mistake, as they were not able to do this during the click task itself. Thus it was not possible to determine precisely when listeners changed their minds (although presumably most speakers only reflected on this when we asked them to in the review task). Nonetheless, it would be possible to trace which features speakers tended to change their minds about. This would be an interesting avenue for further research, although beyond the scope of the present study.

The Voices Project

Voice sample 2

Below is the sound wave you just heard, showing where you clicked in the last task. There is also a transcript of what you heard. You can hear the portion of speech you selected by clicking on the marker in the sound wave. Each click also has a box next to it. If you know what it was that made you click, please write what this was in the relevant box. If you don't know, or you made a mistake, please tick the appropriate box. You can listen to the sample as many times as you like for this portion of the task. If you notice anything else in the course of completing this part of the task, please write about this in the box at the end.

> ▶   out he   joined the Anzacs   and   he   got   wounded
>   ^ Click

What made you click here?

☐ I don't know.   ☐ I made a mistake and didn't mean to click here.

> ▶   got   wounded   at   Gallipoli
>   ^ Click

What made you click here?

☐ I don't know.   ☐ I made a mistake and didn't mean to click here.

> ▶   came   back   to   the
>   ^ Click

What made you click here?

☐ I don't know.   ☐ I made a mistake and didn't mean to click here.

FIGURE 4. The click review screen.

identifying salient linguistic features. This allowed us to be more confident that the listeners' evaluations of the speakers and the location of the samples primed the identification of the specific features. In effect, the ordering of tasks sought to draw attention to the relationship between evaluative and geographical stereotypes associated with the vocal stimuli and the linguistic components of the speech signal, given that 'stereotypes influence which social information gets activated (and therefore which exemplars receive activation as a result)' (Drager & Kirtley 2016:17, reflecting on Hay et al. 2006). In the next section, we explain how the guises enabled us to cue differing stereotypes.

2.2. THE GUISES. Previous findings from the Scilly Voices project have suggested that topic plays a role in the ways variants of the TRAP and BATH lexical sets are socially distributed (Moore & Carter 2015).[11] Additionally, it has been demonstrated elsewhere that listeners' expectations about speaker location can play a large role in perception

---

[11] Moore and Carter (2015) show that intraspeaker variation correlates with topics linked to social activities, and they hypothesize that this is because features have social meanings that reflect the stances a speaker has toward particular topics. However, they do not provide the perception research to substantiate this hypothesis.

(Niedzielski 1999, Hay & Drager 2010, D'Onofrio 2015). In order to prime how our samples were perceived, we created two topically distinct Scillonian guises from the same speaker recorded for the Isles of Scilly Museum's Oral History Archive. The speaker was selected because of the topic diversity in his interview. He was male, born in 1947, and schooled exclusively on the islands (some children continue to be sent away from the islands between the ages of eleven and eighteen). He was interviewed with his island-born wife by a female fellow islander in 2012. His interview covers his life and long family history on the islands, including his family's experience of island farming.

Two guises were edited using Praat (Boersma & Weenick 2015) to produce samples of broadly equal length (forty-eight and forty-nine seconds). The two samples were assembled in order to provide different topics and location cues and to ensure that a similar number of traditional Scillonian features were present in each sample. In the first sample, our speaker discusses farming practices, with no information provided betraying the location of these activities. We refer to this as the 'Farmer' guise. In the second sample, the speaker discusses Scillonian traditions and summer events. We refer to this as the 'Islander' guise. In this guise, the speaker mentions locations based within the Isles of Scilly archipelago, as well as local spatial terminology.[12] Figure 5 shows the full text of each of the guises.

Scilly farmer guise

[Six seconds of 'pips']...So he started to make a life out there. World War One broke out. He joined the Anzacs and he got wounded at Gallipoli. Came back to the UK to recuperate. Father went into pigs and all sorts of green crops and that, you know. The farmhouse up there had carbide gas. There was a little carbide plant in where the boiler house used to be. I also remember, must have been the last one they did, but they used to have an agricultural show as well. Father used to take his bull down and, cos he used to keep a bull here then - registered bull. He usually won first prize with his. Must have been eight or nine at the time.

Scilly islander guise

[Six seconds of 'pips']...I mean the only time we met up with the off-islands was one day a year. Occasionally they came to Samson picnic with us. Samson picnic was funded by May Day. The top class of the boys would go around with collecting tins. And we quite often had a sports day with them as well. Er the last one was down one of the long fields down there. When we were kids, we could go to one of the off-islands and be the only one there, or one of the Eastern Isles and be the only one there. I mean, you can't even do that in the middle of the week now. Everybody's got a boat. And the other thing - kids - we used to lie in bed and listen to, every evening, a weather plane going out.

FIGURE 5. Full text of each of the Scilly guises.[13]

Our Scillonian guises include many nonstandard accent features, as indicated by the summary in Table 1. This list includes some features noted in a historical account of the

---

[12] The cues in the Islander guise were chosen to activate local knowledge among respondents who were already familiar with the islands, but the existence of 'island' place names (e.g. 'the Eastern Isles') also clearly activated spatial associations for those unfamiliar with the Isles of Scilly.

[13] The following Scillonian locations are mentioned in the Islander guise: *off-islands* 'the term used to refer to the inhabited islands other than St. Mary's, which is the largest of the inhabited islands'; *Samson* 'one of the largest uninhabited islands, a popular place for day trips'; *the Eastern Isles* 'a collection of uninhabited islands at the easternmost point of the archipelago, and a popular place for wildlife spotting'.

variety (Thomas 1979). However, Moore and Carter (2015) suggest that Thomas's account is not entirely accurate. Therefore, we identify features in our sample speaker's talk that are present across speakers throughout the archive materials, as reported in Moore & Montgomery 2018. As such, we consider them to be traditional Scillonian features.

FEATURE	TRADITIONAL SCILLONIAN PRONUNCIATION	FARMER GUISE	ISLANDER GUISE
BATH	[aː]	plant, last	class, last
CHOICE	[ɔɪ]	joined, boiler	boys
GOAT	[oʊ] **[oʊ]** *[ɛʊ]*	broke, **show**	only, **go**, **go**, only, *boat*, going
MOUTH	[ɛʉ] **[əʉ]**	**out**, out, **house**, down	**around**, down, **down**, now, out
PALM	[a(ː)]	father, father	can't
PRICE	[ɑɪ] **[oɪ]**	life, carbide, carbide, prize, nine, time	**time**, **off-islands**, by, quite, **off-islands**, Isles, lie
STRUT	[ʌ]	bull, bull	up
TRAP	[a(ː)]	A<u>nz</u>acs, A<u>nz</u>acs, **back**, **that**	Samson, Samson
rhoticity	[ɹ]	started, there, World, War, father, sorts, farmhouse, there, carbide, carbide, where, remember, first, or	year, sports, there, were, there, there, weather

TABLE 1. Accent features in each guise.

As one of the main aims of our project was to examine what listeners react to in natural speech, and when, we did not consider it appropriate to select, manipulate, and splice individual tokens. Consequently, it was not possible to include variants occurring in the exact same linguistic contexts for each lexical set for each sample. Furthermore, Table 1 indicates that there is phonetic variation across the samples for four of the features, in particular: the vowels in the GOAT, MOUTH, PRICE, and TRAP lexical sets. The MOUTH vowels exhibit a raising pattern similar to that found in Canadian varieties of English (see Moore & Carter 2018 for a discussion of this pattern in Scillonian English), such that tokens with following voiceless consonants (e.g. *out*) tend to be raised more than those with following voiced consonants (e.g. most instances of *down*). However, it is important to note that this speaker does not always differentiate between these environments, and even the vowels followed by voiced consonants have audibly raised onsets. The variability in the MOUTH lexical set is present in both samples. This is also true for the variability in the degree to which the off-glide of GOAT is centralized. However, there is an extremely fronted onset in the word *boat*, a variant that is found in the Islander guise only. TRAP and PRICE also have variants that occur in one guise only. The PRICE vowels do not show the same raising pattern observed for MOUTH, with all tokens audibly raised, irrespective of following context—indeed, as Table 1 shows, some tokens followed by voiced consonants (e.g. *time* in the Islander guise) are more raised than tokens followed by voiceless consonants (e.g. *quite* in the same guise). The closer variants are found only in the Islander guise. For the TRAP vowels, although the quality is similar across guises, two of the TRAP vowels in the Farmer guise (*back* and *that*) are especially long.

Of course, given that the speaker is discussing different topics (and it is well established that topic can affect vowel quality; see for instance Rickford & McNair-Knox 1994, Love & Walker 2013), some variation across the guises would be expected, and its presence reflects the sociolinguistic reality of topic and token interaction. However, it makes our task of interpreting the individual effects on our listeners more complicated. Given that a key argument in this article is that effects very rarely operate in isolation, we return to this issue when we discuss the results of the test proper. The analysis

of these results was reliant on data obtained during the calibration test. Consequently, we report on the outcomes of this component of our survey in the next section.

2.3. THE CALIBRATION TEST. Figure 6 shows the instances of (th)-fronting at the point at which they occurred in the calibration test sample, along with respondent click data. In this figure (and all similar following figures), we have 'binned' click data into 0.5 second silos (so, for example, clicks at 10.201, 10.305, 10.335, and 10.499 seconds would be included at the 10.0 second point on the x-axis).

The calibration task could be conceived of as a phoneme-monitoring task (Frauen-felder & Segui 1989, Bigand et al. 2001). However, as we were asking participants to respond to instances of (th)-fronting (i.e. [f] where Standard Southern British English would typically use [θ], not simply any instance of [f]), we were actually asking them to engage in 'variant monitoring' (cf. Levon 2016).

Figure 6 shows that respondents did appear to be able to respond to this task with relative success. The black line spikes after most instances of (th)-fronting, demonstrating that respondents were able to react to linguistic features in real time using the survey interface.
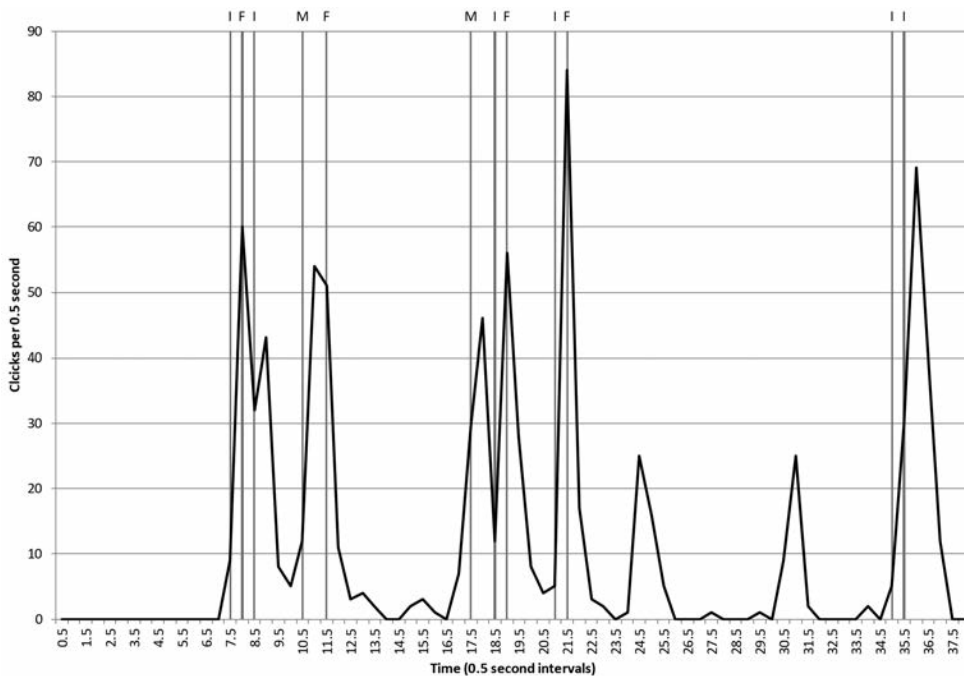

FIGURE 6. Reactions to calibration test sample. The black line indicates total respondent clicks by 0.5 seconds, and gray columns instances of (th)-fronting (labels indicate linguistic environment of each instance of (th)-fronting: I: word-initial, M: word-medial, F: word-final).

There are two further points of note relating to the calibration task: priming effects, and the impact of word position on the perception of (th)-fronting. Possible priming effects can be seen in the spikes in the chart in Fig. 6 at 24.5 and 31 seconds. These clicks appear to be in reaction to the words *with* (at 24.698 seconds into the sample) and *for* (at 30.536 into the sample). Counting all of the clicks within 677 milliseconds of the end of each word, we can observe forty-one respondent clicks for *with* and thirty-four respondent clicks for *for*. The first word did not carry a token of fronted (th) and, although the second word had [f] in word-initial position, this was not an example of fronted (th). It

seems that in both of these instances respondents reacted unreliably. In the first instance, *with* was pronounced as [wɪ]. Thus respondents reacted to a nonstandard variant of (th), just not the one they had been instructed to listen out for. In the second instance, respondents heard [f] in the wrong context and clicked anyway. Both of these instances suggest that a number of respondents had been 'primed' (cf. Niedzielski 1999, Hay & Drager 2010, D'Onofrio 2015) to react in a particular fashion by the task instructions and their previous behavior. In the first instance, they were primed to respond to any nonstandard instance of a specific variable. In the second instance, they seem to have moved from 'variant monitoring' (Levon 2016) to phoneme monitoring over the course of the task. We mention these priming effects because they demonstrate the problems of asking respondents to listen for particular variants. Our test proper did not require respondents to do this, so it provides a way to overcome the limitations of this task.

In addition to these priming effects, there appears to be an influence of word position on the perception of (th)-fronting in the calibration test, with word-final instances receiving fewer clicks than either word-initial or word-medial examples. Stuart-Smith and colleagues (2013:512) show in their data from Glasgow that (th)-fronting is most likely in word-final position, a linguistic constraint that we could reasonably expect to hold across varieties in which (th)-fronting occurs. It is therefore in the positions that are less likely to exhibit (th)-fronting that more respondents noticed the feature. We mention this effect because it appears to support Rácz's (2013:51) theory of 'surprisal' in relation to salience, and we return to this kind of effect in our discussion of the results for our Scilly guises below.

The limitations of the calibration test did not negatively impact its effectiveness in the context of our survey. We used the calibration task not only to train respondents in the survey interface and to verify that they could use it to respond to speech in real time, but also to calculate mean reaction times so that we knew how quickly listeners were likely to react to known stimuli. To do this we selected the most isolated examples of (th)-fronting (10.506: *nothing*; 11.652: *months*; 17.601: *nothing*; 35.44: *think*) and the click times closest to the end of the word carrying each instance.[14] This calculation revealed a mean reaction time to these examples of (th)-fronting of 697 milliseconds, with a standard deviation of 464 milliseconds. This is slower than the c. 350 millisecond reaction time discussed by Pachella (1973:43), and the 215 millisecond reaction time to visual stimuli recorded by http://www.humanbenchmark.com/ (as of 19 January 2016). It is, however, within the two-second window with which Watson and Clark (2013:314) worked in their change-point analyses, and the 1.3-second decision process between stimuli and mouse click observed by D'Onofrio (2015:249).

Of course, as the calibration test and the rest of the experiment were undertaken via the internet, we had little control over the experimental environment. This means that respondents could have been differentially engaged with the task, or as Watson and Clark (2013:314) note, they might have been fatigued prior to completing it, or been under the influence of alcohol or drugs, among a number of other factors that might have produced less robust data. However, it is assumed that any individual's circumstances remained similar over the course of the experiment, and by establishing a benchmark for reaction times using the calibration test we were able to account for unusual reaction times among respondents when we evaluated their responses in the experiment proper.

---

[14] Of course, we are assuming here that the proximity of the click to the instance of (th)-fronting indicates that individuals clicked in response to the (th)-fronting. The regularity of the clicking across all sample participants suggests we can have confidence in this assumption.

In sum, the results from the calibration test provide good evidence that listeners were able to respond to a known type of feature, notwithstanding the priming effects observed above. Note, however, that the conclusions that can be drawn from this task are speculative; for example, although it appears that the listeners are doing variant monitoring (with the exception of the instance discussed above), we cannot be sure that this is the case as we have no way of probing the clicks further. As noted above, this highlights the need for the review element of the survey instrument, especially as we seek to understand the facility of listeners to respond to unspecified features. We consider this below, as we turn to the results from the test proper.

3. RESULTS. After piloting the materials and test interface with our own undergraduate students, the online test was released on 1 May 2014 and gathered data for six weeks. Respondents were contacted via social media (Facebook and Twitter) and asked to complete and share the survey. The final number of participants was 112, of whom eighty were female and twenty-eight were male (four refused to say, or did not define their sex in a binary fashion); of these, six respondents provided no reaction data and their data have not been included in the analysis below. The mean age of respondents was thirty-two, with a highest age of seventy-two, and lowest age of sixteen (the standard deviation for age was 14.2). In order to assess their geographical spread and wider geographical experience, respondents were asked for the postcode of the place where they currently lived and the number of towns/cities they had lived in, and were also asked about travel experience via a question regarding which of ten regions they had visited (based on the Regions of the England (ONS Geography 2010), plus the Isle of Wight and the Isles of Scilly). Respondents were drawn from forty-four of the 124 postcode areas in the UK, had lived in an average of 3.2 places (standard deviation 1.9), and generally had a good amount of travel experience, having visited seven of the regions on average. Figure 7 presents data relating to respondents' postcode distribution on the left, and shows the numbers of those who had visited various regions of England on the right. One of the wider aims of the test was to gather data on Isles of Scilly residents' perceptions of their own variety, and targeted circulation among this population yielded eight respondents who lived in the Isles of Scilly, and a further six who had visited the islands. We discuss where this might have affected responses below.

3.1. GUISES AND LISTENER REGARD. We begin our discussion of the listeners' reactions by considering how the two Scillonian guises were rated relative to each other, using the ratings data gathered prior to the real-time reaction task and the qualitative comments obtained during the real-time task. This provides information that allows us to evaluate the stereotypes activated by the two samples.

The ratings data show statistically significant differences between the ratings assigned to the two guises. Table 2 shows the mean ratings data by evaluative dimension. Paired t-tests run on these data revealed significant differences between the two guises for the ratings factors, with highly significant differences between ratings for Status and Solidarity.

DIMENSION	MEAN: FARMER GUISE	MEAN: ISLANDER GUISE	PAIRED t-TEST
Status	2.39	2.56	***
Solidarity	2.50	2.28	***

TABLE 2. Mean ratings for each guise (* $p < 0.05$, ** $p = 0.01$, *** $p < 0.001$, N/S: not significant).[15]

---

[15] Note that in Table 2, and at future points in the article, ratings data have been transformed in order that higher scores equal 'better' ratings. In the data-gathering task, scales of 1–5 were constructed with values on the left-hand side of the screen (i.e. closer to 1) representing the most positive score.
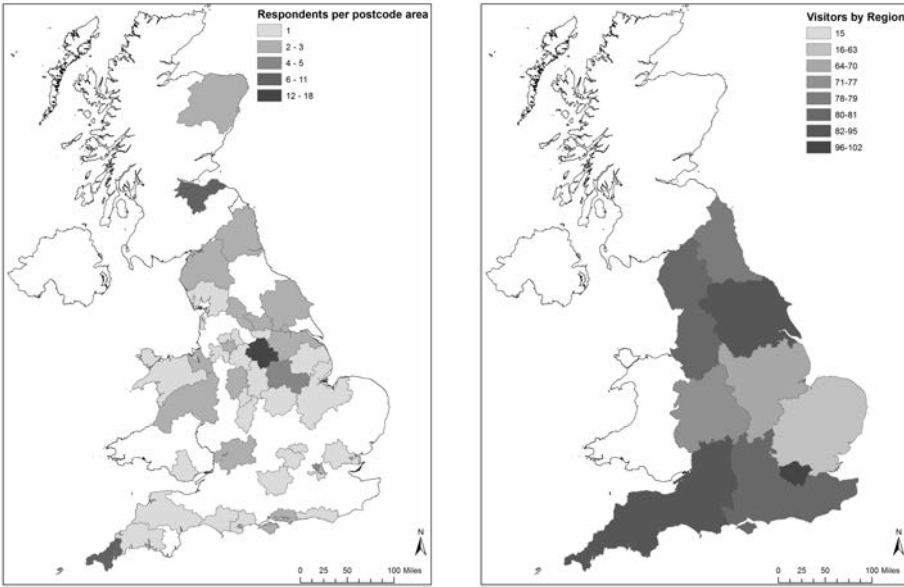
FIGURE 7. Total number of respondents by postcode areas (left) and total number of respondents who have
visited particular regions of England (right). Contains National Statistics data ©Crown copyright and
database right (2012, 2016), OS data ©Crown copyright [and database right] (2012, 2016), and
Royal Mail data ©Royal Mail copyright and database right (2012).
Data from an Ordnance Survey/EDINA supplied service.

The data in Table 2 show that, for the Islander guise, ratings were higher for Status
than Solidarity, and for the Farmer guise, that the ratings were lower for Status than Sol-
idarity. A Tukey post-hoc HSD test[16] demonstrated significant differences between di-
mensions for the Islander guise. The same test showed no significant difference
between Status and Solidarity for the Farmer guise. The differences between the two
guises suggest a more nuanced response to the Islander guise.

The ratings data clearly show very different responses to the two guises, and the
qualitative data suggest possible reasons for this. Unsurprisingly, given its topic, the
Farmer guise resulted in nine free comments relating to farming. This included two ref-
erences to the speaker being from a rural community, and noting 'rural references' in
the talk. One respondent referred specifically to the 'content not voice' of the sample,
presumably differentiating the content of the guise and the accent, while two other lis-
teners did not divorce the two, stating that the speaker 'Sounded very rural' and
'Sounds like a farmer'. Interestingly, some of the comments inferred behavioral traits
from the sample, suggesting stereotypical views of rural locations; for instance, it was
claimed that he 'Sounds like a Hobbit', that he ' … might like a cider', and that he was
a 'Nice country chap'. Furthermore, questions asking more specifically about the type
of settlement the speaker might live in resulted in 84.7% of respondents stating 'Coun-
tryside' for the Farmer guise (versus 21.5% for the Islander guise). In response to the

---

[16] Of course, we cannot be sure that, for example, one unit of 'Status' is equivalent to one unit of 'Solidar-
ity'. However, participants completed the ratings tasks in the same way and used the same five-point scales.
As noted above, the PCA indicated that three factors were the best solution for the data, and we used factor
analysis to group the scales into dimensions. All of this means that we are satisfied that the scales are similar
enough for this test to be run.

question asking respondents to locate the region that the speaker came from, the Farmer guise was correctly allocated to the South West region by 83.7% of listeners (compared to 48.6% for the Islander guise). However, overall, only 8% of respondents correctly identified the Farmer guise as being more specifically from Scilly. As might be expected, the Scilly respondents were slightly better at identifying him more specifically, with one third correctly identifying the speaker as Scillonian.

The Islander guise was clearly viewed quite differently from the Farmer guise, and especially by those listeners from Scilly. Overall, 29% of respondents identified the Islander guise as being from Scilly, but the content of the sample ensured that not only were 88.9% of Scillonian respondents able to identify that the speaker was from Scilly (e.g. 'Lived in Scilly all his life', 'Born and bred on Scilly'), but also, in two cases, listeners claimed to be able to identify the speaker himself. Note that none of the listeners, whether from Scilly or not, mentioned Scilly in the free comments for the Farmer guise. For those listeners not from Scilly, comments relating to the Islander guise are more difficult to classify than those for the Farmer guise. As noted above, 48.6% of respondents identified the speaker as from the South West region. In addition to this being fewer respondents than for the Farmer guise, there was also more ambivalence noted about this designation. For instance, some listeners claimed the speaker had a 'mixed' accent ('He sounds like he has a very mixed dialect/accent'). One listener stated that they thought the speaker had an RP accent, but for the presence of rhoticity, with others also noting this feature (e.g. 'Nice rounded burr to his voice', 'Emphasised "R" … '). Notably, it was quite clear that listeners did not realize that the Islander guise speaker was the same as the Farmer guise speaker, demonstrated by the following comment: 'definitely sounds very different to the previous example from the south west but not sure where from'. In contrast to the Farmer sample, there were no references to rurality or farming in the comments.

The sample of listener comments that we have included above suggests that particular linguistic features were important for the way that listeners regarded the two guises (see, in particular, the mention of rhoticity). In the following section, we detail which features were reacted to in each guise before discussing how such information helps us to further understand listeners' evaluation of the guises.

3.2. REAL-TIME REACTIONS AND PROMINENT FEATURES. Table 3 and Figures 8 and 9 summarize the click data obtained in response to the two Scilly guises, time-aligned with the features listed in Table 1. The data shown in these figures are the raw data, including instances when the listeners reported that they did not know why they had clicked in a particular location, as well as instances in which listeners reported features that were not present in the recordings (for example, when speakers alluded to rhoticity when it was not actually present in the word they indicated in their comments). The data are normalized for each listener based on their mean reaction time for the calibration test sample and displayed in 0.5 second 'bins' as in Fig. 6.[17]

SAMPLE	TOTAL CLICKS	MEAN CLICKS PER 0.5 S.	MAX CLICKS PER 0.5 S.
Scilly Farmer	814	8.39	33
Scilly Islander	810	8.35	33

TABLE 3. Summary of click data.

[17] Of course, it is possible (indeed, likely) that respondents react at different speeds to different kinds of features, and it is difficult to know precisely how the speed of reactions may be affected. We do know, however, that reactions are not instantaneous, and this method of normalization goes some way toward accounting for this.

Figures 8 and 9 show remarkably ordered reactions among listeners. Although there is low-level click activity for both samples from the moment the introductory 'pips' finish (at six seconds), there are numerous spikes in the charts that demonstrate a number of listeners clicking at the same time. This provides evidence that supports the data from the calibration test discussed in §2.3. It indicates that respondents are able to react in real time to voice samples, even when not directed to respond to a specific feature (cf. Watson & Clark 2015). That listeners are responding to the features indicated in Figs. 8 and 9 is supported by the listener commentary on the clicks, which specifically identifies certain features, as we discuss below.
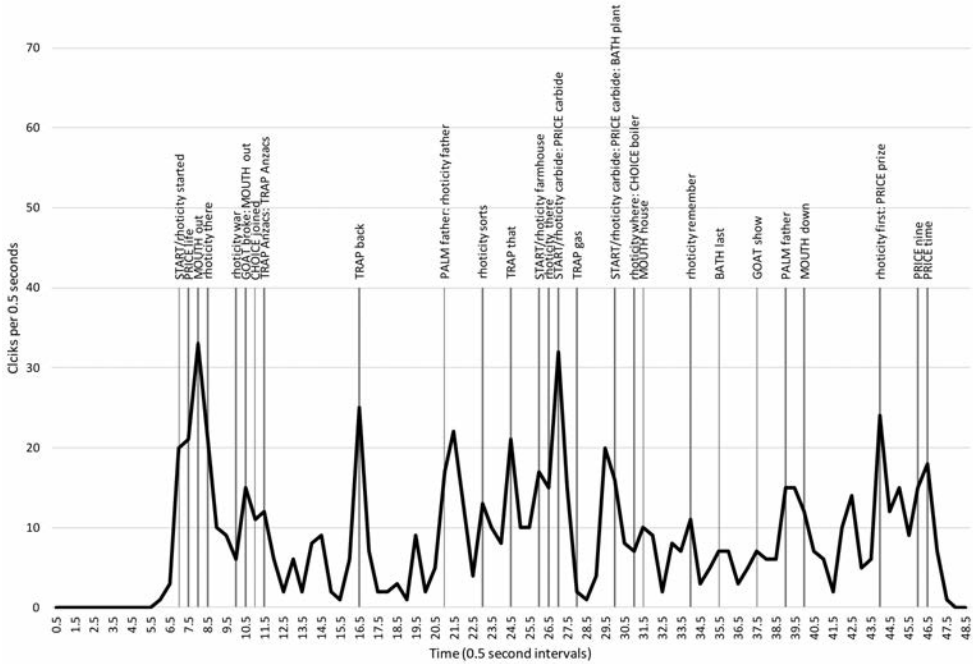

FIGURE 8. Click reactions to Scilly Farmer guise. Black line indicates total respondents by 0.5 second intervals, gray columns instances of features noted in Table 1.

Considering the reactions to the Farmer guise shown in Fig. 8, it appears that some spikes clearly coincide with features previously identified as typically Scillonian in Table 1. The instances of TRAP at 16.5 and 24.5 seconds, the instance of PALM at 39 seconds, the two instances of PRICE at 46 and 46.5 seconds, and rhoticity at 23 and 34 seconds all appear to be examples of collective reactions to specific features. A similar inspection of the Islander guise in Fig. 9 reveals spikes coinciding with instances of GOAT at 7.5 and 39 seconds, BATH at 17 seconds, MOUTH at 27 and 37.5 seconds, and rhoticity at 10, 24, and 33.5 seconds. Note that the token of GOAT with the extremely fronted onset (*boat* at 39 seconds) produces a large spike, but the tokens of PRICE with extremely raised onsets (*time* at 8 seconds, and *off-islands*, occurring twice at 9 and 32 seconds) do not.

In addition to these obvious reaction points, there are also other types of 'spike' in the two figures. Some of these reflect the cooccurrence of features—either where two potentially salient features occur in the same word (for instance, PALM and rhoticity in *father*), or where many features occur in separate words in a short space of time (one such
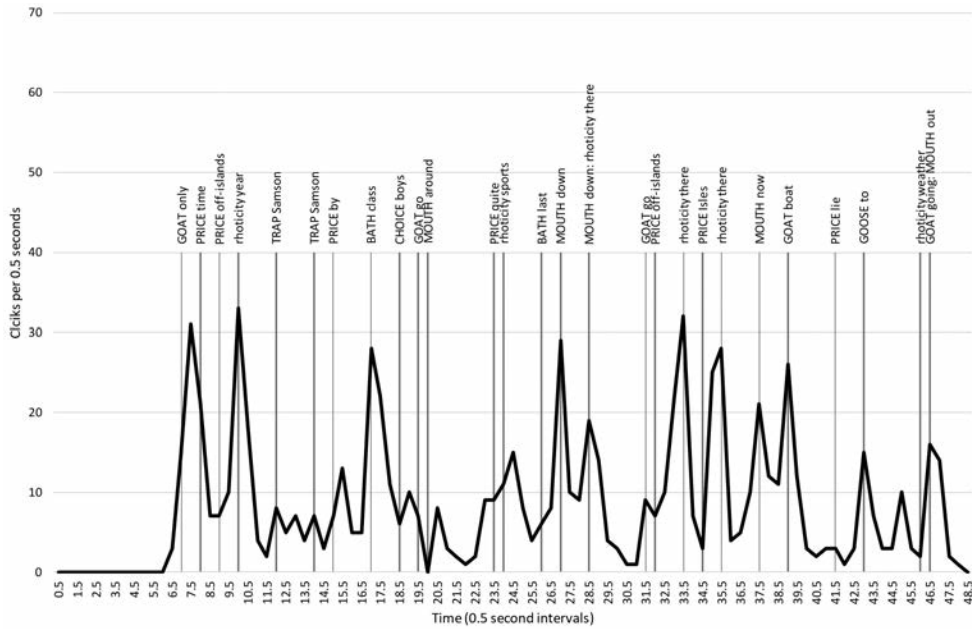
FIGURE 9. Click reactions to Scilly Islander guise. Black line indicates total respondents by 0.5 second intervals, gray columns instances of features noted in Table 1.

example begins at 6.5 seconds in the Farmer guise). In these situations, we could have been faced with the same problem as Watson and Clark (2015:54), who found that simply using reaction data alone meant that it was 'impossible to tease apart exactly which linguistic feature (or a combination of features) listeners are reacting to'. Our survey instrument came into its own in overcoming this issue, as the review function allowed us to gain at least a partial window on the process entered into by listeners when making their reactions.

In order to analyze the review data, each click made by each respondent was coded for one or more of the features listed in Table 1.[18] This was possible because listeners would either report the word itself (which we could then easily classify) and/or mention a particular vowel or consonant sound. For example, comments including 'long a in class', 'down sounds like dowyn', and 'boat is bowt' were made about clicks for the Scilly Islander guise. While coding, clicks that had ambiguous (e.g. 'Intonation, vowels, pronunciation', 'accent') or irrelevant reasons (e.g. 'reference to weather plane') were excluded, as were clicks when listeners stated that they did not know why they clicked when they did, or when there was no reason given. Coding was performed individually by both authors and disagreements flagged. When disagreements remained after further consideration, these clicks were also excluded. After exclusions based on these criteria, 555 clicks remained for the Farmer guise (68.18% of the total clicks) and 504 for the Islander guise (62.22%). Table 4 shows the click-data results after the coding process.

Table 4 demonstrates the variable perception of different features according to guise. This table suggests that one feature, rhoticity, is very frequently recognized in both guises (although less so in the Farmer guise, as we return to below). This frequent recognition

---

[18] Additional features were noted, but inconsistently, which made them unsuitable for quantitative analysis.

FEATURE	INSTANCES	FARMER GUISE POSSIBLE CLICKS	TOTAL CLICKS	% CLICKS	INSTANCES	ISLANDER GUISE POSSIBLE CLICKS	TOTAL CLICKS	% CLICKS
PALM	2	214	65	30.37	1	107	8	7.48
PRICE	6	642	115	17.91	7	749	63	8.41
TRAP	4	428	59	13.79	2	214	3	1.40
BATH	2	214	18	8.41	2	214	65	30.37
MOUTH	4	428	27	6.31	5	535	95	17.76
STRUT	2	214	13	6.07	1	107	2	1.87
GOAT	2	214	12	5.61	6	642	63	9.81
CHOICE	2	214	7	3.27	1	107	12	11.21
rhoticity	14	1,498	239	15.95	7	749	193	25.77

TABLE 4. Coded click data.[19]

of rhoticity is perhaps unsurprising, given the comment data outlined in the previous section, although it also echoes the findings of Clopper and Bradlow (2009) and Wagner, Clopper, and Pate (2014:1080) about the importance of categorical consonant features to perception. Vocalic features tended to exhibit more variable rates of recognition, such that they are much more readily recognized in one guise than the other. In order to test the significance of differences between recognition rates, we conducted a repeated-measures logistic regression using the {lmer4} package in R (Bates et al. 2011) with speaker as a fixed effect and listener as a random effect. Each feature (e.g. MOUTH) was separated into its own analysis, with each of these analyses based on a data table that listed, for each respondent, every opportunity to notice the individual instances of the feature and whether or not the feature had been noted. Table 5 shows the headline results from this process, with the results of each analysis given in the appendix. Vowels from the PALM, PRICE, and TRAP lexical sets are more readily recognized in the Farmer guise, whereas vowels from the BATH, MOUTH, and CHOICE lexical sets are more frequently recognized in the Islander guise. Note also that, while rhoticity is recognized frequently in both guises, it is significantly more recognized in the Islander guise.

FEATURE	FARMER GUISE % CLICKS	ISLANDER GUISE % CLICKS	p-VALUE
PALM	**30.37**	7.48	***
PRICE	**17.91**	8.41	***
TRAP	**13.79**	1.40	***
BATH	8.41	**30.37**	***
MOUTH	6.31	**17.76**	***
STRUT	**6.07**	1.87	N/S
GOAT	5.61	**9.81**	N/S
CHOICE	3.27	**11.21**	**
rhoticity	15.95	**25.77**	***

TABLE 5. Recognition of features present in both guises, with summary of generalized linear mixed-effects regression models for each feature (* $p < 0.05$, ** $p = 0.01$, *** $p < 0.001$, N/S: not significant). Bold figures indicate the higher percentage recognition.

Most of the differences between the recognition of features in Table 5 are highly significant, suggesting robust differences between the two guises. Recall that we noted earlier that there were more extreme tokens of the TRAP vowel in the Farmer guise, and this may explain why this feature is more significantly recognized in this guise, as we

[19] The 'possible clicks' figure is derived from the number of times a feature appeared in the guise multiplied by the number of respondents included in the analysis (107). The 'percentage clicks' is simply the number of clicks made divided by the number of possible clicks that could have been made, multiplied by 100.

return to below. However, the existence of more extreme variants of GOAT and PRICE in the Islander guise does not seem to have resulted in these features being more noticed in this guise. In fact, there is no significant difference in the noticing of the GOAT vowel between the two guises, and the PRICE vowel is actually more significantly recognized in the Farmer guise (despite this guise not including the tokens with the most-raised onsets). Given these findings, and that both guises otherwise contained very similar features, this suggests that guise content and, potentially, the ways in which features cooccur might contribute to how listeners regarded the two guises and the features that they claimed to recognize. We explore this possibility further below as we reflect on our results.

4. DISCUSSION. The evaluation of the two guises, as evidenced through the ratings task and the qualitative data, clearly demonstrates that they were regarded differently. The real-time data and the comments discussed above also point toward differential perception of the same accent features in the two guises. Possible explanations for these results include: the interaction between guise content and the variants, the exact quality of the variants, and the precise placement of the variants.

We begin by considering guise content. Unsurprisingly, as already noted, the Farmer guise resulted in comments related to farming and rural life (for example: 'Sounds like a farmer'; ' … his accent is one I would associate with a farmer'; 'Works the land'). Britain (2017:174) has noted that rural areas are typically stereotyped as 'backward, conservative, boring, dangerous, threatening, "uncultured" and uneducated'. While there are many rural areas in England, the South West of the country is commonly stereotyped in these terms. Figure 10 shows three examples of perceptual dialectology 'draw-a-map' tasks (Preston 1989) from fieldwork undertaken by Montgomery, which show the link between the South West, rurality, and farming.

We have seen the strong associations between the South West, rurality, and farming, and the accurate allocation of the Farmer guise to the South West and to the Countryside by 83.7% and 84.7% of listeners, respectively (see §3.1). Therefore, it is perhaps unsurprising that the four most frequently recognized features in the Farmer guise—vowels in the PALM, PRICE, and TRAP lexical sets and rhoticity—are strongly associated with South Western varieties. Vowels in the PALM lexical set demonstrate the typical South Western pattern in the Scillonian data—that is, [a] rather than [ɑː] (Wells 1982:346); and, while the precise onset of PRICE varies across this region, it has a tendency to be centralized or retracted throughout (Wakelin 1986:27–28). Similarly, while the raised quality of the TRAP vowel is like traditional RP, Piercy (2010:35–37) notes that the lengthening of this vowel is less ambiguously South Western in character (recall that the Farmer guise included lengthened tokens of TRAP in the words *back* and *that*). Finally, rhoticity is geographically restricted in England, occurring only in the South West of the country (in the counties of Cornwall, Devon, Dorset, Somerset, Bristol, Gloucestershire, and Wiltshire) and in some areas of Lancashire and Yorkshire in the North of the country (Trudgill 1999). Despite its existence in some northern counties, Montgomery (2007:250) has shown that nonlinguists particularly associate this feature with the South West, and especially with the farming industry linked historically to these regions.

The exact quality of at least two of these features may also have contributed to their recognition in the Farmer guise. We have already mentioned that the Farmer guise contained lengthened tokens of TRAP, whereas the Islander guise did not. However, while both guises contained 'South Western' PALM vowels, the carrier word in the Farmer
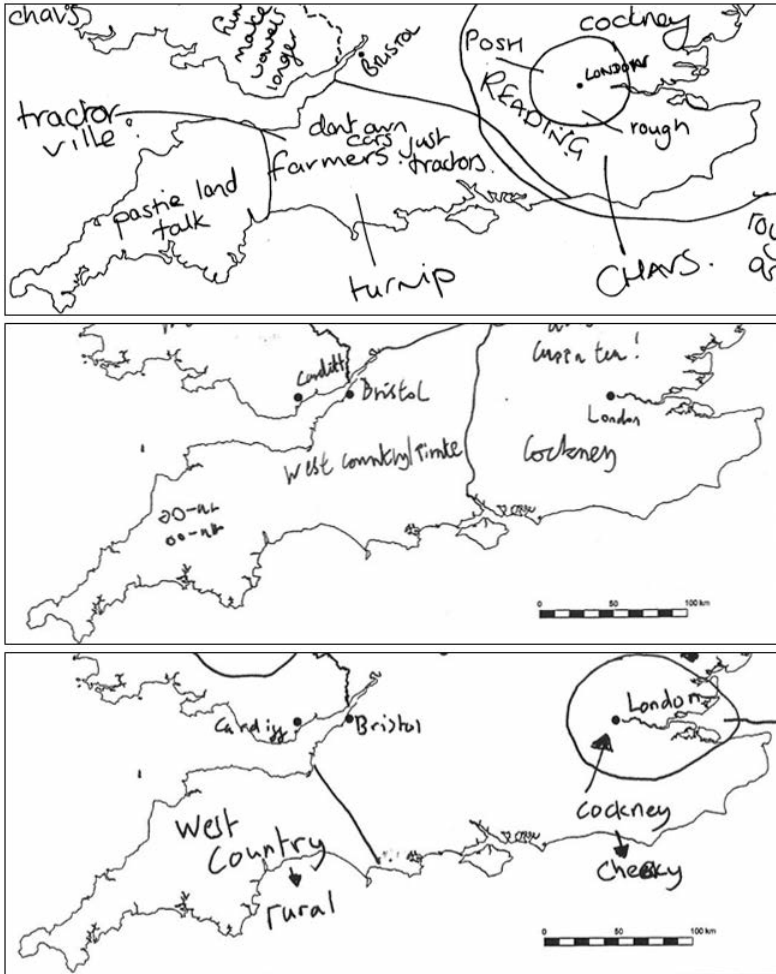
FIGURE 10. Examples of draw-a-map task data from previous fieldwork (see Montgomery 2011, 2012).[20]

guise was a rhotic content word, *father*, which occurred twice (first at 21 seconds and again at 39 seconds). This is in contrast to the Islander guise, where the PALM vowel occurred once at 36.5 seconds in the nonrhotic function word *can't*. Occurrence in a content word and alongside another strongly South Western feature is likely to have affected the way in which PALM was perceived in the Farmer guise, compared to the Islander guise.

Turning now to the Islander guise, the four most frequently recognized features in this guise were the vowels in the BATH, MOUTH, and CHOICE lexical sets and rhoticity. With the exception of rhoticity, these features are either specific to Scillonian English, or at least less markedly 'South Western'. For instance, Wakelin (1986:28) reports that, with regard to CHOICE, '[t]he traditional SW form seems to be /ʌɪ/ but /aɪ/-forms also occur here'—suggesting that a raised onset in CHOICE is more specific to Scillonian En-

---

[20] The phrase 'Oo ar, oo ar' marked on one of the maps represents an example of a stereotypical 'feature' of the variety associated with the South West of England; see numerous definitions (spelled slightly differently) in the Urban Dictionary (http://www.urbandictionary.com/define.php?term=Ooh+arr).

glish (or at least is not recorded as a generally South Western feature). Given how little known Scillonian English is (recall that it was accurately identified by only 8% of listeners in the Farmer guise and 29.4% of listeners in the Islander guise), this form might be notable because it is unusual (cf. Preston 2010, Rácz 2013), but it is unlikely that it carries much social meaning to anyone unfamiliar with this variety. With regard to MOUTH, there is variability in the pronunciation of this form in the South West, with the onset variably centralized, lowered, or fronted, and with centralized forms shared with the South East, according to Wakelin (1986:28). Again, this might make Scilly's raised and centralized onset notable (cf. Preston 2010, Rácz 2013), but less distinctively meaningful in the context of South Western English. This is supported by the following listener's comment about the Islander guise: 'I thought he sounded really west country but then there were bizarre vowels every so often that skewed my perception of him'. Finally, while a fronted BATH vowel can be socially stigmatized relative to RP, it is certainly not restricted to the South West. In fact, it is often more commonly associated with the north of England. This is seen very clearly in the following comment from a listener: ' "class" sounds like a northern "a" but he sounds like a southerner otherwise'.

Given that fewer than half of the listeners (48.6%) identified the speaker of the Islander guise as being from the South West, and even fewer identified him as being from the countryside (26.5%), it is perhaps unsurprising that listeners less readily recognized features simultaneously associated with the South West and rural life.

The potential flaw in this argument is the finding that rhoticity is more frequently noticed in the Islander guise than the Farmer guise. Given its strong association with South West English, we might expect rhoticity to be noticed more in the Farmer guise. However, it may be important to consider not just the associations that variables have but also their ideological loading, and the relationship between ideology and expectation. In contrast to US English, rhoticity is a shibboleth in English English. While the other features of South West English we have identified (those associated with the PALM, PRICE, and TRAP lexical sets) are strongly associated with the region, to our knowledge, none carry the same iconic status as rhoticity (see, for instance, Maguire et al. 2010:97, which stresses the importance of rhoticity in defining the relationships between varieties of English). Consequently, an unexpected occurrence of a fronted PALM vowel, or a raised onset in PRICE, may be reasonably easy to overlook (or even filter out), but an unexpected occurrence of rhoticity is striking. Some support for this hypothesis is found in the comments made by listeners who encountered rhoticity when listening to the Islander guise (for example: 'Most "r's pronounced strongly'; 'Rolling the r's'; 'At first thought he had an RP accent and then picked up on his pronunciation of his r's').

Table 6 shows where the rhotic forms occur in each guise. Note that rhoticity was a highly frequent feature in both guises, occurring fourteen times in the Farmer guise and seven times in the Islander guise. As with all of our features, rhoticity remained salient throughout both of the guises, with listeners clicking in recognition of it until the end of both of the samples. That is to say, listeners did not simply 'turn off' after an early instance of a prominent feature, but continued to attend to it as the sample progressed.[21]

---

[21] This would seem to corroborate Labov and colleagues' (2011) observation on the sociolinguistic monitor, where listeners continued to make judgments about a speaker even after the first instance of the apical (-ing). But they also found that the feature becomes less noticed the further into the sample listeners got. It is important to note, however, that our data set and that used in Labov et al. 2011 are not directly comparable, given that we measure only reaction and not strength of reaction. In addition, our samples were shorter than those used in Labov et al. 2011.

FARMER GUISE			ISLANDER GUISE		
WORD	TIMESTAMP	CLICKS	WORD	TIMESTAMP	CLICKS
started	7.459	11	year	10.404	50
there	8.851	33	sports	24.634	13
World	9.714	12	there	28.956	33
War	10.216	14	were	29.689	4
father	21.450	24	there	33.854	47
sorts	23.506	14	there	35.881	41
farmhouse	26.368	23	weather	46.298	5
there	26.698	9			
carbide	27.626	38			
carbide	30.269	16			
where	31.024	4			
remember	34.383	18			
first	44.215	13			
or	46.031	10			

TABLE 6. Clicks for rhoticity throughout each guise.

Nonetheless, notice that the first occurrence of rhoticity in the Islander guise receives more clicks than any other instance of this form in either of the guises.[22] Compare this with the first (or even the second) instance of rhoticity in the Farmer guise, which receives far fewer clicks. Remember that our listeners completed the click task only on the second hearing of the guises. As we noted earlier, this was a deliberative process which ensured that listeners evaluated the stimuli and that this evaluation primed them for the click task. Given the strong associations between rhoticity and farming, and the strong tendency to associate the Farmer guise with the South West and the countryside, listeners were primed to hear rhoticity in the Farmer guise, but not in the Islander guise. Given the special status of rhoticity as a feature of English English, we raise the possibility that its prominence is increased in the Islander guise. Podesva and colleagues (2015) found that the social meanings that listeners assigned to released /t/ varied according to expectations of how US politicians 'usually' talked, suggesting a correlation between expectation and speaker evaluation. Our data support the role that expectation plays in perception, but suggest that expectation can interact with ideology. If this is the case, it would mean that listeners are able to filter out unexpected features (i.e. those not primed by stereotypes associated with the speaker), but that there may be a threshold for the filtering process, such that some features are simply too iconic to ignore. As we discuss in the conclusion, this would require more empirical testing to verify.

As we have noted, some of the effects in our data may have been the consequence of the exact quality of the variants and the precise placement of the variants in the stimuli. Others point to the importance of topic and discourse context and its ability to prime listeners' expectations. Consequently, it is important to reflect on the interaction between topic and discourse context and the precise task that our listeners engaged with. In particular, the instruction during the click task to 'listen out for anything in the way this

---

[22] It is important to note that one thing not accounted for in our study is the precise sequencing of multiple variables. It is not clear how our proposal about the interactions between expectations and the filtering process operates for variables that are less iconic than rhoticity. For instance, the precise ordering of a BATH vowel and a rhotic, or a PRICE vowel and a MOUTH vowel, may cause a particular effect in a particular guise. As we discuss in the conclusion, the deliberative process of the tasks (having listeners hear the sample in full once before identifying features in the second listening) may have reduced any effects caused by features occurring in different sequences in the two guises. Nonetheless, some sequencing effects may have occurred, and this remains an avenue for future research.

person sounds which makes you wonder where he is from (or confirms where you already think he is from)' likely resulted in attention to features activated by a stereotype associated with place (and its interaction with topic) in one guise. In the other, it seems to have resulted in attention to surprising features in the absence of any stereotypical associations. Therefore, features associated with the South West (the vowels in PALM, PRICE, and TRAP) were significantly more recognized by listeners in the Farmer guise, and those vowels that did not index a recognizable stereotype (the vowels in BATH, CHOICE, and MOUTH) were more recognized in the Islander guise. Rhoticity, with its iconic status, was well-noticed in both guises, but had a higher 'surprisal value' (Rácz 2013:51) to listeners who were not expecting it due to a lack of stereotype priming in the content of the Islander guise.

Our finding that listener expectations can be conditioned by the topic of talk, and that these expectations are strong enough to determine what listeners recognize in the linguistic signal, is a new empirical example of a previously attested phenomenon (e.g. Hay et al. 2006, Preston 2010, 2011, Levon 2011, Rácz 2013). However, our research demonstrates that it is possible to obtain these results without manipulating individual linguistic features or abstracting them from the wider discourse frame in which they occur. More notably, our research has shown the importance of analyzing the perception of multiple features in the speech signal in real time. It would have been possible to isolate each of the variables in Table 1 and test them by varying their occurrence in a controlled sample, but—in addition to being more laborious—this would not have exposed the ways in which variants and message content cohere in relation to a given stereotype. As we have shown, rhoticity is an iconic feature in English English which is especially ideologically loaded and linked to a stereotype of 'South West farmer'. Nonetheless, our data have shown that the occurrence of rhoticity alone does not inevitably result in the percept 'South West farmer'—as evidenced from the fact that this feature not only occurs, but is also recognized, in a guise (that of the 'Islander') where the percept 'South West Farmer' is not generally cued. This suggests that rhoticity means 'farmer' when it occurs in talk that can be associated with farming, AND when the variants co-present in the speech signal do not contradict that evaluation. If this is the case, then it suggests that perception studies that attend to both of these factors—co-present variants and the potential of message content to trigger stereotypes—will better account for the socially situated nature of language perception. Of course, we would not claim that it is not possible to deduce something about the social meaning of a variant from experimental abstraction, but precisely what that 'something' is is an open question, given that listeners in the real world never hear variants in the abstract. If we believe in the 'indexical mutability' of features (as proposed by Eckert 2012:94), then our research needs to do more than reaffirm meaning potentials in the abstract. Abstraction might suggest that a feature has meaning potential, but, as our study has shown, whether that meaning potential is realized depends upon the larger discourse frame in which the feature occurs.

5. CONCLUSION. Over the course of this article we have demonstrated that we are able to capture listeners' real-time reactions to known features (in the calibration test sample), as well as their reactions to features of naturally occurring speech in two Scillonian guises. Listeners demonstrated remarkably ordered responses to the examples of (th)-fronting in the test sample, and similar click reaction times over the course of the Farmer and Islander guises in the test proper suggested that there was a good amount of agreement over the groups of features noticed in each guise. The review mechanism of

the online tool allowed us to examine the reasons for clicks, and this provided further evidence to support our claim that we had accurately identified what listeners were responding to. Consequently, building on the work of others (Labov et al. 2011, Watson & Clark 2013, 2015), we argue that our tool offers a new technique to investigate salience, in addition to being of benefit to all kinds of research on speaker reactions, including work concerned with grammaticality judgments. The discrete nature of mouse clicks, as opposed to slider movements, means that the data presented here were less difficult to interpret than that obtained in other studies.[23] Of course, some parts of our analysis are partial; our click data showed where features were noticed by listeners, and consequently, in the review task, listeners can only have commented on features of which they were consciously aware (cf. Squires 2016). Nonetheless, we contend that this method affords the possibility of further exploring the social meaning of salient features (especially when listeners are presented with naturally occurring speech) due to the ability to link a definite event (a mouse click) to a feature's occurrence in time, and to use listeners' review comments to support the links between clicks and features.

In this way, being able to review the click data added a further analytical layer to our research. It demonstrated that, in the majority of cases, listeners were able to state why they had clicked where they did. There are, of course, some limitations to this approach. For instance, the review window may not have been long enough to capture the specific item that caused a listener to click, or listeners may have been disproportionately distracted by cooccurring features when presented with multiple features in the review window. Furthermore, we proposed that the early (and unexpected) occurrence of rhoticity in the Islander guise may have resulted in increased noticing of the first instance of this form in this guise. More testing, using manipulated samples, would enable us to test this hypothesis. Furthermore, we do not explore how the sequencing of variables that are less iconic than rhoticity affects expectations and subsequent filtering processes. While the deliberative process of the tasks (having listeners hear the sample in full once before identifying features in the second listening) may have reduced any effects caused by other features occurring in different sequences in the two guises, some sequencing effects may have occurred. The implications of general sequencing effects on expectation and filtering processes requires further empirical research. Although the limitations noted above are not insignificant, the coherent way that listeners were able to account for the vast majority of their clicks for the duration of both of the samples supports the utility of this approach when attempting to understand listeners' reactions. Consequently, the methodological approach described in this article provides a medium through which the limitations we have noted could be addressed in subsequent work using different kinds of data.

While we noted earlier that, in our study, both message content and co-present variants do evaluative and perceptual work, it is clear that the guises themselves, with their very different topics, appear to be primarily responsible for the regard in which they are held and the features that are recognized by listeners. As noted above, stereotypes and the expectations they activate appear to play a significant role (cf. Preston 2010, Rácz 2013). For the Farmer guise, features closely associated with traditional, rural, South Western varieties were recognized more readily, and ratings data showed the typical

---

[23] Our approach is slightly different from that taken in other studies, of course, and rather than gathering EVALUATIONS in real time we were collecting data on regional salience and using the separately gathered evaluation data to understand the potential social meaning of linguistic features.

pattern associated with a low-status speaker. By contrast, the Islander guise activated nonstandard features that were less readily stereotyped. Consequently, this guise was rated more in line with a higher-status speaker. This was despite the presence of rhoticity, which was noticed significantly more in the Islander guise due to its unexpected nature. The complex way in which topic, regard, and feature recognition interact supports Clopper and Pisoni's (2004a:44) assertion that 'the process of speech perception involves not only the segmentation of the speech signal into meaningful linguistic units (e.g., words, sentences) and the recovery of the structure of the sound patterns, but also the processing and encoding of indexical information about the talker'. These results are, of course, entirely in line with the findings of Campbell-Kibler 2009, 2010, Pharao et al. 2014, and Podesva et al. 2015, although we note that these studies preselect linguistic features for analysis. Our study suggests that linguistic features work synergistically in the perceptual process. Consequently, it is possible that the preselection of features may serve to reify the importance of some variables over others, ultimately underestimating the extent to which co-present variants contribute to situated meaning.

The importance of situated meaning to studies of perception has been stressed in the work of Campbell-Kibler (2016:128), who notes that the complexity of speaker/hearers' social abilities and responsibilities are far more complex than many models can currently account for. As we noted earlier, exemplar-based approaches are often cited as providing an elegant account of the way that social information is indexed to individual linguistic exemplars. However, their ability to capture a range of situated meanings may be limited. For example, our work suggests the need to model the effects of co-present variants on a given exemplar, not just the social indices of the specific exemplar itself.

The different responses to the Farmer and the Islander guises also adds to our understanding of the effects of enregisterment on the social meanings associated with linguistic features. The noticing of traditional South Western dialect features in the Farmer guise suggests that these features are enregistered, such that they are 'differentiable from the rest of the language (i.e., recognizable as distinct, linked to typifiable social personae or practices)' (Agha 2004:37), the 'typifiable persona' here being that of the stereotypical 'Farmer', as triggered by the content and topic of the extract used for this guise. The less extreme evaluation of the Islander guise suggests that—despite the presence of the same features—the activation of this register is blocked by the absence of 'typifiable' content of the talk itself. Once again, this exposes the fluidity of the social meaning of linguistic features and provides further support for our argument that our perception experiments need to be sensitive to contexts that trigger 'metapragmatic typification' (Agha 2004:29).

Ultimately, this discussion points to the importance of context in the study of the social perception of language. While it was already known that message content and setting can affect how speakers respond to speech samples (Ray et al. 1991, Cargile 1997), much of the research on sociolinguistic perceptions claims to control for context in an attempt to isolate responses to individual linguistic features. We would contend that it is not possible to escape context (even experiments that use single-word stimuli may exhibit effects of co-present phonetic or morphological items). Consequently, it will be essential to more fully reflect on the impact of context if we are to develop a more advanced understanding of the link between language, social meaning, and cognition. This is no straightforward task, given the difficulties in identifying precise effects in complex data. Nonetheless, we hope to have demonstrated in this article one way in which this may be possible.

APPENDIX: SUMMARY OF THE REPEATED-MEASURES LOGISTIC REGRESSION

Data: PALM

AIC	BIC	logLik	deviance	df.resid
317.4	328.7	−155.7	311.4	318

RANDOM EFFECTS

GROUPS	NAME	VARIANCE	SD	
ID	(intercept)	1.435	1.198	

Number of observations: 321; Groups: ID, 106

FIXED EFFECTS	ESTIMATE	SE	z-value	Pr(>|z|)
(intercept)	−1.1017	0.2422	−4.548	5.41e-06 ***
as.factor(Speaker)I	−1.9879	0.4472	−4.445	8.80e-06 ***

Data: PRICE

AIC	BIC	logLik	deviance	df.resid
1008.1	1023.8	−501.1	1002.1	1388

RANDOM EFFECTS

GROUPS	NAME	VARIANCE	SD	
ID	(intercept)	0.8539	0.9241	

Number of observations: 1,391; Groups: ID, 106

FIXED EFFECTS	ESTIMATE	SE	z-value	Pr(>|z|)
(intercept)	−1.7765	0.1559	−11.394	< 2e-16 ***
as.factor(Speaker)I	−0.9451	0.1754	−5.388	7.14e-08 ***

Data: TRAP

AIC	BIC	logLik	deviance	df.resid
377.7	391.1	−185.8	371.7	639

RANDOM EFFECTS

GROUPS	NAME	VARIANCE	SD	
ID	(intercept)	0.6245	0.7903	

Number of observations: 642; Groups: ID, 106

FIXED EFFECTS	ESTIMATE	SE	z-value	Pr(>|z|)
(intercept)	−2.0534	0.2176	−9.438	< 2e-16 ***
as.factor(Speaker)I	−2.4896	0.6034	−4.126	3.69e-05 ***

Data: BATH

AIC	BIC	logLik	deviance	df.resid
387.9	400.1	−191.0	381.9	425

RANDOM EFFECTS

GROUPS	NAME	VARIANCE	SD	
ID	(intercept)	0.6783	0.8236	

Number of observations: 428; Groups: ID, 106

FIXED EFFECTS	ESTIMATE	SE	z-value	Pr(>|z|)
(intercept)	−2.6614	0.3093	−8.603	< 2e-16 ***
as.factor(Speaker)I	1.7006	0.3103	5.481	4.23e-08 ***

Data: MOUTH

AIC	BIC	logLik	deviance	df.resid
658.3	673.0	−326.2	652.3	960

RANDOM EFFECTS

GROUPS	NAME	VARIANCE	SD	
ID	(intercept)	1.962	1.401	

Number of observations: 963; Groups: ID, 106

FIXED EFFECTS	ESTIMATE	SE	z-value	Pr(>|z|)
(intercept)	−3.4628	0.3103	−11.161	< 2e-16 ***
as.factor(Speaker)I	1.3773	0.2507	5.494	3.93e-08 ***

Data: STRUT

AIC	BIC	logLik	deviance	df.resid
113.5	124.8	−53.8	107.5	318

RANDOM EFFECTS

GROUPS	NAME	VARIANCE	SD	
ID	(intercept)	24.04	4.903	

Number of observations: 321; Groups: ID, 106

FIXED EFFECTS	ESTIMATE	SE	z-value	Pr(>|z|)
(intercept)	−6.564	1.546	−4.245	2.19e-05 ***
as.factor(Speaker)I	−1.722	0.944	−1.824	0.0681

Data: GOAT

AIC	BIC	logLik	deviance	df.resid
510.5	524.8	−252.3	504.5	853

RANDOM EFFECTS

GROUPS	NAME	VARIANCE	SD	
ID	(intercept)	0.05106	0.226	

Number of observations: 856; Groups: ID, 106

FIXED EFFECTS	ESTIMATE	SE	z-value	Pr(>|z|)
(intercept)	−2.8453	0.3213	−8.856	< 2e-16 ***
as.factor(Speaker)I	0.6073	0.3260	1.863	0.0625

Data: CHOICE

AIC	BIC	logLik	deviance	df.resid
142.0	153.3	−68.0	136.0	318

RANDOM EFFECTS

GROUPS	NAME	VARIANCE	SD	
ID	(intercept)	1.678	1.295	

Number of observations: 321; Groups: ID, 106

FIXED EFFECTS	ESTIMATE	SE	z-value	Pr(>|z|)
(intercept)	−4.0710	0.8269	−4.923	8.51e-07 ***
as.factor(Speaker)I	1.4141	0.5232	2.703	0.00688 **

Data: rhoticity

AIC	BIC	logLik	deviance	df.resid
2086.4	2103.5	−1040.2	2080.4	2244

RANDOM EFFECTS

GROUPS	NAME	VARIANCE	SD	
ID	(intercept)	0.7446	0.8629	

Number of observations: 2,247; Groups: ID, 106

FIXED EFFECTS	ESTIMATE	SE	z-value	Pr(>|z|)
(intercept)	−1.8927	0.1174	−16.120	< 2e-16 ***
as.factor(Speaker)I	0.6696	0.1144	5.854	4.81e-09 ***

REFERENCES

AGHA, ASIF. 2004. Registers of language. *A companion to linguistic anthropology*, ed. by Alessandro Duranti, 23–45. Oxford: Oxford University Press.

BABEL, ANNA M. (ed.) 2016. *Awareness and control in sociolinguistic research*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139680448.

BATES, DOUGLAS; MARTIN MAECHLER; and BEN BOLKER. 2011. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-42. Online: http://CRAN.R -project.org/package=lme4.

BIGAND, EMMANUEL; B. TILLMANN; BENÉDICTE POULIN; D. A. D'ADAMO; and FRANÇOIS MADURELL. 2001. The effect of harmonic context on phoneme monitoring in vocal music. *Cognition* 81.11–20. DOI: 10.1016/S0010-0277(01)00117-2.

BOERSMA, PAUL, and DAVID WEENINK. 2015. Praat: Doing phonetics by computer. Online: http://www.praat.org/.

Britain, David. 2017. Which way to look? Perspectives on 'urban' and 'rural' in dialectology. *Language and a sense of place*, ed. by Chris Montgomery and Emma Moore, 171–87. Cambridge: Cambridge University Press.

Brooks, Lee R. 1978. Nonanalytic concept formation and memory for instances. *Cognition and categorization*, ed. by Eleanor Rosch and Barbara Lloyd, 3–170. Mahwah, NJ: Lawrence Erlbaum.

Campbell-Kibler, Kathryn. 2006. *Listener perceptions of sociolinguistic variables: The case of (ING)*. Stanford, CA: Stanford University dissertation.

Campbell-Kibler, Kathryn. 2008. I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society* 37.637–59. DOI: 10.1017/S0047404508080974.

Campbell-Kibler, Kathryn. 2009. The nature of sociolinguistic perception. *Language Variation and Change* 21.135–56. DOI: 10.1017/S0954394509000052.

Campbell-Kibler, Kathryn. 2010. The effect of speaker information on attitudes toward (ING). *Journal of Language and Social Psychology* 29.214–23. DOI: 10.1177/0261927 X09359527.

Campbell-Kibler, Kathryn. 2011. Intersecting variables and perceived sexual orientation in men. *American Speech* 86.52–68. DOI: 10.1215/00031283-1277510.

Campbell-Kibler, Kathryn. 2016. Towards a cognitively realistic model of meaningful sociolinguistic variation. In Babel, 123–51. DOI: 10.1017/CBO9781139680448 .008.

Cargile, Aaron Castelan. 1997. Attitudes toward Chinese-accented speech: An investigation in two contexts. *Journal of Language and Social Psychology* 16.434–43. DOI: 10.1177/0261927X970164004.

Clopper, Cynthia G., and Ann R. Bradlow. 2009. Free classification of American English dialects by native and non-native listeners. *Journal of Phonetics* 37.436–51. DOI: 10.1016/j.wocn.2009.07.004.

Clopper, Cynthia G., and David B. Pisoni. 2002. Perception of dialect variation: Some implications for current research and theory in speech perception. *Research on Spoken Language Processing Progress Report* 25.269–90. Online: http://www.iu.edu/~srlweb /pr/25/clopper-269.pdf.

Clopper, Cynthia G., and David B. Pisoni. 2004a. Some new experiments on perceptual categorization of dialect variation in American English: Acoustic analysis and linguistic experience. *Research on Spoken Language Processing Progress Report* 26.29–46. Online: http://www.indiana.edu/~srlweb/pr/26/29-Clopper.pdf.

Clopper, Cynthia G., and David B. Pisoni. 2004b. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32.111–40. DOI: 10.1016/S0095-4470(03)00009-3.

Clopper, Cynthia G., and David B. Pisoni. 2006. Effects of region of origin and geographic mobility on perceptual dialect categorization. *Language Variation and Change* 18.193–221. DOI: 10.1017/S0954394506060091.

D'Onofrio, Annette. 2015. Persona-based information shapes linguistic perception: Valley Girls and California vowels. *Journal of Sociolinguistics* 19.241–56. DOI: 10.1111 /josl.12115.

Drager, Katie, and M. Joelle Kirtley. 2016. Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In Babel, 1–24. DOI: 10 .1017/CBO9781139680448.003.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12.453–76. DOI: 10.1111/j.1467-9841.2008.00374.x.

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41.87–100. DOI: 10 .1146/annurev-anthro-092611-145828.

Ellis, Alexander J. 1890. *English dialects—Their sounds and homes*. London: The English Dialect Society.

Frauenfelder, Uli H., and Juan Segui. 1989. Phoneme monitoring and lexical processing: Evidence for associative context effects. *Memory & Cognition* 17.134–40. DOI: 10 .3758/BF03197063.

Garrett, Peter. 2010. *Attitudes to language*. Cambridge: Cambridge University Press.

Giles, Howard. 1970. Evaluative reactions to accents. *Educational Review* 22.211–27. DOI: 10.1080/0013191700220301.

GILES, HOWARD, and PETER F. POWESLAND. 1975. *Speech style and social evaluation*. London: Academic Press.

HAY, JENNIFER, and KATIE DRAGER. 2010. Stuffed toys and speech perception. *Linguistics* 48.865–92. DOI: 10.1515/ling.2010.027.

HAY, JENNIFER; AARON NOLAN; and KATIE DRAGER. 2006. From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review* 23.351–79. DOI: 10.1515/TLR .2006.014.

HEATH, ROBERT. 1750. *A natural and historical account of the islands of Scilly; Describing their situation, number, extent, soil, culture, produce, rareties, towns, fortifications, trade, manufacture, inhabitants. Their government, laws, customs, grants, records, and antiquities*. London: R. Manby & H. S. Cox.

HOGG, MICHAEL, and GRAHAM VAUGHAN. 2008. *Social psychology*. 5th edn. Harlow: Pearson Education.

JOHNSON, KEITH. 1997. Speech perception without speaker normalization: An exemplar model. *Talker variability in speech processing*, ed. by Keith Johnson and John Mullennix, 145–65. San Diego: Academic Press.

KERSWILL, PAUL. 2003. Dialect levelling and geographical diffusion in British English. *Social dialectology: In honour of Peter Trudgill*, ed. by David Britain and Jenny Cheshire, 223–43. Amsterdam: John Benjamins.

KRISTIANSEN, TORE; PETER GARRETT; and NIKOLAS COUPLAND. 2005. Introducing subjectivities in language variation and change. *Acta Linguistica Hafniensia* 37.9–35. DOI: 10.1080/03740463.2005.10416081.

LABOV, WILLIAM; SHARON ASH; MAYA RAVINDRANATH; TRACEY WELDON; MACIEJ BARANOWSKI; and NAOMI NAGY. 2011. Properties of the sociolinguistic monitor. *Journal of Sociolinguistics* 15.431–63. DOI: 10.1111/j.1467-9841.2011.00504.x.

LAMBERT, WALLACE E.; RICHARD C. HODGSON; ROBERT C. GARDNER; and SAMUEL FILLENBAUM. 1960. Evaluation reactions to spoken languages. *The Journal of Abnormal and Social Psychology* 60.44–51. DOI: 10.1037/h0044430.

LEVON, EREZ. 2011. Teasing apart to bring together: Gender and sexuality in variationist research. *American Speech* 86.69–84. DOI: 10.1215/00031283-1277519.

LEVON, EREZ. 2014. Categories, stereotypes, and the linguistic perception of sexuality. *Language in Society* 43.539–66. DOI: 10.1017/S0047404514000554.

LEVON, EREZ. 2016. Teasing apart social and cognitive factors in sociolinguistic perception. Paper presented at the Sociolinguistic Variation and Language Processing Conference, Virginia Polytechnic Institute and State University.

LEVON, EREZ, and SUE FOX. 2014. Social salience and the sociolinguistic monitor: A case study of ING and TH-fronting in Britain. *Journal of English Linguistics* 42.185–217. DOI: 10.1177/0075424214531487.

LOVE, JESSICA, and ABBY WALKER. 2013. Football versus football: Effect of topic on /r/ realization in American and English sports fans. *Language and Speech* 56.443–60. DOI: 10.1177/0023830912453132.

MAGUIRE, WARREN; APRIL MCMAHON; PAUL HEGGARTY; and DAN DEDIU. 2010. The past, present, and future of English dialects: Quantifying convergence, divergence, and dynamic equilibrium. *Language Variation and Change* 22.69–104. DOI: 10.1017/S09543 94510000013.

MONTGOMERY, CHRIS. 2007. *Northern English dialects: A perceptual approach*. Sheffield: University of Sheffield dissertation. Online: http://etheses.whiterose.ac.uk/1203/.

MONTGOMERY, CHRIS. 2011. Sprachraum and its perception. *Language and space—An international handbook of linguistic variation, vol. 2: Language mapping*, ed. by Alfred Lameli, Roland Kehrein, and Stefan Rabanus, 586–605. Berlin: De Gruyter Mouton.

MONTGOMERY, CHRIS. 2012. The effect of proximity in perceptual dialectology. *Journal of Sociolinguistics* 16.638–68. DOI: 10.1111/josl.12003.

MOORE, EMMA, and PAUL CARTER. 2015. Dialect contact and distinctiveness: The social meaning of language variation in an island community. *Journal of Sociolinguistics* 19.3–36. DOI: 10.1111/josl.12107.

MOORE, EMMA, and PAUL CARTER. 2018. Natural phonetic tendencies and social meaning: Exploring the allophonic raising split of PRICE and MOUTH on the Isles of Scilly. *Language Variation and Change*, to appear.

Moore, Emma, and Chris Montgomery. 2018. The dialect of the Isles of Scilly: Exploring the relationship between language production and language perception in a southern insular variety. *Southern English varieties then and now*, ed. by Laura Wright, 39–73. Berlin: De Gruyter Mouton.

Moore, Emma, and Robert Podesva. 2009. Style, indexicality, and the social meaning of tag questions. *Language in Society* 38.447–85. DOI: 10.1017/S0047404509990224.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18.62–85. DOI: 10.1177/0261927X99018001005.

Niedzielski, Nancy, and Dennis R. Preston. 2003. *Folk linguistics*. Berlin: Mouton de Gruyter.

Niemi, Pekka, and Risto Näätänen. 1981. Foreperiod and simple reaction time. *Psychological Bulletin* 89.133–62. DOI: 10.1037/0033-2909.89.1.133.

Nycz, Jennifer. 2016. Awareness and acquisition of new dialect features. In Babel, 62–79. DOI: 10.1017/CBO9781139680448.005.

ONS Geography. 2010. Regions (former GORs). Newport: Office for National Statistics. Online: http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/administrative/england/government-office-regions/index.html.

Pachella, Robert G. 1973. *The interpretation of reaction time in information processing research*. Technical report. Ann Arbor: University of Michigan Human Performance Center.

Paltridge, John, and Howard Giles. 1984. Attitudes towards speakers of regional accents of French: Effects of regionality, age and sex of listeners. *Linguistische Berichte* 90.71–85.

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. DOI: 10.1017/S0047404513000857.

Piercy, Caroline. 2010. *One /a/ or two? The phonetics, phonology and sociolinguistics of change in the TRAP and BATH vowels in the southwest of England*. Essex: University of Essex dissertation.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. *Frequency and the emergence of linguistic structure*, ed. by Joan Bybee and Paul Hopper, 137–57. Amsterdam: John Benjamins.

Plichta, Bartlomiej, and Dennis R. Preston. 2005. The /ay/s have it: The perception of /ay/ as a north-south stereotype in United States English. *Acta Linguistica Hafniensia* 37.107–30. DOI: 10.1080/03740463.2005.10416086.

Podesva, Robert J.; Jermay Reynolds; Patrick Callier; and Jessica Baptiste. 2015. Constraints on the social meaning of released /t/: A production and perception study of U.S. politicians. *Language Variation and Change* 27.59–87. DOI: 10.1017/S0954394514000192.

Preston, Dennis R. 1989. *Perceptual dialectology: Non-linguists' view of aerial linguistics*. Dordrecht: Foris.

Preston, Dennis R. 2010. Variation in language regard. *Variatio delectat: Empirische Evidenzen und theoretische Passungen sprachlicher Variation (für Klaus J. Mattheier zum 65. Geburtstag)*, ed. by Peter Gilles, Evelyn Ziegler, and Joachim Scharloth, 7–27. Frankfurt am Main: Peter Lang.

Preston, Dennis R. 2011. The power of language regard: Discrimination, classification, comprehension, and production. *Dialectologia* Special issue II.9–33. Online: http://www.publicacions.ub.edu/revistes/dialectologiaSP2011/.

Preston, Dennis R. 2015. Does language regard vary? *Responses to language varieties: Variability, processes and outcomes*, ed. by Alexei Prikhodkine and Dennis R. Preston, 1–36. Amsterdam: John Benjamins.

Rácz, Péter. 2013. *Salience in sociolinguistics: A quantitative approach*. Freiburg: University of Freiburg dissertation.

Ray, George B.; Eileen Berlin Ray; and Christopher J. Zahn. 1991. Speech behavior and social evaluation: An examination of medical messages. *Communication Quarterly* 39.119–29. DOI: 10.1080/01463379109369790.

RICKFORD, JOHN R., and FAYE MCNAIR-KNOX. 1994. Addressee- and topic-influenced style shift: A quantitative sociolinguistic study. *Sociolinguistic perspectives on register*, ed. by Douglas Biber and Edward Finegan, 235–76. Oxford: Oxford University Press.

RYAN, ELLEN BOUCHARD, and HOWARD GILES. 1982. *Attitudes towards language variation: Social and applied contexts*. London: Hodder Arnold.

SHARMA, DEVYANI. 2011. Style repertoire and social change in British Asian English. *Journal of Sociolinguistics* 15.464–92. DOI: 10.1111/j.1467-9841.2011.00503.x.

SQUIRES, LAUREN. 2016. Processing grammatical differences: Perceiving versus noticing. In Babel, 80–103. DOI: 10.1017/CBO9781139680448.006.

STUART-SMITH, JANE; GWILYM PRYCE; CLAIRE TIMMINS; and BARRIE GUNTER. 2013. Television can also be a factor in language change: Evidence from an urban dialect. *Language* 89.501–36. DOI: 10.1353/lan.2013.0041.

THOMAS, CHARLES. 1979. A glossary of spoken English in the Isles of Scilly. *Journal of the Royal Institute of Cornwall* 8.109–47.

TRUDGILL, PETER. 1999. *The dialects of England*. 2nd edn. Oxford: Blackwell.

WAGNER, LAURA; CYNTHIA G. CLOPPER; and JOHN K. PATE. 2014. Children's perception of dialect variation. *Journal of Child Language* 41.1062–84. DOI: 10.1017/S0305000913 000330.

WAKELIN, MARTYN F. 1986. *The Southwest of England*. Amsterdam: John Benjamins.

WATSON, KEVIN, and LYNN CLARK. 2013. How salient is the NURSE~SQUARE merger? *English Language & Linguistics* 17.297–323. DOI: 10.1017/S136067431300004X.

WATSON, KEVIN, and LYNN CLARK. 2015. Exploring listeners' real-time reactions to regional accents. *Language Awareness* 24.38–59. DOI: 10.1080/09658416.2014.882346.

WELLS, JOHN C. 1982. *Accents of English*. 3 vols. Cambridge: Cambridge University Press.

WILLIAMS, ANGIE; PETER GARRETT; and NIKOLAS COUPLAND. 1999. Dialect recognition. *Handbook of perceptual dialectology*, ed. by Dennis R. Preston, 345–58. Amsterdam: John Benjamins.

WOODLEY, GEORGE. 1822. *View of the present state of the Scilly Islands: Exhibiting their vast importance to the British Empire*. London: Longman and Co.

[c.montgomery@sheffield.ac.uk]
[e.moore@sheffield.ac.uk]