Speaker gender and salience in sociolinguistic speech perception: GOOSE-fronting in Standard Southern British English

Roy Alderton

University of Sheffield, UK

Author-accepted manuscript accepted for publication in *Journal of English Linguistics*, SAGE Publications.

Abstract

Listeners' perceptions of sound changes may be influenced by priming them with social information about the speaker. It is not clear, however, whether this occurs for sociolinguistic variables that pass below the level of awareness. This article investigates whether visual speaker gender affects the perception of GOOSE-fronting in Standard Southern British English, a sound change that is led by young women yet does not fulfil criteria for sociolinguistic salience. Participants from across the United Kingdom completed a word identification experiment based on a gender-ambiguous synthesised FLEECE-GOOSE continuum while primed with an image of a man's or a woman's face. The study did not find a significant main effect of priming, but men identified fronter tokens as GOOSE when primed with a woman's face. I argue that sociolinguistic priming effects may be over-stated and that future priming experiments should be designed with maximal statistical power where possible.

Keywords

gender, salience, sociophonetics, speech perception, visual priming

1. Introduction

Research in sociophonetics has demonstrated that people are sensitive to sociolinguistic variation when perceiving spoken language. It has shown that the socio-indexical information contained in the visual appearance of the speaker can prime listeners' perceptions of phonetic variables (Niedzielski 1999; Hay, Nolan & Drager 2006; Hay, Warren & Drager 2006). Some recent work has failed to replicate these priming effects (Squires 2013; Lawrence 2015b; Walker, Szakay & Cox 2019), with one argument being that priming may occur only if the linguistic variable under study has stereotyped or 'salient' social meanings among listeners (Juskan 2016). Accordingly, the present study investigates whether sociolinguistic priming occurs for a low-salience sound change in progress, namely GOOSE-fronting in Standard Southern British English (SSBE). I use a vowel categorisation experiment with participants from around the United Kingdom (UK) to test the effect of visual speaker gender on perceptions of this sound change, which studies suggest is led by women (Tollfree 1999; Williams & Kerswill 1999; Przedlacka 2002). If listeners primed with a woman's face are able to categorise more acoustically fronted tokens as part of the GOOSE set, it would suggest that sociolinguistic priming can influence perceptions of less salient phonetic variation, lending support to models of speech perception that propose that fine social and phonetic detail are stored together (Pierrehumbert 2001; Foulkes & Docherty 2006).

2. Sociolinguistics and speech perception

Various social characteristics of the speaker have been shown to affect speech perception via visual priming, including their age (Harrington, Kleber & Reubold 2007, 2008; Drager 2011), region of origin (Niedzielski 1999; Hay, Nolan & Drager 2006), socioeconomic class (Hay, Warren & Drager 2006; Lawrence 2017), ethnicity (Staum Casasanto 2008), gender

(Johnson, Strand & D'Imperio 1999; Strand 1999; Munson 2011; Winn, Rhone, Chatterjee & Idsardi 2013; Munson, Ryherd & Kemper 2017), and stereotypical persona (D'Onofrio 2018; Villarreal 2018). Studies such as these frequently employ a between-subjects design in which different groups within the participant sample are tasked with categorising the same linguistic variants (e.g., a vowel continuum) while being exposed to different visual primes, such as a photo of the purported speaker that depicts information such as their age or gender. For example, in Drager (2011), participants primed with a photo of a younger speaker heard more raised tokens along the DRESS-TRAP continuum produced by a speaker of New Zealand English as TRAP than those primed with an older speaker; this reflects younger people's greater usage of TRAP-raising in this variety. Some studies have found that even indirect primes such as the names of regions on an answer sheet (Niedzielski 1999; Hay, Nolan & Drager 2006), the accent of the experimenter (Hay, Drager & Warren 2010), or even soft toys in the experiment room (Hay & Drager 2010) can have an effect on where people place the boundary between vowels along a continuum.

The results of these experiments have theoretical implications for understanding how people process sounds and how they associate certain social information with very specific phonetic features. Many studies of sociolinguistic perception give support to exemplar models of speech production and perception (Goldinger 1996; Pierrehumbert 2001; Foulkes & Docherty 2006; Johnson 2006), which claim that the utterances a listener hears are subconsciously stored and categorised in the mind as individual 'exemplars' of an abstracted category (e.g., many tokens of a phoneme like /uː/). Exemplars are stored together with detailed information about the phonetic properties of the speech signal, the social characteristics of the speaker, and other contextual information. When a new utterance is heard, existing stored exemplars are activated and used to categorise the new utterance based upon its similarity to the phonetic and social information of tokens encountered previously (Pierrehumbert 2001). Exemplars are activated faster and more strongly depending on various linguistic, cognitive, and contextual factors, such as how frequently they are usually activated and how acoustically and socially similar they are to the incoming utterance (Foulkes & Docherty 2006; Drager 2015). This theory is supported by many experimental speech perception studies: listeners are more likely to perceive a phonetic contrast in a certain way if the social information in the visual primes matches that of the production patterns to which they have been exposed in their interactions with others (Strand & Johnson 1996; Niedzielski 1999; Hay, Warren & Drager 2006).

Some recent work, however, such as Squires (2013), Chang (2015), Lawrence (2015b), Juskan (2016), and Walker, Szakay, and Cox (2019), has not found clear evidence for priming effects in the expected direction. Walker, Szakay, and Cox (2019) question certain aspects of the priming paradigm itself, with reference to issues of limited working memory on the part of listeners, the artifice of synthetic vowel continua, and a lack of statistical power caused by small numbers of participants and stimuli. Other scholars have interpreted sociolinguistic priming effects (or the absence thereof) as a result of differences in salience between linguistic variables (Drager 2011, 2015; Juskan 2016). The term ‘salience’ is widely used in sociolinguistics but has proved difficult to pin down (Kerswill & Williams 2002). Broadly speaking, there is agreement that salient features ‘stick out’ to some degree as a result of language-internal and language-external factors. This may occur on a cognitive level in the form of a salient feature being unexpected and therefore surprising in that context (Rácz 2013; Jaeger & Weatherholtz 2016), or on a social level via a variant’s ability to index social meaning and potentially become a stereotype in a speech community (Levon & Fox 2014; Schleef 2017).

Drager (2011, 2015) argues that salience may play a role in exemplar activation, as exemplars that are more prominent or surprising given the context require more cognitive attention to process and so are accorded a greater weight (Nosofsky 1986). Hence, more salient variables are said to have a stronger relationship between their phonetic and social information, thus resulting in faster and stronger activation. This also means that variants that are not part of a speaker's own linguistic repertoire (or that of their social group) may be more salient. This is because these 'non-native' exemplars are more surprising in comparison to the speaker's most commonly-encountered exemplars, which strengthens the link between their social and phonetic characteristics.

Juskan (2016), however, situates salience as a key factor in explaining recent failures to replicate the priming effects found in earlier studies. That is, in order for listener perceptions in priming experiments to match the direction of the variation in production, he argues that the stimuli must be based on a highly salient linguistic variable that listeners already stereotypically associate with the visual prime. In addition to high salience, he also suggests (i) that positive priming effects may only be obtained for linguistic variables that vary on a continuum (e.g., most vowels) rather than in a discrete fashion (e.g., most consonants); and (ii) that the variables must not be so radically unexpected in relation to the visual information that listeners will not be convinced by the stimulus-prime combination. This would account for the positive priming effects found for MOUTH-raising in Canada (Niedzielski 1999) and various vowel changes in New Zealand (Hay, Warren & Drager 2006; Drager 2011), which fulfil these criteria. Juskan (2016) claims that British English varieties, however, typically have various distinctive local accent features of different levels of salience that do not always form gradient continua. He suggests that these therefore pose a challenge in 'persuading' listeners via priming that, for instance, /k/ lenition could be produced by a speaker from

Manchester rather than Liverpool. The same situation occurs for the notion that that the BATH-TRAP merger could be produced by a speaker from the South of England rather than the North (as in Lawrence 2015b).

The present study contributes to this discussion by testing the idea that less salient variables do not show significant priming effects. It avoids the kind of 'implausible' priming caused by presenting dissonant auditory and visual stimuli. Instead, it examines the effect of speaker gender on a variable that is used by both men and women, but is led by the latter (see section 3). In doing so, the study also addresses the question of whether sociolinguistic priming effects occur with variables for which the sound change is led by a particular gender, which, to my knowledge, is yet to be tested.

3. Production and perception of GOOSE-fronting

The /uː/ vowel, also known as the GOOSE vowel in Wells' (1982) lexical sets of the vocalic phonemes of English, is traditionally described as a high back rounded vowel. However, over the past few decades, this vowel has been realised increasingly further forwards in the vowel space. This phenomenon is known as GOOSE-fronting. In England, GOOSE-fronting has been studied in various localities, including Nottingham (Flynn 2012), York (Haddican, Foulkes, Hughes & Richards 2013), and Derby (Sóskuthy, Foulkes, Hughes & Haddican 2018). It appears to be most advanced, however, in accents of the South of England, namely in Standard Southern British English (SSBE) (Harrington, Kleber & Reubold 2007, 2008, 2011; Ferragne & Pellegrino 2010; Williams & Escudero 2014), and in related accents of the region such as that of south-east London (Tollfree 1999), Reading and Milton Keynes (Williams & Kerswill 1999), Hastings (Holmes-Elliott 2015), and the Home Counties in general (Przedlacka 2002).

In terms of their acoustic properties, fronted tokens of GOOSE often have very high second formants (F2) that may even overlap with those of the high front vowel FLEECE /iː/ (Harrington, Kleber & Reubold 2008). What distinguishes fronted GOOSE from FLEECE in articulation and acoustics is a matter of debate, as scholars have investigated various factors including the roles played by tongue position, lip rounding, and F2 slope (Chládková & Hamann 2011; Harrington, Kleber & Reubold 2011; Chládková, Hamann, Williams & Hellmuth 2017). In some varieties of British English, diphthongal variants of GOOSE have been identified, though this appears to be more common in northern accents compared to southern ones (Ferragne & Pellegrino 2010; Williams & Escudero 2014).

Previous literature indicates that young people—particularly young women—lead GOOSE-fronting in southern English accents (Tollfree 1999; Williams & Kerswill 1999; Przedlacka 2002; Harrington, Kleber & Reubold 2008; Holmes-Elliott 2015). This mirrors one of Labov's (2001) overall principles of linguistic change, which is that young women are typically the main leaders of change from below the level of consciousness. The matter of speakers' 'consciousness' or 'awareness' of linguistic changes has often been discussed using the notion of salience, as discussed in section 2. In terms of the salience of GOOSE-fronting in SSBE, evidence would suggest that it has low salience regardless of one's definition of salience. Trudgill (1986) lists several 'objective' language-internal factors as criteria for salience, such as whether the variation is reflected in the orthography, shows radical phonetic differences, or is involved in maintaining phonological contrasts. GOOSE-fronting shows none of these, as it is a phonetically gradient change that does not violate any contrasts in English phonology and is not represented in orthography. Regarding extra-linguistic factors, such as a variant's level of stigmatisation or its patterns of sociolinguistic variation (Kerswill & Williams 2002), we have already seen that GOOSE-fronting is socially stratified by age and gender in the South of England. Unlike similar variables like /t/-glottalling, however, it does not seem to attract metalinguistic comment from speakers (Alderton 2019). Based on this evidence, it is reasonable to conclude that GOOSE-fronting is not a very salient sociolinguistic variable in SSBE, even though it is undergoing change. Although listeners may not be consciously aware of its sociolinguistic patterns, it is still possible that these may be stored sub-consciously in the minds of speakers. This is what is posited by exemplar theories and is supported by various priming studies, as described in section 2, so it is worthy of further testing with experimental perception methods.

Existing literature on the perception of GOOSE-fronting offers a mixed picture of listener sensitivity towards its socio-indexical cues. In some studies, GOOSE-fronting is not linked to any regional or social associations in listener perceptions (Fridland, Bartlett & Kreuz 2004, 2005; Fridland 2008). For example, Haddican, Foulkes, Hughes, and Richards (2013) argue that in York, back vowel fronting is so widespread that it does not encode any social meaning, unlike other changes in the city such as FACE and GOAT monophthongisation. In other research, however, a perceptual link does seem to exist between GOOSE-fronting and speakers' social information (Torbert 2004; Lawrence 2017; Villarreal 2018). Listeners in Villarreal (2018) associated GOOSE-fronting, in combination with TRAP-backing, to a Californian regional identity and to a Valley Girl persona. Similarly, Lawrence (2017) finds that listeners perceptually associate fronted GOOSE, alongside fronted GOAT, with middle-class speakers in York. Backed and diphthongal local variants are linked to the working class and to a local stereotypical persona, the 'chav' (a young, anti-social working-class person who wears designer sportswear). However, these social meanings were strongest among younger and more geographically mobile listeners. For older and less mobile participants, they were weaker and in the opposite direction regarding the chav persona. This problematises a straightforward one-to-one relationship between production and perception of GOOSE-fronting while also showing the variable nature of sociolinguistic perception more broadly. In both Villarreal's (2018) and Lawrence's (2017) research, listeners were responding to speakers from their own community, and thus the social meanings of GOOSE-fronting present in listener perceptions may potentially have stemmed from a well-developed awareness of these meanings of local speech. The only work on GOOSE-fronting in SSBE specifically is Harrington, Kleber, and Reubold (2007, 2008), who find that the perceptual boundary between FLEECE and GOOSE in SSBE is fronter for younger speakers than for older speakers. These perception results are mirrored in the participants' production, as younger speakers' GOOSE vowels partially overlapped with the vowel space of FLEECE, while this did not occur for the older speakers.

The literature suggests, then, that GOOSE-fronting in SSBE is not a very salient feature overall but may have social meanings for particular groups depending on the local context. This supports recent research on indexicality which claims that the salience and social meanings of phonetic variables are not fixed properties of a variable but are dynamic and contingent on social and contextual factors (Levon & Fox 2014; Llamas, Watt & MacFarlane 2016; Schleef 2017). The studies that show the strongest indexical associations for GOOSE-fronting involve listeners from specific places rating the speech of their home city or region and the local personae within it, where the change is perceived inconsistently between different groups in the community (Lawrence 2017; Villarreal 2018). This indicates that listeners in these investigations may be responding to locally meaningful combinations of features or styles that are specific to particular groups within that area, rather than universally stereotyped features.

In light of the previous work considered to this point, the present study seeks to answer two main questions. First, does the socio-indexical information contained in a visual prime affect a low-salience sound change like GOOSE-fronting in SSBE? Second, does gender affect the perception of an ongoing sound change in the same way as other drivers of sociolinguistic variation such as age and social class? The analysis is presented in two phases: a confirmatory stage, which directly addresses the effect of experimental condition on listener perception as described above, and an exploratory stage, which examines how the visual prime interacts with social characteristics of the participants, in order to observe more detailed patterns that can be further tested in future work. The advantages of this approach are discussed in section 4.2.

4. Methodology

4.1. Experimental design

The study took the form of a between-subjects binary forced-choice word identification task with stimuli based on a FLEECE-GOOSE vowel continuum. Each participant was randomly assigned to either the 'man's face' or 'woman's face' condition so that the sample was evenly split between both conditions, yet all listeners heard the same set of auditory stimuli. The study design required the use of stimuli that were ambiguous with respect to perceived differences between stereotypical male and female voices. This was done by modifying the fundamental frequency of the recorded speech stimuli so that it lay between the typical male and female pitch ranges, creating gender-ambiguous stimuli. An alternative option is to present participants with both a male and a female voice, although this raises the issue of how to account for the gender prototypicality of each voice (Johnson, Strand & D'Imperio 1999). While the notion that simply changing the fundamental frequency of a recording makes it sound gender-ambiguous has been critiqued (Hillenbrand 2005; Assmann, Dembling &

Nearey 2006), the successful use of this technique by Johnson, Strand, and D'Imperio (1999) and its ability to deal with physiological sexual differences using only one set of stimuli makes it suitable for this study.

Four speakers of SSBE (three men and one woman) were recorded producing FLEECE-GOOSE minimal pairs. The SSBE accent was chosen for several reasons, some discussed earlier: (i) SSBE appears to show increased GOOSE-fronting compared to other British English varieties; (ii) several studies have found that young women lead the change in this and related southern accents; and (iii) the widespread use of SSBE in British broadcast media means that listeners will have regular exposure to it regardless of their regional background (though this is likely to vary; see section 4.2). In order to make the voices sound gender-ambiguous, the fundamental frequency of the stimuli was modified using the Praat phonetics software (Boersma & Weenink 2015). Sex differences in fundamental frequency vary across languages (Traunmüller & Eriksson 1995), and in Henton's (1989) review of several studies of male and female speech, average pitches vary considerably. Most of the studies reviewed by Henton give typical pitch averages for English speakers as around 90-140 Hz for men and 170-260 Hz for women. Simpson (2009) narrows down the average English pitch range to 100-120 Hz for men and 200-220 Hz for women. The pitch of the vowel midpoints in the three men's original recordings fell between 105-130 Hz, so I increased the pitch of these stimuli by 60 Hz, resulting in frequencies around 165-190 Hz, which is situated in between the typical male and female ranges for English, or at least towards the very low end of the female range. A similar procedure was applied to the woman's data, the fundamental frequency of which was reduced from 165-210 Hz by 60 Hz. Individual formant frequencies for F1-F3 were not manipulated despite the correlation between higher fundamental frequency and higher formants (Hillenbrand 2005; Assmann, Dembling & Nearey 2006), since evidence suggests that F0 plays the biggest role in influencing perception of men's and women's voices (Skuk & Schweinberger 2014).

The stimuli were designed to form an eleven-step continuum between the two parts of a FLEECE–GOOSE (/iː/–/uː/) minimal pair in order to convey more fronted and more backed realisations of GOOSE, as in Harrington, Kleber, and Reubold (2008). A synthetic vowel continuum for each speaker was created using a vowel resynthesis script in Praat (written by and described in Lawrence 2015a). The script determines the formant values for the steps of the continuum based on those of two parts of an existing minimal pair, which are then overwritten by the synthetic end-point steps together with the steps in between. It uses linear predictive inverse-filtering to restore the high-frequency component of the sounds after synthesis in order to make them sound more natural, based on work by Alku, Tiitinen, and Näätänen (1999).

The end-points of the continuum were recordings of the minimal pair *bee* and *boo* (*boo* being a member of the GOOSE lexical set) after having undergone pitch shifting. This pair of words was selected as it does not contain any voiceless (or aspirated) sounds, which resynthesis scripts struggle to process due to the lack of resonance caused by vocal fold vibration. It also lacks word-final consonants, which would shorten the duration of the vowel and make it less audible to listeners. Preceding /b/ is a context in which GOOSE is less likely to be fronted than in post-coronal contexts, especially after preceding /j/, but it can still be fronted as long as it does not precede syllable-final /l/ (Sóskuthy, Foulkes, Hughes & Haddican 2018). Since the procedure caused some distortion to the [b] consonants preceding the vowels, the consonants were spliced from the original pitch-shifted recordings into the stimuli to enhance the audibility of these sounds and remove the distortion.

Bee and *boo* are not perfect as stimulus words, as they differ in grammatical function (as a standard common noun versus an interjection or slang common noun respectively) and [b], as with many word-initial obstruents in English, is rarely fully voiced. However, alternative consonant-vowel FLEECE-GOOSE minimal pairs were less viable because they contain voiceless sounds that could disrupt vowel resynthesis, or they are based on very rare words such as *ye* in *ye-you* and *moo* in *me-moo*. In comparison, the difference in frequency between *bee* and *boo* is only 103 occurrences versus 90 respectively in the Spoken British National Corpus 2014 (Love, Dembry, Hardie, Brezina & McEnery 2017). Using non-words would partially deal with this issue, but this has rarely been done in priming experiments, and what might sound like nonsense to one listener may be a real word to another.

The F0, F1, F2, and F3 formant frequencies in Hertz of the vowel midpoints for each step along the pitch-shifted vowel continuum used in the experiment (selected via an online survey, discussed below) are shown in Table 1. F1 and F2 are commonly interpreted to correspond to vowel height and frontness respectively, though this is not a one-to-one mapping, and formants may be influenced by overall vocal tract shape and lip rounding (Docherty, Foulkes, Gonzalez & Mitchell 2018). The GOOSE end-point is somewhat central, resembling an [ʉ] in auditory quality, which is typical in contemporary SSBE. The use of a back [u] vowel for the GOOSE vowel would have allowed for a fuller representation of the front-back dimension of the vowel space, but would not have sounded natural or convincing from the mouth of a young SSBE speaker. The F2 values for the FLEECE and GOOSE endpoints are broadly similar to those reported in previous acoustic studies of the vowels of RP and SSBE speakers (e.g., Deterding 1997; Hawkins & Midgley 2005; Bjelaković 2017). F1 values suggest a lower vowel height than is typical, which may be an artefact of the process of vowel synthesis or an individual idiosyncrasy. In either case, F1 is less important for this study than F2 since it is vowel fronting that is of interest, which is conceived acoustically through the latter. F2 slope, which may act as a perceptual cue distinguishing FLEECE and GOOSE (Chládková, Hamann, Williams & Hellmuth 2017), was consistently flat (i.e., monophthongal) for all steps along the continuum. In Table 1 and in the remainder of the article, a higher continuum step represents a more GOOSE-like realisation. The F0 for all steps along the continuum is between 167 and 169 Hz.

TABLE 1

F0-F3 Formant Frequencies for FLEECE-GOOSE Continuum (Hz)

Continuum step	**F0**	**F1**	**F2**	**F3**	**Difference in F2 versus previous step**
1 (FLEECE endpoint)	168	400	2360	2803	
2	168	400	2247	2787	113
3	167	402	2174	2837	73
4	167	406	2081	2844	93
5	167	408	2006	2923	75
6	167	421	1911	2954	95
7	167	433	1846	3036	65
8	167	428	1763	3045	83
9	167	431	1665	3067	98
10	167	438	1611	3118	54
11 (GOOSE endpoint)	169	473	1535	3132	76

The selection of which set of stimuli to use in the experiment was informed by the results of an online survey. Thirteen British undergraduate students aged 19-31 (ten women, three men, from a variety of regional backgrounds) were asked to rate the four pairs of end-points from the synthesised continua on their comprehensibility, naturalness, and gender ambiguity using Likert scales, shown in Table 2. The speaker whose stimuli were rated closest to 1 for comprehensibility and naturalness (i.e., closest to the 'extremely easy to tell' and 'extremely natural' ends of the scales), and closest to 3 for gender ambiguity (i.e., closest to the middle of the scale), was selected for use in the final experiment. This speaker was a 22-year-old male SSBE speaker from Hampshire (South East England), the pitch of whose speech had been increased by 60 Hz prior to testing, as described above. The mean ratings for his speech were 2 for comprehensibility, 4.4 for naturalness, and 2.9 for gender ambiguity. This suggests that the listeners found it quite easy to tell that the words in the stimuli were *bee* and *boo* and that the gender of the speaker was not obvious. The participants did not consider the stimuli to sound particularly natural, but this is perhaps understandable given the digital manipulation needed to shift the pitch and create the vowel continuum.

TABLE 2

Reproduction of likert scales used in the stimulus quality survey to test comprehensibility, naturalness, and gender ambiguity

	1	**2**	**3**	**4**	**5**	**6**
The words were…	Extremely easy to tell	Quite easy to tell	Slightly easy to tell	Slightly hard to tell	Quite hard to tell	Extremely hard to tell
The speaker's voice sounds…	Extremely natural	Quite natural	Slightly natural	Slightly artificial	Quite artificial	Extremely artificial
Speaker's gender	Definitely male	Probably male	Not sure—could be either	Probably female	Definitely female	

This survey was also used to test the quality of the face photographs used as primes in the experiment. Eight photos (four for each gender) were selected from the Psychological Image Collection at Stirling online database (2013). Since gender was the main speaker characteristic under investigation, it was important that the photos did not drastically vary in terms of other variables such as age and ethnicity, so the images used were all of white people in their early twenties, which is consistent with the speakers who were recorded for the audio stimuli. The survey respondents were asked to rank the four images for each gender according to their prototypicality as a young person from the South of England. The photo of a woman with the highest prototypicality rating was chosen for use in the experiment. The image of a man with the highest prototypicality rating was not considered to be comparable to the woman due to noticeable differences in appearance and facial expression, so the man's photo that came second in terms of prototypicality was chosen.

The experiment itself was run using the software PsychoPy (Peirce 2007) in a quiet room. The same native SSBE-speaking experimenter conducted each session. After being informed of the nature of the task, the participants were given instructions on the computer and shown a photo of a person's face corresponding to the condition to which they were assigned. Listeners heard three instances of each of the eleven tokens, which were pseudo-randomised so that each step was only repeated after the entire set had been played, minimising the risk of immediate repetition of the same step. After hearing a token, they pressed a key on the keyboard which corresponded to the word they thought they heard (*boo* versus *bee*). Distractor stimuli were not required for this experiment as participants' awareness of the target vowels had minimal implications for the results, and their inclusion would have lengthened the duration of the study, risking listener fatigue. The photo of the speaker was shown as part of the instructions before the task started and then remained on the screen for the whole duration of the time during which participants were identifying the tokens. The words *bee* and *boo* and their corresponding keys were also visible for the duration of the experiment, and were kept constant across the participant sample. After finishing the perception task, participants completed a brief questionnaire about their background (age, gender, home town, etc.). This was done after the experiment in order to avoid priming the participants with words referring to social characteristics such as place names (Niedzielski 1999; Hay, Nolan & Drager 2006).

Fifty-two people took part in the study. All were native speakers of English born in the UK, aged between 19 and 30 years old (mean = 22.4), recruited via departmental emails and social media. The participants were all living in Lancaster (North West England) at the time of the experiment but were from a wide variety of geographical origins within the country. Little work has been conducted on the accent of Lancaster and the surrounding area, though research further north in Cumbria suggests that GOOSE-fronting has reached urban, but not rural areas (Jansen 2018). In any case, the vast majority of listeners were not locals of the city and had moved there from elsewhere in the UK to study at university. As noted above, SSBE is frequently encountered in British media broadcasts and through geographical mobility, yet listeners with different regional origins are likely to have different levels of exposure to it and hence different activation rates for their stored exemplars of GOOSE in this accent. Participants from the North of England were over-represented in the sample, particularly in the woman's face condition (see Appendix 1). For these reasons, geographical origin is accounted for as a fixed effect in the exploratory statistical analysis (see section 4.2). The sample was equally split according to gender across the two conditions, so that each condition had twenty-six participants, thirteen of whom were men and thirteen of whom were women.

4.2. Statistical analysis

Experimental studies of speech perception, and other quantitative research in phonetics more generally, are often investigated using statistical methods such as linear mixed-effects regression modelling (Baayen 2008). While these powerful quantitative techniques have been helpful in allowing for rigorous analysis of linguistic data, recent work in the field has identified several issues with how statistical analyses are commonly reported in published material, with particular focus on the distinction (or lack thereof) between confirmatory and exploratory data analysis (Baayen, Vasishth, Kliegl & Bates 2017; Roettger 2019; Roettger, Winter & Baayen 2019). Confirmatory analysis seeks to directly address a specific hypothesis with one model that includes only the terms needed to answer the research question. Exploratory analysis, on the other hand, may use several models to explore all the patterns in the data in order to observe potential effects and generate hypotheses that can be investigated in a future study with its own confirmatory analysis (Roettger, Winter & Baayen 2019). In a world in which positive, confirmatory results are much more likely to be published than negative or exploratory findings, scholars are under pressure to maximise their 'researcher degrees of freedom' (Roettger 2019), one of which is to present exploratory results as confirmatory. This can potentially involve problematic practices such as 'HARKing' (hypothesising after the results are known) and '*p*-hacking', increasing the likelihood of obtaining false positives that are difficult to replicate.

In light of this, I present my analysis in two parts: a confirmatory phase and an exploratory phase. The confirmatory analysis uses one regression model to directly test the extent to which the experimental condition affects listener perceptions of the FLEECE-GOOSE continuum, in order to answer the main research question (see section 3). The exploratory analysis presents additional models that probe other effects on the perception of GOOSE-fronting, such as interactions between experimental condition and other social variables. It is hoped that this approach makes for a transparent and clear study of sociolinguistic speech perception that does not fall into the trap of promoting false positives or conflating two distinct phases of research. Both the confirmatory and exploratory analyses are conducted using generalised linear mixed-effects logistic regression modelling, fitted to the data with the lme4 and lmerTest packages in R (Bates, Mächler, Bolker & Walker 2015; Kuznetsova, Brockhoff & Christensen 2017; R Core Team 2017). The continuous variables were standardised (i.e., scaled and centred) using z-scores. The models included a random intercept by participant and a random slope for continuum step by participant. The threshold for significance is set to 95 percent ($p < 0.05$).[1]

The regression model used for the confirmatory analysis includes CONTINUUM STEP and EXPERIMENTAL CONDITION as independent variables. The former is included as a control variable in order to test whether listeners responded differently to more *bee*-like and more *boo*-like stimuli, which is necessary in order for the vowel continuum to be representative of fronted and backed variants of GOOSE. CONDITION is the main variable under investigation, and is included in the model as a fixed effect with two levels (WOMAN'S FACE and MAN'S FACE, with the former as the reference level). This is the only model that is needed and is appropriate for the confirmatory analysis, as it tests directly for the influence of visual speaker gender on perception of the FLEECE-GOOSE continuum.[2]

The statistical modelling for the exploratory analysis incorporates a greater range of independent variables. In addition to CONTINUUM STEP and EXPERIMENTAL CONDITION, the exploratory modelling adds three socio-demographic characteristics of the participants as predictors, namely their GENDER (woman or man, with the former as the reference level), their AGE in years (standardised), and the DISTANCE between their home town and London in miles (standardised). Two-way interactions between these three factors and EXPERIMENTAL CONDITION were also included. Fixed effects and interactions were removed according to backwards selection if they were not significant, which is appropriate for exploratory analysis. Backwards selection can sometimes produce anti-conservative *p*-values, but all the significant effects in the final model reported in the Results section were also significant in the full model. The predictors that were removed in order to arrive at the final model were participant AGE, DISTANCE from London, and the respective two-way interactions between these two variables and EXPERIMENTAL CONDITION.[3]

The distance between each participant's home town and London accounts for the listeners' region of origin. GOOSE-fronting is often included in the set of features supposedly emanating from Greater London and spreading throughout the South East and beyond, collectively referred to as 'youth norms' (Williams & Kerswill 1999) or 'Estuary English' (Altendorf 2017), so it is important to consider regional variation in the analysis. While the sound change is widespread throughout the UK, it appears to be more advanced in SSBE and related varieties than in northern accents (Ferragne & Pellegrino 2010; Williams & Escudero 2014). In some areas of the far North of England, such as Newcastle upon Tyne and West Cumbria, GOOSE-fronting is absent or at least much less common (Beal, Burbano-Elizondo & Llamas 2012; Jansen 2018). This motivates the use of distance from London as an operationalisation of listener region. Some areas further away from London have historically had fronted GOOSE, such as the West Country, Manchester, and Scotland (Wells 1982), which creates a problem for this way of measuring region. Alternatives, such as coding listener origin into discrete regional categories (e.g., North, South, Midlands, Other),[4] were tested but were found to be less balanced across the sample and do not meaningfully reflect the regional distribution of GOOSE-fronting in the UK any better than distance from London. Interactions between these social variables and EXPERIMENTAL CONDITION are motivated by previous work showing that salience and social meaning (and thus, priming effects) may vary between different groups in a community (Levon & Fox 2014; Juskan 2016; Llamas, Watt & MacFarlane 2016; Schleef 2017).

5. Results

5.1. Confirmatory results

The results in this section address the main research question, which is whether the gender of the face in the visual prime affects listeners' categorisation of the FLEECE-GOOSE continuum. Table 3 shows the fixed effects of a regression model that tests for the effects of CONDITION and CONTINUUM STEP, as explained in section 4.2.

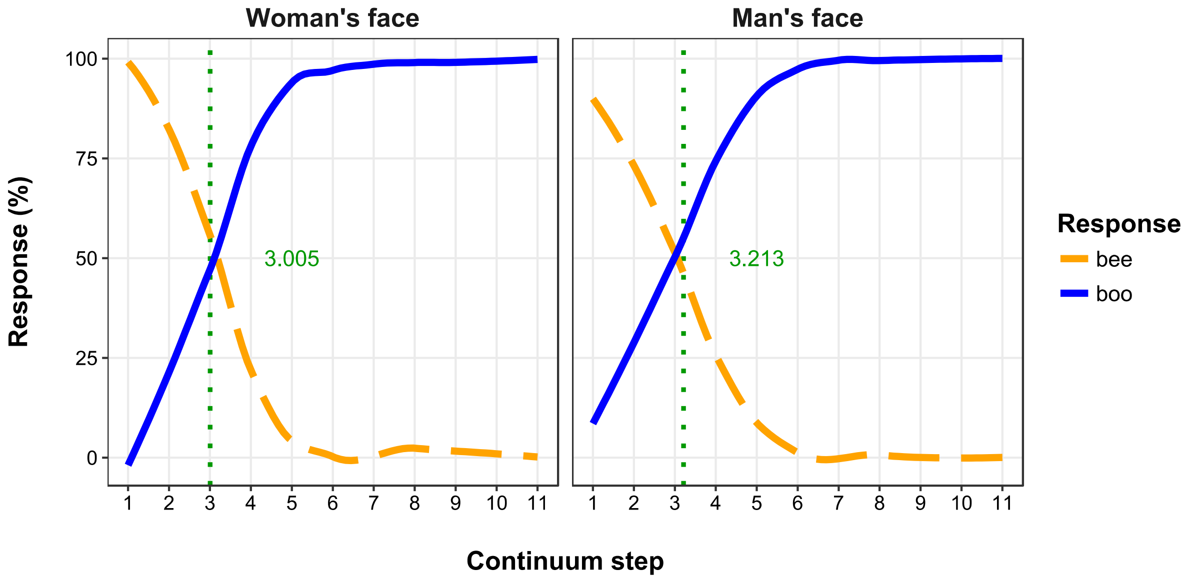
TABLE 3

Coefficients of fixed effects from a confirmatory generalised linear mixed-effects model for FLEECE-GOOSE experiment responses

	β	*SE*	*z*	*p*
(Intercept)	5.70	0.64	8.92	<0.001
Continuum step	6.10	0.57	10.70	<0.001
Condition = man's face	−0.33	0.51	−0.65	0.517

The model shows a highly significant effect of CONTINUUM STEP ($\beta = 6.10$, $p < 0.001$). This is predictable, as it is expected that participants will categorise stimuli containing highly FLEECE-like vowels as *bee* and those with very GOOSE-like tokens as *boo*. This is shown in Figure 1. In all figures used in this article, stimulus 1 represents the most fronted token (most FLEECE-like), while stimulus 11 represents the most backed token (most GOOSE-like).

Figure 1: *Mean* FLEECE-GOOSE *continuum responses by experimental condition*



Note: The vertical dotted lines and accompanying numbers in each cell represent the points at which the rates of bee and boo responses are equal.

The asymmetrical curves in Figure 1 reflect the fact that the GOOSE end-point (stimulus 11) is already somewhat fronted and that the FLEECE end-point (stimulus 1) is a little lower and backer than how this vowel is normally produced in SSBE. Despite this, the near-categorical selection of *bee* for stimulus 1 and *boo* for stimulus 11 shows that the stimuli differentiated effectively between the two vowels at the end-points of the continuum and that participants were able to perceive the differences in vowel frontness. More importantly, however, the effect of CONDITION does not reach significance ($\beta = -0.33$, $p = 0.517$) in the model and shows a minimal difference in Figure 1. In terms of the research question posed in section 3, then, it does not seem that the visual gender of the speaker affects perception of GOOSE-fronting in SSBE.

The implications of this finding will be explained more fully in section 6, but for now, we ought to conclude that there is no difference in perception of GOOSE-fronting in SSBE between listeners primed with a man's face versus those primed with a woman's face. The exploratory analysis in section 5.2 will examine whether a positive priming effect occurs within a limited set of the sample, for listeners of particular social backgrounds. This is for the purpose of further exploring other patterns in the data, which may then form the basis of more robust hypothesis testing in future studies.

5.2. Exploratory results

As discussed in section 4.2, the exploratory model investigates the extent to which a potential effect of CONDITION may be mediated by other social factors by including appropriate interaction terms in the model. Table 4 shows the output for the fixed effects of the regression model. The bigger the β estimate, the greater the likelihood of a GOOSE response.

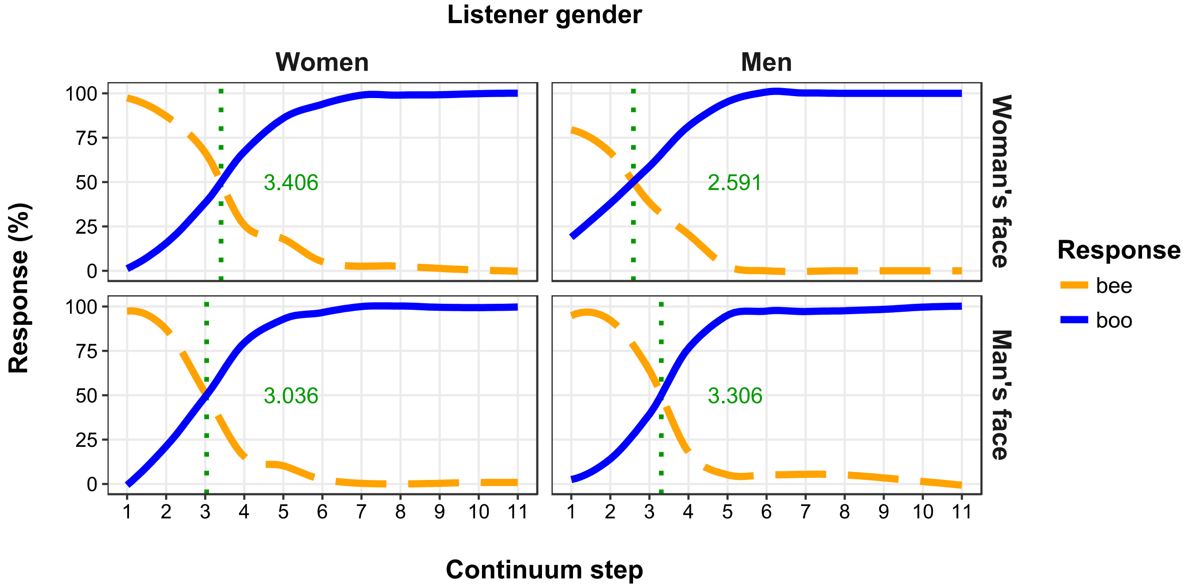
TABLE 4

Coefficients of fixed effects from an exploratory generalised linear mixed-effects model for FLEECE-GOOSE experiment responses

	β	SE	z	p
(Intercept)	4.90	0.68	7.19	<0.001
Continuum step	6.09	0.57	10.72	<0.001
Condition = man's face	0.59	0.67	0.86	0.376
Gender = man	1.63	0.68	2.38	0.017
Condition = man's face × Gender = man	−1.88	0.96	−1.96	0.050

As before, CONTINUUM STEP is significant (β = 6.09, p < .001), showing that listeners were more likely to label a stimulus as *boo* if it had a lower F2. Moreover, there is a significant interaction between GENDER and CONDITION (β = −1.88, p = .050). Men were significantly more likely to perceive a token as *boo* rather than *bee* in the woman's face condition relative to the participants in the other three cells. This is illustrated in Figure 2.

Figure 2: *Mean* FLEECE-GOOSE *continuum responses by listener gender and experimental condition*



Note: The vertical dotted lines and accompanying numbers in each cell represent the points at which the rates of bee *and* boo *responses are equal.*

The graph shows that the results for the women in both conditions and the men primed with a woman's face are all broadly comparable. However, the men in the woman's face condition perceived tokens towards the left-hand side of the FLEECE-GOOSE continuum (that is, the most FLEECE-like tokens) as GOOSE considerably more frequently than the other listeners. This even applies to stimulus 1, which was perceived as GOOSE 20.51 percent of the time by men in the woman's face condition, yet almost never occurred for the listeners in the other three cells. This effect is not restricted to only one or two outlying participants, as six men out of the thirteen in this condition labelled stimulus 1 as *boo* at least once. The result of this is that the overall crossover point at which both stimuli reach 50 percent is further to the left for participants in this condition (2.591 versus at least 3.036 for the other three cells), suggesting that their perceptual boundary between GOOSE and FLEECE may be further forward in the vowel space.

Examining the stimulus boundary points for each individual—that is, the points at which each participant's rates of perception reach 50 percent for both responses—can also be used to assess listeners' differences in perception (Johnson, Strand & D'Imperio 1999; Drager 2011). Boundary points for the data were estimated by obtaining coefficients for each speaker based on their deviation from zero in a generalised logistic regression model. The results were then fitted with a linear regression model using the lm() function in R, with the same fixed effects and model selection process as the previous model, presented in Table 5. In this model, the smaller the β estimate, the fronter the boundary between GOOSE and FLEECE (i.e., the closer it is to stimulus 1, the most FLEECE-like token).

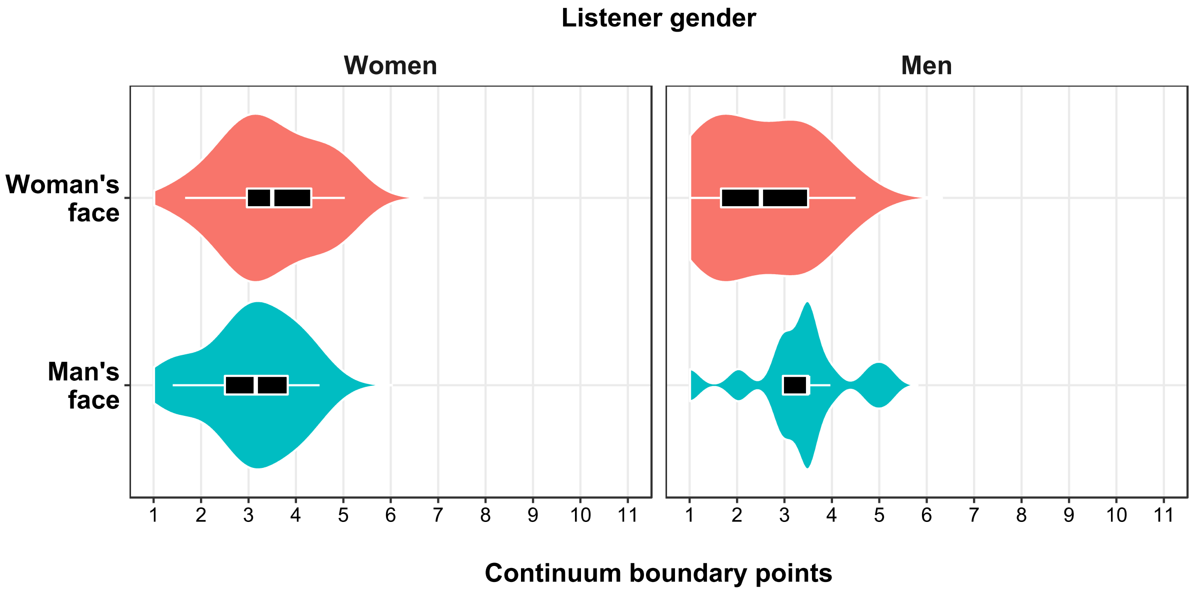
TABLE 5

Coefficients of fixed effects from a linear regression model for FLEECE-GOOSE *stimulus boundary points*

	β	*SE*	*t*	*p*
(Intercept)	3.49	0.29	12.08	<0.001
Condition = man’s face	−0.40	0.41	−0.97	0.356
Gender = man’s	−0.97	0.41	−2.37	0.022
Condition = man’s face × Gender = man	1.21	0.58	2.10	0.041

The boundary point results mirror those for the total responses to the experimental stimuli. The significant interaction between CONDITION and GENDER (β = 1.21, p = .041) is illustrated in the violin plots in Figure 3.

FIGURE 3: *Distribution of boundary points along the* FLEECE-GOOSE *continuum by listener gender and experimental condition*



The violin plots show that there is a much larger density of highly front boundaries (close to stimulus 2) for men in the woman's face condition, reflecting the above results from the responses to the stimuli and the interaction in the statistical model. In the man's face condition, men are highly clustered around the mean with a small number of outliers, while men exposed to a woman's face are more widely dispersed. In contrast, the difference between the two conditions for women is much less dramatic.

6. Discussion

The confirmatory results suggest that visual speaker gender does not affect listener perceptions of GOOSE-fronting in SSBE overall. However, the exploratory analysis indicates that it may be mediated by listener gender, with men showing the expected priming effect of categorising more tokens as GOOSE when exposed to a woman's face. Despite this, the findings ought to be considered with caution, as the significant interaction between CONDITION and LISTENER GENDER only just falls below the (admittedly arbitrary) cut-off for significance at $p = 0.04995$. This does not necessarily mean that the result is a false positive, but that with a more statistically powerful study, the effect may not reach significance. Statistical power in experimental studies can be enhanced by raising the number of participants and/or stimulus items. Sociolinguistic priming studies often suffer from a lack of statistical power, as they usually have small numbers of participants (typically around 20-50 per experiment) and items (typically around one to five). They also tend to use between-subjects designs, which are more efficient for data collection than within-subjects designs but are less statistically powerful and are risky in the sense that null results are difficult to interpret. Scholars in psychology and in phonetics have discussed the issue of power in experimental work and have offered suggestions and tools for ensuring at the design stage that a study is sufficiently statistically powerful to reduce the likelihood of encountering false positives and false negatives (Westfall, Kenny & Judd 2014; Kirby & Sonderegger 2018). The participant sample in the present study is larger than in similar work (e.g., Hay, Nolan & Drager 2006; Harrington, Kleber & Reubold 2008; Hay & Drager 2010; Hay, Drager & Warren 2010; Drager 2011; Lawrence 2015b), but the brittle nature of the results suggests that statistical power must be improved further and taken increasingly seriously in future work in this area. Walker, Szakay, and Cox (2019) provide an excellent recent example of a sociolinguistic priming study that replicates the design of a previous investigation (Hay & Drager 2010) while offering improved statistical power.

Notwithstanding these issues, the lack of evidence for a uniform priming effect across all listeners could be caused by the low social salience of GOOSE-fronting in SSBE. As discussed in section 3, some studies of this sound change have found that any sociolinguistic variation in GOOSE production passes by unnoticed in perception (Fridland 2008; Haddican, Foulkes, Hughes & Richards 2013). This stands in contrast to speech perception literature on more strongly stereotyped variables that show a perceptual link between social information and phonetic variation (e.g., Niedzielski 1999; Hay, Nolan & Drager 2006; Hay, Warren & Drager). The studies that do find a perceptual link between GOOSE-fronting and socio-indexical features have focused on regionally specific meanings and local stereotypical personae among listeners from the region in question (Lawrence 2017; Villarreal 2018), which may not apply to SSBE speech or to a geographically diverse listener sample. This would seem to support Juskan's (2016) view that sound changes that are not salient for listeners are not consistently affected by social priming, though the question remains as to why the men primed with a woman's face in the present study seemed to respond differently from the other participants.

Similar interactions between condition and gender have been found in other research (Hay, Nolan & Drager 2006; Hay & Drager 2010; Drager 2011). Drager (2011) posits that women may be more aware of the relationship between phonetic variation and social information than men, and would hence potentially index more characteristics with each stored exemplar and attach more weight to social information. She also claims that men and women may take more notice of different aspects of the acoustics of the speech signal—namely, that men pay more attention to vowel backness, while women base their vowel categorisation on vowel height. Neither of these explanations, however, has been properly tested yet in sociophonetics.

Hay and Drager additionally suggest that cultural phenomena specific to New Zealand society, such as same-sex school attendance (Drager 2011) and local masculine sporting rivalries (Hay, Nolan & Drager 2006; Hay & Drager 2010), may affect why men and women responded differently to social priming in their research. This points to broader ideas in sociolinguistics regarding how the socio-indexical meanings associated with linguistic variants by different individuals are influenced by a wide range of socio-cultural factors, including levels of exposure, identity construction, and attitudes and ideologies affecting people's views of language, people, and lifestyles (Eckert 2008; Levon & Fox 2014). These factors are argued to have an influence on the strength of exemplar activation in exemplar models (Foulkes & Docherty 2006; Drager 2015) as well as play a role in determining cognitive and social salience (Rácz 2013; Levon & Fox 2014).

In light of this discussion, it is possible that the indexicality and salience of GOOSE-fronting may potentially vary among the listeners in the study, with men having a particular sensitivity towards the female-led nature of the change. However, this interpretation does not fit well with the absence of an effect of listener region of origin. If salience and social meanings vary for different people, we might expect listeners from certain parts of the country to be influenced more by speaker gender than others, especially if GOOSE-fronting is more common in southern accents of England compared to northern ones. It is worth noting at this point that most previous work that has found positive priming effects has involved listeners rating speakers from their own or nearby regions. The sample was made up of people from a wide variety of regional backgrounds, so it is possible that many of the participants' lack of close familiarity with the speaker's accent (SSBE) meant that they did not have sufficient stored exemplars to make the connection between GOOSE-fronting and female speakers. The men in the woman's face condition whose responses were different to those in the other cells did not differ substantially in terms of background from the other listeners, so this result is unlikely to be caused by regional effects. Future work on priming and speech perception may find it useful to design studies in which the participants share the same geographical background in order to avoid these difficulties.

Two more caveats are also worth noting. One is that GOOSE-fronting may have reached the point of such ubiquity that it is no longer significantly stratified by gender in southern English accents (e.g., Alderton 2019). Further work on vocalic variation and change in British varieties is required to test the extent of this across different regions and social groups in the UK. The other is that physiological differences between men and women mean that listeners may tolerate a higher F2 while primed with a woman's face simply because they expect higher formant values for women in general. Differentiating between physiological and social effects is arguably impossible in this kind of perception experiment but ought to be borne in mind in future perceptual work on vowel changes led by a particular gender.

7. Conclusion

The main aim for this study is to answer the question of whether the gender of the face of a perceived speaker has an effect on the perception of GOOSE-fronting in SSBE, as a low-salience sound change. The results of the confirmatory analysis show that the gender of the face does not significantly affect where the participant sample as a whole place the boundary between the end-points of the continua, unlike in other sociolinguistic priming literature (e.g., Johnson, Strand & D'Imperio 1999; Niedzielski 1999; Strand 1999; Hay, Nolan & Drager 2006; Hay & Drager 2010). However, the exploratory analysis shows that men primed with a woman's face accepted more fronted tokens as GOOSE.

These findings echo some recent speech perception experiments that do not report positive and predictable priming effects (Squires 2013; Lawrence 2015b; Juskan 2016; Walker, Szakay & Cox 2019). It has been argued that this could be partly as a result of differences in salience between variables, with linguistic features that do not index any stereotypical social associations failing to produce priming effects (Juskan 2016). It is also possible that sociolinguistic priming effects are simply not as robust as indicated in earlier work. This raises important questions regarding the role of social information in speech perception and how to test this using laboratory methods. Other experimental paradigms such as the Implicit Association Test (Campbell-Kibler 2012) may offer useful alternative routes to study these phenomena. It is also likely that the mixed picture obtained from sociolinguistic priming work to date derives from many studies being statistically under-powered. Future work in this area ought to use a sufficiently large number of participants and items, ideally as part of a within-subjects design, in order to maximise statistical power and produce reliable and replicable findings to further our understanding of the perception of sociolinguistic meaning.

Notes

1. The data and code used for this article are available online as Supplementary Materials.

2. lme4 model formula for the confirmatory model: glmer(response ~ (1 + continuum step | participant) + continuum step + condition, family = binomial).

3. lme4 model formula for the final exploratory model: glmer(response ~ (1 + continuum step | participant) + continuum step + condition + gender + condition : gender, family = binomial).

4. The ‘Other’ category encompassed one participant each from Scotland and Northern Ireland, as well as two participants who had lived in various locations throughout their childhoods. For these latter two participants, distance from London was calculated based on their main caregivers’ home towns.

Acknowledgments

I would like to thank Alexandra D'Arcy, Katie Drager, Peter Grund, Sam Kirkham, Claire Nance, and an anonymous reviewer for their extremely valuable feedback on previous versions of this article. I am grateful to the Faculty of Arts and Social Sciences at Lancaster University for their award of a postgraduate scholarship to enable me to carry out this research. I also thank Daniel Lawrence for supplying the Praat vowel synthesis script used in the study, and audiences at the Southern Englishes Workshop 2016 (London), BAAP 2016 (Lancaster) and ExAPP 2016 (Vienna) for providing comments on earlier stages of this work.

Author Biography

Roy Alderton is a University Teacher in Linguistics at the University of Sheffield, UK. He recently completed his doctoral thesis at Lancaster University, UK, entitled 'Salience and social meaning in speech production and perception'. His research interests are sociophonetics, language variation and change, and the relationship between speech production and speech perception.

References

Alderton, Roy. 2019. Salience and social meaning in speech production and perception. Lancaster, UK: Lancaster University PhD thesis.

Alku, Paavo, Hannu Tiitinen & Risto Näätänen. 1999. A method for generating natural-sounding speech stimuli for cognitive brain research. *Clinical Neurophysiology* 110(8). 1329-1333.

Altendorf, Ulrike. 2017. Estuary English. In Alexander Bergs & Laurel Brinton (eds.), *The history of English*, vol. 5, *Varieties of English*, 169-186. Berlin: Mouton de Gruyter.

Assmann, Peter, Sophia Dembling & Terrance Nearey. 2006. Effects of frequency shifts on perceived naturalness and gender information in speech. *Proceedings of the International Conference of Spoken Language Processing 2006*, 889-892.

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. Harald, Shravan Vasishth, Reinhold Kliegl & Douglas Bates. 2017. The cave of shadows: Addressing the human factor with generalised additive mixed models. *Journal of Memory and Language* 94. 206-234.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1-48.

Beal, Joan, Lourdes Burbano-Elizondo & Carmen Llamas. 2012. *Urban North-Eastern English: From Tyneside to Teesside*. Edinburgh: Edinburgh University Press.

Bjelaković, Andrej. 2017. The vowels of contemporary RP: Vowel formant measurements for BBC newsreaders. *English Language and Linguistics* 21(3). 501-532.

Boersma, Paul & David Weenink. 2015. Praat: Doing phonetics by computer [Computer program]. Version 5.4.09, retrieved June 2015 from http://www.praat.org/.

Campbell-Kibler, Kathryn. 2012. The Implicit Association Test and sociolinguistic meaning. *Lingua* 122(7). 753-763.

Chang, Yung-Hsiang Shawn. 2015. Use of social information in the perception of Mandarin alveolar-retroflex contrast. *Proceedings of the International Congress of Phonetic Sciences XVIII*. London: International Phonetic Association.

Chládková, Kateřina & Silke Hamann. 2011. High vowels in Southern British English: /u/-fronting does not result in merger. *Proceedings of the International Congress of Phonetic Sciences XVII*, 476-479. London: International Phonetic Association.

Chládková, Kateřina, Silke Hamann, Daniel Williams & Sam Hellmuth. 2017. F2 slope as a perceptual cue for the front-back contrast in Standard Southern British English. *Language and Speech* 60(3). 377-398.

Deterding, David. 1997. The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association* 27(1-2). 47-55.

Docherty, Gerard, Paul Foulkes, Simon Gonzalez & Nathaniel Mitchell. 2018. Missed connections at the junction of sociolinguistics and speech processing. *Topics in Cognitive Science* 10(4). 759-774.

D’Onofrio, Annette. 2018. Personae and phonetic detail in sociolinguistic signs. *Language in Society* 47(4). 513-539.

Drager, Katie. 2011. Speaker age and vowel perception. *Language and Speech* 54(1). 99-121.

Drager, Katie. 2015. *Linguistic variation, identity construction and cognition*. Berlin: Language Science Press.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453-476.

Ferragne, Emmanuel & François Pellegrino. 2010. Formant frequencies of vowels in 13 accents of the British Isles. *Journal of the International Phonetic Association* 40(1). 1-34.

Flynn, Nicolas. 2012. A sociophonetic study of Nottingham speakers. York, UK: University of York PhD thesis.

Foulkes, Paul & Gerard J. Docherty. 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34(4). 409-438.

Fridland, Valerie. 2008. Regional differences in perceiving vowel tokens on Southerness, education, and pleasantness ratings. *Language Variation and Change* 20(1). 67-83.

Fridland, Valerie, Kathryn Bartlett & Roger Kreuz. 2004. Do you hear what I hear? Experimental measurement of the perceptual salience of acoustically manipulated vowel variants by Southern speakers in Memphis, TN. *Language Variation and Change* 16(1). 1-16.

Fridland, Valerie, Kathryn Bartlett & Roger Kreuz. 2005. Making sense of variation: Pleasantness and education ratings of regional vowel variants. *American Speech* 80(4). 366-387.

Goldinger, Stephen D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology* 22(5). 1166-1183.

Haddican, Bill, Paul Foulkes, Vincent Hughes & Hazel Richards. 2013. Interaction of social and linguistic constraints on two vowel changes in northern England. *Language Variation and Change* 25(3). 371-403.

Harrington, Jonathan, Felicitas Kleber & Ulrich Reubold. 2007. /u/-fronting in RP: A link between sound change and diminished perceptual compensation for coarticulation? *Proceedings of the International Congress of Phonetic Sciences XVI*, 1473-1476. London: International Phonetic Association.

Harrington, Jonathan, Felicitas Kleber & Ulrich Reubold. 2008. Compensation for coarticulation, /u/-fronting, and sound change in Standard Southern British: An acoustic and perceptual study. *The Journal of the Acoustical Society of America* 123(5). 2825-2835.

Harrington, Jonathan, Felicitas Kleber & Ulrich Reubold. 2011. The contributions of the lips and the tongue to the diachronic fronting of high back vowels in Standard Southern British English. *Journal of the International Phonetic Association* 41(2). 137-156.

Hawkins, Sarah & Jonathan Midgley. 2005. Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association* 35(2). 183-199.

Hay, Jennifer & Katie Drager. 2010. Stuffed toys and speech perception. *Linguistics* 48(4). 865-892.

Hay, Jennifer, Katie Drager & Paul Warren. 2010. Short-term exposure to one dialect affects processing of another. *Language and Speech* 53(4). 447-471.

Hay, Jennifer, Aaron Nolan & Katie Drager. 2006. From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review* 23(3). 351-379.

Hay, Jennifer, Paul Warren & Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34(4). 458-484.

Henton, Caroline. 1989. Fact and fiction in the description of female and male pitch. *Language & Communication* 9(4). 299-311.

Hillenbrand, James. 2005. The role of fundamental frequency and formants in the perception of speaker sex. *The Journal of the Acoustical Society of America* 118(3). 1932-1933.

Holmes-Elliott, Sophie. 2015. London calling: Assessing the spread of metropolitan features in the southeast. Glasgow, UK: University of Glasgow PhD thesis.

Jaeger, T. Florian & Kodi Weatherholtz. 2016. What the heck is salience? How predictive language processing contributes to sociolinguistic perception. *Frontiers in Psychology* 7(1115). URL: https://doi.org/10.3389/fpsyg.2016.01115.

Jansen, Sandra. 2018. Peripheral communities and innovation: Changes in the GOOSE vowel in a West Cumbrian town. In Natalie Braber & Sandra Jansen (eds.), *Sociolinguistics in England*, 297-331. London: Palgrave Macmillan.

Johnson, Keith. 2006. Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34(4). 485-499.

Johnson, Keith, Elizabeth A. Strand & Mariapaola D'Imperio. 1999. Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics* 27(4). 359-384.

Juskan, Marten. 2016. Production and perception of local variants in Liverpool English: Change, salience, exemplar priming. Freiburg, Germany: Albert Ludwig's University PhD thesis.

Kerswill, Paul & Ann Williams. 2002. "Salience" as an explanatory factor in language change: Evidence from dialect levelling in urban England. In Mari Jones & Edith Esch (eds.), *Language change: The interplay of internal, external and extra-linguistic factors*, 81-110. Berlin: Mouton de Gruyter.

Kirby, James & Morgan Sonderegger. 2018. Mixed-effects design analysis for experimental phonetics. *Journal of Phonetics* 70. 70-85.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1-26.

Labov, William. 2001. *Principles of linguistic change*, vol. 2, *Social factors*. Oxford: Blackwell.

Lawrence, Daniel. 2015a. A brief introduction to vowel resynthesis for speech perception research. Unpublished manuscript.

Lawrence, Daniel. 2015b. Limited evidence for social priming in the perception of the BATH and STRUT vowels. *Proceedings of the International Congress of Phonetic Sciences XVIII*. London: International Phonetic Association.

Lawrence, Daniel. 2017. Sound change and social meaning: The perception and production of phonetic change in York. Edinburgh, UK: University of Edinburgh PhD thesis.

Levon, Erez & Sue Fox. 2014. Social salience and the sociolinguistic monitor: A case study of ING and TH-fronting in Britain. *Journal of English Linguistics* 42(3). 185-217.

Llamas, Carmen, Dominic Watt & Andrew MacFarlane. 2016. Estimating the relative sociolinguistic salience of segmental variables in a dialect boundary zone. *Frontiers in Psychology* 7(1163). URL: https://doi.org/10.3389/fpsyg.2016.01163.

Love, Robbie, Claire Dembry, Andrew Hardie, Vaclav Brezina & Tony McEnery. 2017. The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics* 22(3). 319-344.

Munson, Benjamin. 2011. The influence of actual and imputed talker gender on fricative perception, revisited. *The Journal of the Acoustical Society of America* 130(5). 2631-2634.

Munson, Benjamin, Kayleigh Ryherd & Sara Kemper. 2017. Implicit and explicit gender priming in English lingual sibilant fricative perception. *Linguistics* 55(5). 1073-1107.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language & Social Psychology* 18(1). 62-85.

Nosofsky, Robert M. 1986. Attention, similarity, and identification-categorization relationship. *Journal of Experimental Psychology* 115(1). 39-57.

Peirce, Jonathan. 2007. PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods* 162(1-2). 8-13.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan L. Bybee & Paul J. Hopper (eds.), *Frequency and the emergence of linguistic structure*, 137-158. Amsterdam: John Benjamins.

Przedlacka, Joanna. 2002. *Estuary English? A sociophonetic study of teenage speech in the Home Counties*. Frankfurt am Main: Peter Lang.

R Core Team. 2017. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, version 3.4.0, retrieved May 2017 from http://www.r-project.org/.

Rácz, Péter. 2013. *Salience in sociolinguistics: A quantitative approach*. Berlin: Mouton de Gruyter.

Roettger, Timo B. 2019. Researcher degrees of freedom in phonetic research. *Laboratory Phonology* 10(1). 1-27.

Roettger, Timo B., Bodo Winter & R. Harald Baayen. 2019. Emergent data analysis in phonetic sciences: Towards pluralism and reproducibility. *Journal of Phonetics* 73. 1-7.

Schleef, Erik. 2017. Social meanings across listener groups: When do social factors matter? *Journal of English Linguistics* 45(1). 28-59.

Simpson, Adrian P. 2009. Phonetic differences between male and female speech. *Language and Linguistics Compass* 3(2). 621-640.

Skuk, Verena G. & Stefan R. Schweinberger. 2014. Influences of fundamental frequency, formant frequencies, aperiodicity, and spectrum level on the perception of voice gender. *Journal of Speech, Language, and Hearing Research* 57(1). 285-296.

Sóskuthy, Márton, Paul Foulkes, Vincent Hughes & Bill Haddican. 2018. Changing words and sounds: The roles of different cognitive units in sound change. *Topics in Cognitive Science* 10(4). 787-802.

Squires, Lauren. 2013. It don't go both ways: Limited bidirectionality in sociolinguistic perception. *Journal of Sociolinguistics* 17(2). 200-237.

Staum Casasanto, Laura. 2008. Does social information influence sentence processing? *Proceedings of the Annual Meeting of the Cognitive Science Society* 28. 799-804.

Strand, Elizabeth A. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language & Social Psychology* 18(1). 86-100.

Strand, Elizabeth A. & Keith Johnson. 1996. Gradient and visual speaker normalization in the perception of fricatives. In Dafydd Gibbon (ed.), *Natural language processing and speech technology*, 14-26. Berlin: Mouton de Gruyter.

Tollfree, Laura. 1999. South East London English: Discrete versus continuous modelling of consonantal reduction. In Paul Foulkes & Gerard J. Docherty (eds.), *Urban voices: Accent studies in the British Isles*, 163-184. London: Arnold.

Torbert, Benjamin. 2004. Southern vowels and the construction of social salience. Durham, NC: Duke University PhD thesis.

Traunmüller, Hartmut & Anders Eriksson. 1995. The frequency range of the voice fundamental in the speech of male and female adults. Unpublished manuscript.

Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.

University of Stirling. 2013. Psychological image collection at Stirling. http://pics.psych.stir.ac.uk/ (24 June, 2015).

Villarreal, Dan. 2018. The construction of social meaning: A matched-guise investigation of the California Vowel Shift. *Journal of English Linguistics* 46(1). 52-78.

Walker, Michael, Anita Szakay & Felicity Cox. 2019. Can kiwis and koalas as cultural primes induce perceptual bias in Australian English speaking listeners? *Laboratory Phonology* 10(1). 1-29.

Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Westfall, Jacob, David A. Kenny & Charles M. Judd. 2014. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology General* 143(5). 2020-2045.

Williams, Ann & Paul Kerswill. 1999. Dialect levelling: Change and continuity in Milton Keynes, Reading and Hull. In Paul Foulkes & Gerard J. Docherty (eds.), *Urban voices: Accent studies in the British Isles*, 141-162. London: Arnold.

Williams, Daniel & Paola Escudero. 2014. A cross-dialectal acoustic comparison of vowels in Northern and Southern British English. *The Journal of the Acoustical Society of America* 136(5). 2751-2761.

Winn, Matthew, Ariane Rhone, Monita Chatterjee & William Idsardi. 2013. The use of auditory and visual context in speech perception by listeners with normal hearing and listeners with cochlear implants. *Frontiers in Psychology* 4(824). URL: https://doi.org/10.3389/fpsyg.2013.00824

APPENDIX 1

Regional distribution of participants across England by gender and condition. Participants in the 'Other' category were from Scotland, Northern Ireland, or had lived in various places during their childhood.

	Woman's face condition		**Man's face condition**		
	Women	*Men*	*Women*	*Men*	*Total*
North	8	9	4	7	28
South	0	3	4	4	11
Midlands	3	1	4	1	9
Other	2	0	1	1	4