LIMITED EVIDENCE FOR SOCIAL PRIMING IN THE PERCEPTION OF THE BATH AND STRUT VOWELS

Daniel Lawrence

The University of Edinburgh
dlawrenc@staffmail.ed.ac.uk

ABSTRACT

It has been claimed that priming listeners with information about a speaker's regional background can lead to predictable biases in the categorization of their speech [7], reflecting the influence of socially-indexed episodic representations on speech perception [4, 5, 6]. The present study tested for this effect among a group of 40 speakers of Southern Standard British English, attempting to replicate Niedzielski's [7] vowel-matching experiment in a new context. Listeners heard a speaker from Sheffield reading a set of sentences containing BATH and STRUT words utterance-finally. After each sentence, they listened to a continuum of re-synthesized vowel tokens and matched them to the speaker's natural realizations. Half of the participants were told that the speaker was from 'Sheffield, Northern England', and the other were told that the speaker was from 'London, Southern England'. In contrast to previous replications of this experiment, there was no evidence that the regional labels affected listeners' responses.

Keywords: sociophonetics; speech perception

1. INTRODUCTION

Niedzielski [7] provides evidence that a listeners' impression of the social background of a speaker can influence the way in which that speaker's speech is categorized. In the vowel-matching experiment reported in [7], participants from Detroit listened to a series of sentences containing utterance-final tokens of words containing the MOUTH vowel. After hearing each sentence, they listened to a synthetic continuum representing six levels of raising of the vowel onset and were asked to choose the token closest to the one which they had just heard. Half of the participants were told that they were listening to a speaker from Detroit, and half were told that they were listening to a speaker from Canada. Consistent with US sociolinguistic stereotypes, participants who thought they were listening to a Canadian were more likely to identify raised tokens of MOUTH than those who thought they were listening to a Detroit speaker, even though the feature was present in the recordings played to both groups. Similar effects are reported in [5] with reference to the PIN-PEN merger in New Zealand.

Hay et al. [4, 5, 6] explain these 'social priming' effects with an exemplar-based account of speech perception. Under this approach, listeners' language knowledge is claimed to consist of highly detailed episodic representations of speech events. These representations include not only linguistic information, but also information about non-linguistic aspects of individual perceptual events, such as the social characteristics of the speaker. The perceptual effect reported in [7] can thus be explained in terms of spreading activation. The presence of information about a speaker's regional origin, for example, the label 'Canadian', causes the activation of speech exemplars which are indexed as 'Canadian', facilitating their retrieval during perception. It has been suggested that the perceptual system is extremely sensitive to such social effects – Hay et al. [4] present evidence that similar results can be obtained by priming participants with soft toys representing national stereotypes, which were present during the authors' experimental briefing.

The present study aimed to replicate the vowel-matching experiment presented in [7] in a new context, focusing on two vowels which are widely acknowledged as highly salient markers of regional identity in British English – the vowels in BATH and STRUT. These vary in production as follows:

- Southern speakers may realize the vowel in BATH as a low back vowel [ɑː], and the vowel in TRAP as higher and further forward [æ]; speakers of Northern varieties, however, often realize both BATH and TRAP as [a], with [ɑː] used only in the PALM and START lexical sets.
- Southern varieties contrast the vowel in FOOT /ʊ/ with the that in STRUT /ʌ/. Northern speakers, on the other hand, may merge /ʊ/ and /ʌ/ completely, with speakers realizing both as [ʊ], or as an intermediate [ə] - like variant.

Under the exemplar model put forward by Hay et al. [4, 5, 6], assuming sufficient exposure to a

range of Northern and Southern varieties, speakers of Standard Southern British English should posess socially-indexed exemplars of these vowels – BATH words realized with [ɑː] will be indexed as 'Southern' and those realized with [a] will be indexed as 'Northern'; STRUT words realized as [ʌ] will be indexed as 'Southern' and those realized as [ʊ] or [ə] will be indexed as 'Northern'. Since these two vowels are two of the most salient features of Northern and Southern linguistic stereotypes in England, existing accounts of social priming would predict that similar results to [7] would be easily obtained. Priming listeners with information related to Northern and Southern regions should lead to activation of the corresponding linguistic representations, facilitating their retrieval during vowel identification.

2. METHODS

The design of the present study aimed to mirror that presented by Niedzielski [7] and subsequent replications [4, 5] as closely as possible. Stimuli were a set of 40 sentences, each containing a word with a target vowel sentence-finally. These sentences contained only one instance of the target vowel and no instances of the other vowels under study. 10 sentences contained BATH words sentence-finally, and 10 contained STRUT words sentence-finally. The remaining items contained an equal number of FACE and GOAT words, which are excluded from the present analysis due to space constraints. Examples are listed below:

1. The old house was a mess; we had to buy a new **bath**.
2. I think it's important to choose a degree which you really want to **study**.

A male speaker in his 20s from Sheffield was recorded reading these sentences using the built-in microphone of a Zoom H2n portable recorder. This speaker was also recorded reading a series of "h_d" tokens containing the target vowels. A six-step F1 and F2 continuum was then synthesized for each of these, using the analysis and synthesis functions of *Praat* [1] The vowel segments were extracted from the "h_d" tokens from the onset of voicing to stop closure. These were re-sampled to 10Khz., then high-pass filtered with a cut-off frequency of 50Hz.. Formants were estimated using the Burg algorithm with 8 LPC coefficients, a window length of 25 ms., and a prediction order of 10. Inverse filtering was performed by convolving the original signal with the inverse of the LPC filter, yielding an approximation of the glottal source. Modifications were then made to the filters for each vowel, representing the desired formant values for each step of the continuum. Exciting these filters using the inverse-filtered glottal source resulted in a set of vowel tokens with formant values at the appropriate frequencies.

The formant values used in the continua aimed to mirror the approach used in [7]. Token 1 was based on the reference values provided in [3], token 4 was based on the speaker's natural production of the target vowel, and tokens 5 and 6 were two equal bark-domain steps more extreme than token 4. The intermediate tokens were interpolated at equal bark steps, calculated at four equidistant points along the vowel trajectory. Additional modifications were made to the BATH tokens to account for the fact that typical Southern realizations of this vowel tend to be longer than their Northern counterparts. The lengths of the tokens were altered at equal intervals using Praat's *overlap-add* function. Endpoint durations were calculated by measuring the duration of the vocalic portion (from onset of voicing to stop closure) in the speaker's productions of "had" and "hard". Finally, the mean intensity of all stimuli was scaled to 70 dB SPL.

Token	**BATH**		**STRUT**	
	F1	F2	F1	F2
1	728	1228	647	1497
2	750	1362	598	1468
3	773	1507	550	1440
4	797	1666	504	1413
5	820	1839	459	1386
6	844	2030	416	1359

Table 1: Mean formant values for synthesized continua.

Token	**Duration (ms.)**
1	420
2	378
3	336
4	294
5	252
6	210

Table 2: Modified durations of BATH tokens.

Participants were 40 monolingual speakers of Standard Southern British English from London and surrounding areas with no hearing difficulties and normal or corrected-to-normal vision. All were current students at the University of Edinburgh. They were paid £5 for their participation. The experiment took place over headphones in a quiet computer laboratory on a PC running Windows 7. Participants were told that they were taking part in an experiment assessing the similarity of synthesized speech to natural speech. The listeners were presented with

Figure 1: Spectrograms of the resynthesized STRUT continuum.

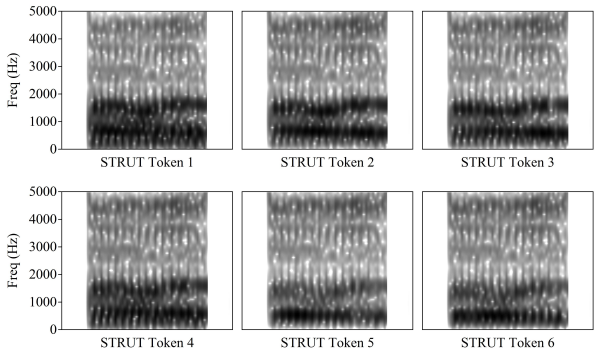
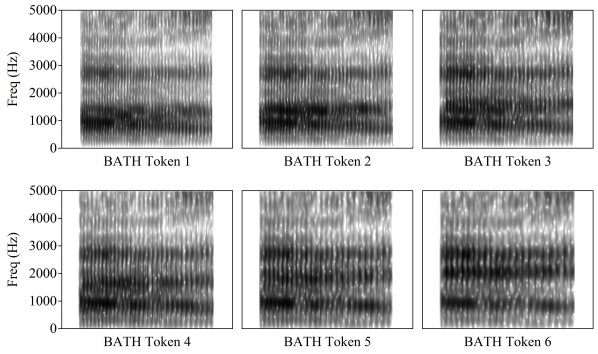


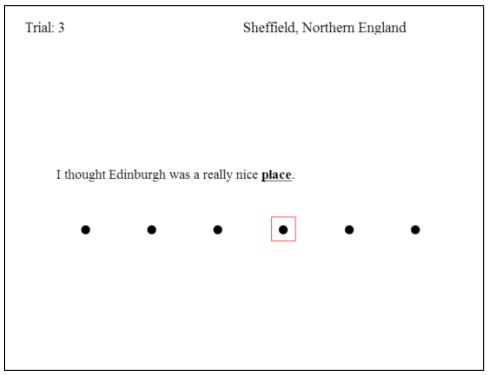
Figure 2: Spectrograms of the resynthesized BATH continuum.



the stimuli sentences in written form, with the target word underlined. They listened to the recorded sentences and were asked to pay careful attention to the underlined word. After a 1-second pause, listeners heard each of the vowel tokens from the appropriate continuum, played one-by-one at 2-second intervals. As each sound played, a black dot appeared on the screen, with dots appearing sequentially from left to right. After the complete continuum had played, participants were asked to identify the token which they felt was closest to the vowel in the previously-heard word by clicking on the corresponding dot (see Fig. 3). It should be noted that this method of presentation differs slightly from that in [4], [5] and [7], where listeners saw the sentences on a sheet of paper and were asked to indicate their response on the answer sheet.

The stimuli sentences were presented in pseudo-random order, with no two instances of the same vowel appearing consecutively. In order to discourage participants from adopting a repetitive pattern of responses, the presentation order of the continua was alternated with each trial of a given vowel; this represents a small divergence from the methods in [4], [5] and [7]. Before beginning the experiment

Figure 3: Experimental interface.



proper, participants took part in four practice trials where they identified non-target vowels.

Participants were split into two conditions who were given different information about the speaker's regional origin. One group saw the label 'Sheffield, Northern England' on the screen throughout the experiment; the other group saw 'London, Southern England'. Participants were also explicitly informed that the listener was from one of these areas in the experimental instructions. Following previous findings, it was predicted that differences in the distributions of token selections would be observed between the two experimental conditions. In the exemplar-theoretic explanation given by Hay et al. [4, 5, 6], the presence of different regional labels should activate listeners' episodic representations related to the concepts of 'London' or 'Sheffield', with activation spreading to linguistic representations associated with those places and facilitating their retrieval. Based on these claims, the distribution of selections from listeners who saw the label 'London, Southern England' was expected to tend toward the Southern end of the continuum. The opposite prediction was made for listeners who saw the label 'Sheffield, Northern England' – a tendency towards the selection of more Northern tokens was expected.

3. RESULTS

Figures 4 and 5 show the proportion of token selections for each vowel (represented by the name of the lexical set), across the two experimental conditions. Overall there appears to be little evidence of the effects reported in [4], [5] and [7]. In the case of BATH, the distribution of responses peaks around the fifth step of the continuum, approximating the mean F1 and F2 values of the vowels in the stimulus sentences (stimulus mean F1=801, stimulus mean F2=1790; token 4 F1=770, token 4 F2=1666; token 5 F1=820, token 5 F2=1839). Listeners who saw the 'Sheffield' label selected the fifth step of the BATH

continuum on 41% of trials while those in the 'London' condition selected it 36% of the time. STRUT selections peak around the fourth token, with listeners in the 'Sheffield' condition exhibiting a slight advantage in selecting the target (22% vs 28%). Chi-squared tests comparing the proportion of selections for each token across the two experimental conditions return non-significant in both cases (BATH: $\chi^2 = 3.0251$, $df = 5$, $p = 0.6961$; STRUT: $\chi^2 = 4.8594$, $df = 5$, $p = 0.4333$)

Figure 4: BATH: proportion of tokens from each continuum step selected across experimental conditions.

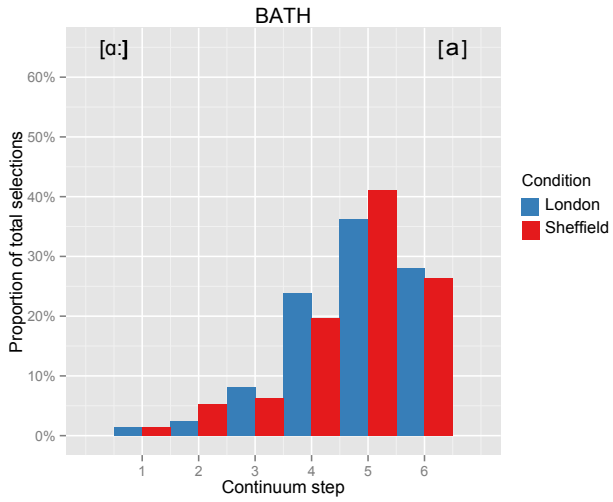
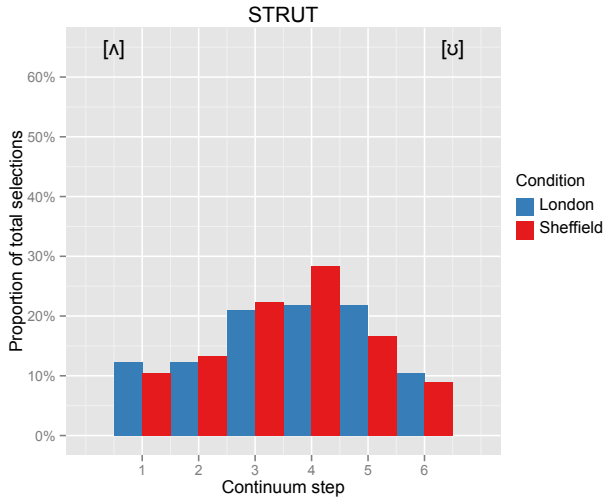


Figure 5: STRUT: proportion of tokens from each continuum step selected across experimental conditions.



4. DISCUSSION

The absence of evidence for any influence of the regional labels is somewhat surprising in light of previous findings. In Niedzielski [7], Hay et al. [5] and Hay & Drager [4], the presence of such social stimuli was claimed to influence listeners' selection of vowel tokens in line with documented production patterns. Under the exemplar-based explanation adopted by Hay et al. [4, 5, 6], it would be expected that the effect would replicate among Southern British English speakers. If the perceptual system is so sensitive to social information that even nationally-themed stuffed toys can lead to priming effects [4], then given the salience of STRUT and BATH as part of Northern/Southern sociolinguistic stereotypes in England [2, 9], it is particularly surprising that no priming effects would be found. However, this null result resonates with the findings reported by Squires [8], who tested for social priming effects in acceptability judgments of variable morphosyntactic features in American English. While speakers assigned consistent judgments of social status to standard and non-standard forms, Squires [8] reports little evidence of an effect of social priming on linguistic judgments. Together with the results of the present work, these findings suggest that the influence of social information on linguistic perception may be more limited than has been previously suggested.

5. CONCLUSION

The present study attempted to replicate the vowel-matching experiment in Niedzielski [7] in a new context, testing the influence of a regional label on Southern English listeners' ability to match resynthesized vowel tokens to previously-heard natural realizations. Under the exemplar-based account of social priming effects put forward in previous work [4, 5, 6], it was predicted that speakers who saw a label reading 'Sheffield, Northern England' would be likely to select more Northern-like tokens of the vowel than those who saw the label 'London, Southern England'. Despite previous evidence regarding the sensitivity of the perceptual system to social priming effects, the presence of the regional labels had no influence on participants' responses. In light of this null result, and the similar findings reported by Squires [8], it is suggested that future research into the role of social information in speech perception should develop testable predictions regarding which variables, contexts, and listening conditions are expected to demonstrate social priming effects, and crucially, the circumstances under which circumstances such effects would *not* be expected.

6. ACKNOWLEDGMENTS

This work was supported by the Economic and Social Research Council [grant number ES/J500136/1]

7. REFERENCES

[1] Boersma, P., Weenink, D. 2009. Praat: Doing phonetics by computer (Version 5.1.05) [Computer program]. Retrieved May 1, 2009.

[2] Chambers, J. K., Trudgill, P. 1998. *1998. Dialectology*. Cambridge: Cambridge University Press.

[3] Ferragne, E., Pellegrino, F. 2010. Formant frequencies of vowels in 13 accents of the British Isles. *Journal of the International Phonetic Association* 40(1), 1–34.

[4] Hay, J., Drager, K. 2010. Stuffed toys and speech perception. *Linguistics* 48(4), 865–892.

[5] Hay, J., Nolan, A., Drager, K. 2006. From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review* 23(3), 351–379.

[6] Hay, J., Warren, P., Drager, K. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34(4), 458–484.

[7] Niedzielski, N. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1), 62–85.

[8] Squires, L. 2013. It don't go both ways: Limited bidirectionality in sociolinguistic perception. *Journal of Sociolinguistics* 17(2), 200–237.

[9] Wells, J. C. 1982. *Accents of English: Volume 1*. Cambridge University Press.