LISTENING WITH STYLE: The effect of bidialectal style-shifting on PRICE vowel perception in Britain's Black Country

Thomas Packer-Stucki

Queen Mary University of London
t.j.packer-stucki@qmul.ac.uk

ABSTRACT

Speech perception research has found that manipulating listener attitudes towards a dialect area can lead to convergence or divergence in production style, as well as a simultaneous convergence or divergence in perception. The present study investigates whether these two effects are connected, such that the changes in production style are the cause of changes in perception. The potential link is interrogated directly through an experiment which pushes participants' stylistic production towards the extreme ends of their repertoires, followed by a perceptual test in which they match vowel tokens to a vowel line-up resynthesized along the same standard-dialect continuum. Analysis of the data shows that style does affect perception: as participants are asked to switch from a more standard to a more dialectal style, their perceptual categories shift towards the standard end of the vowel continuum, i.e., when they speak in dialect, vowels are more likely to be rated as standard-sounding, and vice-versa.

Keywords: Vowel perception, experimental sociolinguistics, bidialectalism, style-shifting

1. INTRODUCTION

1.1. Priming and speech perception

Research in experimental sociolinguistics has shown that a listener's perception of an incoming speech signal can be influenced by external factors, for example, acoustically identical sounds can be perceived differently when listeners are primed with differing biographical information about the speaker [12]. Many of these studies make use of vowel rating tasks, wherein participants hear a target vowel, and must identify the most similar sounding vowel from a synthesised line-up. This methodology was originated by Niedzielski [16], whose study had Detroit participants rate instances of the MOUTH vowel along a unraised-raised continuum, some being told they were hearing a Detroit speaker and others that they were hearing a Canadian. As this vowel is raised in both dialect areas, but stereotyped only in Canada, participants only reported hearing the raising where they expected it, in the Canada condition. While the vowel stimuli were identical across both conditions, the participants' stereotypes of the perceived speaker's dialect biased their perception in the expected direction.

1.2. Attitudinal priming

More recently, Walker et al. [21] uncovered an effect whereby the *attitudes* of New Zealand listeners towards Australia seem to bias their perception of the KIT vowel along a NZE-AuE continuum of possible realisations. Their experiment involved a listening task in which participants heard a series of target KIT vowels embedded within sentences and were asked after each to identify the most similar-sounding vowel from a resynthesized line-up, ranging from an AuE to NZE realisation. The listening task was completed twice, before and after treatment in one of three conditions, which involved the reading of positive or negative facts about Australia, or a control with random facts. Analysis of results revealed a shift in perception to more AuE-like KIT in response to the *negative* facts, while the *positive* facts had the opposite effect. Specifically, Walker et al. interpreted that participants had a defensive reaction to conditioning with positive facts about Australia; Kiwi pride leads to more Kiwi-sounding vowel perceptions. However, it is unclear by which mechanism such attitude-based shifts occur.

1.3. Listening with style

While the motivation for a link between attitude and perception has so far been difficult to identify [21], the link between attitudes and style is well theorised and attested. Research contributing to Communication Accommodation Theory (CAT) has shown that the propensity to converge towards or diverge away from a particular style or variety can be motivated by attitudes towards the speakers of that code [6], [11]. The present paper hypothesises that in the case of Walker et al. it is an attitude-motivated change in style which is responsible for the change in perception. We directly test the possibility of style-motivated shifts in perception by means of an experiment which manipulates production style rather than attitude and observes changes in vowel productions and perceptions. We did this by

following a similar experimental paradigm, but instead of manipulating attitude between each of the listening task iterations, we expressly manipulate bidialectal participants' production style by means of a reading task, pushing them from the most standard to the most dialectal end of their stylistic repertoire. What follows is a presentation of the perceptual side of this larger study involving both production and perception elements.

2. METHODOLOGY

2.1. Participants

While it is attested that all speakers exhibit *style-shifting*, we need to be certain that our participants will shift decisively between two distinct styles, if we are to argue that this shift is responsible for any perceptual effects observed. We should also be certain that such a decisive shift is naturalistic; something which our participants would usually do. To this end, we identified *bidialectal* speakers [20] as the ideal candidates for our experiment – according to Hazen [14] their speech can occupy the, "extreme ends of the style-switching continuum."

In the United Kingdom, one area which retains a unique, productive dialectal vernacular is the *Black Country*, whose dialect forms an important part of local identity, and is in common use today by the majority of the region's 1.2 million inhabitants [3], [17]. Recruitment of participants took place from the pool of presenters at local community radio station *Black Country Radio*. These speakers were chosen as they regularly access both standard and dialect codes. A total of eight participants were recruited for participation – their ages ranged from 27 to 69 (avg. 54) with five men and three women. Participants were not compensated for their time.

2.2. Procedure

The experiment took place in a recording studio at the radio station and lasted an average of 30 minutes. In the reading tasks, participants were asked to read text in either a standard or dialectal style. In the listening tasks, participants saw a series of target sentences with a PRICE-containing word highlighted (see Figure 1); after each they had to select the most similar sounding vowel from a continuum (see Figure 2). The reading and listening tasks were combined to create the following experimental procedure:

Reading Task 1: news bulletins (standard)
Listening Task 1: 4 sentences
Reading Task 2: news bulletins (standard)
Listening Task 2: 3 sentences
Reading Task 3: Poem 1 (dialect)
Listening Task 3: 4 sentences
Reading Task 4: Poem 2 (dialect)
Listening Task 4: 3 sentences

In our analysis, the main effect will therefore be the location of a particular vowel trial within the experiment, following either a standard or dialect reading task. While the reading tasks always appeared in the same order, the order of the 14 listening task sentences was randomised for each participant to mitigate any potential order effect of the stimuli.

2.3. Perception: Listening task

Following Niedzielski [16] and Walker et al. [21], a vowel perception task was developed for use in the Black Country. The PRICE vowel was selected, as differences between BCE and SSBE are notable, with [ɒɪ] given as the typical Black Country realisation and [ʌɪ] appearing in the south [7].

2.3.1. Stimuli

14 sentences containing PRICE were selected from the Bamford-Kowal-Bench list of sentences controlled for lexical frequency and average age of acquisition [4], and were recorded in a broadcast-standard recording booth by a 25-year-old female speaker from Stourbridge, in the south-west of the Black Country. When the sentences were presented to participants, they were shown orthographically on a screen with the target vowel-containing word underlined and played twice through headphones, as shown in Figure 1. Participants then heard the SSBE-BCE PRICE vowel continuum (see section 2.3.2.) and were asked to select the most similar-sounding vowel to the target. Additionally, each stimuli sentence was presented only *once* in the experiment to decrease the likelihood of participants guessing the aim – average vowel ratings between reading task styles will therefore be compared.

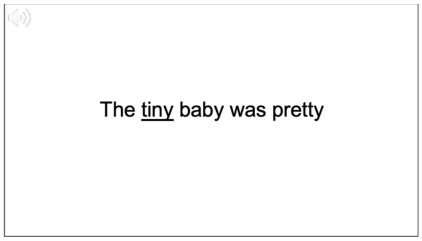

Figure 1: An example stimulus sentence as seen by participants, where the target vowel for matching with the continuum is underlined

2.3.2. Vowel continuum

In order to create the 6-step vowel continuum for PRICE, a bidialectal speaker from the Black Country with phonetics training 'performed' both the SSBE and BCE endpoints for the vowel. For the resynthesis itself, Praat software [5] was used along with a script [15] which takes two natural vowel endpoints, estimates their formants using linear prediction, and then uses inverse filtering to synthesise 6 vowels equidistant from each other in terms of formant values from each endpoint. In the resulting continuum, '1' represented the most SSBE-type vowel, and '6' the most BCE-type vowel. Following presentation of the target vowel, participants saw six boxes on a screen which represented the six vowel choices available (in order). They could then use a mouse to hover over each of the boxes, which would then cause that particular vowel to be played to them. They were asked to select the vowel which sounded most similar to the target, as shown in Figure 2.

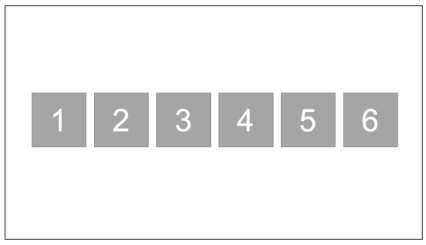

Figure 2: The vowel continuum as seen by participants, follow the presentation of each target vowel

2.4. Production: Reading task

The production element of the study was designed to manipulate the participants' style and aimed to push the participants to the most extreme ends of their stylistic repertoires. The instructions were explicit in doing so: participants were asked to speak firstly in a formal style i.e., for a 'national audience', and then secondly in a dialectal style i.e., 'broad Black Country'. This was done to ensure the biggest shift in style possible, given that in previous research, perceptual shifts tend to be slight even where they are significant [9].

2.4.1. SSBE style script (standard)

Given that the participants are radio presenters, the task used to elicit the participants' most standard style was the reading of two 'fake' news bulletins, adapted from BBC News stories selected for their financial or academic focus. Participants were asked to read these stories with a national audience in mind.

2.4.2. BCE style script (regional dialect)

To elicit the most dialectal style possible, participants were tasked with reading two poems written in Black Country dialect, chosen for their setting in the region and their frequent use of dialect grammar and vocabulary.

Both the standard and dialect tasks, having two parts each, were split in two and interspersed within the listening task sentences to increase the likelihood that participants remained under the influence of each style as they completed the perceptual element. To mitigate potential biases, participants were told that the reading and listening tasks were separate experiments mixed together to prevent the procedure from becoming tiresome.

3. RESULTS

A total of 112 vowel rating tokens were analysed across the preceding production condition (dialect or standard). As shown in Figure 3, on average, participants selected vowels from the opposite end of the continuum compared to the preceding production condition. In essence, when participants produced more dialect, they perceived the same vowels as sounding more standard.

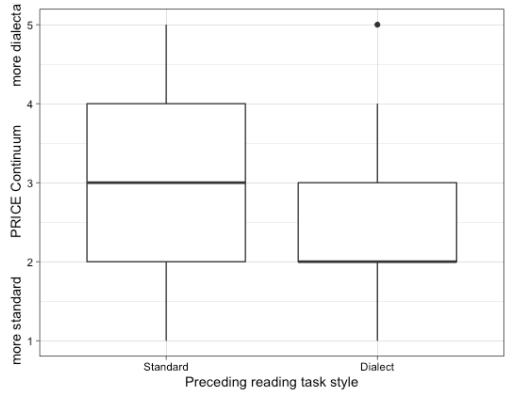

Figure 3: Average PRICE vowel ratings split between preceding reading task style (standard and dialect)

Following previous vowel perception research [13], [12], a gender effect also seems to be present – while the shift in rating between conditions is similar for men and women, the men's ratings seem to occupy a lower part of the scale than those of the women (Figure 4), such that the male participants rated the vowels as more standard-sounding overall.

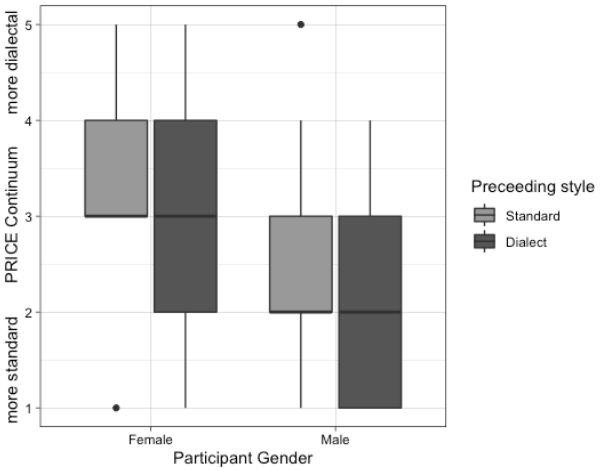

Figure 4: Average PRICE vowel ratings across the two conditions by participant gender

To check the significance of the effect observed, and to control for the potential gender effects, a linear mixed effects regression model was built in R [13] using the 'lme4' package [2], [19], with the vowel ratings as the dependent variable, and the condition and gender as fixed effects, with random intercepts and slopes fitted across the condition for participants.[i] The output model, shown in Table 1, indicates that the *reading task style* is a significant predictor of *vowel rating* response. When participants shift from standard to dialectal production, the coefficient for vowel rating shows a negative difference of -0.37 (Cohen's d = -0.42[ii]), meaning they rate the stimuli vowels as sounding more standard than before.

Fixed effects	Estimate	Std. error	*t*	*p*-value
(Intercept)	3.46	0.27	12.62	<0.0001***
Preceding reading task: dialect	-0.37	0.17	-2.18	0.0317*
Gender: male	-1.02	0.32	-3.19	0.0129*

Number of observations: 112; Groups: Participant (8, SD=0.01)

Table 1: lmer analysis of experimental results

In sum:

- Participants report hearing more dialectal realisations of PRICE when they are asked to produce more standard language and vice versa.
- Male participants report hearing more standard realisations overall, while female participants hear more dialectal realisations, though the direction and extent of the shift is the same.

4. DISCUSSION

The present paper makes a contribution to an emerging field of research at the intersection of language attitudes, speech production and speech perception. We provide evidence that style can be a significant mediator of speech perception, and by extension the interpretation that style may be the source of seemingly attitude-based effects. Additionally, we have presented evidence for such a production-perception relationship alluded to by Walker et al. [21], that as participants are encouraged to shift their production style in one direction, their vowel perceptions shift in the *other*.

While Walker et al. found that manipulation of attitudes produced a change in participants' vowel perceptions, the motivation for such an effect is somewhat unclear. When interpreting their results, the authors refer to previously published production data from the same paradigm [10] and identify that the attitude treatment had a stylistic effect on the participants' own KIT productions. Namely, for the participants who were sports fans, the *positive* facts about Australia led them to produce more *Kiwi* vowels. These findings were interpreted using an *exemplar theory* framework, in which linguistic variables are represented as individual mental entries which must reach a certain level of activation to outcompete others – recent activation can therefore introduce bias as these exemplars have a head start against others [18]. In this way, it could be argued that greater 'activation' of more AuE-type vowel exemplars in production means those same exemplars are more easily activated in perception. However, given that at least *some* of the participants style-shifted to more AuE productions in response to the positive facts, there would remain to be identified a mechanism by which shifts in production can occur in one direction, while shifts in perception occur in the other.

While our data support the intuition that style should be considered a candidate for mediator between attitude and perception, they do not lend themselves to an exemplar-based explanation. While an exemplar model can elegantly explain effects introduced by priming with biographical information such as age or dialect area, it has difficulty in accounting for the findings presented here, which suggest that perception can be biased *away* from rather than *towards* the more recently accessed vowel realisations.

6. CONCLUSIONS

Making use of speakers with especially board stylistic repertoires, the present study contributes evidence that production style seems to influence vowel perception. In this way style, being sensitive to attitudes, may be the 'missing link' where affect seems to influence speech perception. However, it still remains to identify the mechanism by which style can mediate perception.

7. REFERENCES

[1] T. Baguley, 'Standardized or simple effect size: What should be reported?', *British journal of psychology (London, England : 1953)*, vol. 100, pp. 603–17, Dec. 2008.

[2] D. Bates, M. Maechler, B. Bolker, and S. Walker, 'lme4: Linear Mixed-Effects Models using "Eigen" and S4'. 2022.

[3] BBC Archive, 'Black Country dialect', *BBC Local*, 24-Sep-2014. [Online]. Available: https://www.bbc.co.uk/blackcountry/features/2002/12/accents/black_country_dialect.shtml. [Accessed: 11-Aug-2022].

[4] J. Bench, A. Kowal, and J. Bamford, 'The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children', *Br J Audiol*, vol. 13, no. 3, pp. 108–112, Aug. 1979.

[5] P. Boersma and D. Weenink, 'Praat: doing Phonetics by Computer'. 2022.

[6] B. Chakrani, 'Arabic interdialectal encounters: Investigating the influence of attitudes on language accommodation', *Language & Communication*, vol. 41, pp. 17–27, Mar. 2015.

[7] U. Clark and E. Asprey, *West Midlands English: Birmingham and the Black Country*, Annotated edition. Edinburgh: Edinburgh University Press, 2013.

[8] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. New York: Routledge, 1988.

[9] K. Drager, 'Speaker Age and Vowel Perception', *Lang Speech*, vol. 54, no. 1, pp. 99–121, Mar. 2011.

[10] K. Drager, J. Hay, and A. Walker, 'Pronounced Rivalries: Attitudes And Speech Production', *Te Reo*, vol. 53, pp. 27–53, Jan. 2010.

[11] P. Garrett, 'Communication accommodation theory', in *Attitudes to Language*, 1st ed., Cambridge University Press, 2010, pp. 105–120.

[12] J. Hay, A. Nolan, and K. Drager, 'From fush to feesh: Exemplar priming in speech perception', vol. 23, no. 3, pp. 351–379, Oct. 2006.

[13] J. Hay, P. Warren, and K. Drager, 'Factors influencing speech perception in the context of a merger-in-progress', *J. Phonetics*, 2006.

[14] K. Hazen, 'An Introductory Investigation Into Bidialectalism', *University of Pennsylvania Working Papers in Linguistics*, vol. 7, no. 3, Jan. 2001.

[15] D. Lawrence, 'A Brief Introduction to Vowel Resynthesis for Speech Perception Research'. University of Edinburgh, 2015.

[16] N. Niedzielski, 'The Effect of Social Information on the Perception of Sociolinguistic Variables', *Journal of Language and Social Psychology*, vol. 18, no. 1, pp. 62–85, Mar. 1999.

[17] Office for National Statistics, 'Population and household estimates, England and Wales', *Census 2021*, 28-Jun-2022. [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationandhouseholdestimatesenglandandwales/latest. [Accessed: 11-Aug-2022].

[18] J. B. Pierrehumbert, 'Exemplar dynamics: Word frequency, lenition and contrast', in *Typological Studies in Language*, vol. 45, J. L. Bybee and P. J. Hopper, Eds. Amsterdam: John Benjamins Publishing Company, 2001, p. 137.

[19] R Core Team, 'R: A Language and Environment for Statistical Computing'. R Foundation for Statistical Computing, Vienna, Austria, 2022.

[20] J. Smith and M. Durham, 'Bidialectalism or Dialect Death? Explaining Generational Change in the Shetland Islands, Scotland', *American Speech*, vol. 87, no. 1, pp. 57–88, Mar. 2012.

[21] A. Walker, J. Hay, K. Drager, and K. Sanchez, 'Divergence in speech perception', *Linguistics*, vol. 56, no. 1, pp. 257–278, Jan. 2018.

[i] Call formula: lmer(VowelRating ~ Condition + Gender + (1+Condition|Participant), PRICE_data, REML=FALSE)

[ii] It is important to acknowledge the limitations [1] of relying on effect size statistics such as Cohen's [8] *d,* and for this reason interpretation of the effect is also based on the coefficients.