

# Plasticity in vowel perception and production: A study of accent change in young adults

Bronwen G. Evans and Paul Iverson

*Department of Phonetics and Linguistics, University College London, 4 Stephenson Way, London NW1 2HE, United Kingdom*

(Received 30 December 2005; revised 12 March 2007; accepted 12 March 2007)

This study investigated changes in vowel production and perception among university students from the north of England, as individuals adapt their accent from regional to educated norms. Subjects were tested in their production and perception at regular intervals over a period of 2 years: before beginning university, 3 months later, and at the end of their first and second years at university. At each testing session, subjects were recorded reading a set of experimental words and a short passage. Subjects also completed two perceptual tasks; they chose best exemplar locations for vowels embedded in either northern or southern English accented carrier sentences and identified words in noise spoken with either a northern or southern English accent. The results demonstrated that subjects at a late stage in their language development, early adulthood, changed their spoken accent after attending university. There were no reliable changes in perception over time, but there was evidence for a between-subjects link between production and perception; subjects chose similar vowels to the ones they produced, and subjects who had a more southern English accent were better at identifying southern English speech in noise. © 2007 Acoustical Society of America.

[DOI: 10.1121/1.2722209]

PACS number(s): 43.71.Es, 43.71.Ft [ARB]

Pages: 3814–3826

## I. INTRODUCTION

Native speakers of regional accents typically modify their speech when living in a community where that accent is not commonly spoken (e.g., Munro *et al.*, 1999; Sankoff, 2004). Speakers often avoid variants that are markedly regional or unusual to facilitate communication (Trudgill, 1986) and to fit in with their new community (Foulkes and Docherty, 1999). However, they also retain some regional variants in order to show their allegiance to particular social or geographical groups (e.g., Foulkes and Docherty, 1999; Trudgill, 1986). For example, in British English, the focus of the present study, speakers of northern English who live in southern England typically modify their production of the vowel in words such as *luck* from a high back vowel [ʊ] to a centralized vowel, [ə], so that it is closer to how southerners produce this vowel (Evans and Iverson, 2004; Trudgill, 1986; Wells, 1982). However, they also retain [a] when producing words like *bath* just like in their native northern English accent, rather than producing it with the southern vowel [ɑ:] (Evans and Iverson, 2004; Trudgill, 1986; Wells, 1982).

The present study focused on accent change in a group of university students from a small town in the Midlands, U.K., where the local accent is a variety of northern English, a nonstandard accent. In Britain, it is usual for students to attend university in a different area from the one in which they have been raised, and so students come into contact with speakers of a wide variety of accents. In particular, they come into contact with speakers of Standard Southern British English (SSBE), the prestige accent of English and the accent of education.<sup>1</sup> As a result, students from the north of England typically change their accent from regional to educated (i.e., SSBE) norms in order to better fit in with their

new university community. The aim of this study was to quantify these changes in accent among university students from the north of England, and to investigate whether the adjustments in production were accompanied by changes in perceptual processing.

One could imagine that perception and production might be based on the same underlying phonetic categories, and that any changes in production would be accompanied by changes in perception (e.g., Liberman *et al.*, 1967; Liberman and Mattingly, 1985; Fowler, 1981, 1986). However, the evidence for such a strong relationship between perception and production is mixed. Bell-Berti *et al.* (1979) demonstrated that differences in production strategies for the vowel /i/ were significantly correlated with differences in perception for /i/. They took this as evidence for a common mechanism or process that mediates the production and perception of vowels. Newman (2003) also found significant, though small, correlations between listeners' perceptual prototypes and their average voice onset time (VOT) for stop consonants, as well as between their perceptual prototypes and spectral peaks for voiceless fricatives. Other studies have failed to find such correlations. For example, Bailey and Haggard (1973, 1980) found no correlation between average VOTs produced in voiced and voiceless consonants and listeners' perceptual category boundaries for a /g/-/k/ continuum. Similarly, Ainsworth and Paliwal (1984) found no correlation between performance on production and perception tasks for glides, and Frieda *et al.* (2000) for the vowel /i/.

Results from studies of production and perception in second language (L2) learning are also not clear-cut. Flege *et al.* (1997) found evidence for correlations between the perception and production of the four English vowels, /i/, /I/, /ε/

and /a/ by non-native learners of English, but the magnitude of these correlations was small (see also Flege *et al.*, 1999). Other studies of L2 acquisition have found no evidence for the existence of these perception-production correlations. For example, Bradlow *et al.* (1997) demonstrated that Japanese learners improved in both their production and perception of the phonemes /t/ and /l/ after perceptual training, but found no significant correlations between individual learners' improvement in perception and production. That is, even though subjects changed their production and perception, the subjects who showed the most change in production did not always show the most change in perception.

Even if there is no direct link between changes in production and perception, subjects may still adjust their perceptual processes when listening to a non-native accent. For example, studies of talker-specific learning (e.g., Nygaard and Pisoni, 1998) have shown that after only a short amount of experience with an unfamiliar talker, word recognition for that talker improves. Moreover, listeners are able to adapt rapidly to foreign-accented speech (Clarke and Garrett, 2004; Nygaard *et al.*, 2005); adaptation to foreign-accented speech has been shown to require exposure to only two to four sentence-length utterances (Clarke and Garrett, 2004), and listeners appear to be able to shift their VOT categorization boundary for stop consonants to better match a speaker's productions after less than 2-min experience with that speaker (Clarke and Luce, 2005). Based on this evidence, one might expect that in a multidialectal situation, where subjects are regularly exposed to a non-native accent, subjects would be able to easily adapt perceptually to that accent. In the present study, northern listeners who are regularly exposed to SSBE might thus be able to easily adjust their perceptual processes to better adapt to SSBE.

However, our evidence (Evans and Iverson, 2004) suggests that adaptation to a non-native accent may be slower. In this previous work, we investigated whether listeners from the north of England changed their best exemplar locations when listening to speech produced in their native accent and SSBE. Subjects with different backgrounds were tested; northerners who had been living in London (south of England) for a minimum of 1 year, and northerners aged 16–17 years who had been born and raised in the north of England. The results demonstrated that northerners living in the south of England adjusted their best exemplar locations when listening to SSBE speech; listeners chose best exemplar locations for vowels in SSBE sentences that more closely matched how southerners produce these vowels. However, northerners still living in the north of England did not adjust their best exemplar locations; these subjects used their native (i.e., northern) vowel categories when listening to both SSBE and northern English. This was surprising as these northerners had extensive experience of listening to southern English speech through the media. This suggests that, although subjects are able to adjust their perceptual categorizations, these changes take place over a long period of time and are not just determined by short-term exposure to a particular speaker or accent (see also Eisner and McQueen, 2005).

The present study used a longitudinal design to investigate whether subjects changed their production as a result of attending university, and whether these changes were accompanied by changes in perception. Students were tested before beginning university, 3 months later, and on completion of their first and second years at university. At each testing session they completed three experiments. Experiment 1 investigated whether subjects changed their speech production. Experiment 2 created perceptual vowel space maps (i.e., best exemplar locations) in a three-dimensional space that varied  $F1$ ,  $F2$ , and duration for vowels embedded in SSBE and northern English carrier sentences. The aim was to investigate if subjects changed the acoustic qualities of the vowels according to whether the sentence was spoken with an SSBE or northern English accent. Further analysis addressed whether any changes in the perceptual vowel space maps were linked to changes in production. Experiment 3 investigated whether any changes in production and the best exemplar locations had an effect on sentence recognition in noise.

## II. EXPERIMENT 1: MEASUREMENT OF PRODUCTION

### A. Method

#### 1. Participants

Twenty-seven subjects were tested and paid for their participation. All were native monolingual English speakers and reported no speech, hearing, or language difficulties.

Subjects were recruited from Ashby de la Zouch, Leicestershire, UK, a small market town in the Midlands where the local accent can be classified as a variety of northern English (Evans, 2005; Wells, 1982). At the time of recruitment, subjects were aged 17–18 years, and were completing their school education at Ashby Grammar School, the local comprehensive school. All subjects had lived in Ashby since at least age 5 years,<sup>2</sup> and had been educated at local schools. This period of development, characterized by the influence of the peer group, is thought to be particularly important for the development of regional accent (Kerswill and Williams, 2000). All had parents and immediate family local to the area, minimizing the risk that subjects had regular contact with speakers of southern English accents.

Subjects attended different universities in England,<sup>3</sup> and so all testing was carried out in Ashby. Students in Britain typically attend university in a different area from the one in which they have been raised, so the location of the university has little influence on the accent background of the student body as a whole. Thus, subjects attending university in the north of England and those attending university in the south of England were both likely to interact with SSBE speakers. All students lived in university accommodation and reported that they spent the majority of their time interacting with fellow students.

Two subjects did not complete the experiment because they dropped out of university during their first term. Of the remaining 25 subjects, two subjects were dropped from the experiment because their best exemplar locations for the vowels found in experiment 2 were not reliable (i.e., their

best exemplar locations for words that are produced the same in SSBE and Sheffield English accents, such as in the words *bird* and *bead*, differed by more than 2 ERB in the two carrier sentences). This gave a test sample of 23 subjects, 7 male and 16 female. Due to problems in contacting all of the subjects, only 19 of these subjects were tested at the last testing time.

## 2. Stimuli and apparatus

The stimuli consisted of 11 test words in the carrier sentence *I'm asking you to say the word [ ] please*, and a phonetically balanced passage, "Arthur the Rat." The test words were *bad*, *bard*, *bawd*, *bed*, *bird*, *booed*, *bud*, *bead*, *cud*, *could*, and *bath*. This allowed potential changes in subjects' best exemplar locations (experiment 2) to be compared with changes in production. Of particular interest were the vowels in the words *bud*, *cud*, and *bath* that are produced differently in northern and southern English accents (e.g., Wells, 1982). Speakers of SSBE produce words like *bud* and *cud* using the vowel [ʌ]. However, speakers of northern English do not have this vowel in their phoneme inventory; they say words like *bud* and *cud* using the high back vowel [ʊ], such that *bud* is produced with the same vowel as *book*, and *cud* becomes a homophone or near homophone of *could*. Northerners and southerners both have the vowels [a] and [ɑ:] but with somewhat different lexical distributions; words like *bath* are produced with [ɑ:] by southerners, but with [a] by northerners.

All recordings were made in a quiet room and recorded onto DAT with a sampling rate of 44.1 kHz, 16-bit resolution.

## 3. Procedure

*a. Recording.* Subjects recorded two repetitions of each target word in a randomized order and one reading of the passage. Subjects were instructed to read the passage as if they were reading it to a friend. The researcher was the first author, who lived in Ashby from age 2–18 years when she moved away to attend university. Although she still uses some local features in her accent (e.g., short "a," [a], in *bath*), she has changed her accent so that it is more similar to that of the SSBE speakers in the community where she now lives. She conducted all testing and was known to the subjects as the sister of a friend. It was thought that the relationship of the researcher to the subjects and the researcher's own linguistic background would encourage subjects to speak in the way that they would when interacting with friends at university (Milroy, 1987a, 1987b).

*b. Accent ratings.* Six phonetically trained listeners rated for accent samples of subjects' speech from each testing session [i.e., before beginning university (Time 1; T1), 3 months later (Time 2; T2), after 1 year (Time 3; T3), and after 2 years (Time 4; T4)]. All listeners were native British English speakers. They gave ratings on a short sample of the reading passage: *Just then the old captain saw Arthur. "Stop," he ordered the others coarsely. "You are coming, of course?" "I'm not certain," said Arthur, undaunted. "The roof may not come down yet."* This passage was selected

because it was thought to be revealing of accent. That is, it contains examples of the vowels that were expected to change after living in a multidialectal environment (e.g., the vowels in the words *just*, *coming*, and *come*).

The rating sessions were self-paced and listeners could listen to each sample as many times as they wished. Listeners gave their ratings by indicating on a response sheet, on a scale of 1 to 10, how northern or southern they thought the speaker sounded. A rating of "10" corresponded to "very northern" and a response of "1" to "very southern." Stimuli were presented in a random order (i.e., all recording times presented together). Listeners were not made aware of the design of the study. That is, they were not told that they would hear the same speaker more than once, and were unaware that speakers were expected to have changed their accent. All listeners were native British English speakers.

*c. Acoustic analysis.* *F1*, *F2*, and duration were measured for the vowels in each target word spoken in the carrier sentence, giving two sets of measurements for each word. *F1*, *F2*, and duration were averaged across repetitions of each word and these measurements were used in the subsequent analyses. The test words rather than the passage were used because this enabled any effects of consonantal context to be controlled. In contrast, the accent ratings were made on the passage to give a more global measure of each subject's accent.

The acoustic measurements were made in PRAAT (Boersma and Weenink, 2004). The stimuli were located manually, and then *F1* and *F2* were extracted using hand-corrected LPC analyses. Formant frequencies were measured from the midpoint of the steady-state portion of the vowel. The steady-state portion of the vowel was defined as the part of the vowel that was closest to the midpoint and where the formant frequencies were most stable. All duration measurements were taken from the beginning of the *F2* transitions to the end of the *F2* transitions.

So that the data from male and female speakers could be compared, a version of Nearey's (1978) individual log mean procedure was used to normalize the acoustic measurements. The procedure expresses each log-transformed formant frequency as a distance to a reference point, the log mean, which is calculated by averaging the log-transformed *F1* and *F2* values for all vowels for a given speaker. This procedure was chosen because it has been shown to be one of the most effective methods for reducing the effects of anatomical and physiological variation, while retaining the phonemic variation that is important for accent variation (Adank, 2003; Adank *et al.*, 2004). The average formant frequencies for each test word were normalized following the equation:

$$F_{ijk \text{ norm}} = G_{ijk} - \bar{G}, \quad (1)$$

where *i* is the formant, *j* is the vowel being transformed, *k* is the speaker, *G* is the log-transformed frequency of formant *i*, and  $\bar{G}$  is the log mean for a speaker *k*. In this study the log mean was calculated using the vowels that are produced similarly in northern and southern English accents (i.e., *bad*, *bard*, *bawd*, *bead*, *bed*, *bird*, and *booed*) for that talker, averaged over all testing sessions. The vowels in *bud*, *cud*,

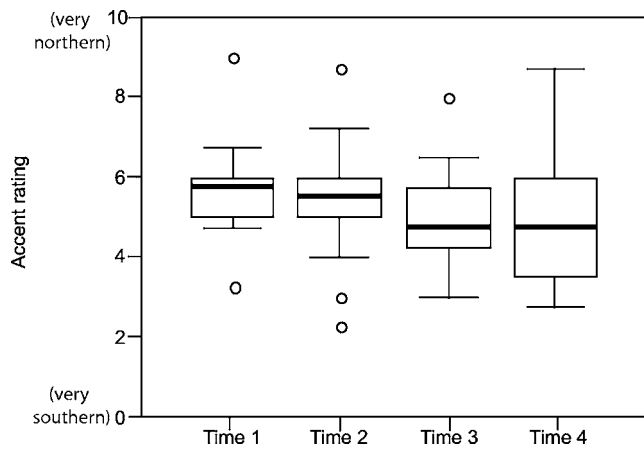


FIG. 1. Boxplot of the accent ratings for subjects at T1, T2, T3, and T4 ( $N=23$  at T1, T2, and T3;  $N=19$  at T4). Boxplots display the interquartile range of scores. The box shows the 25th to 75th percentiles, with a line at the median value. The lower and upper “whiskers,” respectively, show the minimum and maximum values of the data range, with outliers represented by the unshaded circles. Subjects changed their spoken accent over time; subjects were given a more southern (i.e., lower) rating at T4 than at T1.

*could*, and *bath* were omitted from the log mean calculation because these vowels were expected to change over time.

## B. Results

### 1. Perceived accent rating

As displayed in Fig. 1, subjects changed their accent over time. Subjects were judged to sound more southern at T4 than at T1, though there appeared to be little change

between the testing sessions at T3 and T4, and there was more variability in the ratings at T4 than at T1, T2, or T3 (see Table I). The potential differences between the ratings at T1, T2, T3, and T4 were tested in a repeated-measures ANOVA, with time (T1, T2, T3, and T4) coded as a within-subject variable. There was a main effect of time,  $F(1, 20) = 22.94$ ,  $p < 0.001$ , confirming that subjects changed their accent to sound more southern after attending university.

However, as displayed in Table I, there were differences in the amount and direction of accent change between subjects. Although the majority of subjects (18 out of 23) was rated as sounding more southern at the final testing time than at T1, the change in accent rating<sup>4</sup> from T1 to the final testing session was often small; for nine subjects, the change in accent rating was less than 1 point on the 10-point rating scale. Also, three subjects were judged not to have changed their accent at all and one was judged to have a more northern accent at the final testing session than T1.

There were also large individual differences in the accents of the talkers. For example, as can be seen from the data in Table I, one subject had an accent rating greater than 8 at T1, T2, T3, and T4 (i.e., very northern), but another subject had an accent rating between 2.67 and 2.83 at T1, T2, T3, and T4 (i.e., very southern). These differences can be seen clearly when the overall accent rating (i.e., the average of each subject’s rating at T1, T2, T3, and T4 or T1, T2, and T3 when the subject could not be contacted at T4), is calculated (Table I).

A Pearson correlation investigated whether the change in accent rating was related to the overall accent rating, that is,

TABLE I. Individual accent ratings for subjects at T1, T2, T3, and T4, the overall accent rating and change in accent rating.

Sex	Accent rating				Overall accent rating	Change in accent rating (final–first testing time)
	T1	T2	T3	T4		
F	6.50	5.67	6.33	...	6.21	-0.17
F	6.83	7.33	5.67	6.33	6.54	-0.50
F	6.83	5.83	5.50	...	5.92	-1.33
M	6.50	6.67	4.17	4.17	5.38	-2.33
F	5.33	4.50	4.83	3.67	4.58	-1.67
F	6.33	5.83	5.33	4.17	5.42	-2.17
F	6.50	6.33	6.67	6.00	6.38	-0.50
F	5.17	5.50	5.50	6.17	5.58	+1.00
F	5.83	5.83	4.83	5.00	5.38	-0.83
F	2.83	2.83	2.67	...	2.72	-0.17
M	8.00	6.83	5.83	5.83	6.63	-2.17
F	6.67	6.33	5.67	5.50	6.04	-1.17
F	6.50	6.17	5.83	6.33	6.21	-0.17
M	6.67	6.33	5.83	5.50	6.08	-1.17
M	5.83	6.33	5.83	5.00	5.75	-0.83
M	...	8.00	7.50	6.33	7.28	-1.67
F	3.33	5.33	3.67	3.33	3.92	0.00
F	6.17	5.83	4.83	6.17	5.75	0.00
F	6.17	3.67	5.50	...	5.21	-0.67
F	5.33	5.83	4.83	4.67	5.17	-0.67
F	5.83	4.00	5.33	2.83	4.50	-3.00
M	8.67	8.33	8.17	8.67	8.46	0.00
M	5.00	3.33	3.83	3.83	4.00	-1.17

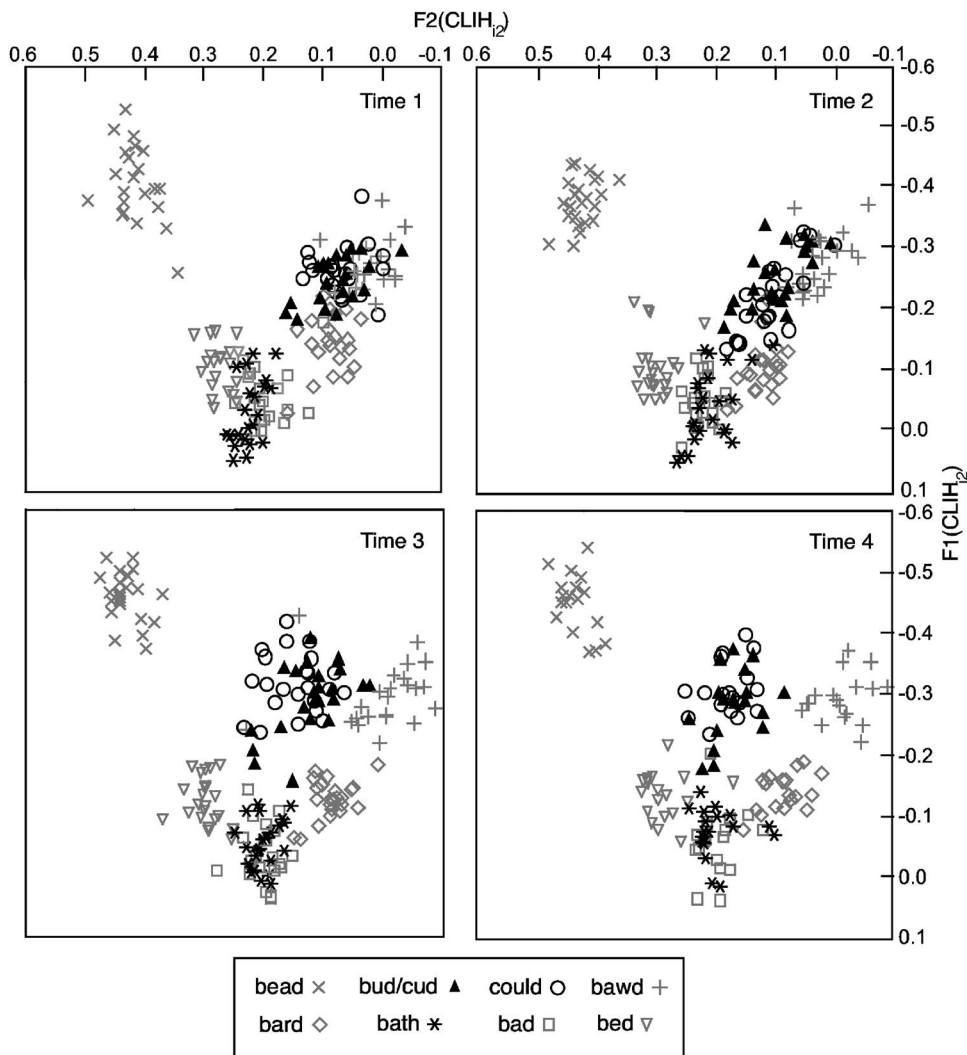


FIG. 2.  $F_1$  and  $F_2$  formant frequency plots of subjects' productions of target words at T1, T2, T3, and T4. *Bud* and *cud* (averaged), *could* and *bath* (i.e., the vowels that were expected to change with accent), are shown in black symbols. All other vowels (i.e., those that are produced similarly in both northern and southern English accents) are shown in gray. Measurements are given in Nearey's log mean units (CLIH<sub>12</sub>). This measure was used to normalize the data so that male and female speech could be compared. Subjects changed their production of *bud*, *cud*, and *could* so that it was more centralized at T4. This change occurred in both the  $F_1$  and  $F_2$  dimensions. No changes occurred in any other vowels.

whether subjects who were judged to sound more northern at T1 changed their accent more than those who were judged to sound more southern. The results demonstrated that there was no significant relationship between the overall accent rating and the change in rating,  $r=0.35$ ,  $p>0.05$ .

## 2. Acoustic analysis

Based on our previous work (Evans and Iverson, 2004), test words were divided into groups that would be likely to change (i.e., *bud*, *cud*, *could*, and *bath*) and groups that would likely not change (i.e., all other words). This enabled us to avoid multiple statistical tests.

*a. Bud and cud.* As displayed in Fig. 2, subjects changed the way in which they produced the vowels in *bud* and *cud*. At T1, subjects produced these words with a high-back vowel, as is typical for speakers of northern English accents. At T2, some, but not all, subjects had begun to centralize these vowels. This centralization process continued over time, so that at T3 and T4 almost all subjects produced these words with a more central vowel that is acoustically similar to the vowel in *bird*. The shift appeared to occur in both the  $F_1$  and  $F_2$  dimensions, though the amount of change appeared to differ amongst subjects.

The potential changes in  $F_1$  and  $F_2$  were tested in separate repeated-measures ANOVA analyses, with word (*bud* or *cud*) and time (T1, T2, T3, and T4) coded as within-subject variables. For  $F_1$ , there was a main effect of time,  $F(3,48)=6.80$ ,  $p<0.01$ , demonstrating that subjects produced this vowel with a higher  $F_1$  at T4. There was also a main effect of time for  $F_2$ ,  $F(3,48)=29.33$ ,  $p<0.001$ , demonstrating that subjects produced this vowel with a higher  $F_2$  at T4. There was a significant interaction of word and time for both  $F_1$ ,  $F(3,48)=8.22$ ,  $p<0.01$ , and  $F_2$ ,  $F(3,48)=10.08$ ,  $p<0.001$ . Inspection of the data revealed that this was because changes in  $F_1$  and  $F_2$  for *cud* occurred faster than those for *bud*; at T1 *bud* was produced with a higher  $F_1$  and  $F_2$  (i.e., more central) than *cud*, but at T2 *cud* was produced with a higher  $F_1$  and  $F_2$  than *bud*. There was also a main effect of word for  $F_2$ ,  $F(1,16)=13.42$ ,  $p<0.01$ ; overall  $F_2$  was lower (i.e., further back in the vowel space) for *cud* than for *bud*.

As displayed in Table II, there was little change over time in duration for *bud* and *cud*, although subjects appeared to be producing a shorter vowel at T1, T2, T3, and T4 for *cud*. A repeated-measures ANOVA for duration revealed that there was a main effect of word,  $F(1,16)=38.24$ ,  $p<0.001$ , indicating that subjects were producing a shorter vowel for

TABLE II. Average durations (ms) of vowels in test words at T1, T2, T3, and T4.

Word	T1	T2	T3	T4
bud	135.67	133.65	136.67	120.11
cud	119.24	117.62	117.38	93.67
could	124.48	123.62	123.10	95.06
bath	121.38	122.43	121.71	115.78
bad	160.85	170.62	159.42	139.21
bard	268.53	297.02	266.00	272.10
bawd	262.72	320.99	262.74	247.86
bead	239.13	292.84	238.38	283.40
bed	140.55	184.62	140.07	121.31
bird	269.51	332.52	268.02	248.87
booeed	252.64	304.81	251.03	234.32

*cud*, but there was no significant main effect of time, and no significant interaction of word and time,  $p > 0.05$ .

*b. Could.* As displayed in Fig. 2, subjects changed the way in which they produced the vowel in *could*. At T1, subjects produced this word with a high-back vowel, as is typical for speakers of northern and southern English accents. At T2, some, but not all, subjects had begun to centralize this vowel, and by T3 and T4 almost all subjects produced this word with a more central vowel. This is surprising as both northerners and southerners produce this word with the same vowel, [ʊ]. As for *bud* and *cud*, the shift appeared to occur in both the  $F1$  and the  $F2$  dimension, though the degree of change appeared to differ among subjects. The differences in  $F1$  and  $F2$  were tested in separate repeated-measures ANOVA analyses, with time coded as a within-subject variable. There was a significant effect of time for  $F1$ ,  $F(1, 16) = 30.48$ ,  $p < 0.001$ , and  $F2$ ,  $F(1, 16) = 39.00$ ,  $p < 0.001$ . As displayed in Table II, there did not appear to be any change in duration. A repeated-measures ANOVA confirmed that there was no significant effect of time,  $p > 0.05$ .

*c. Bath.* As displayed in Fig. 2, there appeared to be some changes in the production of *bath* in the  $F1$  and the  $F2$  dimensions; some subjects seemed to be producing this word with a lower  $F1$  and lower  $F2$  at T3 and T4, so that it was closer to how they produced the vowel in *bard*. The differences in  $F1$ ,  $F2$ , and duration were tested in separate repeated-measures ANOVA analyses, with time coded as a within-subject variable. There was a significant effect of time for  $F1$ ,  $F(1, 16) = 7.11$ ,  $p < 0.05$ , and  $F2$ ,  $F(1, 16) = 9.39$ ,  $p < 0.01$ , confirming that subjects changed their production of this vowel in these dimensions. For duration, there was no significant effect of time,  $p > 0.05$ , indicating that subjects produced this vowel with a similar duration at each testing session.

*d. Other vowels.* As displayed in Fig. 2, subjects did not appear to change their production of any other vowels, although there was a possible shift in *bawd* in the  $F2$  dimension. The potential differences in  $F1$  and  $F2$  were tested in separate repeated-measures ANOVA analyses, with word (i.e., all words other than *bud*, *cud*, *could*, and *bath*) and time coded as within-subject variables. There was a main effect of word for  $F1$ ,  $F(1, 16) = 20.50$ ,  $p < 0.001$ , and  $F2$ ,  $F(1, 16) = 7.26$ ,  $p < 0.05$ , demonstrating that subjects produced these

words with different formant frequency values, but there was no main effect of time and no significant interaction of word and time,  $p > 0.05$ . The shift in *bawd* was thus not reliable.

As displayed in Table II, subjects chose different durations for different words, but there did not appear to be any change in duration from T1 to T4. A repeated-measures ANOVA analysis confirmed that there was a main effect of word,  $F(1, 16) = 20.04$ ,  $p < 0.001$ , demonstrating that subjects produced these words with different durations, but there was no main effect of time and no significant interaction of word and time,  $p > 0.05$ .

*e. Comparison with accent ratings.* To verify that the changes observed in the  $F1$  and  $F2$  frequency measurements for *bud*, *cud*, *could*, and *bath* were related to the change in accent ratings, individual accent ratings from each testing session were correlated with  $F1$  and  $F2$  values for *bud*, *cud*, *could*, and *bath*. Significant correlations were found between the accent ratings and  $F1$  for *bud*,  $r = -0.44$ ,  $p < 0.05$ ,  $F2$  for *bud*,  $r = -0.61$ ,  $p < 0.05$ ,  $F2$  for *cud*,  $r = -0.55$ ,  $p < 0.05$ , and  $F2$  for *could*,  $r = -0.5$ ,  $p < 0.05$ . This indicates that the observed changes in formant frequencies corresponded to the changes in perceived accent.

### 3. Discussion

The results demonstrated that subjects changed their accent after attending university. Although the change in accent rating was relatively small, the acoustic analysis demonstrated that there were substantial changes in how these speakers produced *bud*, *cud*, *could*, and *bath*. For *bud*, *cud*, and *could* this shift occurred in both the  $F1$  and  $F2$  dimensions; subjects produced a more fronted and lower vowel in these words at T4 than at T1, which is closer to how southerners produce the vowel in words like *bud* and *cud*. For *could* this was particularly surprising because northerners and southerners produce this word using the same vowel, [ʊ]. It is possible that subjects produced *could* with a more centralized vowel, just like the vowel they used to produce *bud* and *cud*, because they were unable to split a category that is merged in their native accent. In SSBE, words such as *cud* and *could* are produced with two different vowels, [ʌ] and [ʊ]. However, northern English speakers do not have the vowel [ʌ]; they produce these words with the same vowel [ʊ], such that *cud* and *could* are homophones. One could imagine that these subjects had changed their underlying [ʊ] category so that it was closer to how SSBE speakers produce words like *bud* and *cud*, and that they used this new vowel category in all words containing [ʊ] in their native accent, even when this did not match how those words are produced in SSBE (see also Sankoff, 2004; Evans and Iverson, 2004). Equally, subjects may have acquired a new category alongside their native [ʊ] category, but were hypercorrecting (Evans, 2005; Sankoff, 2004; Wells, 1982). That is, subjects may have used a more central vowel in words where they would use the lower back vowel [ʊ] in their native accent in order to “posh-up” their accent (Wells, 1982, p. 353).

Subjects also modified their production of *bath*; they changed their  $F1$  and  $F2$  formant frequency locations over time. This was surprising as northerners typically retain their native pronunciation of this vowel when adjusting to SSBE

speech (Trudgill, 1986; Wells, 1982). However, the amount of change in the  $F1$  and  $F2$  dimension was much smaller than that in *bud*, *cud*, and *could*, and there was no change in duration. If subjects had been adjusting their pronunciation to match SSBE speakers, then there should have been a significant change in duration; southerners produce words like *bath* with the long vowel [ɑ:] but northerners use the short vowel [a]. This indicates that although subjects were making small changes to their production of this vowel, they were still using a vowel that would be categorized as [a]. It is perhaps more likely that subjects were retaining [a] in words like *bath*, but were making small changes to this category in order to “soften” their native accent (i.e., to sound less northern).

### III. EXPERIMENT 2: AN INVESTIGATION OF PERCEPTUAL VOWEL SPACES

In experiment 2, subjects created perceptual vowel space maps (i.e., best exemplar locations) in a three-dimensional space that varied  $F1$ ,  $F2$ , and duration for vowels embedded in SSBE and northern English carrier sentences. The aim was to investigate if subjects changed the acoustic qualities of the vowels according to whether the sentence was spoken with an SSBE or northern English accent. Subjects heard synthesized vowels embedded in natural carrier sentences that were produced in either a northern English (i.e., native) or SSBE (i.e., non-native) accent. They gave goodness ratings on the vowels and a computer program iteratively adjusted the  $F1$ ,  $F2$  and duration values on successive trials until a best exemplar was found. Of particular interest were the vowels in *bud*, *cud*, *could*, and *bath*, as subjects had changed their production of these words to better fit in with southerners.

#### A. Method

##### 1. Participants

Same as experiment 1.

##### 2. Stimuli and apparatus

The stimuli were the same as those used in Evans and Iverson (2004). They consisted of synthesized vowels in the phonetic environments /b/-V-/d/, /b/-V-/θ/, and /k/-V-/d/, embedded in natural recordings of the carrier sentence *I'm asking you to say the word [ ] please*. This carrier sentence was chosen because it contained an overt cue to accent; the word *asking* is produced with a long, back vowel [ɑ:] in SSBE, but in Sheffield English with a short vowel, [a]. There were 11 test words; *bad*, *bard*, *bawd*, *bed*, *bird*, *booed*, *bud*, *bead*, *cud*, *could*, and *bath*.

The carrier sentence was produced in both Sheffield English and SSBE accents by the same male speaker ([əmaskɪnjə?ser?w?d---plɪzɪŋ] in Sheffield English and [ɑmɑ:skɪŋju:təseiðwɜ:d---plɪzɪŋ] in SSBE). The speaker had lived in Sheffield until the age of 19 years when he moved to the south of England to attend university, where he had lived for 7 years at the time of recording. This speaker was selected because he was able to switch between accents at will, and was able to produce versions of both accents that were judged by trained phoneticians to sound like those of

native speakers. In addition to the carrier sentences, the speaker was recorded reading a 2-min passage from a novel in both accents. This was used for familiarization.

CVCs were embedded in the carrier sentences. The bursts, fricatives, and aspiration were spliced from the sentence recording, and the voiced portions were synthesized in advance using the cascade branch of a Klatt synthesizer (Huckvale, 2003; Klatt and Klatt, 1990). Each stimulus had a middle portion in which the formant frequencies were static, and had formant transitions appropriate for the consonants. All transitions were linear. The stimuli varied in terms of  $F1$  and  $F2$  frequencies and duration of the middle portion.  $F1$  frequency was restricted so that it had a lower limit of 150 Hz and an upper limit of 950 Hz.  $F2$  was restricted to have a lower limit of  $F1+50$  Hz, and had an upper limit defined by the equation

$$F2_{\text{upper-limit}} = 3000 \text{ Hz} - 1.7 * F1. \quad (2)$$

The duration of the middle portion was restricted to be greater than 20 ms and less than 403 ms.  $F3$  frequency was fixed at 2500 Hz. All other synthesis parameters were chosen to mimic the natural speech. After synthesis, the CVC stimuli were processed using a multiband filter to fine-tune the match between the long-term average spectra of the synthetic and natural speech (see Evans and Iverson, 2004; Evans, 2005, for a fuller description of this process).

The entire range of possible vowels was synthesized with a resolution of 0.5 ERB in  $F1$  and  $F2$ . Duration was quantized in 16 steps on a log scale. There were a total of 7616 stimuli synthesized for each of the CVC contexts.

The stimuli were played at a sampling rate of 11 kHz using a computer sound card, over headphones (Sennheiser HD 414) in a quiet room.

#### 3. Procedure

The procedure was the same as that described in Evans and Iverson (2004). There were two testing sessions, one for each accent. The order of sessions was counterbalanced across subjects. Sessions were conducted on separate days (i.e., at least 1 day apart) to minimize the risk that subjects would be influenced by the fact that the speaker was the same in both conditions.

At the start of each session, subjects listened to a short passage read by the speaker in order to familiarize them with the accent. They then found the best exemplar for one practice word (*kid*), and best exemplars for 11 test words in succession: *bad*, *bard*, *bawd*, *bed*, *bird*, *booed*, *bud*, *bead*, *cud*, *could*, and *bath*. The order of presentation was randomized across subjects. To find the best exemplars, the Goodness Optimization procedure was used (Evans and Iverson, 2004; Iverson and Evans, 2003; Iverson *et al.*, 2005). Subjects heard a synthesized word embedded in a carrier sentence on each trial, and rated whether it was close to being a good exemplar of the target word that was displayed orthographically on a computer screen. They were instructed to make their rating based on what sounded like a good version of the word to them in the sentence context. They gave their response by positioning and clicking a computer mouse on a continuous scale from *close* to *far away*. Based on their judg-

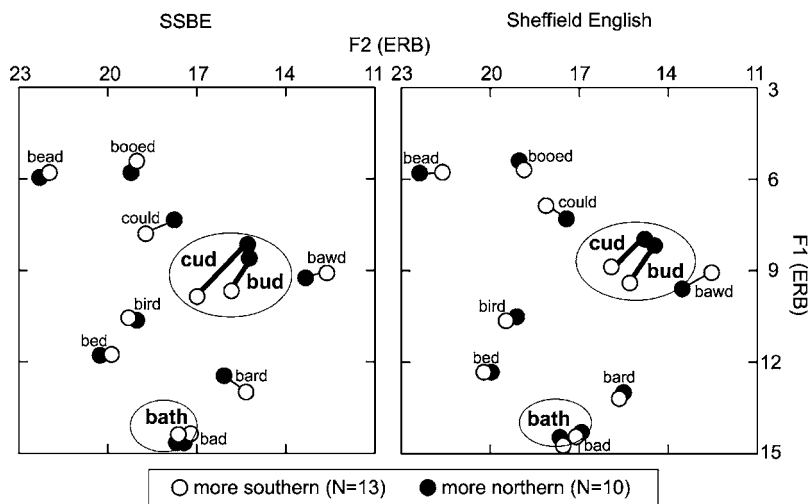


FIG. 3. Average  $F1$  and  $F2$  formant frequency (ERB) locations for best exemplars in SSBE and Sheffield English carrier sentences. The open circles represent the average formant frequency locations at T1, T2, T3, and T4 for subjects who had an overall accent rating lower than or equal to the median overall accent rating (i.e., had a more southern accent overall;  $N=13$ ). The filled circles represent the average formant frequency locations for subjects who had an overall accent rating greater than the median overall accent rating (i.e., had a more northern accent overall;  $N=10$ ). Subjects who had a more southern accent overall chose more southern vowels, and subjects who had a more northern accent overall chose more northern vowels.

ments, the computer program iteratively adjusted the acoustic parameters ( $F1$ ,  $F2$ , and duration) to converge on a best exemplar. The acoustic parameters were adjusted over 24 trials with a total of 6 trials per stage, with subjects having the opportunity to repeat stages if needed. Briefly, the procedure adjusted  $F1$  and  $F2$  in stage 1, starting along a path that passed through the average  $F1$  and  $F2$  frequencies that the speaker of the carrier sentence had used for that word, averaged over the two accents. This meant that the procedure would be likely to approximate a best exemplar quickly. Stage 2 adjusted  $F1$  and  $F2$  along a straight-line path that was orthogonal in the  $F1/F2$  plane to the stage 1 path and included the best exemplar found in stage 1. The procedure then adjusted the more secondary dimension of duration in stage 3 and fine-tuned the best exemplar location in stage 4 by covarying  $F1$ ,  $F2$ , and duration along a single vector.

#### 4. Results

*a. Bud and cud.* Separate repeated-measures ANOVA analyses for  $F1$ ,  $F2$ , and duration tested whether subjects changed their best exemplar locations for *bud* and *cud*. Word (*bud* or *cud*), time (T1, T2, T3, and T4) and sentence context (SSBE or Sheffield English) were coded as within-subject variables. To investigate whether any changes in production were accompanied by changes in perception, the change in accent rating (experiment 1) was entered as a continuous factor. The overall accent rating (experiment 1) was also entered as a continuous factor in order to investigate if the overall differences in spoken accent were accompanied by individual differences in perception.

For  $F1$ , an inspection of the data indicated that some subjects were choosing a more southern vowel (i.e., higher  $F1$ ) for *bud* and *cud* after attending university. However, there was no significant main effect of time or word for  $F1$ ,  $p > 0.05$ . This indicates that subjects were not changing their preferred  $F1$  frequencies over time, and were choosing similar  $F1$  frequencies in both words. There was also no main effect of sentence context,  $p = 0.05$ , indicating that subjects were choosing similar  $F1$  formant frequencies for SSBE and Sheffield English carrier sentences.

However, there was a between-subjects effect of overall accent rating for  $F1$ ,  $F(1, 14) = 6.65$ ,  $p < 0.05$ . This can be

seen in Fig. 3;<sup>5</sup> subjects who were judged to have a more southern accent overall chose more southern best exemplar locations for *bud* and *cud* in both SSBE and Sheffield English sentence contexts. These subjects chose a lower vowel (i.e., higher  $F1$ ), which does not exactly match what native speakers of SSBE or Sheffield English would produce, but which matches what northerners produce when interacting with southerners (Evans and Iverson, 2004; Trudgill, 1986). Subjects who were judged to have a more northern accent overall chose more northern best exemplar locations for *bud* and *cud*. This demonstrates that individual differences in production were accompanied by individual differences in best exemplar locations.

There were no significant main effects or interactions for  $F2$ ,  $p > 0.05$ . This indicates that subjects were not changing their best exemplar locations for *bud* and *cud* in this dimension, and that they were choosing similar  $F2$  locations in both northern English and SSBE carrier sentences.

There were also no significant main effects or interactions for duration,  $p > 0.05$ . This can be seen in Table III: There are few changes in duration over time, and all subjects chose similar durations in both SSBE and Sheffield English carrier sentences.

*b. Could.* Separate repeated measures analyses for  $F1$ ,  $F2$ , and duration tested whether subjects changed their best exemplar locations for *could*, and whether any changes were linked to the change in accent rating or overall accent rating. Time (T1, T2, T3, and T4) and sentence context (SSBE or Sheffield English) were coded as within-subject variables; change in accent rating and overall accent rating were coded as continuous factors.

There was some evidence in the raw data to suggest that subjects chose a different best exemplar in SSBE and Sheffield English sentences; some subjects changed their best exemplar location in SSBE carrier sentences after attending university so that it was similar to the vowel they chose for *bud* and *cud*. However, the analysis demonstrated that there were no significant main effects or interactions of sentence context, time, change in accent rating or overall accent rating for  $F1$ ,  $F2$ , or duration,  $p > 0.05$ . This suggests that there were no reliable differences or changes in subjects' best exemplar locations for *could*.

TABLE III. Average duration (ms) of the vowels in best exemplars for listeners in SSBE and Sheffield English sentences at T1, T2, T3, and T4.

Word	SSBE				Sheffield English			
	T1	T2	T3	T4	T1	T2	T3	T4
bud	72.4	70.6	75.4	76.1	74.0	64.4	72.3	71.9
cud	59.0	70.2	71.8	68.8	59.1	68.3	62.9	65.2
could	73.0	88.0	81.3	64.8	76.2	71.7	78.9	70.2
bath	80.8	95.7	79.6	97.3	70.7	79.3	79.6	80.4
bad	87.0	103.1	95.0	83.3	67.5	86.5	89.2	85.3
bard	157.7	164.3	166.6	157.3	148.8	148.5	157.2	164.1
bawd	155.6	145.7	145.0	155.0	140.8	144.3	145.9	152.8
bead	125.9	135.8	128.1	150.1	118.7	136.3	130.4	154.8
bed	70.3	84.7	74.6	61.8	60.3	67.1	62.7	65.2
bird	125.7	131.4	151.4	145.8	116.2	144.0	134.7	129.8
booed	152.4	132.7	146.9	142.8	123.9	146.9	135.3	144.2

*c. Bath.* Separate repeated-measures analyses for  $F1$ ,  $F2$ , and duration investigated whether subjects changed their best exemplar location for *bath*. Time (T1, T2, T3, and T4) and sentence context (SSBE and Sheffield English) were coded as within-subject variables, and change in accent rating and overall accent rating were coded as continuous factors.

The analysis demonstrated that there were no significant main effects or interactions for  $F1$  and  $F2$ ,  $p > 0.05$ , indicating that subjects did not change their best exemplar location for *bath* over time and that they did not normalize this vowel for accent (i.e., they chose similar best exemplars for *bath* in both sentence contexts). There were also no significant between-subject effects of overall accent rating or change in accent rating,  $p > 0.05$ , demonstrating that the changes in production in *bath* in experiment 1 were not accompanied by changes in best exemplar locations. There were also no significant main effects or interactions for duration,  $p > 0.05$ .

*d. Other vowels.* Separate repeated-measures analyses for  $F1$ ,  $F2$ , and duration tested whether there were any changes for all other words. Word (i.e., all words other than *bud*, *cud*, *could*, and *bath*), sentence context, and time were coded as within-subject variables, and change in accent rating and overall accent rating as continuous factors. There was a main effect of word for  $F1$ ,  $F(1, 14) = 51.91$ ,  $p < 0.001$ , and  $F2$ ,  $F(1, 14) = 13.69$ ,  $p < 0.001$ , demonstrating that different words had different formant frequency values. However, there were no main effects of sentence context or time, and no between-subject effects of change in accent rating or overall accent rating,  $p > 0.05$ , indicating that subjects did not change their best exemplar locations over time for any other words.

For duration there were no significant main effects or interactions,  $p > 0.05$ ; subjects chose similar vowel durations for each target word in both SSBE and Sheffield English carrier sentences at T1, T2, T3, and T4. This was a result of a lack of power in the test; when the change in accent rating and overall accent rating continuous factors were omitted, there was a highly significant main effect of word,  $F(1, 16) = 37.33$ ,  $p < 0.001$ , indicating that subjects chose different durations for different words. This can be seen in Table III; subjects chose vowels with a longer duration for words like

*bard* and *bird*, but vowels with a shorter duration for words like *bad* and *bed*.

## 5. Discussion

Experiment 1 demonstrated that subjects changed their spoken accent, and the aim of experiment 2 was to investigate whether the changes in production were accompanied by changes in perception. The results demonstrated that there were no reliable changes in subjects' best exemplar locations; subjects chose similar vowels at T1, T2, T3, and T4 for *bud*, *cud*, *could*, *bath*, and all other vowels. However, there was a between-subjects link between production and perception. That is, subjects chose similar best exemplar locations for *bud* and *cud* to those that they produced; subjects who produced more southern vowels overall chose more southern best exemplars, and subjects who produced more northern vowels overall chose more northern best exemplars.

Our previous work (Evans and Iverson, 2004) demonstrated that northerners who had been living in London for a minimum of 1 year chose different best exemplar locations for *bud* and *cud* in SSBE and northern English carrier sentences; northerners chose best exemplar locations in their native accent that matched what native speakers of that accent produce (i.e., a high-back vowel), but chose more centralized best exemplars in SSBE sentences that are closer to what SSBE speakers produce. In contrast, northerners living in the north of England, and who were less experienced with SSBE speakers, chose best exemplars for *bud* and *cud* that would be appropriate for northern English speakers in both northern English and SSBE carrier sentences. Based on this evidence, we had expected that after experience of living in a multidialectal environment and interacting with SSBE speakers, subjects in this study would choose different best exemplars in SSBE and northern English sentences. However, there was no evidence that, even after experience of living in a multidialectal environment, subjects, as a group, were learning to normalize. This suggests that rather than being a process that is rapidly learned (e.g., Clarke and Garrett, 2004), the ability to normalize for a non-native accent may require a great deal of experience with a particular accent, perhaps being immersed in that accent for a long period of

time. Alternatively, it is possible that the subjects in this study had not had the same experience with a non-native accent as those in our previous study. Subjects in our previous study were living in London, a multidialectal community where they regularly came into contact with speakers of a wide variety of different accents. Although a university environment is multidialectal, it is possible that the range of accents that subjects encountered was not as wide as that encountered by our subjects in London, and that this affected their ability to normalize. Another possibility is that subjects had learned to normalize but that the testing location affected their responses. Subjects in the present study were tested in Ashby. It is possible that if they had been tested at university, a multidialectal environment where SSBE is regularly used, normalization effects may have been found.

#### IV. EXPERIMENT 3: SENTENCE RECOGNITION IN NOISE

Experiment 3 further investigated whether the changes in production found in experiment 1 were accompanied by changes in perception. The aim of the experiment was to investigate whether changes in production had an effect on a more basic measure of speech perception, the ability to recognize speech in noise.

##### A. Method

###### 1. Subjects

Same as in experiment 1.

###### 2. Stimuli and apparatus

The stimuli were recordings of the BKB sentences (Bench *et al.*, 1979). The BKB sentences are a standardized sentence list that is widely used as an assessment tool in clinical and nonclinical tests of speech perception. There are 21 lists of 16 sentences, and each sentence contains 3, 4, or 5 highly familiar keywords that are identified by the listener, e.g., “The *house* had *nine rooms*,” “They are *buying some bread*” (keywords are in italics). Only the sentences with 3 keywords were used in this experiment. The sentence lists were recorded by two female speakers of different accents, a northern English speaker and a SSBE speaker. The northern English speaker had been born and raised in Ashby-de-la-Zouch. The stimuli for the SSBE speaker were taken from existing recordings made at University College London. All recordings were made in a sound-isolated booth. Stimuli were recorded at a sampling rate of 44.1 kHz and then downsampled to 16 kHz. The speech was mixed with white noise; the noise level was fixed to 71 dBA, and the level of the speech was varied adaptively.

Stimuli were played using a computer sound card. Subjects listened over headphones (Sennheiser HD 414) in a quiet room.

###### 3. Procedure

There were six blocks of testing, three blocks for each speaker. Three blocks were completed on the first day of testing, and three on the second day. In each three-block testing session, the speakers were presented in alternate

blocks (i.e., subjects did not hear the same speaker in consecutive blocks), with the order of presentation counterbalanced across subjects. The maximum number of trials per block is 20 and so, at each testing session, subjects were randomly assigned a set of 60 sentences for each speaker (120 sentences in total). List 1 was used for familiarization. Subjects were assigned a different set of sentences for each speaker at T1, T2, T3, and T4, so that for each speaker, each sentence was only identified once.

A modified Levitt procedure (Baker and Rosen, 2001) was used to find subjects’ noise thresholds. The procedure started with an easy stimulus with an SNR of +10 dB (i.e., above threshold) in order to enable subjects to tune in to the talker. The SNR then decreased in 8 dB steps after each correct response (i.e., becomes more difficult), until the first reversal (i.e., an incorrect response). After the first reversal the SNR changed in steps of 2 dB for a further eight reversals.

A one-up/one-down procedure was used, with sentences scored as correct when they repeated all three keywords aloud to the researcher. If subjects only repeated one or none of the keywords, then the sentence was scored as incorrect. If subjects repeated two keywords then the SNR remained the same and this was not counted as a reversal. The procedure thus converged on a 66.6% identification level. The test terminated when subjects completed eight reversals or after 20 stimuli had been presented.

##### 4. Results

A repeated-measures analysis tested whether the observed differences in performance were linked to the change in accent rating and the overall accent rating. Speaker (northern or SSBE) and time (T1, T2, T3, and T4) were coded as within-subject variables, and change in accent rating and overall accent rating were coded as continuous factors.

There was a main effect of speaker  $F(1, 15)=136.61$ ,  $p<0.001$ ; all subjects performed better with the SSBE speaker (see Fig. 4<sup>3</sup>). The northern speaker was from Ashby-de-la-Zouch, Leicestershire, and one of the features of the local accent is that it has a flat intonation contour. This difference, combined with the fact that subjects were highly familiar with SSBE through the media (Foulkes and Docherty, 1999), may have resulted in the SSBE speaker being more intelligible in noise than the northern speaker. There was no main effect of time and no interaction of time and speaker,  $p>0.05$ , indicating that subjects did not perform better with SSBE after the experience of living in a multidialectal community.

There was a significant interaction between speaker and overall accent rating,  $F(1, 18)=10.78$ ,  $p<0.01$ . As displayed in Fig. 4, subjects who were judged to have a more southern accent performed better with SSBE speech than those who had a more northern accent; all subjects performed similarly with northern speech. However, there were no interactions between the change in accent rating and speaker,  $p>0.05$ , demonstrating that subjects who were judged to have changed their accent to sound more southern had no advantage over subjects who had not changed their accent.

## 5. Discussion

As in experiment 2, there was a between-subjects link between individuals' production and perception. Subjects who produced more southern vowels overall were better able to identify SSBE speech in noise than those who produced more northern vowels overall. This indicates that individual differences in speech production were accompanied by differences in perceptual processing. However, there was no evidence to suggest that subjects' ability to recognize SSBE speech in noise changed over time.

## V. GENERAL DISCUSSION

The results of this study demonstrated that subjects changed their accent as a result of attending university. For example, subjects changed their production of *bud*, *cud*, and *could* so that they produced them with a more fronted (higher  $F2$ ) and lower (higher  $F1$ ) vowel at T4, which is closer to how southerners produce the vowel in words like *bud* and *cud*. Although subjects changed their production, there were no reliable changes in perception; there were no changes in subjects' best exemplar locations in experiment 2, and no change in subjects' ability to recognize SSBE speech in noise in experiment 3. However, there was a between-subjects link between individuals' overall production and perception. That is, subjects who produced more southern vowels chose more southern best exemplar locations for *bud* and *cud*, and subjects who produced more northern vowels chose more northern best exemplars for *bud* and *cud*. Likewise, subjects who produced more southern vowels performed better with SSBE speech in noise than did those who produced more northern vowels.

Even though subjects were selected so that they were from a homogeneous background (i.e., they had all been exposed to the same regional accent until leaving for university and had attended local schools since age 5 years), it is not surprising that we found overall differences in production. Previous research has shown that within a homogeneous community, structured differences in production emerge as a result of sociolinguistic influences. For example, Eckert (1989, 2000) found that high-school students in Detroit used phonetic variation to define their social groups (e.g., "jocks" and "burnouts") and construct their own identity. The idiolectal variation in the present study may have likewise emerged as markers of how each individual fitted within their community. The production differences may thus have reflected how these individuals chose to present themselves to the world, rather than being reflective of underlying differences in perceptual experience (i.e., the phonetic content of the speech that they had heard during their lifetime).

Given that these individual differences in production are likely due to sociolinguistic influences, it was surprising that they were directly linked to differences in perception. That is, even though spoken accent is a marker of social identity, perception is essentially private and it would be adaptive for individuals to be able to understand as wide a range of speakers as possible. Our subjects likely had very similar perceptual experiences when growing up, but the idiolectal variation in production had influences on both subjective rat-

ings (i.e., best exemplar locations in experiment 2) and speech-in-noise recognition. It thus appears that the socio-phonetically driven differences found in production had an impact on perceptual processing, suggesting that there is a strong perception-production link.

Given this link between perception and production in terms of overall individual differences (i.e., averaged across time), it is unclear why perception and production were not linked in terms of changes over time. For example, individuals who changed their spoken accent to sound more southern did not improve in their ability to recognize SSBE speech in noise. It is notable, however, that the changes in accents were relatively small (e.g., the changes in accent ratings were smaller than the overall individual differences). In order to understand speech from a range of talkers (e.g., different accents and idiolects), listeners' category representations presumably need to be tolerant to a range of acoustic variation. It is possible that the changes in production fell within this range of tolerance. That is, listeners may have had some degree of freedom to change their spoken accent without making corresponding changes to their underlying category representations or their perceptual processing.

Although we found no changes in perception over time, previous studies of L1 perception have found that subjects are able to rapidly adapt to foreign accented speech (e.g., Clarke and Garrett, 2004). One difference between the present study and previous work on perceptual adaptation (e.g., Clarke and Garrett, 2004; Nygaard *et al.*, 2005) is that our study investigated adaptation to a regional accent, not to a particular talker or group of talkers. For example, Nygaard *et al.* (2005) found that American English listeners were most accurate at transcribing items spoken by Spanish-accented talkers after a short amount of training with the same speakers; listeners who were exposed to Spanish-accented speech produced by different talkers were less accurate. In our study, subjects heard a range of accents while attending university, but were not regularly exposed to the specific talkers that they heard in our experiments. It is possible that we did not see the same perceptual adaptation effects that have been found previously (e.g., Clarke and Garrett, 2004; Nygaard *et al.*, 2005) because such changes may be talker-specific and not easily generalizable to other speakers of that accent. Adaptation to accent-general patterns of variation may thus require more long-term experience with speakers of that accent.

In summary, the present study demonstrated that subjects were able to change their spoken accent at a late stage in their language development, young adulthood, although the changes in production were not accompanied by changes in perception. There was evidence for a between-subjects link between production and perception; subjects who produced more southern vowels chose more southern best exemplar locations and were better at identifying SSBE speech in noise. Such individual differences in production and perception were particularly remarkable because subjects were selected so that they had a relatively homogeneous background. These findings suggest that individual differences in production are accompanied by fine-grained differences in

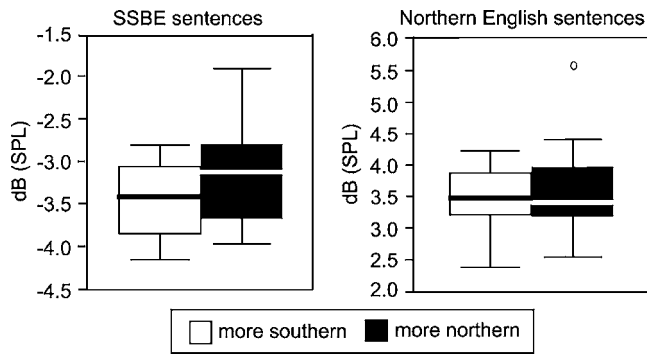


FIG. 4. Boxplots of subjects' average SNR threshold for SSBE and northern English sentences. Subjects are grouped according to their overall accent rating; the open boxes represent subjects who were judged to sound more southern overall ( $N=13$ ) and the filled boxes represent subjects who were judged to sound more northern overall ( $N=10$ ). Subjects who had a more southern accent overall performed better with SSBE speech than subjects who had a more northern accent overall.

perceptual representations, and that these small individual differences have an impact on speech processing.

## ACKNOWLEDGMENTS

This research was supported by an EPSRC Doctoral Training Award to the first author. We are grateful to Richard Dewire, Natasha Fish, Graham Fisher, and Graham Hobbs for help with recordings, and to Jackie and Alexandra Evans for help with subject recruitment.

<sup>1</sup>The term SSBE instead of Received Pronunciation (RP) is used here to denote the standard, prestige British English accent. This is to avoid confusion with a more narrow definition of RP (see Wells, 1982 for a full discussion of the definition and scope of RP, as well as other varieties of RP).

<sup>2</sup>Of the 27 subjects tested, 26 were born and raised in the East Midlands. One subject was born in Ireland but his mother, his primary caregiver, was originally from Ashby and they returned to Ashby when he was 5 years old.

<sup>3</sup>All subjects attended university in England. Consequently, they did not come into regular contact with speakers with Scottish, Welsh, and Irish accents.

<sup>4</sup>As a result of the fact that not all subjects could be contacted at T4, the change in accent rating was calculated by subtracting each subject's accent rating at the final testing session (T3 or T4) from the accent rating at T1.

<sup>5</sup>For the purposes of Figs. 3 and 4, subjects were divided into two groups based on the accent ratings (experiment 1). Those who had an overall accent rating lower than or equal to the median overall accent rating of 5.75 were assigned to the "more southern" group ( $N=13$ ). Those who had an accent rating greater than the median overall accent rating were assigned to the "more northern" group ( $N=10$ ).

Adank, P. (2003). "Vowel normalization: A perceptual-acoustic study of Dutch vowels," Unpublished Ph.D. dissertation, University of Nijmegen, The Netherlands.

Adank, P., Smits, R., and van Hout, R. (2004). "A comparison of vowel normalization procedures for language variation research," *J. Acoust. Soc. Am.* **116**, 3099–3107.

Ainsworth, W. A., and Paliwal, K. K. (1984). "Correlation between the production and perception of the English glides /w, r, l, j/," *J. Phonetics* **12**, 237–243.

Bailey, P. J., and Haggard, M. P. (1973). "Perception and production: Some correlations on voicing of an initial stop," *Lang Speech* **16**, 189–195.

Bailey, P. J., and Haggard, M. P. (1980). "Perception-production relations in the voicing contrast for initial stops in 3-year-olds," *Phonetica* **37**, 377–396.

Baker, R. J., and Rosen, S. (2001). "Evaluation of maximum-likelihood

threshold estimation with tone-in-noise masking," *Br. J. Audiol.* **35**, 43–52.

Bell-Berti, F., Raphael, L. J., Pisoni, D. B., and Sawusch, J. R. (1979). "Some relationships between speech production and perception," *Phonetica* **36**, 373–383.

Bench, J., Kowal, A., and Bamford, J. (1979). "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children," *Br. J. Audiol.* **13**, 108–112.

Boersma, P., and Weenink, D. (2004). PRAAT: Doing phonetics by computer (Computer software) (University of Amsterdam, The Netherlands).

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. (1997). "Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production," *J. Acoust. Soc. Am.* **101**, 2299–2310.

Clarke, C. M., and Garrett, M. F. (2004). "Rapid adaptation to foreign-accented English," *J. Acoust. Soc. Am.* **116**, 3647–3658.

Clarke, C. M., and Luce, P. A. (2005). "Perceptual adaptation to speaker characteristics: VOT boundaries in stop voicing categorization," in *Proceedings of the ISCA Workshop on Plasticity in Speech Perception*, London, U.K.

Eckert, P. (1989). *Jocks and Burnouts: Social Categories and Identity in the High School* (Teachers College Press, New York).

Eckert, P. (2000). *Linguistic Variation as Social Practice* (Blackwell, Oxford).

Eisner, F., and McQueen, J. M. (2005). "The specificity of perceptual learning in speech processing," *Percept. Psychophys.* **67**, 224–238.

Evans, B. G. (2005). "Plasticity in Speech Perception and Production: A study of accent change in young adults," Unpublished Ph.D. thesis, University College London: University of London, U.K.

Evans, B. G., and Iverson, P. (2004). "Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences," *J. Acoust. Soc. Am.* **115**, 352–361.

Flege, J. E., Bohn, O.-S., and Jang, S. (1997). "The production and perception of English vowels by native speakers of German, Korean, Mandarin, and Spanish," *J. Phonetics* **25**, 437–470.

Flege, J. E., Mackay, I. R. A., and Meador, D. (1999). "Native Italian speakers' perception and production of English vowels," *J. Acoust. Soc. Am.* **106**, 2973–2987.

Foulkes, P., and Docherty, G. J. (1999). "Urban Voices—Overview," in *Urban Voices: Accent Studies in the British Isles*, edited by P. Foulkes and G. J. Docherty (Arnold, London).

Fowler, C. A. (1981). "Production and perception of coarticulation among stressed and unstressed vowels," *J. Speech Hear. Res.* **24**, 127–139.

Fowler, C. A. (1986). "An event approach to the study of speech perception from a direct-realist perspective," *J. Phonetics* **14**, 3–28.

Frieda, E. M., Walley, A. C., Flege, J. E., and Sloane, M. E. (2000). "Adults perception and production of the English vowel /i/," *J. Speech Lang. Hear. Res.* **43**, 129–143.

Huckvale M. (2003). *Speech Filing System* (computer software) (University College, London, UK).

Iverson, P., and Evans, B. G. (2003). "A goodness optimization method for investigating phonetic categorization," Paper presented at the International Congress of Phonetic Sciences (ICPhS), Barcelona, Spain.

Iverson, P., Smith, C. A., and Evans, B. G. (2005). "Vowel recognition via cochlear implants and noise vocoders: The effects of formant movement and duration," *J. Acoust. Soc. Am.* **120**, 3998–4006.

Kerswill, P., and Williams, A. (2000). "Creating a new town koine: Children and language change in Milton Keynes," *Language in Society* **29**, 65–115.

Klatt D. H., and Klatt L. C. (1990). "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *J. Acoust. Soc. Am.* **87**, 820–857.

Liberman, A. M., and Mattingly, I. G. (1985). "The motor theory of speech perception revised," *Cognition* **21**, 1–36.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). "Perception of the speech code," *Psychol. Rev.* **74**, 431–461.

Milroy, L. (1987a). *Observing and Analysing Natural Language* (Blackwell, Oxford).

Milroy, L. (1987b). *Language and Social Networks*, 2nd ed. (Blackwell, Oxford).

- Munro, M. J., Derwing, T. M., and Flege, J. E. (1999). "Canadians in Alabama: A perceptual study of dialect acquisition in adults." *J. Phonetics* **27**, 385–403.
- Nearey, T. M. (1978). "Phonetic Feature Systems for Vowels" (Indiana University Linguistics Club, Bloomington, IN).
- Newman, R. S. (2003). "Using links between speech perception and speech production to evaluate different acoustic metrics: A preliminary report," *J. Acoust. Soc. Am.* **113**, 2850–2860.
- Nygaard, L. C., and Pisoni, D. B. (1998). "Talker-specific learning in speech perception," *Percept. Psychophys.* **60**, 355–376.
- Nygaard, L. C., Sidas, S., and Duke, J. (2005). "Perceptual learning of accented speech," in *Proceedings of the ISCA Workshop on Plasticity in Speech Perception*, London, U.K.
- Sankoff, G. (2004). "Adolescents, young adults and the critical period: Two case studies from 'Seven Up,'" in *Sociolinguistic Variation*, edited by C. Fought (Oxford University Press, Oxford).
- Trudgill, P. (1986). *Dialects in Contact* (Blackwell, Oxford).
- Wells, J. C. (1982). *Accents of English: The British Isles* (Cambridge University Press, Cambridge), Vol. 2.