Journal of Phonetics 39 (2011) 121–131

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

The physiological, acoustic, and perceptual basis of high back vowel fronting: Evidence from German tense and lax vowels

Jonathan Harrington *, Phil Hoole, Felicitas Kleber, Ulrich Reubold

Institute of Phonetics and Speech Processing, Ludwig-Maximilians University of Munich, Schellingstr. 3, D 80799 München, Germany

ARTICLE INFO

Article history:
Received 20 June 2010
Received in revised form
10 December 2010
Accepted 13 December 2010
Available online 17 January 2011

ABSTRACT

High back rounded vowels are prone synchronically to fronting in a fronting context and diachronically they are more likely to front than high front vowels are to retract. In order to shed light on the reasons for this back–front asymmetry, tense and lax vowels produced in three place of articulation contexts by seven first-language speakers of German at two speech rates were analysed physiologically, acoustically, and perceptually. An articulographic analysis showed greater magnitudes and peak velocities of horizontal tongue dorsum movement in CV transitions for /u:, o:/ than for /e:, i:/. A second experiment showed that the difference between tense and lax vowels in the tongue dorsum's horizontal position was greater for back than front vowels. A third experiment showed /u:, ʊ/ were more likely to encroach on the /y:, ʏ/ spaces than the other way round for measurements based both on the horizontal tongue dorsum position and on spectral slope; a similar pattern of results emerged in a forced-choice perception experiment. The general conclusion is that high back vowels that are as peripheral as those in German have a high articulatory cost which may explain both the diachronic tendency for back vowels to front and why the absence of a high back vowel often contributes to asymmetric vowel distributions in the world's languages.

1. Introduction

Sound changes that are found in many languages and that have some relationship to the way that the sounds of the world's languages are distributed can often be explained using physiological principles of speech production and how these are linked to the acoustic signal and perception (Ohala, 1993). Consider in this regard the fronting of back vowels and in particular the diachronic development of fronted variants of back vowels. This type of sound change has been incorporated as part of one of Labov's (1994) general principles of chain shifting and it is reported to occur in languages as structurally diverse as Swedish (and East Norwegian), Proto-Southern Yiddish, Albanian, and Akha, a Lolo-Burmese language (see Labov, 1994, for further details). In addition, the fronting of /u/ (lexical set, GOOSE) and of its lax counterpart /ʊ/ (FOOT) have been shown to be sound changes in progress in the last 50 years in Australian English (Cox, 1999; Cox & Palethorpe, 2001), in many varieties of North American English (Fridland, 2008; Labov, Ash, & Boberg, 2006), and in Standard Southern British (Harrington, Kleber, & Reubold, 2008; Hawkins & Midgley, 2005). Moreover, there is an asymmetry in the extent to which high vowels are prone to diachronic change along a front–back phonetic dimension: thus, although the diachronic retraction of high /ɪ/ front vowels is by no means unattested (e.g. the shift in the last 50 years of /ɪ/ (HID) to a more central position in New Zealand English: Maclagan & Hay, 2007), this type of change seems to be much less common than the diachronic fronting of back vowels. There may be some rather more limited evidence for a corresponding asymmetry in the pattern of distribution of vowels in the world's languages. Thus /u/ occurs in 28 fewer of the 451 languages of the UCLA Phonological Segment Inventory Database (UPSID) than does /i/ and it is also slightly less common than a back mid-vowel (Maddieson, 1984). Moreover, when a language's vowel system is defective from the point of view of not having a symmetrical distribution of vowels in the vowel space, then this is most likely to be occasioned by the absence of /u/ (Maddieson, 1984). Compatibly, Schwartz, Boë, Vallée, and Abry (1997) showed in their study of UPSID that when languages have a non-symmetrical distribution of vowels along a front–back dimension, they were more likely to be left (i.e. with a greater number of front vowels) than right dominant.

Of course, /u/ is nevertheless the third most frequent vowel of the UPSID database and there are phonetically grounded reasons based on quantal (Stevens, 1989, 2003) and maximal dispersion (Liljencrants & Lindblom, 1972; Lindblom, 2003) theory for why this should be so. However, we still lack empirical evidence from most languages of the world on the extent to which /u/ really is produced as a peripheral high *back* vowel, i.e. with a quality close to cardinal vowel 8. Thus, it may be that the vowel that is auditorily labelled as /u/ may be more fronted than in the minority of languages such as French, German or Swedish that have a phonemic contrast between high front rounded /y, ʏ/ and high back rounded /u, ʊ/ vowels.

* Corresponding author. Tel.: +49 89 2180 2758.
E-mail address: jmh@phonetik.uni-muenchen.de (J. Harrington).

0095-4470/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2010.12.006

Certainly, empirical studies of languages with a small number of oppositions such as those of Australia show a much less peripheral distribution than implied by the point vowels /i, u, a/ (Butcher & Anderson, 2008).

The different coarticulatory effects of place of articulation on high vowels could provide some explanation for the greater propensity for back vowels to front than for front vowels to retract diachronically. Thus a more advanced constriction for /u/ is likely in a coronal context such as /tut/ because the tongue body is dragged forward resulting acoustically in a raised second formant frequency (Flemming, 2001, 2003; Öhman, 1966). An explanation for diachronic /u/-fronting could also be motivated acoustically and perceptually. If the release of a consonantal /t/ overlaps substantially with the onset of /u/, then high frequency energy due to the /t/ release would occur simultaneously with the low frequency resonant energy due to the back vowel. The combination of the two would cause the overall spectral centre of gravity to shift upwards in frequency towards that of a high front vowel (potentially resulting, following Ohala's (1993) model, in a perceptual reinterpretation by the listener of /u/ as a fronted vowel). On the other hand, the place of articulation that is most likely to cause /i/ to shift acoustically towards /u/ due to F2-lowering is labial which is less likely to interfere with the palatal constriction, given that the lips and the tongue are produced with independent sets of articulators: thus, the perturbation of the vowel target due to context may be less likely in /pip/ than it is in /tut/. And while either the retracted tongue body of uvular or the retracted tongue tip of retroflex places of articulation (Flemming, 2003) may induce considerable F2-undershoot in high front vowels, these places of articulation may be rarer in the world's languages than coronals that have a marked fronting effect on high back vowels. In addition, given that the tongue body of velars tends to shift with the phonetic backness of the vowel (e.g. being advanced to post-palatal in English *key* and retracted to post-velar in *court*), then velars may be no more likely to produce tongue-dorsum undershoot in /i/ than in /u/. Thus the different coarticulatory effects of consonantal place of articulation may provide some phonetic basis for the greater tendency for high back vowels to front than for high front vowels to retract diachronically. Finally, Alfonso and Baer (1982) have shown in their physiological study of American English vowels that the tongue's movement towards the target begins earlier in back than in front vowels which they attribute to the inherent sluggishness in the back-raising movement of the tongue dorsum: thus the potentially greater time taken for the tongue manoeuvre to be completed in high back vowels may contribute to the greater likelihood for their targets to be undershot than in high front vowels, especially at faster rates or in less formal speaking styles. Compatibly, Tabain and Perrier (2007) show that when /u/ occurs at the beginning of weak prosodic boundaries, F2 is higher than when /u/ is domain-initial at strong boundaries. Since they showed that the lip constriction was tighter at weak than at strong boundaries, and given that a closer lip approximation causes formants to lower, then the raised F2 that they also found at weak boundaries must have been due to a considerable degree of tongue fronting (see also Tabain (2008) for a similar finding and a comparison with Australian English vowels).

The task in this paper is to begin to explore why high back vowels might be so prone to fronting or more so than high front vowels seem to be to diachronic retraction. For this purpose, we investigated vowels in German both because tense /iː/, /yː/, and /uː/ are phonetically peripheral, i.e. very close to cardinal vowels in quality, and secondly because German has a tense–lax contrast that provides another index for comparing the peripherality of high front and back vowels. We first tested whether there are properties of the tongue dynamics for high back vowels that might favour fronting. One possibility is that the production of /uː/ in an acoustically stable region with a low F1 and F2 may require the tongue dorsum to be displaced from a central, neutral configuration to a greater extent than it is in the other direction for /iː/. If this is so, then high back vowels might be especially prone to fronting

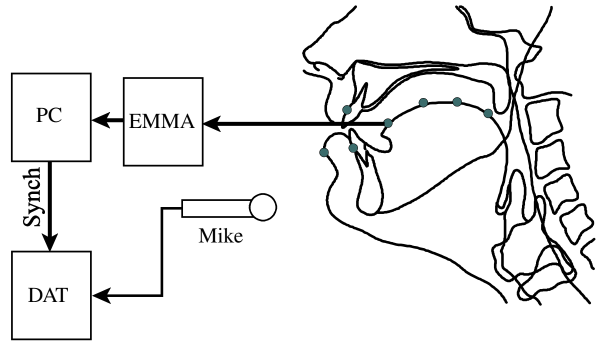

Fig. 1. The experimental set-up for recording the EMMA data showing a cross-sectional view of the vocal organs with the approximate position of the sensors. The articulatory data in this study was taken from the fourth furthest back tongue sensor positioned closest to the uvula.

at faster rates of speech and the articulatory distance between the tense and lax categories could be greater in high back than in high front vowels. These are two of the main issues that were investigated with a physiological database of German vowels in Experiments I and II. Moreover, if there is a phonetic basis to diachronic high back vowel fronting, then the confusions between high vowels due to processes of reduction (such as context or fast speech) should be asymmetric: more specifically, we tested in Experiment III whether /uː/ was more likely to encroach articulatorily and acoustically on the /yː/ space than the other way round. For the same reason, we predicted an analogous bias in listener confusions: that is for speaking styles that are likely to produce some form of hypoarticulation, there may be a perceptual asymmetry such that high back vowels are more likely to be misperceived as high front vowels than in the opposite direction. This was tested in a listening experiment with stimuli derived from the same corpus that was analysed physiologically and acoustically.

2. Experiment I: Tongue movement in tense front and back vowels

2.1. Method

The data analysed in these experiments were taken from the same corpus of vowels produced by seven first-language German speakers described in Hoole (1999) and Hoole and Mooshammer (2002). The data were acquired using electromagnetic midsagittal articulometry (EMMA; AG100 Carstens Medizinelektronik Göttingen) with four sensors attached to the surface of the tongue (Fig. 1), and one each on the jaw and lower lip (plus one sensor each on the upper incisors and bridge of nose to compensate for head movements). For the present paper, we analysed the horizontal (TD_X) and the vertical (TD_Y) positions of the sensor positioned on the midline and as far back on the tongue (up to 6 cm back from the tongue tip) as the subject could tolerate (Fig. 1).

The data were rotated so that they were parallel to the occlusal plane that was estimated by having a subject bite onto a bite-plate. The articulatory data were sampled at 250 Hz and low-pass filtered with a FIR filter (cut-off frequency 35 Hz). The synchronised acoustic waveform was digitised at 16 kHz. These procedures were carried out in Matlab and the output stored in self-documented Matlab files. All of the data were converted into an Emu compatible format and analysed in the R programming language (Harrington, 2010).

Seven speakers who spoke a variety of Standard German as their first language with only minor regional colouring produced symmetrical CVC sequences for C = /p, t, k/ and all the German monophthongs embedded in the target word and carrier phrase *ich habe* /gəCVCə/ *gesagt* (literally *I have* /gəCVCə/ *said*). With the exception of /ɛː/,

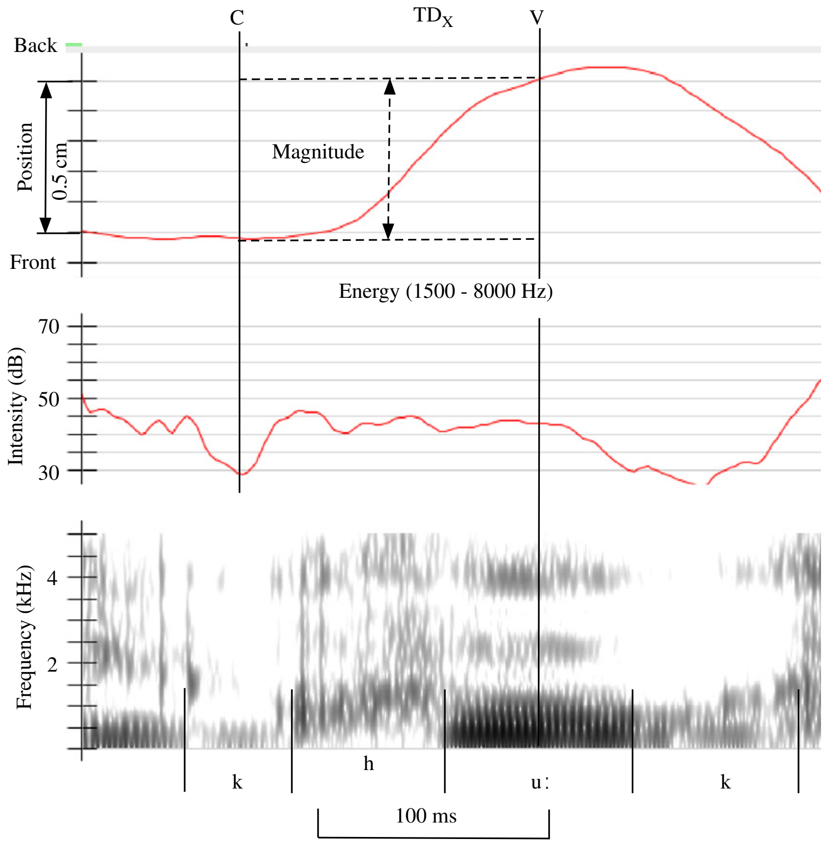

Fig. 2. The horizontal movement of the sensor positioned furthest back on the tongue (top trajectory, TD_X), intensity calculated after high-pass filtering the signal leaving energy above 1500 Hz, and a wideband spectrogram of the target syllable /ku:k/ with superimposed acoustic phonetic segment boundaries. C marks the energy minimum in the closure; V is at the temporal midpoint between the acoustic onset and offset of voicing for the vowel. The magnitude of the TD_X movement in the CV transition was defined as the difference in the positions at these C and V time points.

German vowels that occur in rhythmically strong syllables can be grouped into four front (/i:, ɪ/, /y:, ʏ/ /e:, ε/, /ø:, œ/), one central (/ɐ:, ɐ/)[1] and two back (/o:, ɔ/, /u:, ʊ/) tense–lax pairs. The phonetic qualities of the vowels denoted by these transcriptions correspond approximately to those of the international phonetic alphabet (thus /e:/ is phonetically close to a long cardinal vowel 2, etc.).

The carrier phrase was produced with a nuclear accent on the target non-word /gəCVCə/ and in most cases with a falling intonational melody. The carrier phrases were repeated five times, randomised separately for each subject, and presented individually on a computer monitor in the corresponding orthography (e.g. for /pɐ:p/ *ich habe gepaape gesagt*, for /kɪk/ *ich habe gekikke gesagt*, etc.). The entire corpus was produced twice on separate occasions, once at a normal speech, then once at a fast speech rate. In order to ensure some degree of consistency of rate across speakers, a pre-test was conducted in which the speakers were asked to read samples of the material at various self-selected rates. These were subsequently analysed for vowel duration and the rate that was chosen – i.e. the rate at which the subject was asked to speak in the fast rate – was the one in which the vowel durations of the tense vowels most closely matched those of their lax counterparts at the slow rate. At regular intervals during the sessions, examples of this chosen rate were played back to the subject to remind the speaker of the target rate that was to be achieved. The total number of presented sentences was 3 (places of articulation) × 15 (vowels) × 2 (rates) × 5 (repetitions) × 7 (speakers) = 3150. In addition, the subjects produced separately three steady-state versions each of all tense vowels except /ε:/.

The speech materials were segmented manually to mark the onset of aspiration of the first stop and the offset of the closure of the second stop in each target word. The acoustic onset of voicing was calculated algorithmically using a measure based on the probability of voicing. We also calculated the point of minimum energy in the closure of the initial stop firstly by filtering the acoustic signal so that only energy above 1500 Hz remained, then by calculating the dB-RMS energy with a window size of 20 ms and a frame shift of 2 ms in this high-pass filtered data, and finally calculating within an interval of 100 ms preceding the aspiration onset the point at which the energy minimum occurred in the signal (Fig. 2). Given the inherent difficulty of finding reliable turning points in the articulatory data for all speakers, vowel contexts and both rates, the tongue positions (of the most back sensor, i.e. of TD_X and of TD_Y as defined above) were instead extracted at two acoustic landmarks only: specifically at the time of the energy minimum in the closure, as defined above, and at the temporal midpoint of the vowel (halfway between the acoustic onset and offset of periodic vowel voicing). We then defined the gestural magnitude as the difference between the horizontal positions of the tongue at these time points (Fig. 2). The peak velocity of TD_X was calculated as the velocity maximum (in tongue backing movements) or minimum (in tongue fronting movements) within this same interval.

2.2. Results

The aim in this analysis was to test whether the TD_X magnitude and peak velocity for /u:, o:/ at the normal rate were more extreme than for other vowels, and in particular than for /i:, e:/.

[1] In broad phonemic transcriptions of Standard German, these vowels are conventionally transcribed with a Cardinal Vowel 4 symbol. However, we have not used CV4 for these vowels to emphasise (see Fig. 5) that they are phonetically *central* in Standard German.

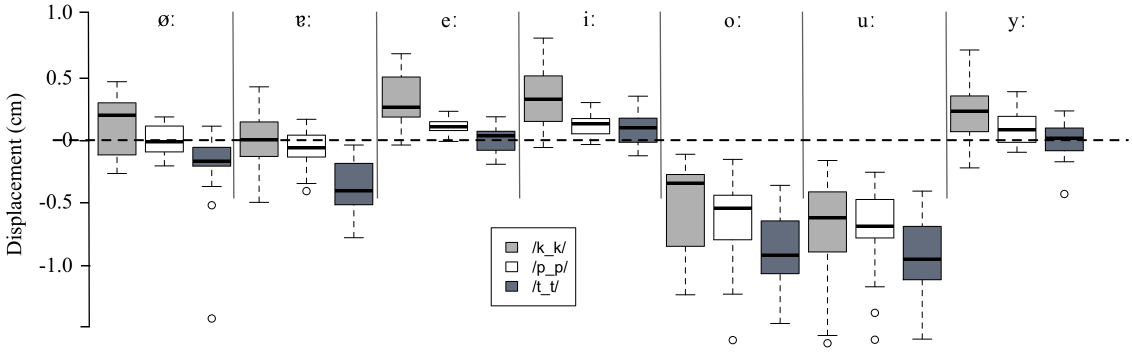

Fig. 3. The magnitude of displacement in the CV transitions (see Fig. 2) pooled across all seven speakers for tense vowels at the normal rate shown separately by context. Positive/negative values are for a forward/backward TD_X movement from C to V.

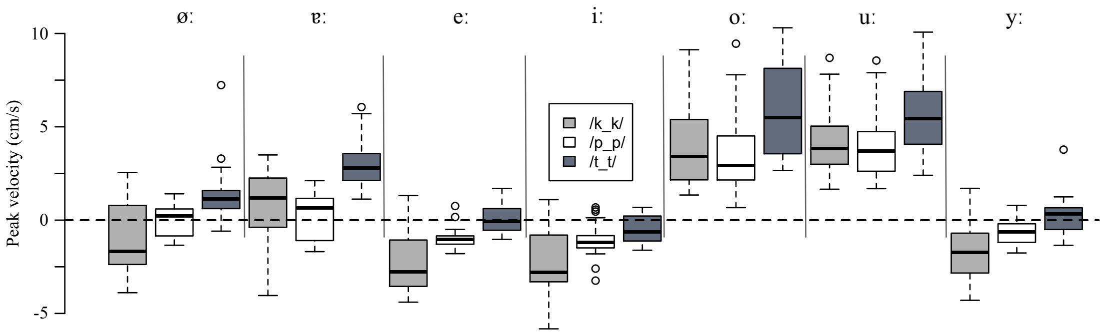

Fig. 4. The peak velocity calculated over the same CV interval as for the data in Fig. 3.

The results in Fig. 3 show that the TD_X magnitude (calculated between the energy minimum of the initial stop closure and the acoustic vowel midpoint—see Fig. 2) over the CV interval was greatest for the back vowels. The much greater TD_X range shown by the boxplots (pooled across all 7 subjects) for back than front vowels suggests that back vowels are subject to a high degree of variation on this articulatory parameter, possibly because the tongue-dorsum movements are so extreme and therefore more prone than for other vowels to undershoot. The boxplots in Fig. 4 show that /o:, u:/ also had the highest peak velocities over this CV interval.

We ran two repeated measures ANOVAs to test for differences between the categories *front* including /i:, e:/ and *back* including /u:, o:/. The independent variables were Vowel (two levels: front vs. back, as defined above), Rate (two levels: normal vs. fast) and Place (three levels: /p, t, k/). The results with the absolute TD_X magnitude as the dependent variable showed a significant effect for Vowel ($F(1, 6)=74.4$; $p<0.001$) and no significant effects either for any of the other variables or for their interaction with Vowel. For the absolute peak velocity as the dependent variable, there was once again a significant effect for Vowel ($F(1, 6)=99.9$; $p<0.001$), a significant effect for Rate ($F(1,6)=11.0$; $p<0.05$), and a significant Vowel × Rate interaction ($F(1, 6)=21.0$; $p<0.01$). *Post-hoc* Bonferroni *t*-tests showed that there were significant differences between front and back vowels both at the slow ($t=5.5$; $p_{adj}<0.001$; df$=20$) and at the fast ($t=9.2$; $p_{adj}<0.001$; df$=20$) rates. There were also significant differences between the slow and fast rates for back vowels ($t=5.3$; $p_{adj}<0.001$; df$=20$) but not for front vowels.

2.3. Discussion

Both the absolute magnitude and peak velocity of TD_X over the CV transition were greater for /u:, o:/ than for /i:, e:/: thus for back vowels, the tongue dorsum moved through a greater distance and it did so at a faster rate over this interval. For these reasons, TD_X may be prone to undershoot under conditions in which vowel duration is compromised. Although there was no evidence from these data of gestural magnitude reduction in /u:, o:/ compared with /i:, e:/, there was a significant increase in articulatory velocity from the normal to the faster rate for the back but not the front vowels: thus, to the extent that peak velocity is an index of articulatory effort (Moon & Lindblom, 1994; Nelson, 1983), it can be tentatively concluded that the effort to maintain the articulatory distance through which the tongue dorsum travels in CV transitions is greater for /u:, o:/ than for other vowels at the faster rate.

In the next experiment, we studied the relationship between tense vowels and their lax counterparts. Based on the results so far, we predicted that the articulatory distance on the horizontal dimension between tense and lax vowels would be greater for back, than for front, vowels.

3. Experiment II: Tense and lax vowels

3.1. Method

In order to gain an overall impression of the relative distribution of the vowels in an articulatory tongue dorsum space, we expressed their position in terms of a number of standard deviations from the speaker mean. The speaker mean was quantified as follows. We calculated for each utterance the mean value, m_{Xi}, of TD_X across all of the frames separately between the start and end times of the *i*th utterance produced by the speaker. The global speaker mean, m_X, and the global speaker standard deviation, s_X, were defined as the mean and standard deviation of these utterance means in (1) and (2), respectively

$$m_X = \frac{\sum_1^i m_{Xi}}{i} \tag{1}$$

$$s_X = \sqrt{\frac{\sum_1^i m_{Xi}^2}{i} - m_X^2} \tag{2}$$

The z-score transformed value, z_X, of a given TD_X frame of data, f_X, was obtained using the following equation:

$$z_X = \frac{f_X - m_X}{s_X} \tag{3}$$

Exactly analogous calculations were obtained for the vertical position of the tongue dorsum to obtain z-score transformed values, z_Y. Thus the point [0,0] in this z-transformed space represents a global average for a given speaker of the horizontal and vertical positions of the tongue dorsum; positive and negative values are the number of standard deviations from this mean. The utterances 1, …, i in (1) and (2) from which this global speaker space was calculated included only those containing either front unrounded vowels /iː, eː, ɪ, ɛ/, back vowels /uː, ʊ, oː, ɔ/ or the open central vowels /ɐː, ɐ/: that is, we excluded utterances containing the front rounded vowels from these calculations to ensure an equal balance between front and back (since otherwise the mean would be skewed towards the front of the vowel space just because there would have been more front than back vowels).

The statistical analyses were not calculated in this normalised space, but on the raw TD_X data. The aim here was to test whether the tense–lax pairs in back vowels were further apart from each other on the horizontal movement of the tongue dorsum than for front vowels. We did this for TD_X values extracted at the acoustic temporal midpoint of the vowels by calculating separately for each tense–lax pair and separately for each speaker two parameters: d_L, the absolute distance of each lax vowel (V_L) to the tense mean (m_T); and d_T, the absolute distance of each tense vowel (V_T) to the lax mean (m_L):

$$d_L = |V_L - m_T| \tag{4}$$

$$d_T = |V_T - m_L| \tag{5}$$

Thus for the /uː, ʊ/ tense–lax pair, (4) and (5) provided a distribution of how far the /ʊ/ tokens were to the mean of /uː/ and how far the /uː/ tokens were to the mean of /ʊ/ (based on TD_X values at the vowels' acoustic temporal midpoint).

3.2. Results

The mean positions of the vowels in the z-transformed $TD_X \times TD_Y$ space pooled across all speakers is shown in Fig. 5. Recall that [0,0] represents the global mean calculated separately per speaker from the means obtained between the start and end time of each carrier phrase. Note that what are typically labelled as open mid-vowels /ɐː, ɐ/ (e.g. *lahm/Lamm*: *lame/lamb*) in German were appropriately quite close to $z_X=0$. By contrast, $z_Y=0$ was somewhat above the central vowel position presumably because the calculation to obtain this z-transformed space, being based on the entire utterance, also included several stop consonants with complete closures (and therefore high TD_Y positions).

Fig. 5 shows that the front vowels were closer to $z_X=0$ than were the back vowels: thus, tense /uː, oː/ were some four standard deviations away from the mean compared with around two standard deviations for tense /iː, eː/. Indeed, /uː, oː/ were as peripheral in the articulatory vowel space on the horizontal z_X dimension as was the open tense vowel /ɐː/ on the vertical z_Y dimension. The second striking feature about these mean positions is that tense /uː, oː/ were much further from their lax counterparts /ʊ, ɔ/ (some two standard deviations) on the horizontal dimension than were tense /iː, eː/ from their lax counterparts /ɪ, ɛ/ (roughly one, or fewer, standard deviations).

In order to test whether tense–lax pairs were nearer to each other in front than in back vowels, we ran a repeated measures ANOVA with distance (defined by (4) and (5)) as the dependent variable and with the following independent variables: Vowel (two levels: front and back collapsed across /iː, ɪ, eː, ɛ/ and /uː, ʊ, oː, ɔ/, respectively), Place (three levels: /p, t, k/), Rate (two levels: normal vs. fast) and Tensity (two levels: tense including /iː, eː, uː, oː/, vs. lax including /ɪ, ɛ, ʊ, ɔ/). The results showed a significant effect for Vowel ($F(1, 6)=21.4$; $p<0.01$) which means that front and back vowels differed on the distance between tense and lax vowels, a significant effect for Place ($F(2,5)=7.1$; $p<0.05$) and a significant Vowel × Place interaction ($F(2, 5)=32.2$; $p<0.01$). *Post-hoc* Bonferroni t-tests showed a significant difference between front and back vowels in /p/ ($t=3.7$; $p_{adj}<0.05$; df=27) and in /t/ ($t=4.3$; $p_{adj}<0.01$; df=27) but not in /k/ contexts: these results mean that the greater tense–lax vowel distance in back, compared with front, vowels was in evidence only in /p, t/ but not /k/ contexts.

Finally, recall that there were two ways in which the distances between tense and lax vowels were calculated: firstly, by measuring the distances of all lax vowels to the corresponding tense centroid (e.g. all /ʊ/ tokens to the centroid of /uː/); and secondly, by measuring the distances of all tense vowels to the corresponding lax centroid (e.g. /uː/ tokens to the centroid of /ʊ/). These two modes of calculation had no impact on the distance between tense and lax vowels, as shown by the finding of no significant effect for Tensity and no significant Tensity × Vowel interaction.

Thus the main conclusion from these results is that the tongue dorsum's horizontal distance between tense and lax vowels was

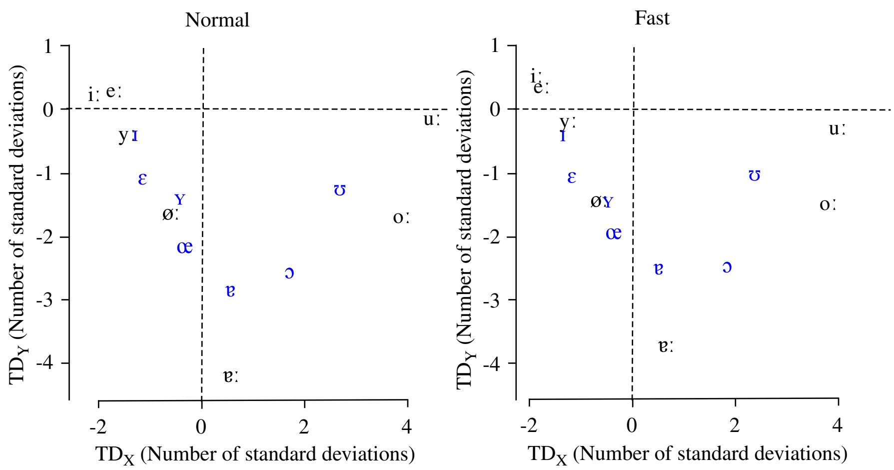

Fig. 5. The mean positions of all vowels pooled across all speakers calculated at the temporal midpoint of the vowel in a speaker normalised space of $TD_X \times TD_Y$ position for normal (left) and fast (right) rates.

greater for back, than for front vowels, in two (/p, t/) of the three consonantal contexts analysed here.

3.3. Discussion

The results from both experiments so far show that /u:, o:/ had more peripheral positions on the horizontal position of the tongue dorsum than their front counterparts: compared with front vowels, back vowels had greater CV magnitudes, greater CV peak velocities, their articulatory velocity increased significantly from a normal to a faster rate (Experiment I) and they were more distant from their lax counterparts (Experiment II). These findings suggest that back vowels should be more prone to target undershoot on this horizontal tongue dorsum parameter in speaking styles that favour hypoarticulation: that is, if at extreme forms of hypoarticulation the vowels shift towards the centre of the vowel space, then this shift should be greater for back than for front vowels because the former are, as Fig. 5 shows, several standard deviations away from the mean on this parameter.

We cannot test directly this hypothesis using the present corpus because the hypoarticulation-induced shifts due to rate differences are quite small. However, even in laboratory speech of the present kind, it should follow that, if back vowels are more peripheral than front vowels and if there is a tendency for vowels to shift towards the mean due to factors such as context or rate, then they should stray more easily into the front vowel space than front vowels do into the back vowel space. More specifically, we predicted that /u:, ʊ/ were more likely to move towards /y:, ʏ/, respectively, than were /y:, ʏ/ towards /u:, ʊ/. In the next experiment, we tested this hypothesis once again physiologically on the TD_X parameter as well as acoustically using a perceptually transformed measure of spectral slope.

4. Experiment III: Relative articulatory and acoustic distances between front and back vowels

4.1. Method

For each context vowel and separately for each speaker, we calculated the logarithm of the distance ratio between it and the sustained, isolated productions of /u:/ and /y:/. The details of this calculation are as follows. For each sustained /u:/ and /y:/, the mean across all TD_X frames in time was calculated resulting in one mean per sustained vowel. We then calculated (separately for each speaker) the mean of these means resulting per speaker in a single TD_X value, one for sustained /u:/ (henceforth T_u) and one for sustained /y:/ (T_y). We interpret T_u and T_y as idealised targets: they are the most hyperarticulated variants of these vowels from which actual values due to context or to hypoarticulation speaking styles deviate (see also Flemming (2001) and Moon and Lindblom (1994) for similar approaches). Where p_i is a single TD_X position taken from one of the same speaker's /u:, ʊ, y:, ʏ/ vowels in a CVC context between the acoustic stop release and vowel offset (periodic offset of the vowel), we calculated d_i the ratio of its absolute distance to T_u and T_y given in the following equation:

$$d_i = \frac{|p_i - T_u|}{|p_i - T_y|} \tag{6}$$

d_i was calculated for each of the N points, $p_1, p_2, \ldots, p_N$ between the acoustic stop release and acoustic offset of the vowel, resulting in N such distances, $d_1, d_2, \ldots, d_N$ per context vowel. We then calculated in (7) the log. Euclidean distance ratio, *Eratio*, which is the logarithm of the mean of these distances:

$$Eratio = \log\left(\frac{1}{N}\sum_{i=1}^{N} d_i\right) \tag{7}$$

When *Eratio* is zero, then the TD_X trajectory between the stop release and acoustic vowel offset is on average equidistant between the two steady-state, sustained targets T_u and T_y. When it is negative, then the TD_X trajectory over this interval is closer to the steady-state back vowel target, T_u; and when *Eratio* is positive, then it is nearest to the steady-state front vowel target, T_y.

We carried out an analogous calculation in the acoustic domain: in this case, the parameter was the linear slope calculated over a Bark-scaled spectrum in the frequency range 300–3500 Hz (calculated with a window size of 32 ms). This frequency range was chosen because it contains most of the acoustic information for separating /u:/ from /y:/. Because F2 is low for /u:/ but higher for /y:/, then /u:/ has its energy concentrated principally in the lower part of this frequency range and, as a result, the linear slope calculated across this range should fall much more steeply for /u:/ than for /y:/: as the boxplots of the linear spectral slopes from all seven speakers in Fig. 6 confirm, steady-state,

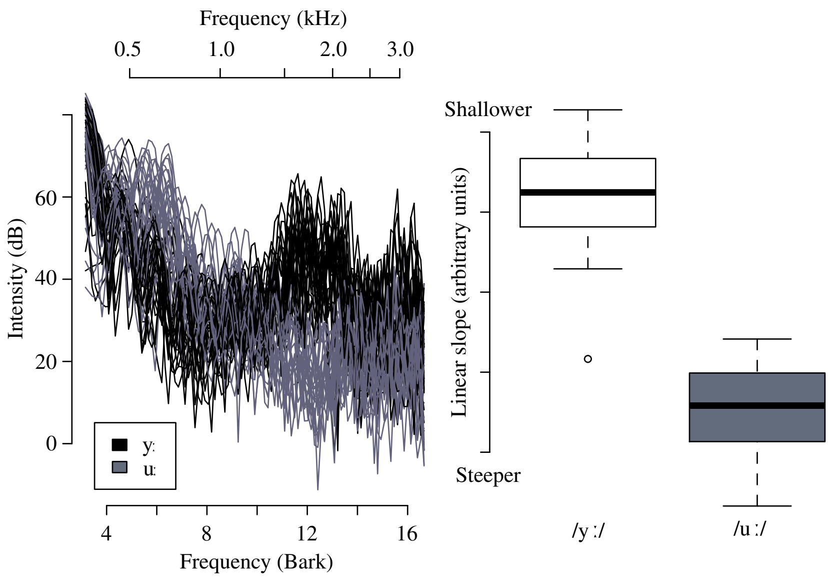

Fig. 6. Ensemble spectra of sustained, steady-state /y:/ and /u:/ for all speakers at the normal rate calculated at the temporal midpoint of the vowel (left). The distribution of the slopes of these spectra is shown on the right.

sustained /u:/ and /y:/ were indeed very effectively separated on this acoustic measure.

We calculated spectral slopes in this frequency range for the same vowels that were analysed physiologically and then applied (7) to obtain corresponding *Eratio* values in the acoustic domain over the same interval between the acoustic release of the stop and the acoustic vowel offset.

4.2. Results

The results in the top row of Fig. 7 of the relative proximity of tense /u:/ and tense /y:/ to each other on the parameter TD_X shows that at both rates /u:/ was more likely to stray into the /y:/ space than the other way round: in both the velar and alveolar contexts, the interquartile range extends over the line $Eratio = 0$ which means that many back vowel tokens were closer to T_y than to T_u in these contexts. There are similar and even more marked effects for lax vowels (row 2 of Fig. 7): at both slow and fast rates, even the median of lax /ʊ/ in an alveolar context is closer to T_y than to T_u. These trends were much more pronounced for the measures based on spectral slope shown in Fig. 8: there is fairly clear evidence that on the spectral slope derived from a Bark-scaled spectrum in the 300–3500 Hz range, high back vowels were much more likely to stray into the front vowel space than the other way round. Fig. 8 also shows that the median of /u:/ was greater than zero for all three contexts (grey-shaded boxplots) and at both rates.

We compared the distributions of high back and front vowels using a repeated measures ANOVA with independent variables Vowel (two levels: front, /y:, ʏ/ vs. back, /u:, ʊ/), Tensity (two levels: tense vs. lax), Rate (two levels: normal vs. fast), and Place (three levels: /p, t, k/). We did not want to test whether the distributions of front and back vowels were different (which will inevitably be the case, because front and back vowels predominantly have values that were different in sign as Figs. 7 and 8 show) but whether one distribution was closer to zero than the other. In order to do this, we flipped the values of the front vowels (by multiplying them by -1), which is equivalent to reflecting the boxplots for /y:, ʏ/ about the horizontal zero line. The test was then to compare whether front and back vowels differed significantly from each other. Our prediction was that they do because, after performing this reflection, the distributions for the back vowels would be closer to the horizontal zero line (the point equidistant between T_u and T_y) than those of the front vowels.

The results with the relative position of TD_X as the dependent variable showed a significant effect for Vowel ($F(1,6)=10.0$; $p<0.05$), for Place ($F(2, 5)=62.1$; $p<0.001$), and for Tensity ($F(1, 6)=388.2$; $p<0.001$). There was also Vowel × Place interaction ($F(2, 5)=8.0$; $p<0.05$). *Post-hoc* Bonferroni *t*-tests showed that the difference between front and back vowels was not significant for the /p/ context, not quite significant for the /k/ context ($t=3.1$; $p_{adj}=0.06$; df=27) and highly significant for the /t/ context ($t=4.8$; $p_{adj}<0.001$; df=27). Thus, there was a relatively greater encroachment of the back on the front vowel space primarily in a /t/ context, possibly marginally in a /k/ context, but not in a /p/ context.

Turning now to the acoustic data in which the independent variables were the same but in which the dependent variable was the relative position between T_u und T_y of the Bark-scaled spectral slope, there was a significant effect for Vowel ($F(1,6)=27.2$; $p<0.01$), Place ($F(2, 5)=13.7$; $p<0.01$), and Tensity ($F(1, 6)=56.6$; $p<0.001$), with no significant interactions between Vowel and any of the other independent factors. The acoustic data provided therefore stronger evidence than the articulatory data that back vowels were more likely to encroach on the front vowel space than *vice-versa*.

4.3. Discussion

Although /u:/ was shown to be highly peripheral as far as the TD_X magnitude, its peak velocity (Experiment I), its steady-state

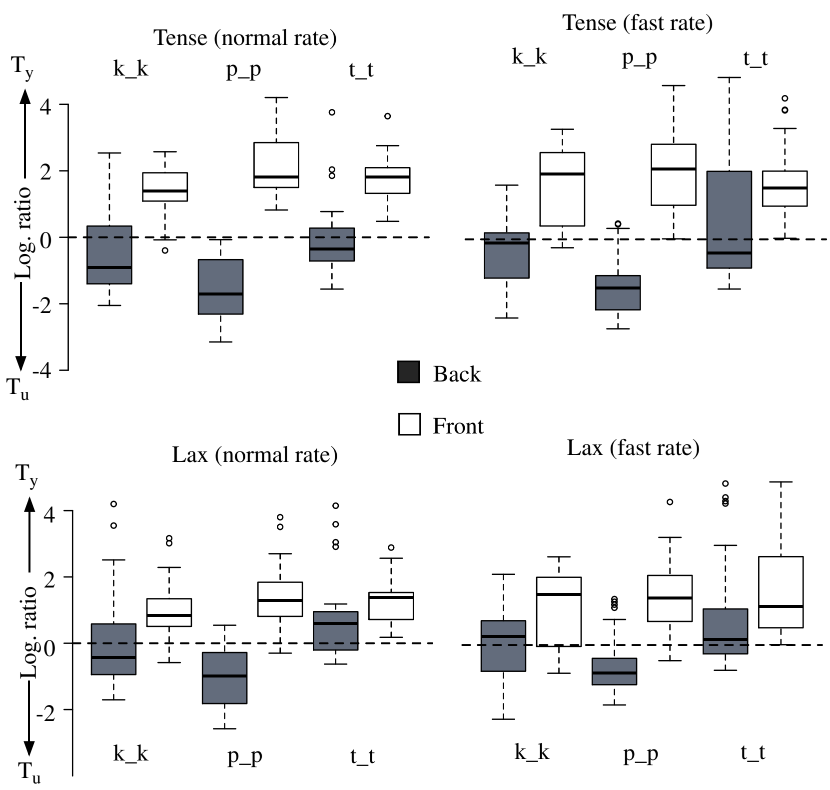

Fig. 7. Boxplots for all speakers of the relative distance between sustained, steady-state /u:/ (T_u) and /y:/ (T_y) based on the TD_X position calculated from the entire trajectory between the stop's release and acoustic vowel offset. The results are shown separately for high front (white) and high back (grey) tense (first row) and lax (second row) vowels at normal (left) and fast (right) rates. (A value of zero means that a vowel is equidistant between T_u and T_y.)

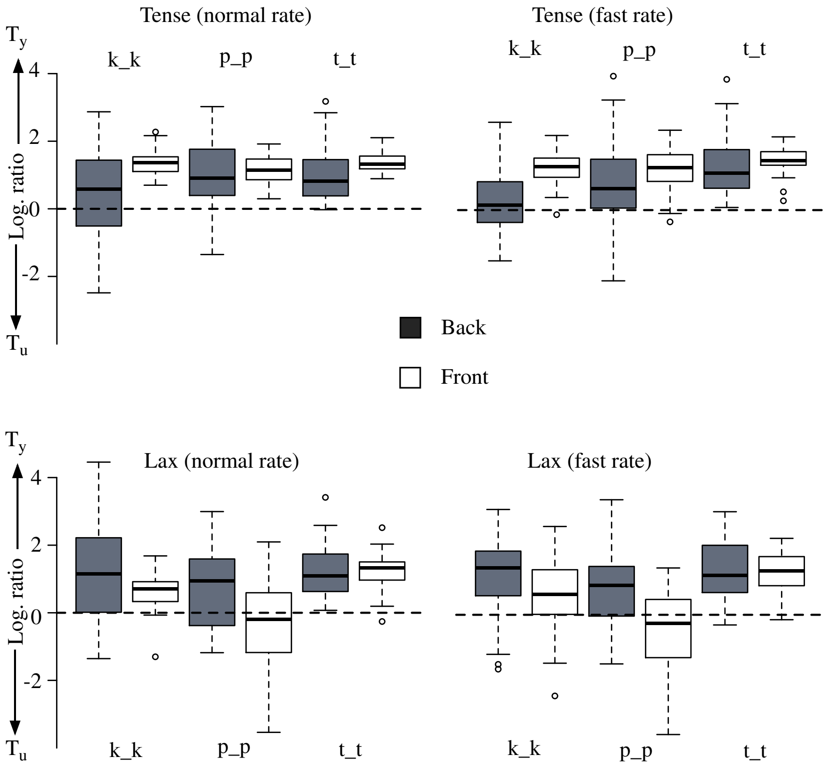

Fig. 8. Boxplots for all speakers of the relative distance between sustained, steady-state /u:/ (T_u) and /y:/ (T_y) based on the spectral slope calculated from the entire trajectory between the stop's release and acoustic vowel offset. The results are shown separately for high front (white) and high back (grey) tense (first row) and lax (second row) vowels at normal (left) and fast (right) rates. (A value of zero means that a vowel is equidistant between T_u and T_y.)

position (Experiment II) and the distance from its lax counterpart (Experiment III) compared with /y:/ are concerned, the results of the present experiment showing a greater propensity for /u:, ʊ/ to encroach both articulatorily and acoustically on the /y:, ʏ/ spaces suggest that high back rounded vowels are also quite unstable (see also Hoole & Kühnert, 1995).

It is also clear that the shift of /u:, ʊ/ in the direction of /y:, ʏ/ was much greater for the acoustic than for the articulatory data. The reason for this is that the signal over which the distance metric in (6) and (7) was applied included not only the voiced part of the vowel, but also the preceding frication/aspiration stage, although some of the information about place of articulation especially in alveolars (whose frication stage is characterised by high frequency energy) will have been excluded, because the calculation was restricted to the frequency range 300–3500 Hz. Nevertheless, given that this transition from the plosive and in particular from the alveolar stop to the vowel includes high frequency energy, and since /y:, ʏ/ spectra have more energy in higher frequencies than those of /u:, ʊ/, then the greater tendency for high back vowels to shift acoustically into the space of high front vowels than the other way round may, to a certain extent, be a consequence of the high frequency energy at the transition from the frication/aspirated stage of the stop into the voiced part of the vowel: thus the large acoustic shift of the back vowel into the front vowel space cannot be attributed to the effects of fronting the back of the tongue alone.

In the next experiment, we investigated perceptually the relative confusability between high front and back vowels from the same corpus: based on the results from Experiment III, our prediction was that the confusion would be correspondingly asymmetric in which high back vowels were more likely to be perceived as their front counterparts than *vice-versa*.

5. Experiment IV: Perceptual confusions between /ʊ/ and /ʏ/

5.1. Method

In order to test whether the confusion between front and back rounded vowels was asymmetric, the stimuli for the perception experiment were taken exclusively from the lax vowels spoken at the fast rate, given that a pilot experiment had shown almost completely correct identifications for the other categories (either /y:/ vs. /u:/ produced at either rate or /ʏ/ vs. /ʊ/ at the normal rate of speech). The lax /ʏ/ and /ʊ/ vowels produced by all seven speakers at the fast rate were excised from the speech signal between their acoustic periodic onset and offset, randomised and presented once to 17 first language adult listeners of German at intervals of 1 second: their task was to identify the vowel as either /ʏ/ or /ʊ/ with a mouse click on the corresponding orthographic symbol (*Ü* or *U*) that appeared on a computer monitor. The total number of responses was 2 (/ʏ/ or /ʊ/) × 3 (/p, t, k/ contexts) × 7 (speakers) × 5 (speaker repetitions) × 17 (listeners) = 3570. The hypothesis to be tested was that there would be a greater proportion of misclassifications of /ʊ/ as /ʏ/ than the other way round.

5.2. Results and discussion

The results in Fig. 9 show a greater proportion of /ʊ/→/ʏ/ than /ʏ/→/ʊ/ misclassifications. However, the same figure also shows very clearly that these results were context-dependent. In the (symmetrical) /k/ context, the proportion of misclassifications was about the same for the two vowels, whereas in the /p/ context there were more /ʏ/→/ʊ/ and in the /t/ context substantially more /ʊ/→/ʏ/ misclassifications than confusions in the other direction. Given that the results of the confusions are in opposite directions in two contexts, three generalised linear mixed models were applied to

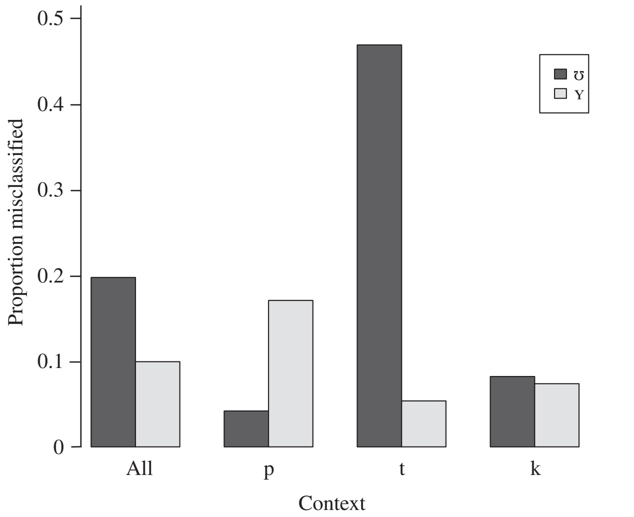

Fig. 9. The proportion of misclassifications of /ʊ/ and /ʏ/ (as a function of total responses) from all listeners shown separately in the three contexts and pooled across contexts (*All*). The responses were based on sections of these vowels excised between the voicing onset and offset at the fast rate.

the data separately for each /p, t, k/ context. In all cases, the dependent variable was classification (correct or incorrect), and the independent variables were Vowel (two levels: /ʊ, ʏ/) and Place (three levels: /p, t, k/); the listener (17 levels) and the speaker (7 levels) from which the stimuli had been derived were random factors. The results showed a highly significant influence of Vowel on classification in a /p/ ($z=8.8$; $p<0.001$) and /t/ ($z=14.4$; $p<0.001$) but not in the /k/ context. Thus, there was a significantly greater likelihood of /ʏ/→/ʊ/ than /ʊ/→/ʏ/ confusions in a /p/ context, and a significantly greater likelihood of /ʊ/→/ʏ/ than /ʏ/→/ʊ/ confusions in a /t/ context, but the proportion of confusions was about the same in the /k/ context. Finally, and compatibly with Fig. 9, the results of a mixed model with the same factors as before showed that the proportion of /ʊ/→/ʏ/ confusions in a /t/ context was significantly greater ($z=11.2$; $p<0.001$) than the proportion of /ʏ/→/ʊ/ confusions in a /p/ context.

The results therefore support the hypothesis that the confusions are asymmetric: /ʊ/ was more likely to be misperceived as /ʏ/ than the other way round (left most barplot in Fig. 9). However, this came about both because of the very high proportion of misperceptions of /ʊ/ as /ʏ/ in an alveolar context and because this proportion was higher than that of the misperceptions of /ʏ/ as /ʊ/ in a labial context. In general, these results are consistent with the acoustic data from Experiment III—the main basis for comparison here is with the lower right panel of Fig. 8, which shows the results for the relative distance of lax vowels at the fast rate between sustained /y:/ and /u:/. Thus as for the present perception experiment, the extent to which /ʊ/ shifted into the /y:/ space acoustically was greatest in the alveolar context (as shown by the boxplot which is well above the horizontal line, *Eratio*$=0$ in Fig. 8) and the extent to which /ʏ/ shifted into the /u:/ space was greatest in the labial context. The main discrepancy between the acoustic and perceptual data is for the /k/ context: contrary to the pattern of results from the perception experiment, the shift acoustically of /ʊ/ towards /y:/ was greater than that of /ʏ/ towards /u:/. That there should not be complete parity between the results of the acoustic and perceptual experiments is not surprising, given that the acoustic study included the initial voiceless transitions, whereas in the present experiment, responses were elicited from the acoustically voiced part of the vowel.

Finally, we investigated the extent to which the perceptual responses could be related to the articulatory and acoustic data of Experiment III. We did this only for the two contexts in which the perceptual responses were the most equivocal, i.e. for /pʏp/ and /tʊt/. To do so, we ran two separate generalised linear mixed models, one for each context, with the categorical response as the dependent variable and with the relative distance between the steady-state targets T_u and T_y as the independent variable. The speaker and listener were entered into the model as random factors. The results for the /tʊt/ context showed that the categorical perceptual responses – that is, whether the listeners responded with *Ü* (front) or *U* (back) – were predictable from the relative position between T_u and T_y in data from both the back of the tongue ($z=2.00$, $p<0.05$) and from spectral slope ($z=3.53$, $p<0.01$). As these results show, the fit to the spectral slope was evidently better which suggests that there was information in the acoustic signal beyond that contributed by the shift of the back of the tongue that had an influence on listeners' judgments. For /pʏp/, the categorical perceptual responses were found to be predictable from the relative position of the back of the tongue ($z=3.20$, $p<0.01$) but not from that of spectral slope ($z=0.47$, $p>0.05$). A possible interpretation for these findings lies in observing that both the proportion of /ʏ/→/ʊ/ misperceptions (Fig. 9) and the relative shift of the back of the tongue (Fig. 7, lower right panel) were small in this labial context, whereas the shift towards T_u due to spectral slope (Fig. 8, lower right panel) was considerable. Consequently, even though the acoustic shift towards the idealised /u:/ target was large for /pʏp/, listeners' judgments of the vowel in this context were not swayed by it. Thus the ambiguity in spectral slope (that is, productions in which the spectral slope of /ʊ/ resembles that of /ʏ/ or *vice-versa*) appears to have been more detrimental to the identification of the intended vowel in /tʊt/ than in /pʏp/.

6. General discussion

Taken together, the four studies of this paper can begin to provide a phonetic explanation for the greater likelihood of high back vowels to front diachronically than of high front vowels to retract. The first two experiments showed that high back vowels had the largest and fastest change in position of the tongue dorsum in CV transitions. The deviation of the tongue dorsum in tense vowels along a horizontal axis both from the centre of the vowel space and from the corresponding lax vowels was greater for back than for front vowels: in general, these experiments suggest that high back vowels that are as peripheral as those in German have a high articulatory cost. These results are consistent with those from an electromyographic study of American English vowels in Alfonso and Baer (1982) who showed that the horizontal movement of the tongue began some 500 ms before voicing onset in /u/ and around 250 ms earlier than the horizontal movement in producing /i/. They attribute these differences in part to the intrinsically slower backing movement effected by the styloglossus muscle compared with the more rapid fronting and raising movements that are both brought about by genioglossus activity in /i/.

The second set of experiments showed that under conditions that are likely to induce hypoarticulation such as context or faster speech, the confusion between front and back vowels was asymmetric: high back vowels were more likely to stray into the front vowel space than the other way round and /ʊ/ was more likely to be misidentified by listeners as /ʏ/ than /ʏ/ was as /ʊ/. This synchronic pattern is consistent with the corresponding diachronic asymmetry favouring a fronting of high back vowels. The predisposition of high back vowels to be misidentified as high front may explain why in languages that contrast high front rounded and high back rounded vowels, the extent of tongue dorsum retraction and therefore CV magnitude and velocity need to be quite extreme for /u:/, in order to ensure that it is produced such that it is acoustically distinct from /y:/.

Compatibly with many other studies (Flemming, 2001; Hillenbrand, Clark, & Nearey, 2001; Lindblom, 1963; Oh, 2008; Stevens & House, 1963), alveolars were found to have a marked coarticulatory fronting effect on high back vowels and were responsible for the especially high proportion of perceptual misclassifications of back as front vowels. At the beginning of this paper, we had suggested that the acoustic and perceptual deviation of high vowels due to coarticulatory influences was likely to be greater for back vowels in an alveolar than for front vowels in a labial context, because although the latter induces a reduction of F2, there is no articulatory retracting influence of the tongue dorsum in the same way that coronals induce a lingual fronting in back vowels. Our results in experiments III and IV are consistent with this idea. Thus while /ʏ/ is most likely to stray acoustically (see Fig. 8, lower row) and perceptually (Fig. 9) into a back vowel space in a labial context, the acoustic and perceptual /ʊ/ misclassifications as /ʏ/ in an alveolar context are far more dramatic. At the same time, it must be recognised that these results cannot be easily extended to perceiving spontaneous speech, given that listeners do not usually hear vowels devoid of any consonantal context as they did in this experiment. Nevertheless, the results suggest how confusions could arise if listeners fail to attribute coarticulatory effects in vowels to consonantal context, which has been argued to be a condition that triggers sound change (Ohala, 1993).

It is interesting to consider in the light of these results whether high back vowels occur proportionately more often in alveolar or coronal contexts than high front vowels in retracting contexts, since this could then be another factor that contributes to the likelihood of diachronic back vowel fronting. This question is of course very difficult to answer because, with the exception of a small number of well-studied languages, there is a lack of both language corpora that include measures of phonotactic frequency and of speech corpora containing sufficient information to assess the effect of speaking style on vowel centralisation. However, it is possible to produce some relevant descriptive statistics from the UCLA Phonological Segment Inventory Database (UPSID) of 451 languages (Maddieson, 1984). For this purpose, we defined a consonantal context *acute* that included the three categories dental, dental/alveolar, alveolar which are most likely to induce back vowel fronting; and a consonantal context *grave* that included five place categories bilabial, labiodental, retroflex, labial-velar, and uvular that are likely to cause some degree of front vowel retraction. This backing influence on front vowels is likely to come about because of a spectral shift in the vicinity of F2 (Fig. 6) in the context of the three labial places of articulation and because the curled tongue-tip position in retroflex and the very back tongue position in uvular consonants are likely to perturb the fronted tongue dorsum position in palatal vowels. The question we asked of UPSID was this: is the proportion of *acute* greater than *grave* in languages that have either high front or high back vowels? If so, then the composition of a language's sound system may also contribute to the greater tendency for high back vowels to front compared with high front vowels to retract. Clearly, such an analysis cannot provide a decisive answer to this question: for this, we would require statistics of type and token frequencies of words containing combinations of these consonants with high vowels which are not available for most of the languages surveyed in UPSID.

In order to answer this question, we used the web-interface to UPSID (Reetz, 2010) to search on all languages that had, according to the available criteria of the search engine, either normal voiced, high back rounded monophthongs, normal voiced high front monophthongs, or one of the eight places of articulation listed under *acute* and *grave* as defined above. The vowel and consonant searches were made with no other qualifications. These searches therefore resulted in eight lists of segment occurrences in the separate languages often containing a very large number of variants: for example, laryngealised, pharyngealised, fricative, overshort and nasalised variants of [u]; labiodentals with various voice qualities (voiceless, voiced, and breathy) manner (fricative, affricate, and stop) and secondary (labialised, palatalised, and prenasalised) articulations. Scripts in the R programming language were then used to determine various conjunctions of these sets and to tabulate the proportions for the two vowel types.

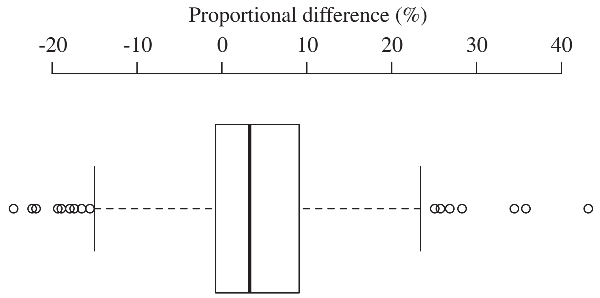

Fig. 10. A boxplot showing the difference between the proportion of *acute* and *grave* consonants (as a function of the total number of sounds) in 375 languages that have both high front and high back vowels (a value of zero indicates that a language has the same number of *acute* and *grave* consonants).

The results of this analysis showed that on average *grave* consonants made up 20.3% of the sounds in a language from the 411 languages with high front vowels; the corresponding proportion of *acute* consonants in the 383 languages with high back vowels was 24.3%, i.e. there is approximately a 4% bias towards *acute*. In order to test whether this difference was significant, we examined only the 375 languages that had both high front and high back vowels and then summed the proportion of *acute* and *grave* consonants in each language. The distribution of the difference between these two categories calculated separately for each language in Fig. 10 shows once again a bias towards *acute*. The results of a paired sample *t*-test with Language as the independent factor and proportion as the dependent variable showed that this result was highly significant ($t=8.2$; $p<0.001$; $df=374$). Thus when a language contains both high front and high back vowels, then the proportion of *acute* consonants (as defined here) is likely to be some 4% greater than the proportion of *grave* consonants. This slight preference for tongue-tip consonants combined with their greater coarticulatory influence on back vowels than that of *grave* (or at least of labial) consonants on front vowels are likely to be contributing factors both to the slight left–right asymmetry in the distribution of vowels in the languages of the world (Schwartz et al., 1997) and to the greater tendency for back vowels to front than front vowels to retract diachronically (Labov, 1994). Thus, the vowel space is probabilistically slightly skewed towards the front.

Based on these findings, our general conclusion is that a peripheral [u] with a low F1 and F2 and wide spacing between F2 and F3 may not be as common as auditory impressions have suggested, especially in the vast majority of languages that do not make a phonological contrast between front and back rounded vowels: that is, most languages may have high back rounded vowels that are either [ʉ] or intermediate between [ʉ] and [u]. However, given the availability of experimental evidence and speech corpora for only a handful of the 375 languages represented in Fig. 10, this conclusion must remain speculative.

Acknowledgments

We thank Associate Editor Kenneth de Jong and two anonymous reviewers for very helpful comments on an earlier draft of this manuscript. This research was supported by German Research Council Grant HA 3512/3-2.

References

Alfonso, P., & Baer, T. (1982). Dynamics of vowel articulation. *Language & Speech*, *25*, 151–173.

Butcher, A. R., & Anderson, V. B. (2008). The vowels of Australian Aboriginal English. In *Proceedings of the ninth annual conference of the International Speech Communication Association* (pp. 347–350).

Cox, F. M. (1999). Vowel change in Australian English. *Phonetica*, *56*, 1–27.

Cox, F. M., & Palethorpe, S. (2001). The changing face of Australian English vowels. In D. B. Blair, & P. Collins (Eds.), *Varieties of English around the world: English in Australia* (pp. 17–44). Amsterdam: John Benjamins Publishing.

Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, *18*, 7–44.

Flemming, E. (2003). The relationship between coronal place and vowel backness. *Phonology*, *20*, 335–373.

Fridland, V. (2008). Patterns of /uw/, /ʊ/, and /ow/ fronting in Reno, Nevada. *American Speech*, *83*, 432–454.

Harrington, J. (2010). *The phonetic analysis of speech corpora*. Chichester: Wiley-Blackwell.

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

Hawkins, S., & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association*, *35*, 183–199.

Hillenbrand, J. M., Clark, M. J., & Nearey, T. M. (2001). Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America*, *109*, 748–763.

Hoole, P. (1999). On the lingual organization of the German vowel system. *Journal of the Acoustical Society of America*, *106*, 1020–1032.

Hoole, P., & Kühnert, B. (1995). Patterns of lingual variability in German vowel production. In *Proceedings of the 13th international conference of phonetic sciences*, Stockholm (Vol. 2, pp. 442–445).

Hoole, P., & Mooshammer, C. (2002). Articulatory analysis of the German vowel system. In P. Auer, P. Gilles, & H. Spiekermann (Eds.), *Silbenschnitt und Tonakzente* (pp. 129–152). Tübingen: Niemeyer.

Labov, W. (1994). *Principles of linguistic change, Vol. 1: Internal factors*. Oxford: Blackwell.

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English: Phonetics, phonology, and sound change*. Berlin: Mouton de Gruyter.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulations of vowel quality systems: The role of perceptual contrast. *Language*, *48*, 839–862.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, *35*, 1773–1781.

Lindblom, B. (2003). Patterns of phonetic contrast: towards a unified explanatory framework. In *Proceedings of the 15th international congress of phonetic science (Barcelona)* (pp. 39–42).

Maclagan, M., & Hay, J. (2007). Getting fed up with our feet: Contrast maintenance and the New Zealand English "short" front vowel shift. *Language Variation and Change*, *19*, 1–25.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America*, *96*, 40–55.

Nelson, W. (1983). Physical principles for economies of skilled movements. *Biological Cybernetics*, *46*, 135–147.

Oh, E. (2008). Coarticulation in non-native speakers of English and French: An acoustic study. *Journal of Phonetics*, *36*, 361–384.

Ohala, J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical linguistics: Problems and perspectives* (pp. 237–278). London: Longman.

Öhman, S. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, *39*, 151–168.

Reetz, H. (2010). *User interface to the UCLA Phonological Segment Inventory Database (UPSID)*, 〈http://web.phonetik.uni-frankfurt.de/upsid_info.html〉. Retrieved 14.6.2010.

Schwartz, J.-L., Boë, L., Vallée, N., & Abry, C. (1997). Major trends in vowel system inventories. *Journal of Phonetics*, *25*, 233–254.

Stevens, K. (1989). On the quantal nature of speech. *Journal of Phonetics*, *17*, 3–45.

Stevens, K. (2003). Acoustic and perceptual evidence for universal phonological features. In *Proceedings of the 15th international congress of phonetic science (Barcelona)* (pp. 33–38).

Stevens, K. N., & House, A. S. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research*, *6*, 111–128.

Tabain, M. (2008). Production of Australian English /ʉ:/ language-specific variability. *Australian Journal of Linguistics*, *28*, 195–224.

Tabain, M., & Perrier, P. (2007). An articulatory and acoustic study of/u/ in preboundary position in French: The interaction of compensatory articulation, neutralization avoidance and featural enhancement. *Journal of Phonetics*, *35*, 135–161.