CONTROL OF LARYNX HEIGHT IN VOWEL PRODUCTION

Philip Hoole & Christian Kroos

Institut für Phonetik und Sprachliche Kommunikation
Ludwig-Maximilians-Universität München
Schellingstr. 3
D-80799 Munich, Germany
email: hoole@phonetik.uni-muenchen.de
kroos@phonetik.uni-muenchen.de

ABSTRACT

Digital video filming of the thryroid prominence was used to measure larynx height in German vowels, with focus on contrasts involving front unrounded, front rounded and back rounded vowels. The study aimed to provide a foundation for interpreting the acoustic consequences of articulatory manoeuvres not only at the larynx but also elsewhere in the vocal tract. Results showed the expected pattern of lower larynx position for the rounded vowels. However no clear preference emerged for the same, more, or less larynx lowering on front rounded versus back rounded vowels. Coarticulatory effects of the flanking consonants were weak. The most striking result was that the magnitude of the differences between vowels varied substantially over speakers. This reinforces the contention that interpretation of vertical laryngeal gestures must be embedded in speaker-specific analysis of downstream articulatory manoeuvres.

1. INTRODUCTION

Together with the lips, the vertical position of the larynx is the major determinant of vocal tract length, and thus has a fundamental impact on the acoustic properties of the tract. This aspect of laryngeal articulation is thus an important element in a full account of vowel production. Some basic regularities are well-known. In particular, rounded vowels such as /u/ have a lower larynx position than unrounded vowels such as /i/. The differences may be substantial; studies by Ewan & Krones (1974) for English, Riordan (1977) for French, and Petersen (1983) for Dutch indicate that the differences between /i/ and /u/ may be of the order of 1cm, though all these studies also indicate that a fair amount of interspeaker variability in the magnitude of the effects is to be expected.

Languages such as German with front rounded vowels (e.g /y/) are of particular interest. It seems to be generally agreed that they pattern like back rounded vowels, and Wood (1986) has advanced the very specific acoustically-motivated hypothesis that lower larynx position in front rounded vowels serves to ensure that the acoustic consequences of potential perturbations to the main linguo-palatal constriction are similar for /y/ and /i/ (a difference of 15mm between low-larynx and high-larynx vowels was used in this modelling study). However, measurements explicitly comparing the larynx height of front unrounded, front rounded and back rounded vowels for a large corpus of data are extremely scarce.

Thus it is by no means clear for German whether the assumed differences are equally clear for multiple speakers, for different degrees of vowel height, for both tense and lax vowels, for different consonantal contexts, and when recorded with a sufficient number of repetitions to allow statistical analysis. This in turn means that it is currently unclear how tightly speakers control this aspect of vowel production, and thus how central an ingredient it is in the bundle of articulatory maneuvers making up the motor plan for specific vowels.

In short, there were essentially two motivations behind the present study, the preliminary stages of which we will report on here. Firstly, by examining a large corpus of German data it should be possible to achieve a better understanding of motor representations in vowel production. Secondly, the study should be seen within the wider context of investigations of articulatory-acoustic relationships in vowel production. For the speakers taking part in this study, recordings of tongue and lip activity in vowel production are either available or planned. For such speakers, an interesting question is, for example, whether the acoustic consequences of variability in lingual positioning for vowels fit in with the predictions of current vocal tract models. However, questions of this kind are difficult to address without some knowledge of the length of the vocal tract for the sounds under consideration. This approach can also be turned around: If interspeaker differences in laryngeal behaviour occur (as seems quite likely), then it may ultimately be possible to resolve some of these differences through consideration of articulatory behaviour further downstream in the vocal tract.

2. MATERIAL AND PROCEDURE

All 15 monophthongal vowels of German (/iː,ɪ;yː,ʏ;eː,ɛ;ɛː;øː,œ; ɑː,a;oː,ɔ;uː,ʊ/) were spoken in 3 symmetrical consonant contexts (/p, t, k/), embedded in the carrier phrase "*Ich habe geCVCe gesagt*". Five repetitions of each CVC combination were produced by 3 male speakers with large thyroid prominences.

The laryngeal area of the subjects was filmed in profile using a high-quality CCD camera and a telecentric lense (f=55mm). Field of view was approximately 13*10 cm, the full frame resolution of the camera was 748 horizontal by 576 vertical. Video sequences of 1 second duration, triggered by hand to catch the target portion of each utterance, were digitized online during the experiment at 50 fields per second. The speech signal was recorded on DAT tape. A second computer was used to present the speech stimuli to the subject. This computer also generated a synchronization pulse that was recorded on the second channel of the DAT tape and remotely controlled a FOR.A VTG55 video timer to place in each video field a series of digits indicating the trial number and the time in milliseconds relative to this pulse. This timing information was then extracted from the digitized images using a pattern recognition algorithm. After processing the speech signal with a waveform editor, the timing information was used to select the video field closest to the midpoint of each target vowel for further analysis (in future work we will look at the time course of laryngeal movement over the complete CVC sequences).

A contour recognition algorithm was then used to extract the contour of the neck as well as the centroid coordinates of three small circular spots attached to the neck. The latter were used to compensate for slight changes in the position of the neck over the course of the experiment. The neck contours were rotated so that the average orientation of the contours was vertical. The vertical position of the laryngeal prominence was then determined semi-automatically as the vertical coordinate of the contour corresponding to where the most anterior horizontal coordinate occurred.

3. RESULTS

For the purposes of the present paper we will focus on similarities and differences in laryngeal positioning for vowels grouped according to tenseness and phonological vowel height. In other words we form the following 4 groups of vowels: 1. tense, high (/i:, y:, u:/), 2. tense, mid (/e:, øː, o:/), 3. lax, high (/ɪ, ,ʏ, ʊ/) and 4. lax, mid(/ɛ, œ, ɔ/). For each of these 4 groups (and for each speaker) two-way ANOVAs were run with factors VOWEL and CONSONANT (i.e /p/ vs. /t/ vs. /k/). The results are summarized in Table I-III. The principal preliminary expectation was that the main effect of VOWEL was most likely to be significant for the tense groups and the main effect of CONSONANT for the lax groups.

Speaker SR		Vowel	Consonant	V*C
tense	high	***	*	n.s.
	mid	***	***	n.s.
lax	high	**	***	n.s.
	mid	**	***	n.s.

Table I Results of the two-way ANOVA with factors VOWEL and CONSONANT for subject SR: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Speaker KH		Vowel	Consonant	V*C
tense	high	***	n.s.	n.s.
	mid	**	*	n.s.
lax	high	n.s.	***	n.s.
	mid	*	*	n.s.

Table II Results of the two-way ANOVA with factors VOWEL and CONSONANT for subject KH: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Speaker FS		Vowel	Consonant	V*C
tense	high	***	n.s.	n.s.
	mid	***	n.s.	n.s.
lax	high	***	n.s.	n.s.
	mid	***	**	n.s.

Table III Results of the two-way ANOVA with factors VOWEL and CONSONANT for subject FS: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

By and large this expectation is fulfilled. However, even before we go into specific results for individual sounds it is apparent that the patterns vary quite substantially over speakers. For example, the effect of CONSONANT is particularly strong for speaker SR and particularly weak for FS (while FS has a very strong VOWEL effect, even on the lax vowel groups). Fortunately, interpretation of the results is not further complicated by significant VOWEL*CONSONANT interactions.

On the background of these ANOVA results, and for the sake of brevity, we will restrict consideration of laryngeal position for individual vowels to the tense vowel groups, and consideration of the consonantal coarticulation effects to the lax vowel groups.

Fig.1 shows for each speaker, and for each tense vowel group, the results for the individual vowels averaged over all three consonantal contexts relative to the overall average height of all vowels of the speaker, including /ɛ:/, /ɑː/, and /a/[1]. As was to be expected from previous investigations the unrounded vowel consistently attracts the highest position. The more striking feature of the results is the very large difference in the magnitude of the effects over speakers (for instance compare differences of /e:/ and /o:/ for speaker FS (~ 9mm) and speaker KH (~ 1mm). Speaker-specific features are apparent not just in the overall magnitude of the effects but also in the relative position of the front and back rounded vowels, with FS showing a higher position for front rounded than back rounded, SR showing the reverse, and KH, for whom the all effects were very weak, showing no clear pattern. Note that there is, however, intra-speaker consistency, in the sense that all speakers show the same pattern for both the high and the mid vowel groups.

Turning now to the consonantal effects, Fig. 2 shows for the two groups of lax vowels the laryngeal position averaged over all vowels for the three consonantal contexts, again relative to the overall average height of the speaker. Although some clear contextual effects are to be found, e.g lower position of the larynx in /p/ context and higher position in /k/ context for speaker SR and FS, it is nonetheless apparent that both the magnitude and consistency of contextual effects are weaker compared to the intrinsic effects of vowel identity shown in Fig. 1.

---

[1] Speaker-specific patterns were also apparent for the vowel /ɑː/, which has not been considered explicitly here. For speaker FS it was located between /i:/ and /y:/, for SR it was substantially higher than /i:/ (and consequently had the highest position of all vowels), for KH it was about the same as /i:/.

4. DISCUSSION

The main aim of the present study was to examine whether front rounded and back rounded vowels show equally clear differences in larynx height from front unrounded vowels. In the face of substantial interspeaker differences no clear answer to this specific question could be given, though rounded vowels as a whole did show the expected lower larynx position compared to unrounded vowels. The observation of messy interspeaker differences is of course not particularly useful by itself. However, the magnitude of these effects (though observable in other studies) does not seem to have been remarked on explicitly before. Thus speakers such as FS and SR could have of the order of almost 1cm MORE difference between a rounded/unrounded pair than a speaker such as KH. This must result in a substantial difference in the acoustic properties of their vocal tracts and serves to emphasize the point made in the introduction that such investigations of larynx height must be embedded in investigations of articulatory strategies employed for these sounds at other locations in the vocal tract.

Wood's modelling studies indicate that laryngeal depression cannot be used as a simple compensatory mechanism for lip-rounding undershoot in front rounded vowels. Thus we would not necessarily predict that speakers with particularly strong laryngeal depression show weaker lip-rounding for rounded vowels. In fact the reverse hypothesis may be more likely, namely that speakers with weak lip-rounding for front rounded vowels may show weak laryngeal depression since the lingual constriction will not then have been displaced into an acoustically instable region.

A further, almost equally speculative possibility, is that some speakers simply may not adopt strategies that acoustic theory suggests to be optimal in terms of stability. One direction for future research is thus to link up multiple measures of articulatory variability with accompanying acoustic variability.

A further clear necessity is to increase the number of speakers analyzed. This is now quite straightforward on the basis of the methodological developments carried out for this pilot study. Direct digitization of video data gives excellent image quality and

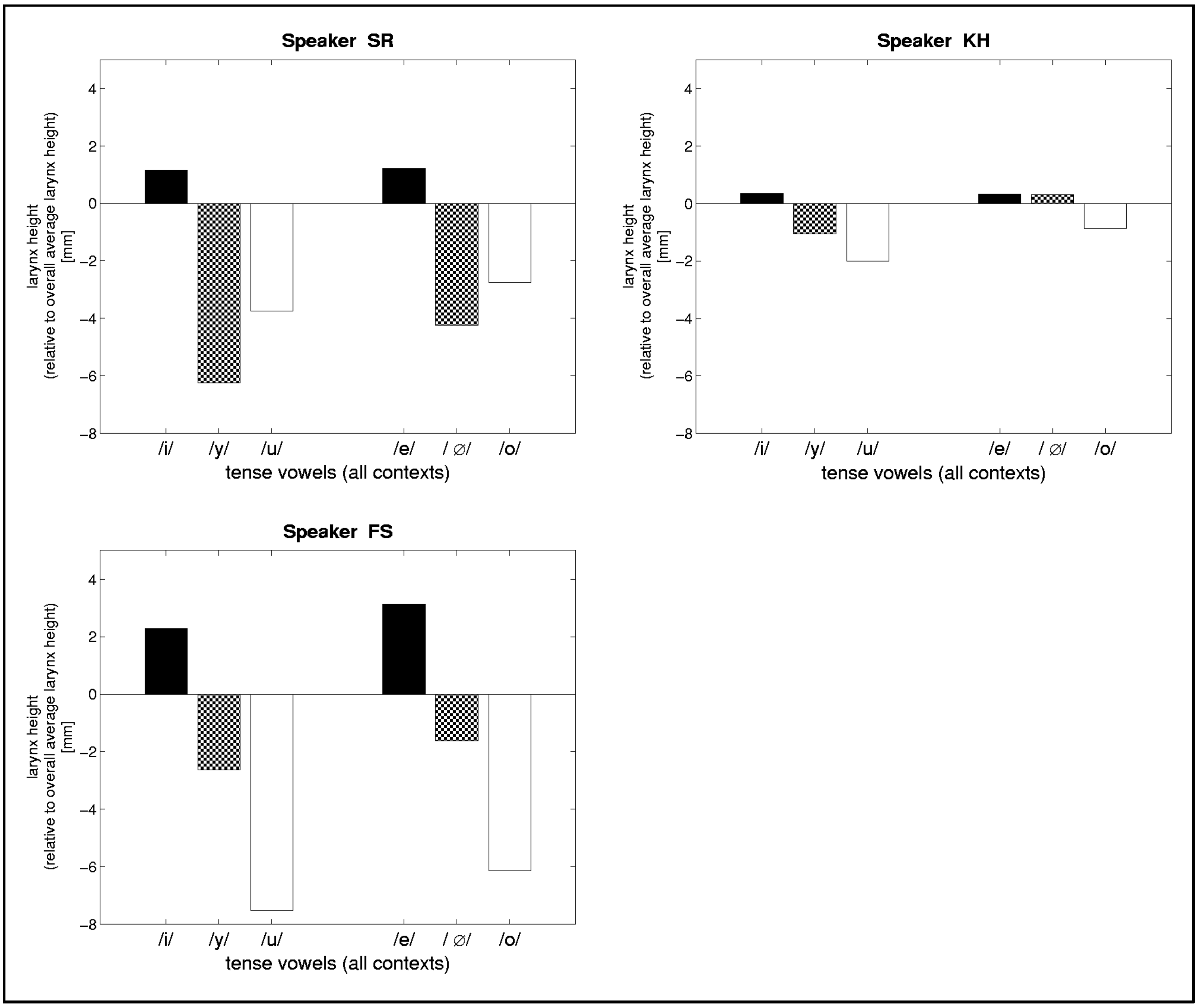

Figure 1: Larynx height for tense vowels (averaged over all consonantal contexts) relative to overall average larynx height

the requirements on disk storage and main memory for lengthy experiments of this kind are no longer a major stumbling block. The actual image processing algorithms are fairly straightforward and the analysis can be largely automated.

One caveat remains: Using this approach the speaker population will always be biased towards males with large thyroid prominences. We are currently investigating whether alternative technology, such as ultrasound and magnetic resonance scans, can be used to record less protuberant populations.

REFERENCES

Ewan, W.G. & Krones, R. (1974). "Measuring larynx movement using the thryroumbrometer". J. Phonetics, 2, 327-335.

Petersen, N.R. (1983). "The effect of consonant type on fundamental frequency and larynx height in Danish". ARIPUC, 17, 55-86.

Riordan, C.J. (1977). "Control of vocal-tract length in speech". J. Acoust. Soc. Am., 62, 998-1002.

Wood, S. (1986). "The acoustical significance of tongue, lip and larynx maneuvers in rounded palatal vowels". J. Acoust. Soc. Am., 80, 391-401

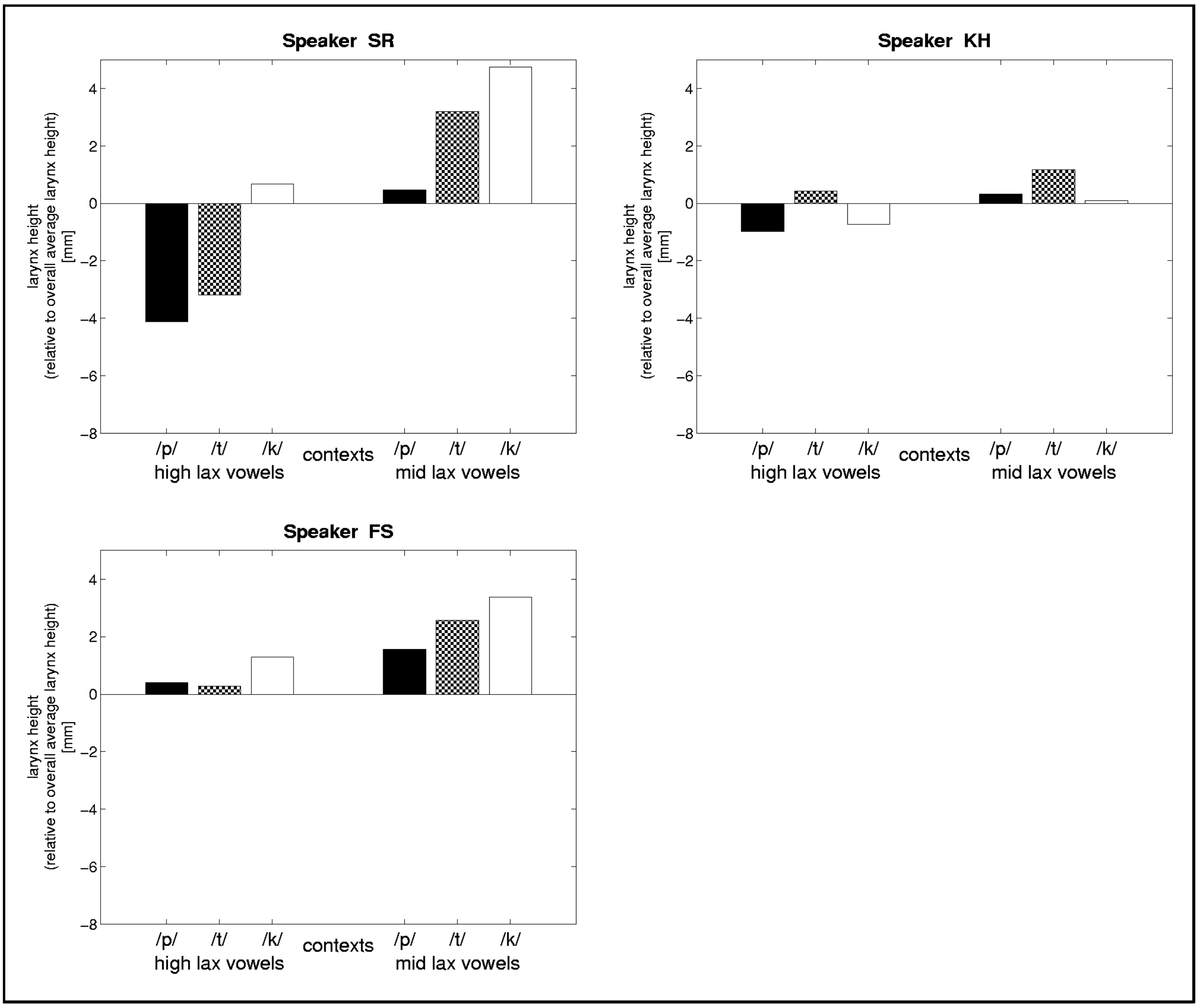

Figure 2: Mean larynx heights of the two lax vowel groups depending on context (relative to the overall averaged larynx height)