An EMMA/EPG study of voicing contrast correlates in German

Susanne Fuchs [1,2] and Pascal Perrier[3]
[1] ZAS, Berlin, Germany; [2] QMUC, Edinburgh, United Kingdom;
[3] ICP & Université Stendhal Grenoble, France
E-mail: fuchs@zas.gwz-berlin.de, perrier@icp.inpg.fr

ABSTRACT

Except for cavity enlargement strategies there is not much consensus about the involvement of supralaryngeal movements in the production of the voicing contrast. In order to study supralaryngeal stop production mechanisms we investigated the kinematics of tongue tip and jaw as well as tongue-palate contact patterns for four German subjects. We took alveolar stops in word medial (Cm) and word final position (Cf) into account. Results from Cm provide evidence that even though acoustic results exhibited consistently a longer closure duration for the voiceless stops, speaker-dependent articulatory mechanisms were involved. In word final position the rule of final devoicing applies in German, i.e. voiced stops are neutralised to voiceless. Results from acoustics and EPG generally showed complete neutralisation, but some differences, particularly in jaw position at the consonantal target and in tongue-jaw coordination, are still maintained.

1. INTRODUCTION

Laryngeal and supralaryngeal movements participate in the production of the phonological voice–voiceless distinction for stops. Much attention has been dedicated to laryngeal adjustment, i.e. glottal opening for voiceless stops and closed glottis for voiced stops. There are less consistent results with regards to supralaryngeal movements (Löfqvist and Gracco, [6]). On the one hand there seems to be general agreement about strategies for cavity enlargement. These strategies contribute to limit the increase of intraoral air pressure in order to maintain voicing during oral closure (Westbury, [9]). On the other hand only weak or non-significant effects were found with respect to differences in tongue movement, at least for alveolar stops (Kent and Moll [5]). Some trends in jaw involvement (jaw velocity peaks) were found by Fujimura and Miller [2]. Since tongue and jaw are connected articulators, their coordination could also help to facilitate the voicing distinction.

However, results from tongue palate contact patterns in alveolar stops differ considerably, yet surprisingly, they have been given similar interpretations. Dagenais et al. [1] found a consistently greater midline length of contact for /d/ compared to /t/ averaged over data from 10 American English speakers. It was suggested that the voiced stops can be associated with a more relaxed tongue posture due to a lower intraoral pressure, whereas for /t/ the tongue may have greater tension in order to resist the air pressure. The opposite was found for 2 Norwegian speakers who had larger contact in /t/ than in /d/ (Moen et al. [7]). Moen et al. proposed that such differences would covary with intraoral pressure, which they recorded too. High intraoral pressure was associated with /t/ and more tongue-palate contact.
The aims of the current study are: first, to test the hypothesis that there are supralaryngeal differences in tongue and jaw kinematics associated with the voicing contrast for alveolar stops in German. Second, to test whether there are noticeable differences in tongue-palate contact patterns during oral closure. In German the /d/-/t/ distinction is based on phonation when the stop is located in word medial position, whereas in word final position the rule of final devoicing applies and /d/ becomes voiceless (Jessen [4]). To study different aspects of the contrast we will consider both positions.

2. METHOD

2.1. EXPERIMENTAL SETUP

In order to study supralaryngeal correlates of the voicing contrast we investigated tongue tip and jaw movements together with tongue-palate contact patterns by means of simultaneous EPG (Reading EPG3) and EMMA (AG 100, Carstens Medizinelektronik) recordings. Tongue tip movement was associated with the movement of a sensor placed midsagittally approximately 1 cm behind the tip. Jaw movement was associated with a sensor at the lower incisors (all other sensors placed on the tongue and the lips will not be considered here). Two sensors served as reference points to compensate for helmet movements, one at the nasion and one at the upper incisors. Speech signals were recorded on DAT. Sampling frequencies were 16 kHz for the acoustic data, 100 Hz for EPG and 200 Hz for EMMA data respectively. Four German subjects were recorded, three males (CG, DF, JD) and one female (SF). The speech material consisted of nonsense words “geCVC^{m}e” and “geCVC^{f}”, where C was either /t/ or /d/. Cm was defined as the consonant in the poststressed (i.e. the consonant occurred after the stressed vowel) word medial position and Cf as the consonant in the poststressed word final position. The stressed vowel preceding Cm or Cf was always tense /a/. Target words were embedded in the carrier phrase “Ich habe geCVCe nicht geCVC erwähnt.” (I said geCVCe not geCVC.). Each sentence was repeated 10 times. We should point out that the measured tongue tip sensor signals are composed of active tongue and jaw movements since decomposition is not a straightforward process.

 ISBN 1-876346-48-5 © 2003 UAB

2.2. LABELING CRITERIA AND DATA ANALYSIS

Acoustic data: Three time landmarks were labeled to calculate closure and noise duration: *tF2off* as the offset of the second formant in /a/, *tb* as the first burst and *tnoisoff* as the offset of high-frequency noise on the spectrogram. Closure duration *(clos)* was calculated as *tb-tF2off* and noise duration *(nois)* as *tnoisoff-tb*. *Nois* and *clos* could not be computed for the cases where burst or high frequency noise were missing.

EPG data: Regarding EPG data, 2 time landmarks were taken into account. Closure onset *tcloson* was defined as the time where at least 2 contacts were "on" in the most central region (column 3 to 6) of the anterior part (row 1 to 4) of the artificial palate. Closure offset *tclosoff* was defined as the last EPG pattern before oral release. For incomplete closure we chose the time point before one contact in the anterior region was missing. However, we also made a note, where this incomplete closure occurred (in the front or in the lateral region). From the tongue-palate contact patterns we computed 3 parameters: the percentage of contact in the anterior region (*ant*), in the posterior region (*post*) and the centre of gravity index (*cog*) following Hardcastle et al.[3]. The *cog* parameter is a weighted EPG index of the main concentration in the front–back dimension. The higher the *cog* value, the more anterior the sound is articulated.

Since we were especially interested in changes of tongue-palate contact patterns during oral closure, we compared /d/ and /t/ with respect to changes of the *ant, post* and *cog* parameters. Concerning changes in tongue-palate contact patterns we applied the following procedure: for each subject and each repetition the whole closure duration was normalised in time, and in a second step we over-sampled the values of the relevant EPG parameter so that for each parameter 10 values were computed during each normalised closure duration.

EMMA data: The movement of the closing gesture from the stressed vowel to the following stop was defined on the tangential velocity signal of the relevant sensor (tongue tip and jaw). We labeled 5 time landmarks: *ton, tlow1, tmax, tlow2* and *toff*. *Ton* and *toff* correspond to the velocity minima, *tmax* to the velocity peak, *tlow1* was defined using a 20% threshold criterion in the velocity signal (= the onset of the closing gesture) and *tlow2* was defined with a 20% threshold criterion too (= offset of the closing gesture).

Several temporal as well as spatial articulatory correlates from *tlow1, tmax* and *tlow2* were taken into account:

- the duration of the overall closing gesture (*clg_dur_tt* defined on the tongue tip tangential velocity signal and *clg_dur_j* defined on the jaw tangential velocity signal) as the difference between *tlow2* and *tlow1* of the relevant sensor,

- the movement amplitudes for tongue tip and jaw movements *(amp_clg_tt, amp_clg_j)* as the integral in the tongue tip closing gesture interval from *tlow1* to *tlow2* of the tongue tip tangential velocity signal and additionally in the EPG defined closure interval as the integral from *tcloson* to *tclosoff* of the tongue tip tangential velocity signal (*amp_clos_tt*) and of the jaw tangential velocity signal (*amp_clos_j*),

- the velocity peaks of the closing gesture for tongue tip (*peak_tt*) and jaw movements (*peak_j*),

- the x-values (horizontal dimension) and y-values (vertical dimension) for tongue tip and jaw sensors at the beginning of the tongue tip closing gesture *tlow1* and when the tongue tip reaches its consonantal target at *tlow2*

- the latencies between onset of tongue tip and jaw movements (on = tlow1_tt-tlow1_j) and offset of tongue tip and jaw movements (off = tlow2_tt-tlow2_j) of the closing gesture.

3. RESULTS

3.1. ACOUSTIC RESULTS

In order to test the significance of the results the General Linear Model (GLM) in SPSS version 11.0 was used at a .05 significance level. Closure and noise duration were used as dependent variables and consonant (i.e. /d/ versus /t/) as the independent variable, split by subject (CG, DF, JD, SF) and position (Cm, Cf).

In word medial position closure duration was significantly shorter for /d/ compared to /t/, whereas it did not differ in word final position. Similar to closure duration, noise duration was significantly longer for /t/ in word medial position. In most cases the measured noise duration in /d/ consisted of the duration of the weak burst signal. Surprisingly, for subject SF 8 out of 10 /d/'s were rather devoiced in word medial position.

In word final position three subjects (DF, JD, SF) do not show any significant differences with regards to noise duration, whereas CG showed a highly significant effect with longer noise duration for /t/. However, we assume that results from CG are in agreement with recent work on final devoicing and neutralisation with respect to regional variations of German (Piroth and Janker [8]). CG grew up near Lake Constance in southern Germany. Since Piroth and Janker show that Bavarian speakers do not totally neutralise voiced stops in word final position, it is most likely that subjects from another southern German region close to Bavaria produce a similar incomplete neutralisation.

3.2. RESULTS FROM EMMA

The following table provides an overview of the calculated supralaryngeal correlates which could participate in the voicing distinction. Again, we used the GLM in SPSS, where each of the movement amplitudes, velocity peaks, x- and y-positions were the dependent variables and consonant (/d/ versus /t/) the independent variable, split by subject and position.

Word medial position: Even though there was a consistent significant difference regarding acoustic closure duration, the duration of the relevant tongue tip closing gestures did not differ for 3 out of 4 subjects. However, jaw movement duration was shorter in /d/ for DF and SF. One could expect that a shorter articulatory closure duration in /d/ is accompanied by a smaller movement amplitude. This was true for JD and DF concerning tongue tip and jaw movement amplitudes during closure and jaw amplitudes for CG and SF, but did not hold for tongue tip amplitudes (CG, SF). Tongue tip velocity peaks were higher in /d/ than /t/ for CG and SF.

subject position	CG Cm/Cf	DF Cm/Cf	JD Cm/Cf	SF Cm/Cf
temporal				
clg_dur_tt	=/=	=/=	++/=	=/=
clg_dur_j	=/---	-/-	=/=	---/---
movement amplitudes				
amp_clg_tt	++/+	=/=	--/++	=/=
amp_clg_j	---/=	=/=	+/=	--/=
amp_clos_tt	+++/=	--/=	---/=	=/=
amp_clos_j	=/=	---/=	---/=	=/=
velocity peaks				
peak_tt	++/=	=/=	---/+	+/+
peak_j	---/=	=/=	=/+	=/=
y-positions				
tlow1_tt_y	-/-	=/=	---/---	-/=
tlow1_j_y	=/--	=/=	=/--	=/=
tlow2_tt_y	+++/=	---/=	---/--	=/=
tlow2_j_y	---/-	-/=	+++/--	--/---
x-positions				
tlow1_tt_x	-/=	---/=	--/-	--/=
tlow1_j_x	=/=	=/=	=/=	-/=
tlow2_tt_x	---/=	--/=	---/=	---/=
tlow2_j_x	=/=	=/=	+/=	-/=

Table 1: Results from GLM with respect to differences of /d/ and /t/, split by subject (CG, DF, JD, SF) and position (Cm = word medial, Cf = word final); "-" corresponds to /d/ which was smaller, shorter, lower or more retracted compared to /t/; "=" no significant difference; "+" opposite to "-"; "+++"or "---" $p<0.001$; "++" or "--" $p\leq0.01$; "+" or "-" $p\leq0.05$; for further explanation see text.

More consistent results can be seen for y- and x-positions. Tongue tip started with a lower (CG, JD, SF) and more retracted position (all) in /d/ compared to /t/, whereas jaw positions did not differ. At the consonantal target tongue tip still had a lower (DF, JD) and retracted position (all) and the jaw was lower too for CG, DF and SF. Since there seemed to be different involvement of tongue and jaw, we further calculated tongue-jaw latencies at the onset and offset of the closing gesture.

In word medial position it was found that tongue and jaw started the closing gesture relatively synchronously with some exceptions in JD's data. At the consonantal target JD and SF still exhibited relative synchrony for /d/, but not for /t/. In /t/, tongue tip reached its target first, followed by a delayed jaw (comparable with the example in figure 1).

Word final position: Although most of the articulatory characteristics analysed here were completely neutralised, some significant differences concerning jaw closing gesture duration (shorter for /d/), velocity peaks for tongue tip (higher for /d/), y-position at closing gesture onset for tongue tip and jaw (lower for /d/), and y-position at the consonantal target for the jaw (lower for /d/) maintained.

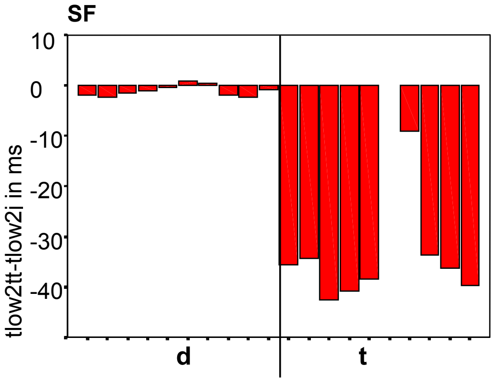

Figure 1: Example for SF's tongue-jaw coordination in ms at the consonantal target for /d/ (left) and /t/ (right); each bar corresponds to one token; negative values = jaw delay

Concerning tongue-jaw coordination at closing gesture onset, generally similar differences are found as in word medial position. At the consonantal target a considerable delay of the jaw was found for /t/, whereas in /d/ tongue and jaw are more synchronised (CG, SF, and a trend in DF; for an example see figure 1).

3.3. RESULTS FROM EPG

Electropalatographically defined closure duration showed similar significant effects to acoustically defined closure duration, but the first were slightly shorter. This could be a result of sampling frequency and labeling. Time normalised interpolated EPG results provide evidence that in most cases /d/ and /t/ were produced with differences in *ant, post* and *cog* in word medial position (Figure 2), whereas in word final position they were almost identical, hence we will focus on Cm.

The *ant* was significantly lower during the whole closure period in /d/ for DF and JD. For CG the /d/-/t/ difference lies rather in changes of the *ant* than in maximum contact. There was also a general trend in all subjects' /t/ production towards a small increase of *ant* at the beginning of the closure, followed by a relatively stable phase. Changes in /d/ are more speaker dependent. The *post* increased continuously during closure for /t/, whereas in /d/ it increased slightly less. A lower percentage of contact in the posterior region during /d/ closure was associated with cavity enlargement strategies. Differences of the *cog* were only marginal.

Additionally it was found, that /d/ was produced with incomplete anterior closure in Cm (2x for DF and 6x JD) and even with lateral opening in CG's data (all /d/ in Cm and 4x in Cf). Note, CG has the longest EPG palate. Concerning cavity enlargement strategies such incomplete closures should not be neglected since they do not permit increasing intraoral pressure.

ISBN 1-876346-48-5 © 2003 UAB

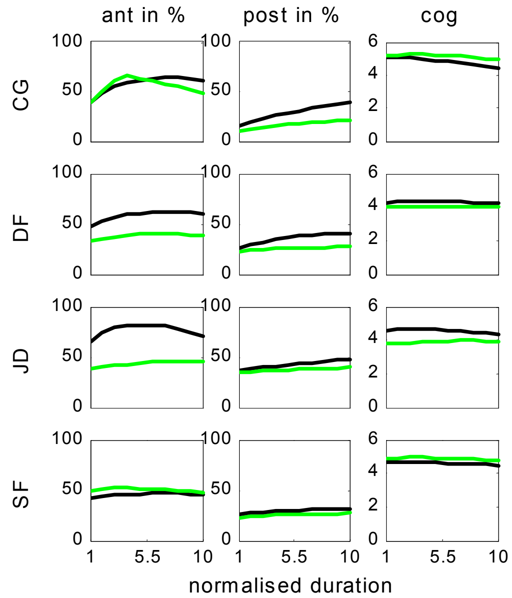

Figure 2: Interpolated ant, post, cog (from left to right) for CG, DF, JD, SF (from top to bottom) produced in word medial position, black line = /t/, grey (green) line = /d/

4. CONCLUSIONS

By means of simultaneous EMMA and EPG recordings we investigated supralaryngeal kinematics of alveolar /d/ and /t/ in order to search for articulatory correlates of the voicing contrast in word medial and final position. Generally speaking we found more consistency in the acoustic than in the articulatory results. In word medial position voiced stops were produced with a shorter closure duration. Articulatorily, speaker-dependent strategies did occur. Two subjects (JD; DF) produced the shorter closure in /d/ with shorter movement amplitudes which probably caused the smaller amount of anterior contact. One subject (CG) realised a very dynamic gesture for /d/ which resulted in a rapid increase in anterior contact and a slow decrease. Subject SF's results did not show significant effects with respect to anterior contact. Her differences in closure duration were rather an effect of tongue-jaw coordination, i.e. a synchronisation for /d/ and a jaw delay for /t/. Additionally, tongue tip started lower (3 subjects), and more retracted (all), and ended more retracted (all). These results are interpreted with regard to the longer vowel duration often found before voiced stops. The lower jaw position at the consonantal target for /d/ was assumed to increase the oral cavity actively to prevent quickly rising intraoral pressure. The higher jaw position in /t/ can be interpreted as a strategy to increase intraoral pressure and to provide a close constriction for the salient burst (Mooshammer et al. at this conference [10]).

We conclude that in word final position neutralisation of devoiced /d/ in comparison to /t/ is not complete regarding supralaryngeal kinematics. In particular the jaw has a higher position in /t/ at the consonantal target and tongue-jaw coordination can differ as well. The two articulators are highly synchronised in /d/ and less in /t/. For /t/ the jaw still moved when the tongue had already reached its velocity minimum at the consonantal target. However, at the same time results from acoustics showed complete neutralisation in 3 out of 4 subjects. Hence, we will further investigate perceptual correlates of the observed articulatory differences in the auditory and visual domains.

ACKNOWLEDGMENTS

This work was supported by a grant from the German Research Council GWZ 4/8-1, P.1. Thanks to C. Mooshammer, P. Hoole, B. Pompino-Marschall and F. Windsor for helpful comments to this paper. We thank J. Scobbie for long discussions about variances and invariances.

REFERENCES

[1] P.A. Dagenais, L.C. Lorendo & M.J. McCutcheon, "A study of voicing and context effects upon consonant linguapalatal contact patterns", *Journal of Phonetics 22*, pp. 225-238, 1994.

[2] O. Fujimura & J.E. Miller, "Mandible height and syllable-final tenseness", *Phonetica 36*, pp. 263-272, 1979.

[3] W.J. Hardcastle, F. Gibbon & K. Nicolaidis, "EPG data reduction methods and their implications for studies of lingual coarticulation", *Journal of Phonetics*, vol. 19, pp. 251-266, 1991.

[4] M. Jessen, *Phonetics and phonology of tense and lax obstruents in German,* Amsterdam/Philadelphia: John Benjamins Publishing Company, 1998.

[5] R.D. Kent & K.L. Moll, "Vocal-tract characteristics of the stop cognates" *Journal of the Acoustical Society of America 46(6),* pp. 1549-1555, 1969.

[6] A. Löfqvist & V.L.Gracco, "Tongue body kinematics in velar stop production: influences of consonant voicing and vowel context", *Phonetica,* vol. 51, pp. 52-67, 1994.

[7] I. Moen, H.G. Simonsen, M. Huseby, & J. Grue, "The relationship between intraoral air pressure and tongue-/palate contact during the articulation of Norwegian /t/ and /d/", Eurospeech CD-ROM. paper 265, 2001.

[8] H.G. Piroth & P.M. Janker, "Speaker-dependent differences in voicing and devoicing of German obstruents", *Journal of Phonetics*, (in press).

[9] J.R. Westbury, "Enlargement of the supraglottal cavity and its relation to stop consonant voicing", *Journal of the Acoustical Society of America 73(4)*, pp. 1322-1336.

[10] C. Mooshammer, A. Geumann, P. Hoole, P. Alfonso, P. van Lieshout, S. Fuchs, "Coordination of lingual and mandibular gestures for different manners of articulation", (this conference).