Temporal development of compensation strategies for perturbed palate shape in German /sch/-production

Jana Brunner, Phil Hoole, Pascal Perrier, Susanne Fuchs

To cite this version:

Jana Brunner, Phil Hoole, Pascal Perrier, Susanne Fuchs. Temporal development of compensation strategies for perturbed palate shape in German /sch/-production. 7th International Seminar on Speech Production, Dec 2006, Ubatuba, Brazil. pp.247-254. hal-00403289

HAL Id: hal-00403289
https://hal.science/hal-00403289

Submitted on 10 Jul 2009

Temporal development of compensation strategies for perturbed palate shape in German /ʃ/-production

Jana Brunner[1 2 3*], **Phil Hoole**[4†], **Pascal Perrier**[2] , **Susanne Fuchs**[3]

[1]Humboldt-Universität zu Berlin,

[2]Institut de la Communication Parlée, INPG, CNRS & Univ. Stendhal Grenoble

[3]Zentrum für Allgemeine Sprachwissenschaft, Berlin

[4]Institut für Phonetik und Sprachliche Kommunikation der Universität München

`brunner@zas.gwz-berlin.de`

Abstract. *The palate shape of four speakers was changed by a prosthesis which either lowered the palate or retracted the alveoles. Subjects wore the prosthesis for two weeks and were recorded several times via EMA. Results of articulatory measurements show that speakers use different compensation methods at different stages of the adaptation. They lower the tongue immediately after the insertion of the prosthesis. Other compensation methods as for example lip protrusion are only acquired after longer practising periods. The results are interpreted as supporting the existence of different mappings between motor commands, vocal tract shape and auditory-acoustic target.*

1. Introduction

Since there is a many-to-one relation between between articulation and acoustics, there are sometimes several articulatory strategies available to produce a certain sound (motor equivalence). An example for that is the American English /r/ (e.g. Westbury *et al.* (1998)), which can be produced as ”bunched r” with a lamino-postalveolar constriction or as the ”retroflex r”. Another example is /u/ which can either be produced with a constriction at the lips together with another one in the velar region or without a constriction at the lips and a pharyngeal constriction (Savariaux *et al.* (1995), Savariaux *et al.* (1999)).

Observations like these support the hypothesis of auditory-acoustic targets (e.g. Guenther *et al.* (1998)). This means that what is represented in the speakers brain are the auditory or acoustic characteristics of a sound. During speech development, speakers acquire such a target for each phoneme of the language.

---

[*] Supported by the Deutsche Forschungsgemeinschaft, grant PO 334/4-1

[†] Supported by the Deutsche Forschungsgemeinschaft, grant HO 3271/1-1

However, in order to produce speech, this auditory-acoustic target needs to be linked to other domains (other *reference frames* as they are called in the motor control domain (see for example Guenther *et al.* (1998)) involved in speech production , for example the state of the articulators which produce a certain auditory-acoustic target or the motor commands needed to produce a certain articulator configuration. Coming back to the example of the American English /r/, the speaker has a certain auditory-acoustic target, but, in case he uses both ways of production, he has two mappings between articulation and auditory-acoustic target.

As long as the articulation is not perturbed, each set of motor commands is linked to a certain shape of the vocal tract (Boë *et al.* (1992)). If the normally fixed borders of the vocal tract are changed, however, for example due to a dental device, the old sets of motor commands are not linked to the same vocal tract shapes any longer. Speakers therefore have to acquire a new mapping between motor commands and constriction in order to reach the same target.

Speakers usually have experience in speaking in perturbed conditions. For example, most speakers manage to reach the targets of the phonemes rather well while smoking or eating at the same time as they speak, or while wearing dental devices. There are several possible explanations for why speakers manage to do this. They might have acquired a new mapping between motor commands and vocal tract shape. In this case they would try to produce the same vocal tract shape with different motor commands. For instance, a speaker wearing a dental device lowering the palate might lower the tongue in order to reach the same degree of constriction. Another possibility would be a different mapping between the articulation and the acoustic target. Using the fact that there are several possibilities to reach the same acoustic output a speaker might for example lower the larynx if he happens to be unable to produce lip protrusion during the production of /u/ while smoking. This compensation strategy has been found by Riordan (1977)[1].

Looking at the temporal development of adaptation for perturbed articulation might provide information about what kind of mapping is changed. At first speakers will probably try to produce the same vocal tract shape by using different motor commands. Only after having tried out several vocal tract shapes and their relation to the acoustic output they will find new ways of articulation which produce a similar acoustic output. At this point in time, they will have acquired two new mappings, one between motor commands and vocal tract shape and another one between vocal tract shape and auditory-acoustic target.

In the experiment which is going to be described in this article, the palate shapes of speakers were perturbed during the production of initial and medial /ʃ/ (surrounded by /a/) for a period of two weeks. Speakers were expected to change their compensation strategies over this time. First they would try to produce the same vocal tract shape and thereby continue using the old mapping between vocal tract shape and auditory-acoustic target. Only later they would develop a new mapping and produce different vocal tract shapes leading to the same acoustic output.

---

[1] Tuller and Fitch (1980), however, contradict this finding.

2. Methods

For four speakers[2] palate prostheses were made of acryl. For two of them (speakers TP and KD) these prostheses moved the alveoles to the back, for the other two the palatal vault was filled out so that the palate became flatter (speakers OP and SK). Subjects wore the palates for 14 days and were recorded via Electromagnetic articulography (Carstens AG 100) regularly over this period under different conditions:

- Day 1:
    - session 1/1: without artificial palate
    - session 1/2: with artificial palate, with auditory feedback masking due to white noise over headphones (session missing for two speakers)
    - session 1/3: with artificial palate with full auditory feedback
- Day 8:
    - session 2/1: with artificial palate and full auditory feedback
- Day 15:
    - session 3/1: with artificial palate and full auditory feedback
    - session 3/2: without artificial palate and full auditory feedback

The target sound /ʃ/ was embedded in the nonsense words /'ʃaxa/ and /'daʃa/ which were produced in the carrier phrase *Ich sah ... an.* (I looked at ...). The sentences were repeated 20 times per session in randomised order[3].

Three sensor coils of the AG 100 were placed on the tongue, one below the lower incisors in order to track jaw movements, one at the upper lip and one at the lower lip. Two coils at the upper incisors and the bridge of the nose served as reference sensors to compensate for head movements. A parallel acoustic recording was carried out on a DAT recorder.

Beginning and end of the target sound (friction onset and offset) were labelled in the acoustic signal using PRAAT 4.2.17. For the analysis of the articulatory data sensor positions at the articulatory targets (highest point of the tongue tip) for the different subsessions were calculated. Figure 1 shows the results of this step for the medial /ʃ/ of one speaker with an alveolar prosthesis. Each subplot represents the subsession which is given in the header of the subplot (see list above for explanation of subsessions). The solid lines above the markers represent the palate contour used in this session. The sensor denotations are given in the first subplot only: *ttip*: tongue tip, *tdor*: tongue dorsum, *tback*: tongue back, *jaw*: jaw, *llip*: lower lip, *uplip*: upper lip, *upinc*: upper incisors. For one repetition of each session the tongue sensors are connected by straight lines in order to demonstrate further steps of analysis.

Comparing the vertical position, the tongue is a bit higher in the unperturbed than in the perturbed condition (compare e.g. 1/1 and 1/2). Furthermore, there are differences in constriction width, e.g. at the tonge tip sensor. If one compares the horizontal difference between the sensor of the upper incisor and the one at the tongue tip in sessions 3/1 and 3/2 one can notice that in 3/1 the tongue is a bit retracted. Comparing the difference in the position of the lower lip sensor and the jaw sensor in 3/1 with the difference between the

---

[2] Results for additional four speakers will be presented at the conference.

[3] together with other items which are not analysed here

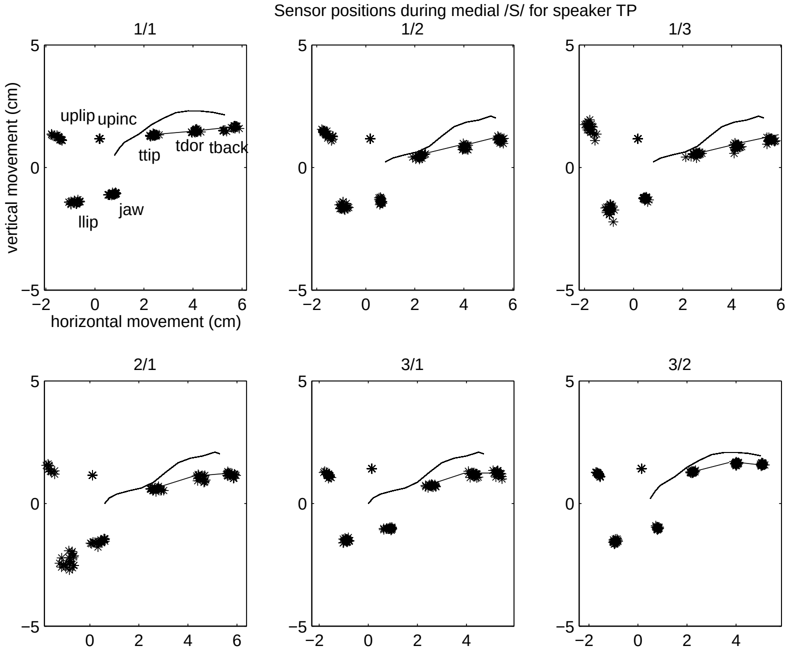

Figure 1. Sensor positions for 20 repetions per session for one speaker with an alveolar prosthesis (medial /ʃ/). See text for further details.

two sensors in 2/1 could suggest that there is some more lip rounding in 3/1 than in 2/1. Furthermore, the overall sloping of the tongue changes. Whereas the three tongue sensors are more or less at equal vertical position in 1/1, the tongue back sensor has a higher position in comparison to the tongue tip sensor in session 1/2. If one looks at the angle between the three tongue tip sensors (marked by the lines connecting the tongue sensor markers), one can see that it is smaller than 180 degrees in 3/1 (the tongue is bunched) but around 180 degrees in 1/1 (the tongue is flat).

Following these observations, it was assumed that six parameters were involved in the adaptation process: a) vertical position of the tongue in relation to the natural vocal tract shape, b) constriction size or vertical position of the tongue in relation to the palate which was used in the session, c) horizontal position of the tongue, d) lip protrusion, e) slope of the tongue and f) bunching of the tongue. In order to capture these parameters numerically, the following calculations have been carried out for each repetition:

Vertical position of the tongue in relation to the natural palate. Since the sensor at the upper incisors could be assumed to be placed at the same position during each recording, this sensor was taken as a reference. A mean of the y-coordinates of the three tongue sensors was calculated, afterwards the difference between this value and the y-coordinate of the sensor at the upper incisors was calculated.

Constriction size. For the second parameter at first the palate contours, natural and artificial, were estimated. In order to do so, the midsagittal coordinates of the palate were measured from the prostheses (for the artificial palate) and the dental cast (for the natural contour). Afterwards, all positions of the tongue which had been recorded during a certain subsession were plotted and the measured palate contours were fitted to the contour which appeared as the upper border of the plotted positions. Finally, the difference between the y-coordinate of each tongue sensor and the corresponding point at the palate (sharing the x-coordinate with the tongue sensor position) was calculated.

Horizontal position of the tongue. The horizontal position was most difficult to measure. This has to do with the experimental procedure. For each recording day, the sensors were glued to the tongue anew. The tongue tip sensor was placed around one centimeter behind the tongue tip, the tongue back sensor at the place were the tongue touches the border between palate and velum in rest position, and the tongue dorsum sensor halfway in between the two. However, in contrast to, for example, the sensor at the upper incisors which is placed exactly midsagittally above the incisors, it is hard to find landmarks on the tongue which enable the experimenter to find the same position for the sensors at different days. We therefore did not rely on single tongue sensors but calculated a mean of the x-values of the three in order to reduce the influence of one slightly differently placed sensor. Afterwards, the difference between this mean value and the x-value of the sensor at the upper incisors was calculated.

Lip protrusion. Lip protrusion was estimated as the difference in the x-coordinates of the lower lip sensor and the jaw sensor. For both sensors landmarks existed so that they were glued to exactly the same position during the three experiments. The jaw sensor was placed below the lower incisors and the lower lip sensor at the lower border of the lower lip.

Slope of the tongue. This parameter gives information about whether the tongue is rather horizontal in the mouth (slope around 0) or sloped to the front (slope is positiv) or to the back (slope is negativ). In order to calculated this parameter the ascent of the line connecting the position of the tongue tip sensor and the tongue back sensor was calculated.

Bunching of the tongue. In order to estimate whether the shape of the tongue is rather flat or bunched, the angle between the three tongue sensors was calculated (method adapted from Tiede *et al.* (2005)). If this angle is close to 180 degrees the tongue is rather flat, if it is smaller, the tongue is bunched.

3. Results

If speakers at first try to reach a configuration similar to the vocal tract shape they used to have before the perturbation and only later develop strategies involving different vocal tract shapes which result in similar acoustic outputs, one can expect that some of the above parameters will change immediately whereas others might take longer. Since the prosthesis makes the vocal tract narrower, the vertical position of the tongue should change immediately in order to reach the same degree of constriction. For speakers with an alveolar prosthesis the alvoles are moved to the back. In response to that these speakers should retract the tongue immediately in order to preserve the shape of the constriction. For the other speakers with the flattening prosthesis there should be an immediate change in the shape of the tongue (less bunching). The other parameters (lip protrusion, degree of constriction, slope) can be expected to change later since they create a new vocal tract shape.

Figures 2 and 3 present the results for the six articulatory parameters.

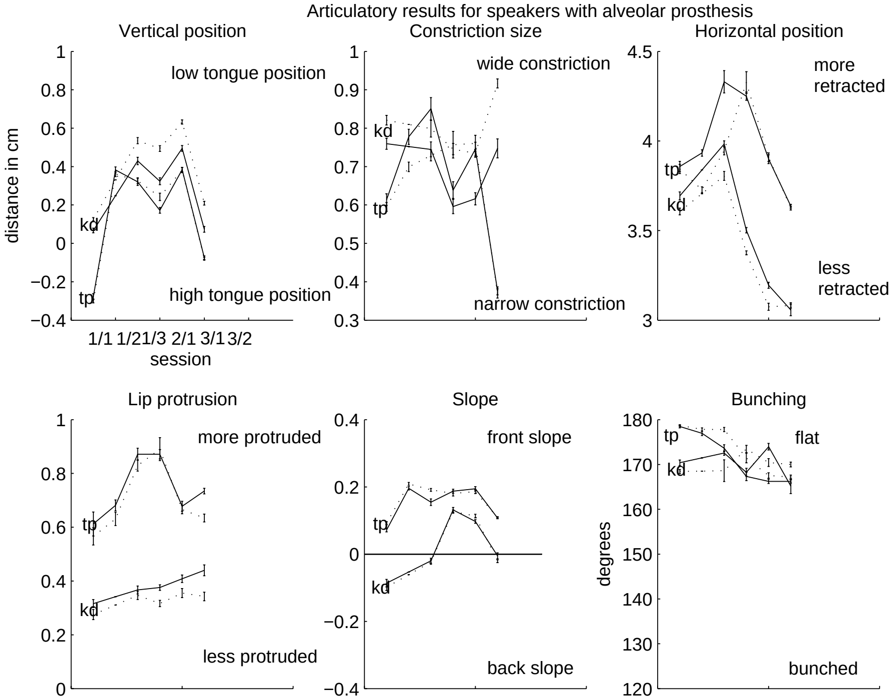

Figure 2. Results for the six compensation parameters for the speakers with an alveolar prosthesis. Solid lines represent the results for the medial /ʃ/, dashed lines the ones for the initial /ʃ/. Error bars show standard error.

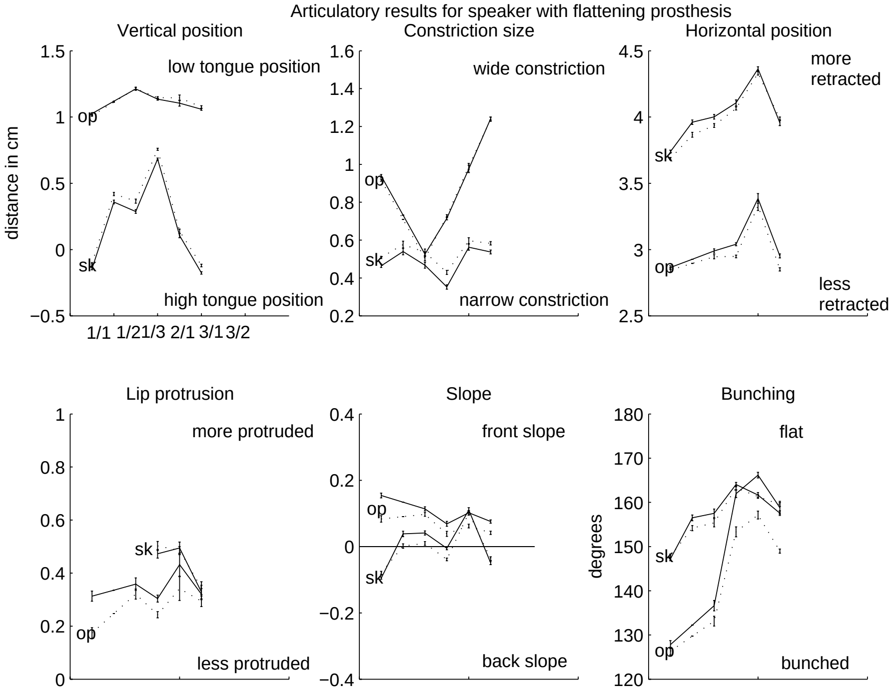

Figure 3. As figure 2, but for speakers with a flattening prosthesis

Vertical position of the tongue. Independently of the prosthesis, all speakers significantly lower the tongue in relation to the natural vocal tract in the perturbed condition (p=0.000). Furthermore, the change takes place immediately after inserting the prosthesis.

Constriction size. The constriction size changes rather individually. Speaker KD makes the constriction a little narrower up to session 3/1, speaker OP immediately produces a very narrow constriction which becomes wider and wider, speaker SK first has a wider constriction, which becomes very narrow in session 2/1 and wider again in 3/1, speaker TP has constrictions which are a little wider in the perturbed condition than in the unperturbed. Since speakers show different degrees of constriction one can suppose that they are trying to improve the productions all through the two weeks. The differences between the perturbed and unperturbed condition are significant for speakers TP, OP and KD.

Horizontal position of the tongue. In general, speakers have a tendency to retract their tongue. However, the expected difference between the two kinds of prostheses can be found. Whereas the speakers with the alveolar prosthesis retract the tongue very soon after insertion of the prosthesis (in order to preserve the degree of constriction), the other two speakers develop more and more tongue retraction throughout the time of adaptation. The differences between the sessions are significant except for session 3/1 vs. 3/2 for speaker KD, 1/1 vs. 3/2, 1/3 vs. 3/2 and 1/3 vs. 2/1 for speaker OP, 1/2 vs. 1/3 vs. 3/2 for speaker SK, 1/1 vs. 1/2 vs. 3/1 for speaker TP.

Lip protrusion. In general, speakers have a tendency to develop more and more lip protrusion until (except for session 3/1 for speaker TP). For speaker KD only the first session differs significantly from sessions 1/3-3/2. For speaker TP there are significant differences for session 1/1 and 1/2 vs. 1/3 and 1/4 and vs. 1/5 and 1/6. Similarly for speaker OP the differences between 1/1, 1/3 and 3/1 are significant. For speaker SK there are no results for sessions 1/1 to 1/3 because of a technical problem, but for the remaining sessions the differences between the perturbed and the unperturbed condition are significant. The development is for all speakers rather slow and consistent over the sessions.

Slope of the tongue. Speakers TP, KD and SK produce a tongue shape which is sloped more to the front in the perturbed than in the unperturbed condition. Speaker TP acquires this sloping immeditately after the insertion of the prosthesis, speakers KD and SK develop it later. Speaker OP starts with a tongue which is sloped to the front in the unperturbed condition and develops a very horizontal position which can be explained by the flattening palate shape. Except for OP, who has a strong after effect, the differences between the perturbed and the unperturbed condition are significant.

Bunching of the tongue. The speakers with the flattening prosthesis develop less bunching immediately after the insertion of the prosthesis. The other speakers, however, do not change the bunching very much and also not immediately. For speaker OP sessions 1/1, 1/3 and 2/1 differ significantly, for speaker SK session 1/1 differs from 1/3 and those two from 2/1 significantly. For speaker KD only sessions 3/1 and 3/2 differ significantly. For speaker TP session 1/1 differs significantly from 1/3, those two from 2/1 and those three from the two last sessions.

4. Discussion: Temporal development of the parameters

All the parameters have been found to be involved in the adaptation. However, there are clear differences in the temporal development. Whereas parameters which help to develop a vocal tract shape similar to the one of the unperturbed condition are changed immediately (vertical tongue position, bunching for the speakers with the flattening prosthesis, horizontal position for the speakers with the alveolar prosthesis), changes in the other parameters need longer to develop. An explanation for this late development could be that a new mapping between vocal tract shape and acoustic output is involved which can only be acquired by trying out several vocal tract configurations and therefore needs practise.

Acknowlegements: This work was supported by the ”Programme de Recherche en Rseaux Franco-Allemand” (”POPAART” Project) funded by the CNRS and the French Foreign Office.

References

Boë, L., Perrier, P., and Bailly, G. (**1992**). “The geometric vocal tract variables controlled for vowel production: proposals for constraining acoustic-to-articulatory inversion”, Journal of Phonetics **20**, 27–38.

Guenther, F. H., Hampson, M., and Johnson, D. (**1998**). “A theoretical investigation of reference frames for the planning of speech movements”, Psychological Review **105**, 611–633.

Riordan, C. (**1977**). “Control of Vocal Tract Length in Speech”, Journal of the Acoustical Society of America **62**, 998–1002.

Savariaux, C., Perrier, P., and Orliaguet, J.-P. (**1995**). “Compensation strategies for the perturbation of the rounded vowel [u] using a lip-tube: A study of the control space in speech production”, Journal of the Acoustical Society of America **98**, 2428–2442.

Savariaux, C., Perrier, P., Orliaguet, J.-P., and Schwartz, J.-L. (**1999**). “Compensation strategies for the perturbation of the rounded vowel [u] using a lip-tube II. Perceptual analysis”, Journal of the Acoustical Society of America **106**, 381–393.

Tiede, M. K., Gracco, V., Shiller, D. M., Espy-Wilson, C., and Boyce, S. E. (**2005**). “Perturbed palatal shape and north american english /r/ production (Presentation at the 149th Meeting of the Acoustical Society of America)”, .

Tuller, B. and Fitch, H. L. (**1980**). “Preservation of vocal tract length in speech: A negative finding”, Journal of the Acoustical Society of America **67**, 1068–1071.

Westbury, J., Hashi, M., and Lindstom, M. (**1998**). “Differences among speakers in lingual articulation of american english /r/”, Speech Communication **26**, 203–226.