The use of sensory feedback in the adaptation of perturbed /s/

Jana Brunner[1,2,3,4], Phil Hoole[4] & Pascal Perrier[3]

[1] Humboldt-Universität zu Berlin, [2] ZAS Berlin,
[3] DPC/Gipsa-lab, INP Grenoble, CNRS & Université Stendhal,
[4] Institut für Phonetik und Sprachverarbeitung München
E-mail: brunner@zas.gwz-berlin.de

Abstract

The study investigates the contribution of tactile and auditory feedback in the adaptation of /s/ in presence of a palatal prosthesis. Five speakers were recorded via EMA (1) without the prosthesis, (2) with the prosthesis and auditory feedback masked, and (3) with the prosthesis and auditory feedback available. The results show that the initial adaptation attempts without auditory feedback are dependent on the prosthesis type and directed towards reaching the original tongue palate contact pattern. In a second adaptation step with auditory feedback available speakers reorganised tongue and jaw movements in order to produce more subtle acoustic characteristics of the sound such as the high amplitude noise which is typical for sibilants.

1   Introduction

Previous work on perturbation has shown that sensory (auditory and tactile) feedback is essential in order to adapt for a perturbation of the articulation. Jones & Munhall [2], for example, show that speakers need auditory feedback in order to adapt the acoustic characteristics in /s/ produced with extended upper incisors. Honda & Murano [1] tested the influence of both auditory and tactile feedback during adaptation to an inflatable palate. They report the best results in adaptation when both kinds of feedback were available, less good results when auditory feedback was masked, even worse results when tactile feedback was masked and the worst results when both kinds of feedback were unavailable.

The present study investigates the influence of auditory feedback on adaptation when speech is perturbed with different palatal prostheses. Our expectation was that when the auditory feedback is masked and speakers have only tactile feedback available, adaptation will depend on the prosthesis type, since speakers will try to achieve the tongue palate contact pattern they used to have in the unperturbed condition. When both kinds of feedback are available, speakers will change the adaptive strategy and try to produce the acoustic output (spectrally characterised by the same acoustic centre of gravity) they used to have, even if this involves a change of the tongue palate contact pattern.

2   Methods

A perturbation experiment was carried out where speakers' palate shape was modified with a palatal prosthesis. Two types of palatal prostheses were used, one which lowered the palate and moved the alveolar ridge posteriorily ("alveolar prosthesis"), and one which lowered the palatal surface by filling out the palatal vault ("central prosthesis"). Five German subjects took part in the study. Three of them, TP, AP and

DS had an alveolar prosthesis, the other two, BP and SK, had a central prosthesis. The articulatory movements of the speakers were recorded via electromagnetic articulography (three speakers with the AG 100 and two with the AG 500). Sensors were placed midsagittally on the tongue tip, tongue dorsum and tongue back, jaw and both lips. For the present purpose the data of the tongue tip and the jaw sensor were analysed.

In a first session speakers were recorded without the perturbation (to be called *unperturbed condition*). Then the artificial palate was inserted and speakers' auditory feedback was masked with white noise (to be called *white noise condition*). Afterwards, the masking noise was removed and speakers were recorded while adapting with auditory feedback available (to be called *auditory feedback condition*). /s/ was recorded in the nonsense word /ˈzasa/ which was spoken in a carrier phrase: *Ich sah sassa an* ("I looked at zassa."). There were 20 repetitions (randomised with other material) in each session. Acoustic recordings were carried out as well.

The fricative was acoustically segmented (friction onset to friction offset) in each utterance. The acoustic centre of gravity (COG) for a band pass filtered signal (700 Hz to 12 kHz) was measured. Furthermore, the horizontal position of the tongue tip sensor and the vertical position of the jaw sensor in the middle of the acoustically measured consonant interval were estimated. In order to judge whether the measured parameters differed significantly across sessions repeated measures ANOVAs were calculated for data split by speaker.

3   Expectations

Speakers with an *alveolar* prosthesis were expected to retract the tongue in the white noise condition, in order to keep the alveolar place of articulation. This would result in lower COG values due to a lengthening of the front vocal tract cavity. In the session thereafter (with auditory feedback available) these speakers

were expected to adapt the tongue position so that the original COG would be reached.

In contrast to that, speakers with a *central* prosthesis were expected not to retract the tongue, since their alveolar ridge was not moved posteriorily by the prosthesis. The centre of gravity should therefore stay the same in all sessions for these speakers.

Both groups of speakers were expected to lower the jaw when speaking with the prosthesis in order to avoid a complete vocal tract closure.
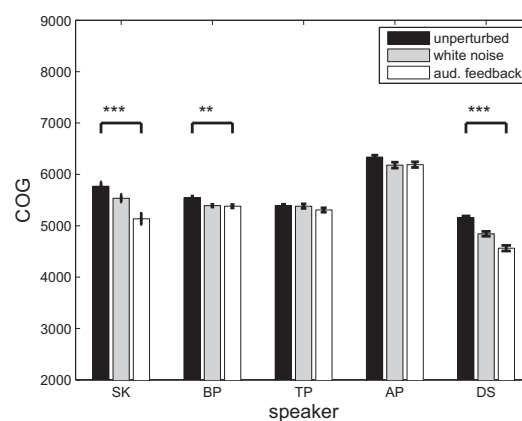
4   Results

Figure 1: *Mean COG values in the different sessions. Error bars show standard error. When the difference between the unperturbed and the auditory feedback session is significant this is signalled by a bracket above the bars. Significance levels: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.*

Acoustic centre of gravity. The results for the measurements of the centre of gravity are shown in figure 1. Each bar shows the mean over one session. Bars belonging to the same triple refer to one speaker and they are in the order of the recordings: the first (black) bars of a triple correspond to the unperturbed session, the second (grey) ones to the white noise condition and

the third (white) ones to the auditory feedback condition. The two leftmost speakers (SK and BP) had central prostheses, the other three had alveolar prostheses.

For four speakers (SK and BP with a central prosthesis and AP and DS with an alveolar prosthesis) the COG is lower in the white noise condition than in the unperturbed condition. For three speakers (SK with a central prosthesis and TP and DS with an alveolar prosthesis) the COG is lower when auditory feedback is available than when it is not. The difference between the unperturbed and the auditory feedback condition are significant for three speakers.
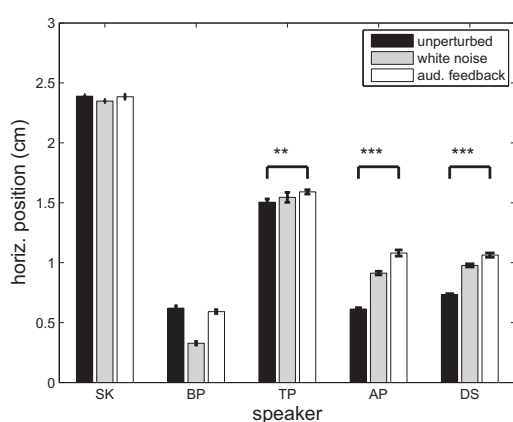

Figure 2: *As figure 1 but showing mean horizontal position of the tongue tip. Higher values denote more retracted positions.*

Horizontal tongue position. Figure 2 shows the results for the horizontal tongue position. Again the abscissa gives the speaker, the first bar of a triple refers to the unperturbed session, the second to the white noise condition and the third to the auditory feedback condition. The ordinate gives the horizontal position of the tongue tip sensor. Higher values represent a more retracted position.

The two speakers with a central prosthesis (SK and BP) protrude the tongue in the white noise condition but correct for this change afterwards. For the three speakers with an alveolar prosthesis (TP, AP, DS) a

consistent retraction of the tongue tip over the three sessions was found (but only one of those speakers, DS, also has a consistent decrease in COG across the three sessions.

Vertical jaw position. Figure 3 shows the results for the vertical jaw position. Higher values signify a higher jaw position, lower values a lower one. For all speakers the jaw position is lowered when the prosthesis is inserted, but raised again when auditory feedback becomes available. This difference between the white noise and the auditory feedback condition is significant for all speakers. Speakers with a central prosthesis seem to raise the jaw a little less than speakers with an alveolar prosthesis.
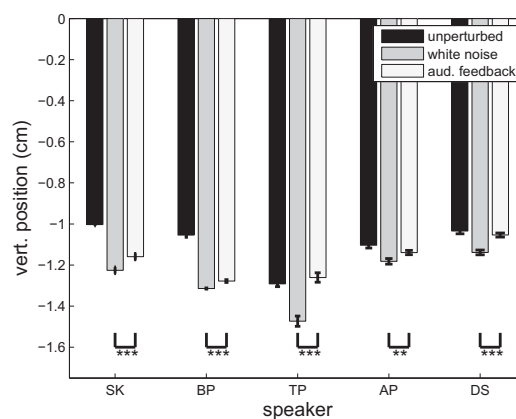

Figure 3: *As figure 1 but showing mean vertical position of the jaw sensor at the consonantal target position in /s/. Higher values denote higher positions. When the difference between the white noise condition and the auditory feedback condition is significant this is signaled by a bracket above the bars.*

5   Summary of the results

As expected, we found a difference in the adaptive behaviour of speakers with different prosthesis types: Speakers with an alveolar prosthesis retracted

the tongue in the white noise condition whereas speakers with a central prosthesis protruded it. However, a correction of this positional change in the auditory feedback condition could only be found for speakers with a central prosthesis.

Furthermore, the expected relation between COG and horizontal position of the tongue could not be found. For a retracted tongue the COG did not always decrease, nor did it increase consistently for a protruded tongue. This suggests that the COG is affected by a further mechanism apart from the constriction position.

As expected, independent of the prosthesis type the jaw was lowered for all speakers when the prosthesis was inserted. Contrary to our expectations, however, it was raised in the third session.

6   Discussion

Even if the results thus match our expectations only partly, they can be interpreted in the following way. When speakers insert the prosthesis (white noise condition) they might try to keep the same global articulatory position in reference to the palate: For the speakers with an alveolar prosthesis the tongue is retracted because the alveolar ridge is retracted; for the other speakers the tongue is protruded probably as a mechanical effect of the prosthesis.

When auditory feedback becomes available speakers' first aim should be to check the acoustic output they produce and to correct it. However, no adaptation of the COG can be found. Hence COG could be an inappropriate parameter to characterize the perception of the fricatives, which was already suggested by Jones & Munhall [2].

However, the adaptation efforts could be directed towards another, more prominent acoustic characteristic of the sound. /s/ is a sibilant, and an important acoustic characteristic of sibilants is the high amplitude noise produced by directing the air jet against an obstacle (the incisors) in front of the constriction (Shadle [4], Shadle [5]). This, however, necessitates a high jaw

position, otherwise the distance between the incisors is too large and no turbulences are created (Mooshammer et al. [3]). This higher jaw position can in fact be found in the auditory feedback condition. When auditory feedback becomes available speakers can thus be assumed to notice that they are lacking the high amplitude noise typical for /s/, and this might lead them to raise the jaw. The retraction of the tongue which can be found in the auditory feedback condition could be seen as a consequence of the higher jaw position: When the jaw is high the space available for the tongue in the alveolar region can be assumed to be no longer sufficient so that the tongue is moved to a more retracted position where there is more space.

To conclude, our results show that certain properties of the sound, such as the constriction size, can be adapted by using tactile feedback only. More subtle acoustic characteristics of /s/, however, require a reorganisation of the articulatory strategy, i.e. the positional relation of tongue and jaw has to be changed. For these articulatory changes auditory feedback becomes essential, as shown by the changes observed in the third session.

References

[1] M. Honda and E. Murano. Effects of tactile and auditory feedback on compensatory articulatory response to an unexpected palatal perturbation. *Proceedings of the 6th International Seminar on Speech Production*, 2003.

[2] J. Jones and K. Munhall. Learning to produce speech with an altered vocal tract: The role of auditory feedback. *Journal of the Acoustical Society of America*, 113(1):532–543, 2003.

[3] C. Mooshammer, P. Hoole, and A. Geumann. Interarticulator cohesion within coronal consonant production. *Journal of the Acoustical Society of America*, 120(2):1028–1039, 2006.

[4] C. Shadle. The acoustics of fricative consonants. Technical Report 506, Research Laboratory of Electronics, MIT Cambridge, MA., 1985.

[5] C. Shadle. The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics*, 19:409–424, 1991.