Lund University, Dept. of Linguistics
Working Papers 49 (2001), 108-111

Identifying places of articulation in coronal consonants: a combined EPG and listening test

Peder Livijn and Olle Engstrand
Department of Linguistics, Stockholm University
peder@ling.su.se, olle@ling.su.se

Abstract

A listening test was carried out to explore Swedish listeners' capability of distinguishing place contrasts in the dental/alveolar region. The speech material consisted of recordings of /t d n l/ in a V_V context, during which place of articulation was shifted continuously from dental to postalveolar. 'Dental', 'alveolar' and 'postalveolar' stimuli were selected on the basis of simultaneous EPG records. Whereas listeners identified the dental and post-alveolar stimuli well above chance – but with a considerable degree of individual variation – there was a marked uncertainty about the alveolar stimuli. In consequence, the results do not allow the conclusion that Swedish listeners are capable of identifying more than two places of articulation for coronals in a cross-language or cross-dialectal speech material.

1 Introduction

For the purpose of labelling large speech databases, such as the IRIS language database or the SWEDIA dialect database, it is of some interest to determine to what extent variable productions of 'the same' phonological category can be reliably detected auditorily. For coronal consonants, for example, a number of languages of the world, e.g., Malayalam, feature a three-way contrast in the dental/alveolar region, involving both place of articulation, extension of contact area and tongue shape (Dart & Nihalani 1999). For another example, observations of the Swedish dialects suggest the existence of systematic, dialect-dependent shifts in place and also tongue shape in coronal consonants such as /t/, /d/, /n/ and /l/ (Livijn & Engstrand 2001).

It is reasonable to assume that native speakers of a language with a dental/alveolar contrast should be relatively sensitive to place shifts in the dental/alveolar region, at least to the extent that it is of an order of magnitude used in the language. Even though Swedish does not display an underlying dental/alveolar contrast such as that of Malayalam, sequences of /r/+coronal are usually realised as more or less retroflex alveolars in Central Standard Swedish and many regional varieties of the language. This assimilation effect appears within and across words and morphemes, giving rise to minimally contrasting pairs such as *våta* [vo:ta] 'wet' vs. *vårta* [vo:ʈa] 'wart'; *modet* [mu:det] 'the courage' vs. *mordet* [mu:ɖet] 'the murder'; *vana* [vɑ:na] 'habit' vs. *varna* [vɑ:ɳa] 'warn'; *påla* [po:la] 'pile' vs. *porla* [po:ɭa] 'ripple'; and *mossa* [mɔs:a] 'moss' vs. *morsa* [mɔʂ:a] 'say hello'. Swedish speakers would thus be expected to be relatively reliable in detecting dental/alveolar contrasts in coronal consonants.

The experiment to be reported here was carried out to explore Swedish listeners' capability of distinguishing phonetic contrasts along an articulatory continuum comprising dental, prealveolar and postalveolar places of articulation for coronals. A successful

accomplishment of this task would suggest that Swedes are capable of auditorily classifying coronals with a high degree of place resolution. Such a result would encourage us to believe that a cross-language or cross-dialect auditory analysis of coronals could be undertaken at this level of detail.

2 Method

A female and a male native speaker of Central Standard Swedish read VCV utterances with V=/ɑ i ɛ ɔ u/, V_1=V_2 and C=/t d n l/. Simultaneous audio and electropalatographic (EPG) recordings were made while the speakers were repeating the utterances several times trying to vary the respective consonants continuously between a dental and a postalveolar place of articulation. The EPG palates, which extended from behind the upper incisors to the posterior part of the hard palate, had 8 rows of electrodes with 6 electrodes in the first row and 8 electrodes in the remaining 7 rows (Hardcastle 1989). The electrodes are activated when coming into contact with the tongue. An EPG record thus shows the time-varying contact pattern between the tongue and the roof of the mouth.

For both speakers, the resulting averaged EPG data displayed three maxima with respect to degree of activation of electrode rows. These maxima were taken to reflect three distinct places of articulation: 1) dental (with the first and last full contacts pertaining to the consonant occurring on the second row); 2) alveolar (with the first and last full contacts occurring on the third row); and 3) postalveolar (with the first and last full contacts occurring on or behind the fourth row).

One representative of each of these EPG categories was used for each consonant /t d n/ and /l/ and vowel /ɑ i ɛ ɔ/and /u/ as a stimulus for the listening experiment. The dental, alveolar and postalveolar stimuli are henceforth referred to as the 'D', 'A' and 'P' stimuli, respectively. The stimuli consisted of the respective utterances as produced by the two speakers. The intended number of stimuli was, thus, 120 (5 vowels × 4 consonants × 3 places × 2 speakers), but in 17 cases speakers failed to produce exemplars meeting the criteria.

Ten phoneticians, staff or students at the Department of Linguistics, participated in the listening test. Eight of the subjects were native Swedish speakers, all using the dental/alveolar contrast discussed above, and two were native speakers of German and Finnish, respectively, but also fluent in Swedish.

The stimuli were played back from a PC using the 'Judge' tool included in the Swell Soundfile Editor (AB Nyvalla DSP, 87-98). The subjects took the test individually and listened to the stimuli via headphones. A forced choice was given between the three possible responses D, A and P. The stimuli produced by the female and the male speakers were presented separately. Speaker order was varied such that half of the subjects listened to the female speaker first, and half listened to the male speaker first. Stimulus order was automatically randomised for each listener. Subjects were allowed to listen to each stimulus repeatedly, and could go back to earlier stimuli and make changes.

3 Results

In the following confusion matrices, 'correct' responses are indicated in the leftmost column. Table 1 shows the response patterns, averaged over all listeners, to the stimuli produced by the female speaker. It can be seen, for example, that the D stimuli were perceived as dental in 76.6 percent of the cases, whereas 14.9 and 8.5 of the D productions were perceived as alveolar and postalveolar, respectively. Of the A stimuli, 44.1 percent were judged to be alveolar, while almost the same proportion of the A stimuli (37.5 percent) were heard as dentals. The latter pattern was even clearer in the responses to the male speaker, a majority of the A stimuli being marked as dentals. In general, the D and P stimuli reflected the EPG patterns more reliably than did the intermediate A stimuli.

Table 1. Confusions matrix showing average listener responses to stimuli produced by the female speaker (percent).

	Answer		
Correct	D	A	P
D	76,6	14,9	8,5
A	37,5	44,1	18,4
P	1,2	36,0	62,7

Table 2. Average listener responses to stimuli produced by the male speaker (percent).

	Answer		
Correct	D	A	P
D	77,5	12,6	9,9
A	51,2	27,7	21,1
P	2,3	27,8	69,9

Listener responses did not display any systematic effects of vowel context. In contrast, the consonantal influence was quite substantial. Results per consonant, again averaged over all listener responses, are summarised in Table 3 (consonant indicated on the top left). It can be seen that the D stimuli for /t/ were identified most reliably (94.9 percent), followed by /d/ (87.9 percent). The least reliable response to the D stimuli were those to /n/ (71.8 percent) and /l/ (55.8 percent). In contrast, the P stimuli displayed the reversed order with the highest average score for /l/ (87.7 percent) followed by /n/, /d/ and /t/ in that order. Again, the lowest score were obtained for the intermediate A stimuli. In summary, then, when these responses were averaged over listeners, D and P stimuli tended to agree reasonably well with the EPG classification, but responses to the intermediate A stimuli appeared to be quite random.

Table 3. Averaged listener responses by consonant (percent).

/D/	Answer		
Correct	D	A	P
D	87,9	5,2	6,9
A	42,7	35,4	22,0
P	2,8	35,2	62,0

/L/	Answer		
Correct	D	A	P
D	55,8	33,7	10,5
A	46,2	35,2	18,7
P	0,0	12,3	87,7

/N/	Answer		
Correct	D	A	P
D	71,8	11,8	16,5
A	42,5	30,0	27,5
P	4,8	19,0	76,2

/T/	Answer		
Correct	D	A	P
D	94,9	3,8	1,3
A	34,3	44,8	20,9
P	2,2	41,6	56,2

The amount of individual variation is illustrated in Tables 4 and 5. The table shows the two 'best' (Table 4) and two 'worst' (Table 5) listener results for all stimuli and both speakers combined. The left and right 'best' listeners are authors PL and OE, respectively. The figures indicate that the best listeners were usually more consistent than average in identifying the D, A and P categories with the corresponding places of articulation.Note in particular OE's 60 percent correct identification of the A category. The A figures for the worst listeners were clearly below chance level.

Table 4. Listener responses given by the two 'best' listeners (PL, left, and OE, right).

Conf. Sub.5

	Answer		
Correct	D	A	P
D	93.5	6.5	0.0
A	27.3	39.4	33.3
P	0.0	33.3	66.7

Conf. Sub.6

	Answer		
Correct	D	A	P
D	87.2	12.8	0.0
A	13.3	60.0	26.7
P	0.0	14.7	85.3

Table 5. Listener responses given by the two 'worst' listeners.

Conf. Sub.2

	Answer		
Correct	D	A	P
D	80.0	20.0	0.0
A	61.8	20.6	17.6
P	9.3	33.3	57.4

Conf. Sub.4

	Answer		
Correct	D	A	P
D	25.0	25.0	50.0
A	69.0	21.4	9.5
P	0.0	40.0	60.0

4 Summary and conclusions

In conclusion, whereas the average listener identified the respective D and P stimuli as dentals and postalveolars well above chance, there was a marked uncertainty about the A stimuli. Interestingly, however, there was a high degree of individual variation. In particular, the scores obtained by the two experimenters were clearly above average. This may be a matter of motivation, a sign of individual differences in auditory-phonetic sensitivity to the dimension in question, or a result of temporarily tuning in to the dimension. It should also be noted that the EPG-based criteria used here were solely based on place. Thus, they may not have reflected natural articulatory categories that would also have involved tongue shape. Nevertheless, the present experiment does not allow the conclusion that even the most reliable listeners would be capable of safely identifying more than two places of articulation for coronals in a cross-language or cross-dialectal material. It is conceivable, however, that further practice in listening to contrasts in the dental/alveolar region would improve the results to some extent. This would be of interest also in connection with second language acquisition.

Acknowledgements

This work was supported by the Bank of Sweden Tercentenary Foundation (Riksbankens Jubileumsfond), grant 1997-5066:03, 04 to the project *De svenska dialekternas fonetik och fonologi år 2000 (SWEDIA 2000)* [*Phonetics and phonology of the Swedish dialects around the year 2000.*]

References

Dart, S.N. and P. Nihalani 1999. 'The articulation of Malayalam coronal stops and nasals.' *Journal of the International Phonetic Association* 29(2): 129-142.

Hardcastle, W. (ed.) 1989. 'Clinical linguistics and phonetics.' 3(1) Special issue: Clinical applications of electropalatography.

Livijn, P. and O. Engstrand 2001. 'Place of articulation for coronals in some Swedish dialects.' *This volume.*