Perceptual experiments with nasals

Nord, L.

journal: STL-QPSR
volume: 17
number: 2-3
year: 1976
pages: 005-008

http://www.speech.kth.se/qpsr
B. PERCEPTUAL EXPERIMENTS WITH NASALS

L. Nord

Abstract

A limited experiment with nasals was performed in order to get an estimate of the relative importance of the nasal murmur compared to the vowel transitions as a cue for the place of articulation of the nasal consonant. The Swedish nasals /m/ and /n/ were studied in an /a-a/ context. The speech wave was edited with the help of a computer program. New stimuli were formed with occlusions and parts of the transitions exchanged.

Introduction

This report is part of a project concerning the analysis of nasals. The aim is to produce synthetic nasals of good quality for a rule-synthesis program.

Earlier studies with synthetic VC-stimuli at Haskins Laboratories, Liberman et al (1954), have shown that the F2-transition alone could serve as a cue for distinguishing the place of articulation of the nasal consonant when a fixed nasal segment was used for /m, n, ɳ/. However, the F2-transition was not so effective for the discrimination between nasals as between stops. Malécot’s experiment (1956) with spliced and edited natural speech showed that transitions in CV- and VC-syllables give the essential place information while the nasal resonance primarily indicates the manner of articulation. However, for the nasals in final position (VC-syllables) the nasal segment also seemed to contribute somewhat to the place information.

The results of these experiments motivate a synthesis strategy where the nasal segment is one and the same for both labial, dental, and velar articulation, and the formant transitions alone determine the place of articulation.

An analysis of nasals of course reveals differences in their spectral envelopes, see Fujimura (1962) and Fant (1960), although it is difficult to extract any consistent features of the spectra, especially when a comparison is made over several speakers. The question thus is, how much does the occlusion contribute to the identification in comparison with the transitions? And furthermore, will the segmental quality of the synthetic speech improve if the nasal identity to some extent would depend on the
occlusion too, and not only the transition? As a start it could be convenient to see what happens in natural speech when the nasal segment in VCV-words is exchanged for another nasal. The following experiment is thus a repetition of Malécot's study but with a smaller word material and another technique for editing the speech wave.

**Experiment**

Two words with not too pronounced vowel transitions, namely /ama/ and /ana/ were chosen, see Fig. I-B-1a. The following manipulations were made on the digitized speech wave with the aid of a computer program*. /ama/ and /ana/ with approximately equal stress on both syllables and of approximately the same duration were taken out of two nonsense sentences. The nasal segment or the nasal segment together with a part of the CV- and/or VC-transition were exchanged between the words so that new items were formed, see Fig. I-B-1b. The computer program makes it possible to place the boundaries between segments at zero crossing of the speech wave pitch synchronously, which gives the stimuli a very good quality with no clicks at the boundaries. The stimuli were presented in randomized order to a group of listeners. The task was to identify the consonant as /m, n, or /s/. Each stimulus was presented 3 times with 1 sec interval and with a 5 sec interval to the next group. The listening sessions took place, with a few days interval, via earphones and via a loudspeaker in an ordinary office room at a distance of about 2.5 meter at a comfortable listening level.

**Results**

Five subjects listened twice to every stimulus, i.e. every point in Fig. I-B-2 corresponds to 10 answers given. In 2% of the cases no answer was given. The small size of the test gives the curves a somewhat uneven appearance. Some tendencies nevertheless can be found. Fig. I-B-2 shows that the nasal resonance seems to contain some information of place of articulation, since an exchange of only this segment gives 80% /n/ response and 70% /m/ response for stimulus /a_m a/ and /a_n a/, respectively. When listening through loudspeaker on the other hand, the nasal segment cannot change the identity unless a few

---

* The editing program was written by Rolf Carlson and Björn Granström.
Fig. I-B-1a. The test words taken from two sentences.
Fig. I-B-1b. The middle parts of /ama/ and /ana/ taken from two sentences. Segmental boundaries between nasal and vowels are marked. The exchanged glottal pulses are numbered.
Fig. I-B-2. Identification scores for VCV stimuli where the nasal segment and a number of glottal pulses in the preceding and/or following vowel have been interchanged. 10 responses / data point.
glottal pulses from the following vowel have been added. Both with ear-
phones and loudspeaker the VC-transition, that is, the glottal pulses be-
fore the nasal contribute very little to the place information, while
the CV-transition contains almost all the information. One could argue
that part of the transition already is included in the 'occlusion' as that
segment alone will change the identity when listening through earphones.
But it is difficult to find a better segment boundary than that chosen,
see Fig. I-B-1b. The vowel transitions as well as the spectral differ-
ence between /m/ and /n/ are relatively small for these samples, see
Fig. I-B-3, with a higher zero due to a smaller mouth cavity shunt for
/n/ compared to /m/.

The result that the CV-transition contributes more than the VC-transi-
tion in these VCV-words agrees well with Malécot's finding concerning
CV- and VC-syllables, that in the VC-syllables, the nasal segment car-
rries more information than in the CV-syllables.

Complementary tests with [ am:a] and [ an:a] spoken in isolation
were made. Here the segment boundaries are more distinct, see Fig.
I-B-4, but still some place information seems to go with the nasal oc-
clusion, see Fig. I-B-5.

Some unpublished data from a similar experiment on isolated CV-
syllables containing the long vowel [a:] made by Carlson and Granström
are shown in Fig. I-B-6. (For additional material on stops, laterals,
and nasals, see Carlson et al (1972).)

The reason for the large variation in identification score of the stim-
uli where only the occlusion had been switched is difficult to find and not
the aim of this study. Probably factors such as word material, speaker
dependence, pronunciation etc. have to be taken into account.

Conclusions

The experiment with edited natural speech shows that for the chosen
words, the nasal resonance contains some information on the place of
articulation of the nasal consonant. The transition into the next vowel is
the most important part, while the transition from the preceding vowel
is of considerably less importance for place information. It is not pos-
sible to quantify any exact relations between the amount of place informa-
tion contained in the nasal segment compared to the transitions, as that
seems to vary strongly with context, speakers etc.
Fig. I-B-3. Spectral envelope of the nasal segments corresponding to Fig. I-B-2.
Fig. I-B-4. Parts of the speech waves of /ama/ and /ana/ spoken in isolation. Segmental boundaries between nasal and vowels are marked. The exchanged glottal pulses are numbered. Note that parts of the nasal segments are omitted in the figure.
Fig. I-B-5. See Fig. I-B-2. Identification scores for VCV stimuli spoken in isolation.
Fig. I-B-6. See Fig. I-B-2. Identification scores for CV syllables.
However, from the experiments the conclusion can be drawn that when synthesizing nasal sounds the quality and intelligibility will improve if also the nasal segment contains some information about the place of articulation of the nasal consonants.

References: