EXPERIMENTS WITH THE PERCEPTION OF NASALITY

CORNELIS W. KOUTSTAAL

1. INTRODUCTION

In the literature, various accounts and estimates of nasality are available. These accounts and estimates may be categorized as clinical and experimental. The clinical material explores the significance of velo-pharyngeal closure, the coupling of oral and nasal cavities and other functional aspects of the speech mechanisms. The experimental material has been devoted to analyzing and acoustically measuring the differences between normal and nasalized speech. Generally we find the terms HYPER-NASALITY, NASALITY, and NASALIZED SPEECH used interchangeably and meaning 'the quality of speech sounds when the nasal cavity is used as a resonator, especially when there is too much nasal resonance'.

The present study attempted to combine some of the clinical ability of judging nasality and laboratory techniques of analyzing and/or modifying nasality. Two areas were specifically explored. The first was whether clinicians, in a laboratory perceptual task, would generate data which would agree with those reported in the literature. The second was whether other clinicians, without this laboratory experience, would demonstrate agreement in their judgments regarding the laboratory data. In order to attain these objectives, a series of experiments were constructed.

2. EXPERIMENT I

The first question asked was whether clinicians can manipulate reliably the speech spectra on the basis of auditory monitoring to decrease the amount of perceived nasality.

1. *Method.* — Speech samples of six post-operative cleft palate speakers were recorded. All cases were diagnosed and under treatment for hypernasality. Because of the age range of these subjects, four to 21 years of age, the speech material consisted of saying the digits one through twelve and saying the names of the colors.

Six graduate student clinicians in speech pathology and audiology at Bowling Green State University served as the first set of listeners. They were instructed to

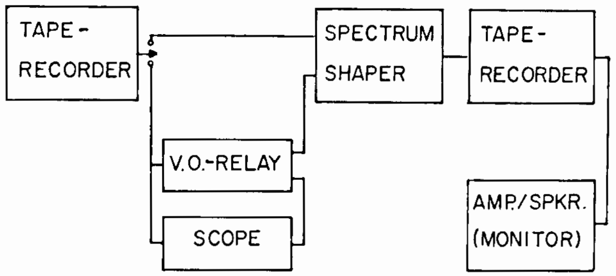

Fig. 1. Schematic of instrumentation used in modifying nasal speech of 6 cleft-palate speakers.

listen to the nasalized speech and then to manipulate the attenuators of the 1/3-octave filters on a B & K spectrum shaper. The frequency values of the filters were obscured. Their task was to arrive at attenuator settings which would eliminate the hyper-nasal quality.

After they had achieved this goal, according to their own judgment, the modified speech signal was re-recorded. To obtain the speech spectrum envelopes, the peak voltage values for each syllable from each 1/3-octave was plotted.

2. *Results.* — When all voltages of the syllable peaks were obtained they were averaged for the conditions of non-filtered nasal speech and filtered nasal speech. These values were plotted and two distinctly different spectral envelopes resulted. The greatest amount of modification occurs between the frequencies 400 Hz and

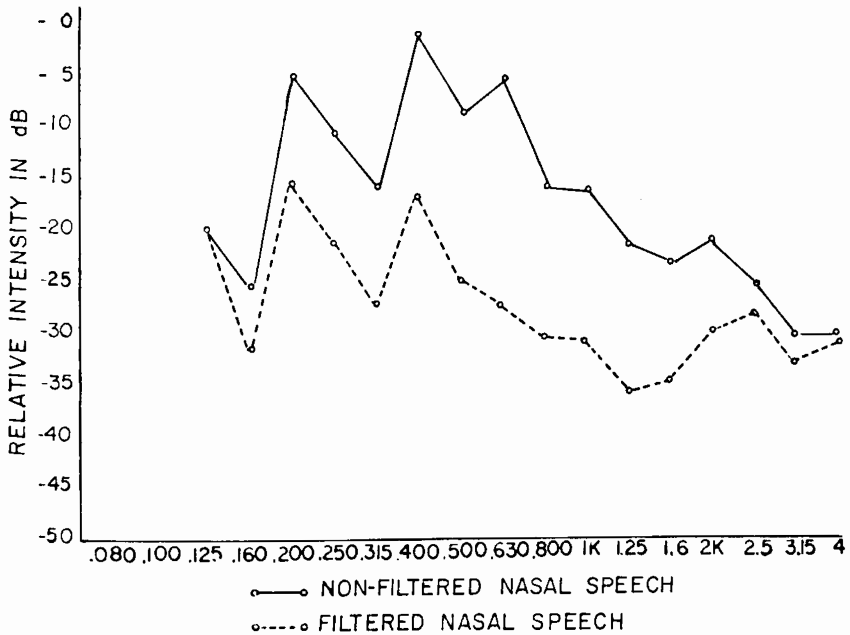

Fig. 2. $^{1}/_{3}$ Octave Filter Frequencies in kHz.

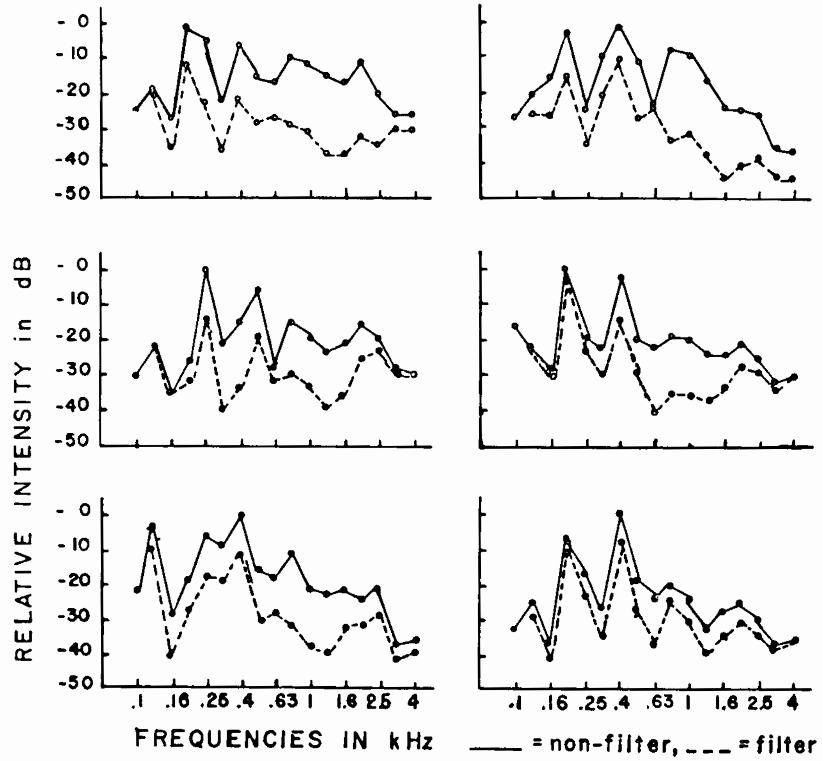

Fig. 3. Filtered and non-Filtered Nasal Speech Spectra of 6 Cleft-Palate Speakers.

1600 Hz. Also, the roll-off appears steeper in the filtered condition than in the non-filtered condition. These findings appear consistent with those reported in the literature. The agreement of the student clinicians regarding the attenuator settings was tested with a Kendall-W (W = .78) coefficient of concordance. When tested for significance, a confidence level of .999 was obtained. On the basis of these data it was concluded that clinicians can reliably modify the speech spectra of hypernasal speakers.

3. EXPERIMENT II

The second question asked was whether clinicians can recognize different levels of nasality in the non-filtered and filtered speech of cleft palate speakers.

1. *Method.* — For the listening task 30 undergraduate students in speech pathology and audiology were recruited. They were instructed in the use of a nine-point rating scale to rate the perceived amount of hypernasality of each of the speech samples. Prior to the experiment they listened to other speech material than that contained in the experimental samples produced by the same speaker in order to acquaint them with the range of nasal quality. There were three experimental listening conditions.

Condition *A* was the original speech sample, Condition *C* was the modified speech spectrum for each speaker and for Condition *B* all filter conditions were averaged.

2. *Results.* — Systematic changes in the amount of perceived hypernasality were observed. The test was administered twice; each time to a group of 15. Between group reliability was .98. When the differences between mean ratings were tested with an *F* ratio no significance was observed.

TABLE 1

Average perceived hyper-nasality for five speakers under three listening conditions

Speaker	No-Filter	Mean Filter	Individual Filter
1	4.32	3.71	3.48
2	8.1	7.48	7.16
3	5.45	4.71	4.55
4	4.87	3.94	3.58
5	5.30	5.52	4.45

In re-thinking the procedure, it was decided that the listening conditions for the six clinicians modifying the spectra and those judging the modifications were not the same. The clinicians listened to modified speech of which the signal strength had been attenuated as a function of the filtering. On the other hand the 30 listeners listening to the recording of the modified speech listened to speech that had been pre-amplified to peak at zero for recording. Another recording was made reflecting these differences in signal to noise ratios. The signal level differences ranged from 4.9 to 9.6 dB. When the recordings of modified nasal speech reflecting these signal to noise ratio differences were rated by another group of listeners ($N = 33$) for hypernasality the samples with 7 dB or more intensity difference were rated significantly lower.

TABLE 2

Differences in dB between signal level before and after filtering and average perceived hyper-nasality for two signal levels of filtered speech

Speaker	Difference in dB	Average perceived hyper-nasality	
		No Attenuation	Attenuation
1	9.6	4.33	2.91 *
2	7.5	7.64	7.15 *
3	4.9	8.51	8.24
4	9.9	5.24	3.64 *
5	7.0	4.97	3.54 *
6	5.8	4.30	4.27

* Differences significant ($\alpha = .05$) when tested with t.

On the basis of these data it was concluded that listeners can concur with the six clinicians who modified the speech spectra. One does wonder, however, whether the different ratings for perceived hypernasality were given on the basis of different spectral relationships or merely on the basis of modified *S/N* ratios.

In the experimental tape, the listening conditions had been presented in the same order for each of the speakers in sequence. Thus the decrease in perceived hypernasality might be argued to represent an order and sequence effect or in other words represent a form of adaptation. Also, subjective comments from clinicians and other listeners suggested that there were accompanying features that were audible. Such features were labeled as glottal pulse, clicks and nasal emission. They primarily referred to the onset of the syllables in addition to the resonance characteristics. To take these two arguments into consideration, another listener experiment was constructed.

4. EXPERIMENT III

The third question asked was whether a delay in onset of the syllable by itself and in combination with filtering would affect the ratings of perceived hypernasality significantly.

1. *Method.* — With the use of a voice operated relay and monitored on a dual trace oscilloscope the speech samples were re-recorded with a delay of onset time of 40 msec. Four listening conditions were constructed: Condition *A* — the original hypernasal speech, Condition *B* — the use of the voice-operated relay which delayed the onset of the signal by 40 msec, Condition *C* — the use of filtering, and Condition *D* — the combination of delay of onset and filtering.

Three groups of listeners were used. Groups 1 and 2 listened to the twenty-four segments in randomized order and group 3 listened to each speaker in sequence but with the conditions *ABCD* in randomized order.

2. *Results.* — The mean ratings for perceived hypernasality for listener groups 1 and 2 showed no systematic changes in average ratings. When these data are compared with the averages obtained from listener group 3 there is an apparent difference. Listener group 3 showed systematic differences. Apparently sequencing the speakers rather than randomizing all samples is of significance. The data were not subjected to a test with a parametric statistic because of the presence of sequence and order effects.

5. SUMMARY AND CONCLUSIONS

On the basis of this series of studies, several conclusions were drawn. Although all subjects had previously some experience and knowledge of cleft palate speech, the definition of hypernasality seemed to vary from subject to subject. In a laboratory setting consistent and significant results can be obtained which do agree with reports in the literature. Listener judgments in the traditional research methodology of randomization will not render significance. It would appear from this that the laboratory tasks and rating or clinical judgments are incompatible. Because of the limitations in the number of speech features that could be modified and the unlimited time the clinicians had to modify the nasalized speech a certain adaptation to other speech features may have taken place. Apparently other speech characteristics than nasal resonance alone were incorporated in the clinical judgment. Terminology should be carefully reviewed and further research carried out to more accurately and/or completely specify the speech features which constitute hypernasality.

TABLE 3

Average perceived hyper-nasality for six speakers under four conditions obtained from random presentation to two listener groups (N = 26 *and* 31)

	Conditions			
Speaker	Original	Delay	Filter	Filter & Delay
1	4.04	4.65	4.00	4.42
	3.52	4.55	4.03	4.58
2	8.19	8.23	7.92	8.19
	7.58	7.48	7.65	7.84
3	6.96	7.31	7.73	8.12
	6.81	7.26	7.23	7.81
4	3.69	4.96	3.65	5.77
	3.71	4.61	3.00	5.45
5	2.85	3.77	3.15	3.96
	2.52	3.39	3.03	4.16
6	2.52	5.04	2.73	4.38
	3.45	5.35	3.26	4.97

TABLE 4

Average perceived hyper-nasality for six speakers under four conditions where conditions were randomized for each speaker obtained from 25 listeners

Speaker	Original	Delay	Filter	Delay & Filter
1	5.72	5.12	4.12	3.72
2	7.52	7.32	7.48	6.76
3	5.56	4.96	5.56	5.20
4	4.24	3.84	2.44	2.76
5	3.72	3.04	3.04	3.04
6	4.80	5.28	3.36	3.48

Communication Sciences Laboratory
Bowling Green State University
Bowling Green, Ohio

REFERENCES

Bradford, L.J., A.R. Brooks, and R.L. Shelton
1964 "Clinical Judgment of Hypernasality in Cleft Palate Children", *Cleft Palate Journal* 1:329.

Koepp-Baker, H.
1957 "Pathology of Cleft palate and Cleft Lip", in *Handbook of Speech Pathology*, Travis, ed. (New York, Appleton-Century-Crofts, Inc.), see Chapter 19.

Luchsinger, R. and G.E. Arnold
1959 *Lehrbuch der Stimm- und Sprachheilkunde* (Wien, Springer-Verlag).

Peterson, G.E., and H.L. Barney
1952 "Control Methods Used in a Study of the Vowels", *Journal of the Acoustical Society of America* 24:175.

Schwartz, M.F.
1968 "The Acoustics of Normal and Nasal Production", *Cleft Palate Journal* 5:125.
1971 "Acoustic Measures of Nasalization", *Cleft Lip and Palate*, W.C. Grabb, S.W. Rosenstein, and K.R. Bzoch, ed. (Boston, Little, Brown and Company) pp. 798-804.

Sherman, D.
1952 "The Merits of Backward Playing of Connected Speech in the Scaling of Voice Quality Disorders", *Journal of Speech and Hearing Disorders* 17:198-212.

Van Riper, C. and J.V. Irwin
1958 *Voice and Articulation* (Englewood Cliffs, N.J., Prentice-Hall).

Wood, K.S.
1957 "Terminology and Nomenclature", in *Handbook of Speech Pathology*, Travis, ed. (New York, Appleton-Crofts, Inc.) see Chapter 2.

DISCUSSION

COLEMAN (Portland, Oreg.)
How were the clinicians trained to use the scale; were certain examples given with a certain rating?

KOUTSTAAL
The listeners were given speech samples of all speakers. No examples of rating were provided. In other words, the scale was not anchored. Listeners were asked to use an 'internal' standard of comparison.

SINGH (Washington)
Would you elaborate on the idea of multiple features?

KOUTSTAAL
Because it is apparent from the data that many other things were operating, I propose to give the judges the opportunity to attend to several variables. On the basis of the literature (experimental and clinical) I propose at least the following features: (a) distinctiveness of speech (or, speech clarity); (b) articulatory proficiency, or correctness; (c) intonation; and, (d) quality or resonance.

HUCKLEBERRY (Muncie, Ind.)
How did the judgments of the young clinicians agree with the author's judgment?

KOUTSTAAL
They did not agree, because I had become sufficiently biased to recognize differences.

HUCKLEBERRY
What are the values in the training of young clinicians?

KOUTSTAAL
Incorporating the strategy of multi-dimensional scaling, we might be able to acquire agreement between clinical and laboratory data.