Dept. for Speech, Music and Hearing
Quarterly Progress and Status Report

Acoustical and perceptual evaluation of speech training in post-operative cleft palate patients

Ericsson, G. and Fant, G. and de Serpa-Leitao, A.

KTH Computer Science and Communication

journal: STL-QPSR
volume: 14
number: 1
year: 1973
pages: 025-028

http://www.speech.kth.se/qpsr

II. MEDICAL APPLICATIONS

A. ACOUSTICAL AND PERCEPTUAL EVALUATION OF SPEECH TRAINING IN POST-OPERATIVE CLEFT PALATE PATIENTS

Gärda Ericsson*, G. Fant, and A. de Serpa-Leitao

Abstract

Spectrographic and subjective evaluation tests show improvements in speech quality as a result of secondary operation of cleft palate patients. The spectrographic analysis has concentrated on nasalization cues in vowels. The main improvements come after a period of speech training following operation. This is a small pilot study which exemplifies the multilevel approach planned for the continuation.

Introduction

One of the most apparent characteristics of cleft palate speech is the nasal quality of vowels. The main cause of this degradation is supposed to be insufficient velo-pharyngeal closure. Cleft palate patients generally undergo a primary operation at a very early age and a secondary at a higher age. The latter pharyngeal flap operation usually ensures a velo-pharyngeal closure competence as revealed from post-operative inspection. Nevertheless, the improvement in vowel quality is not as great as one would expect. The nature of the remaining insufficiency is not very well known. It could be that earlier acquired compensatory articulation patterns become inadequate with the change in functional constraints and interfere with a correct programming of velo-pharyngeal closing and opening movements or that a correct synchronization of such movements never has been learned. It may also be that certain aspects of the impression of nasality are related not to insufficient velo-pharyngeal closure but to deviant articulatory modes.

There are other well-known aspects of cleft palate speech that will not be taken up in this report. These are consonant insufficiency and detoriation, incorrect place of articulation, and voice source deficiences.

The ultimate scope of our joint work is to study the improvement in speech quality following secondary operation and speech training and to establish an acoustically correlated physiological model of changes in speech patterns during this period. An important problem is to find

---

* Regionssjukhuset, Linköping.

out what parameters in speech spectrograms and other acoustical measurements are most closely correlated to improvements.

This is a long-time project involving cooperation between several Swedish groups; the Phoniatrics and the Radiology clinics of the Regional Hospital in Linköping, the Plastic surgery and the Phoniatrics clinics of the Karolinska Sjukhuset, Stockholm, and the Department of Speech Communication at the Royal Institute of Technology (KTH) in Stockholm and the Department of Medical Electronics at the Technical University of Linköping. The present report presents a small pilot study of spectrography and speech quality evaluation of a speech material collected by Gärda Ericsson in Linköping and analyzed at KTH.

The subjects' speech 1,5 years after operation in September 1971 showed a substantial improvement which was enhanced in a test six months later.

Spectrographic control

Voice print broad-filter spectrograms from four occasions are shown in Fig. II-A-1a and sections in Fig. II-A-1b. These occasions are:

A. Before operation. December 1968.

B. After operation. March 1970.

C. After a period of training following the operation. September 1971.

D. After additional training. August 1972.

The sentence selected for this illustration is "Mimmi bakar kakor". Comparing A and B it may be seen that the effect of the surgery is to lower the frequencies of nasal resonances in the low and medium range of the nasal consonants [m]. In the [a] vowels there is improvement in the second formant which comes out clearer. These effects are probably the direct consequence of the decreased velo-pharyngeal aperture. The remaining nasalization characteristics are apparent, e.g., the extra formants in the region of F1, see especially sample (2) of recording B.

The visually most apparent changes come after the period of speech training, as seen by comparing C and B. The first formant peak comes

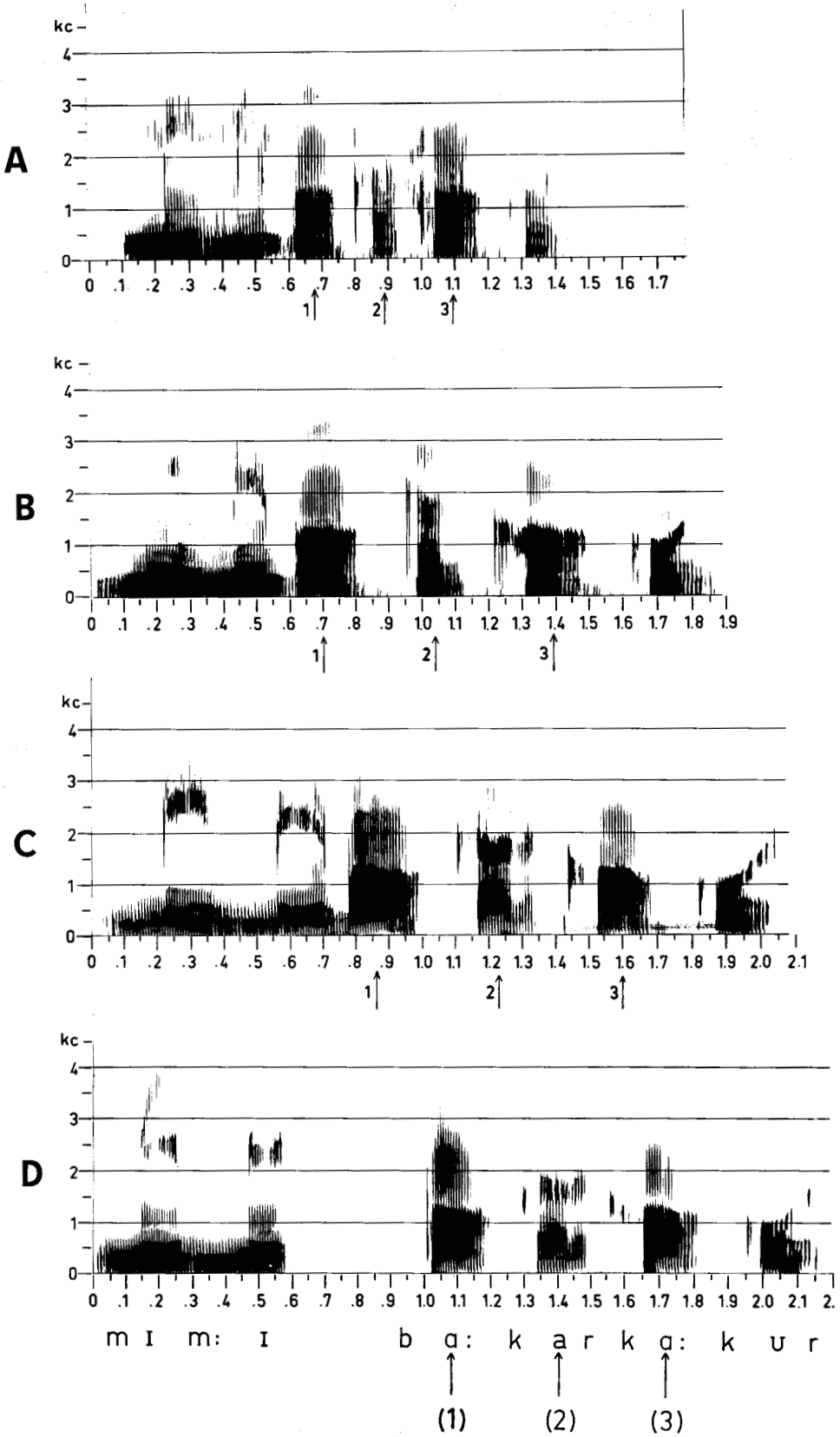
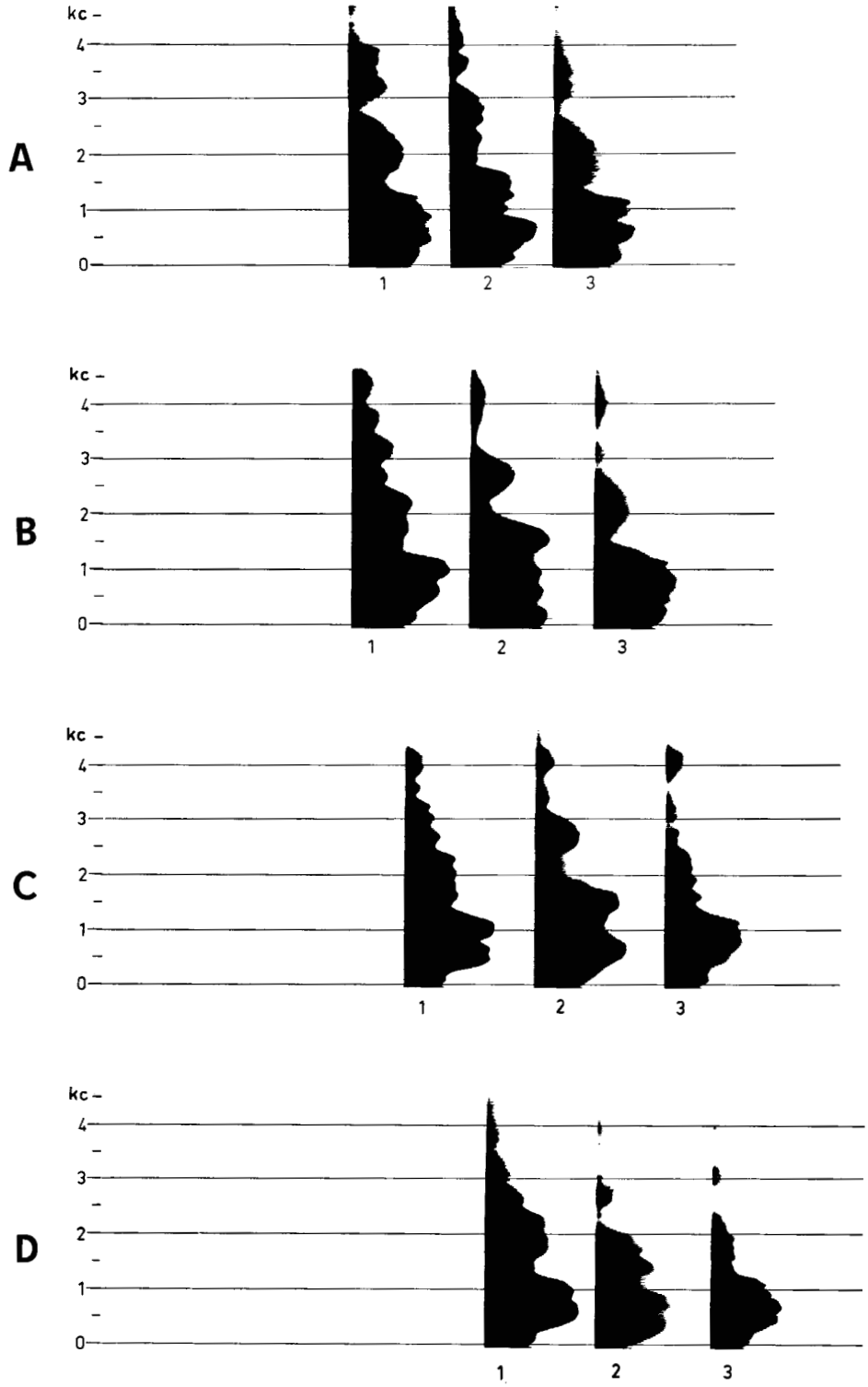


Fig. II-A-1a. Spectrograms of the sentence "Mimmi bakar kakor" spoken:
A before and B after the secondary operation,
C after a following period of voice training, and
D a follow up test.



Sections on (1) = [ɑ] in bɑ:kar
(2) = [a] in ba:kar
(3) = [ɑ] in kɑ:kur

Fig. II-A-1b. Amplitude-frequency sections of vowels from Fig. II-A-1a.

out clearer and the domain below F1 shows less traces of extra formants, see section (1) and (2) of C. The improvements in D are of the same kind. It must be stressed that this is a very limited material and that the observations are qualitative only. However, we have several other sentences spoken by the same subject and other subjects as a basis for our evaluation. The most consistent phenomenon is the improvement after operation and subsequent speech training.

Clinical studies of speech changes after cleft palate operation and training have generally not included spectrographic analysis and detailed discussions of nasalization cues. The most illustrative source is still that of Nylén (1961) to be compared with normal nasal coupling studied by Björk (1961). Nylén found that the second formant intensity of [i]and [e]increased substantially after operation, which conforms with our observation here on the [a] vowel.

The general theory of nasalization in normal speech has been outlined by Fant (1960). The further studies of Fujimura&Lindqvist (1971) and Lindqvist&Sundberg (1972) need to be pursued further. Work along these lines is in progress.

Listening tests

A small listening test was performed to evaluate the overall subjective quality of the patients' speech from the four periods of rehabilitation. The same sentence as illustrated for spectrographic analysis, Fig. II-A-1, was selected. Each recording was repeated six times and mixed in random order with the other samples and played to a group of 13 university students which had to grade the quality in a seven point scale according to degree of detoriation. Recording A before operation gave an average score of 4.4, condition B after operation gave 3.1 and subsequent speech training 2.3 for condition C and 1.7 for condition D. These data support the overall impressions of the results of the rehabilitation and correlate positively with the visual correlates discussed above.

References:

Björk, L. (1961): "Velopharyngeal function in connected speech. Studies using tomography and cineradiography synchronized with speech spectrography", thesis work. Acta Radiologica, Suppl. 202, Stockholm, 94 pp.

Fant, G. (1960): Acoustic Theory of Speech Production, Mouton & Co., 's-Gravenhage; 2nd edition 1971.

Fujimura, O. and Lindqvist, J. (1971): "Acoustic properties of the nasal tract", J.Acoust.Soc.Am. 49, pp. 541-558.

Lindqvist, J. and Sundberg, J. (1972): "Acoustic properties of the nasal tract", STL-QPSR 1/1972, pp. 13-17.

Nylén, B. (1961): "Cleft palate and speech. A surgical study including observations on velopharyngeal closure during connected speech, using synchronized cineradiography and sound spectrography", thesis work. Acta Radiologica, Suppl. 203, Stockholm, 124 pp.

Acknowledgments

The listening test was organized and evaluated by R. Carlson and B. Granström.