Hirschberg, J. 1990. 'Using discourse context to guide pitch accent decisions in synthetic speech'. *Proceedings of the ESCA workshop on speech synthesis (Autrans, France)*, 181-184.

Horne, M. 1987. 'Towards a quantified focus-based model for synthesizing English sentence stress'. *Lingua* 75, 25-54.

Horne, M. 1991. 'Why do speakers accent "given" information?'. *Proceedings of Eurospeech 91:2nd European conference on speech communication and technology*, Vol. 3, 1279-1282.

Kohler, K. J. 1987a. 'The linguistic functions of Fo peaks'. *Proceedings XIth Int. Cong. of Phonetic Sciences, Tallinn, Estonia.* Vol. 3, 149-152.

Kohler, K. J. 1987b. 'Categorical pitch perception'. *Proceedings XIth Int. Cong. of Phonetic Sciences, Tallinn, Estonia.* Vol. 5, 331-333.

Lyons, J. 1977. *Semantics*. Cambridge: Cambridge University Press.

Pierrehumbert, J. B. 1980. *The phonology and phonetics of English intonation*. Unpublished dissertation, MIT. (Distributed by Indiana University Linguistics Club, 1988).

Pierrehumbert, J. B. & J. Hirschberg. 1990. The meaning of intonation contours in the interpretation of discourse. In P.R. Cohen, J. Morgan & M. E. Pollack, *Intentions in communication*, 271-311. Cambridge, Mass.: MIT Press.

Pierrehumbert, J. B. & S.A. Steele. 1989. Categories of tonal alignment in English. *Phonetica* 46, 181-196.

Sigurd, B. 1990. *SWETRA - a multilanguage translation system. Praktisk lingvistik* 14 (Dept. of Ling., Lund University).

Lund University, Dept. of Linguistics
Working Papers 38 (1991), 115-32

Fundamental Frequency Control and Voice Quality in Cochlear Implant Users

David House and Ursula Willstedt*

This paper presents results comparing the speech production of three postlingually deafened speakers before and after activation of their cochlear prostheses. Three different types of results are presented: 1) voice quality evaluation by two trained speech pathologists, 2) F0 measurements of the production of different moods in speech, and 3) results of a listening test conducted with two groups of naive listeners using tokens of different moods taken from recordings of one of the three speakers. The speakers also received speech training immediately following activation of their implants. The results demonstrate improved voice control and changes in the use of fundamental frequency immediately following activation and voice training. Dramatic changes in the use of fundamental frequency to signal speech mood are shown after a period of six months following implant activation. Results of the listening test are also consistent with the analysis results.

Introduction

As an increasing number of postlingually deafened individuals receive cochlear implants, there has been a rising interest in the effect of implants on speech production. In a number of recent studies, speakers fitted with cochlear implants have shown improvements in their speech production following activation of the implants. Improvement has been observed primarily on the suprasegmental level such as increased control of voice quality and fundamental frequency (Plant & Öster 1986, Öster 1988, Ball & Faulkner 1989) and modification of speech breathing (Lane, Perkell, Svirsky & Webster 1991).

As part of a larger project concerning hearing impairments and the perception and production of mood in speech, we have so far recorded nine speakers prior to and at intervals after being fitted with the Nucleus multichannel cochlear prosthesis (see e.g. Waltzman & Hochberg 1990, Skinner, Holden, Holden et al. 1991). The speakers read a short text passage

*Also at the Department of Audiology, Lund University Hospital

(89 words), isolated sentences designed to elicit differences in focal accent and intonation, lists of isolated words, and isolated vowels. The speakers were also presented with situations in which they were asked to read semantically neutral sentences in different moods.

The use of different moods as speech production material was carried out for two reasons. Firstly, it has been shown that listeners with moderate to severe hearing loss have difficulties in identifying the mood of speakers in test sentences, particularly confusing happy with angry and sad (Öster & Risberg 1986, House 1990). Implant users show similar confusions (House 1991), and one of the goals of the project is to relate problems in perception to possible problems in production. Secondly, the use of the production of moods appears to be a promising evaluation tool for ascertaining changes in speech production particularly concerning control of fundamental frequency, intensity and voice quality as these are primary acoustic cues for signalling speaker mood (see e.g. Hadding-Koch 1961, Gårding & Abramson 1965, Williams & Stevens 1972, Fónagy 1981, Bruce 1982, Scherer, Ladd & Silverman 1984, House 1990, Granström & Nord 1991).

In this paper, results are presented showing changes in voice quality and control of fundamental frequency in three implant subjects. Voice quality was evaluated auditorily by two trained speech pathologists based on recordings of the text passage. Fundamental frequency (F0) mean and range were measured using recordings of the different moods. Finally, a formal listening test was conducted with two groups of naive listeners using tokens of different moods taken from recordings of one of the three speakers.

Method

Subjects and Speech Materials

The subjects were three of the first patients to receive a Nucleus multichannel cochlear prosthesis within the cochlear implant program at the Department of Audiology, University Hospital, Lund. All subjects are speakers of dialects of southern Swedish, two are females in their early 40's (S1 and S2) and one is a male in his late teens (S3). Prior to receiving their implants, S1 and S2 had been profoundly deaf for more than five years while S3 had been profoundly deaf for between 1 and 2 years.

The subjects were recorded in a sound-treated booth using a high quality professional tape recorder at the Department of Logopedics and Phoniatrics, University Hospital, Lund. Recordings were made shortly

before surgery and at two weeks, six months and a year after implant activation. Future recordings of the subjects will also be made at yearly intervals. The first three recordings were used for the results presented here.

The speech material used for voice quality evaluation consisted of a short standard text passage (89 words). For the F0 study, the speakers were presented with situations in which they were asked to read semantically neutral sentences in different moods. The sentence used for analysis was *Nu flyttar jag* 'Now I'm going to move' describing four different situations (neutral, angry, happy and sad). Each sentence was read three consecutive times in each of the four moods.

Voice training

In conjunction with processor activation and adjustment, each subject received intensive voice training once a day for two weeks. The training consisted of general voice training followed by individually adapted voice exercises. The general training was designed to provide the subjects with speech and voice awareness leading to greater voice control. Included in the general training were relaxation exercises and exercises to improve coordination of breathing and phonation. The individually adapted exercises for the three subjects were aimed at lowering fundamental frequency, improving breathing coordination in longer phrases, and improving control of intensity and nasality.

Voice quality evaluation

Two speech pathologists with considerable experience in voice quality evaluation listened to the recordings of the 89-word text passage without being informed of the order of the recordings. Six laryngeal and three supralaryngeal parameters were rated using a 7-point scale where 0 represents no deviation from normal and 6 represents extreme deviation (Hammarberg 1986). Voice pitch was rated on a non-numerical scale having endpoints 'far too low' and 'far too high'. A general impression of voice quality and articulation was also given for each recording on a non-numerical scale having endpoints 'normal' and 'abnormal'.

Fundamental frequency measurements

A fundamental frequency contour was extracted for each of the three repetitions of the four moods for each speaker using the LuPP program

(Lund Prosody Parser, Eriksson 1990) on a Macintosh II computer. Two intonation parameters were of particular interest in the analysis, namely F0 mean and F0 range, since these two parameters can be exploited in different ways to signal different moods (cf. House 1990). While methods of quantifying these parameters is a matter of discussion (see e.g. Gårding 1991), we chose a method of approximating the parameter values using measured F0 maximum and minimum values in each utterance. In several instances these values were obtained by measuring individual period lengths especially where voice source perturbations created problems for the F0 extraction algorithm.

For F0 mean, which corresponds to an approximation of the average fundamental frequency for the utterance in Hz, we simply calculated the arithmetic mean of F0 min and F0 max over the three repetitions for each mood. For F0 range, corresponding to Gårding's 'R=global range' (Gårding 1991), we divided F0 max by F0 min giving us a 'range quotient'. This method has been used previously by Touati (1987) for comparing individual accent patterns. This method clearly has limitations for calculating both F0 mean and F0 range and would not be suitable for utterances containing several accents. However, for our short test sentence *Nu flyttar jag* with only one accent, the method supplies us with a quick and suitable approximation of both F0 mean and F0 range.

An example of F0 mean and F0 range values for four moods produced by a trained speech pathologist is presented in Figure 1. The recordings were used in a speech mood perception test and were identified at a level of 98% by a group of normal hearing listeners (House 1991). From these measurements, it is clear that the speaker uses both F0 mean and F0 range to differentiate between happy and angry, while neutral and sad are differentiated by means of F0 range. It is interesting to note that while angry and sad have the most similarity in terms of F0 for this speaker, they represent opposite extremes on a relative intensity scale, angry being produced with high intensity and sad with low (cf. House 1990).

Listening tests

As a complement to the F0 analysis procedures presented above, formal listening tests were carried out using the second of the three repetitions of each mood produced by S1 prior to implant activation (Test A) and six months following implant activation (Test B). This gave a total of eight different test tokens, two tokens of each of the four intended moods. Each

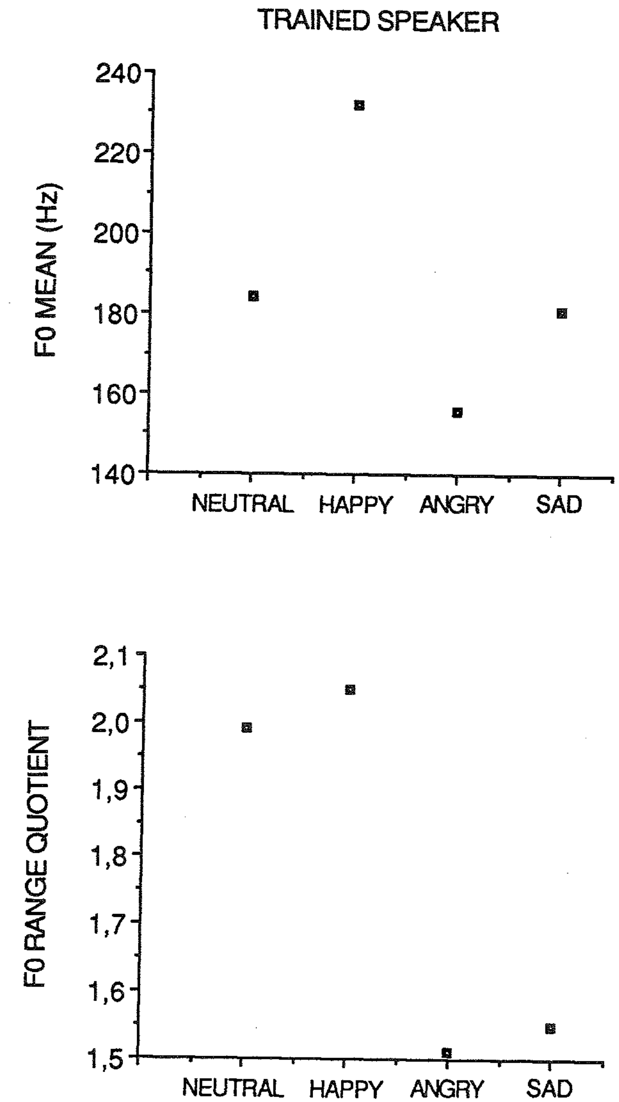

Figure 1. Plots of F0 mean and F0 range for four different moods of the test sentence *Nu flyttar jag* 'Now I'm going to move' produced by a trained speech pathologist (see text for measurement details).

token was repeated six times giving a total of 48 responses for each listener and both tests.

To correct for possible order and learning effects, two independent groups of listeners served as subjects (Group 1, n=12 and Group 2, n=15). Both groups of listeners were comprised of beginning students in linguistics who voluntarily participated in the tests. The stimuli for Test A were divided into two blocks of 12 stimuli each. The stimuli were randomized within each block, each stimulus occurring three times. The same procedure was applied to Test B. The tests were presented to Group 1 in the following order: A B B A. In other words, the listeners heard half of Test A followed by all of Test B and then the other half of Test A. The tests were presented to Group 2 in the reverse order: half of Test B, all of test A, and then the other half of Test B, i.e. B A A B.

In a forced choice task, the listeners were asked to categorize each stimulus as one of four possible moods (neutral, happy, angry or sad). The same situations and contexts were presented in the instructions to the listeners as had been presented to the speaker (S1). All eight test tokens were presented in a pre-test buffer block to acquaint the listeners with the task. Each stimulus was followed by a five-second response interval, and a longer pause (about ten seconds) followed each block. The tests were presented via a high quality tape recorder and loudspeakers. The listeners marked their responses on a prepared answer sheet.

Confusion matrices were made for each group and each test. Differences in the distributions of the confusions were tested statistically for each intended mood comparing Test A with Test B and Group 1 with Group 2 using the standard chi-square test of independence.

Results

Voice quality

Rating differences between the three recordings for each speaker were mainly found in the following parameters: *grating, hyperfunctional* and *hypernasality*. Figure 2 presents changes in mean ratings for S1, S2 and S3. The results here are to be seen as preliminary as the recordings are also being evaluated by a larger group of speech pathologists. The greatest improvements in rated voice quality occurred after two weeks in the parameters *grating* (S1) and *hyperfunctional* (S3). For S2, however, the mean rating for hypernasality increased by one point after six months.

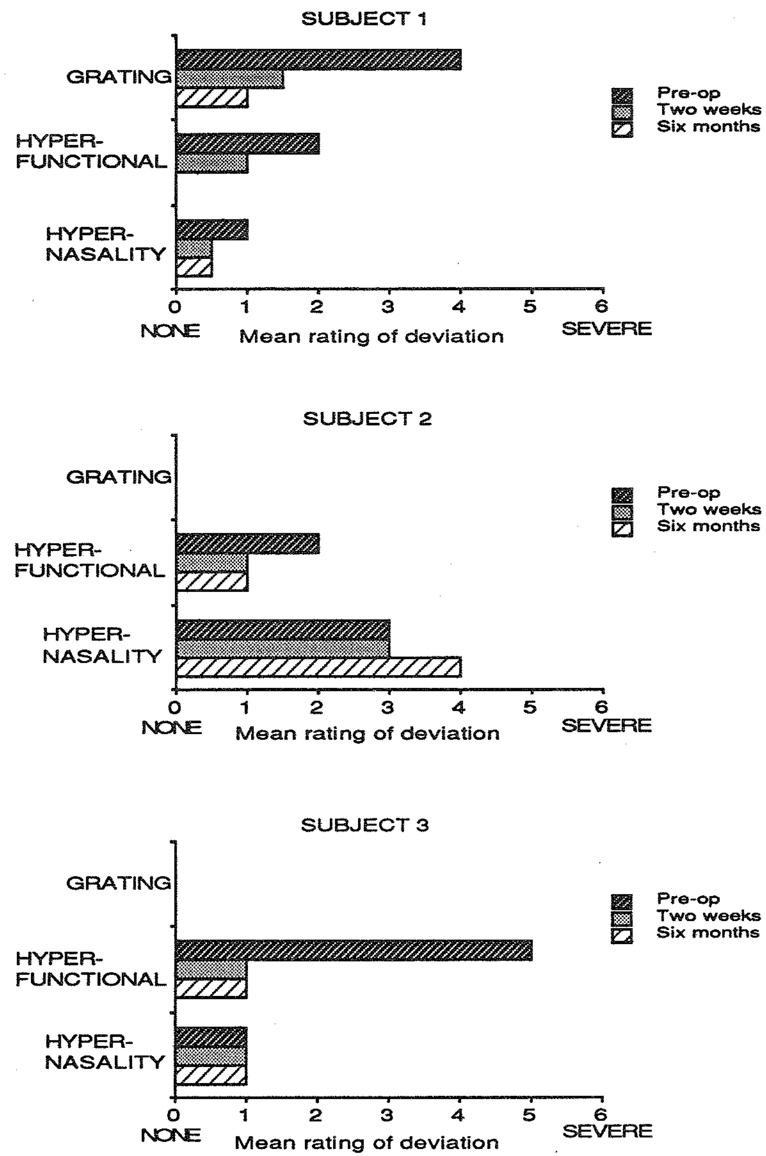

Figure 2. Results of the voice quality evaluation showing mean ratings of three parameters for the three subjects.

S1 was judged as having normal pitch and received a nearly normal general impression rating in all three recordings. S2 was judged as having normal pitch, but with a general impression rating falling about midway between normal and abnormal for all three recordings. For S3 pitch was judged as 'far too high' in the first recording (pre-operation) but nearly normal for the second and third recordings. The general impression for S3 was evaluated as lying about midway between normal and abnormal for the first two recordings and falling slightly closer to abnormal for the six-month recording.

F0 mean and range

Figures 3-5 show F0 mean and F0 range for the three speakers and the four intended moods for three recording sessions. After two weeks, S1 (Figure 3) lowers her F0 mean primarily for the neutral mood. This demonstrates use of F0 mean to separate neutral from the other three moods. After six months, however, she achieves a high degree of mood separation using F0 mean. Particularly noteworthy is the separation of happy and angry involving a substantial lowering of F0 mean for angry. It is also interesting to note that these two moods are among the most difficult for hearing impaired listeners to differentiate in perception (see Introduction, above). For the same speaker, no dramatic change is seen in F0 range after two weeks. However, after six months, the speaker is clearly able to use F0 range to help separate happy from the other moods. On the other hand, the similarity in F0 range between angry and neutral is also witnessed in confusion patterns in the listening test results presented below.

Results for S2 (Figure 4) are remarkably similar to those of S1. The greatest separation of all moods for both F0 parameters occurs six months after processor activation. As is the case for S1, F0 mean for angry and happy coincide after two weeks but are separated after six months as the speaker produces a higher F0 for happy and a lower F0 for angry. F0 range for happy also increases dramatically after six months.

Results for S3 (Figure 5) are not as uniform as for the other speakers. He does demonstrate a substantial lowering of F0 mean for neutral after six months and also produces a separation between angry and happy by lowering F0 mean for angry. Of the three speakers, S3 used the most variety of speech parameters (e.g. duration, intensity and voice quality) to distinguish between the moods. He also applied different strategies for

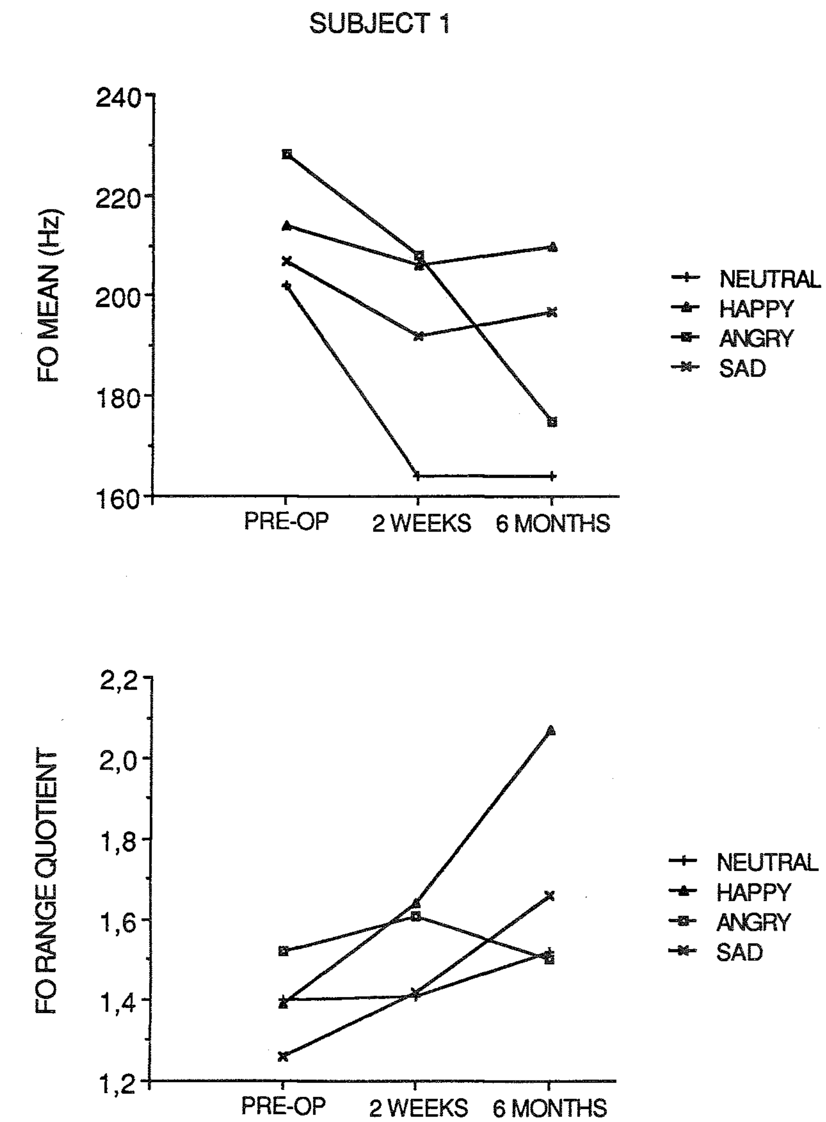

Figure 3. Plots of F0 mean and F0 range for four different moods of the test sentence *Nu flyttar jag* 'Now I'm going to move' produced by S1 showing changes as a function of time.

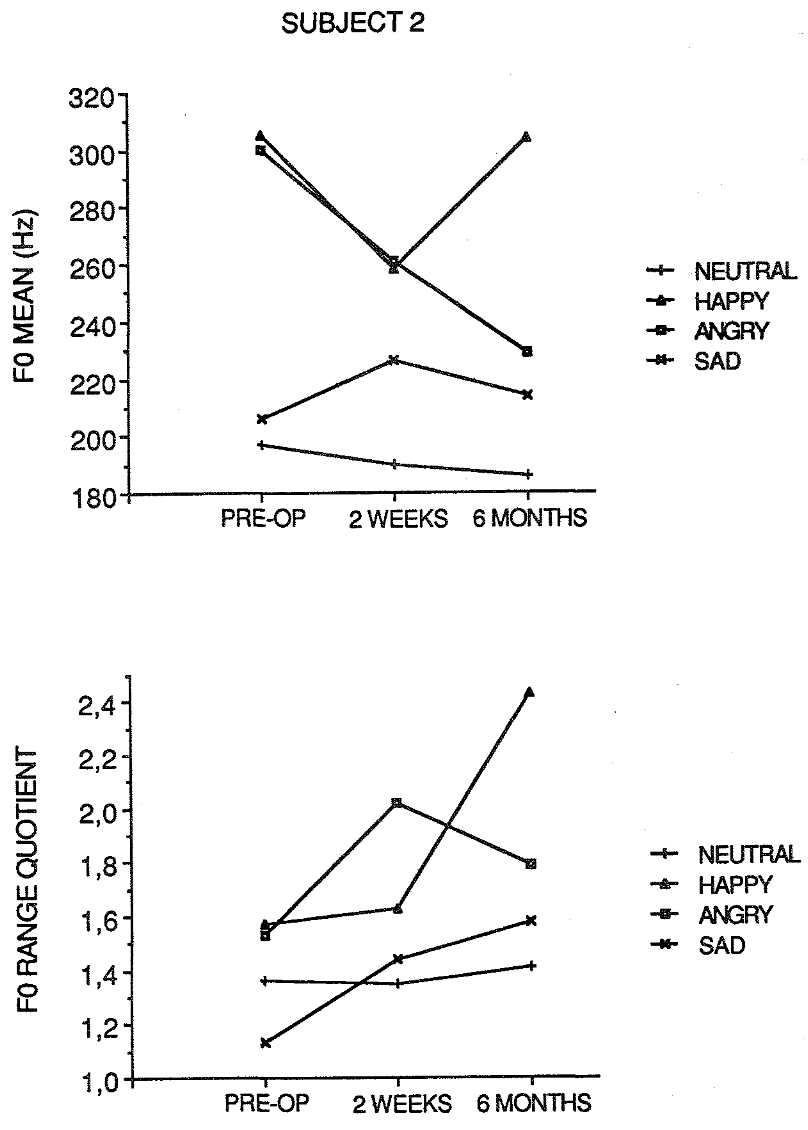

Figure 4. Plots of F0 mean and F0 range for four different moods of the test sentence *Nu flyttar jag* 'Now I'm going to move' produced by S2 showing changes as a function of time.

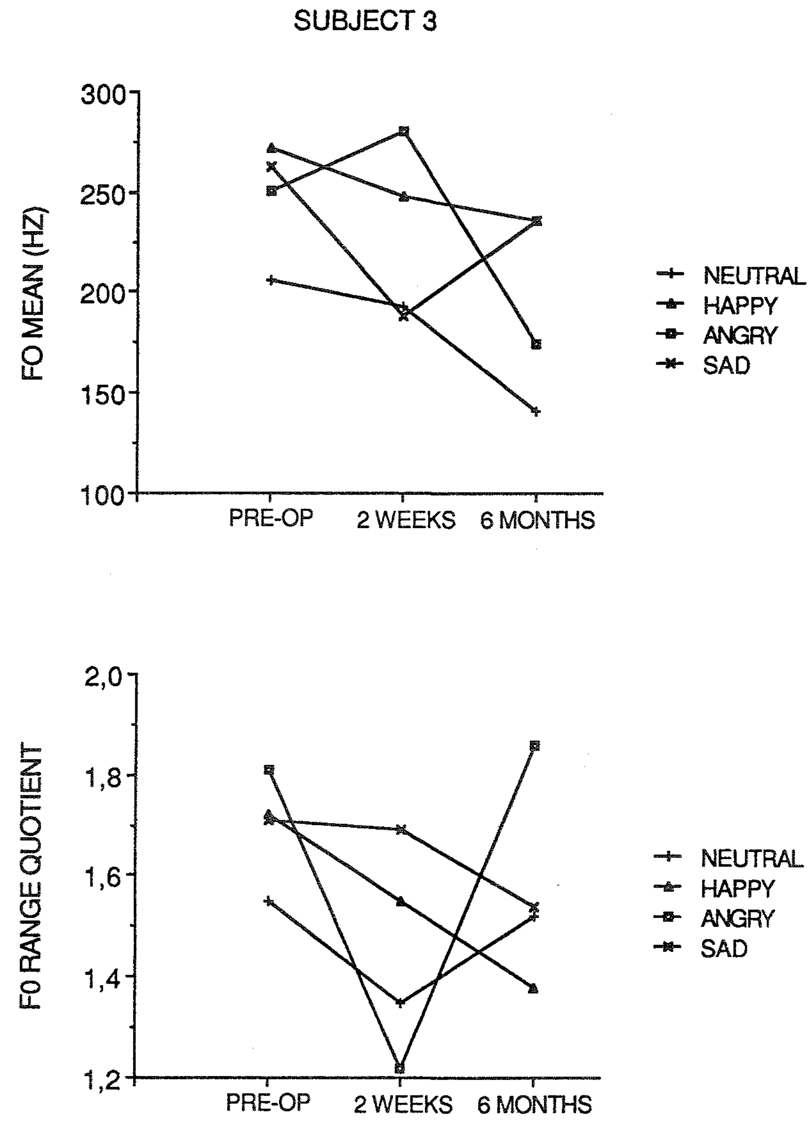

Figure 5. Plots of F0 mean and F0 range for four different moods of the test sentence *Nu flyttar jag* 'Now I'm going to move' produced by S3 showing changes as a function of time.

different recording sessions and made use of body language to help convey the different moods.

Listening test

Confusion matrices for the two listener groups (1 and 2) and the two tests (A and B) are presented in Figure 6. In Test A (pre-implant, left column),

RESPONSE IN %

GROUP 1 (A B B A) n=12

Pre-op (TEST A) — 53%

STIMULUS	A	H	S	N
A	60	5	35	0
H	14	11	10	65
S	0	0	82	18
N	36	1	3	60

6 months (TEST B) — 71%

STIMULUS	A	H	S	N
A	39	6	1	54
H	0	98	1	1
S	0	1	98	1
N	46	3	0	51

GROUP 2 (B A A B) n=15

Pre-op (TEST A) — 68%

STIMULUS	A	H	S	N
A	89	0	11	0
H	5	32	12	51
S	0	0	86	14
N	32	0	1	67

6 months (TEST B) — 73%

STIMULUS	A	H	S	N
A	42	1	8	49
H	7	93	0	0
S	0	0	100	0
N	42	1	1	56

Figure 6. Confusion matrices for the two listener groups and the two test presentations. Total scores for each group and test are given in percent under each matrix. A=Angry, H=Happy, S=Sad, and N=Neutral.

both listener groups had difficulties identifying the intended happy mood and the intended neutral mood. Group 1, however, had more difficulty with the angry mood than did Group 2. Total identification scores for Test A differed between the two groups with Group 1 attaining 53% correct identification compared to 68% for Group 2.

In Test B (six months, right column in Figure 6), the confusion distributions for both listener groups were very similar with both groups showing little confusion for the happy and sad moods but considerable confusion for angry and neutral. Total scores were also very similar with Group 1 attaining 71% correct identification compared to 73% for Group 2.

A comparison of total identification scores for Test A with Test B by group shows a considerable improvement for Test B and Group 1 with a great increase in the identification score for the happy mood contributing the most to the improvement. The total score increase for Test B was much less for Group 2, however, where a substantial decrease in the identification score for angry in Test B offset the increase in the happy mood identification. It is clear that Test A is more sensitive to order and learning effects than is Test B. However, the bidirectionality of results for the happy and angry moods renders the use of total identification scores less revealing than the patterns of confusion.

For both groups, the patterns of confusion are much cleaner for Test B than for Test A. The essential confusion in Test B was symmetrically divided between angry and happy (i.e. intended angry confused with neutral and intended neutral confused with angry), while for Test A, confusions differed between intended and perceived moods. (e.g. intended angry was confused with sad, intended happy and sad with neutral, and intended neutral with angry).

These differences between the two tests are also confirmed by statistical analysis using the chi-square test of independence. Table 1 shows levels of significance comparing the confusion distributions between Test A and Test B and between listener Group 1 and Group 2. Distribution differences between the two tests were highly significant for both groups for all moods except neutral. Differences between listener groups were only significant for angry and happy in Test A.

Table 1. Results of the standard chi-square test of independence comparing confusion distributions between Test A and Test B (upper panel) and between Group 1 and Group 2 (lower panel). Degrees of freedom are indicated in parentheses. Where possible, cells having an expected frequency of less than 5 were grouped together with the neutral category.

TEST A vs. TEST B

Stimulus	Group 1		Group 2	
ANGRY	χ^2 (2) = 24	p<.001	χ^2 (1) = 41	p<.001
HAPPY	χ^2 (2) = 108	p<.001	χ^2 (3) = 84	p<.001
SAD	χ^2 (1) = 7	p<.01	χ^2 (1) = 12	p<.001
NEUTRAL	χ^2 (1) = 1	p>.05 ns	χ^2 (1) = 2	p>.05 ns

GROUP 1 vs. GROUP 2

Stimulus	Test A		Test B	
ANGRY	χ^2 (2) = 20	p<.001	χ^2 (2) = 1	p>.05 ns
HAPPY	χ^2 (3) = 14	p<.01	χ^2 (1) = 1	p>.05 ns
SAD	χ^2 (1) = .2	p>.05 ns	χ^2 (1) = 1	p>.05 ns
NEUTRAL	χ^2 (1) = .1	p>.05 ns	χ^2 (1) = .09	p>.05 ns

Discussion

The results of this study document a change in the use of fundamental frequency and an improvement in voice quality characteristics for three postlingually deafened speakers after the activation of their cochlear prostheses. Although preliminary, the voice quality evaluations show the greatest improvement in the voice parameters *grating* and *hyperfunctional* after two weeks with less or no improvement between two weeks and six months (see Figure 2). With the exception of *hypernasality* for S2 all voice quality parameters are already approaching a rating of no deviation after two weeks leaving little room for substantial improvement at six months. It is also important to note that with the exception of *hypernasality* for S2 there is no increase in deviation between two weeks and six months. This indicates that the speakers are able to exercise long term control over these voice quality parameters.

While it is of course impossible to say whether these effects are a result of voice training or the implant, we prefer to see voice training as an integral part of the fitting of an implant. Changes in the particular voice parameters studied are probably a result of voice training made possible by the auditory feedback provided by the implant. While auditory feedback alone can result in changes in speech behavior (cf. Lane, Perkell, Svirsky & Webster 1991) these changes can most likely be accelerated by voice training. Many of our patients have also expressed initial insecurity concerning the perception of their own speech. Here, voice training can play an important role in the initial adaptation and recoding of auditory feedback.

As contrasted to the voice quality effects, the pitch analyses of the production of different moods shows the most dramatic changes occurring between two weeks and six months. Here we can conjecture that voice training initializes a learning process which continues with the use of the implant. The separation of the four moods exploiting fundamental frequency is not an uncomplicated task as witnessed by an informal experiment where we asked untrained students to produce the four moods. The expression of the moods involves a conventionalized use of intonational parameters and requires a considerable amount of control over fundamental frequency. Two weeks is probably not long enough to sufficiently master the F0 parameters required for mood separation. It is also interesting to note that the two moods which coincide most often in the pre-operative and two-week production of S1 and S2 are happy and angry, two moods which hearing impaired listeners and implant users have considerable difficulty in perceiving (cf. Öster & Risberg 1986, House 1990, 1991). Separation of these two moods in production may facilitate separation in perception.

The listening test, while only using production from one speaker (S1) clearly shows the increase in control of F0 from the pre-operative recording to the six-month recording. Although the speaker did not succeed in separating angry from neutral in the six-month recording, the patterns of confusion are much simpler and consistent than those from the earlier recording. The dramatic improvement in expressing happiness is certainly noteworthy, and the social implications of the expression of this mood are not to be underestimated.

In conclusion, the combined results of this study are encouraging for the users of cochlear implants. It is clear that a combination of voice training and a cochlear prosthesis can result in considerable improvement in voice quality and fundamental frequency control for postlingually deafened speakers.

Acknowledgments

This research was carried out under a grant from the Bank of Sweden Tercentenary Foundation to the Department of Linguistics and Phonetics, Lund University, Gösta Bruce, principle investigator. Recordings and speech training were included in and financed by the Cochlear Implant Program, Department of Audiology, Lund University Hospital. Sten Harris fitted the prostheses, and Bengt Almquist activated and adjusted the processors. We would like to thank all the members of the Cochlear Implant Team for valuable discussions, our three subjects for participating in the experiments, the speech pathologists who evaluated voice quality, the students who participated in the listening tests and Jan-Olof Svantesson for advice on statistical analysis.

References

Ball, G. & A. Faulkner. 1989. 'Speech production of postlingually deafened adults using electrical and acoustic speech pattern prosthesis'. In *Speech, Hearing and Language: Work in Progress* 3, 13-32. University College London.

Bruce, G. 1982. 'Developing the Swedish intonation model'. *Working Papers* 22, 51-116. Department of Linguistics, Lund University.

Eriksson, L. 1990. 'New Phonetic Programs for Macintosh'. *Working Papers* 36, 73-80. Department of Linguistics, Lund University.

Fónagy, I. 1981. 'Emotions, Voice and Music'. *Research Aspects of Singing* 33, 51-79. Swedish Academy of Music.

Gårding, E. 1991. 'Intonation parameters in production and perception'. In *Proceedings of the Twelfth International Congress of Phonetic Sciences*, 1:300-304. Aix-en-Provence, France.

Gårding, E. & A. Abramson. 1965. 'A study of the perception of some American English intonation contours'. *Studia Linguistica* 19, 61-79.

Granström, B. & L. Nord. 1991. 'Ways of exploring speaker characteristics and speaking styles'. In *Proceedings of the Twelfth International Congress of Phonetic Sciences*, 4:278-281. Aix-en-Provence, France.

Hadding - Koch, K. 1961. *Acoustico-phonetic studies in the intonation of Southern Swedish.* Lund: Gleerup.

Hammarberg, B. 1986. *Perceptual and acoustic analysis of dysphonia.* Doctoral dissertation, Huddinge University Hospital, Dept. of Logopedics and Phoniatrics, Stockholm.

House, D. 1990. 'On the perception of mood in speech: Implications for the hearing impaired'. *Working Papers* 36, 99-108. Department of Linguistics, Lund University.

House, D. 1991. 'On hearing impairments, cochlear implants and the perception of mood in speech'. *PERILUS* XIII, 125-128. Institute of Linguistics, University of Stockholm.

Lane, H., J. Perkell, M. Svirsky & J. Webster. 1991. 'Changes in Speech Breathing Following Cochlear Implant in Postlingually Deafened Adults'. *Journal of Speech and Hearing Research* 34, 526-533.

Öster, A-M. 1988. 'Changes in speech with use of an implant'. *Speech Transmission Laboratory - Quarterly Progress and Status Report (STL-QPSR)* 4/1988, 13-22. Department of Speech Communication and Music Acoustics, Royal Institute of Technology, Stockholm.

Öster, A-M., & A. Risberg. 1986. 'The identification of the mood of a speaker by hearing impaired listeners'. *Speech Transmission Laboratory - Quarterly Progress and Status Report (STL-QPSR)* 4/1986, 79-90. Department of Speech Communication and Music Acoustics, Royal Institute of Technology, Stockholm.

Plant G. & A-M. Öster. 1986. 'The effects of cochlear implantation on speech production. A case study'. *Speech Transmission Laboratory - Quarterly Progress and Status Report (STL-QPSR)* 1/1986, 65-86. Department of Speech Communication and Music Acoustics, Royal Institute of Technology, Stockholm.

Scherer, K.R., D.R. Ladd & K.E.A. Silverman. 1984. 'Vocal cues to speaker affect: Testing two models.' *Journal of the Acoustical Society of America* 76:5, 1346-1356.

Skinner, M.W., L.K. Holden, T.A. Holden, R.C. Dowell, P.M. Seligman, J.A. Brimacombe & A.L. Beiter. 1991. 'Performance of Postlinguistically Deaf Adults with the Wearable Speech Processor (WSP III) and Mini Speech Processor (MSP) of the Nucleus Multi-Electrode Cochlear Implant'. *Ear and Hearing* 12, 3-22.

Touati, P. 1987. *Structures prosodiques du suédois et du français.* Lund: Lund University Press.

Waltzman, S. & I. Hochberg. 1990. 'Perception of Speech Pattern Contrasts Using a Multichannel Cochlear Implant'. *Ear and Hearing* 11:1, 50-55.

Williams, C.E. & K.N. Stevens. 1972. 'Emotions and speech: Some acoustical correlates'. *The Journal of the Acoustical Society of America* 52:4, 1238 -1250.

Lund University, Dept. of Linguistics
Working Papers 38 (1991), 133-54

A Minimalist Parser for Fast Partial Analysis

Christer Johansson

This work is based on Blank 1985, 1989, presenting a method for making a machine that can provide fast analysis of natural language with bounded resources. The machine is a new kind of finite automaton and the formalism is called Register Vector Grammar, RVG. The machine is claimed to provide analysis of natural language in linear time. Linear time is a considerable improvement compared to polynomial time, which is what is currently available for parsers. In order to make linear time, the depth of the analysis is restricted. The machine has been constructed to model performance rather than competence thus the approach introduces restrictions on the input language. The restrictions can be motivated by the fact that humans can be viewed as limited devices, in particular their short-term memory is limited. As an example, an RVG grammar for agreement in Swedish nominals is developed.

Introduction

A great deal of language analysis can be done with a simplified model that does not have the overhead of context-free grammar. It can be desirable to split up the big analysis module needing context-free power into several smaller analyzers. One reason is that the smaller modules need less computer capacity; another, more important reason, is that a small module is easier to maintain. In any big analysis module there will always be conflicts among the rules and in the lexical database. It is therefore better to build up the analysis in small stages. There have been several experiments with late assigning of grammatical roles and incremental parsing at the Dept. of Linguistics, Lund (Sigurd et al. 1989, Sigurd 1990) and more recently, experiments with the RVG model. These experiments have searched for a psychologically plausible model for human language performance.

The parser which this paper presents is heavily influenced by Register Vector Grammar as described by Blank 1985, 1989. This specific implementation of RVG is, however, constructed to aid limited partial analysis.