PROSODY AND VOICE QUALITY IN THE EXPRESSION OF EMOTIONS

Elisabeth Zetterholm

Lund University
Dept. of Linguistics and Phonetics
Helgonabacken 12
S-223 62 LUND
Sweden
E-mail: Elisabeth.Zetterholm@ling.lu.se

ABSTRACT

Terms for voice quality or phonation types for use in normal speech often come from studies of pathological speech (laryngeal settings) and it is hard to describe voice quality, especially the variations of a normal voice. In normal speech we use different voice qualities both for linguistic distinctions in some languages, prosodically as a boundary signal, socially depending on social and regional variants and paralinguistically in attitudes and emotions. This paper shows some reference types of voice qualities, recorded by a trained phonetician, and their acoustic correlates. In a pilot study a male actor recorded four attitudinally neutral sentences using five different emotions which are being compared to his neutral voice. It is evident that voice quality, as well as rhythm and intonation, plays an important role in giving the impression of different emotions.

1. INTRODUCTION

The human voice is flexible and everybody changes his/her voice quality in different social contexts and situations. Phonetic habits are also a part of our personal style. It is possible to change the voice by changing the phonetic settings, but there are some anatomical difficulties which are almost impossible to overcome and such components of voice quality are innate and outside the speaker's control. We use our voices to express emotions and we change our voice quality to strengthen the impression of emotions. We change the prosody (speech rhythm and intonation), but we even use voice quality as a phonetic cue for the listener. Most studies, focusing on the expression of emotions, have concentrated on the prosody, F0 and duration patterns (Vroomen, Collier & Mozziconacci 1993, Mozziconacci 1995, Mozziconacci and Hermes 1998). The results of their experiments showed that it is possible to express emotions in synthetic speech by manipulating pitch and duration, but the aspect of voice quality must be taken into consideration as well. The importance of voice quality is not fully considered in these studies, but it is clear, though, that voice quality and phonation type is fundamental in the expression of emotions.

2. VOICE QUALITY IN NORMAL SPEECH

We use different voice qualites in normal speech depending on the situations, but there are also linguistic distinctions in some languages. Creaky voice is one common quality with linguistic distinctions, as in some languages in Africa where it is used for phonological contrast to distinguish types of consonants and vowels, from sounds with normally voiced phonation (Laver 1994). Nasal voice also has a linguistic distinctions, in different languages, Hindi and Yoruba for example (Laver 1994).

Prosodically voice quality can be used as a boundary signal, and for example creaky voice is an 'end-of-utterance' phenomenon (Henton and Bladon 1988, Laver 1994, Bruce 1998). Variation in voice quality is then typically used together with other phonetic cues as final lengthing, drop in F0, decreased intensity and possibly pausing.

Wardhaugh (1998) talks about 'human speech patterns' and discusses if there are some unique sets of items or patterns, which make it possible to identify regional and social dialects. Elert and Hammarberg (1991) have researched some Swedish dialects and found out that there are obvious differences between the regional dialects in both average F0 and voice quality. Interesting studies have been done of social markers in speech focusing both on grammatical features, prosody and voice quality (Labov 1974, Esling 1978, Laver and Trudgill 1979, Scherer 1979, Henton and Bladon 1988, Laver 1991, Pittam and Scherer 1993, Pittam 1994). Results have shown that phonological systems, pronunciations and how often they are used differ between social groups and social status often correlates with a difference in laryngeal settings. Creaky voice is often said to be a male characteristic voice quality and breathy voice seems to be more common among female speakers (Lindblad 1992). Nasal voice is sometimes correlated to high social prestige, in RP English for example.

Voice quality is used paralinguistically in attitudes and emotions. To signal bored resignation a creaky voice is used in the RP dialect in English and breathy voice appears to be signalling intimacy in many languages (Lindblad 1992, Laver 1980, 1994). In many cultures harsh voice is used paralinguistically as a signal of anger and aggression (Lindblad 1992, Laver 1980, 1994).

3. METHODOLOGY

For the recording and analysis of the expression of emotions it is common to use an actor, but it can result in exaggeration and unnatural speech because he is likely to be imitating the expressions. It would be interesting to study emotional utterances in spontaneous speech, but it may then be difficult to make comparisons with the speaker's neutral voice. Both auditory and acoustic analysis should be done and the waveform, as well as fundamental frequency, intensity, formant frequencies and voice quality must be analysed. The imitation of emotions by an actor may be analysed like imitation and impersonation are analysed, with pitch, formant frequencies, rhythm and voice quality (Zetterholm 1997). Analysis by synthesis and perceptual tests may be necessary to understand the production and perception of emotions. It is important to consider serveral aspects, such as social, regional and psychological in the study of voice quality. Only then can we get a complete picture of the vocal communication of emotions.

4. TYPES OF VOICE QUALITY

Terms for voice quality or phonation types for use in normal speech often come from studies of pathological speech (laryngeal settings) and it is hard to describe voice quality, especially the variations within the range of a normal voice. Articulatory settings and voice quality settings and their acoustic correlates have been studied by, among others, Honikman (1964) and Laver (1980, 1991, 1994). de Krom (1994) has studied the acoustic correlates of breathiness and roughness. One of his concluding remarks is that 'more knowledge about the nature and magnitude of voice quality variations as they occur in the conversational speech of individual speakers' is important for voice quality evaluation. In the description of voice quality it is important to consider both laryngeal and supralaryngeal settings.

The most common types of phonation types (laryngeal settings) by Laver (1980) are:

- modal or normal voice
- breathy voice, as opposed to tense or strained voice, with a high rate of air-flow
- creaky voice with very low frequency and usually irregularly spaced in time
- harsh voice with a normal fundamental frequency but aperiodicity or noise in spectrum
- tense or strained voice with a low rate of air-flow (often described like a 'metallic voice')

There are also a number of compound phonation types and the combined breathy and tense voice is an example in my reference types. In spite of the suitability of these terms we still lack a more complete typology for description of voice quality in normal speech.

Recordings of different voice qualities (phonation types) have been made by a trained Swedish phonetician. There are obvious differences already in the waveforms for the six voice qualities (modal, breathy, creaky, harsh, tense and the compound breathy and tense voice). These represent reference types for my study and are shown in figures 1-6 (the same utterance).

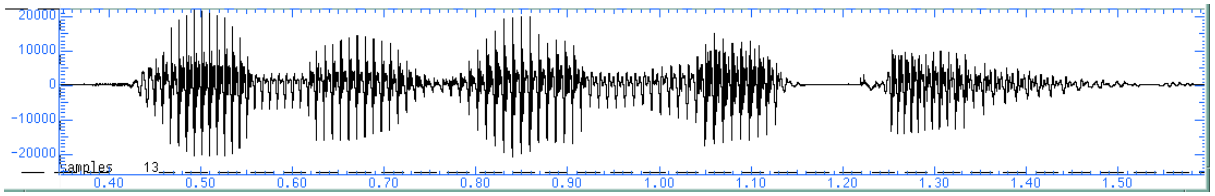

Figure 1: Modal voice

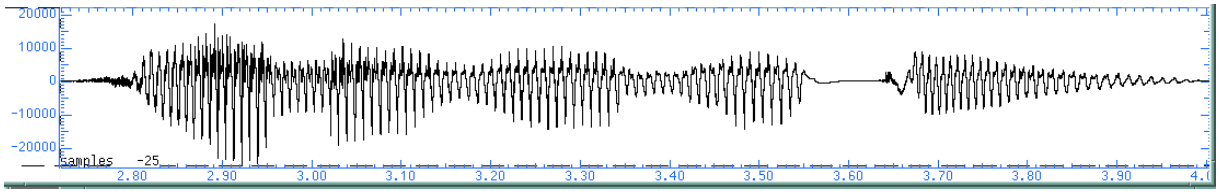

Figure 2: Breathy voice

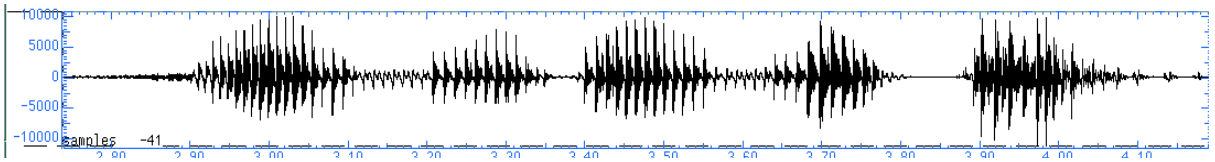

Figure 3: Creaky voice

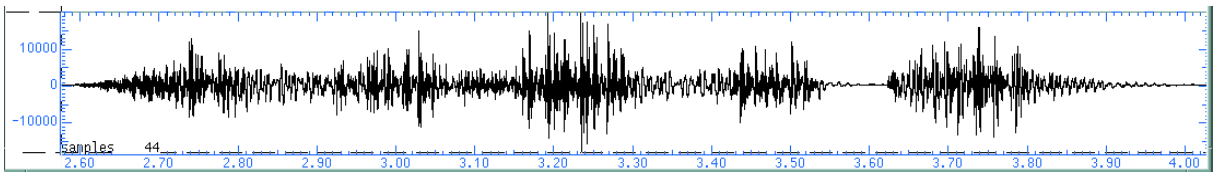

Figure 4: Harsh voice

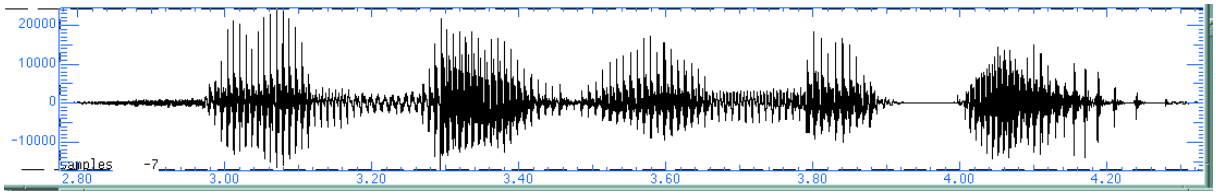

Figure 5: Tense voice

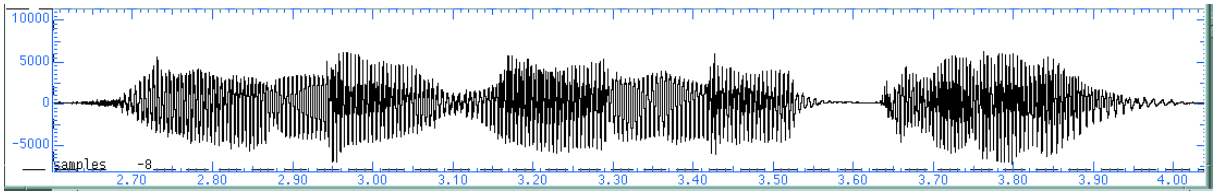

Figure 6: Compound breathy and tense voice

5. A STUDY OF EXPRESSION OF EMOTIONS

A male actor recorded four attitudinally neutral sentences using five different emotions (anger, anxiety, indifference, joy and sadness), in addition to his neutral voice. The recordings were made with a DAT-recorder. They are being analysed both auditorily and acoustically using the analysis programme ESPS/Waves+.

The analysis of the recordings of this male actor, so far, shows that voice quality, as well as pitch and speech rhythm, plays an important role in giving the impression of different emotions, in accordance with other studies. In the emotion of anger he uses a rather high pitch and a high intensity. The impression is strengthened by the slow speech rate and the very distinct articulation. The speaker's voice quality is taut but it sounds like he is gritting his teeth as if he is irritated. When people are anxious or in a weak position they often use a higher pitch, than they usually do, and some times even a breathy voice quality. This actor uses this attribute for the expression of anxiety in these recordings. Indifference and sadness have a low pitch and intensity and there is only a small pitch range. The speech rate is rather high with a lax articulation and he sounds really depressed. He uses a creaky voice quality and that strengthens the impression of the emotion. In the expression of joy these recordings show variations in intonation and a rather high pitch compared to the recordings with his neutral voice. The auditory impression is that he is smiling because even the rounded vowels sound unrounded, and an analysis of the formant frequencies will show how they are affected.

Figures 7-9 show three examples of voice qualities in the expression of emotions referring to my reference types. ('I kväll ska vi gå på bio' ('We are going to the cinema this evening')

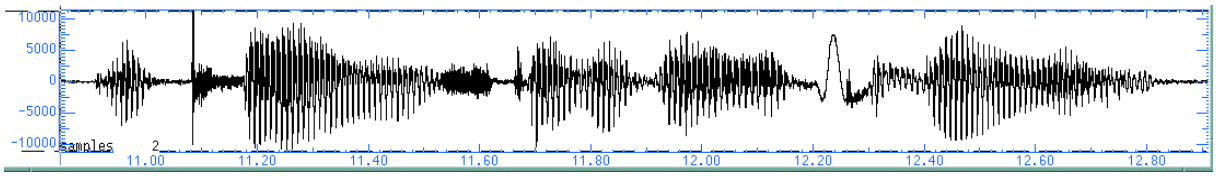

Figure 7: Neutral voice

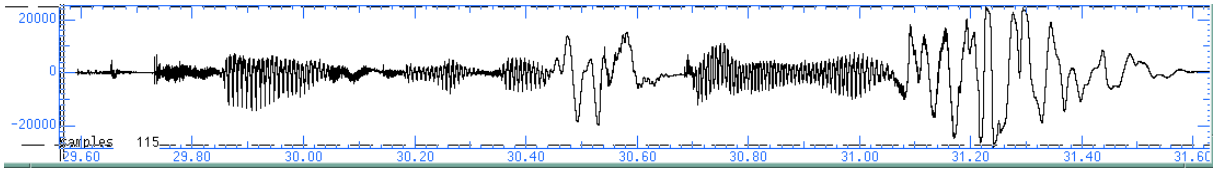

Figure 8: Anxiety - breathy voice

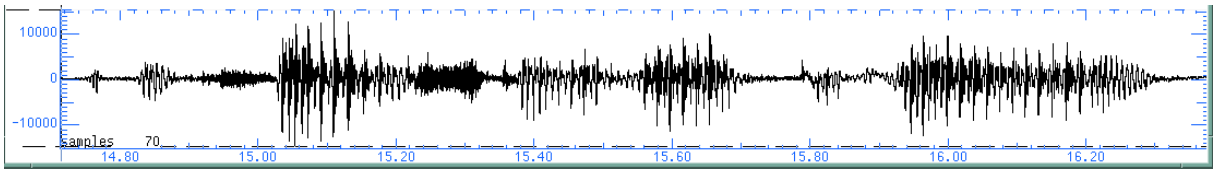

Figure 9: Sadness - creaky voice

6. REFERENCES

1. Bruce, G., *Allmän och svensk prosodi*, Praktisk lingvistik 16, Department of Linguistics and Phonetics Lund University, 1998.
2. Elert, C., and Hammarberg, B., ''Regional voice Quality Variation in Swedish,'' *Proc XIIth Intern Congr Phonetic Science*, Aix-ex-Provence, Frankrike, 4:418-420, 1991.
3. Esling, J., *Voice Quality in Edinburgh: A Sociolinguistic and Phonetic Study*, Unpublished doctoral dissertation, Edinburgh University, Edinburgh, 1978.

4. Henton, C. , and Bladon, A., ‘‘Creak as a Sociophonetic marker,’’ *Language, speech and mind. Studies in honour of Victoria A. Fromkin*, L. Hyman & C. Li (eds), Routledge, London, 1988.

5. Honikman, B., ‘‘Articulatory settings,’’ *In honour of Daniel Jones*, D. Abercrombie et al. (eds). Longmans, London, 1964.

6. de Krom, G., *Acoustic Correlates of Breathiness and Roughness. Experiments on voice quality*, Research Institute for Language and Speech, Utrecht University, 1994.

7. Labov, W., ‘‘Logiken i “Nonstandard English’’,’’ *Barnspråk i klassamhälle*, B Loman. LiberLäromedel, Lund, 1974.

8. Laver, J., *The Phonetic Description of Voice Quality*, Cambridge University Press, 1980.

9. Laver, J., *The Gift of Speech. Readings in the Analysis of Speech and Voice*, Edinburgh University Press, 1991.

10. Laver, J., *Principles of phonetics*, Cambridge University Press, 1994.

11. Laver, J., and Trudgill, P., ‘‘Phonetic and linguistic markers in speech,’’ *Social markers in speech*, K.R. Scherer & H. Giles (eds). Cambridge University Press, 1979.

12. Lindblad, P., *Rösten*. Lund: Studentlitteratur, 1992.

13. Mozziconacci, S., ‘‘Pitch Variations and Emotions in Speech,’’ *Proceedings ICPhS95*, Stockholm, 1, 178-181, 1995.

14. Mozziconacci, S., and Hermes, D., ‘‘Study of Intonation Patterns in Speech Expressing Emotion or Attitude: Production and Perception’’, *IPO Annual Progress Report*,1998.

15. Pittam, J., *Voice in Social Interaction*. SAGE Publications, California, 1994.

16. Pittam, J., and Scherer, K.R., ‘‘Vocal Expression and Communication of Emotion’’, *Handbook of emotions*, M. Lewis & J. M. Haviland (eds). The Guilford Press, New York, 1993.

17. Scherer, K.R., ‘‘Personality markers in speech’’, *Social markers in speech*, K.R. Scherer & H. Giles (eds). Cambridge University Press, 1979.

18. Vroomen, J., Collier, R., and Mozziconacci, S., ‘‘Duration and Intonation in Emotional Speech’’, *Proceedings Eurospeech* 93, Berlin, 577-580, 1993.

19. Wardhaugh, R., *An Introduction to sociolinguistics*, Blackwell Publishers, Oxford, 1998.

20. Zetterholm, E., ‘‘Impersonation: A Phonetic Case Study of the Imitation of a Voice’’, *Working Papers 46*, Lund University, 1997.