Fonetik 2000, Swedish phonetics meeting in Skövde, May 24-26, 2000

(Expanded version, internal TMH report)

Half a century in phonetics and speech research

Gunnar Fant

Department of Speech, Music and Hearing, KTH, Stockholm, 10044

Abstract

This is a brief outlook of experiences during more than 50 years in phonetics and speech research. I will have something to say about my own scientific carrier, the growth of our department at KTH, and I will end up with an overview of research objectives in phonetics and a summary of my present activities.

Introduction

As you are all aware of, phonetics and speech research are highly interrelated and integrated in many branches of humanities and technology. In Sweden by tradition, phonetics and linguistics have had a strong position and speech technology is well developed and internationally respected. This is indeed an exciting field of growing importance which still keeps me busy. What have we been up to during half a century? Where do we stand today and how do we look ahead? I am not attempting a deep, thorough study, my presentation will in part be anecdotal, but I hope that it will add to the perspective, supplementing the brief account presented in Fant (1998)

The early period 1945-1966

KTH and Ericsson 1945-1949

I graduated from the department of Telegraphy and Telephony of the KTH in May 1945. My supervisor, professor Torbern Laurent, a specialist in transmission line theory and electrical filters had an open mind for interdisciplinary studies. My thesis was concerned with theoretical matters of relations between speech intelligibility and reduction of overall system bandwidth, incorporating the effects of different types of hearing loss.

This work paved the way for my employment 1945-1949 at the acoustics laboratory of the Ericsson Telephone Company. They needed basic knowledge of the formant structure of Swedish speech sounds as a support to intelligibility tests and for predictions of effects of selected frequency band elimination in telephony. So I was given free hands to construct speech analysis equipment and go ahead with studies of the time-frequency-intensity distributions of Swedish speech sounds and sentences. My tools were primitive but effective. I constructed a wave analyser with manual continuous variation of the centre frequency and a choice of narrow or broad bandwidth setting. The subject was seated in a proper an-echoic chamber. A special feature, usually not retained in present day's routines, was absolute intensity calibrations in all recordings. Vowels and sonorants were analysed with the subject sustaining the sound for 3 seconds. Time-frequency-intensity spectrograms were manually compiled from oscillograms of intensity versus time in a substantial number of frequency bands. These quantified patterns closely resembled the spectrograms I had just learned about from the Bell Labs Visible Speech.

Except for early internal Ericsson reports, these data were not published until 1959, in an Ericsson Technics monograph, Fant (1959), which also included basic speech synthesis theory, and was added to my two other thesis publications, (Fant, 1960) and Fant (1958). One outcome of the Ericsson work was a graph in the frame of an audiogram of the major frequency-intensity formant regions of Swedish speech sounds at a specified talking distance. This has become a standard reference in audiology referred to as the "formant banana". The acoustic-phonetic data acquired at that time was included in a compendium on Swedish Phonetics (Fant 1957) which was adopted for phonetic courses at the Stockholm University.

MIT 1949-1951

My Ericsson work included detailed spectral sections of the burst of stop consonants. When I came to America in late 1949 these data were presented at a seminar at Harvard University. In the audience was Roman Jacobson who found my data to fill a missing link in his feature theory, the realisation of compactness and gravity in consonants. This is how our co-operation started. It was followed up by the teamwork of Jacobson, Fant, Halle, (1952). My later views on distinctive features were published in Fant (1973).

The stay at MIT 1949-1950 was extremely rewarding. This was a truly pioneering era in speech research as an outgrowth from linguistics, electrical circuit theory, psychoacoustics and information theory. A fascinating experience was the Pattern Play back at Haskins laboratories. It had a pedagogical value unsurpassed by recent year digital technology. With my internalised knowledge of spectral representations in Swedish I could paint a sentence and have it replayed. It was monotone but understandable. My work on speech production theory had a good start at that time. With the support of Roman Jacobson and Morris Halle X-ray studies of sustained articulations of a Russian subject were pursued. These became the foundation for my book, Fant (1960).

The Speech Transmission Laboratory, 1951-1966

Back in Sweden in 1951 the nucleus of the Speech Transmission Laboratory was formed. A project already started in 1948 when I was part time employed at KTH was to carry out statistics of the relative occurrence of letters, phonemes and words in written material and in telephone conversations. This data was now processed by my first employee, Marianne Richter, who also became the secretary and finance officer of the group. Results from our survey were published in the proceedings of the Linguistic conference in Oslo 1958. The main part is retained in my speech communication course compendium of 1967.

It is of some historical interest that in 1947 when we acquired the telephone speech material from the Swedish Telecom lab, it was stored on an old-fashioned wire recorder. After we had transcribed the text the wire had to be cut into pieces to preserve the privacy.

Our first synthesiser OVE 1, (Fant, 1953) was capable of producing quite natural vowels and simple sentences made up of transitional voiced gestures. Its name was given in a program over the Swedish radio in 1953. The interviewer asked if the machine had a name. Well, I could make it produce the name OVE and so it was christened. Later it was interpreted as "Orator Verbalis Electricus". Our present OVE 1 designed by Johan Liljencrants in the 1970 still maintains a substantial pedagogical value.

OVE II, capable of producing unrestricted connected speech, appeared in 1961, (Fant and Martony, 1962). It was programmed from function lines drawn with conducting ink on a plastic sheet and employed vacuum tube electronics. A transistorised version was later developed by Johan Liljencrants. During the 1962 international speech communication seminar in Stockholm it created a substantial interest. Janos Martony and our guest John Holmes had demonstrated high quality copy synthesis.

An incident is worth while mentioning. One night during the conference, without our consent, a representative of the Melpar Company in USA, Joseph Campanella, took photographs and measurements of the function generator, which enabled them to produce a rather exact copy. A year later their synthesiser EVA was on sale. Arne Risberg who had been involved in our original design and I now made a claim of compensation, which was granted by a small royalty. However, not many Eva systems were sold and our revenue was small.

At this time an important object of speech technology was vocoder systems for bandwidth reduction in telephony to be combined with secrecy coding. We were consulted by the Ericsson Company in the design of their channel vocoder. An alternative approach was formant coded analysis-synthesis systems, which potentially is closer to production theory. But formant vocoders were not a success. A well known system, that of the Melpar Company. sounded natural but was quite unintelligible, an instance of a failure of both the model and the formant tracking. More sophisticated systems are now considered in bandwidth reduction schemes.

The later period, 1967- present days

A major break-through in Speech Technology was text-to-speech synthesis. It was developed by Rolf Carlson and Björn Granström who worked out a language and phonetics program interfacing the OVE II (Carlson and Granström, 1975, 1991). Applications in reading aids for the blind and speaking aids for vocally handicapped have been of great importance and the prototype has paved the way for applications in general information services.

A financial support was received in 1959-1973 through American grants from the US Air Force and US army and from the National Institute of Health. These made possible a substantial increase of our basic research, which continued, with increasing support from Swedish funding agencies.

Research in speech production, speech synthesis and speech perception including auditory modelling was promoted, and a wide range of handicap applications were considered. Music acoustics developed strongly and gradually attained its important role in general and clinically oriented research in the human voice.

Our Speech Transmission Laboratory Quarterly Progress and Status Report, STL-QPSR, now TMH-QPSR, was started in 1960 and has been distributed free of charge, at present to over 800 individuals and institutions within 50 countries Over the years we have had a large number of contributors. Phonetically oriented research was initiated by Björn Lindblom and Sven Öhman already in 1960. The most productive authors in terms of number of contributions, see the cumulative index in STL-QPSR 1/1995, have been: Eva Agelfors, T.V-Ananthapadmanabha, Mats Blomberg, Rolf Carlson, Kjell Elenius, Gunnar Fant, Frans Franson, Karoly Galyas, Jan Gauffin-Lindqvist, Björn Granström, Brita Hammarberg, Sharon Hunnicutt, Inger Karlsson, Johan Liljencrants, Qi-guang Lin, Björn Lindblom, Janos Martony, Lennart Nord, Arne Risberg. Karl-Erik Spens, Sven Öhman: In music acoustics the main contributors were Anders Askenfelt, Frans Fransson, Erik Janson, Johan Sundberg and Sten Ternström, the latter two also contributing to studies of voice and speech production. The largest number of contributions 1960-1995 are those of Sundberg (97) and Fant (77).

Working contacts were established with Russsian (St. Petersburg) and French (Grenoble) groups through symposia and frequent scientific visitor exchanges. The editor of the STL-QPSR was Si Felicetti who also had an important role in the organisation of scientific meetings in Sweden and abroad

The STL-QPSR made possible publication with a minimum of delay to a large number of recipients, but a disadvantage still remains. It often becomes a substitute for refereed publications in established journals. We have several major contributions in the STL-QPSR on speech production theory, speech analysis and speech perception which have not been published elsewhere. However, a major publication from that time is a chapter in Malmberg´s Manual of Phonetics, (Fant 1968) and the collection of articles in Fant (1973).

An activity well represented in the period 1960-1980 was auditory modelling (Carlson, Granström and Fant, 1970, Carlson and Granström, 1982, Carlson, Fant and Granström, 1975, Fant, 1978). Apart from studies in speech prosody we nowadays lack projects in auditory mechanisms. The symposium volume edited by Carlson and Granström (1982) contains several articles that deserve a renewed attention.

A peripheral problem which engaged me during this early period was to explain the distortion encountered in the speech of divers when breathing different gas mixtures adjusted for operation at a certain depth. While the gas mixture accounts for a simple Donald Duck linear frequency transposition, the pressure component at high depths produces a non-linear frequency scale compression perceived as a nasal quality (Fant and Lindqvist-Gauffin, 1968).

A minor, phonetically oriented, project I happened to get involved in was inspired by Arne Risberg. It was a phonetic feature structured hand-finger alphabet for non-hearing subjects (Fant, 1972). It has been tested in a Japanese school for the deaf.

There has always been a tension between forensic science and speech research. There exists a seeming but deceptive analogy between fingerprints and voice prints which would motivate spectrographic analysis as legal evidence. However, such applications have been refuted by a majority of speech scientists as being premature and insecure. In 1970 Stockholm was visited by the head of the FBI who gave an interview in Dagens Nyheter on the use of voiceprints for identifying terrorists. The newspaper turned to me for comments, which were quite negative. On the first page of Dagens Nyheter of January 16 our controversy was reviewed. There were two photographs, one of Edgar J. Hoover head of the FBI, and one of Gunnar Fant, suggested as a possible FBI enemy number one.

The close ties between linguistics, phonetics, speech technology, auditory research and music acoustics and their importance for a number of applied areas such as handicap aids, auditory and speech rehabilitation have been especially well developed in Sweden. A recognition was the funding through state research grants of a project named “Tal, ljud och hörsel” (Speech, sound and hearing) which sponsored some research and annual meetings in the years 1983-1985. When the TLH support terminated our annual Swedish phonetics conferences took over as a forum for these interdisciplinary meetings.

In recent years the major recipient of funding within our department has been the Center for Speech Communication (CTT). It has had a productive start, to a substantial part sponsored by Swedish industry. The main activity is development of interactive man-computer information systems. It involves work in speech synthesis with an added talking head display, speech recognition and language engineering.

But what about the basic research? It still prevails but not with the same breadth as in earlier periods. On the other hand, we have extended contacts with phonetics departments in Sweden, e.g. in prosody research and dialect studies. This brings me to the second part of my survey. Some thoughts about topics in phonetics and my own research interests.

Research objectives and trends

It would carry too far to perform a detailed review of developments and the status of Swedish phonetics. I will have the opportunity to contribute to a historically oriented survey, planned by Olle Engstrand. Over the years there has been a substantial progress and a close co-operation with speech technology. The most prominent area is prosody, which by tradition has a firm foundation in Swedish phonetics and is of considerable importance for text-to-speech synthesis and in language teaching. Co-operation networks involving KTH, and the Stockholm, Lund and Umeå universities are well established.

In the last 50 years, research in phonetics has expanded over a wide range of areas reflecting diverse interests, backgrounds and methods of approach. I shall take the liberty of executing a personal view of structures and trends.

One is the apparent diffusion of phonetics into many sub-areas of narrowly defined problems. On the other hand, there is a search for universal principles of language development and unifying principles in human behaviour, e.g. the relations of speech and language to other human activities.

However, the central object of general phonetics is the speech communication chain, the many stages of message encoding within the speaker, the transmission medium and the receiving partner. It can be studied in functional details, for instance with respect to vocal tract acoustics, articulatory gestures and properties of the auditory system, including transformational relations between links in direct or reverse order, e.g. articulatory interpretation of spectrograms. The latter technique is of great importance as a supplement to direct studies of articulatory dynamics. Our insight in brain functions of the speaker and the listener is limited and restricted to indirect observations and functional analogies.

All these aspects of the function of the speech chain can be referred to by a major category labelled MECHANISMS which is largely a matter of physiology, physics and sensory psychology. A second major category is introduced by adding the linguistic function and linguistic competence. This is the SPEECH CODE, the relation between message units and their realisation in the speech chain, involving both discrete linguistic units as in phonology and continuous as related to our social code of expressing attitudes, denotations and emotions (Fant, 1989).

A challenge for future research, and the ultimate aim of general phonetics, is to force the speech code, i.e. to predict the articulatory, acoustic and perceptual manifestation of any utterance given the message transcript and the particular language, dialect, speaker and situational context. One can conceive of this task as a very advanced project of deriving rules for text-to-speech synthesis. Our present knowledge of the speech code is incomplete and in part hidden in text-to-speech programs. We need an extension of code oriented analysis such as in the comprehensive early studies of Dennis Klatt for American English and of Carlson and Granström for Swedish.

In recent years, large data banks of spoken material have been collected, e.g. for studies of Swedish dialects and dialect independent speech recognition. They have been used mainly in automatic computer training, but have not been much exploited in acoustic-phonetic studies. However, in quest of the speech code, we need comprehensive analysis of all acoustics correlates, segmental and prosodic, which puts a limit to the size of a corpus. Up till now most attempts of this kind have been fragmentary, i.e. directed to some detail aspect only.

Irrespective of the aim of a project it can be pursued with more or less emphasis on the knowledge to be gained or to instrumental tools, either those available or new tools developed within the project. This dichotomy has been referred to as "knowledge driven" versus "instrumentally driven" research. It is temping to choose an established algorithm for deriving outputs from known inputs. An example is the plotting of locus equations, which are primarily of statistical interest and lack perceptual significance. Knowledge driven projects e.g. for multi-parameter investigations of the speech code, are more demanding.

We have a lack of knowledge but there is also a lag in knowledge transfer to consider. Present systems, especially for speech recognition, are not designed to make use of available knowledge on prosody and segmental constraints. The question is now whether we shall be able to effectively handle all new developing information in explicit rules or if we have to continue to rely on computers to learn the code for speech recognition? Furthermore, will concatenated synthesis remain a prerequisite for high quality text-to-speech systems, or can we expect articulatory oriented parametric systems to take over? Some kind of compromise emerges already. Much statistically relevant information can be stored in neural networks, but we should not give up the challenge of securing a profound acoustic-phonetic insight. My forecast for the future is that a more solid and integrated view of speech and language structure will develop and find its way also into speech recognition and synthesis work.

Recent engagements

During the last 14 years a substantial part of my time has been spent in studies of prosody (Fant, Nord and Kruckenberg, 1986, Fant and Kruckenberg, 1989, 1994, 1996, 1999a, 1999b, 2000A, 2000B, Fant, Kruckenberg and Liljencrants, 2000) Studies of poetry reading have also been carried out (Fant, Kruckenberg and Nord, 1991, Kruckenberg and Fant, 1993. Another main area of interest has been properties of the voice source in the frame of a speech production model (Fant, 1993, 1995, 1997, Fant and Lin, 1988 Fant, Liljencrants and Lin, 1985).

A novelty, increasing the insight in the speech chain, is to supplement traditional speech analysis displays of oscillogram spectrograms, intensity and F0 curves with respiratory data in terms of sub-and supra-glottal pressures, and a continuously scaled parameter of perceived syllable and word prominence, labelled Rs, Fant, Hertegård, Kruckenberg (1962), Fant and Kruckenberg (1996). Fant etc

The inclusion of the prominence parameter (Fant and Kruckenberg, 1989, 1994, 1999a, Fant, Kruckenberg and Liljencrants, 2000) is of special importance for relating degrees of stress and emphasis to all relevant physical parameters, including voice source properties.

A recent development within our group is studies of intonation in prose reading (Fant and Kruckenberg, 1999, 2000). We have data on F0 patterns of five subjects reading of one minute of speech, from which we have derived tentative rules for F0 predictions. There are several novelties in our approach. One lies in a system for F0 data extraction with notations from Bruce's canonical system but with conventions added for sampling of quantitative data. All F0 data are expressed in a logarithmic scale of semitones versus 100 Hz. Furthermore, individual subjects data are normalised with respect to their average F0 level which minimises male-female differences.

Predictions of local accent 1 and accent 2 F0 patterns are based on the prominence parameter Rs and relative positions within a primary or a secondary clause. Each such unit usually associated with a breath group is associated with a particular baseline for unstressed syllables characterised by F0 onset, rise, declination and a final fall. The declination drop tends to be independent of the total length of the group. As a consequence, when estimating positional consequences, the location of a syllable is expressed on a relative scale from 0 to 1. Once the local accent patterns have been derived-a special set of rules takes over for smoothing and syntactical coherence and the distributed effects of focal prominence..

The size of the F0 reset at a juncture is found be proportional to the pause duration, which in turn depends on the sequence of syntactic structures. Special rules apply to the realisation of junctures of lower probability. Rules for final lengthening and trends of quantal steps in pause duration (Fant and Kruckenberg, 1996, Fant, Kruckenberg and Nord 1991) are also incorporated.

It is our intent to test our accumulated knowledge of prosodic realisations not only in concatenated synthesis, which nowadays has gained dominance, but also in formant coded synthesis, revised to include higher level rules for articulatory continuities and prosodic modifications of segmental structures and voice source characteristics. The prominence parameter will have an important role for prosodics as well as segmentals.

The speech code. Final remark

A main point in my look ahead has been the challenge to force the speech code. This is not a matter of a single intellectual undertaking to find the key. It requires co-ordinated multilevel investigations and integrated modelling over long periods of time. The seeming lack of invariance which has discouraged so many investigators ceases to be a problem if we are able to structure the variability as a part of the code (Fant, 1985).

The complexity is great but, as time goes by, research proceeds towards increasing insights. Mankind is making much progress in mapping the genetic code. We need some of the same patience and persistence in mapping the speech code. The reward will be a more solid theoretical basis of phonetics as well as new methods in speech technology design and improved quality of performance.

References

Carlson, R. and Granström, B. 1975. A text-to-speech system based on a phonetically oriented programming language. *STL-QPSR 1/1975*, 1-4.

Carlson, R. and Granström, B. 1982. Towards an auditory spectrograph. *In* R. Carlson and Granström, B.(eds) *The repesentation of speech in the auditory system*. Elsevier Biomedical Press, Amsterdam, 109-115.

Carlson, R. and Granström, B.1991. Performance rules in a text-to-speech system. *In* J. Sundberg, L. Nord and Carlson, R. (eds) *Music Language, Speech and Brain,* MacMillan Press, 121-131.

Carlson, R., Fant, G. and Granström, B. 1975. Two-formant models, pitch, and vowel perception. *In* G. Fant and. Tatham M. (eds.).*Auditory Analysis and Perception of : Speech* London: Academic Press Inc, 55-82.

Carlson, R., Granström, B. and Fant, G. 1970. Some studies concerning perception of isolated vowels *STL-QPSR 2-3/1970*: 19-35..

Fant, G. 1953. Speech communication research. *IVA, Royal Swedish Academy of Engineering Sciences,* Stockholm 24: 331-337

Fant, G. 1957. Den akustiska fonetikens grunder. *KTH, Inst. för Telegrafi-Telefoni, Rapport nr 7, Taltransmissionslaboratoriet* (61 pages).

Fant, G. 1958. Modern instruments and methods for acoustic studies of speech, *Acta Polytechnica Scandinavica,* No 1, 1-81.

Fant, G. 1959. Acoustic analysis and synthesis of speech with applications to Swedish. *Ericsson Technics* 1 1959, 1-106.

Fant, G. 1960. *Acoustic theory of speech production*. Mouton, the Hague. 2nd ed., 1970

Fant, G. 1968. Analysis and synthesis of speech processes". *In* B. Malmberg (ed) *Manual of Phonetics,* Chapt. 8, 173-276 Amsterdam: North-Holland Publ. Co.

Fant, G. 1969. *Kompendium i talöverföring, del 1*. KTH, inst för talöverföring 1967

Fant, G. 1973. *Speech Sounds and Features*. Cambridge, MA, USA The MIT Press.

Fant, G. 1978. Vowel perception and specification. *Rivista Italiana de Acustica* II. 69-87.

Fant, G. 1985. Features - fiction and fact". *In* J. Perkell et al. (eds). *Invariance and Variability of Speech Processes, Brain and Reading* Lawrence Erlbaum Ass. Publ.

Fant, G. 1989: The speech code. *In* C. von Euler, I. Lundberg and G. Lennerstrand G. (eds.) *Brain and Reading* MacMillan, London, 171-182, 1989.

Fant, G. 1990: Människans röst och tal i *Forskning i ett föränderligt samhälle, Stiftelsen Riksbankens Jubileumsfond 1965-1990*, sid. 328-348. I engelsk version: Man's Voice and Speech, in *Swedish*

Research in a Changing Society, The Bank of Sweden Tercentenary Foundation 1965-1990, 339-360.

Fant, G., 1993 Some problems in voice source analysis, *Speech Communication* 13, 7-22

.Fant, G. 1995 The LF-model revisited. Transformations and frequency-domain analysis. *STL-QPSR* 2-3/1995, 119-156.

Fant, G. 1997 The voice source in connected speech. *Speech Communication 22:* 125-139.

Fant, G. 1998 Half a century with speech science. *STL-QPSR* 3/1998, 1-3.

Fant, G., Hertegård, S. and Kruckenberg, A. 1966. Focal accent and subglottal pressure. *TMH-QPSR* 2/1966, 29-32

Fant, G. & Kruckenberg, A. 1989: Preliminaries to the study of Swedish prose reading and reading style, *STL-QPSR* 2/1989, pp. 1-83.

Fant, G. and Kruckenberg A. 1994. Notes on stress and word accent in Swedish. *Proceedings of the International Symposium on Prosody, Sept 18 1994, Yokohama.* Also published in *STL-QPSR* 2-3/1994, 125-144.

Fant, G. and Kruckenberg A. 1996a. On the quantal nature of speech timing. *Proceedings of the International Conference on Spoken Language Processing. ICSLP* 1996, 2044-2047.

Fant, G. and Kruckenberg A. 1996b. Voice source properties of the speech code. *TMH-QPSR* 4/1966, 45-56.

Fant, G. and Kruckenberg, A. 1999a. F0-patterns in text reading. *Proc of Fonetik -99, The Swedish Phonetics Conference,* Göteborg University, 53-56.

Fant, G. and Kruckenberg, A. 1999b. Syllable and word prominence in Swedish, *Proc of Fonetik -99, The Swedish Phonetics Conference,* Göteborg University, 57-60.

Fant, G. and Kruckenberg, A. 2000. F0 analysis and prediction in Swedish prose reading.. To be published in Travaux du Cercle Linguistique de Copenhague, (Editors Nina Grϕnnum and Jϕrgen Rischel)

Fant, G., Kruckenberg, A. and Nord, L. 1991. Stress patterns and rhythm in the reading of prose and poetry with analogies to music performance. *In* J. Sundberg, L. Nord and Carlson, R. (eds) *Music Language, Speech and Brain,* MacMillan Press, 380-407.

Fant, G. and Lin, Q. 1988 Frequency domain interpretation and derivation of glottal flow parameters, *STL-QPSR* 2-3/1988,. 1-21.

Fant, G. and Lindqvist-Gauffin, J. 1968. Pressure and gas mixture effects on divers_ speech. *STL-QPSR* 1/1968: 7-17.

Fant, G. and Mártony, J. 1962. Instrumentation for parametric synthesis (OVE II synthesis strategy, and quantization of synthesis parameters. *STL-QPSR* 2/1962: 18-24.

Fant, G. & Richter, M. 1958. Some notes on the relative occurrence of letters, phonemes, and words in Swedish. Proc. of the VIIIth Intl. Congress of Linguists, Oslo 1958. 815-816.

Fant, G., Liljencrants, J. and Lin, Q. 1985: A four-parameter model of glottal flow, STL-QPSR 4/1985. 1-13**.**

Fant, G., Nord, L. and Kruckenberg, A. 1986. Individual Variations in Text Reading. A Data-Bank Pilot Study, *STL-QPSR* 4/1986 1-17 and in RUUL 17, 1987. 104-114.

Fant, G., Kruckenberg, A. and Nord, L.1991c. Stress patterns and rhythm in the reading of prose and poetry with analogies to music performance. *In* J. Sundberg, L. Nord, R. Carlson (eds) *Music, Language, Speech, and Brain*, Wenner-Gren International Symposium Series, Vol. 59, 1991. 380-407.

Fant, G., Kruckenberg, A. and Nord, L. (1992b: Prediction of syllable duration, speech rate and tempo, . *Proceedings of the International Conference on Spoken Language Processing. ICSLP* 92, Banff, Vol 1 667-670.

Fant, G., Kruckenberg, A and Liljencrants, J. 2000a. Acoustic-phonetic analysis of prominence in Swedish. Antonis Botinis (ed*) Intonation. Analysis, Modelling and Technology*,. Kluwer, Academic Publishers, pp.55-86

Fant, G., Kruckenberg, A. and Liljencrants, J. 2000b. The source-filter frame of prominence. *Phonetica*-57, pp.113-127..

Jakobson, R., Fant, G. and Halle, M. 1952. Preliminaries to speech analysis. The distinctive features and their correlates. *Acoustics Laboratory, Massachusetts Inst. of Technology*, *Technical Report* No. 13. *MIT press*, seventh edition, 1967.

Kruckenberg, A. and Fant, G. 1993 Iambic versus trochaic patterns in poetry reading, *Nordic Prosody* VI, Stockholm, 1993. 123-135.