Dept. for Speech, Music and Hearing

Quarterly Progress and Status Report

Gunnar Fant 60 years

Watanabe, A., Felicetti, S., Hedström, B., Surjadi, G., Tannergård, G., Tegerstedt, I., Wejnebring, B., Wetterling, M-B., Andersson, L., Hallsten, L., Kaunisto, M., Murray, T., Eriksson, H., Haapakorpi, M., Karlsson, I., Nord, L., Stålhammar, U., Elenius, K., Blomberg, M., Liljencrants, J., Carlson, R., Granström, B., Risberg, A., Spens, K-E., Agelförs, E., Boberg, G., Mártony, J., Tunblad, T., Öster, A-M., Galyas, K., Gauffin, J., de Serpa-Leitão, A., Askenfelt, A., Jansson, E. and Sunberg, J.

journal: STL-QPSR
volume: 20
number: 2
year: 1979
pages: 001-045
http://www.speech.kth.se/qpsr

KTH Computer Science and Communication

GUNNAR FANT 60 YEARS

On October 8th, 1979, Gunnar Fant reached the age of 60 years. To all the readers of the STL-QPSR Gunnar´s name as the author of important contributions to speech research is well known. Another, equally important contribution is the fact that Gunnar has created the Speech Transmission Laboratory, now called the Department of Speech Communication. We, who are working together with Gunnar in this laboratory, got the idea of publishing a special report "from inside the lab" in honor of Gunnar. The report you have in your hands is the result of our efforts. It contains different articles written in a more or less personal and definitely informal style by members of the lab community.

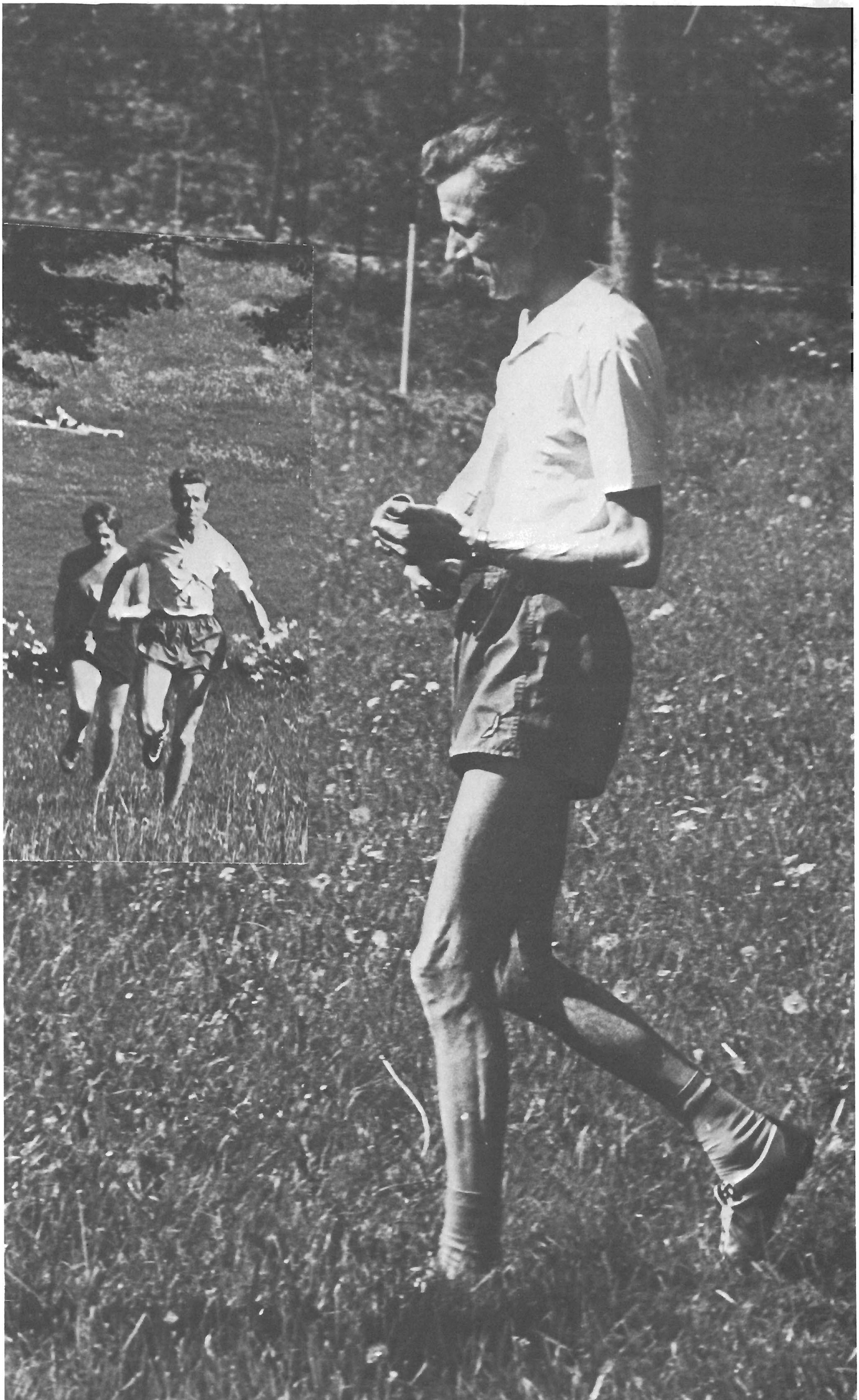
Many names of the staff members regularly appearing in the first pages in the STL-QPSR represent familiar faces to the readers, but others are names without faces. Particularly, perhaps, this is true for those of us who belong to the administration and technical staffs. They form an indispensable part of the smoothly working lab organisation. In this report, therefore, action shots of all of us have been included.

The first issue of the Speech Transmission Laboratory Quarterly Progress and Status Report was published in 1960. This is the 63rd issue. During these years more than 300 articles have been published in the STL-QPSR by staffmembers and guest researchers. Several of these articles have not been published elsewhere. In this issue we have included a cumulative index of all STL-QPSR articles. We feel that this index will be of help to the readers.

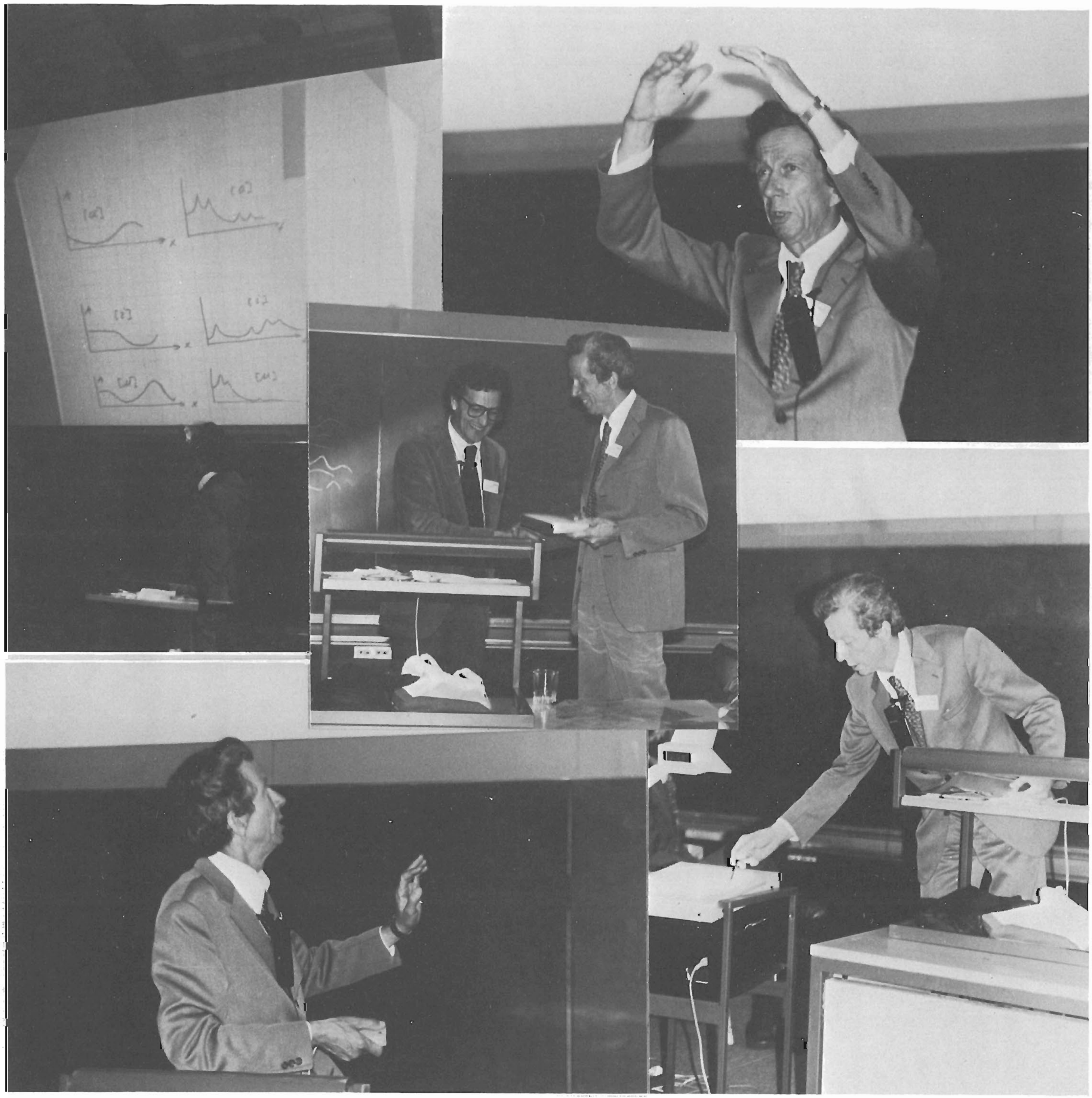
As a rule an issue of this type will contain a photo of the celebrator. We started by trying to find a photo where Gunnar in his best suit looked significant and dignified, but for some reason we failed to find a picture of this kind. Instead we found some photos of Gunnar in full activity. The first is from one of the days after the 1979 meeting in Copenhagen. It shows Gunnar in our lunchroom during an intense discussion with a scientist from Peking. In the foreground there are utensiels from the "porridge eaters". The other photos show Gunnar singing, talking, jogging, lecturing. This is how we see and like to see Gunnar.

Finally we all like to express our sincere thanks for all the years that we have had the opportunity to work together with you, Gunnar. There might be other places offering good opportunities to work but probably few that give such a good combination of stimulating ideas and good atmosphere. You have created this laboratory. We all congratulate you on your 60th anniversary and look forward to work with you in the future.

GUNNAR AS WE SEE HIM — AND LIKE TO SEE HIM



Ninth International Congess of Phonetic Sciences , Copenhagen 1979



Frontiers of Speech Communication Research
is handed over by Björn Lindblom

GUNNAR FANT'S NEW LIFE

I would like to tell you that what Professor Gunnar Fant has shown me by his whole character during my stay at the Department of Speech Communication that is a seriousness to the studies and a friendship without any discrimination and much greater than what I expected.

According to a calendar in our country, 60 years is the longest period of the circulation. So that it is said since long that 60 year old people are at a starting point towards their new lifes.

I hope that Gunnar in his new life will be healthy and add even more and new ambitious reports to his great research production.

I am quite sure that all the visiting guest researchers at the lab will join me in saying CONGRATULATIONS!

Akira Watanabe

During the last 20 years or so we have been visited by many guest researchers who have stayed for a week or perhaps for a year - and many of them returning back now and then - and we think that Akira's letter to Gunnar expresses what all our friends from so many different countries would like to say.

AHRENS, C., BRD
ANTHONY, J.K.F., Great Britain
ARNOLD-HEINZ, J., USA

BASTIN, K.C., Australia
BASU, D.K., India
van den BERG, Jw., The Netherlands
BLADON, A., Great Britain
BLESSER, B., USA
BOROVIČKOVÁ, B., Czechoslovakia

CALDERSMITH, G., Australia
CAMPANELLA, S.J., USA
CARRÉ, R., France
CARTERETTE, E.C., USA
CHIANG, H.T., USA
CHISTOVICH, L., USSR
CLARK, J.E., Australia
CLEVELAND, T.F., USA
COLEMAN, R.O., USA
COOKER, H.S., USA
COULTER, D.C., USA

DELBRIDGE, A., Australia
DERKACH, M., USSR
DOLANSKÝ, L., USA

FERRERO, F., Italy
FIBIGER, S., Denmark
FINTOFT, K., Norway
FIRTH, I., Australia
FIRTH, I.M., Great Britain
FLANAGAN, J.L., USA
FUJIMURA, O., Japan, USA
FUJISAKI, H., Japan

GALVAGNY, M-H., France
GAY, T., USA
GRONDSTRA, J.W., USA

HAGGARD, M.P., Great Britain
HALLE, M., USA
HEINZ, J., USA
HIKI, S., Japan
HIERONYMUS, J., USA
HOLLIEN, H., USA
HOLMES, J.N., Great Britain
HULTZEN, L-S., USA
HURME, P., Finland
HYDE, S.R., Great Britain

INGEMANN, F., USA
ISHIZAKA, K., Japan
ITAHASHI, S., Japan

JANHUNEN, O., Finland
JANOTA, P., Czechoslovakia
JASSEM, W., Poland

KACPROWSKI, J., Poland
KLATT, D.H., USA
KNAUS, R.G., USA
KOZHEVNIKOV, V.A., USSR
KRINGLEBOTN, M., Norway
KROKSTAD, A., Norway

LADEFOGED, P., Great Britain, USA
LEEDS, C., USA
LEHISTE, I., USA
LIBERMAN, A.M., USA
LISKA, J., Czechoslovakia
LISKER, L., USA
LUBKER, J.F., USA
LUNDH, P., Denmark

MACHONIN, V.A., USSR
MALAČ, V., Czechoslovakia
MATTINGLY, I.G., Great Britain
MEHNERT, D., DDR
METTAS, O., France
de MORI, R., Italy

OJANEN, J., Finland
ONDRAČKOVÁ, J., Czechoslovakien

PARKER, A., Great Britain
PETERSON, G.E., USA
PETTORINO, M., Italy

PFEIFFER, L.L., USA
PICKETT, J.M., USA
POLLARD, H.F., Australia

RAMISHVILI, G., USSR
RIGAULT, A., Canada
RISCHEL, J., Denmark
RISKA, T-B., Finland
ROCHETTE, C., Canada
ROSEN, S., USA
ROTHENBERG, M., USA

SAKAI, T., Japan
SAKUMA, A., Japan
SHUPLJAKOV, V., USSR
SLAWSON, W., USA
SMITH, S., Denmark
SONDHI, M.M., USA
SOROKIN, V., USSR
STEVENS, K.N., USA
SUGAR, J., Hungary

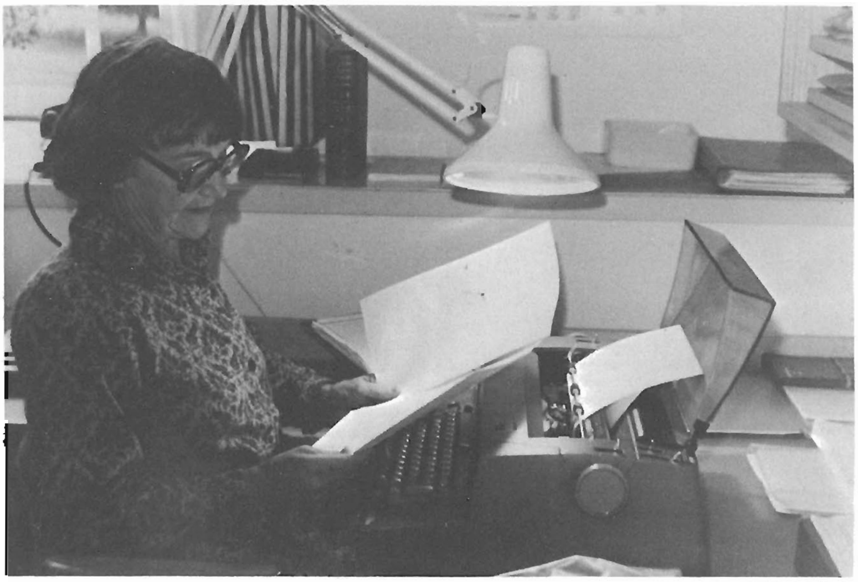
TAPPERT, C.C., USA
TARNOCZÝ, T., Hungary
TATHAM, M., Great Britain
TOMIC, T., Yugoslavia
TRUBY, H.M., USA
TSCHESCHNER, W., DDR

UHR, H., Israel

VERRILLO, V., USA

WAKITA, H., Japan, USA
WANG, W. S-Y., USA
WATANABE, A., Japan
WEBSTER, J., USA
WHANG, H.Y., Korea

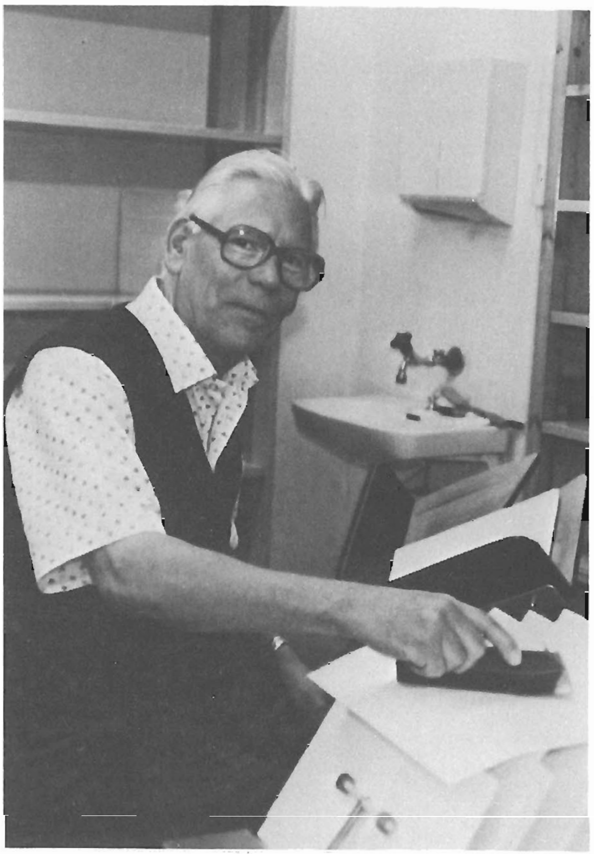
ZAGORUIKO, N.G., USSR
ZEMLIN, W.R., USA
ZLAMAL, J., Czechoslovakia



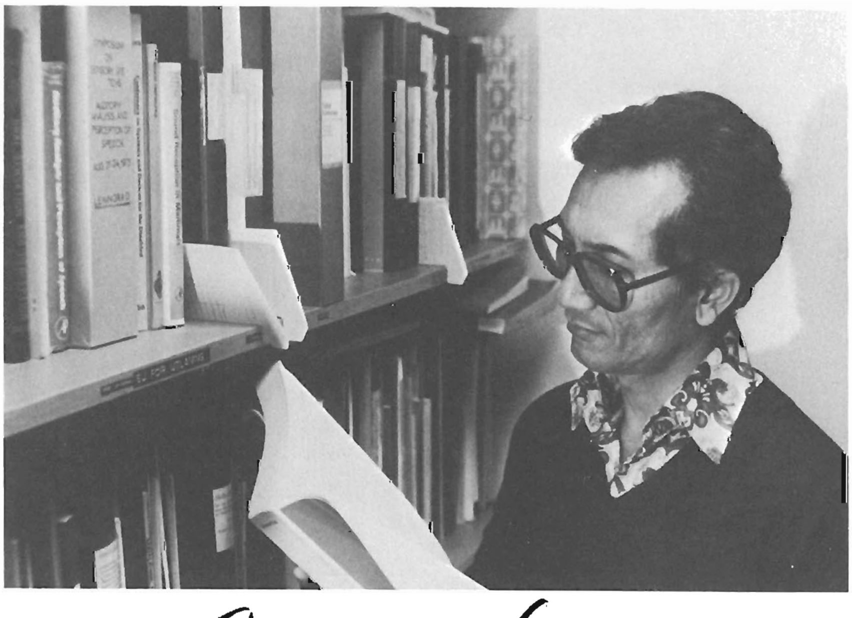
Inga

Gudrun

Ing. Dlix



Bertil Hedström



THE ADMINISTRATION GROUP

From the very beginning it was Gunnar Fant. Then it was Gunnar and Marianne Richter, who now has left us and lives in the south of Sweden. And Gunnar made research and Marianne made research too and sometimes she typed a letter. But the lab was not meant to be a two-person lab, it had to grow. Arne Risberg and Bengt Wejnebring and other people joined the lab.

In the beginning of 1960 we already were a rather big group: Si Felicetti, employed to make statistical work and measuring on sonagrams, found herself busy with typing reports and Marianne took care of the economy and other people took over some parts of the administrative work. When Gunnar got his chair Inga Tegerstedt came to help him to administer the students and the piles of paper that every week come from the Institute. Today the administration group functions as follows:

Inga takes care of the papers and the students. Si takes care of the reports, and the correspondence is shared between them. Maj-Britt Wetterling with the aid of Gudrun Tannergård handles all these millions of crowns raining over us. Some years it is only a drizzling rain but other years the situation is bright. Gudrun also makes the beautiful drawings you have seen in our reports and articles.

When our QPSR is to come out, Bengt and Si are discussing how to solve different problems, how to make the lay-out and so on. Our QPSR - our flagship - is distributed to over 800 persons and institutions in more than 40 countries. You should see the printing department when Bengt and Bertil Hedström are in the final working phase of the reports. Papers everywhere, thousands and thousands.

To start with our few books and reports were taken care of by Si. But then we had good cooperation with about 40 universities/institutions in not only Sweden but also in many other countries and every day lots of papers such as progress reports, thesis works, reprints etc. arrived. Of this reason Gunnar took an initiative to found a library to organize all those papers and this is now the biggest one at any department of KTH. We also got a librarian, Gorda Surjadi, known by all as Surr and famous for his never failing good temper.

We all constitute a group and without that group we all know that our lab cannot function efficiently.

The administrative group all the time has tried to modernize the lab. The traditional women-making-coffee-period never existed, neither are the women automatically made secretaries at the monthly sessions we have with all the lab and in the board. Everything is rolling in alphabetic order, every person making coffee during one week, every person being secretary at one session, every person being chairperson at those meetings. We never carried this through against Gunnar´s will. On the contrary. He is as enthusiastic as a young boy when we are testing new ways to make everything function to everybody´s satisfaction.

CONGRATULATIONS AT YOUR 60TH ANNIVERSARY.

Si Felicetti Bertil Hedström Gorda Surjadi

Gudrun Tannergård Inga Tegerstedt Bengt Wejnebring

Maj-Britt Wetterling

FREEDOM, EQUALITY, AND FRIENDSHIP

There must be a very good reason if a person stays at the same working place for more than 20 years. I believe the reason is that what Gunnar has offered me and all of us is freedom, equality, and friendship.

Freedom. In the 1960th when I tried to be a good mother and at the same time fulfil my duties at the laboratory I had enormous difficulties to come to my work at 8.30 in the morning. Sometimes I came at 9 or sometimes even at 10 o´clock. Very often I went to Gunnar and said: "I´m so sorry but the bus never came today" or "An accident happened so I was delayed" or "I couldn´t wake her up this morning, because she was so tired". And Gunnar always smiled and said "It doesn´t matter". But when it had happened rather many times he said to me: "But why don´t you start somewhat later in the morning? When the children are more grown up you could come earlier again". That was all, but it is very typical for Gunnar, to him the quality of our work is important, not the hours we are here.

And what concerns the equality I think that most of our guests have noticed that if Gunnar invites them for lunch in our luncheon room, he himself prepares the food (sometimes strange Swedish food like herrings or blood-pudding with jam) and there are no women running around to facilitate his efforts. Sometimes it happens that a new employee tries to be a male chauvinist but he soon stops when he realizes that we are not looked upon as servants and that we receive a total respect from Gunnar´s side.

And finally, the friendship. The friendship is the lab, is the people Gunnar has collected, people fitting well together, knowing each other well and knowing that we are a group, a working group where the results are depending on what every single person can contribute.

Thank you Gunnar. My 23 years with you have been good years.

Si Felicetti

Workshop and Computer Facilities.

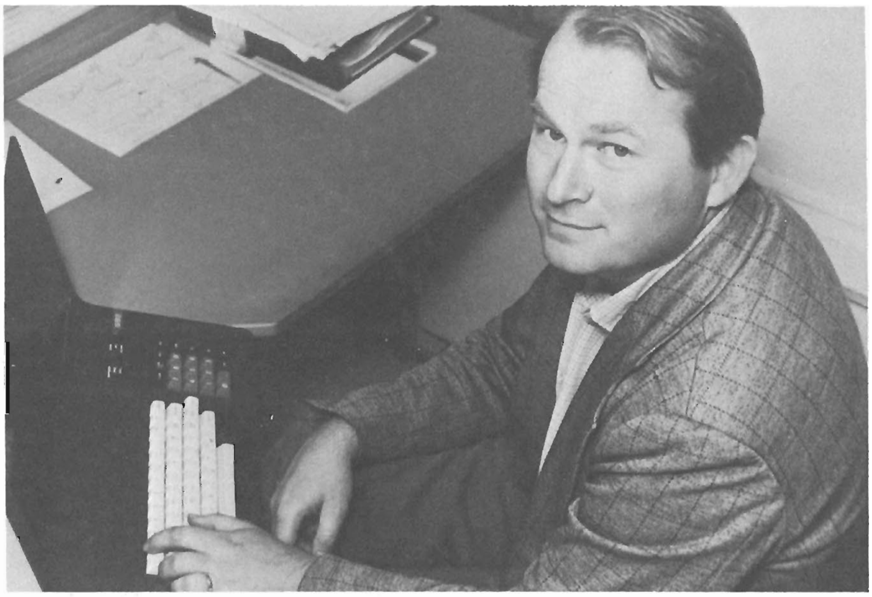
To measure is to know, but to measure you have to have measuring instrument designed for the purpose. In speech research of the early years much time had to be devoted to developing and maintaining special measuring instruments of various kinds. The workshop and development group of the Speech Transmission Laboratory of the early years was therefor a substatial part of the lab. and many different equipments for the analysis and synthesis of speech were built.

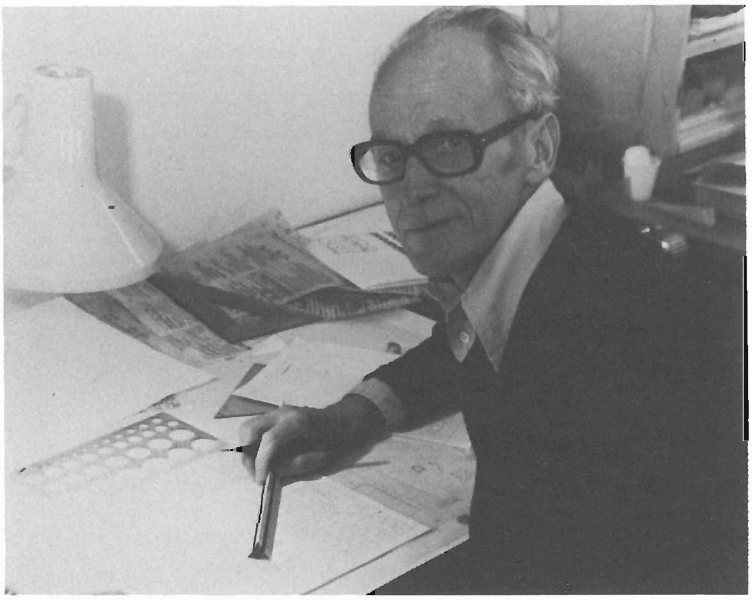
The work that earlier was made by specialised equipment is now taken over by computers. We got our first computer in 1967, a CDC/700, and we had a hard time to get it to do the things we liked it to do for us. The CDC-computer is now ready for retirement, and to replace it we have got an Eclipse that Thomas has been working on to get it shape to do all the things (and many more) that the old system could do. At the same time the minicomputers and microprocessors have made their way into the lab. Computers are now used in a synthesis by rule system for use as an aid for the blind and for non-vocal children. The microprocessor is at present mostly a sparetime interest for some of the members but we already see its applications in analysing equipment and aids for the deaf.

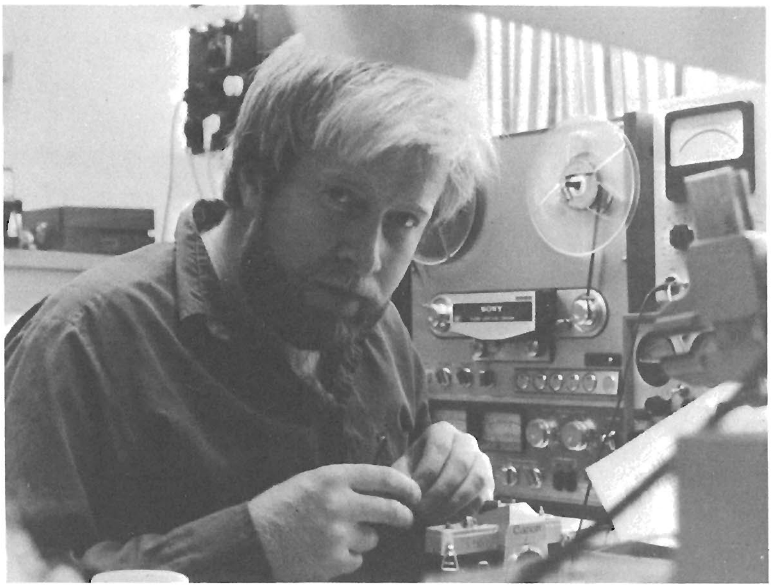
Even if a computer can do a lot of things they are costly and in many situations special purpose analog instruments are still the best. The workshop and development group is always building new instruments and reparing old, not to speak about all the time that is spent in building special interface units to our computers. The handicap groups constantly need help by the workshop. So even if the computer programming has taken over much of the development of new analysing and synthezising techniques the workshop is still an important part of the laboratory. Often we even think that we ought to extend it.

Lennart Andersson
Lennart Hallsten
Matti Kaunisto
Thomas Murray
Herman Eriksson
Markku Haapakorpi

Thomas Murray







TO GUNNAR FANT

The Speech Analysis Group

We hereby want to honour Gunnar Fant, who has just reached his 60th birthday. As we in 'the analysis group' have been working in close contact with Gunnar during the last decade allow us to make some remarks about what we have found are the distinctive features of Gunnar, a person who for many years has been and still is one of the leading scientists in the vast cross-disciplinary field of speech communication.

Gunnar's never ending enthusiasm combined with his great knowledge creates an astonishing flow of ideas. Often when you sit down to discuss your own work with Gunnar you will find yourself getting involved in some new projects as well before the talk is over especially if he is newly home from a visit to other research laboratories or a conference.

Gunnar also manages to convey his never abating curiousity and engagement in the subject and if it was not so, it would be easy to feel inferior particularly when you find that some of your own new ideas are already described in a footnote somewhere in Acoustic Theory of Speech Production (already published in 1960 by Mouton).

We have also realised that Gunnar's learning is combined with a sometimes very good memory, he will mostly be able to tell you who wrote about what in which publication and about when this was done and he is also embarrassingly good at citing your own arguments raised in a discussion some years back.

These characteristics combined with Gunnar's eminent fund-raising capacity have been the origin of what we think is one of his greater feats: the creation of the Speech Transmission Laboratory and the Department of Speech Communication, where research in almost all areas of speech is carried out and where people tend to stay as it is a stimulating and extremely friendly place to work at. We have signed outselves 'the Speech Analysis Group' even though one of the agreeable characteristics of our laboratory is the cooperation between everybody despite our diverse main interests.

Gunnar, our most sincere wish for Your future is: More time for research and less administration work. The area of speech communication is so huge and the total knowledge limited, we need Your insight and experience and we hope to be able to go on working together with You for many years to come. We wish You many more years of inspiring research.

From Inger, Lennart and Ulf who have spent the last nine years striving to learn and doing research in the area of Acoustic Phonetics under Gunnar's patient guidance.

Inger Karlsson Lennart Nord Ulf Stålhammar

Lennart Ulf Inger

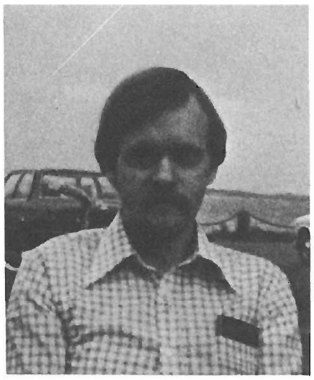
PERSPECTIVES FROM THE POINT OF VIEW OF FORMANT STRUCTURES

One might have supposed that the analysis of formant structures would have been easy and pleasant because of three factors: (1) because of the imaginative beauty therein displayed, the formant structure being regarded as the acoustic correlate of the phoneme, (2) because of the sophisticated electronic equipment at our disposal, and (3) because of the association with Gunnar Fant. Undoubtedly, the study is utterly pleasant, very much so due to the latter factor, though it is obviously not an easy one. After decades of analysis the study of formant structure poses ever new and ever more profound questions. The deeper we enter into the secret realm of formant structure the more we discover fresh vistas and mysterious problems of which we did not even suspect the existence. Every new structure to be analysed reveals features which may modify our views. Even if we multiply spectrographic samples we do not feel sure that convincing deductions can be made to reward the labour involved in the attempt. And yet - at the base of every formant structure there shines the existence of the phoneme and above it all the soul and spirit of our jubilee-celebrator. Formants have a distressing tendency to arrange themselves in a somewhat chaotic manner. Alàs, that is why I in a previous context have referred to the vowel as a spectral chameleon. It seems as if we are here encountering an apparent absence of logic which makes this particular study delicate. We can hardly deal in generalities. Can we even assert that there are any general ideas or governing principles? Is not every formant structure a specific case with its own special rules? This promotes the fundamental questions: Ought we to study the formant structure? Can any information from such a study be gained that could benefit a deeper understanding about decoding mechanisms in Humans? Could it as a spin off supply us with more efficient algorithms to be included in our Automatons? Let me be optimistic and state the following: If it were a case of impossibility it would be an extraordinary fact of which we might indeed be proud. All topics are examined.

All branches of human knowledge have their manuals and their courses. Our branch, in addition, has Gunnar! Why alone should the formant concept evade the inquiries of science? This is a challenge! We have to, by constant and systematic experimenting, introduce order into the formant chaos. We have to, as Gunnar Fant so exquisitely has done, throw the light of theoretical reason upon it. And ultimately, we have to remember that even a chameleon, however spectral, must have a decodable nature!

Gunnar - Thank You!

J.U.J. Stålhammar



RESEARCH ON SPEECH RECOGNITION

Automatic recognition of speech is a complex problem but encouraging results have been accomplished during its short history of about thirty years. A lot of the progress has been due to the rapid development of computer technology, whereas the specific use of speech knowledge has not been elaborated to the same extent. The efforts made have pointed out numerous areas where future research is essential for improving the state of the art.

In a recent review of the ARPA Speech Understanding Research Project (Lea & Shoup 1979), a majority of the interviewed researchers in the field regarded the acoustic phonetic analysis as the most important area of future work. Together with prosodics it is also where the recognition researchers consider their own qualifications insufficient and where other expertise is needed.

The major requirements for such an expert are broad experience and a profound insight into all the various aspects of speech. Speech researchers all agree upon Gunnar Fant's qualifications in this matter. His work has had a major influence on the understanding of the acoustics of speech and further development of speech recognition will benefit a lot from Gunnar's work concerning speech production and perception.

As members of the Speech Recognition Group at the laboratory we have benefited very much from Gunnar's great knowledge and experience in acoustic phonetics which makes him able to see whether a proposed strategy is worth testing. We have received a lot of good ideas concerning for example spectrum parametrisation and phoneme segmentation and classification that we have implemented in our recognition scheme. There are still more ideas that we would like to use which we have not yet transformed into recognition strategies.

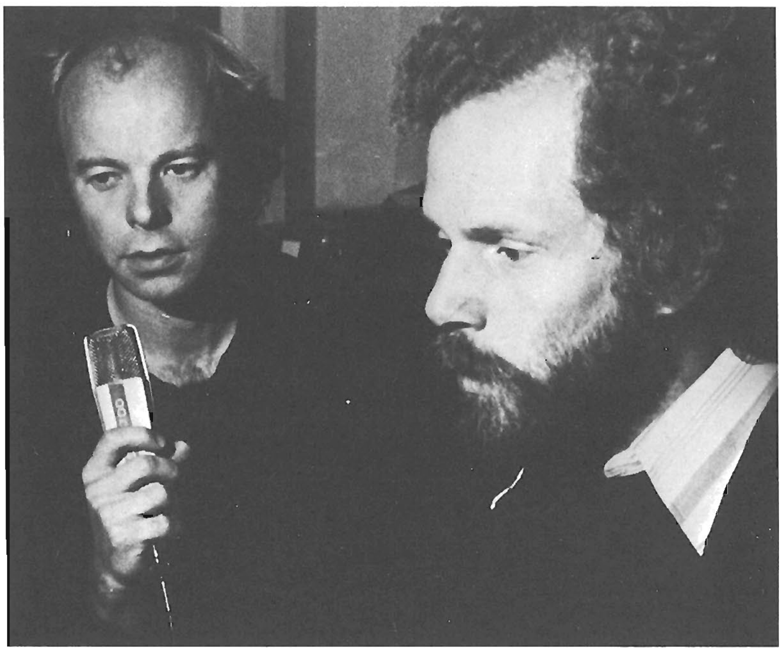
Gunnar's general opinion about important issues of speech research has had great influence on our work. He early realized the value of perceptual considerations in speech recognition.

For example, he has often stressed the importance of the dynamic properties of the speech patterns. Implementing the idea in form of time derivatives of the speech parameters across phoneme boundaries halved the number of errors in a word recognition experiment compared to static measures only. In the future we hope to include more of these perceptually oriented parameters according to Gunnar's ideas.

The open and friendly atmosphere at the laboratory is very much a result of Gunnar's personality and is a fertile ground of inspiration and exchange of ideas. We very much enjoy working in this environment and wish to continue benefiting from Gunnar's past, present, and future contributions to speech research.

W.A. Lea & J.E. Shoup: "Review of the ARPA SUR project and survey of current technology in speech understanding", Final Report, ONR, Jan. 16, 1979

Mats Blomberg Kjell Elenius



Kjell Elenius Mats Blomberg

THE CONJURER'S APPRENTICE

When I some twenty years ago started out at Gunnar's lab as a thesis student there was a big turmoil going on. At last it seemed that all the essential problems in speech analysis and synthesis had been ironed out theoretically, and to make the perfect vocoder was just a practical matter of designing the proper circuits. Well, perhaps the pitch measuring was a bit hard to do, but would certainly be a matter of a few months more work, - and then .... Of course Gunnar knew better than the student, and the latter was trapped into 18 months' work on the thesis instead of a projected three months, including also a firm employment. That is a problem with Gunnar: he has so many good ideas that you just have to check out. And moreover, he creates such an atmosphere that it is inevitable that his cooperators come up with even more. Then as time is irreversible, you have to make a selection. At the time much of the vocoder work was put outside the lab at FOA and later L M Ericsson, and Gunnar's principles led to a channel vocoder design that performed very well, and it still stands up though the competition is tough. As vocoders had proven to be practical and useful the military curtain of secrecy went down and the whole business went out of sight for a decade or so.

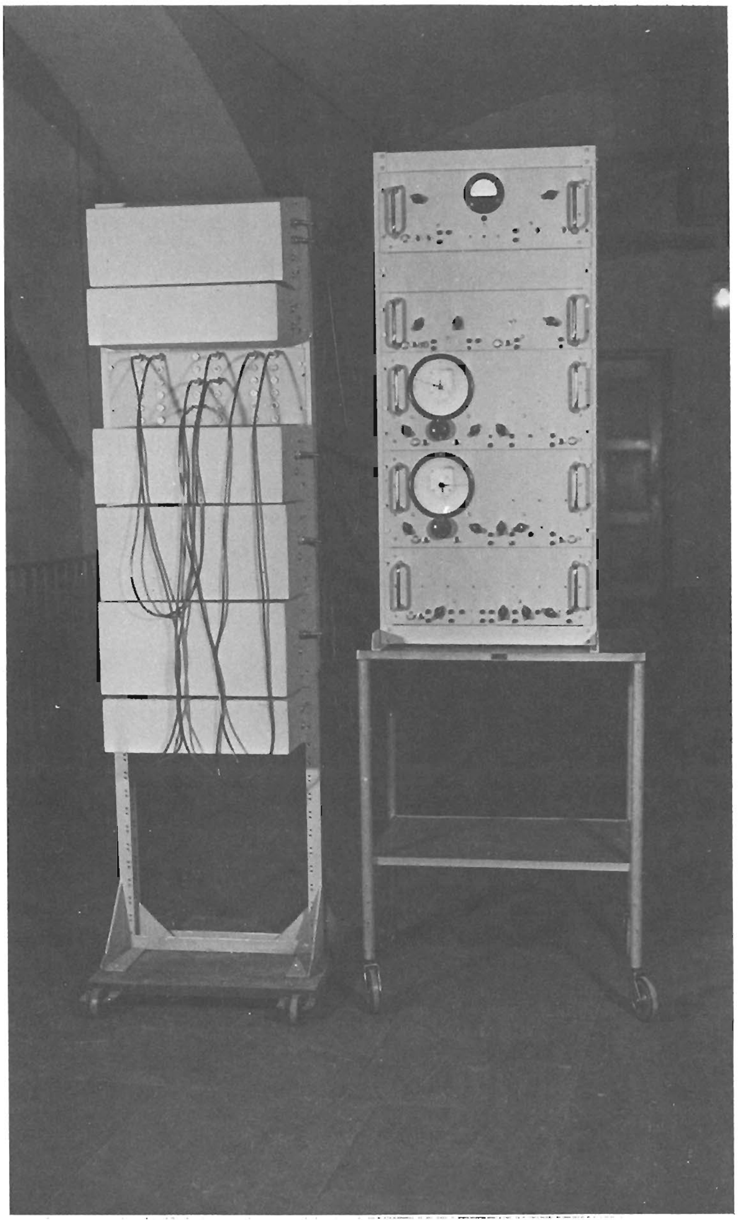
At that time I became aware of the re-discovery and re-invention cycle. It seemed very obvious from the vocoder research reports that in almost every vocoder development performance had been improved by the introduction of some trick. But comparing to older literature the same tricks had been applied by somebody else, say five or ten years ago. Things seemed to be improving fast all the time, but still Dudley's original vocoder from the late thirties sounded good by comparison. How was it possible that one could make so much progress without really getting anywhere further? Gunnar often points out the important thing: you must know what you are dealing with, you must have a profound knowledge of the real world, not only of the simplified models. The vocoders were but only a little sideline among Gunnar's interests. Gunnar's exhaustive work on the theory of speech production had some years earlier materialized in the manually controlled OVE I synthesizer, a couple of heavy electronics boxes and a big drawing-board with a mechanism to control pitch and formants. I remember hearing that Gunnar had taken the apparatus to an ASA-meeting in USA but had trouble in the customs. Finally he got through with his instruments classified as conjurer's equipment...

The synthesizer work had been successfully pursued, and when I met Gunnar the OVE II synthesizer had just been completed.

It could produce synthetic speech of an amazingly good quality. With one important proviso, you had to spend a month or so to produce 2 seconds of good sounding speech, a ratio of several 100,000 times real time. It seemed necessary to perfect the analysis instrumentation in order to get better data than could be obtained with the fragile Sona-Graph of yesterday.

In the sixties comparatively good opportunities were to raise funds for instrumental aids to basic research. A group was formed in the lab that got heavily involved in the design of an advanced filter bank spectrometer. We started out with Rasslan, a feasibility prototype with six parallel bandpass filters and a stepwise controllable frequency translation system. With it we could get reliable spectrum sections on a Mingograph in just a few minutes and with a wide range of analysis parameter settings. Rasslan paved the way for the second main project in the early sixties: the 51-channel filter bank. This was made to include the novel digital technique of recording the output data. We then started thinking about computers, a natural consequence of Gunnar's and his colleagues' many international inspirations. While the computer funding paperwork ground its way the OVE III speech synthesizer was developed as a match to the filter bank, and in 1967 we inaugurated the new toy as the second institution at KTH with an in-house dedicated computer.

Then came the hard years when all the work had to be done over again with a new technique. We got the things working and it became practical to start work groups for synthesis by rules and for speech recognition. To the latter the filterbank was quite useful because it works in real time. Also we got hold of the new sharper tools of FFT spectrum computation and digital filtering



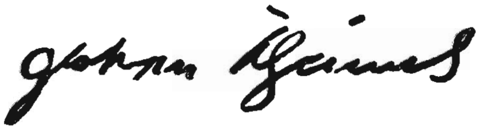
Heterodyne filter constructed by Gunnar Fant 1950.

using the z-transform. One of the selections made at this time was to stay with the basic OVE II approach using cascade synthesis rather than parallel because of its inherent superiority in modelling the vocal tract.

In the seventies fundamental research has had a rather tough position for its funding, there has been a lot of emphasis on that practical applications have to come out. Gunnar has always been very persistent that you must keep fundamental research behind the practical things, even if it will have to be done in the back yard at times. Today we are again going through the ordeals of putting a new computer to work - all over again, now loaded with LPC analysis. The perennial question rises: is it a good model of reality, or is it just a good model?

The state of the art improves as always, but many basic problems still remain. Really, why should it be so difficult to get the naturalness in synthetic speech, especially with female voices. I think Gunnar's approach is significant: start again from the beginning, let us get a better voice source. The first simulations of it sound encouraging...

Johan Liljencrants



SOME NOTES ABOUT GUNNAR, OVE, AND SPEECH SYNTHESIS

In 1969 we joined the friendship called Department of Speech Communication. Gunnar, you told us about research, hard work, and moments with the feeling of success. You also told us a little about the possibilities to go to conferences and to meet the people in the world-wide community called speech research. Since we at that time had an interest to go to Alaska we mentioned that shyly. "Speech conferences in Alaska are unfortunately very rare", you added with a smile.

We never went to Alaska but we found something else far more important. We found a group of kind and stimulating people that you had gathered around you, which, for us, is one of your most important contributions. We found people working on different subjects and with different backgrounds heading for the same goal - to understand something about the communication process.

It took a long time before we realized your position in the scientific world. Gunnar was Gunnar. The only irritating thing was that you always predicted the true outcome of the experiments before we made them or you had already written about it in "Acoustic Theory of Speech Production".

In this synthesis part of the report in honor of your 60th birthday, we want to make a small review of the past and present work on speech synthesis at KTH. We will also touch a little on our plans for the future.

Speech synthesis has always played an important role in the work at the lab, partly because of your important contribution in this area and partly because of the way this method explicitly tells us about our present knowledge of speech production and speech perception. Although we will not discuss it here, we would like to emphasize the important work on speech analysis that always is behind the quality of speech synthesis and is an integrated part of the speech synthesis project.

Let us first go down to the historic cellar where the old OVE:s are standing with painted sheets, tubes, and coils.

Orator Verbis Electricis

"I love you". A number of resonances coupled in cascade and tuned in a proper way by Gunnar's gentle hand in the beginning of the fifties (see figure 1). It all started when Gunnar worked at L M Ericsson (1945-1949) and made analyses of Swedish vowels and consonants. OVE I, presented in 1953, was not the first cascade type of synthesizer but was the first to include the important "higher pole correction", which until then had been neglected. Gunnar had already been engaged in building synthesis models at MIT (1949-1951) together with K.N. Stevens and had also worked with the Pattern Play-Back synthesizer at Haskins Laboratories. Analysis, synthesis, and description of speech sounds are main subjects of interest both in Gunnar's scientific work and in the history of the Department of Speech Communication. A thorough description of the formant concept was presented in 1956, and in the classical publications "Acoustic Theory of Speech Production" and "Acoustic analysis and synthesis of speech with applications to Swedish", the general theory was set.

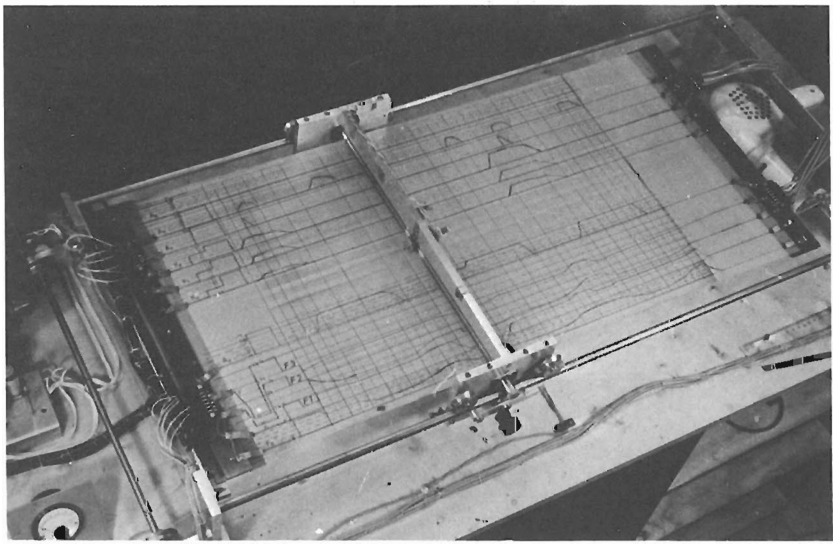
In parallel with the construction of PAT by Walter Lawrence in Edinburgh, the cascade machine OVE II was developed. The uniqueness of OVE II was not only the ability to synthesize all kinds of speech sounds but also the means of controlling it. A system of conducting lines on a sheet was scanned by a rolling potentiometer, a technique developed together with Arne Risberg. At the Speech Communication Seminar in Stockholm in 1962 (see figure 2) speech was presented by Fant, Risberg, Rengman, and Mártony of a high quality that seldom is reached by present-day synthesis systems.

Modern techniques with integrated circuits were the base for OVE III, designed by Johan Liljencrants. The synthesizer was controlled by a computer program developed by Johan Liljencrants utilizing the ideas of smoothed step functions in an elegant fashion.

At an early stage the importance of the glottal source was realized, and detailed work by John Holmes and Jan Lindqvist-Gauffin guided us to a new model that increased the naturalness considerably.

Gunnar Fant and OVE I

Janos Martony , Gunnar Fant and OVE II



In the end of this review of synthesis work at KTH, we conclude that the main synthesis technique was developed by Gunnar, Walter Lawrence, Franklin Cooper, John Holmes ... already in the fifties. What has happened after that?

Synthesis has changed from being a mixture of detailed analysis, testing of basic models of the vocal tract and artistic skill to more abstract models of the whole speech communication process.

The synthesis by rule technique of today tries to describe speech production as production rules or networks, and begins with ordinary texts or even concepts as input. Much has been added to model linguistic behavior, e.g. work on prosody by Björn Lindblom. However, on the segmental level of these models, we have to admit that the knowledge or the art of the pioneering work has been too much forgotten. The increase in quality has not been as rapid as was hoped at that time. Much could be explained by the fact that the art was left behind, and was replaced by generality. But something was forgotten in this transformation. We have to turn back to the basic principles of speech production and once again achieve this knowledge. Gunnar has continued to remind us of that since we started on this project in 1969. It is apparent that the secrets of the basic units are still uncovered after 25 years in speech synthesis research.

Speech synthesis by rule; present and future

At present, work on speech synthesis is under way at different levels. The methods are the same as they always have been: analysis, models, and perceptual evaluations. A new area, "applications", has also become highly relevant. The rule synthesis system at KTH is now a complete text-to-speech system that accepts ordinary spelt text as input and produces speech as output. The system is described by a sequence of rules utilizing the feature concept. These rules transform the text to a phonetic transcription, to a prosodic frame, and to a parameter representation that controls a modern version of OVE III.

Let us leave the details and discuss our personal view on the system as well as our goals for the future.

We think that speech synthesis work should be more clearly divided between applications and basic speech research. The present system is good enough to be used in certain applications but the technical solutions have to be improved. As we will discuss later, our system has already been subject to a long-term evaluation as a speech prothesis but the needs that guide the work are quite different from the need to change the present general synthesis strategies. We think that our philosophy behind the system is well suited for such a division between research and applications. Since the whole transformation from text to speech is described by rules, these could be improved in advanced research systems and then transferred to the user's system. Consequently, we should develop a small and cheap system that accepts rules, but is built without considerations of the content of these rules. The main effort should be spent to make the system well suited for different kinds of users, and flexible enough to accept input of different sorts, e.g. from optical readers and databases.

We think it is time to realize fairly good quality synthesis systems that could be used in several applications on a large scale. On the other hand, we have to go back to the basic issues of speech production and perception, the main themes in Gunnar's work, to be able to make major improvements in the quality.

Speech synthesis in use

Since we work in a technical university and have received much of our support from the Swedish Board for Technical Development (STU), one important goal for speech synthesis research has always been to design practical systems that could be used in different kinds of applications. Speech communication is the most basic means of communication between humans and ought to be used even in the increasing sphere of man-machine communication. Very often, visual communication, like text or light signals, could only be used in a limited way. The reasons for this limitation could be either technical problems such as distance from the source either the user, or a physical handicap such as blindness. General speech synthesis is clearly not the way to go in all cases. Several solutions to generate speech could be chosen depending on the size of the vocabulary or the quality.

We will briefly mention two experiments that are both related to the use of speech synthesis as an aid for the handicapped.

In the spring of 1976, thesis work was carried out with the goal of studying speech synthesis in use by the visually handicapped. We could conclude that the synthetic ideolect was learnt after a short period of listening. This knowledge was retained for a long time. Furthermore, we received very positive reactions that encouraged us to continue to develop a transportable prototype that could be used outside the laboratory.

In the spring of 1978, a prototype was running. The first one to use the system was Mikael, a 14-year-old CP- boy in a school for physically disabled in Gothenburg.* He was motivated to work with the equipment. The possibility to "speak" for the first time was a very stimulating experience.

Enthusiastic reactions were received also from his teacher in the Swedish language. When using this new way of communication, she could establish a much better contact with her student than before. She also noted that Mikael's language use was improved after a short time. He was very eager to explore new ways of expressing himself and even fooled the system with "misspellings" when some words were mispronounced.

Since this time, the speech synthesis has been used by more children in the same school with essentially positive reactions. One interesting field of research is how to give the children that are regarded as non-verbal the access to spoken language. Quite a few in this category are now using the Bliss symbol system for

* The work in Gothenburg is a joint effort by Rolf Carlson, Karoly Galyas, Björn Granström, Björn Larsson, Johan Liljencrants

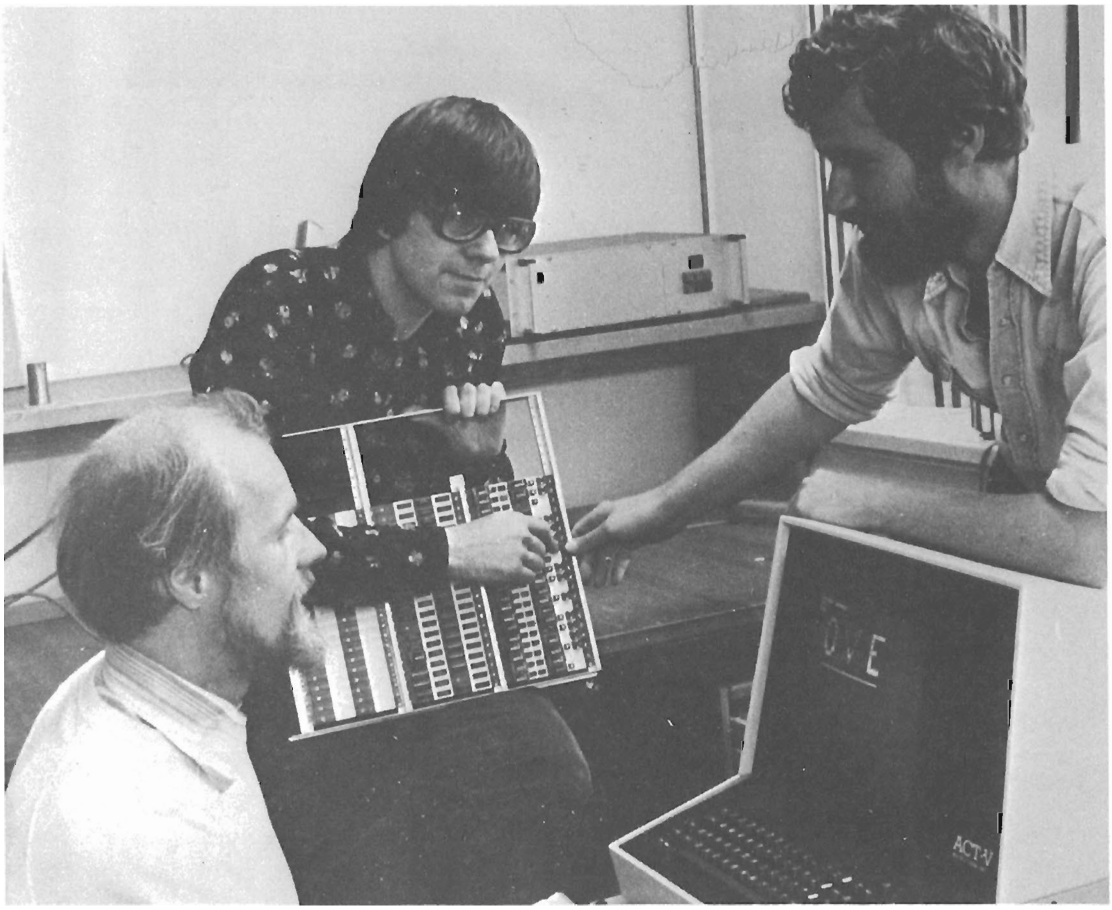
communication and translation programs to synthetic speech are being developed.

Speech synthesis is hence no longer just a research tool, and besides applications for the blind and non-vocal people, we will soon hear synthetic speech of improved quality in our daily life. It is of great importance that this development does not dehumanize our environment, rather that it helps us to gain access to new services and information.

We wish you a happy 60th birthday and thank you for our many warm and rewarding experiences during our time at the Department of Speech Communication.

Rolf Carlson Björn Granström

The Synthesis Prototype System 1978



Björn

Rolf

The working group on aids for the deaf

Activities around aids for the handicapped continued to be mainly a matter of thesis studies, either at the Royal Institute of Technology or at the Teacher Training College for teachers of the deaf up to 1964, when a separate working group was established at the laboratory specialized in these problems. The group was from the beginning small and constituted mainly by Arne Risberg and one engineer. The problems of this group were, however, fascinating and rewarding and gradually several from the other groups also started to work on projects around the deaf. One of them was János Mártony, who attacked the problems in a very thorough way by starting his career in the group by going through a training as a teacher of the deaf. The direct contact with the school that his training gave us has been very important. János specialized in the speech of the deaf and in speech training.

Later Karl-Erik Spens joined the group. After a period of work with frequency transposition and the development of visual displays for speech training he gradually specialized in tactile speech transmission. In his article he is describing how he today looks at the problem in this interesting but difficult field.

Many others have also worked in the group for longer or shorter periods. Two who have stayed are Eva Agelfors and Göran Boberg. Both have a background in the field of the auditory handicapped. Eva from a training as an auditory measurement technician and Göran as engineer in a factory for manufacturing of group amplifiers and other aids for the deaf.

Our work in the group has gone through three phases:

A. The instrumental phase.

B. The learning phase.

C. The integrating phase.

During the instrumental phase the main interest was in trying to adopt the different analyzing and presentation techniques developed in speech research in aids for the deaf. This phase ended around 1967 and resulted in a series of visual aids for speech correction

Our contacts with the teachers and the children in the schools for the deaf have made us aware of that we had very little knowledge about the real problem that a deaf child encounters and what it really meant to be a teacher for these children. We also realized that the problem of deafness had been treated in a too simplified way by many investigators in the past. This resulted in that we as a group entered into a learning phase, where we tried to get a better understanding of the theoretical problem both in speech production and in speech perception. In speech production the result was a separate project around the speech of the deaf financed by the Bank of Sweden Tercentenary Foundation and the Swedish Board of Education. In speech perception we started a series of measurements of severely and profoundly deaf children's ability to perceive speech sounds, the prosodic features of speech, discrimination of different simple signals, studies of lip-reading etc. In the area of tactile speech perception this phase of our work was characterized by studies where we tried to find an optimum way to transmit signals through the sense of touch.

During this learning phase our contact with the rest of the work at the lab sometimes was rather limited and we think that many, including Gunnar, wondered what we really were doing. And probably also the Swedish Board for Technical Development that financed the main part sometimes wondered how relevant our studies were for technical development. They, both Gunnar and the Board for Technical Development, have however during these years always given us full support for which we are very thankful.

Gradually all the work in the group has shifted over to a phase that we can call the integration phase. It is apparent that we still have a very limited knowledge about deafness and the effect of deafness on speech and language learning, on speech perception etc but we have now entered into a phase where we would like to concentrate more of our efforts on the application of what we have learnt both in the development of new technical aids and in the development of new teaching and diagnostic methods. Our present view on the global problems of education of the deaf is described by János Mártony in his article.

We hope that this phase of our work will result in that we get a closer contact with the work in the other groups at the laboratory. The problem of deafness is a multidisciplinary problem and can only give results in research groups that have close contacts with researchers from many other fields, such as linguistics, phonetics, psychology, medicine etc and also have a close contact with the schools for the deaf. Technology cannot solve all of the problems in the education of the deaf but technological development has given substantial help to many severely hard of hearing persons and it is our belief that modern speech research together with modern technology also can give help to many of the profoundly deaf. The Speech Transmission Laboratory established by Gunnar is one of the places in the world where a multidisciplinary research team has been established and a close contact has also been established with technological development. We look forward to the work together with Gunnar during this new phase in our work.

Arne Risberg Karl-Erik Spens Eva Agelfors Göran Boberg

Karl Erik Spens

Arne Risberg

Eva Agelfors

Göran Boberg

AIDS FOR THE HEARING IMPAIRED

Gunnar and I

My first contact with Gunnar was in 1955. I was a student at the Royal Institute of Technology and was looking for a theme for a thesis study. A friend of mine, Björn Arfwedsson, took me down to the cellar region of the institute. There we found a tall man in a dark room working together with a woman with a copying machine. The man was Gunnar Fant and the woman Marianne Richter. These two then constituted the Speech Transmission Laboratory.

When Gunnar heard that we were interested in a thesis study he started to explain the principle of the channel vocoder and gave us some reprints, asked us to read them and then come back. This contact resulted in that I started to develop a channel vocoder and my friend to develop a device for visual and tactile presentation of speech.

These two projects illustrate Gunnar's range of interest, a range of interest that since the beginning has characterized the Speech Transmission Laboratory and later the Department of Speech Communication. Gunnar is primarily interested in basic research but likes to see the results of his research applied in different areas. This interest has resulted in many contacts with people from the linguistic, medical, and educational fields. Many of these contacts have had a profound effect on the work in these areas in Sweden. As an example the work on auditory training of deaf children by Erik Wedenberg (1954) can be mentioned.

Björn Arfwedsson's work resulted in the first prototype of the visual spectrum indicator for speech training with deaf children. The tactile display resulted in the first tactile device developed at the laboratory. My thesis work resulted in the first channel vocoder built in Sweden and also in that I became a member of the staff of the Speech Transmission Laboratory. Twice I have left the laboratory to see other parts of life, but after a couple of years I have returned. The research area defined by Gunnar's interests and the group of people that Gunnar has collected around himself seems to be one of these waterholes that you always return to.

Ref. Wedenberg, E. (1954): "Auditory training of Severely hard of Hearing Preschool Children". Acta Oto-Laryng Suppl. 110.

Arne Risberg

PERSONAL EXPERIENCES AND THOUGHTS ABOUT WORKING WITH TACTILE AIDS AT GUNNAR´S LABORATORY

In this paper feelings are more reliable than are facts.

Gunnar covers an extraordinary broad field as a speech researcher. A good example of this is that he already during the fifties initiated research on visual and tactile communication of speech signals to deaf and profoundly hard of hearing persons. At the Speech Transmission Laboratory (the name before it became a department) a 16-channel filterbank which could control 16 neon discharge tubes or ten vibrators was built under Gunnar´s guidance. The light- or vibratory patterns generated by the speech could in this way be interpreted either visually or tactually by the fingers.

During a visit 1962 at the lab while I was a student just these speech-controlled lightening patterns generated by the "Lucia" to me personally became the first strong and colorful evidence that there existed a "speechtransmission laboratory" and that people there worked with such humanitary things as aids for the handicapped.

I did not know then that some years later I would be a member of the staff and that I would be busy to make a production prototype of a "Lucia".

Positive results from experiments with these tactual and visual aids for the deaf were presented by Lövgren & Nykvist in 1959 and by Pickett in 1963. It was shown that information extracted from the acoustic speech wave and displayed tactually could support and complement lipreading, the only channel that otherwise was open to totally deaf people. There was a lot of optimism and it was hoped that via aids for training of speech production and speech perception together with an improved knowledge of speech and sophisticated technology it would be possible to reach the ultimate goal within a not too distant

Spectral speech information presented visually and tactually, 1961.

future. This ultimate goal was a tactile hearing prothesis which, if not alone, at least in combination with lipreading should be capable to convey most of the information in the acoustic speech signal. It was a wonderful and inspiring time in the handicap group at Gunnar's lab and a long sequence of more or less new aids for the deaf and hard of hearing were built. Most of them were speech correction aids like the s-indicator, the pitch indicator, the nasal indicator, a teddybear with speech-controlled glowing eyes, and a speech-controlled toy car etc. and also an aid for speech perception, a two-vibrator system conveying the voiced-voiceless features. All these aids are today commercially available.

Towards the end of the sixties development of aids for the deaf became more difficult for several reasons. (1) The faith in electronics aids was pushed aside by a growing interest in sign language. (2) Aids for tactile speech did not seem to improve speech perception significantly on the sentence level. (3) The theory of Liberman et al about speech being a code which only could be interpreted by an intact auditory system did not support tactile speech displays.

Consequently the research on tactile aids for the deaf was slowed down all over the world. At our laboratory some comparative experiments were performed with visual and tactile information of voiced-voiceless and voiced-voiceless-stop features in combination with lipreading. The experiments showed that the tactile mode of presentation was superior to the visual mode for patterns of sentence length and also that a lot of training was needed before one really could benefit from the aid. Meanwhile, promising reports about the Optacon, a tacile reading aid for the blind, were published which shed light on the relatively high information receiving capacity of the skin (brain) if an optimal display mode was chosen. Also Kirman's (1973) review on tactile communication of speech was very encouraging. Very convincingly he concluded that it was at least not disproved that speech information could be conveyed tactually, and that one of the most important variables was the choice of display mode. The commercially available Optacon made it possible for me to make some experiments with tactile speech displays with extended time windows which indicated that the long time window resulted in a higher accuracy.

My own opinion about realizing a tactile ear prothesis in the future is that it will be possible to make a tactile aid which at least will support and complement lipreading a lot. However, it will take a considerable amount of training to learn to benefit to one hundred per cent from the information presented on a complex vibratory or electrotactile display. Therefore, maybe the most important thing at present is to concentrate on the developing of a simple tactile aid which is as acceptable to the deaf user as a conventional hearing aid. The exposure time to the aid could then be extended to be a part of real life and not only visits to a laboratory. That would also enable us to collect data from long term training periods with tactile speech reading aids of which there are very few today.

My stay at Gunnar´s lab has not only given me a very stimulating place to work at but also the possibility to meet almost all colleagues in the world in this field and that has given me good reasons to feel optimistic about the future development of a tactile prothesis. This is also a very good reason to feel very pleased and satisfied with life.

Thank you very much, Gunnar!

Ref.: Kirman, Jacob H.: "Tactile Communication of Speech: A Review and an Analysis". Psychological Bulletin 1973, Vol. 80, No.1, 54-74.

Karl-Erik Spens

SPEECH TRAINING AND TRAINING AIDS FOR THE DEAF

Some personal ideas

Characteristic for our place of work - and for Gunnar´s way of working - is the fact that we come from many different parts of the field, and that we struggle to find application for our research work. The problem of the hearing impaired and deaf has interested Gunnar for a long time. This interest has been shared by many co-workers here, and also by me when in 1964 I left the synthesis work to become a member of Arne Risberg´s group in order to work with speech production and perception difficulties of the hard of hearing and the deaf.

Children with prelingual deafness or severe hearing loss have extreme difficulties in learning to speak. To make the speech learning process somewhat easier, technical aids could be used. These aids should in some way compensate the hearing loss and give the speaker a visual or tactual information of other people´s speech and a feed-back of his own speech. The idea of using visual aids is not new; already A.G. Bell had some ideas but he was not very successful. The first more successful indicator was a voice pitch indicator constructed by Coyne in 1938 and used in some South-African schools for the deaf. In the middle of the forties Lorentzen and Nielsen built their "s"- and intensity indicators which still are in use in some schools. Many different constructions followed after the first pioneers´ work, for instance the first prototype of Lucia, a spectrum indicator built under the guidance of Gunnar Fant in the middle of the fifties.

Today there is a great variety of different, technically sophisticated visual and tactual speech training aids, but many of them are not integrated in the teaching programs for verbal language in the schools for the deaf.

During the last five years our work (János's, Tekla's and Anne-Marie's) has concentrated on the characteristical speech errors of deaf speakers. The initial aim of our work was to learn more about the speech of deaf children, and why they have these typical speech habits and to use this knowledge to develop better speech training aids and suitable pedagogical methods.

In this research work (sponsored by the Bank of Sweden Tercentenary Foundation and the Swedish Board of Education) we have learnt a lot about speech production mechanisms in deaf speakers, and we have proposals for speech training programs. But we think that we should look even further, raise questions and get answers on language, language development, bilingualism etc.

Not until we have raised these questions and got their answers, can we explain the failure of many speech training aids, and find ways eventually to integrate speech training aids and speech work in general in the schools for the deaf.

Before articulating these questions, in my opinion, the following statement is essential: Deaf and hard of hearing persons should be bilingual. They must have full competence in both lanuages, in the genuine sign language and in a verbal language. The manual language is not only the natural language for a deaf person but usually the first language learnt at early childhood for obvious reasons which need not be discussed more. Deaf people have to master a verbal language too, or else they cannot receive knowledge and information from the hearing society in which they live. Speech is only one possible carrier of the linguistic information of the verbal language, but it is an important part in language acquisition. Full competence of a verbal language for the deaf cannot (as for normally hearing persons) include normal speech perception and production, but it includes some of it and full competence in reading and writing.

Speech teaching and training come in a third place in language learning, partly chronologically because the manual language is the first learnt, and because language knowledge is a presumption for speech.

Some of these questions that must be answered are:

(A) What is the influence of sign language on verbal language?

Two more or less simultaneously learnt languages can have positive or negative influence on each other. Studying this influence it must be taken into account that most people (parents and teachers) use signed Swedish, i.e. a manual form of a verbal language.

Contrastive linguistic studies are important for improving the teaching of grammatically correct verbal language.

(B) At what age should speech acquisition and training start?

If the first language of the deaf child is a manual language, it may happen that the learning of perception and production of the verbal language starts too late so that sensory deprivation makes learning and using of speech much more difficult. To avoid this, research in the following areas is necessary: sensory deprivation, the optimal period of learning, necessary proficiency level in speech production and perception at an early age, the increase in learning difficulties if speech acquisition does not start at an optimal age. Research in these topics should result in a model on language acquisition (for manual and verbal language) in prelingually deaf persons. On the basis of this kind of model, it would be possible to plan longtime curriculums in language and speech work.

(C) What types of strategies are optimal for speech learning and speech training?

Some methods start on the segmental level (synthetic approach) to build up speech, phoneme by phoneme;

other methods start at a general pattern of a phrase or a word (analytic approach). In the first case, prosody, coarticulation, i.e. the dynamic aspects, are disregarded in the initial stage; in the second method speech sound quality is neglected. In both cases the deaf child learns and gets used to erroneous pronunciations which later on and with much work has to be changed, corrected.

Speech teaching programs should be constructed in a careful balance between the two methods based on a basic knowledge of speech and language.

(a) What features of speech are most important for transmitting linguistic information? These important features should get a priority in speech training to improve the intelligibility.

(b) How can these features be transmitted in an auditive, visual or tactual way to a prelingually, severely hard of hearing or deaf young child?

A model of speech perception (based on information theory) is a presumption for pedagogical methods for training and also a presumption for constructing technical aids (auditory, visual and tactual) to increase speech perception ability. Construction of useful technical aids can only be possible if a basic knowledge of speech acoustics and speech analysis is coupled with pedagogical needs.

(c) How should speech production training be coordinated with speech perception and language training, and last, but not least, how can young children be motivated for oral communication?

<u>(D) What are the basic causes for speech errors</u>?

Expecting "perfect" speech from prelingually deaf persons is somewhat Utopian, but we can expect better speech with less amount of work. Therefore we have to find the correct diagnosis of the speech errors. In our research we have worked with these problems and we have some answers, but more work is needed.

According to us speech errors can be caused by:
(1) Insufficient knowledge of the linguistic and phonological rules of the language.
(2) Insufficient knowledge of speech production.
(3) No internal control of its own production.

In diagnosing the deaf speaker's speech not only the errors but also underlying difficulties have to be indicated. Speech training aids can be used to establish the internal feedback, i.e. they can be used in correction of errors of type (3). Indicators which give good and stable pedagogical results are those which make it possible to build up an internal feedback.

To avoid or correct errors we need a set of different speech training aids, suitable for different ages, hearing losses and speech proficiency levels.

(E) Which errors have the greatest impact on the naturalness and intelligibility of speech?

A hierarchical order of errors should help pedagogues in their correction work. Such an order should probably be the same as the orders of important features in speech (C.a.).

The problems which have to be solved are of variable types and cover different sciences as e.g. linguistics, phonetics, pedagogics, psychology, speech research, technology... At the Dept. of Speech Communication, under the guidance of Gunnar, we have had many opportunities to do interdisciplinary research, which was of great importance, and also most stimulating.

To solve the practical technical-phonetical problems, as for speech teaching of the deaf and as for developing technical aids, it is most essential that there are possibilities for applied as well as for basic research. Gunnar - you gave us these possibilities.

János Mártony Tekla Tunblad Anne-Marie Öster

János Márton

Tekla Tunblad

Anne-Marie Öster

WHAT CAN SPEECH RESEARCH AND TECHNOLOGY OFFER TO THE SPEECH IMPAIRED?

When Gunnar for the first time in 1971 talked to me about aids for the speech impaired, he had a vision of several possible technical solutions which could improve communication ability of handicapped persons. Some of the ideas gave rise to research and development projects, others have been dropped after a critical examination and there are still others awaiting further studies. Our work is conducted by a firm conviction that speech research and technology can offer a lot of advantages to the speech impaired. The question in the head-line is difficult to answer properly without a detailed knowledge of what the needs are and what kinds of solutions are acceptable for the handicapped. I only have fragments of that knowledge and therefore I'll not try to cover the whole topic. I'll rather summarize my general view on this field of research and development and describe our activities and my personal belief on future development.

The spoken word constitutes the primary mode of communication between human beings. Those who cannot make themselves understood adequately are also handicapped in their ability to communicate personal attitudes and to establish satisfactory social relations. When speech therapy cannot provide a satisfactory speech for every communicational need, technical aids might help the individual in expressing personal thoughts and feelings.

Speech-impaired persons form an extremely heterogeneous group of handicapped individuals. Very little is known about what aids the different categories of handicapped need. Certain types of impairment require some support to the person's own speech, whilst non-vocal individuals might need aids which generate artificial speech.

Speech therapy is another field where technical aids can be useful in making the treatment more effective and meaningful to the patient. Speech research has been concerned with the diagnosis of speech disorders for a long time. There is an interest in instrumentation for objective measurements of voice characteristics, nasality, and so on.

When Gunnar initiated our first development project, the first objective was to define the requirements in cooperation with the Swedish Association for the Laryngectomized and the Swedish Institute for the Handicapped. We tested the available speech amplifiers and artificial larynges. A study by two speech pathologists provided useful information on the user's views and requirements. Development work could then be started along rather well-defined guide-lines.

A new speech amplifier has been developed and received favorable judgements in a field evaluation program. The production started in 1977 and it is known as the Speech Amplifier TF-2.

Further work was directed to the development of a new artificial larynx. The evaluation of commercially available models has emphasized some drawbacks like the large weight and the lack of intonation control. We were aware of the fact that prosody carries a lot of useful information. Intonation alone would be sufficient provided that controlling was able to learn. We have built a control unit to be operated by one hand. Tests with non-laryngectomized subjects first; later with two laryngectomees, have shown that it is possible to learn to imitate normal intonation patterns. Building portable prototypes was the next step to test the performance in real life situations. We used a separate vibrator unit to be attached to the throat by an elastic band. Rechargeable batteries and electronics were placed in a separate box which was too big for a pocket. Two devices were in use for several months by three laryngectomees. They were pleased to be able to vary their intonation but the aid was too large and the speech level was found too low. In 1977 we built two new prototypes, small enough to carry around in a pocket. The remaining complaints on the low level of speech, vibrator noise and some difficulties in using the hand control unit, put requirements on changes in the design. The most serious problems concern the vibrator unit. Attached to the throat, there is no counter-weight to increase the sound output. Yet, there are communication situations where this speech level is sufficient, e.g. telephone conversation, a quiet room with only a few people etc.

With a one-piece unit, with the mass of the device and the user's hand behind the vibrator, a loud voice and a good speech quality can be achieved at the cost of the tiresome use of a heavy device. I think there are needs for both kinds of artificial larynges. I am planning to design a new one-piece device with intonation control and to try to improve the separate vibrator unit.

Ever since synthetic speech was produced with an acceptable quality, hopes ran high that it would become a useful aid for the non-vocal. There was, however, very little known about how to implement it and what requirements the users would put on aids which speak for them. It was also hard to predict how this kind of communication would be accepted by the environment of the user, the handicapped him- or herself and the general public. In order to test the use of synthetic speech in different communication situations we built a system which was easy to move around. I got the responsibility to take part in the field trial and keep in touch with the therapists and teachers involved. It is very stimulating to cooperate with people interested in new approaches to solve the communication problems of the non-vocal.

In the special school for mobility handicapped children, Bräcke-Östergård in Gothenburg, there are about 15 children who cannot make themselves understood by their own speech. Some of them communicate by writing or selecting symbols, others can only express their most basic needs. Mikael who had learned reading and spelling was the first one to try the Synthesizer. He learned to use it very quickly, using his mouth stick to type his messages. He realized from the beginning that he could communicate with synthetic speech. The experiments with him gave us valuable information.

Three 8-year old children were involved in another experiment where synthesized speech was used to stimulate their learning of reading and spelling. The first effect was a remarkable increase of their motivation. They behaved very differently which could partly be depending on differences in their ability to hit the keys on the terminal keyboard. Only one of the children could use his hand, the other two used head sticks with big difficulties.

This first year of trials confirmed the usefulness of synthetic speech for the non-vocal. The children listened critically to the sound and they had their own ideas of how certain words should sound. Mikael often implemented variations on pronunciation by using phonetic spelling.

We have also realized the importance of a good fitting of the input key-board to the user.

The technology for portable speech synthesizers already exist, but there is not sufficient knowledge of software for obtaining good speech quality, of the psychological factors occurring when communicating with synthetic speech and of the real needs of the handicapped. That is why our approach of making field trials feels important to carry out. It is the only way to find out how people react and in which communication situations synthetic speech is acceptable.

I am very grateful to Gunnar for guiding me into this fascinating work and for supporting me with great interest. The positive atmosphere at the laboratory is a good prerequisite for us to make valuable contributions. I am convinced that increased knowledge of human speech and progress in technology will lead to development of new and improved aids, facilitating communication of many handicapped, thus making their lives more meaningful.

Karoly Galyas

ACOUSTICAL LOGOPEDICS AND PHONIATRICS OF THE DEPARTMENT OF SPEECH COMMUNICATION

In another article in this special report our laboratory is described as a big flowering garden with many different plants and with Gunnar as the chief gardener. We want to follow this metaphor and give a brief presentation of one of these plants which thrives in the rich humus of the garden.

For the senior of the above authors it began already back in 1961 when he went through the garden gates believing that he should work with tactual communication aids for the deaf. After half an hour with the chief gardener he was convinced that the only right thing to do was to study the voice source in speech by means of inverse filtering methods. Initially the goal was to learn more about voice production in order to improve the speech synthesis, but in the mid sixties the idea rose to apply the knowledge and techniques established in the lab to the field of phoniatrics and logopedics.

At that time Gunnar Bjuggren at Sabbatsbergs Hospital had created courses for education of "logopedists", a kind of highly qualified speech therapists. In connection with these courses some ideas were tested by initiating thesis works, where e.g. inverse filtering technique, long-time average spectra, and statistic distribution of the fundamental frequency were used in order to evaluate normal and pathological speech. At that time all these methods were very cumbersome and time-consuming to work with since we did not even have the access to a computer and all data had to be collected with simple measuring technique and evaluated by hand.

However, the seed was planted but it was not much looked after for many years due to other research activities. Then very much happened. The fruitful cooperation with Johan Sundberg and the musical acoustic group in the lab was itensified A rewarding, continuous cooperation was established with Björn Fritzell and Britta Hammarberg, head of, and logopedist at the Phoniatric Clinic of Huddinge Hospital. Other groups were also open for cooperation as e.g. Rolf Leandersson and Bengt Nylén at Karolinska Institutet and Gärda Ericsson at Linköping Hospital.

On the technical side there has also been great improvements in the measuring technique and computer facilities have been established. The introduction of the microcomputers has made it possible to construct better research tools that can be applied to clinical situations since they allow design of compact and cost effective instruments that can perform complex acoustic analyses, make statistics, and present the results directly in the form of numbers, curves or diagrams.

There are many questions to answer before meaningful research can be planned regarding adoption of modern technique for acoustic analyses to the field of phoniatrics and logopedics. First of all we have to face the question if there is a need at all to introduce objective acoustic measurements as a clinical tool. Clearly the answer is YES. This is demonstrated by the vast number of reports and other publications related to this field. The need has also been explicitly defined for instance in the report from the conference on early detection of laryngeal pathology held in Gainesville, Florida in 1973. Among other things the report stressed that the speech pathologist and laryngologist need compact instruments, easy to handle, which can be used clinically as aids in evaluating vocal and laryngeal disorders. As yet no standardized measurement exists to evaluate the voice quality and the laryngeal function, there is for instance no counterpart of the electrocardiograph or the audiometer in the phoniatric field. It was also recognized that there is even a lack of adequate terminology to describe voice disorders and laryngeal behavior and that there exists a strong need in the medical profession to enhance the sensitivity and awareness of the relation between vocal deviation and laryngeal pathology. These needs were also acknowledge at the symposium held at Huddinge Hospital in August 1978. Above we have chosen to stress those needs relating to problems around which we are concentrating our research.

To be able to satisfy any of the mentioned needs we have to gain a better understanding of the relationship between the normal and abnormal function of the voice mechanisms and the acoustic signal, and between the different acoustic features and our perception. This type of knowledge is not only applicable in the area concerned in this presentation but also in other fields such as speech synthesis and speech recognition. However, a major point is that we are not only satisfied to increase the theoretical knowledge about voice production and perception, we also want to bridge the gap between research and practical, clinical work.

To take care of plants in a garden fertilizers and tools are needed. It is not necessary here to mention all the facilities our gardener has created; they are well known to all. Instead we want to demonstrate the active part that he takes in the cultivation of this specific plant in his big garden by referring as an example to his previous and present work about the voice source. We find it very stimulating to be able to work in the fascinating, flowering garden of the Department of Speech Communication at KTH.

Thank you, Gunnar.

Jan Gauffin Antonio de Serpa-Leitão

MUSICAL ACOUSTICS & SPEECH COMMUNICATION

Speech communication research and research in the acoustics of music have been joined in a perfect marriage at the Department of Speech Communication, KTH since 1961. The starting point of this marriage was mere chance, as is indeed the case with many other marriages. A colleague of Gunnar Fant happened to be a neighbor of Frans Fransson, who had an ardent and productive interest in the field of acoustics of music. Moreover, Frans Fransson happened to be an exceedingly fine personality, who had no trouble in successfully entering a new carreer at the age of 70. Soon Fransson expanded the music acoustics research activity in the lab by attracting two more scolars, who later on passed their doctor exams. Today the future of musical acoustics at KTH looks brighter than ever, as a personal professorship in acoustics of music has been given to Johan Sundberg and a long-term research position has been given to Erik Jansson. The main motor behind these achievements has been Gunnar Fant. But his significance to musical acoustics is more profound than this. Our research has profited in several important respects from the work done by Gunnar and his staff at the lab, e.g. regarding theoretical framework, research tools and methods, and questions of mutual interest. We feel it to be adequate and important to give some examples of this at this moment.

Speech and music are both interhuman communication by means of sound waves. The message is coded into sound which is generated by the sender: the speaker or the musician. In the case of speech the human voice organ serves the function of a sound generator, while in music it is the instrument which generates the sound. A scientific understanding of both speech and music requires that the sound generation process is properly understood. In the human voice organ as well as in musical instruments acoustic resonators play a crucial role. In his Acoustic Theory of Speech Production Gunnar applied acoustic theory to the vocal tract resonator. Given the dimensions of the resonator, its resonance frequencies could be computed. It turned out that Gunnar's equations could also be used to calculate the resonance frequencies of organ pipes, a task which was treated in the doctoral dissertation of one of the present authors. Thus, the theoretical framework needed was available.

Obviously, the human voice organ can be used as a musical instrument. In our research on the singing voice, Acoustic Theory of Speech Production has typically served as a sort of starting point, one of the main questions being: is this theory capable of predicting the acoustic characteristics of singing, or must we accept inexplicable mysteries, such as the serious belief of some singers that the vocal folds are not used in singing? As yet no inexplicable acoustic characteristics have been revealed, though. It is not a question of politeness that Acoustic Theory of Speech Production is a standard reference in our research.

As speech and music are both cases of communication by means of sound, many research tools developed within speech science have been found useful in acoustics of music. There are several examples of this. We may point at procedures for measuring fundamental frequency, which was explored and developed in a three-year project aiming at automatical notation of tape recorded melodies. Starting with this project Anders Askenfelt joined our group. Another example is the long-term-average-spectrum (LTAS) which turned out to be a useful tool for mapping certain aspects of instrument timbre. The idea of analysis by synthesis has been successfully adopted for musical purposes even with respect to the analysis of musical style, as in the generative theory of music. And for obvious reasons nearly all of our research in the singing voice has taken advantage of existing facilities developed and available at the lab. The text-to-speech program has been modified to a notation-to-singing version. The speech synthesizer family OVE has gotten a musical cousin called MUSSE. In these and other cases musical acoustics started its work in a workshop that was already well equipped.

The fact that both speech and music are human communication by means of sound implies that several questions have a direct relevance to both research fields. Both speech and music are received and processed by the same sensory organ, the ear. Thus, the choice of acoustic coding of information is dependent on and thus reflects the properties of hearing. Therefore, a complete scientific understanding of speech as well as music requires a thorough understanding of hearing processes. As a sequel of this several investigations have been jointly made by members of the speech and music groups.

Another aspect is worth mentioning here, which hopefully will generate future joint research efforts. In good music performances the important feature is that an emotional content is communicated to an attentive listener. The most apparent example is perhaps singing, where a skilled singer's performance mirrors the emotions of the poem. But instrumental music certainly contains the same or a related component. It seems likely that the coding of emotions in music can be understood by a listener because he is acquanted with this code from his non-musical experience, e.g. from speech. In other words there would be common ways of coding emotions in speech and in music. The principles of this coding can probably be most efficiently analyzed by a comparative examination of emotional speech and singing.

Above we have mentioned some more or less randomly chosen examples of the ties between speech and music research. These examples all demonstrate that our research in music has profited from the close contact with the lab. However, the most decisive advantage has not been mentioned, namely the fact that we form a part of the community at this lab. The frequent discussions of findings with Gunnar himself and with the people which he has managed to collect in the lab has been of a great educational value. It has helped us to acquire and maintain a sense of quality in research. There are also personal aspects on this. Several of our closest friends are in fact our colleagues at the lab. The personal and scientific atmosphere at the lab has been important for our research.

Gunnar Fant is a rather well known person in this country. Some time ago one of us met a member of the Parliament and mentioned Gunnar's name. The reaction was "Oh, that's the man who built a machine singing HELAN GÅR (a Swedish aquavitae tune) with a metallic voice quality!". It is now many years since the speech synthesizer was brought to sing that song. We hope that future will bring opportunities to increase a reversed traffic between acoustics of music and speech science, so that acoustics of music does not only make use of research results of speech research but also offers speech research a greater number of useful results. (E.g. by improving both voice timbre and musical quality in speech synthesis!)

As previously mentioned, the starting point of the marriage between speech science and acoustics of music at the lab was a fruit of mere chance. Above we have pointed at several facts demonstrating that the preservation of this marriage is far from being a fruit of mere chance. It is natural in view of the numerous points of common interests. - But thinking a bit more we arrive at this point to the conclusion that it has not been a perfect marriage at all. For instance, the speech research lives in a rather polygamic way at the lab, housing several other research groups than our music group. We experience this polygamy as positive, but we find it hard to accept that polygamy is prosperous in marriage. So, at this point we would prefer thinking of the lab as a greenhouse, where Gunnar is the successful master gardener. We congratulate ourselves to the dignity of being his plants and we congratulate Gunnar to his 60th anniversary!

Anders Askenfelt Erik Jansson Johan Sundberg

Johan S

Anders C

Erik

Jesús Alonso