BNS-7722305 | STANFORD UNIV | PLEASE RETURN BY 21 SEP 1977
PRINCIPAL INVESTIGATOR: CHOWNING JOHN | NSF PROGRAM: SENSORY PHYSIO & PERCEPTN

EXPERIMENTS IN TIMBRE PERCEPTION

'he timbre perception proposal is superb! It is a pattern for the kind of research that should be emphasized today. It is one of the too rare cases where scientific studies are making real contributions to an area of great applied importance. There are plenty of areas where good scientific methods are discovering interesting new things about the world, but often the new information does not have clear importance in applications. Likewise, there are plenty of important practical problems, but too often scientific methods can make only weak contributions to their solutions. The fundamental studies proposed here can produce information of enormous importance to music, both for immediate applications and for the far future.

'he proposal both demonstrates that great progress in understanding timbres has already been made and that there are many potent questions which can now be attacked with presently available methods. The proposal could not have been written five years ago. The results pointed out in the review of timbre research are principally from the last five years. Likewise, the methods of study are new in this period.

t is proposed to focus three techniques on timbral questions. These are:

- New methods of analysis of sound waveforms, specifically heterodyne filters and phase vocoders. These techniques promise to provide better spectral information about sounds than has ever before been available, in particular the ability to track subtle changes in the frequencies of overtones. The heterodyne filters were developed by the Stanford group itself. Phase vocoders were developed at Bell Labs, but have been utilized for sound analysis for the first time by Stanford.
- Completely general sound synthesis methods. The pure digital synthesis techniques now available at Stanford allow any sound to be immediately and conveniently synthesized with precise control of all its parameters. Comparison of synthesized sounds and naturally produced sounds is one of the essential tools in the Stanford arsenal. It makes possible identification of the perceptually important factors in the timbres.
- Multidimensional scaling and the identification of the perceptual similarity spaces in which people subjectively classify timbres. This very powerful new data analysis technique enables one to systematically order the relationships between various timbres. This is essential because timbral research has reached the stage where we can study whole groups of timbres rather than individual sounds. It is also essential to understand the relations between sounds in order to utilize them in practical applications.

hese three research methods mutually interact and reinforce each other, and it is their combined use that promises much progress in understanding timbres.

In past proposals the National Science Foundation has had to consider whether it was supporting music as an art or whether it was supporting a scientific study. In this proposal, there is no question about the focus of the effort. The work is entirely of a scientific nature; only the applications concern music. The personnel are appropriate for scientific work. Chowning, Grey, and Moorer are recognized leaders on purely scientific bases and Moore has just obtained his PhD in EE. Furthermore, the work is likely to have important implications in related scientific fields such as speech studies and multidimensional scaling methodology.

(Continued on Page 2)

ATIONAL SCIENCE FOUNDATION

UNIVERSAL REVIEWER FORM (PAGE 2)

NSF TEST FORM X

OPOSAL NO.	INSTITUTION	PLEASE RETURN BY
BNS-7722305	STANFORD UNIV	21 SEP 1977

INCIPAL INVESTIGATOR	NSF PROGRAM
CHOWNING JOHN	SENSORY PHYSIO & PERCEPTN

EXPERIMENTS IN TIMBRE PERCEPTION

lthough it is not central to the proposal, it is appropriate to point out why imbral studies are currently so important musically. In the past, timbres ere limited to those producible by existing mechanical musical instruments. oday, digital techniques and cheap, complex integrated circuits have removed ll restrictions from the control and production of sound. We can make any ound we can understand and our only limitations are in our understanding. he Stanford group, in a separate effort, has contributed greatly to the evelopment of these digital techniques.

he Stanford proposal is very inexpensive relative to the benefits likely to be roduced. I have been involved with setting up a French laboratory for the tudy of musical acoustics (Institute for the Research and Coordination of coustics and Music). The institute has ten times the personnel than the tanford group. They are attacking many similar problems. The Stanford group s presently far ahead of the Paris institute and given modest support, it is ikely to stay ahead. The reason for the efficiency of the American group, n my opinion, is the very strong technical environment supporting them which ontains readily available expertise in a full range of subjects from digital ignal processing to experimental psychology.

would like to suggest that parts of the Stanford proposal be published. It ould serve as a model for proposals for applied scientific research. But in dition, the review of timbre research is one of the best summaries of this ea that I have read and the clear questions for study which have been oposed could have an important impact on researchers in this field everywhere.

ATTENTION: Sherman Lee GUTH

NATIONAL SCIENCE FOUNDATION

Proposal No. BNS-7722305
Institution: Stanford University
Principal Investigators: John CHOWNING & John GREY
NSF Program: Sensory Physiology and Perception

EXPERIMENTS IN TIMBRE PERCEPTION

GENERAL IMPRESSION:

The proposal is highly meritorious and presents numerous opportunities for major advances in our understanding of the perception of complex auditory material. Grey, Moore, and Moorer are well trained scientists and have demonstrated considerable resesrch productivity. (They also know a great deal about music.) Chowning and Rush are well trained musicians with considerable background and experience in scientific investigation. Based on past performance and this propossl, I strongly support continued funding of their project. They have assmbeled an impressive set of tools and methodology and have produced with them a number of useful results.

SPECIFIC COMMENTS:

1) The proposed analysis-based synthesis methods for the production of naturalistic stimuli, exemplified by heterodyne filter analysis and additive synthesis (Appendix IV), the modified linear-predictor (Appendix III), and the phase-vocoder (Appendix V), are mathematically sound and computationally practical. I can attest to the fact that these procedures do indeed work by virture of the fact that they have been made operational on IRCAM's PDP-10 by Moorer. The stimuli produced by these methods are of extremely high quality. I found the discussion in Appendix III on the relationships between modeling speech and music useful. I believe that the Stanford group's insistence on synthesis of high acoustic quality will provide beneficial fall-out upon the speech researchers who often because of a lack of good synthesis schemes must be satisfied with speech of poor quality. The Interactive Acoustic Manipulation Language described in Appendix VI should also prove very useful in speech research.

2) One especially important feature of some of the proposed lines of research is the use of temporal contexts. Nearly all psychoacoustical investigations of timbre have used isolated tones. The results of such studies may not generalize to the richer contexts as found even in simple music. It is important and timely to begin an assessment of the effects of such contexts. Grey reported an extensive experiment on the effects of context upon the discriminability of different timbres at the IRCAM musical psychoacoustics symposium held in Paris in July. Grey found that the perceptual significance of certain cues for discrimination was systematically modified by the introduction of a melodic context.

3) It is significant that the proposed scaling experiments are not just one-shot correlational type studies but involve systematic modifications of the stimuli and rescaling for verification of the interpretation given the solutions. A detailed example of this approach is provided in Appendix XI by the Grey and Gordon experiment on the perceptual effects of spectral modifications. Another example can be found in the experiment where Grey performed a rescaling of a set of tones with tones constructed by interpolating between members of the original set added (Appendix XII).

4) The proposed studies on categorical perception are also timely now that such effects have been found with a variety of non-speech auditory stimuli. It appears that categorical like perception is not special to speech and probably depends on general perceptual mechanisms. Pisoni and Jim Miller have recently presented results supporting this view. The analysis-synthesis approach provides for systematic interpolation between complex natural sounds and should considerably advance the demystification of categorical perception.

5) As I am now at IRCAM, I can report upon the rather special and powerful influence that the Stanford group is having upon research and development here and upon auditory and speech research in Europe. The PDP-10 system here at IRCAM is equipped with much of the special and quite powerful software from the Stanford Artificial Intelligence Project: the monitor, the FAIL and SAIL languages, the E text editor to name a few. A large portion of the sound analysis and systhesis programs

described in the proposal have been made operational at IRCAM. The Stanford developments are well documented and reliable and will doubtlessly be used by a number of prominent musicians and auditoy researchers. At a recent IRCAM Symposium on Musical Psychoacoustics the Stanford team made four different and impressive presentations (program enclosed). As a result of the extensive cooperation with Stanford, IRCAM has become the European outlet for many of the products of their research.

6) The Stanford group has provided other investigators with the data from the analysis of instrumental timbres. At Michigan State University we received a set of amplitude and frequency functions for additive synthesis and experienced no difficulty in generating and modifying the tones on our PDP-11. We successfully replicated the scaling experiments and then used the tones to study the perception of auditory analogies of the form: tone A is to tone B as tone C is to tone X. The choices made for tone X, the analogy solution, were predicted from the scaling solutions.

7) The budget seems in good order. The addition of a disk drive seems justified considering the large amounts of digital data involved.

NATIONAL SCIENCE FOUNDATION
WASHINGTON, D.C. 20550

March 6, 1978

Dr. John M. Chowning
Dr. John M. Grey
Department of Music
Stanford University
Stanford, CA. 94305

RE: BNS77-22305

Dear Drs. Chowning & Grey:

Please find enclosed verbatim copies of the reviews of your recent proposal to the National Science Foundation. I hope you will find them informative and helpful.

Sincerely,

Sherman L. Guth
Program Director
Sensory Physiology
and Perception

Enclosures

77-22305

The panel summary was: "There was agreement that this was an excellent proposal qua proposal. Two concerns were discussed at length, (1) the extent to which this work will make a contribution to perception as opposed to the fields of music or aesthetics and (2) the need to cut the budget substantially perhaps primarily in personnel."

NATIONAL SCIENCE FOUNDATION

PROPOSAL RATING SH

PROPOSAL NO.	INSTITUTION
BNS 77-22305	STANFORD UNIV.

PRINCIPAL INVESTIGATOR: CHOWNING, JOHN C

NSF PROG

TITLE

EXPERIMENTS IN TIMBRE PERCEPTION

COMMENTS (CONTINUE ON ADDITIONAL SHEET(S) AS NECESSARY)

I am impressed with the thoroughness of this proposal. If anything it is too redundant. However, concerning your points of reference:

1. Sci. importance. I've worked off & on in this field since my MA theses (1941 State U. of Iowa) & have followed the work of PLOMP & others & find the work very important in the framework of psychoacoustics & more broadly in the use of multicimensional scaling. Did rate it E.

2. Soundness. It appears they have covered the gamut; analysis, synthesis, scaling techniques, computer know-how & avilability. With several papers already in the best journal in this field (J. Acoust. Soc. Am.) & with the availability & cooperation of the San Francisco Symphony & with the development of direct digetalization, they are certainly sound in their proposal. I'd rate it V/E
3. Capacity of applicants. I don't know them personally, nor have I heard them talk at the Acoustical meetings (usually don't have time to go to musical acoustic sessions, unfortunately), but I have read their papers & do know their consultants. Also they have the WHOLE team approach. I'd rate them V/E.
4. Aware of recent developments. They are as aware, if not more so, than I am. They know of the phase vocoder, the newest scaling techniques, etc. You'd have to get a better judge than me to fault hem on keeping up with the literature. In my opinion, E.
5. Importance. To this field it is imperative. They have zeroed in on the problem, temporal relationships & have the ability, interest, know how & equipment (computer). They sould proceed post haste. E

I don't claim any knowledge of budget matters. The sum seems large but so is the imporatnce of the work & the wherewithall to do it. I would hope NSF could sponsor them.

RATING: ☐ EXCELLENT X ☐ VERY GOOD ☐ GOOD ☐ FAIR ☐ POOR

UNIVERSAL REVIEWER FORM FORM X

PROPOSAL NO.	INSTITUTION	PLEASE RETURN BY
BNS-7722305	STANFORD UNIV	

PRINCIPAL INVESTIGATOR	NSF PROGRAM
CHOWNING JOHN	SENSORY PHYSIO & PERCEPTN

EXPERIMENTS IN TIMBRE PERCEPTION

Although it took some time to read Chowning's proposal, I can be brief in my summary evaluation of the proposal.

Without hesitation I put the proposal in the category "Excellent."

The issues addressed under the rubric timbre are fundamental and solutions to associated problems should have wide applicability. The project is generally sound, and one should expect good productivity given the past performance of Chowning and the other principals.

It would be very hard to find another group with both the technical skills and the intellectual ability to carry out the proposed work, at least in the United States. The only competitor is Pierre Boulez' group at IRCAM in Paris, and even they have made heavy use of American talent, from Stanford, UCLA, Bell Labs and Michigan State, at least. The Bell Laboratories group have been widely influential on the technical side but their contributions to musical theory and perception rather less. Of course this remark does not apply to the enormously important work of Bell Laboratories in psychoacoustics and speech, only to the musical theory and perception.

The project is well planned and the principals are certainly aware of recent developments in the field, though sometimes they appear not to be, as when a few straw men are set up in the prolegomena. The least appealing part of the project is that on categorical perception, a controversial problem which keeps evaporating as it is examined in increasingly intense logical light beams. The work of Grey (referred to in the proposal) doesn't find categorical perception, no doubt in part because Grey is aware of the pitfalls due to method and tries to avoid them. (Incidentally, the issues of categorical perception concerned with the criterion problem are illuminated by Macmillan et al. in the very latest August (?) 1977 Psychological Review.) Naturally I don't blame Grey for the existence of the 15-year old issues of categorical perception. Many respectable investigators have been caught up in it. I suppose it is important to clarify the issue in musical perception as well.

Chowning's project is important because it addresses a very broad problem with connections from aesthetics and technology to perception, performance and ~~listening for pleasure. Music, far more than art, is the realm where the objective~~ and subjective can be tied together with scientific and logical rigor. Musical perception is connected to every aspect of hearing--temporal, spectral, binaural, spatial and timbre (which is so ill-defined as to be a virtually unlimited domain) and deserves investigation.

If Chowning's group can keep its impetus it may come to be a national facility for musical/psychoacoustics research much as Haskins Laboratories has been for speech researchers. Thus they could serve many researchers whose own

Report on BNS-7722305, John Chowning's proposal to the NSF for "Experiments in Timbre Perception" (Continued)

facilities are far less grand. Obviously the Stanford Center would be an unparalleled situation for musicians, or psychologists, or engineers who wish to learn essential complementary methods and ideas.

NATIONAL SCIENCE FOUNDATION

UNIVERSAL REVIEWER FORM

NSF TEST FORM X

PROPOSAL NO.	INSTITUTION	PLEASE RETURN BY
BNS-7722305	STANFORD UNIV	

PRINCIPAL INVESTIGATOR	NSF PROGRAM
CHOWNING JOHN	SENSORY PHYSIO & PERCEPTN

EXPERIMENTS IN TIMBRE PERCEPTION

On the review summary I have noted this proposal as excellent in all categories and my confidence in these ratings is high. I do not feel that this is excessive.

I believe that the only places where there are personnel facilities and work comparable to those at Stanford are IRCAM in Paris (see proposal, pp. 51-52 and 56-57) and Bell Laboratories. This is not meant as a criticism of excellent work relevant to the scientific aspects of music and timbre by leading acousticians such as Plomp and Wessel.

At Stanford, Bell Laboratories and IRCAM excellent people with excellent facilities are involved in programs whose motivation is at once scientific and musical. I have followed the work at these three places closely; as a one-time employee and present consultant at Bell Laboratories, as a member of a scientific Advisory Committee of IRCAM, which I visited in 1976 and 1977, and through a number of visits to Stanford. In coherence of program and variety, and competence of talent available, I would put Stanford first among the three.

The present proposal is by far the best available account of musical timbre. It brings together and compares past work relevant to timbre perception and discusses the techniques available both for analysis of responses to the timbre sounds by multidimensional scaling and by clustering.

Our understanding of the perception of timbre is incomplete. The proposed research is a well-chosen continuation, extension and amplification of outstanding work that has already been done. All the areas proposed for further research, analysis by synthesis, multidimensional scaling and categorical perception, are those most reasonable in seeking to unravel the problems of timbre percetion. The strength of the attack lies in the facts that the work deals with real musical sounds and that it is to be carried out by people who have high musical competence and excellent understanding of both psychophysical methods and who have excellent facilities for analyzing and synthesizing sounds and for analyzing data.

I regard the budget as very reasonable for the people involved and the work proposed. To do comparable work anywhere else would be very costly, because Stanford has equipment and expertise beyond that available elsewhere.

REVIEW SUMMARY

(Optional trial form -- please submit in addition to, but not instead of, standard review sheet.)

Rating Scales:

Proposal: E (Excellent), V (Very Good), G (Good), F (Fair), P (Poor)
Confidence: H (High), M (Medium), L (Low)

CATEGORY	Your Rating of Proposal (E, V, G, F or P)	Your confidence in your own judgment (H, M or L)
Potential Scientific Contribution	E	H
Competence of Investigator	E	H
Project Feasibility (Is proposed duration appropriate, are facilities adequate, etc.)	E	H
Productivity of Investigator	E	H
Reasonableness of Budget	E	H
Other (Please specify)		
OVERALL RATING (should be the same as shown on regular rating form)	E	H

Comments on this trial form and/or on our general review procedures:

NATIONAL SCIENCE FOUNDATION

UNIVERSAL REVIEWER FORM

NSF TEST FORM X

PROPOSAL NO.	INSTITUTION	PLEASE RETURN BY
BNS-7722305	STANFORD UNIV	

PRINCIPAL INVESTIGATOR	NSF PROGRAM
CHOWNING JOHN	SENSORY PHYSIO & PERCEPTN

EXPERIMENTS IN TIMBRE PERCEPTION

This is a very well written proposal. It includes an excellent review of the literature, a detailed outline of the problems, the proposed experiments, techniques of analysis and the possible importance of the results. The investigators clearly have the expertise to carry out the research. This expertise is required in several areas covering psychoacoustics, music, electrical engineering, and multi-dimensional scaling. The investigators have been productive and creative in areas related to the various phases of this research. They are sufficiently tooled up for this kind of work, that the research outlined is clearly feasible. There are very few laboratories in the country capable of executing this research. There seems to be little advantage in going through the ritual of explaining the manner in which this proposal is excellent in all of these regards.

There remains only one remaining question, namely the scienfific importance of this research. I do not work in this area and may not be able to judge, and certainly not with confidence, the impact of expected results. I can well imagine that individuals working in synthetic music and related areas would be very interested in this research. I also know that many excellent psychoacousticians are interested in some closely related problems of the perception of complex signals. In many regards, one area's applied problem may be another area's theoretical problem. I would think that this research will make more of a contribution to multidimensional scaling than it will to perception, although the topic is nominally timbre. It will help us specify what kinds of stimuli will lead us to say tones differ in timbre or some other words that are dorrelated with it. If research of this topic has high priority no one will do it better than this group.

1 on proposal

2 on importance

NATIONAL SCIENCE FOUNDATION

PROPOSAL RATING SHEET

NSF FORM 990
TEST (1-77)

PROPOSAL NO.	INSTITUTION	PLEASE RETURN BY
BNS 77-22534 305	STANFORD UNIVERSITY PENNSYLVANIA STATE UNIV.	

PRINCIPAL INVESTIGATOR	NSF PROGRAM
PUSCKS, BENJAMIN H. CHOWNING JOHN M.	DIV. OF BEHAV & NEURAL SCI SENSORY

TITLE

AFFERENT MENTAL PATHWAYS PERCEPTION DISCRIMINATION

COMMENTS (CONTINUE ON ADDITIONAL SHEET(S) AS NECESSARY)

This is a very impressive proposal. Indeed, it is difficult for me to imagine how this investigation of timbre could be significantly improved. As I began reading this proposal, I was skeptical about how the phenomena might be over simplified in order to force them into particular linear analytic molds, but after reading this proposal I am impressed with the investigators' flexibility and competence in developing analytic tools that might adequately reflect the complexity of the phenomena. The investigators' broad technical sophistication in music theory, acoustic signal analysis, and psychophysics are certainly impressive. So far as I can tell, they are at the state of the art in all of these fields. Additionally, this project has general implications for other perceptual phenomena--not only for speech perception as the investigators point out, but also for other cases of context effects in the perception of components of dynamic event.

The most obvious question about this project concerns the budget. Although this project has been productive, and although a considerable amount of research is proposed for two years, the personnel budget is still unnecessarily large. Even though the many competent investigators and staff personnel must have contributed to this proposal, they should not all be needed to be actively involved in conducting this research.

RATING: [X] EXCELLENT [ ] VERY GOOD [ ] GOOD [ ] FAIR [ ] POOR

PRINCIPAL INVESTIGATOR COPY