IPO ANNUAL PROGRESS REPORT

no 1 1966

Editor: A. van Katwijk  Adviser: H.J.A. van Beckum  Secretary: Miss W.L. Akroyd

INSTITUTE FOR PERCEPTION RESEARCH - INSTITUUT VOOR PERCEPTIE ONDERZOEK

INSULINDELAAN 2 EINDHOVEN  TELEPHONE: 040 - 60000 - 6605 / 040 - 33222 - 2485
SUPERVISORY BOARD I.P.O.

Prof. dr. C. Zwikker (Chairman) - Eindhoven
Drs. J.H. Bannier - Den Haag
Dr. ir. F.W. van Berckel - Bunnik
Prof. dr. H.B.G. Casimir - Eindhoven
Dr. G.J. Fortuin - Eindhoven

SCIENTIFIC BOARD I.P.O.

Prof. dr. H.B.G. Casimir (Chairman) - Eindhoven
Prof. dr. M.A. Bouman - Utrecht
Prof. dr. M.C. Colenbrander - Leiden
Prof. dr. J. Droogleever Fortuyn - Groningen
Prof. dr. L.W.B. Jongkees - Amsterdam
Prof. dr. J. Koekebakker - Leiden
Prof. dr. ir. H. Mol - Amsterdam
Prof. dr. A. Oldendoroff - Eindhoven
Prof. dr. A.J.B.N. Reichling - Amsterdam
Prof. dr. L.H. van der Tweel - Amsterdam
Prof. dr. A.J.H. Vendrik - Leiden
Prof. dr. ir. C.J.D.M. Verhagen - Delft
Prof. dr. P.M. van Wulfften Palthe - Den Dolder
Prof. dr. ir. A. van Wijngaarden - Amsterdam

DIRECTOR I.P.O.

Prof. dr. J.F. Schouten - Eindhoven
In the nine years of its existence, the Institute for Perception Research published 26 Progress Reports at four-monthly intervals. These "Voortgangsrapporten" were written in the Dutch language and aimed at keeping our Supervisory Board, our Scientific Board, our sponsors and those immediately connected with our institute informed on the totality of our activities. Hence it was a mixture of an administrative and a scientific report.

As our activities grew, this mixture was felt increasingly to be a disadvantage. Hence it was decided to split our reporting into two different publications: a mainly administrative report in Dutch at half-yearly intervals and an Annual Progress Report in English dealing with our scientific and technical activities, to appear at the end of each year.

Many befriended laboratories in the world issue such Progress Reports, which always arouse an eager interest as a survey of the totality of their recent activities. The less specialized the scope of these laboratories is, the more this general survey tends to be lost by publications in widely different scientific journals.

The contents of our present first attempt can be divided into three categories:

1. Summaries of scientific and technical contributions for professional journals or congress proceedings.
3. Short notes on simple methods or handy instrumentations which so often are the very gist of the efficiency of research and which yet, all too often, remain unpublished.

We hope that our many friends will enjoy receiving our Progress Reports and, most of all, that it will serve its purpose of useful communication with all those who are working in similar fields of research.

J.F. Schouten
2. CONTENTS

1. INTRODUCTION

2. CONTENTS

3. PERSONNEL

4. THE INSTITUTE FOR PERCEPTION RESEARCH

5. RESEARCH

a. AUDITORY PERCEPTION
   - J.L. Goldstein
     Auditory Spectral Filtering and Monaural Phase Perception
   - J.L. Goldstein
     Auditory Nonlinearity
   - G. Domburg
     The Just Noticeable Difference for Loudness
   - R.J. Ritsma, B.L. Cardozo, G. Domburg, J.J.M. Neelen
     The Build-Up of the Pitch Percept
   - R.J. Ritsma
     The "Octave Deafness" of the Human Ear
   - B.L. Cardozo, R.J. Ritsma, G. Domburg, J.J.M. Neelen
     Unipolar Pulse Trains with Perturbed Intervals; Perceptibility of Jitter
   - R.J. Ritsma
     The Pitch of Sinusoids and Complex Signals as Affected by Masking Noise
   - B.L. Cardozo, Th.A. de Jong, R.J.J. Boddeke
     Improvement of Intelligibility by the Application of Directional Microphones in a Simulated Hearing Aid Situation
   - J.F. Schouten, J. Dijk
     Rustle Noise
   - J.F. Schouten, G. Domburg
     Pitch and Anticipation

b. SPEECH
   - A. Cohen, A. van Katwijk
     The Perception of Phonemes as a Function of Acoustical and Distributional Cues
   - I.H. Slis
     A Model for the Distinction between Voiceless and Voiced Consonants
   - A. Cohen
     Errors of Speech and their Implication for Understanding the Strategy of Language Users
   - G. Forrez
     Relevant Parameters of the Voiced Fricative /z/
   - J. 't Hart
     Perceptual Analysis of Dutch Intonation Features
   - A. van Katwijk
     On Perceptual Units in Speech

c. VISUAL PERCEPTION
   - H. Bouma, J.J. Andriessen
     Anchor Effects in the Perception of Slant
   - H. Bouma, J.J. Andriessen
     Induced Changes in Perceived Orientation
   - H. Bouma
     Visual Form Perception as Approached by a Technique of Selective Suppression
I.G.H. Ishak, H.J.J. van Bussel
Magnitude Estimations of Colour Attributes 63

J.A.J. Roufs
On the Relation between the Threshold of Short Flashes, the Flicker Fusion Frequency and the Visual Latency 69

J.F. Schouten
Subjective Stroboscopy and a Model of Visual Movement Detectors 78

d. PERCEPTUAL & MOTOR SKILLS

W.G. Koster
Can Perceptual Load be Measured? 81

W.G. Koster
The Psychological Refractory Period 84

F.F. Leopold, H. Stremmelaar
A Comparative Study of Human Performance with Dial and Keyset Telephones 86

J.J. Andriessen, J. Vredenbregt
Measurements on Pin Assembly Tasks 90

J. Vredenbregt, W.G. Koster
Some Aspects of Muscle Mechanism in Vivo 94

J.F. Schouten, J.A.M. Bekker
Reaction Time and Accuracy 100

e. MISCELLANEOUS

J.F. Schouten, W.G. Koster
Electrocardiophony 101

J.F. Schouten
Industrial Evolution 102

6. INSTRUMENTATION & TECHNIQUES

G.J.J. Moonen, C.A. Lammers
A Preset Cascade Counter 104

D.J.H. Admiraal
Segmentators for Sound Analysis 106

B.L. Cardozo
A Sequential Up-and-Down Method 110

D.J.H. Admiraal
A Sound Spectroscope with High Scanning Frequency 114

L.F. Willems
IPOVOX II: A Speech Synthesizer 120

L.F. Willems
The Intonator 123

D.J.H. Admiraal
An Infra-Red Pupillometer Based on Television Scanning Technique 126

7. SELECTED RESEARCH TOOLS

J. 't Hart, I.H. Slis
Tape Loop Arrangement 135

J.J. Andriessen, H.S. Fuchs, J. Vredenbregt
Electro-Mechanical Histometer 136

J. Vredenbregt
Suction Cup Electrodes for Electromyography 137

J. Vredenbregt
An Accelerometer for Use at Low Frequencies 138

G.J.J. Moonen, C.A. Lammers
Lining Up Information in a Shift Register 138

F.F. Leopold, H.S. Fuchs
Conversion of a Typewriter for Typing on Paper Tape 140

F.F. Leopold, H.S. Fuchs
Paper Tape Reading Device 141

L.F. Willems
Vowel Synthesizer 142

8. I.P.O. PUBLICATIONS
Personnel

Director
Prof. dr. J.F. Schouten

Research Associates

D.J.H. Admiraal
Dr. H. Bouma (for one year at McGill University, Montreal, Canada)
Dr. A. Cohen
Drs. A.F.V. v. Katwijk
Ir. W.G. Koster
Drs. B. Lopes Cardozo
Drs. S.G. Nooteboom
Dr. R.J. Ritsma
Ir. J.A.J. Roufs
Ir. C. Schumm
Drs. J.M. Westhoff (till 30.4.1966)
Ir. L.F. Willems

Research Staff

M.A. Alewijnse
J.J. Andriessen
J.A.M. Bekker
H.J.J. van Bussel
G. Domburg
J. 't Hart
Th.A. de Jong
F.F. Leopold
H.J. Meulenbrugge
G.J.J. Moonen
I.H. Slis
J.C. Valbracht
J. Vredenbregt

Research Assistants

P.B.W. Boonstra
C.H.J.M. Goudeseune
P.F.R. Kalshoven
J.J.M. Neelen

Technical Assistants

C.A. Lammers
A.R.M. Loonen
W.H. Noordermeer
B.J.J.E.M. Poulissen
M. van Zwam

Secretaries

Miss W.L. Akroyd
Miss L.M.W. Bögemann
Miss E.E. Moorman
Miss I.J.M. van der Valk

Workshop

H.S. Fuchs
C.G. Basten
H.E.M. Mélotte
E.W.L. Kanters
C.W.P.M. Mikkers
J.W.F. Minten
C.J.M. Salij
H.W.W. Smulders

Guest Researchers

Dr. G.H.J. Forrez 6.10.1965 to 31.1.1966; Katholieke Universiteit Leuven, Belgium.
Dr. J.L. Goldstein till 14.10.1966; Harvard University, Cambridge, U.S.A.
Dr. G.H. Ishak 19.6.1966 to 28.9.1966; Ain Shams University, Cairo, U.A.R.
Ir. C.G.H. Scholten, Philips International Institute, Eindhoven, Holland.
4. THE INSTITUTE FOR PERCEPTION RESEARCH

Foundation and Organisation

The Institute is a joint undertaking of Philips Research Laboratories and the Technological University at Eindhoven. Legally it is a separate foundation (September 12, 1957) linked by two non-exclusive agreements (May 31, 1958) to its sponsors towards providing staff, equipment and housing. The foundation is governed by a Supervisory Board in which the sponsors as well as the Netherlands Organisation for Pure Scientific Research are represented. A Scientific Board acts as an adviser to the Supervisory Board and the Director.

Preparatory action started in June 1956. Research began in July 1957 in provisional lodgings. On September 19, 1958 the newly erected laboratory was officially inaugurated.

By now the total staff consists of 45 members, both from Philips Research Laboratories and from the Technological University. In addition the institute accommodates a floating staff of students and guestworkers from Holland and abroad.

Scope

The initial impetus to found our laboratory arose from the management of Philips Research Laboratories. It was felt that a great many industrial products and methods are judged not only by their technological merits but, increasingly, also by their merits in terms of the human capabilities and restrictions of observation, interpretation and manipulation. It was thus decided to found a special laboratory dealing with these problems of a more subjective nature. The Technological University at Eindhoven, founded in 1956, fell in with these ideas, since students in technology should be made aware of the fact that in balanced engineering both the technological and the subjective aspects are to be taken into account.


In Auditory Perception the focus of interest is the mechanism of hearing and in particular the properties of perceptual sound analysis and the perception of pitch and timbre. Applied research is carried out on hearing aids and on the perceptual evaluation of sound reproduction.

In Speech the relevant physical parameters of phonemes and higher linguistic structures are studied. Improved instrumental methods of speech analysis and electronic speech synthesis are used as research tools. Applied research deals with speech transmission and language teaching.

In Visual Perception the original scope aimed at the dynamics of the visual system: adaptation, visual delay and pupil behaviour. Later the perception of form and of movement were added, as well as colour vision and eye movements. Applied research deals with the perceptual evaluation of visual conditions (lighting) and visual aids e.g. a simple entoptic pupillometer for use in ophthalmology.

In Perceptual and Motor Skills manual operations and reaction times are studied together with, as an offspring, the dynamics of muscular action.
Applied research deals with a variety of problems related to industrial ergonomic design and, in an incidental fashion, with aids for the invalided.

In Instrumentation a great variety of instruments, as yet commercially unavailable, are developed for our own purposes of research. At times these instruments find applications in other fields like language teaching and medical electronics.

In spite of the seemingly heterogeneous character of the themes involved, a strong similarity in experimental approach and method should be noted.

Great stress is laid upon the individual human faculties of perception. This involves a critical introspection of the reliability and reproducibility of the observed phenomena and of their correlation with objective stimuli. The universality of the particular perception can be checked by group experiments.

The essence of this approach is that the human perceiver is used as a measuring device, determining not only thresholds, equalities and just noticeable differences of perception, but also expressing the magnitude of a perception in terms of a subjective scale.

When, in 1859, Fechner introduced the science of psychophysics he defined it as the study of phenomena of a psychological nature by physical methods. Consistent with this definition the first three items of our scope could have been expressed in terms of psychoacoustics, psycholinguistics and psychoptics.

The psychophysical method should be used in close cooperation with neurophysiology and, both together, with sensible modelmaking.

J.F. Schouten
5.a. AUDITORY PERCEPTION

J.L. Goldstein

Auditory Spectral Filtering and Monaural Phase Perception. 1)

Monaural phase perception phenomena reflect the limited resolution of auditory spectrum analysis. Perception is dependent upon phase relations when the frequency separations between spectral components in a coherent sound are sufficiently small.

A detailed study was made of the relationship between various monaural phase phenomena and the properties of the auditory spectrum analyzer. Old and new data were considered for experiments employing carrier tones with sine amplitude modulation (AM) or quasi frequency modulation (QFM); these stimuli are related by a phase transformation. One experiment considered AM, QFM and FM modulation thresholds. A second experiment considered the quality difference between AM and QFM.

A unified psychophysical account for these phenomena is given with a model of auditory spectrum analysis which was derived from classical work. In this model each characteristic frequency in the analyzer consists of a quasi linear bandpass analyzing filter followed by ideal envelope detection and concluding with either a peak-to-minimum or peak threshold decision. Psychophysical formulations with this model provide new insights on the functional importance of the analyzing filter stopband properties and demonstrate that the concept of an ideal rectangular analyzing filter is inadequate.

(A manuscript describing this work has been submitted for publication).

1) This work is based primarily on the author's Ph.D. dissertation, An Investigation of Monaural Phase Perception, University of Rochester, N.Y., 1965 (University Microfilms, Ann Arbor, Michigan, Pub. no. 66-6852).

J.L. Goldstein

Auditory Nonlinearity

Combinations tones (CTs) produced by two-tone stimuli ($f_1$ and $f_2$) at relatively low sound levels contradict the classical view that auditory mechanics is an essentially linear process which suffers significant percentage distortion only at high levels. CT level and phase behavior were measured extensively with pitch cancellation and loudness balancing methods. Relative level of the most prominent CT $2f_1-f_2$ was nearly unaffected by stimulus level but decreased sharply with increasing frequency interval, being typically 15-20% for $f_2/f_1 = 1.10$. 
In contrast, the difference tone was audible only for stimulus levels above 50 dB SL, grew to 15% relative distortion only for estimated stimulus levels exceeding 100 dB SL, and was relatively insensitive to frequency interval.

Other CTs of the form \( f_1-n(f_2-f_1) \) were heard and these decreased sharply in level with increasing integer \( n \). The fact that CTs above the stimulus frequencies were inaudible is not caused by masking, but reflects a mechanical frequency selectivity in the nonlinear source.

Thus although auditory mechanical analysis is not an essentially linear process, the concept that the cochlea performs a limited resolution frequency-place transformation is supported. Physical studies of the cochlea should elucidate the nature of the hypothesized essential cochlear mechanical non-linearity.

(A manuscript describing this work has been prepared for submission for publication).

G. Domburg

The Just Noticeable Difference for Loudness

Introduction

The just noticeable difference for loudness \( \Delta S \) has been measured by many investigators. It has turned out that the loudness level itself is an important parameter contributing to variations in \( \Delta S \).

Various techniques for the assessment of \( \Delta S \) are available. The schedule below presents a rough indication of the more common techniques.

![Various techniques of signal presentation.](image)

Method B was chosen, as it was our aim to study the effect of the duration \( T \) of the stimulus on the just noticeable difference for loudness. We only know of one other instance where \( \Delta S \) was measured as a function of \( T \) (Garner et al., 1944). However, our experiment was not confined to a study of the duration effect.
The effect of loudness was, partly by way of familiarization, also measured. This was, in fact, our first experiment. In a second experiment we investigated the duration effect.

Both sinusoids and noise have been used as stimuli.

**Apparatus**

The apparatus is shown in figure 2 below.

![Block diagram of the apparatus](image)

**fig. 2**: Block diagram of the apparatus

The subject is listening binaurally with PDR 8 headphones in a sound proofed booth. When pressing a button (a) he starts the preset cascade counter (cf. Moonen & Lammers, 1966) (casc. counter) that serves as a programme unit. It opens and closes a gate (P.W.M. gate) twice according to the schema shown in fig. 3.

![The stimulus cycle](image)

**fig. 3**: The stimulus cycle

The silent interval between the two stimuli was 500 msec in all our experiments. The durations of the stimuli were always equal. The loudness of one stimulus was $S$ (dB S.L.) while the other was $S + s$. The random programme of a tape unit (Tape unit) operating on a switching relay (relay) determined whether $S$ preceded $S + s$ or vice versa.

The loudness difference $s$ was controlled by one of the attenuators (Att. I and II).
Procedure

In the first experiment the subjective loudness level \( S \) was found in the following way. First the threshold of the weaker signal was determined by adjustment of attenuator I. Then this attenuator was set at the appropriate level \( S \). Then attenuator II was set so as to make the other stimulus clearly stronger \((S + s)\) to the subject, upon which the trials would begin.

The subject was obliged to tell whether the stronger tone came first or second after each presentation of a stimulus pair. The loudness difference \( s \) was controlled, according to the sequential up-and-down method, in such a way that the 75% correct response level (\( \Delta S \)) was attained in about a hundred trials (cf Lopes Cardozo, 1966).

Results

Two subjects (B.L.C. & G.D.) performed the first experiment with a stimulus duration of 1000 msec and a sinusoid of 1 kHz. Results are presented in figure 4. It is obvious that one subject performed better than the other. The results of the better subject show good agreement to the data presented by Montgomery (1935).

![Figure 4: The j.n.d. for loudness as a function of loudness S. Notice that the vertical dB scale is magnified by a factor 40.](image)

The second experiment was performed by two subjects (T.J. & G.D.) with a sinusoid of frequency 1 kHz, \( S = 30 \) dB S.L. and durations \( T \) ranging from 4 up to 1000 msec. Also a few measurements were made by G.D. with a noise band of 0-4 kHz instead of the sinusoids. These results are given in figure 5.

It is clear that points of one kind can be fitted by a straight line with one notable exception viz. the filled circle at 1024 msec. We do not think that the weight of one single point is sufficient to abandon the linear representation of the data. Moreover, at these very low values of the j.n.d. deficiencies of the apparatus are more likely than at normal values.

The points for noise systematically lie higher than the corresponding points for the sinusoid. The difference between subjects is again striking.
The data given by Garner and Miller (1944) have not been plotted but show reasonable agreement, although their measuring method was different (method A).

The slopes of all the lines in figure 5 are more or less the same. This leads to the general expression

$$\Delta S \cdot T^{0.3} = \text{constant},$$

where the constant depends on the choice of subject, frequency and loudness.

References


Lopes Cardozo, B. 1966 Sequential up-and-down method. This issue.


Moonen, G.J.J., Lammers, C.A. 1966 A Preset Cascade Counter. This issue.
R.J. Ritsma, B.L. Cardozo, G. Domburg and J.J.M. Neelen.

The Build-Up of the Pitch Percept.

Introduction

Intonation (i.e. the behaviour of pitch as a function of time) is an important phenomenon in speech. Cohen and 't Hart (1965) investigated intonation by means of a gating technique and hit upon the problem of how long a segment had to be made in order to give rise to a definite pitch sensation. They found that for their purposes 30 ms was a practical value.

When passing a limited number of pulses through a band pass filter that suppresses the fundamental and a number of its harmonics and listening to this signal over a pair of headphones, one hears a high pitch if only one pulse is presented. In two pulses one sometimes observes something lowish, while normally a low pitch is heard when three or more pulses are presented. All this is subject to a proper choice of the pulse interval and of the pass-band of the filter, both of which should be within the existence region of the tonal residue. (Ritsma, 1962).

In this qualitative way one easily establishes that the pitch percept builds up in about 2 full periods of sound capable of eliciting a residue.

We then tried to get data of a more quantitative nature by the use of matching experiments.

Apparatus and Procedure

Most of our experiments were made with the apparatus presented below in fig. 1.
The subject listens binaurally in a sound-proofed booth. Stimuli are given at 30 dB S.L. by means of Permoflux P D R headphones. The subject starts a preset impulse counter ("Casc. counter" in the diagram) (cf Moenen, 1966) that fires periodic impulses (impulse rate \( g_1 \)) into a filter \( (f_1) \). A gate (P W M) opens and closes in synchronism with the pulses. The duration \( D_1 \) of the stimulus, its repetition rate \( g_1 \) and its filter frequency \( f_1 \) remain fixed during a matching series. Another signal, the matching stimulus, is available to the subject during the time he presses a button (B), provided the first stimulus is not sounding. This second stimulus consists of a periodic pulse with repetition rate \( g_2 \), filterfrequency \( f_2 \), and a duration \( D_2 \) that will, as a rule, be different from \( D_1 \). The subject adjusts \( g_2 \) for a best match of the pitch of the second stimulus to that of the first one.

The relative r.m.s. difference of the pulse repetition rates:

\[
h = \frac{\left( g_1 - g_2 \right)^2}{g_1}
\]

may be taken as a reciprocal indication of the discernibility of pitch.

Results and Conclusion

In one experiment the data reproduced in fig. 2 below were obtained by two trained observers (G.D. and J.J.M. N.). The repetition rates \( g_1 \) were chosen more or less in accordance with values occurring in speech, being:

![Graph showing relative standard deviation as a function of the duration of the pulse trains.](image)

Fig. 2: Relative standard deviation as a function of the duration of the pulse trains.
100 and 533 impulses per second. The notation 100/1000-1400, for example, designates an experiment with $g_1 = 100$ pulses per second, $f_1 = 1000$ Hz and $f_2 = 1400$ Hz.

In a second experiment the same subjects produced the data represented in fig. 3.

![Graph showing relative standard deviation as a function of the duration of the pulse trains.](image)

**Fig. 3:** Relative standard deviation as a function of the duration of the pulse trains.

As a rule, the subjects showed rather similar behavior so that the average values of the two do not mask significant discrepancies.

In both graphs there is a horizontal part for long duration $D_1$ where the relative root mean square deviation of pitch settings $h$ is independent of $D_1$. With short durations, on the other hand, $h$ is strongly dependent upon $D_1$. The transition between both regions is not always a sharp one and its position appears to be influenced by experimental conditions such as the magnitude of the difference of the filter frequencies $f_2 - f_1$. The transition is not found to coincide with a time corresponding to two periods which has been mentioned earlier as a qualitative result. A greater number of periods (6 - 15) is needed for best accuracy of the pitch settings.

Returning to the problem in experimental phonetics mentioned at the beginning, it is clear from our data that an accuracy of one per cent or better can be obtained with pitch matchings using speech segments of 30 ms duration, as the pitch in speech usually is higher than 100 Hz (cf. B. Lopes Cardozo et al., 1965).
It is a well known fact that the ear is especially sensitive for the relation "octave of". A pure tone can be matched to the same frequency as well as to a frequency which is nearly an octave of the first. Both matching situations can be distinguished from each other, as in one case both signals are identical in all their aspects. In matching two different types of periodic complex sounds with respect to pitch an octave setting is quite normal; the subject cannot distinguish whether both signals have the same pitch or differ by an octave. This can easily be demonstrated by singing a tone and then whistling a tone with the same pitch. A difference of one or two octaves between both fundamentals will occur. Only for similar sounds will the timbre aspect give so much information that an octave-difference can be recognized. In this paper a more quantitative examination of the relation "octave of" is given. This is done by measuring the accuracy with which equal settings and octave settings can be made for pure tones and complex sounds.

For a periodic complex sound it could be declared that the pitch is correlated with the periodicity in the sound, whereas the timbre is correlated with the spectral envelope (formant regions) (Schouten, J.F., Ritsma, R.J., Lopes Cardozo, B., 1962). For vowels the quality is in first order determined by the formant regions, whereas the pitch is given by the periodicity in the vowel. To measure the standard deviation, \( \sigma \), in pitch measurements for complex sounds in an isolated form, experiments were done with amplitude modulated sinusoids (AM-signals). For an AM-signal the centre frequency and the periodicity (the modulation frequency) are independent parameters. This implies that an AM-signal consisting of the frequencies 2400-2800-3200 Hz can be matched as to pitch with an AM-signal consisting of the frequencies 1200-1600-2000 Hz. Both signals have the same (residue) pitch (400 Hz); however, they clearly differ in timbre. \( \sigma \) in the match of both signals will be a measure for \( \sigma \) in pitch perceived because there is no possibility of having spectral components coincide (Ritsma, R.J., 1963, 1965). On the strength of this principle \( \sigma \) in pitch detection for this type of signals has been determined for a pitch range from 60 to 600 Hz. The results are illustrated in figure 1 (curve B).
Fig. 1: $\sigma$ for equal settings and for octave settings of pure tones (curve A, • x) and of complex tones (curve B, o +).

For comparison $\sigma$ in frequency of a pure tone is given in curve A.

If a pure tone is matched to an AM-signal with respect to its residue-pitch $\sigma$ in the matchings goes up by a factor 2 - 3. This is a quantitative confirmation of our previous findings, in accordance with others, that pitch matchings between a periodic complex sound and a pure tone are difficult (Schouten, J.F., Ritsma, R.J., Lopes Cardozo, B., 1962).

To confirm the subjective relation "octave of" for an AM-signal, matchings were also made for AM-signals with periodicities which differed by an octave. The subjects made these octave judgments without any difficulty. Judgments were made for the combinations 125-250 Hz; 200-400 Hz; and 250-500 Hz. The results are plotted in Figure 1 as o and +. It turned out that there is no significant difference between $\sigma$ in equal pitch matchings and in octave pitch matchings. From this it follows that the mechanism in the ear responsible for the periodicity detection will be octavedeaf.

For comparison, $\sigma$ in octave judgments of trained subjects with respect to pure tones are also plotted in Figure 1. Closed circles represent the mean values of our subjects; crosses represent the mean values of the subjects taking part on the experiments by Ward (Ward, W.D., 1954). From these results it follows that there is a large difference between the equal settings and the octave settings for pure tones. This is not surprising since in the case of equal settings both signals can be made identical. Secondly it was found, that the octave settings of a pure tone are no longer possible when the octave frequency exceeds about 5000 Hz. This frequency is also the upper limit for the centre frequency of an AM-signal in which one can hear a residue-pitch (Ritsma, R.J., 1962).

It is unlikely that two mechanisms, "place" detection and "periodicity" detection, based on different working principles both possess the same property of octave sensitivity.
One must decide that a pure tone is not only detected on the principle of place detection but also on the principle of periodicity detection. In the case of octave settings of pure tones, only periodicity detection operates, while in the case of equal tone settings, place detection predominates.

References


B.L. Cardozo, R.J. Ritsma, G. Domburg & J.J.M. Neelen

Unipolar Pulse Trains with Perturbed Intervals: Perceptibility of Jitter.

Introduction

It is a well-known fact, that consecutive pitch periods in speech sounds are not exactly equal. In the first place the intonation movements cause the pitch period to vary from one instant to another. Moreover, the involuntary irregularities due to imperfections of the vocal cords cause the pitch period to fluctuate in a more or less random manner. It is known that gross imperfections of the vocal cords may lead to an audible hoarseness of the voice (Ph. Lieberman, 1962).

This led us to question what magnitude of perturbation of the pitch period is at all perceptible to the human ear. Various approaches are possible. We shall report here on three experiments.

First Experiment: The just noticeable delay of one single pulse.

In the first approach a single, non-random perturbation was considered. The apparatus is schematised in figure 1. The subject is listening binaurally with PDR8 headphones while sitting in a sound proofed booth. When pressing the start button he brings a preset impulse generator (casc. counter II, cf. Moonen et al. 1966) into action, thereby starting the programme for the presentation of one stimulus pair. One of these stimuli is the unperturbed pulse train consisting of 8 equally spaced pulses, which are fired by another preset impulse generator (casc. counter I). These pulses are 100 microseconds wide. They are filtered by means of a band pass filter with central frequency f, bandwidth 0.33 f and slopes of 34 dB per octave. The other stimulus is identical to the one just mentioned except that one of the eight impulses is delayed by a time Τ milli-
seconds. A tape unit with a random programme determines whether the perturbed stimulus precedes or follows the unperturbed one.

The stimuli used were JJJ/2800 (i.e. repetition rate JJJ impulses per second, filter centre frequency 2800 Hz), 100/2800 and 100/1000. In each case the just noticeable delay for each of the 8 pulse positions was determined.

Procedure. Changes in the delay were made in steps of a factor 2. Along this logarithmic delay scale the sequential up-and-down method was used for bracketing the 72% correct response level (cf B.L. Cardozo, 1966).

The results of this first experiment are presented in Figure 2. Two subjects (G.D. and J.J.M. N.) obtained data that were essentially similar. Their logarithmic means have been plotted.

When discussing these results the following points are to be made:

1. In the 100/1000 stimulus, which is atonal, the just noticeable delay is of the order of one millisecond; it is generally and significantly greater than in both other stimuli, which do have tonal quality (pitch). (cf Ritsma, 1962).
2. The central impulses show just noticeable delays that seem to be independent of their position in the pulse series. For the tonal stimuli the just noticeable delay is about 1.5 per cent of the pitch period.
3. The leading pulses in the total stimuli show just noticeable delays of about 10 per cent of the pitch period. This is considerably more than would be expected on the ground that a delay of a boundary pulse affects only one period whereas a central one impairs two periods.
4. Notice that durations were very short.
Second Experiment: Just Noticeable Jitter in Unipolar Pulse Trains.

To the first experiment one may object that it was limited to stimuli of very short duration. Therefore this second experiment was devised to yield some information on the perceptibility of perturbations of the periodicity in longer pulse trains.

The apparatus for bringing about the perturbations is schematized in Figure 3a, while Figure 3b shows how in a number of stages the original mixture of a sinusoid and a noise is transformed into a train of jittered impulses. We used unipolar pulses of 100 microseconds width, filtered them afterwards so as to get some similarity (be it a vague one) to the vowel sounds in speech and passed them through a gate that cut out segments of the required time duration. The rest of the apparatus was similar to that of the first experiment. Also the subjects and the method for determining the j.n.d. were the same as in the first experiment.
Fig. 3a: Circuit used for the production of pulse trains with random gaussian jitter.

Fig. 3b: Consecutive stages in the formation of jittered pulse train. cf figure 3a.
Fig. 4: Computed probability distributions of the moment of the first rising zero of a mixture of a sinusoid and noise.

The quasi-gaussian character of the jitter can be arrived at by computation (cf. figure 4), with the proviso that the signal-to-noise ratio in the original signal be sufficiently large. An experimental verification is exemplified in Figure 5 showing a histogram of 600 interpulse intervals.

Fig. 5: The measured histogram based on 600 individually determined pulse intervals.

Figure 6 shows the "relative jitter" as a function of the signal-to-noise ratio, which was the parameter that was varied in the experiment for bracketing the just noticeable jitter. Jitter in this experiment is defined as the standard
6: Experimental values of relative jitter $\tau/T$ as a function of the signal-to-noise ratio in the sinusoid-noise mixture at the input of the jitter circuit.

Fig. 7: Just noticeable jitter as a function of duration
deviation of the interpulse-interval, being \( \sqrt{2} \) times the r.m.s. timing error of individual pulses.

Pooled results for both subjects are shown in Figure 7. The notation 100/1400 again designates a 100 p.p.s. pulse train, filtered at 1400 Hz (same filter as experiment one).

A comparison with the first experiment shows just noticeable jitter in experiment II to be smaller by a factor 2 roughly, taking the durations into account.

With short durations the just noticeable jitter rises slightly more steeply than inversely proportional to duration. With long durations the just noticeable jitter gradually levels off. The transition is not very sharp.

Discussion

1. The discrepancy of a factor 2 between both types of experiment seems to be slightly too large to be attributed to statistical effects. A possible way of bridging the gap would be to suppose that the hearing mechanism bases the discrimination on the largest perturbation that occurred. This hypothesis would indeed reduce the factor 2 to something of the order of 1.4. However, this hypothesis is untenable since it predicts a slope for the left hand parts in Figure 7 that is much flatter than the experimental one. Somehow, then, the ear seems to integrate over a greater number of perturbations.

2. The asymptotic values of the just noticeable jitter for long duration are rather widely different for 100 pulses per second and for 333 pulses per second. It is reasonable to attribute at least part of this to the former signal being effectively extinguished at the end of a 10 msec interpulse-interval while the latter signal exhibits a discontinuity in the waveform itself that wanders to and fro as a result of the jitter.

3. If one accepts 50 microseconds speculatively as a measure for the timing accuracy of the pitch extractor, then this would have a consequence for the accuracy with which jittered pulse trains can be matched. An experiment along this line is under consideration.

4. A comparison of our results with those of A.E. Rosenberg (1966) is hardly possible. Rosenberg uses jitter as an expedient for smoothing the medium and high frequency part of the spectrum. He finds that a relative jitter of 5 to 25 per cent does the job of rendering unipolar and alternating pulse trains indistinguishable. Our experiment starts with two identical pulse trains and, by introducing to one of them a relative jitter of 0.15 to 0.7 per cent, makes them discriminable.

5. When applying our data to phonetics one has to keep in mind that, owing to intonation, the speech signal seldom has constant periodicity. Only during comparatively short times is there a more or less invariant pitch period. Thus the shorter pulse trains are probably more indicative of the order of magnitude at which perturbations in periodicity become manifest in speech. As an extra caution it should be remarked that hoarseness, which is the typical phonetic aspect, is not heard in the very faint (but just noticeable) jitter phenomena of the psycho-acoustic experiment.

Third Experiment: Just Noticeable Jitter in Unipolar Pulse Trains with Sweeping Pitch

This an extension of the second experiment. We tried to bring "intonation" into the stimulus and economized on some of the other parameters.
In producing the stimulus we substituted a saw-tooth function for the original sinusoid. The saw-tooth was made to vary its frequency linearly with time. A sweep over one octave was studied as well as a sweep over two and a third octave. These sweeps were larger than most intonational pitch changes encountered in speech. Sweep rates covered a range of a factor 10 and were moderately to extremely fast, taking normal intonation patterns in speech as a standard.

A consequence of the pulse rate being a linear function of time is that the pitch interval T was a hyperbolic function of time. Hence the relative increments of the pitch period were proportional to this period squared.

The jitter was checked to be random and gaussian like in experiment II. Jitter was applied in a way that maintained a constant ratio between the root-mean-square perturbation of the pitch interval and the instantaneous value of this interval. Hence the relative jitter was invariant during a pitch sweep. It was varied from trial to trial in order to bracket the just noticeable value under a range of stimulus parameters.

These parameters were:

\[ f = \text{filter frequency, viz. 1000 Hz, 2000 Hz, 2800 Hz and 4000 Hz.} \]

\[ D = \text{duration, viz. 100 ms, 250 ms and 500 ms.} \]

\[ \text{Sweep range, viz. from 400 p.p.s. down to 200 p.p.s. and from 500 p.p.s. down to 100 p.p.s.} \]

The method for determining the discriminability threshold was the same as in the previous experiments and so were the subjects.

Results are summarized in Figures 8, 9, 10 and 11. Figure 8 is meant to compare results of the present experiment with those of experiment II.

---

**Fig. 8:** Just noticeable jitter percentage as a function of duration. Plotted circles stem from experiment III. Lines have been added from experiment II for the purpose of comparison.
The six circles represent mean values for the two subjects with a sweep from 400 p.p.s. to 200 p.p.s., so that the mean pitch period is close to 3 milliseconds. The closed circles all lie above the full line representing the just noticeable jitter percentage in stationary pulse trains with a period of 3 milliseconds and a filter frequency of 2800 Hz. The same trend is seen at a filter frequency 1400 Hz. The faster sweep rates (500 to 100 p.p.s., not shown in Fig. 8) exhibit the same phenomenon to a higher degree. Our first conclusion, therefore, is that changing pitch makes it more difficult to perceive jitter but not very much so.

Fig. 9: Just noticeable relative jitter as a function of filter frequency. Each point is an average over two subjects and three durations.

Figure 9 shows that the just noticeable percentage jitter increases exponentially as the filter frequency rises. Two remarks are in order here.

In the first place the bandwidth of the filter is proportional to \( f \), so that a higher frequency goes with a smaller time constant. The second point of the discussion of experiment II is applicable here.

In the second place it is a legitimate question whether the jitter discrimination is made on the basis of certain spectral properties or whether it is the time function that reveals the jitter. If it were the spectral properties, then one would not expect a sharp rise with the filter frequency, because the jitter affects the high frequency parts of the spectrum more than it does the low frequency regions. So there is reason to suppose that the time function is responsible for the weakest jitter perception. It should be restated here that the hoarse or husky sound quality comes in at stronger jitter. It is quite possible that hoarseness does rest on some spectral property. We have no data on this.

Figure 10 depicts how jitter is more easily perceived as the duration \( D \) is increased. This is seemingly at variance with the results of experiment II (cf. figure 7), where the effect of \( D \) is almost negligible beyond 0.1 second. In the present experiment, however, the duration affects the sweep rates and consequently the systematic increments of the pitch interval. No two pitch intervals are equal and it is the task of the hearing organ to decide whether the increments are due to the intonation or to the jitter.
Fig. 10: Just noticeable relative jitter as a function of duration. Each point is an average over two subjects and four filter frequencies.

It appears reasonable to assume that this task is easiest when the systematic relative increments are smallest. For this reason we plotted the just noticeable relative jitter as a function of the minimal relative increment occurring in each type of stimulus. Figure 11 demonstrates that the just noticeable relative jitter is of the same order of magnitude as the minimal relative increment of the pitch interval.

Fig. 11: Just noticeable jitter as a function of minimal relative increment. Each point is an average over two subjects and four filter frequencies.
Conclusion

Three experiments on the perceptibility of perturbations of the pitch period have been described. The method, the criterion, the subjects were the same, but the stimuli were different.

We found that perturbations in short tonal stimuli are "just noticeable" when they are of the order of a per cent. Perturbations between one per cent and one part in a thousand are just noticeable in stimuli longer than about 0.1 sec. In stimuli with changing pitch the just noticeable relative jitter appears to be of about the same size as the smallest relative changes in the pitch interval.

References

Lopes Cardozo, B. 1966 The Sequential Up-and-down Method. This issue.
Moonen, G.J.J., Lammers, C.A. 1966 A Preset Cascade Counter. This issue.

The Pitch of Sinusoids and Complex Signals as Affected by Masking Noise

R.J. Ritsma 2)

Experiments were carried out to establish the effect of a change in the pitch by masking a tone with various bands of white noise. Presence of white noise generally raises the pitch of a pure tone whose frequency is between 500 and 4000 Hz. The effect is more pronounced at low loudness levels and is progressively larger as the frequency of the tone rises. (cf. figure 1). Applying filtered noise and pure tones of low frequency the effect of pitch rise becomes larger when the high cut-off frequency of the masking noise abuts on the frequency of tone.

For a complex tone no change in pitch was found during masking with various bands of white noise.

From this it may be concluded that the pitch mechanism for a pure tone differs from the pitch mechanism for a complex tone.

For pure tones pitch discrimination is mainly defined by the place of maximum stimulation along the basilar membrane, whereas for complex tones a mechanism is in operation to determine the periodicity in the signal.
Introduction

Directional localisation of sounds depends primarily on binaural hearing. It is an important factor in situations, occurring so often in everyday life, where one has to concentrate on one speech message while simultaneously other, competing messages are also audible. Apart from the binaural effect there is an important monaural directivity pattern resulting mainly from the sound shadow of the head.

Fig. 1: Pitch shift of a sinusoid resulting from the addition of partly masking white noise. The frequency of the sinusoid is plotted along the abscissa.


B.L. Cardozo, Th.A. de Jung & R.J.J. Boddeke

Improvement of Intelligibility by the Application of Directional Microphones in a Simulated Hearing Aid Situation
It is the purpose of the present report to investigate whether this 'non-binaural' directivity could be improved by introducing a directional microphone into a hearing aid. For this reason intelligibility measurements have been made in order to compare omnidirectional and directional microphones and the potential use of the latter for the hard of hearing.

Method

Subjects were presented monaurally with a mixture of two competing signals. One signal was a series of test sentences, the other signal was a voice babble. The loudness ratio of the two signals was varied with an attenuator so as to cover the range between test sentences being well intelligible and test sentences being effectively masked by the voice babble. Both signals were played back from magnetic tape. They had been recorded separately. In some of these recordings the hearing aid situation was simulated to some extent.

In a first set of recordings a person was seated in an anechoic room. A microphone was mounted against the pinna of one of his ears. This microphone was facing forward. At a distance of 1.5 m from the person's head a loudspeaker emitted a signal. When this was the test signal, then the loudspeaker was exactly in front of the person. When the babble signal was being emitted, then the loudspeaker was in one of four directions (at 0, 90, 180 or 270 degrees with respect to the forward direction of the person). This person was entirely passive; his only contribution to the experiment was the shadow of his head.

A second set of recordings were made in a reverberation room (roughly 5 m by 5 m, non-parallel concrete walls, reverberation time 8 seconds at 1 kHz) that produced a more or less isotropic sound field. In this set, recordings were made without using a head as a microphone stand. The loudspeaker always was opposite to the microphone.

The babble consisted of a mixture of four series of sentences, which were, in fact, the original test sentences played four times over incoherently. Sinusoids were recorded with each signal for calibration purposes. Thus a test-to-babble ratio 0 dB means that the test signal as measured at the loudspeaker had the same volume as one of the four messages constituting the babble signal.

All of the above recordings were made twice: one time with an omnidirectional microphone (Brüel & Kjaer, type number 4132, condensor microphone) and a second time with a directional one (a Philips EL6091, electrodynamic microphone).

5 Subjects, students of the Technische Hogeschool, Eindhoven, all possessing normal hearing performed the intelligibility tests.

Results

The results are depicted in figure 1. Points represent averages over 5 subjects. Figure 1 shows that in the 0 degree situation the omnidirectional microphone seems to have a slight advantage over the directional one. The possibility that this is due to the overall qualities of this condensor microphone can not be disregarded. However, in the 180 degree situation the directional microphone has an advantage over the omnidirectional one. Also in the perpendicular situations the directional microphone shows an advantage, though but a very small one when the loudspeaker is at the same side as the microphone.
Fig. 1: Intelligibility as a function of the test-to-babble ratio. This ratio is equal to 0 dB when the volume of the test sentences was equal to the volume of one of four messages constituting the babble. Intelligibility was scored by computing the percentage of significant words repeated correctly. The simulated listening situations in the anechoic room are symbolized above. In the reverberation room only one geometrical situation was used.

In the reverberation situation hardly any babble is tolerable. In fact, with the omnidirectional microphone the best intelligibility is but slightly over 50 per cent, with virtually no babble added. Here again the directional microphone shows its usefulness attaining nearly 100 per cent in the best condition. In conditions of somewhat poorer test-to-babble ratio the advantage is still some 4 dB.

Conclusions

Attention has been called for the improvement in intelligibility that can be obtained by substituting an omnidirectional microphone by a directional one (pointing forward). It is some 18 dB in the most favourable geometrical si-
Rustle Noise

General

In acoustics the different types of noise are usually distinguished with respect to their spectral envelope. In obvious analogy with visual colours white noise is defined by a constant spectral density in the frequency range considered. Similarly, coloured noise is defined by a non-uniform density function of frequency. Coloured noise can be obtained by passing white noise through a band pass filter.

It struck us that in describing many of the noises encountered in nature or in the laboratory a time parameter $\theta$ should be introduced in addition to the usual colour aspect.

Take for instance the ticks of a Geiger counter responding to a radioactive source. The ticks occur at random and the probability density of an interval $t$ between two successive ticks is given by the zero order Poisson distribution

$$q(t) = \frac{1}{\theta^t}$$

in which $\theta$ is the mean interval.

If the ticks are sufficiently narrow this Geiger noise is "white" independent of the value of $\theta$.

Yet when decreasing $\theta$ from say one second to below one millisecond the auditory impression will change from ticking through hiss and hiss to a more or less constant hiss which is indistinguishable from acoustic white noise proper.

Evidently, the properties of our auditory system come into play, in that the ear can follow phenomena sufficiently spaced in time, but fails to do so above a rate characteristic of that system.

We therefore introduce the concept of rustle noise. Rustle noise consists of random pulses characterized by the rustle time $\theta$ (the mean interval) and by their shape. The latter determines the colour. Normal noise is rustle noise with a rustle time $\theta$ lower than some critical value characteristic of the measuring system considered.
Rustle noise is readily obtained from a driving gaussian white noise in
triggering a pulse generator by the amplitude peaks surpassing a pre-set value
$y$. The higher this level $y$ the higher the rustle time $\Theta$. Colour is then intro-
duced by filtering.

Mathematical

Let the driving noise have the amplitude distribution $g(x)$. The probability
of an amplitude surpassing $y$ (fig. 1) is given by

$$p(y) = \int g(x) \, dx$$

(2)

Some interesting properties of rustle noise can be obtained in terms of $p$,
irrespective of the choice of the amplitude distribution.

We divide the time axis in segments $\Upsilon$ such that there is no correlation be-
tween the amplitude of the driving noise in neighbouring segments.

The output of the rustle generator is then a sequence of pulses (+) or non-
pulses (-) like:

\[ - \quad - \quad + \quad - \quad - \quad + \quad - \quad - \quad + \quad - \quad + \quad + \]

\[ \text{cycle} \quad \text{cycle} \quad \text{cycle} \]

On the average there will be a fraction of $p$ plusses and $(1-p)$ minuses.

The probability of a plus series having a length $k$ is given by

$$P(k) = p^{k-1}(1-p)$$

(3)

which is an exponentially decreasing function of $k$. The mean length of a plus
series is

$$\bar{k} = \frac{1}{1-p}$$

(4)

Similarly the mean length of a minus series is

$$\bar{1} = \frac{1}{p}$$

(5)

and the mean length $n$ of a "cycle"

$$\frac{1}{n} = \frac{1}{\bar{k}+\bar{1}} = \frac{1}{p(1-p)}$$

(6)
Fig. 2: Sonagrams of rustle noise
Three recordings for $\theta = 10.2$ and $0.5$ msec. Occasionally intervals considerably larger than $\theta$ occur.
The latter value can be immediately obtained from the probability of a positive zero crossing (\(+\)) which is obviously equal to \((1-p)p\).

For large values of the cyclelength \(n\) can be proved to approach an exponential function like \((1)\).

The rustle time is given by

\[ \Theta = \frac{n \tau}{p(1-p)} \]  

\((7)\)

Experimental

The minimum discernible \(\Theta\)

Rustle noise with a rustle time as small as \(\Theta = 1\) msec. is still distinguishable from the hiss of normal noise. This need not be so surprising, since the slight irregularities discernible in comparison with constant hiss are caused by incidental long intervals (see fig. 2). At \(\Theta = 0.3\) msec. white rustle noise is indiscernible from white noise proper.

The rustle diagram for verbs

\(\Theta\) was varied between 0.1 msec and 1 sec. The centre frequency of a two octave bandpass filter was varied between 200 and 6400 c/sec. Subjects were asked to name the sound in terms of known verbs. The diagrams for English and Dutch are reproduced in fig. 3.

---

**Fig. 3: Rustle Diagrams for English and Dutch verbs**

Each verb is plotted at the location of its rustle time \(\Theta\) and average frequency band \(f\).

Note the extra syllable in the region of \(\Theta = 10\) msec, the plosive consonants for isolated events (\(\Theta \geq 100\) msec) the fricative consonants for continuous events (\(\Theta < 1\) msec) and the systematic change both in vowels and in consonants along the frequency axis.

It is evident to the listener that the rustle noise with its two parameters, in many cases, is but a poor imitation of the name given. Even so the general trend of the location of the verbs in the rustle diagram seems fairly warranted.

It is interesting to note the onomatopoeic character of the verbs denoting noises of various character.
For the English verbs we point to the following general trends:

1. Along the vertical rustle time axis:
   - For \( \theta \) above 100 ms: verbs of one syllable containing plosive consonants indicating isolated events.
   - For \( \theta \) between 2 and 100 ms: verbs of two syllables indicating repetitive events.
   - For \( \theta \) below 2 ms: verbs of one syllable containing fricative consonants indicating a continuous event.

2. Along the horizontal frequency axis:
   - With increasing \( f \) the vowels change in accordance with their increasing formant frequencies.
   - The consonants tend to change from \( b, d, m \) to \( t, k, s, f, z \) and \( sh \).

For the Dutch verbs similar trends can be observed. The verbs have one more syllable than the corresponding English ones.

J.F. Schouten & G. Domburg

Pitch and Anticipation

When listening to two subsequent short tone pips of 1000 Hz with a duration of some 4 - 8 ms it struck us that the observed pitches are not necessarily equal. On the contrary, depending upon our anticipation, the second pip could be heard higher, equal or lower at will. The phenomenon was demonstrated to many naive observers who, when asked: "Can you hear the second pip at a lower pitch than the first?", would answer in the affirmative but equally when asked the opposite question.

In a systematic experiment we presented two tones separated by 200 ms as a pair. The interval between pairs was 2 sec. The duration of the pips was \( \tau = 2, 4, 8, 16, 32 \) and 128 ms. The frequency of the first pip was kept constant at \( f_1 = 1000 \) Hz. The frequency of the second pip was chosen equal or higher than the first. The frequency rise \( \Delta f = f_2 - f_1 \) was set at \( \Delta f = 0, 5, 10, 15, 20, 35, 50, 70 \) and 100 Hz. For each setting of \( \tau \) and \( \Delta f \) the listener was required to determine how many times out of 10 subsequent presentations he was able to hear the pitch change as requested and anticipated.

Both authors acted as listeners. Initially, the task given was more difficult for the second author. This is probably due to the fact that he has extensive prior experience in "unbiased" listening to tone pips. Yet, the results of both listeners were in surprising agreement.

Experiment I. "Higher Pitch?"

The pitch of the second pip was heard higher in 99% of all presentations, including \( \Delta f = 0 \).

Experiment II. "Equal Pitch?" (see fig. 1)

It amazed both listeners that equality of pitch can be heard for values of \( \Delta f \) and \( \tau \) which, to the unbiased ear, are easily and consistently heard as a jump upwards in pitch. For the second author the standard deviation in pitch
Fig. 1: Frequency Changes and Pitch Changes

Black bars: frequencies. The second is higher than the first.
Striated bars: pitches.
Above: The pitches are heard equal at some value intermediate between $f_1$ and $f_2$.
Below: The pitch is heard to drop. The observed interval is even larger than the interval of the increase of frequency.

discrimination in unbiased condition was determined to be about $\Delta f$. $T = 50$ if $T$ is expressed in milliseconds.

In the present experiment, for each $T$ the range of $\Delta f$'s for which equality was observed in more than 50% of the presentations extended as far as two to three times this unbiased standard deviation.

Experiment III. "Lower Pitch?" (see fig. 1)

Again it was amazing how large $\Delta f$ may be taken whilst yet hearing the anticipated decline in pitch. Curiously enough the 50% borderline is no smaller than the one obtained for equality of pitch. Both listeners had the impression that their anticipation worked by imagining a pitch pair, both absolute and relative, such that both pips would fulfil anticipation.

Experiment IV. "What pitch interval do you hear?"

This experiment was carried out by the first author only. He has no faculty of absolute pitch whatsoever but prides himself in having a rather keen ear for pitch intervals.

In previous experiments it struck him time and again that he always hears small frequency intervals, say below 2%, as half a tone (6%) in pitch or a quarter of a tone (3%) at the very best. In the present experiment with short tone pips pitch intervals between half a tone (6%) and a fourth (13%) were heard, both and in equal measure when anticipating a rise or a fall in pitch. The results can be summarized by the empirical rule that the observed intervals of pitch are about three times as large as the intervals of frequency.
Discussion

It is found that by mere anticipation the pitch of short tone pips can be heard higher or lower at will. The range over which this phenomenon occurs exceeds the standard deviation of unbiased pitch discrimination by factor two to three. The pitch interval anticipated and heard is far larger than the objective frequency interval. For a pitch change of a sign equal to that of the frequency change this would mean a mere exaggeration in subjective scaling. For a pitch change of a sign opposite to that of the frequency change this is quite a baffling result. It would be interesting to find similar effects of anticipation upon other subjective evaluations like loudness, brightness, colour, etc. Hitherto we were unable to find any such analogous effect. The obvious question in measurements involving judgments of pitch is whether the "unbiased" observer ever is as unbiased as one would like him to be.

The experience of the second author, who has proved to be a very keen observer in unbiased conditions, seems to warrant that the anticipation, though evidently dangerous in principle, need not disturb the well-trained observer, from whom an unbiased but keen listening is required.

Perception: By speech perception we understand the interpretation of incoming signals in terms of a mapping, a plan, with which we have learned to operate.

Mapping: By the term 'mapping' we understand the linguistic rules structuring speech into linguistic signs, and, in a wider sense, the structuring of the outside world by linguistic signs so as to provide the language users with a means to communicate with each other.

Learning: In the process of learning the mother tongue, the child first uses its capacity to generate any amount of sounds some of which will be strongly reinforced by people in the environment; thanks to its capacity to hear it will gradually learn to make approximations to the sounds it hears in an effort to imitate others. Next, words are acquired as acoustic shapes through which a certain meaning transpires when embedded in a specific concrete situation. Subsequently, such forms are handled as linguistic signs in linguistic contexts according to the rules of the language. Once they have been firmly established in use, they can open the way to accommodate new words which are introduced as meaningless forms through which the meaning once more has to percolate through use. This acquisition of new words can be seen as a very active process since the child is now capable of using this means of getting to grips with the world both outside and within himself.

This whole process takes a number of years filled with prolonged and daily exercises. Once a certain proficiency has been achieved it should cause no surprise that the knowledge acquired can be put to secondary uses, isolated and derived from the primary use in speech communication.

Acoustic cues

The way has now been cleared to lay the acoustic form open to analytical treatment by language users. It is possible linguistically to establish the inventory of e.g. vowel phonemes in a particular language. On this basis we may assume that the users of this language operate with a built-in pattern as a projection of the acoustic vowel space which enables them to distinguish and produce the vowels of their language. It now turns out to be possible to make this pattern explicit by presenting subjects with a large number of acoustic shapes, consisting of F1/F2-combinations and asking them to judge whether these are acceptable or not as vowels of their language. Not only will they be able to give verdicts on the colour perceived but also on a variety of durations. We can represent these findings graphically by plotting on a F1/F2-field the various synthetic sounds that were presented and the interpretation in terms of vowel judgment of the language concerned. If the same stimuli are used in experiments involving users of a different language the mapping will lead to significant differences. Thus along the /i/-/y/-/u/ line a Japanese listener, e.g., will accept more /i/-judgments and after a cloudy stretch will start by coming up with /u/-judgments.
Distributional cues

Evidence of the possibility of interpreting correctly meaningless acoustic shapes which are nevertheless based on the patterns of a particular language can be found in the so-called PB lists. They actually represent a mixture acoustic and distributional cues since they are composed so as to accommodate all the phonemes of a language with due regard to the statistical rules governing their use in actual speech and to contextual constraints.

So consonants that normally occur only prevocally will not be included in the material postvocally. We once carried out a very simple experiment in which we asked subjects to complete tape recorded CV-combinations. Not only were the subjects able to perform this task, but their reactions quite faithfully reflected the statistical properties of the vocabulary; they found it easier to supply words starting with /z/ and /v/ than with /s/ and /f/; some of the vowels were more amenable to the completion task than others.

This may be taken as an indication that frequent combinations are stacked away so as to be more readily available, when necessary, than less frequent ones.

Another instance of the facility the language user has at his disposal with respect to mere acoustic forms divested of meaning, but nevertheless occurring as parts of larger meaningful wholes, is provided by his ability to offer rhyme words or alliterations.

A study of lapses, errors in speaking, shows that some phonemes may easily get displaced at the expense of others. Vowels may get in the way of other vowels, consonants of other consonants but also certain clusters obtrude often at the expense of other clusters or indeed of single consonants.

Numerous instances can be quoted to effect that in exposure to a foreign language the acoustic shapes will be interpreted in terms of the known structures resulting in errors of perception which show that the distributional cues may prevail over mere acoustic ones.

Thus a combination of Dutch /s/ and /x/ initially, in Scheveningen presents an awkward problem to e.g. German speakers, who know both phonemes from their own language, but not in this order.

So far we have not been able to adduce evidence to the effect that the so-called distinctive features as such play a large part in perception. It seems attractive to try to specify by what phonemic characteristics language users are able to identify and distinguish the linguistic signs, such as words, of their language. But as we have attempted to prove, an explanation couched in terms of acoustic cues only will hardly account for the observed behaviour of language users in the situations cited. Even if we stick for a moment to a consideration of vowel phonemes, which can largely be sufficiently characterised by acoustic cues, there is more to them than mere 'otherness' to recognise them: even if the same amount of distinctions are made, i.e., the same number of divisions, of the vocalic acoustic space, a mapping that is slightly shifted with respect to one's own, will readily be interpreted as dialectal deviation.

With respect to consonants one may observe that for instance liquids and nasals always receive some support from adjoining speech sounds, whereas in the distinctive feature approach it is taken for granted that they can be fully characterised, as it were, in isolation.

Yet, since these phonemes always occur in conjunction with vowels, one need not be surprised if recognition on acoustic cues only, may easily fail. Thus /m/ and /n/ are often confused in Dutch in spite of meaning and situational cues.

To sum up: phoneme recognition is made possible through a long process of training in and making use of phonemic combinations in meaningful wholes,
words, that have received the stamp of approval by the linguistic conventions of a specific language.

The interpretation of incoming sound waves is indeed a multidimensional operation in which the phonemes can be assumed to play the part of perceptual entities by means of acoustic and distributional cues that can be studied experimentally and used as a means to learn more about the perception of speech as well as to test the adequacy of a particular phonemic analysis.

   Phonetica, 2(1963), 65-78.

2) van Katwijk, A., On Perceptual Units in Speech. 
   This issue.

3) Slis, I.H., Vergelijking van Taalgebruik en Taalstatistiek. 
   I.P.O. Minor Report No. 46.

4) Cohen, A., Errors of Speech and their Implication for understanding the Strategy of Language Users. 
   This issue.

I.H. Slis

A Model for the Distinction between Voiceless and Voiced Consonants

1. Introduction

In the relevant literature a great number of features can be found to contribute to the distinction between voiceless and voiced consonants, as: moment of voice onset, duration of the consonant, duration of the previous vowel, intensity of the consonant noise, intensity of adjacent vowels, presence of voice during the consonant, intonation-pattern during adjacent vowels, vowel-formant-transitions, intra-oral-pressure, air-flow out of the mouth, volume of the lower pharynx, position of the larynx, state of the vocal cords.

This vast amount of features seems to stand in no relation to the simple binary features as used in phonological descriptions. This incompatibility in a way is a challenge to try to reduce the number of relevant parameters as being due to one origin. In this report a tentative articulatory model is proposed accounting for the difference voiced-unvoiced, by showing interactions of the aspects mentioned above.

By means of the model we hope to get a better insight in the connection between the various acoustic distinctions. A number of experiments, carried out to test a few hypotheses experimentally, will be described.
2. The articulatory model

We think a possible indication might be found by assuming that the difference voiceless-voiced is originated by the state of tension of the pharyngeal constrictors. From this state of tension all the relevant articulatory and acoustic features will be derived.

The middle pharyngeal constrictor arises at the major and minor horns of the hyoid bone and the inferior pharyngeal constrictor originates from the sides of the cricoid and thyroid cartilages. Therefore, it is likely that the position of the larynx will be influenced by this state of tension. The condition of the vocal cords will in their turn be influenced by the situation of the larynx.

First the articulatory situations will be described corresponding to the production of voiceless consonants, next to voiced consonants.

3. Voiceless consonants

Suppose the voiceless character of a consonant embedded in vowels (V₁CV₂) originates in a high tension of the pharyngeal constrictors, then the following process would take place. During the preceding vowel (V₁) the throat wall begins to tighten; the larynx will be pulled up by these muscles and rotates slightly during this movement around an axis perpendicular on the plane of symmetry. By this rotation the vocal cords are stretched and consequently acquire a less favourable condition for effective vibration, like the falsetto voice of untrained singers; the vocal cords will be tense and slightly apart. Because of this the pitch rises and the intensity falls.

At the end of V₁ the mouth starts its closing movement, which results in a gradually growing resistance for the air flow; the air pressure in the mouth-throat cavity will rise. The volume of the throat will be small because of the rigid wall and the high position of the larynx, and the resistance of the glottis low because of the slight abduction of the vocal cords. As a consequence the build-up of the intra-oral pressure will be quick. As the condition for vibration of the vocal cords is poor, they will stop vibrating when the pressure drop across the glottis gets lower. It will stop even before the mouth finishes its closing movement. Supposing the vowel formant transitions are an acoustic reflection of the movement of the articulators, then the transitions will only be partly finished at the moment the voice stops. The acoustic result will be: a short formant-transition over a small frequency range.

With the fricatives a strong airflow will pass out of the small mouth opening, because of the high intra-oral pressure. This results in a noise segment of high intensity. With the plosives the mouth closes entirely which results in a completely soundless interval. At the moment the mouth opens a strong airflow will pass out and a powerful noise burst results. With fricatives as well as plosives the air pressure in the mouth decreases and the pressure drop across the glottis rises as the opening movement of the mouth goes on.

The vocal cords will start vibrating again as soon as the air pressure difference under and above the vocal cords is high enough. As the pharyngeal constrictors did not yet relax, the vocal cords still are in a poor vibration condition. Consequently the pitch will be high and the intensity of the vowel will be low. The moment of voice onset will be late and the formant transitions will be short and move over a small frequency range as could also be observed in the preceding vowel. During the vowel the muscles relax and the pitch drops gradually.
Because of an early moment of voice offset, a late moment of voice onset and a high intra-oral pressure obtained in a short time, the fricatives will have a long duration, the voiceless pause with plosives will be long and the noise burst with plosives will be long.

4. Voiced consonants

Unlike the situation sketched above, the pharyngeal constrictors will be lax with voiced plosives and fricatives.

We might expect the following events:

The larynx will stay in its low position and consequently the vocal cords will not be stretched. (If at rest, the vocal cords would lie loosely together). The condition for vibration is very good, like the chest voice of untrained singers, a small airflow will be sufficient to keep the vocal cords in vibration; because of this the pitch will be low and the intensity high.

At the end of the vowel when the mouth closes the intra-oral pressure rises. This rise will be slower than with the unvoiced consonants, because the throat wall is lax and consequently stretches like a balloon, and the larynx stays in its low position; the volume of the mouth-throat cavity will be big. Moreover, the resistance in the closed glottis is high.

At the moment the closing movement stops, the pressure in the mouth is not nearly equal to the subglottal pressure, the vocal cords will not stop vibrating and the vowel formant transitions will be acoustically realised. This results in long, extensive formant-transitions.

With the fricatives a small airflow will leave the mouth because of the low pressure in the mouth. The fricative noise might even be interrupted because the pressure will be built up pulse wise (during the open phase of the vocal cords), between these pulses the air leaks out of the mouth and the pressure drops below the "friction threshold". With the voiced plosives no air flows out of the mouth, nevertheless the maximal air pressure (equal to the subglottal pressure) will not be reached. Therefore vibration of the vocal cords goes on during the closed period.

The moment the mouth opens, the noise of the fricatives stops because the pressure in the mouth drops quickly. With the plosives a short noise burst occurs to discharge the overpressure in the mouth. The vibrations of the vocal cords continue making for an immediate start of the vowel, and long formant transitions result. The amplitude of the vowel will be high, the pitch will be low because of the condition of the vocal cords (chest voice).

Because of the continuing voice and the slow build-up of a low intra-oral pressure the fricatives will be short, the pause of the plosives will be voiced and short and the noise bursts of the plosives will be short.
5. Summary of acoustic distinctions

The acoustical differences between voiced and unvoiced are summarized in the following table.

<table>
<thead>
<tr>
<th>part of the utterance</th>
<th>acoustic parameter</th>
<th>voiced</th>
<th>unvoiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>preceding vowel</td>
<td>duration</td>
<td>long</td>
<td>short</td>
</tr>
<tr>
<td></td>
<td>fundamental frequency</td>
<td>low</td>
<td>high</td>
</tr>
<tr>
<td></td>
<td>intensity</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td></td>
<td>duration of transition</td>
<td>long</td>
<td>short</td>
</tr>
<tr>
<td></td>
<td>duration</td>
<td>big</td>
<td>small</td>
</tr>
<tr>
<td></td>
<td>excursion of transition</td>
<td>long</td>
<td>short</td>
</tr>
<tr>
<td></td>
<td>intensity</td>
<td>big</td>
<td>small</td>
</tr>
<tr>
<td>fricative</td>
<td>packets of noise</td>
<td>low voice bar</td>
<td>continuous noise</td>
</tr>
<tr>
<td></td>
<td>intensity</td>
<td>low voice bar</td>
<td>voiceless</td>
</tr>
<tr>
<td></td>
<td>character of pause</td>
<td>short</td>
<td>long</td>
</tr>
<tr>
<td></td>
<td>duration of pause</td>
<td>low</td>
<td>long</td>
</tr>
<tr>
<td>plosive</td>
<td>character of pause</td>
<td>low voice bar</td>
<td>voiceless</td>
</tr>
<tr>
<td>following vowel</td>
<td>duration of noise burst</td>
<td>short</td>
<td>long</td>
</tr>
<tr>
<td></td>
<td>intensity of noise burst</td>
<td>low</td>
<td>high</td>
</tr>
<tr>
<td></td>
<td>fundamental frequency</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td></td>
<td>intensity</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td></td>
<td>duration of transition</td>
<td>long</td>
<td>short</td>
</tr>
<tr>
<td></td>
<td>excursion of transition</td>
<td>big</td>
<td>small</td>
</tr>
</tbody>
</table>

6. Initial /t/ and /d/. Experiment I

Suppose the model is valid for initial consonants. The difference between initial and embedded consonants will lie in the way the intra-oral pressure is built up. In embedded position the intra-oral pressure is brought about by an already existing subglottal pressure; in initial position the subglottal pressure has to be built up and consequently the intra-oral pressure rises more slowly.

With initial plosives it will be sufficient in a first order approximation to suppose that the acoustic cues of the preceding vowel can be omitted. It should be possible to change a voiceless initial plosive into a voiced one by adding a preceding voice bar and by shortening the duration of the noise burst. This is done by means of electronic gates. This test succeeded for Dutch initial voiceless plosives. The lack of extensive transitions in the following vowel was not perceived. A voiced initial plosive did not yield a voiceless plosive by omission of the voice bar, because of a low intensity noise burst and large vowel transitions. By additionally removing 30 ms of this transition, just after the noise burst, it was possible to obtain a voiceless plosive. However, vital information concerning place of articulation was lost.

We can conclude that a combination of acoustic parameters had to be changed to get an acceptable change from a voiceless into a voiced plosive (adding voice before the noise burst and decreasing duration of the burst). The same condition holds for the change from voiced to voiceless (removing the voice lead and vowel transitions).

7. Speech synthesis. Experiment II

By means of the speech synthesizer IPOVOX-II 1) a series of VCV combinations were obtained in which the acoustic parameters valid for a voiceless consonant (/t/ and /s/) were changed one by one in the parameters valid for a voiced consonant (/d/ and /z/). While listening to those series one got the impression of a gradual change from voiceless to voiced. Here again the conclusion is that a combination of parameters is perceptually needed.
8. Whispering, Experiment III

Another experiment can be done if we assume that this model can be used for whispered speech, and if we assume that the difference between normal and whispered speech lies in the formation of a so called "whispering triangle" at the back part of the vocal cords. The vocal cords cannot be supposed to vibrate because the air passes through the whispering triangle which has a lower resistance than the front part of the vocal cords.

We are aware this is all very speculative, but should this view be right, we might expect that only the information originating in the voice is lacking; consequently the change-over between vowel and consonant will be less distinct, and it is reasonable to assume that differences concerning the preceding and following vowel will not be significant any more. Because of a lower resistance in the glottis (because of the whispering triangle) especially in the case of the "voiced" consonants, the build-up of the introral pressure will be faster with whispering than with normal speaking.

In a listening experiment it appeared that the expectation based on this supposition that voiceless consonants will be better perceived than voiced consonants (because the voiced consonants lose more of their character) was confirmed by a higher intelligibility of voiceless consonants (85%) than voiced consonants (58%). Moreover, the fricatives are better perceived (73%) than the plosives (68%); this can be explained by the perceptual help of the duration of the fricative noise.

These experiments do not prove the model to be right but they might at least be interpreted to support the tenability of the model as accounting for the presence of a combination of acoustical parameters and for their common origins.

1) Willems, L.F., IPOVOX-II, A Speech Synthesizer. This issue.

A. Cohen

Errors of Speech and their Implication for Understanding the Strategy of Language Users 1)

The general framework against which the errors reported should be seen is as follows: in generating a speech message we assume that there are at least 4 levels at issue -

1) A language potential, an intricate network of relations of language material based on semantic and formal correspondences.
2) A plan involving a selection from this potential as determined by the situation in which the speaker finds himself and the intention he wishes to convey.
3) A programme, controlled by the sequential rules of the language and the stylistic needs of the moment.
4) The actual performance or production of acoustic speech signals.
The primary object of this paper is to see whether a study of the readily observable spontaneous speech errors may throw some light on one or more of the more hidden aspects of the generation of speech. To strengthen the evidence collected rather haphazardly, means were sought and, I hope, found for experimental verification of the observed regularities. In the following, 2 sets of data will be presented; the first made up of observations in ordinary speech situations, "spontaneous errors", the second consisting of errors collected under experimental conditions.

Three types of spontaneous errors were distinguished to be called

I Anticipations; where a speech segment was produced clearly before its turn, thereby suppressing another segment which should have been produced. E.g. expiration in experiments, instead of inspiration in experiments. The segment size involved in this error was a syllable.

II Perseverations; when a speech segment that had already been pronounced turned up once more afterwards, thereby suppressing the segment that was due at that place in the utterance. E.g. I prefer to preserve this term, which should have been to reserve.

III Transpositions or metatheses, when 2 segments completely changed places. E.g. to shut him court for to cut him short.

I distinguished 5 types of segments, namely: single vowel, single consonant, consonant cluster, syllable, and word. Of 600 errors reported in this survey the divisions are given in Table I.

<table>
<thead>
<tr>
<th>Type of segment involved in speech errors</th>
<th>Groups</th>
<th>single vowel</th>
<th>single consonant</th>
<th>consonant cluster</th>
<th>syllable</th>
<th>word</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Anticipations</td>
<td>469 78%</td>
<td>164</td>
<td>121</td>
<td>34</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>II Perseverations</td>
<td>22 15%</td>
<td>39</td>
<td>25</td>
<td>4</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>III Transpositions</td>
<td>9 7%</td>
<td>21</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

The most striking feature is the predominance of errors of group I, those of anticipation, about 80% of the material. Clusters were particularly interesting in that they seemed to follow the lines of single phonemes and could be interchanged with them. Thus a collocation like vowel triangle would come out as trowel triangle.

An interesting case showing automatic substitution of a cluster caused by a single consonant was provided by Prof. Liberman's lecture in Moscow where he used the phrase: hence perception becomes limped to a system etc. where the /m/ of becomes or of system caused the /nk/ cluster of linked to appear as /mp/. Distances between the cause and the occurrence of the errors were expressed in terms of the number of syllables. No error was found to occur within the span of a single syllable. The distribution looks quite normal. The percentage of spans of two, adjacent, syllables was 35%, for three syllables likewise 33%, for four syllables 15%, for five 10%. This applies to errors of group I.

A similar distribution was found for perseverations, both with respect to segment size involved and to syllable span. While in group I, anticipations, the initial parts of words, notably with respect to the consonants, were mostly affected, this was not the case with errors of perseveration. An interesting sideline in perseveration errors is that there seems to be a correlation with physical fatigue.
Group III is too small to warrant further comment.

At first no errors produced under other conditions than those attending ordinary speech communication were taken into account. It transpired, however, from noticing errors produced by readers of news bulletins on radio and TV that these errors seemed to be of a nature very similar to the spontaneous ones. So written texts were used in the experimental situation.

Phrases, meaningful as such, but without mutual semantic connections were typed without capitals or punctuation marks on a paper ribbon. This was shown moving from right to left to the subject who received the instruction to read aloud the moving text while paying particular care in using proper intonation to mark off one phrase from the next.

The test material in one experimental run consisted of 100 phrases derived from the collection of spontaneous errors. Although a number of additional errors were produced, the main outcome was that not only similar errors to the spontaneous ones could be induced in this way, but even exact replicas occurred, sometimes even several times over with different subjects. Again, anticipating errors abounded. The error span also was similar to that in spontaneous speech.

Returning to the framework sketched in the beginning, it seems fair to conclude that integrative processes are at issue well before the actual production of speech, stage 4. The segments involved were not necessarily of phoneme size and the spans were at least two or three syllables. This would agree well with the notion, coined by George Miller of "Chunks of speech", which the listener can handle at one go in interpreting the incoming acoustic speech signals.

In the majority of cases, speakers did not correct their errors and even when they did they often proved to have been unaware of their own slips of the tongue as well as of their corrections. This may be interpreted, perhaps, in terms of a perceptual feedback loop that is highly peripheral.

In this context I would like to stress once more the criterion of awareness of errors in deciding about the mode of speech perception. In ordinary conversation this awareness of errors is fairly low both in the speaker and in the listener. I think this state of affairs may well be characteristic of what has been called the speech mode. In e.g. the interpretation of family names, such as the well known Ditter, this awareness must be taken to be much higher, and so may well be taken to represent a different mode of perception.

A final point about the reading aloud test: the similarity of errors made under experimental conditions and in ordinary speech production may have a bearing on the connection between the visual and auditory patterns of language. The errors produced in either case must have arisen at a stage where the two modalities, reading and speaking were clearly comparable, which would mean not higher than at the programming stage, since stage 2, the plan, was presented ready-made to the reader, whereas the spontaneous speaker would have to devise his own plan.

This type of work is, I believe, in accordance with George Miller's plan for working towards a theory of language usage to come into being next to the already extant theory, or theories, of language.

1) Text of a paper read to the Seminar on Speech Production and Speech Perception at Leningrad August 17, 1966.

G. Forrez

Relevant Parameters of the Voiced Fricative /z/. ¹)

As a contribution to the building of a speech synthesizer (IPOVOX II) a study was made of the various ways in which a noise and a periodic signal can be combined.

The ultimate object was to find acoustic specifications for synthesizing acceptable voiced fricatives, particularly /z/. Three methods were tried; adding, multiplying and a combination of these two. The three series of signals were presented for perceptual evaluation to a panel of listeners.

The combined method proved to be optimal.

¹) Summary of I.P.O. Report V. 55.

J. 't Hart

Perceptual Analysis of Dutch Intonation Features

This report will deal with the various ways in which we have tried to tackle the problem of pitch movements in speech. Pitch is assumed to be the largest contributing factor in the sensation of intonation.

The first step in our attempt to describe the contribution of pitch in Dutch intonation was to select a method for obtaining data about such pitch movements in speech.

As can be seen from the enormous problem of building a reliable pitch extractor, the speech signal does not lend itself too easily for instrumental analysis of a parameter that can be taken to reflect the pitch impression to the human listener. This can have two reasons:

a. the apparatus is not capable of unconditionally measuring the parameter it is designed for, because of the great complexity of the speech signal;

b. the apparatus does measure, in a perfect manner, the wrong parameter.

Since most of the troubles can be reduced to imperfect performance of the instrument, the second possible reason cannot even be checked.

In our attempt to find another method to measure pitch in speech, we looked back to the perceptual analysis of speech sound qualities. In the previous years, this technique had been employed in order to find perceptual segments which were hypothesized to be the constituent parts that were relevant for the perception of speech. The analysis procedure was followed by speech synthesis in order to test this hypothesis (Cohen and 't Hart, 1962).

As perceptual analysis of speech sound qualities had been carried out with an electronic gate which, by isolating the desired parts of speech, helped the listener to judge their quality, once more such a gating technique was opted for, this time with a view to analyzing pitch contours.
The demand for a stimulus duration that would be optimal to allow concentration on pitch, led to a gate width of 30 ms. This is a compromise between too short for a listener to hear any pitch at all, and too long for a quasi-stationary pitch impression.

One important difference between this and the earlier perceptual analysis is constituted by the fact that a trained listener, though being capable of judging vowel quality in an absolute manner, may not possess absolute pitch. Therefore, in pitch measurements, the isolated parts are matched against an equally long artificial sound with approximately the same timbre and loudness, whose periodicity can be adjusted and subsequently measured on a frequency counter. In this matching technique a trained listener can reach an accuracy of about 1% (Cohen and 't Hart, 1965).

Comparison of pitch curves obtained from perceptual analysis and the contours of the fourth or fifth harmonic in a narrow band spectrogram with enlarged frequency scale showed no discrepancies. However, in many stretches where the perceptual matching technique was able to yield pitch curves, the spectrographic method was not. Moreover, though the procedure is less time consuming, the accuracy is much less in such spectrograms. On the other hand, we do not know if greater accuracy necessarily leads to better insight in intonation.

Though we can be certain to have measured pitch in this way, we cannot be so sure about the perceptibility of all the details in the total pitch contour of, e.g., a syllable.

To study this, we employed "intonation synthesis", according to the same idea that had led to speech synthesis before. In an instrument built for this purpose, the fundamental frequency of a voice source can be controlled by nine independently adjustable dc-voltages that are to be triggered successively by the nine available outputs of the IPO cascade counter (Moonen and Lammers, 1966). In practice, the lengths of the steps were often taken 30 ms. In this way, with this special voice source substituting for the ordinary one in our speech synthesizer, intonation patterns on single syllables could be generated. Starting from a copy of the measured pattern e.g., the perceptual tolerance for variations of this pattern could be tested.

Remarkable discrepancies between the impression of a particular curve and the actual contour of the fundamental frequency were observed (see fig. 1 next page).

E.g., with \( F_0 \) rising initially during 80 ms and, after remaining constant for 30 ms, falling over 100 ms, the impression is that of a mere fall (Fig. 1a). Furthermore, the inverse pattern, in which \( F_0 \) begins to fall and continues with a rise, is interpreted as rise-fall-rise (Fig. 1b). Thus, depending on the total structure of the pattern, rises may be omitted or added subjectively.

Although this procedure seemed to be a fertile one to learn something about relevant and non-relevant pitch movements in single syllables, or perhaps even in polysyllabic words, it turned out to be impossible to extrapolate from here to longer utterances. In other words, expectations about the pitch contour in a sentence could not be based upon knowledge, obtained so far, about the intonation of isolated words.

It seemed inevitable to undertake the laborious task to analyze, with the matching method, such longer utterances themselves. In making such analyses we encountered a new problem, viz. the fact that, in comparison with single syllables, longer utterances contain relatively fewer voiced stretches.
It should cause no surprise that in this way very little was learned about general characteristics of intonation. Still one important feature was observed which we had not meet before: it turned out that all inflections could be described as deviations from an overall gradual downward slope, to be called declination.

A promising step forward was made once we accepted the challenge to try, on the basis of this declination and all the other features we seemed to have discovered so far, to predict pitch contours for small sentences (of about ten syllables' length). In comparing the predictions with actual measurements, and particularly in being confronted with wrong predictions, we learned, in a relatively short time, quite a lot of what could be taken to be the most essential part of the intonation contour. Thus, the prediction technique appeared to be a powerful tool for analysis.

Now again the need was felt for intonation synthesis, this time on the level of sentence-length utterances. Complete artificial generation of the phonetic content seemed too laborious, and therefore we made use of an apparatus, called the Intonator, which is based on the principles of analysis-synthesis telephony (Willems, 1966). In this apparatus frequency control of an artificial voice source takes place by external means, in our case by a function generator, especially designed to generate functions according to those rules that had proved to be successful in the prediction technique.
Again, the Intonator, though originally being meant as a synthesizer, turned out to be helpful for the purpose of analysis. If, e.g., one wants to know what are the essential movements of pitch in a sentence, one can try to construct a maximally simple contour that is still acceptable as natural to listeners.

In an article, submitted for publication in Lingua, a survey of the results will be given as well as an ample discussion about them. Some of these results may be briefly summarized here.

First of all it seems feasible to distinguish three different layers: major, minor and micro layer. Major refers to movements that can be found in any utterance, minor features occur or do not occur depending on the incidental word usage; both major and minor features are under conscious control of the speaker, whereas micro features have an involuntary character and depend solely on articulatory constraints.

---

![Fig. 2: Hat pattern as being observable in pitch measurements in the sentence **hij is niet veel veranderd sinds vroeger** (he did not change much since formerly).](image)

![Fig. 3: Illustration of various minor features. I shows a hat pattern with an anticipatory dip (a) and a dent (b). I en II form a complex of two hat patterns, with a linking feature in between (c). II ends in a post-hat upward wriggle (d), different from c, since the new hat pattern III starts at an arbitrary level.](image)
In the major layer a hat-like pattern seems to constitute the canonical form. Between two collinear stretches of declination an upward shifted piece of declination line can be seen, with a rise and a fall at either side (Fig. 2).

In the minor layer such a hat can be preceded by an anticipatory dip (Fig. 3a); an extra fall and rise may cause a dent in the hat (Fig. 3b), this situation being distinguishable from a complex of two hats through different conditions for the fall in the dent and the fall at the end of a hat pattern. Moreover, in complexes of hat patterns a linking feature will be found, consisting of an upward wriggle followed by a pitch on the extension of the declination line (Fig. 3c). Other upward wriggles that may be followed by an arbitrary pitch must be taken to be "post-hat" features, without bearing a cue for continuation (Fig. 3d).

The micro layer shows roughly slight elevations for all full vowels, but the preceding consonant influences the shape of these elevations.

In conclusion, it should be pointed out that, being interested in perceptually relevant features, we always applied some technique of pure analysis, followed by some kind of synthesis, which, at its turn, appeared to be applicable as a new type of analysis. Originally, by perceptual analysis we meant some kind of pure analysis, in which use is made of the analyzing faculties and constraints of the ear. However, on the strength of our experience in our attempts to describe intonation from a perceptual point of view, we think it better to call generally "perceptual analysis" such a whole complex of methods applied, including both analysis and synthesis.

References

Cohen, A., 't Hart, J.
1962 Segmentation of the Speech Continuum.
Proc. 4th Int. Congress on Acoustics, Copenhagen, G.51.

Cohen, A., 't Hart, J.
1965 Perceptual Analysis of Intonation Patterns.
Proc. 5th Int. Congress on Acoustics, Liège, A.16.

Moonen, G.J.J., Lammers, C.A.
Willems, L.F.
1966 A Preset Cascade Counter.
This issue.

A. van Katwijk

On Perceptual Units in Speech

The constructs in terms of which what is known as linguistic competence is described, owe their relevance to the classification they allow of linguistic material. In language usage, however, utterances are produced and perceived with no view to comparison of similar utterances — rhyming and making puns, though evident instances of one utterance evoking another, being apparently too tiresome for continued practice. Apart from artificial usage, it is not immediately clear what part linguistic constructs play in organising linguistic performance.
Distributional phenomena in phonology certainly belong to the study of linguistic competence, even though we attempt to relate them to performance, as will be tried below.

First, rules will be presented describing distributional constraints in Dutch phoneme sequences, then experimental data will be given with a view to demonstrating in what respects distributional frames coincide with perceptual units (canonical forms, as it were) which will, in conclusion, be confronted with data from the study of lapses by Cohen, 1966.

Distribution of Dutch phonemes

Certain segments of speech with a definite perceptual identity, lend themselves well for transcription in the I.P.A. alphabet. Dutch utterances will be considered to be constructed from sequences of these, which for the sake of convenience are called phonemes.

The rules describing the set of grammatical phoneme sequences are given in the now familiar form of rewrite instructions. They describe the fact that in Dutch, grammatical phoneme sequences (PS) may open with a consonant or consonant group (C₁) or with a vowel (V); having opened with C₁, the sequence will then have V. V can be followed by an interval / /, or by a second consonant or consonant group (C₂), which in its turn may be followed by / /. Instead of / /, a new cycle may follow.

Formally the rules run as follows:

1. Utterance PS
2. PS → (C₁)V(C₂)(PS)

In rule 2, the symbol PS recurs at the end, thus making for repeated application of rule 2., which will result in sequences longer than one (C₁)V(C₂).

It will thus be noticed that the pair of rules define a recurring constituent that equals the size of a syllable. As will be argued below, this constituent is of sufficient size to allow the description of distributional constraints, i.e., the dependencies among phoneme occurrences can be stated within the frame of one (C₁)V(C₂) (with the exception of assimilation phenomena).

In Dutch, the set C₁ consists of 17 individual consonants (cf Cohen et al., 1961), about 33 consonant pairs (out of 272 theoretically possible permutations), and about 4 consonant triads (3080 permutations).

V contains 16 vowels; C₂ comprises 13 individual consonants, approximately 40 consonant pairs, 24 triads and 4 tetrads. Triads in C₁ and tetrads in C₂ always have an /s/ among their constituting consonants, (C₁ in its first and C₂ in its third or fourth position). For this reason, in these cases, one might speak of "Nebensilbe"/s/ or /st/, a term coined by Sievers, 1901.

As regards transitional constraints, there is hardly any restriction in V, given a preceding element from C₁. Some non-occurring sequences, such as /ji/, are to be rejected on intuitive grounds; other sequences, such as /wy/, occur in the name of a village (Wuustwezel) while /wʃ/, for instance, though not existing in any actual form, is intuitively admissible.

Restrictions are numerous between V and C₂. Moulton, 1962, amongst others, based his classification of Dutch vowels on these restrictions. As V determines C₂ to a certain extent, and vice versa, the information of V is slight given C₂, and the information of C₂ is slight given V.

After C₂ a new cycle can start with any element from C₁.

If (C₁)V(C₂) is a cycle, it is a peculiar one in that the order of quite a number of permissible sequences of elements from C₁ is mirrored with respect to V in element sequences from C₂.
Such a cycle is sketched in the following figure: speech might be regarded as continued running through this cycle.

Intelligibility of interrupted speech

By electronic means a specimen of running speech was stretched by the introduction of gaps, first between all phonemes (i.e., between all segments that were identified as phonemes), and then between all \( (C_1)V(C_2) \)'s. In the former case the utterances had become unintelligible, in the latter possibly even more intelligible than the original. These results can be interpreted to indicate that speech perception is not a phoneme-by-phoneme identification process, but that it is rather the syllable that plays a part in speech perception as a unit of intelligibility.

The next problem was to find out whether the infra-structure of \( (C_1)V(C_2) \) as manifested in distributional constraints could be traced in intelligibility experiments. A well-known procedure in this type of experiments is repeated interruption of speech (Miller & Licklider, 1950). In a comparable set-up, Huggins (1964) found that the intelligibility of disrupted speech was minimal when interruptions occurred three times per second (on a basis of a 1:1 speech-to-silence ratio). Huggins preferred to take this not as a property of our hearing faculty, which could be supposed to need some processing time, but rather of speech itself, and notably the syllables uttered and perceived at the rate of three per second.

On this cue borrowed from Huggins an experiment was carried out for which running speech was recorded that was synchronised with clicks in the headphones of the speaker. Thus we obtained a syllable at every 360 ms on one track and synchronising clicks on another. Every syllable was partially made audible (200 ms) and the rest suppressed (160 ms).

The audible parts were arranged according to four orientations, coinciding more or less with the following representation:

orientation (1): \( (C_2)(C_1) \) \( V \) ("V" denotes a part of \( V \))
orientation (2): \( (C_1) \) \( V \)
orientation (3): \( V (C_2) \)
orientation (4): \( Y (C_2)(C_1) \)

After some trial runs on similar material eight subjects were asked to "shadow" the disrupted speech in each of the orientations. The numbers of correct syllables reproduced were scored.

In orientation (1) 84% of the 960 syllables were correctly reproduced, in (2), 89%; in (3), 58%; and in (4), 25%.

These results suggest that the onset of the syllable, i.e., \( (C_1) \) plus \( V \) contributes more to the perception of speech than the coda, i.e., \( V \) plus \( (C_2) \).
This is in agreement with what was observed when subjects were asked to divide a text into syllables. There turned out to be a tendency to favour C₁ at the expense of C₂. For instance, in Dutch, verste is preferred to versste.

As regards distributional restrictions, in the first two orientations the audible part coincides with the part that is distributionally free and has a high information value, while the reverse is true of orientations (3) and (4). Moreover, it can be observed that C₂ in the first orientation does not contribute much to the intelligibility of the next cycle, as little as C₁ does to that of the preceding cycle.

In conclusion, (C₁)V(C₂) can be said to play a part in the perception of speech, which part is partially determined by the distributional structure within.

Errors of speech related to (C₁)V(C₂)

The confrontation of ideal (C₁)V(C₂) frames with actual usage poses many problems, one of which is the problem of lapses of speech (Cohen, 1966). At first sight the occurrence of errors of speech seems to be incompatible with the hypothesis that (C₁)V(C₂) is a perceptual unit, for the span of cause and effect of such errors always exceeds one syllable. On second thoughts it turns out to be quite feasible to extrapolate quality and quantity of errors to some significant extent in terms of the (C₁)V(C₂) structure.

a. Given the fact that errors of speech do occur, and given the hypothesis that (C₁)V(C₂) is a canonical form of some sort in language usage, it would be more surprising if both cause and effect of errors occurred within this unit, which would give rise to loss of its identity.

b. The error-causing element should, on the strength of the supposed relevance of (C₁)V(C₂), be expected to belong to the same set as the error itself, i.e., interaction between cause and effect is expected to occur only between elements of C₁, or of V or of C₂ in subsequent units.

c. The distributional freedom of C₁ with respect to C₂, the evidence of the greater perceptual load of C₁ as compared with C₂, to which can be added the notion of speech being an active selection of building blocks from some available store, together make errors in C₁ more likely than in C₂.

d. With regard to the mirrored order of a number of consonant sequences in C₁ as compared with C₂, V takes a central position in (C₁)V(C₂) with respect to which C₁ takes a regressive position, and C₂ a progressive one. On the basis of this representation, one might expect that the most natural error for C₁ would be one in which regression is "miscalculated" by the speaker and is carried too far, resulting in anticipations. The same holds for the progressive position of C₂, making way for perseverative errors. From these considerations one can expect anticipations to be more numerous in C₁, and perseverations more numerous in C₂ as compared with alternative error qualities.

e. From the considerations given under d., the central position of V in (C₁)V(C₂) should lead one to expect that the most likely error to be made in V would be of the transposition type.

Expectations a. to d. are supported by the data to be found in Cohen's paper, while e. does not materialise, for anticipations of vowels turn out to be more numerous than transpositions.

Though the results coincide to a significant extent with the expectations, it should be kept in mind that not all suppositions on which the expectations were based, are justified by the results.
Still, the assumption that $(C_1)V(C_2)$ is a canonical form both in language and in speech seems to be validated by the facts adduced.

Cohen, A. 1966 Errors of Speech and their Implication for Understanding the Strategy of Language Users. This issue.


5.c. VISUAL PERCEPTION

H. Bouma and J.J. Andriessen

Anchor Effects in the Perception of Slant.

This investigation is concerned with one aspect of the spatial elaboration of retinal signals that occurs in the visual system, supposedly as part of form perception processes. Guided by the recent electrophysiological evidence that retinal orientation is among the first isolated spatial cues, we have used a psychophysical method to study how perceived orientation and geometrical slant of a line segment are related. (Bouma, H. & Andriessen, J.J., 1966)

In the experiments a line segment with a definite inclination or slant was projected on a screen by means of an optical system, while a dot, projected on the screen as well, had to be adjusted in order to get it in line with the line segment by using a set of appropriate push buttons. Perceived orientation has been defined operationally as being represented by the position of the dot. Usually the line segment was flashed (flash duration 100 ms). The line segment was presented under 15 different angles of slant. Three trained subjects took part in the experiments.

Results indicate that perceived orientation of oblique line segments is closer to the nearest horizontal or vertical than is geometrical slant, the difference being of the order of 5°. For horizontals and verticals, orientation and slant almost coincide, the adjustments showing minimum standard deviations, (about 0.5°). Standard deviations approach the values that visual acuity can be assumed to permit.

It does not seem unlikely that the existence of "horizontal" and "vertical" as preference orientations (or subjective anchors) is at the root of the observed phenomena. Up till now, electrophysiological evidence for the existence of preference orientations seems still lacking.

Further, we have tried to find evidence of what type of information is actually used in the visual system to arrive at "orientation". For this purpose we have varied the length of the line segment between 14 and 360 minutes of arc and, also, we have replaced it by two dots at the extremities of the line segment. Results are consistent with the hypothesis that a foveal part of the line segment of about 60 minutes of arc supplies the information. However, the hypothesis that the extremities bring about the information cannot quite be ruled out.

As to the underlying processes, the hypothesis that there exists a set of orientation filters each tuned to one particular retinal slant, seems sufficiently flexible to cover present results.

Reference
H. Bouma & J.J. Andriessen

Induced Changes in Perceived Orientation

It has long been known that perceived orientation depends not only upon geometrical slant but on other factors in the configuration as well. Many visual illusions provide clear demonstrations of these effects (Fig. 1).

(a) Poggendorff's illusion. The lines x and y are continuous.
(b) Hering's variation of Zöllner's illusion. The diagonal lines are parallel.
(c) Hering's variation of Zöllner's illusion. The 4 intersected lines are straight and vertical.

The most common explanation is in terms of a counterbalance effect that the visual system imposes automatically upon geometrical slants in an attempt to restore in perspective scenes the true geometrical angles. Recently, it has been proposed to look for an explanation rather in terms of interaction of line detectors that presumably exist in the visual system (see the discussion between Gregory and Day, 1965).

We aimed at isolating the effect experimentally in a very simple configuration. In a plane perpendicular to the line of sight, subjects observed a flashed oblique stimulus line that ended just at its intersection with a second vertical induction line (Fig. 2).

As reported in "Anchor Effects etc.", elsewhere in this issue, we asked our subjects to adjust a dot P until it appeared in the extension of the stimulus line. As a measure for the change in perceived orientation brought about by the induction line we used the angular difference $\Delta \theta$ between settings with and without the induction line being present. Regular values for the various parameters were:

stimulus line, $\alpha = 315^\circ$, $l = 21$ minutes of arc, $t = 100$ ms;
induction line, $\theta = 90^\circ$, $m = 360$ minutes of arc, continuous presentation.
Induction line m

Stimulus line l

Fig. 2: Adjustment of a dot P in line with the stimulus line l of slant $\alpha$. The induction line m, having a slant $\delta = 90^\circ$, is placed against the stimulus line. $\beta$ is the angle deviation of the adjusted dot P from the geometrical position.

Fig. 3: The angular differences $\Delta \beta$ between settings with and without induction line as functions of slant $\alpha$. The induction line is oriented vertically. The perceived orientation of oblique stimulus lines tends towards a perpendicular position with respect to the induction line.

Fig. 4: The angular differences $\Delta \beta$ between settings with and without induction line as functions of slant $\alpha$. The induction line is oriented horizontally. For values of $\alpha$ between 360° and 330° the perceived orientation of oblique stimulus lines tends towards the position of the induction line.
The influence of the slant $\alpha$ of the stimulus line is depicted in Fig. 3. Values of $\Delta \beta$ are generally positive and go up to some 10°. This indicates that the change in perceived orientation is away from the induction line towards a perpendicular orientation. Assuming that perpendicularity acts as an internal reference or anchor angle, for which there is some preliminary evidence, we may reformulate the effect in terms of an attracting influence exerted by the anchor angle. This has the advantage of bringing it in line with the effects observed for isolated line segments, for which perceived orientation is attracted by the anchor orientations horizontal and vertical (see Anchor Effects etc.).

Contrary to the above findings, we observed a rotation towards the induction line only for the case of an horizontal induction line combined with a stimulus line and at an angle $\alpha = 360^\circ > \alpha > 330^\circ$. (Fig. 4).

A study of variations of the lengths both of the stimulus line and of the induction line has led us to believe that the greater part of the induction effect arises from an area of about 1° around the point of intersection. The connection between both lines stands out as the most critical parameter. This was established experimentally by varying the distance $p$ between stimulus line and induction line. Close connection between both lines shows a maximum influence of the induction line. (Fig. 5).

![Graph](image)

**Fig. 5:** The angular difference $\Delta \beta$ as functions of the distance $p$ between stimulus line and induction line. The induction line is oriented vertically. Close connection between both line segments appears to show the greatest influence of the induction effect.
<table>
<thead>
<tr>
<th>References</th>
<th>Year</th>
<th>Title</th>
<th>Journal, Volume, Pages</th>
</tr>
</thead>
</table>
In psychophysical investigation of visual processes, threshold determination has proved to be a useful technique, considering the wide application it has received. Usually a luminous object is superimposed on a homogeneous field. The luminance of this object is then lowered until the subject reports no difference from the background. Since the threshold criterion is both sharp and constant, this technique yields quite reliable results.

It is essential that the influence of the background is selective, certain test objects being more suppressed than others. For example, when increasing the diameter of a circular object, threshold luminance decreases up to a certain diameter, beyond which no differences occur. This critical size can be interpreted as the size of the retinal area within which luminance signals interact (receptive field). These particular receptive fields are of a circular shape, the diameters increasing from a few minutes of arc in the fovea to about one degree in the retinal periphery.

Recently, the scope of this technique of selective suppression has been widened by Marsh and Cherry (1966). Rather than using a homogeneous and steady background, they chose a background filled with light and dark dots in continuous variation. Such a field can most readily be produced by providing the video input of a TV monitor with a random signal of sufficiently high cut-off frequency. In analogy with the noise sound obtained by feeding such a signal into a loudspeaker, the term visual noise is commonly used for such a field full of contrasting dots, which perceptually has some resemblance to whirling snow. Suppression by means of visual noise turns out to have a different effect from that of a homogeneous background of the same average luminance. The receptive fields are line shaped rather than disc-shaped with the threshold of lines decreasing until a length of some 30 has been reached. Taking two parallel lines as a test object, the threshold is quite similar to that of a single line, down to distances of some 6 minutes of arc. When the line segment is gradually curved its threshold is found to increase. We have observed that a slightly curved line looks quite straight when it is just above threshold level. These results are independent of the geometrical slants of the test lines.

These results then seem to indicate the existence of receptive fields having a line structure with lengths up to 30. This property of the visual system to select certain configurations rather than others is supposedly part of the spatial analysis making for form perception. Many parameters of background and test fields still have to be investigated, among which the sizes of the noise dots.

A next logical step seemed to be to choose a background field full of straight lines, which might be supposed to produce a general stimulation at the next higher level of visual analysis. Accordingly, the configuration of the most resistant test fields might again indicate spatial analyzers of a still higher type of organization.

Technically, such a background field can most readily be produced by lowering the HF cut-off frequency of the random video signal, in which case a line-structured field is obtained in which all lines run in the scanning direction of the monitor.
The test fields that are most resistant against such a background field are lines running perpendicular to the background lines, whereas parallel lines have the highest thresholds. These results also hold for more complex configurations, for example the character A is perceived as A. The selectivity of the suppressing effect runs over a considerable angular range, differences between lines running at 45 and at 90 degrees still being clearly expressed. By rotating the monitor and the test line independently, it has been ascertained that slant differences rather than absolute slants are the relevant parameters. In particular, no preference for the horizontal and the vertical has been obtained.

As far as these experiments go, they seem to indicate a chain of analyzing processes in which successively (a) local luminosity (b) local contrast, and (c) contrast lines of various retinal orientations are sorted out. Though the evidence is of psychophysical nature, the line of thought is certainly influenced by the physiological evidence afforded by Hubel and Wiesel (1962) about spatial analyzers in the visual pathways of cats and monkeys.

A next step in this line of approach seems to be to use as a background lines running in various directions. Also, one might investigate to what extent line length as such can be selectively suppressed. Much supplementary work remains to be done. Thus, a better physical specification and control of the various noise signals seems desirable. We attempt this by using randomly appearing signals to trigger a square pulse of specified dimensions, thus obtaining randomly appearing line segments of specified lengths and intensities.

We are also trying to conceptualize line detectors or "orientation filters" that would behave as the visual data indicate.

In conclusion, the technique of selective suppression seems to allow of considerable extension towards unravelling analyzing processes that occur in human visual form perception.

Experimental work for this survey has mainly been carried out by students H.A. Dortmans (Onderzoek van vormwaarneming met behulp van ruis. IPO Minor Report 71), Y. Hirai (Detection of Simple Configurations in Visual Noise, IPO Minor Report 79), P.J.M. Savelsburgh, J.G. Pennings.

References

Marsh, A.H., Cherry, C. 1966 Preliminary experiments on the perception of visual forms in noise. 28 Intern. Congr. of Psychology, Moscow, Symp., 16, 44-61.
Magnitude Estimations of Colour Attributes

Introduction

Many cases have been reported in the literature on colour, where the C.I.E. colour specifications do not correlate well with the colour appearances of surface colours. And, since it is a matter of great theoretical and practical importance to specify a colour in a manner which accords with its appearance, so much work has been done on the problem of colour appearance. A comprehensive review of this subject was given by Judd (1965) in the opening lecture of the Lucern International colour meeting.

However, what is relevant to the subject of this investigation is the methods which have been used to assess the colour appearance of an object, and how the differences in the colour appearance, if any, under different conditions were evaluated. The methods used in such studies were either the binocular or the memory matching techniques. A brief account of each of them and of the method which was used in the present work is given below.

Binocular matching technique

In this method an instrument is used, by which part of the field of view (testfield) is seen by one eye, say the left eye, while the other part of the field (matching field) is seen by the right eye.

If it is required to measure the difference in the colour appearances of an object when it is seen under two different conditions, A and B, two matches are obtained as judged by the right eye whose state of adaptation is assumed to be unchanged. The first match is obtained when the object is seen by the left eye under condition A and the second is similarly obtained for condition B. The difference between the two matches is taken to be corresponding to the difference in the colour appearance of the object as seen naturally under A and B.

Memory matching technique

In this method the observer is trained to recognize the colour appearance of surface colours of any colour system, say the Munsell system, to remember their notations for hue, saturation and lightness.

If he is shown any object under any condition, he will give its notation relying on his memory.

Magnitude estimations method

In this method, the subject is simply presented with the testsample and is asked to assign numbers to that sample, so that these numbers are proportional to the subjective magnitude of the attribute under observation.

Comparison between binocular and memory methods

Although, the binocular method gives higher reproducibility than the memory method, yet its validity is questionable. It has been assumed by authors using the binocular method that changing the state of adaptation of one eye will have little or no effect on the colour characteristics of the second eye. This assumption is not true if the change in the state of adaptation is large.
For example, a binocular field which looks yellow, when the stimulus falls on one eye (say the left eye) only, may change to yellowish green if the right eye is stimulated by a red adapting field falling in a retinal area surrounding that which corresponds to the stimulated area of the left eye.

The memory method has the advantage that the viewing conditions are the same as those of every day life; but the reproducibility of its results is rather low.

The aim of the present investigations

The purpose of this undertaking is twofold. First, to try the magnitude estimations method in assessing the colour attributes of surface colours and compare the reproducibility of its results with those of the binocular and memory methods. Second, to test the validity of the assumption, which is frequently used in the study of colour appearance problems, that the Munsell hue, chroma and value, corresponds to our perception of the colour attributes hue, saturation and lightness respectively.

Experimental results and conclusions

Two observers assessed the colour attributes of 60 Munsell samples as seen against seven different backgrounds, (black, grey, white, red, yellow, green and blue). Part of the results obtained are represented in figs. 1-4.

Fig. 1a: Hue and saturation estimations for observer HvB, using same samples shown against black (•), red (x), yellow (Δ), green (+), and blue (o) surrounds. The lines join estimations of the same sample for the chromatic surrounds and for the black surround.

Fig. 1b: The same as fig. 1a but for obs. PB.
Fig. 1a gives the hue and saturation changes of the colour appearance of some samples as seen against black compared with the red, yellow, green and blue backgrounds for observer HvB, and fig. 1b gives the same for obs. PB.

The shifts are in accord with what is known from previous investigations about the effect of chromatic backgrounds on the colour appearance of surface colours, with few exceptions for the blue background. A comparison has been made in the value of the standard deviation of a single match, between our results using the magnitude estimation method and those obtained by Helson et al. (1952) using the memory matching technique and those obtained by Wassef (1955) who used a kind of binocular method. The average values for the standard deviation of a single match for hue judgements if calculated in terms of the 400% magnitude scale in the three methods are $\sigma_m = 9.8, 13.0, 22.8$ for magnitude estimations, binocular and memory method respectively. (cf. supplementary remarks)

This indicates that the magnitude estimations method is an adequate one for assessing colour attributes.

Fig. 2; Relation between log magnitude estimation for lightness (L) and log reflectance of the samples. The percentages reflectance are obtained from Newhall et al. (1943).

Fig. 2 shows the relation between the reflectance of the samples used and their magnitude estimations plotted on a log-log diagram. The relation on this plot is a linear relation which means that the magnitude estimations for lightness are related to reflectance by a power function, a finding which is the same as obtained by Stevens and Stevens (1960) for both brightness and lightness. The exponent in our case is lower than that in Stevens' results. However, no linear relation can be plotted to represent the relation between the magnitude estimations for lightness and the Munsell value.
This means that lightness belongs to the class of sensory continua which Stevens calls prothetic continua. And since from Stevens' work, for prothetic continua, the magnitude estimation is the one which correlates with our subjective judgments, therefore, the Munsell values are not linearly related to our perception of lightness.

Fig. 3: Relation between saturation estimations (S) and Munsell chroma.

Similarly fig. 3 gives the relation between the Munsell chroma and the magnitude estimation for saturation. It is easily seen from the figure that there is no linear relation, therefore Munsell chromas are not linearly related to our perception of saturation.

Fig. 4a and b give the relation between estimations of hue and Munsell hue for the observer HvB and PB respectively. Although the curves which fit the results are not straight lines, yet, within the spread of the results obtained, straight lines can be drawn. Therefore, it could not be decided from the hue measurements whether or not hue belongs to the prothetic or the metathetic classes of sensory continua.

To sum up, it can be concluded that:

a. The magnitude estimations method is an adequate one for assessing colour attributes, and the reproducibility of the results obtained by this method is at least as good as that of the binocular method, while both of these methods are higher in their reproducibility of their matches than the memory method.
Fig. 4a: Hue estimations (H) for obs. HvB for samples having chroma 6 or higher shown against the Munsell hue scale.

Fig. 4b: The same as Fig. 4a but for obs. PB.
b. Munsell value and chroma are not linearly related to our perception of lightness and saturation. The Munsell hue estimations may be linearly related to our hue perception.

References


1) An extensive evaluation will be published elsewhere.

Supplementary remarks

On closer examination the comparison between the reproducibility of the three methods has to be considered more precisely, because their standard deviations (9.8, 13.0 and 22.8) are not comparable directly. From our measurements, calculations show a significant difference between the intra- and the intersubjective standard deviations. These values are 7 and 17 respectively. Therefore, the individual differences between the observers may be responsible for the rather high value of the standard deviation given by Helson and we have to compare that value 22.8 with the value 17 from our method. The value 13.0 from Wassef's results is derived from a set of observations involving many parameters. However, a comparison is still possible if we use the standard deviation 2.14 which she gives for the particular case of a set of red samples matched by one observer. In terms of the magnitude scale this means 8.6, which value shows no significant difference with our intrasubjective standard deviations.

H.J.J. van Bussel.
Time is an important variable in many psychophysical experiments on vision. Much literature can be found about several types of dynamic measurements, but little is known about their interrelation. Mutual quantitative relations would increase the understanding of the processes which govern the dynamic behaviour of the visual system.

The purpose of the investigation, reported briefly here, is to explore possible relations among

1. the flicker fusion boundary,
2. the flash threshold, and
3. the visual latency.

Re (1). At the boundary of flicker fusion the amplitude of modulation of a sinusoidally modulated light, has a value at which flicker is just perceived. This amplitude depends on the frequency of the modulating signal and the mean level of stimulation. Fig. 1 shows the stimulus intensity as a function of time. Fig. 2 illustrates how results are often visualised. The amplitude sensitivity, i.e., the reciprocal of the modulation amplitude at flicker threshold, is plotted as a function of the modulation frequency on a double logarithmic scale (De Lange characteristic). The line, which separates the domain where flicker is perceived from a domain without flicker, we shall call the flicker fusion boundary. Experimental results for one subject are shown in Fig. 3, the mean level $E$ being the parameter. As in all experiments to be described the left eye is stimulated by white light, the stimulus having a diameter of one degree and being situated foveally.

---

**Fig. 1:** Stimulus intensity as a function of time. $E$ is the mean level, $e$ the amplitude of the modulating signal.

**Fig. 2:** Amplitude sensitivity as a function of frequency.
Re (2). The threshold for rectangular flashes is the minimum intensity required for perceiving the flash. For short flashes the product of duration and intensity at threshold is a constant (Bloch's law). For long flashes the threshold is, roughly speaking, only determined by the intensity. The transition between both situations is characterised by the critical duration.

Fig. 3: Amplitude sensitivity as a function of frequency, the mean level $E$ being the parameter. The symbols are experimental results obtained by subject H.J.M. (The dotted lines are theoretical curves for a one-parameter model not explained in this brief report).
Fig. 4 illustrates how the experimental results usually are plotted. In this graph is the product of the duration $J$ and the intensity amplitude $E$ above the mean level $E$. The critical duration is defined by the intersection point of the two tangents. As in the case of the flicker fusion boundary, the experimental results are functions of the background intensity. With a negligible loss in accuracy this is also the mean level.

Re (3). The visual latency is the difference in time between the onset of the stimulus and the moment the subject is aware of it. This time lag is a function of stimulus step and its background intensity. The higher the stimulus intensity, the shorter the latency. As this delay is an integral part of the reaction time, the measurement of the latter provides a means to find the variation in latency as a function of stimulus luminance.

(This method is not trivial, see Roufs, 1963).*)

Fig. 5: Reaction time as a function of the relative intensity expressed in threshold units. Each dot represents the mean of 60 reaction times of subject H.J.M. (Background intensity 0.4 td) The dotted line gives the latency of the theoretical model.

Fig. 5 shows the variation in reaction time with respect to a rectangular intensity step for one subject and one background level. The step intensity is expressed in Fechner's (i.e. the intensity divided by the step threshold intensity).

It is obvious from fig. 4 that the flash threshold curves belonging to the different levels E can be made to cover each other simply by a horizontal and a vertical translation. The shift along the horizontal axis can be obtained by dividing the duration values at each level by the critical duration. To shift the ordinate values one may divide them by the threshold amplitude for long flashes or, which is the same, multiply them by its reciprocal, the sensitivity F. In other words: The time constant $T_c$ and the sensitivity F determine the position of a curve of known shape.

For the flicker fusion boundary we try the same kind of reduction. The amplitude sensitivity is divided by the sensitivity S which is defined by the maximum value of the amplitude sensitivity at a certain level (see fig. 2). The position of the curves with respect to the horizontal axis may be characterised by a dominant frequency $f_d$, which is defined as the frequency at which the amplitude sensitivity is decreased to half the value of the maximum (fig. 2). If we divide the frequency values by $f_d$ or multiply them by the reciprocal $T_d$, we get reduced curves as in fig. 6. In this case, too, we may say in first order of approximation: the dominant time constant $T_d$ and the sensitivity S determine the position of a curve of known shape.

In search of aspects which flash threshold and flicker fusion may have in common, we first compare both characteristic time constants as functions of the mean level. The functions are plotted double-logarithmically in fig. 7.
Fig. 6: Reduced amplitude sensitivity as a function of reduced frequency for subject H.J.M. The symbols have the same meaning as in fig. 3. (The full lines represent the extreme situations of the analogon referred to in fig. 3).

Fig. 7: The variation of the dominant time constant $T_d$ and the critical duration $T_c$ with the mean level $E$. Subject H.J.M. and subject R.K. The dotted line runs at the theoretical distance from $T_c$. 
Log $T_d$ and log $T_c$ run parallel at a distance of about 0.3 log unit. The constant ratio of about 2 for these time constants strongly suggests a fundamental relation.

The second set of magnitudes, suitable for comparison, are the sensitivities $S$ and $F$. They are plotted as functions of the mean level in Fig. 8. Again the resemblance is striking.
What is at the basis of the obvious relation between the two kinds of measurement? Fig. 9 shows a model which could explain this. The input signal is first processed in a (quasi) linear filter L of which the parameters are determined by the level E. An amplitude detector which has a threshold value \( d \), is connected to the output of system L. Only if the output signal of L exceeds the value \( d \) does the subject see a change in the visual field. Stimulation with sinusoidally modulated light causes the subject to see flicker only when the maximum of the sinusoidal output of L exceeds the value \( d \). This implies that the fusion boundary of fig. 2 represents the modulus of the transfer-function of filter L, divided by the unknown constant \( d \). By assuming that the filter L is a minimum-phase system we are able to calculate any response, but for the factor \( d \). For the response of a short rectangular pulse with an area \( a = \xi \cdot A \) the following expression can be derived:

\[
u(t) = a \, u_f(t) = a \, \frac{\beta}{T_d} \, d \, S \, f(t, E) \quad (1)
\]

\( \beta \) is a constant which is slightly dependent on the shape of the flicker fusion boundary and has approximately the value 2.

Fig. 9: Visualisation of the relation between flicker fusion, flash threshold and latency (see text).
\( f(t, E) \) is a time function with unity as the maximum value, and a shape which depends on level \( E \).

A flash is seen when the maximum of the response reaches the value \( \beta \).

\[
\beta \frac{dS}{Td} = \beta
\]

Thus at threshold

\[
a = \frac{Td}{\beta} - \log \frac{Td}{2S} \quad (2)
\]

This is another way of stating Bloch's law, because only the product \( \epsilon \beta \) appears to be relevant to the threshold value. The formula gives a key to the verification of the theory, because on the right-hand side of the formula only properties of the flicker fusion boundary occur, whereas the left-hand side is the threshold value for short flashes. Rewriting formula (2) and taking the logarithm we obtain:

\[
\log a + \log S - \log Td = - \log \beta \quad (3)
\]

---

Fig. 10: Fig. a and b give the normalised pulse responses of the analogon for \( E = 3400 \) td and \( E = 0.4 \) td respectively. Fig. c and d are responses to long rectangular pulses for the two levels. Fig. e shows the product of amplitude and duration for constant maximum output of a rectangular pulse as a function of its duration.

In fig. 11 the experimental values are compared with the threshold ones. Although the separate terms are large compared with the sum, the fit is not bad. The mean deviation from the theoretical values is 0.3 and 0.15 log unit for the subjects HJM and RK respectively. These deviations are of the order of the day-to-day variation.

Increasing the pulse duration the expression for the response becomes more complicated. If the product \( \epsilon \beta \) is calculated as a function of the duration for a constant maximum output, one obtains curves like fig. 4. In fig. 10 the results of such a calculation are drawn for one level (0.4 td). Going upwards from \( E = 0.4 \) td to \( E = 25,000 \) td, the ratio \( Td/Tc \) varies only from 2.0 to 2.2.
Because of the spread in the results the mean value 2.1 can be taken as a good approximation. In fig. 7 the dotted line is drawn at the theoretical distance from $T_c$. The deviation between theory and experiment is very small here.

The analytical expression for the step response is also rather complicated. Fig. 10 illustrates the responses to rectangular pulses which are sufficiently long to produce two nearly independent successive transients.

The time between the onset of the stimulus and the moment when the step response reaches the threshold value $d$ is the visual latency. The higher the intensity of the step, the shorter the latency (fig. 9). The dotted line in fig. 5 shows how latency varies with the step amplitude.

In order to compare the theoretical values with the reaction time measurements the latency curve had to be shifted vertically by 125 ms, to obtain the best possible fit. The shapes of the theoretical and experimental curves are much alike.

Thus on the basis of the phenomenological model presented, one finds an intimate relation among the three visual phenomena discussed.
Subjective Stroboscopy and a Model of Visual Movement Detectors

Phenomenologically the perception of movement is very akin to that of brightness, colour, heat, etc. During viewing the movement seems to slow down and after viewing pronounced negative after-images of movement can be observed. The phi-phenomenon is another example of an illusion of movement.

Neurophysiologically neurons responding to movement in a particular direction have been known since 1959 (Movement detectors). We occupied ourselves with the problem of how to conceive a simple suitable model to account for the properties of a movement detector.

The detector, in order to respond to movement in a particular direction and not to that in the opposite direction, should be hooked up with at least two retinal receptors R1 and R2, spaced a distance $\lambda$ apart, or to any more central projection of these receptors (fig. 1). The detector is supposed to fire if first receptor R1 is stimulated by the light L and then, after or within a time $\tau$, receptor R2. This simple model would account for the phi-phenomenon and, if the detector is subject to fatigue, to the phenomena of slow-down and negative after-image.

The consequence of this model is that it can be tricked to give a false response. If a regular pattern of black and white bars, with a spacing S slightly larger than the spacing $\lambda$ of the two receptors, is moved in the opposite direction it is possible for receptor R1 to be first stimulated by one bar L1 and then for receptor R2 by the neighbouring bar L2 (fig. 2).

Fig. 1: Model of a Movement Detector (Normal operation)
The movement detector MD is supposed to react if the light L first strikes receptor R1 and then, after or within a time $\tau$, receptor R2, spaced a distance $\lambda$ apart.

Fig. 2: Model of a Movement Detector (Abnormal operation)
A regular pattern L1, L2, ..., with a spacing S slightly larger than $\lambda$ is moved in the opposite direction. L1 will strike R1 first. Then L2 will strike R2. The movement detector now reports a movement in the opposite direction to that of the objective stimuli (subjective stroboscopy).

This would result in the perception of movement in a direction opposite to that of the objective stimulus. Thus an illusion of counter-movement should be observed just as it is under stroboscopic illumination.

This illusion of movement, which we called subjective stroboscopy, was observed indeed and seems not to have been described in literature before.

Rather independent of the angle at which the regular pattern of bars is seen, the counter-movement is seen if the frequency at which the bars pass any point on the retina is around 30 Hz. Lighting level has a negligible influence upon this frequency. The effect is strongest in sunlight.

The effect was first studied with rotating discs with black and white sectors (fig. 3). Later it was also observed with railway sleepers and with regular white stripes on roads.

Fig. 3: The arrangement of nine sector discs
Nine discs with black and white sectors ranging between 15 and 90 were mounted on a board and connected to a driving motor in such a way that they rotated with a speed inversely proportional to the number of sectors. Hence the frequency at which the sectors pass any point is the same for all discs. At a frequency around 30 Hz the counter-movement is seen for all discs.
In viewing the rotating discs many effects which bear definite or probable relation to involuntary eye movements, like saccades and irregular eye movements, were observed.

We now distinguish between 3 types of subjective stroboscopy:

1. $\alpha$-stroboscopy, at 10 Hz: a partial stand-still of the disc.
2. $\beta$-stroboscopy, at 30 Hz: the counter-movement predicted by the model.
3. $\gamma$-stroboscopy, between 40 and 100 Hz: a quivery stand-still of a greyish sector pattern, independent of frequency and possibly due to a cumulative perception of micro-saccades.
It is well known how a subject's metabolism is influenced when he is doing dynamic muscle work. The choice of the criteria for the measurement of the load of these tasks is justified by the knowledge of the mechanism. As a consequence of mechanisation and automation more and more muscle work is taken over by machines, and the workers' task is to check and adapt the performances of the machines. Little is as yet known about the mechanisms that play a role when mainly perceptual and mental demands are made on the subject. Yet there is a constant demand for measuring techniques to evaluate the load of such tasks. Consequently, in the literature a number of methods are described, based on common sense or on some empirical findings. However, the problem of measuring perceptual load is so complex and has so many parameters that we can hardly expect all methods to measure the same thing.

Bartenwerfer (1960) has made some investigations in which he studied the sensitivity and discriminability of several methods. He does not, however, mention whether all these methods yield the same order of tasks with respect to their perceptual load. The order of tasks is of great importance in actual practice. In the experiment to be described here four methods are compared, three of which are based on the double-task principle (Bornemann 1959, Brown 1964), the other being a measurement of physiological data. One well-trained subject performed seven tasks. The results of each of the four methods are compared with the order the subject attributed to the tasks. The load of each of the tasks was evaluated with the aid of each of the four methods. Each situation was repeated five times. More details are published elsewhere (Taverne and Koster 1966, Koster 1966).

The methods used were the following:

1. **Reaction task as a loading task**
   The subject has to react alternatively to two tones presented in random order in a paced condition. If the rate of presentation of the tones is increased, he will increasingly neglect his normal work. This diminishing efficiency is taken as a measure of the load (Schouten et al. 1962).

2. **Tracking task as a subsidiary measuring task**
   The subject has to compensate a continuously changing electric signal by means of a pedal. If compensation is complete no sound will be heard. In the case of overcompensation a high tone, and in the case of undercompensation a low tone is produced. The intensity of the tone depends on the difference between presentation and compensation voltages. If the subject performs the tracking task only, a lower mean difference is found than if he does some other work at the same time. The relation of these differences is taken as a measure of the workload.

3. **Tapping** as a subsidiary task

The subject taps a rhythm with his foot choosing his own rate. If, at the same time, he performs his normal work the irregularity in the rate of tapping increases.

This increase is taken as a measure of the workload. (Michon 1964)

4. **Sinus arrhythmia**

A subject at rest shows a pronounced irregularity in his heart rate. If he executes a perceptual task, the irregularity decreases, the mean frequency remaining approximately constant. This decrease is taken as a measure of the workload (Kalsbeek et al. 1963).

**Subjective impression**

The subject evaluates the different tasks according to his impression of the load.

To compare the different methods three criteria were chosen:

a. Correlation between the results obtained with the various methods and the impression of the subject. Both the measurements and the impressions of the subject allow of a classification according to load. To obtain an idea of the degree of correspondence, Spearman's rank-correlation coefficients were calculated.

b. To see whether these calculations have any practical value, an analysis of variance was made. The ratio of the variance due to the tasks to the residual variance gives an impression of the discrimination strength of a method. If a high discrimination factor is found, the corresponding correlation coefficient is highly important.

The ratios calculated are expressed on a four-point scale ranging from "very good" if the corresponding F value is highly significant, to "poor" if the F value is not significant.

c. As a minimum outcome of such investigations any method should be expected to discriminate between situations of work and rest. In this case, too, an analysis of variance of the data provides a measure for the discrimination.

In the following table the results of the experiments are summarized. For each of the four methods are indicated:

- the rank correlation with the subjective impression;
- the discrimination strength of the method between the tasks;
- the discrimination strength between a rest and a task situation;
- the influence of the measurement on the task performance.

<table>
<thead>
<tr>
<th>Methods</th>
<th>reaction</th>
<th>tracking</th>
<th>tapping</th>
<th>sinus arrhythmia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman's coeff. of rank-correlation.</td>
<td>.68</td>
<td>.61</td>
<td>.86</td>
<td>-.64</td>
</tr>
<tr>
<td>critical value</td>
<td>.71 (n = 7, α = .05)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>correlation</td>
<td>moderate</td>
<td>moderate</td>
<td>very good</td>
<td>very good</td>
</tr>
<tr>
<td>subjective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>impression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>discrimination</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>between tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>discrimination</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with respect to</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>influence on</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>task performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The hypothesis was that the irregularity in the heart rate is decreased whenever a subject is performing a perceptual task. The higher the load, the greater the decrease. Averaged over the tasks the decrease appeared to be very small: about 15 percent. The higher the subjective estimate of the load, the smaller the decrease. This explains the negative rank-correlation coefficient. The correlation of the subjective impression with the results obtained by measuring sinus-arrhythmia is rather high; moreover the discrimination strength appeared to be very satisfactory.

The discrimination with respect to rest, however, is poor, which is due to the fact that the performance of some tasks gives a higher irregularity score than the score measured on the subject at rest. This phenomenon may be dependent on the individual. It can scarcely be attributed to an insufficient period of rest, because particular attention was paid to this point. It appeared that the sinus arrhythmia score and the mean value of the heart frequency were correlated; both yielded nearly the same results.

Each method produces a different order of tasks. The only significant rank-correlation coefficient is found between the results of the tapping task and the order indicated by the subject. None of the rank-correlation coefficients between the methods are found to be significant. So we may conclude that the tapping task yields the best results. To investigate the influence of individual factors, it would be worthwhile to repeat these experiments with several subjects.

**Conclusion**

The methods used to measure perceptual load are based on common sense and empirical findings. Contrary to the physiological load, neither the mechanism nor the factors that influence that mechanism are known. So it is not surprising that the methods used should deliver quite different results. As long as we have only a vague notion of perception and even less of perceptual load, the answer to the question: "Can perceptual load be measured?" is: "No, not yet".

**Literature**


Brown, I.D. 1964 The measurement of perceptual load and reserve capacity. Transactions of the Assoc. of Industr. officers, 14, 44-49.


The Psychological Refractory Period

When two stimuli are presented in rapid succession and a subject has either to react to both or to the second stimulus only, the second reaction time increases with decreasing interval. This increase persists as the stimuli are presented to different sense modalities separately. Davis (1957) and Fraisse (1957) presented one stimulus to the eye, the other to the ear of a subject; in both cases an increase in reaction time was found with decreasing interstimulus interval.

Several theories have been put forward in the literature to explain this phenomenon, which cannot but be of central origin.

One theory, the expectancy theory, states that the subject does not expect a stimulus immediately after another one.

Davis (1965) has published results of an experiment that makes the expectancy hypothesis highly improbable. With experiments of our own we could support and reinforce Davis' arguments (Koster and Bekker, 1966). Another theory has been put forward by Welford (1952). In this a one-channel information transmitting system is postulated in which information transmission cannot overlap. In such an intermittency system information has to be held in store when the central system is dealing with previous information. With such a theory, delays in reaction time at intervals shorter than a normal reaction time can be explained. To account for delays at longer intervals Welford postulates feedback signals from the effectors to the central system.

With experiments in which the subject has to react to either both of the stimuli or the second only, approximately the same delays were found (Davis, 1959; Koster and Bekker, 1966). If present, feedback signals appear to have a negligible influence on the delays measured.

A variant of the intermittency hypothesis is proposed by Davis (1957). He expects the central system to be refractory after the passage of a signal. With this theory he predicts delays both with intervals shorter than the normal reaction time and with longer intervals.

To check Davis' theory we carried out several experiments with some highly trained subjects. As an illustration the data of one of these experiments are given in the figure. Two stimuli were presented to the subject who was instructed to react as quickly as possible to the second stimulus only.

There appears to be a clear discrepancy between the experimental results and the data predicted by Davis' model. Examining our data obtained under different experimental conditions and data of other authors, the same discrepancy was found between experimental and predicted data. It appeared that all data form an exponential curve, as is illustrated in the figure (broken line). This curve-fitting indeed proves nothing, it is only mentioned to illustrate that models predicting an exponential curve of the reaction times seem more adequate than the existing models (see WELFORD 1952, p. 2). Perhaps the analogy with the physiological refractory phase as put forward by TELFORD (1931), is not as loose as some authors suggest. We might assume that the sensitivity of the central system is reduced after the passage of a signal. As time passes, this sensitivity will be restored. A signal arriving during this period will be dealt with in a less efficient way and a delay will result.
Figure: The reaction time of a highly trained subject to the second of a pair of stimuli is plotted (blocks). The full straight lines represent the predicted values according to Davis' model with constant central times (sharp edge) or variable central times (rounded edge). The broken line represents the best-fitting exponential curve.

References

Davis, R. 1957 The human operator as a single channel information system. Quart. J. Exp. Psychol., 2, 119-129.

Davis, R. 1959 The role of "Attention" in the psychological refractory period. Quart. J. Exp. Psychol., 11, 211-220.

Davis, R. 1965 Expectancy and intermittency. Quart. J. Exp. Psychol., 17, 75-78.


A Comparative Study of Human Performance with Dial and Keyset Telephones

Introduction
Since 1958 there have been several investigations concerned with the use and operation of dial and keyset telephones. They demonstrated beyond doubt that selection times with keyset telephones are about 50% shorter than with dial sets. With regard to selection errors made during dialling and keying, the results are not so clear. The laboratory investigation described here is aimed at a more detailed analysis of the selection errors made in the use of dial and keyset telephones under different conditions. An error is defined as the dialling or keying of a combination of exchange number and subscribers number which would have resulted in a wrong connection or no connection.

The telephone sets used were both made in the Netherlands (fig. 1).

Fig. 1: The telephone sets used in the experiment.

*) Student guest worker from the Rijksuniversiteit, Groningen.

1) Based on a paper presented at the Third International Symposium on Human Factors in Telephony.
The telephone numbers, too, correspond to those generally used in the Netherlands.

In addition to selection times and errors, the pulse times involved in the operation of keyset telephones were investigated, and also the influence of using the end-of-selection key.

Assuming that the number to be selected can be read clearly from a telephone directory, the selection errors made can be divided into errors of memory and errors of manipulation.

Operationally these errors are indistinguishable. By altering the experimental conditions one can either emphasize the demands made on the memory or minimize them.

These considerations prompted us to split the investigation into two experiments. In the first experiment the visually presented number was no longer visible when the subject began to dial or key the number, and in the second experiment the number remained visible the whole time. Since both experiments also differ in other respects, it seems desirable to discuss the arrangement and the results consecutively.

Experiment 1

Experimental arrangement

The subject had in front of him a book containing the names and addresses of the subscribers to be called. After having read a particular name, the subject was able by pressing a button to make a part of the relevant page from the telephone directory visible, projected onto a screen (fig. 2).

The telephone number consisted of an exchange number of five digits (the first digit always being an 0) and a subscriber's number, likewise of five digits. In order, with the dialling set, to keep the total dial-return time equal for all numbers selected, each 10-digit number contained, after the 0, all digits from 1 to 9. In this way the mathematical sum was always 45, and the total dial-return time 5.5 seconds.

The subject was allowed to look for as long as he liked at the projected page from the telephone directory, but only once and not while selecting a number. Having observed the number, he lifted the receiver and dialled or keyed the number. Recordings were made of the time of the first dialling tone (i.e. the time between taking off the receiver and dialling or keying the first digit), of the times for the individual digits and for the selected numbers as such. In one session and for each apparatus ten telephone numbers were selected by dialling and keying. The subjects were 22 male undergraduates. None of them had any previous experience with the keyset telephone.

Results

The results of experiment 1 are presented in the table below.

<table>
<thead>
<tr>
<th>Set</th>
<th>Average presentation time</th>
<th>Average time 1st dialling tone</th>
<th>Average selection time</th>
<th>Percentage wrong exchange numbers</th>
<th>Percentage wrong subscriber numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dial</td>
<td>37.9</td>
<td>1.8</td>
<td>20.4</td>
<td>44.5%</td>
<td>21.4%</td>
</tr>
<tr>
<td>Key</td>
<td>38.8</td>
<td>1.7</td>
<td>11.3</td>
<td>43.6%</td>
<td>20.0%</td>
</tr>
</tbody>
</table>

No relation could be found between the presentation time and the total of wrong numbers selected. Nor did there appear to be any relation between the total selection time and the percentage of errors.
**Results**

The results of experiment 2 are contained in the table below.

<table>
<thead>
<tr>
<th>Set</th>
<th>Average selection time in seconds</th>
<th>Percentage wrong numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dial</td>
<td>14.5</td>
<td>3.1%</td>
</tr>
<tr>
<td>Key</td>
<td>6.8</td>
<td>5.2%</td>
</tr>
<tr>
<td>Key + E</td>
<td>7.3</td>
<td>4.0%</td>
</tr>
</tbody>
</table>

The higher percentage of errors with the keyset telephone does not differ significantly from that found with a dialling set. The percentage of incompletely selected numbers was 2.3% for both cases.

In regard to the selection times a training effect is found with both sets (fig. 3).

![Graph showing selection times over 6 days](image)

**Fig. 3:** Mean selection times for 6 subjects during 6 successive days. D = Dial, K = Keyset, K+E = Keyset + End-of-selection key.

The subjects learned the use of the end-of-selection key quickly. It was forgotten in only 1.9% of the total numbers selected. The total selection time increased by 0.5 sec. to 7.3 sec. On the sixth day of this second experiment, 50 pulse times were recorded and measured for each of the six subjects. The frequency distribution of these pulse times is shown in fig. 4.
<table>
<thead>
<tr>
<th>Account bel cons</th>
<th>VERENIGING v KINDERUITZENDING</th>
<th>NEDERLANDSE KATH Heekel 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>38668</td>
<td>32416</td>
</tr>
<tr>
<td>Gasselstr 10</td>
<td>Idem</td>
<td>35723</td>
</tr>
<tr>
<td>3MIJ De Citadell</td>
<td>VER v Rkie GSINZVOODGDLJ en PATRO-</td>
<td>33254</td>
</tr>
<tr>
<td></td>
<td>NAGE Ray's-Hertogenbosch Vught-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>terstr 260</td>
<td>38880</td>
</tr>
<tr>
<td>Vughterstr 51</td>
<td>VEREN tot SAMENSTELLENG van</td>
<td>31912</td>
</tr>
<tr>
<td></td>
<td>NEDERL. BEDEVAARTEN (VND) Bur</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ophoveniui 116</td>
<td>34546</td>
</tr>
<tr>
<td>Huis Markt 17</td>
<td>VERFande GLASIND's-HERT BOGLA</td>
<td>32112</td>
</tr>
<tr>
<td>industrie engros</td>
<td></td>
<td>32308</td>
</tr>
<tr>
<td>1 63</td>
<td>VERGEEEST J J Petelaarssew 112</td>
<td>35020</td>
</tr>
<tr>
<td>14</td>
<td>VERHAAREN J Boekbind Bereeuwstr</td>
<td>34892</td>
</tr>
<tr>
<td>Cerkstr 30</td>
<td>VERJAAGE J Prov' bur v d CEC afd N</td>
<td>35716</td>
</tr>
<tr>
<td>31maal 14</td>
<td>Brah Biomeastr 20</td>
<td>33096</td>
</tr>
<tr>
<td>hospelstr 14</td>
<td>VERHAGEN A, D kapper Vughteirst 11</td>
<td>30731</td>
</tr>
<tr>
<td>genl 105</td>
<td>VERHAGEN A A Loogh bedr Adelheidstr</td>
<td>30731</td>
</tr>
<tr>
<td></td>
<td>38490</td>
<td>35498</td>
</tr>
<tr>
<td>EEs Schaanhuis-</td>
<td>VERHAGEN A M Dir NV P v Hoekel en</td>
<td>36540</td>
</tr>
<tr>
<td></td>
<td>34643</td>
<td>35434</td>
</tr>
<tr>
<td>Ikseonel 102</td>
<td>VERHAGEN H Vughteirst 202</td>
<td>31235</td>
</tr>
<tr>
<td>Orthenborstr 56</td>
<td>VERHAGEN H Passementfabr Emmaple 10</td>
<td>34004</td>
</tr>
<tr>
<td>rstra 2</td>
<td>VERHAGEN M Aann bedr Herven 42</td>
<td>34892</td>
</tr>
<tr>
<td>oremborstr 114</td>
<td>Orthen</td>
<td>37387</td>
</tr>
<tr>
<td>Iksew 83 Orthen</td>
<td>VERHAGEN Th H Aann Schanswetering</td>
<td>31597</td>
</tr>
<tr>
<td></td>
<td>38017</td>
<td>34892</td>
</tr>
<tr>
<td>Markt 12a</td>
<td>VERHAGEN en Zn Fa M Dir keet Adr</td>
<td>32356</td>
</tr>
<tr>
<td>str 548</td>
<td>Willaerstr</td>
<td>30569</td>
</tr>
<tr>
<td>Torenstr 21</td>
<td>VERHAGEN= d MEULEN A P M Gre-</td>
<td>32112</td>
</tr>
<tr>
<td>str 21</td>
<td>goringsumsing 3</td>
<td>32356</td>
</tr>
<tr>
<td>bedr Rijksw 202</td>
<td>VERHAGEN= OORENCOT NV MIJ</td>
<td>36540</td>
</tr>
<tr>
<td>Knusstr 22</td>
<td>tvdzv C L Boschdijkstr 49</td>
<td>35434</td>
</tr>
<tr>
<td>'s Havenfonti 13</td>
<td>VERHEES Am Hert Janstr 9</td>
<td>36540</td>
</tr>
<tr>
<td>Torenstr 18</td>
<td>VERHEES A C Foto-kinoh fotografot en fotop art Huinhamerstr 73</td>
<td>35434</td>
</tr>
<tr>
<td>1</td>
<td>Vughterstr 51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VERHEES B P M Econ des Dommeiffat</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Buitenhaven 64</td>
<td>31235</td>
</tr>
</tbody>
</table>

Fig. 2: Part of a page from the directory. The subscriber to be selected was Verhappen, J.C., his number was 01726-35498.

The percentage of incompletely selected numbers (i.e. where the subject stopped selecting the number, possibly making a second attempt) amounted to 7.3% and 6.8% for dial and keyset telephones, respectively.

Experiment 2

Experimental arrangement

In this second experiment the subject was instructed to dial and key as quickly as possible a series of thirty telephone numbers. The numbers to be selected remained on display during the dialling and keying. The composition of the numbers was the same as that in the first experiment. In order to trace possible training effects, the experiment was repeated on six successive days. The subjects were six male members of the institute staff, with no or only very little experience of the use of keyset telephones. In separate sessions with the keyset telephone the subject was asked to use the end-of-selection key after keying the whole number (Key + E).
Conclusion

Summarizing, it can be said that, while the selection time is about 50% shorter with the keyset telephone, the percentage of errors in the conditions of both experiments was the same for both types of set. The use of the end-of-selection key has no influence on the percentage of errors. There proves to be a difference in affinity to the two types of set, that is to say the subject who dials a number relatively quickly is relatively slower with the keyset. The converse may also be found. This may conceivably lead to a subjective preference for a particular type. The pulse time recordings show that very short pulse times may also occur in operation by ordinary users.

J.J. Andriessen & J. Vredenbregt

Measurements on Pin Assembly Tasks

Of many manufacturing tasks it is hard to indicate if they are more or less difficult and for what reason, even of such a simple task as fitting a pin into a hole. For an important part the difficulty will depend on the clearance between pin and hole. Fitts (1954, 1964) and Annett et al. (1958) have done research work in this field. The smallest clearance used was 11% of the hole diameter. In our experiments, briefly reported before by Schouten et al. (1960), the smallest clearance is 0.05% of the hole diameter. We will call this percentage the relative clearance. The hole diameters used were 16, 8, 4, 2 and 1 mm.
The subjects were requested to assemble the pin into the hole at maximum speech, keeping the pin in a vertical position during the experiments. The starting point of the movement was on a 50 mm dia. circle, concentric with the hole. The time, called assembly time, necessary to transport the pin from the starting point to the bottom of the hole, was measured electronically. From the experiments it turns out that

a. the average assembly time is proportional to the logarithm of the relative clearance and the slope of the curves is the same for all hole diameters.

b. the assembly time is proportional to the logarithm of the hole diameter.

The curves in fig. 1 are nearly equidistant (Andriessen, 1959a, Vredenburg, 1959).

![Graph showing assembly time in relation to relative clearance for various hole diameters.](image)

**Fig. 1:** The assembly time in relation to the relative clearance for various hole diameters.

One has often tried to explain such relations in terms of information theory, as done by Fitts. But the results of the experiments, in which the total assembly movement was recorded in relation to time, show that the theory is not applicable to these results. Measurements of the durations of the successive operations of the pin assembly tasks, viz. transporting, locating and inserting the pin, show that the duration of transport is independent of the relative clearance. A very small influence of the hole diameter upon the duration of transport could be noticed. The time for locating the pin depends on the relative clearance, while the insertion only depends on the absolute clearance (Andriessen, 1959b, Andriessen et al. 1964).

A practical application of the relations found is the interpolation and extrapolation of the assembly table of the Work Factor System (Quick et al., 1962).
In this table are shown a rather rough arrangement of the series of possible diameters of the hole and the ratio between pin and hole. Moreover, the Work Factor System gives only one assembly time for all relative clearances below 6.5% for the various groups of diameters of the hole. In view of the present state of affairs in industry it seems recommendable to us, to give a further differentiation of the table of times mentioned, especially for small pin diameters and hole diameters.

The experiments first described in this report were repeated at various assembly speeds. The relation between assembly time and relative clearance as well as the slope of the curves were found to be equal to the results obtained at maximum speed of assembly. Only the absolute values of the assembly times differ, depending on the assembling speed (fig. 2). Four experiments gave the same results. (Andriessen, 1959c, 1960, 1962).

Fig. 2: The relation between the assembly time and the relative clearance for a 4 mm diameter of the hole, as obtained during experiments at three different speeds of assembly.
They show that when assembly times found at various speeds are converted into standard times at standard speed, this conversion by means of multiplication, though often carried out in the practice of work remuneration, has limited applicability.

References

Fitts, P.M. 1954 The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psychol., 47, 381-391.
Some Aspects of Muscle Mechanics in Vivo.

It is hard to understand human motions without knowledge of the muscle characteristics in vivo. Most of the research in muscles has been carried out on isolated muscles. The data obtained under these conditions cannot be used to predict muscle behaviour in vivo, since the isolated muscle is not stimulated by its own nerve system. Moreover, its natural feedback loops are cut off.

The aim of our investigations is to find a description of the mechanical behaviour of muscles under two conditions, viz,
- static at which a muscle contracts at constant length,
- dynamic at which a muscle shortens during the contraction, causing a movement of the limb.

Extensive experiments have been carried out on the forearm flexors of one subject. The experimental conditions, as well as the apparatus used, have been described by Vredenbregt (1966). The experimental set-up is shown in fig. 1.

Fig. 1: General view of the apparatus for measuring simultaneously the degree and rate of contraction, the acceleration and the force of the muscle at the wrist.
The different phenomena, measured simultaneously during the contraction, are:
- the force, exerted by the muscle
- the acceleration
- the degree of contraction
- the rate of contraction
- the electromyographic activity

The first four phenomena are measured at the wrist in a direction parallel to the biceps muscle. The electromyogram (E.M.G.) is picked up with two silver suction cup electrodes (Vredenbregt, 1966b) and is measured by a specially designed high quality electromyograph.

In the static experiments the forearm is fixed at various angles between forearm and upper arm. In this position the subject has to contract the forearm flexors as fast as possible from zero to maximum effort. The mechanical response, as well as the E.M.G. and the integrated E.M.G. are shown in fig. 2.

![Force-time curve, E.M.G., and integrated E.M.G.](image)

Fig. 2: The force-time curve, the E.M.G., and the integrated E.M.G. as they are recorded during a contraction under static conditions.

It can be noticed that the electromyographic activity is almost immediately at a constant value as shown by the integrated E.M.G., which forms a straight line. The force, however, begins about 20 milliseconds later, compared with the E.M.G. and rises slowly in contrast with the electromyographic activity. The same force-time curves have been found by Hill (1949) and Wilkie (1950). Considering the E.M.G. as a measure of muscle activity and the force as the response to this muscle activity, the shape of the force-time curve clearly shows the existence of elastic and damping properties of the muscle.

Comparing the results obtained at different muscle lengths, three remarks can be made:
1. The maximum force that can be exerted decreases with smaller muscle length which is in agreement with data of Wilkie (1949).
2. The rate of increase of the force is smaller for a shorter muscle (Vredenbregt et al. 1962).
J. The E.M.G. shows the same shape and value, independent of the muscle length.

In the dynamic experiments the forearm is moved during contraction. In fig. J a recording is shown.

Fig. J: A recording of a contraction under dynamic conditions. In this figure the registered force is not equal to the actual one. The actual force is obtained by adding the inertial force - for accelerating the mass of the forearm - to the registered force.

As can be seen from this recording the shape and value of the E.M.G. do not differ from those under static conditions.

Comparing the force-time curves measured under static and dynamic conditions a great difference can be ascertained. Because the muscle length changes during the dynamic contraction, a correct comparison between the dynamic and static force can only be made at corresponding muscle lengths and points of time, counted from the beginning of the contraction. A comparison between the dynamic force and the corresponding static force still shows a considerable difference except at the beginning of the contraction, where the change in muscle length is very small. The difference mentioned appears to be a function of the rate of contraction (Vredenbregt et al. 1961). To account for this difference in forces two possible explanations will be mentioned

1. The difference in force might be due to a different muscle activity.
2. The difference in force is caused by a force acting in the opposite direction of the muscle force.

The first explanation is improbable because, as pointed out already, the E.M.G. at maximum effort has the same shape and value under static and dynamic conditions and is independent of the muscle length. As for the second explanation there are two possibilities or a combination of them. They are

1. Activity in the forearm extensors (triceps muscle) during the passive extension due to the contraction of the flexor muscles.
2. Friction in the muscle. As pointed out already by Buchthal et al. (1951), Roberts (1963) and Raek (1966), friction in the muscle has to be taken into account.

Verification of the activity of the triceps muscle revealed that under static conditions no electromyographic activity could be ascertained. However, during the dynamic contractions electromyographic activity could be established, the magnitude of which depends on the rate of contraction of the flexors. With respect to the E.M.G. we presume that the envelope of the E.M.G. is very similar to the active state curve.

Fig. 4: The value of the E.M.G. plotted as a function of the force, as obtained under static conditions at various muscle lengths.
To evaluate the amount of resistive force from the electrical activity, the relation between force and E.M.G. has to be known. This relation has been determined for the flexor as well as for the extensor muscles under static conditions at steady state levels of activity and at various muscle lengths. The force is measured at the wrist, while the E.M.G. is picked up at the muscle itself. When the level of electrical activity is plotted as a function of the exerted force, various curves are found for various values of the angle $\alpha$ be-

![Graph](image)

Fig. 5: The value of the E.M.G., plotted in relation to the ratio between force and maximum force.
tween fore- and upper arm, i.e. for various muscle lengths. This is shown in fig. 4. Plotting the value of electrical activity as a function of the ratio between the force exerted at various levels of activity and the maximum force at the same muscle length all curves appear to coincide without increasing standard deviation (fig. 5). This shows again that the electrical activity is independent of the muscle length. The force depends on the level of activity and the muscle length. The corresponding relation for the triceps muscle shows the same shape at different absolute values.

In contrast with Lippold (1952) a non-linear relation has been found. Repeating Lippold's experiments on the calf muscles, a non-linear relation between force and E.M.G. was also found. However, the non-linearity of this curve was less than that found at the flexors and extensors of the forearm. Our results are in agreement with those of Buchthal (1942) and Bottomly (1964).

Apart from the active force, the muscle can exert an elastic force due to passive elongation. This relation has been determined by using the weight of the forearm in which the elastic force appeared to be an exponential function of the elongation (Vredenbregt et al., 1966). This relation is similar to those found on muscles and muscle fibres in vitro by various authors and to the result found by Ralston (1947) on muscles of amputees.

Some internal friction in the muscle was also measured during these experiments. When the arm was caused to vibrate at various frequencies, which was done in a pilot experiment, a phase relation between change in muscle length and force could be determined, which also indicates friction.

All the results together will contribute to our describing the mechanical behaviour of the human forearm flexors.

References

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottomly, A.H.</td>
<td>1964</td>
<td>The control of muscle.</td>
</tr>
<tr>
<td>Buchthal, F., Kaiser, E.</td>
<td>1951</td>
<td>The rheology of the cross striated muscle fibre.</td>
</tr>
<tr>
<td>Hill, A.V.</td>
<td>1949</td>
<td>The abrupt transition from rest to activity in muscle.</td>
</tr>
<tr>
<td>Rack, P.M.H.</td>
<td>1966</td>
<td>The behaviour of a mammalian muscle during sinusoidal stretching.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>J. Physiol, 182, 1-14.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Amer. J. Physiol., 151, 612-620.</td>
</tr>
<tr>
<td>Roberts, T.D.M.</td>
<td>1963</td>
<td>Rhythmic excitation of a stretch reflex, revealing (a) hysteresis and (b) a difference between the responses to pulling and to stretching. Quart. J. Exp. Physiol., 48, 328-345.</td>
</tr>
</tbody>
</table>

Vredenbregt, J. 1966a A method for measuring simultaneously the force, the acceleration and the degree and rate of muscle contraction. IPO Minor Report no. 81.

Vredenbregt, J. 1966b Suction cup electrodes for electromyography. This issue.


J.F. Schouten & J.A.M. Bekker

Reaction Time and Accuracy

The fraction of errors in a binary reaction experiment is found to depend upon reaction time. This effect, already apparent in a free reaction time experiment, is corroborated by using the method of forced reaction time. In this method the subject is instructed to react in coincidence with an additional command signal. The binary stimulus then serves as an information signal only.

The dependence of the fraction of errors upon reaction time seems to rule out any theory based upon the assumption that the subject reacts if and when he has obtained a fixed certainty regarding the quality of the stimulus.

In an electrocardiogram of normal subjects little energy is measured in the region above 100 cycles per second (Fig. 1).

An electrocardiogram of a subject with a myocardial infarction shows irregularities ("notches") in the R-waves up to several hundred cycles per second (Durrer 1961, Langner 1963) (Fig. 2).

As these frequencies fall in the sensitivity region of the human ear, we suggested listening to the corresponding sounds. Using the ear as a very sensitive analyser we may expect to be able to distinguish clearly between the sound of a normal electrocardiogram and that of an electrocardiogram of a patient with a myocardial infarction.

In cooperation with Prof. Durrer and Dr. Meyler (Dept. of Cardiology, Wilhelmina Gasthuis, Amsterdam) this hypothesis was tested on a partly infarcted heart of a dog.

Indeed, the R-waves were heard as low bumps when the electrode was placed on the unaffected part of the heart, and as definitely higher-pitched roughish sounds when placed on the locus of the infarction.

Clear differences were also found with physical frequency analysers (Koster 1966).

The high frequency notches are invisible on recordings of normally available electrocardiographs. Since, however, these very notches are audible when reproduced by a loudspeaker or headphone we suggest "electrocardiophony" (listening to the higher frequencies of the electrical phenomena in the heart) as a simple and powerful tool for the diagnosis of myocardial infarction. Since the signal is audible it can be used in telephonic communication between cardiologists by placing the headphone in front of the microphone.
In a working group on the industrial impact of ergonomics the author put forward the following considerations.

For many years the Mechanisation and automation of human work have been referred to as the first and the second industrial revolution respectively. Both are as old as industry itself and both have, time and again, been the cause of resentment of the workers.

In order to treat this problem in a more balanced way the concept of humanisation of human work should be included in the considerations. This humanisation, in terms of better working conditions, working times, social security etc. became economically feasible thanks to the increase of productivity, which in its turn could be achieved by the proper introduction of mechanisation and automation.

Hence, it is suggested to talk in terms of Industrial Evolution, the cornerstones of which are the mutually related mechanisation, automation, and humanisation of human work.

Fig. 1: Diagram of three aspects of the evolution of human work
This evolution is symbolized in the diagram. Each of the three aspects, be it the introduction of a new machine, of an automaton or of new working conditions, affects the other aspects, be it positively or negatively. If industry manages to develop the three aspects in mutual balance, then one may speak of industrial evolution, if not of industrial revolution in the ominous sense of the word.

(*) Abstract of "De Mechanisering, de Automatisering en de Humanisering van de Menselijke Arbeid", I.F.O. Memorandum No. 37, January 1964.
6. INSTRUMENTATION & TECHNIQUES

G.J.J. Moonen & C.A. Lammers

A Preset Cascade Counter

The preset cascade counter has been designed to be used as an impuls-pattern generator or as a timing device. It is a fully transistorized decimal counter with a capacity up to $10^4$ periods. Nine outputs are available, each of which can be set to one of the $10^4$ counting positions. The 10th output makes it possible to reduce the counting capacity to values between $10^4$ and 10 (variable feedback). The counter can give an impulse pattern either repeatedly or single shot. In the latter case it is possible (if necessary by remote control) to start a new pattern.

The timing or clock frequency must be fed externally and may be 10 kc/s at most. The driving amplifier is D.C. coupled with the input, so that the cascade counter can be driven with very slowly shifting signals.

For a quick survey a display is coupled to all 40 flip flops as well as to the 10 outputs.

The output amplifiers are built to provide a signal and its complement. Therefore it can be coupled to transistorized units as well as to tube equipped units. In the following figure the block diagram is given. The actual counter receives its driving frequency through AND-gate g. Flip flop G determines whether or not amplified input signals from amplifier M are fed to the counter, i.e., whether the fed impulses are to be counted or not.
The 10 amplifiers (emitter-followers) of every decade are connected to ten 10-position switches, each consisting of 4 sections which can be set independently. One section of every switch is connected to the 10 unit amplifiers of the counter; another section of every switch is connected with the tens and so on for the hundreds and thousands. Now every switch can be set to a position between 0000 and 9999. The four common contacts belonging to each switch are connected to the corresponding AND-gates (a-k). So when the counter is in the position corresponding with the switch position a signal is generated at the output of the AND-gate. By changing the position of the switch the signal will appear at another moment.

With the ten sets of switches ten impulses can be given at any position between 0000 and 9999. The output signals of the AND-gates a-k are fed to an amplifier followed by a long-tailed pair circuit. To keep the diagram within manageable bounds the long-tailed pair belonging to switch K has been omitted. The long-tailed pairs provide complementary signals to short-circuit proof emitter followers, which are connected to the 20 outlets.

Switch K with gate k have the added function to shorten the counting range (counting capacity). By setting the switches SW 1 and SW 2 in the position drawn, the counter will stop when the position given by the switch K is gained, since the signal of gate k will reset flip flop G, thus interrupting the clock frequency by means of g. The intended cycle has then been completed once. To start the next cycle it is necessary first to reset the counter to 0000. By pushing the reset-start button the reset amplifier is operated, thus bringing the counter in 0000 position and one-shot multivibrator R is operated.
The reset amplifier also resets flip flop G, which has no effect because flip flop G is already reset by the signal from gate k.

After the delay time of one-shot multivibrator R flip flop G is set again, which opens gate g and the next cycle starts. Should one wish to repeat the impulse pattern continuously then switch SW 2 is switched to the other position. The signal of gate k will then operate the reset amplifier immediately and the next cycle is started as before. With switch K in position 999 it is not possible to count any further than 9999. If counting till $10^4$ periods is required, switch SW 1 has to be switched over but as a consequence no reset will be obtained. It is evident that the output signal at each output is only present as long as the counter is in the position corresponding with the switch concerned.

Common practice has shown that it is desirable to incorporate one or more flip flops which can be operated by the output signals of the counter to obtain a signal of any duration within the limit of the counter.

Recently available switches, with built in codes (Binary Coded Decimal switches) make it possible to perform the same function with 4 flip flops per decade instead of 10 flip flops present in the subsequent version.

D.J.H. Admiraal

Segmentators for Sound Analysis

At our Institute interest is shown from the phonetic point of view in segmentation of the speech continuum. Thus e.g. one would like to be able to analyse the diphthongs into their constituent elements by applying some technique of segmentation. This calls for the need of an instrument that enables the experimenter to vary time parameters of an acoustic signal.

In view of such studies several comparable segmentators have been constructed in many laboratories with the aid of which it is possible to make audible only a segment i.e. a small part of a sound, for instance a spoken word. Most of these instruments are rather complicated and never leave the laboratory stage.

In our laboratory two types of segmentator have been made in the form of simple and easily adjustable instruments.

Of the many possible applications we may mention: phonetic speech analysis, for which purpose the apparatus was in fact designed, dialect studies, language instruction, speech training and generally in all cases where it is desired to isolate part of an acoustical or electrical vibration pattern.

The sound or the word in question has been recorded on an endless tape with the help of an ordinary tape recorder, so that the signal to be analysed is fed to the segmentator with a regular repetition frequency.

Just before the word is recorded a tap is given on the microphone with a pen or pencil to record a pulse on the tape, the marking or timing pulse. Then the word is recorded and the recorder immediately stopped.

When the tape is played back, the volume control on the recorder is turned up until a lamp on the front panel of the segmentator becomes brighter. What now happens is shown in fig. 1. Fig. 1a shows the time function of the complete signal on the tape, as this is fed to the segmentator. The timing pulse triggers a one-shot flip-flop, called the "guard", which is so adjusted that its switch-over time (750-2,500 ms) is slightly less than the cycle time (about
2,200 ms) of the tape loop, fig. 1b. This bridges parasitic pulses and provides a fixed starting point. The action of this device is controlled by the above-mentioned lamp.

In the Analogue Segmentator, the simpler of the two, the "guard" triggers a second one-shot flip-flop, called the "position", fig. 1c, which determines the place in the word where or the moment when a gate opens to allow a segment to pass, consequent upon the triggering of a third mono (switch-over time 35-330 ms) by the trailing front of the position pulse, fig. 1d. The square-wave output voltage of the gate mono is rounded off somewhat (RC time = 4 ms) and this modified signal, fig. 1e, controls the gate which thus opens and closes gradually, so that switch-over clicks are not audible. The output of the gate, fig. 1f is coupled via an adjustable attenuator (volume control) to an audio-frequency amplifier with a loudspeaker.

---

**Fig. 1 Operation of the analogue segmentator**

*Left: time-sequence diagram  
Right: block diagram*
A special monitor knob has the following features and advantages:

**Top position.** Gate permanently open. In this position the recording can be checked. One can also check whether the lamp of the "guard" does indeed light up after the timing pulse, which is now audible.

**Middle position.** The segment is let through unattenuated, but the signal before and after this segment is attenuated by 20 dB. This feature enables the experimenter to establish qualitatively and quickly the position of the gate within the cycle.

**Bottom position.** Only the segment is let through, all other signals preceding or following it being attenuated by a good 80 dB.

Philips is making a provisional series of 100 of this type of segmentator.

A more complicated instrument especially developed for laboratory use is the Digital Segmentator.

The above-mentioned position mono is so adjusted that it resets at the moment the word begins. Between this mono and the gate mono a digital device has been added consisting of two decimal counters with slightly different counting capacities. By these means the segment automatically shifts through the word at a speed depending on the difference in capacity of the counters.

The instrument has a positive pitch (forward travel) of maximum $5 \times 10^{-3} = 50$ ms per cycle and a maximum negative pitch (backward travel) of $2 \times 10^{-3} = 20$ ms per cycle.
The position of the audible segment with regard to the beginning of the word is shown in tens of milliseconds by counting tubes.

The rise and fall-off time of the gate control voltage can be adjusted over a wide range. With this automatic segmentator the attention of the experimenter can be fully concentrated on the perception of the sound.

Digital Segmentator

References


Introduction

In determining a difference limen, binary choice methods have the advantage of greater reliability and accuracy than the method of adjustment. (cf. Cardozo, 1965) They tend to be time-consuming, however.

A sequential up-and-down method has been devised with the object of primarily, cutting down the time consumed by the subject, secondly, presenting a game situation to the subject that keeps him interested in his own performance and, thirdly, avoiding the necessity of an experimenter (or an automaton) continuously changing the stimulus.

A number of statistical strategies have been proposed for an efficient estimation of psychophysical limens. Without pretending to give complete coverage of the field, mention must be made of Wetherill's paper (1963) and an internal (B.T.L.) report by Levitt (1964).

The present method is merely another strategy. It resulted from practical experience with psychophysical experiments. In it some ideas of the up-and-down method (cf. Dixon and Mood, 1948) and of Wald's sequential sampling technique have been combined, but admittedly, not according to maximal statistical efficiency. Nevertheless it is felt that the method is interesting from a practical point of view.

Practical Description of the Method

Fig. 1: The panel for indicating the odds of the subject.
The subject is seated in front of the panel shown in fig. 1. His left hand rests on a box (not shown) containing five push buttons. These have the following functions (from right to left):

1. (Thumb) : Randomization of the order of the stimulus pair.
2. (Index) : Presentation of the stimulus pair.
3. (Middle finger) : Response: Standard stimulus preceded the variable one.
4. (Ring finger) : Response: Standard stimulus followed the variable one.
5. (Little finger) : Reset the panel for a new game.

In the initial position the panel shows a light in its upper left corner. A correct response takes this light one position to the right, an incorrect response takes the light one slanting step to a position in a lower row (cf. fig. 3). After 4 to 12 responses the light will reach a border position. Three of these have green lights indicating that the subject has won the game. He then proceeds to a more difficult value of the stimulus parameter. The lower border positions have red lights and indicate that the subject has produced too high a proportion of incorrect responses. Having lost the game he must go back to an easier stimulus. One yellow position along the border marks the transition from red to green. It tells the subject to reset the apparatus and to repeat the game with the same value of the stimulus parameter.

According to success or failure in the games, the subject makes steps up and down along the scale of the stimulus parameter, which will be designated by $X$. The starting value $X_0$ is taken so as to present a very easy situation to the subject. After having successfully terminated the first game, the next value of the stimulus will be $\frac{3}{2}X_0$. In general, after $n$ successes, a stimulus $X_n = X_0 / 2^n$ is chosen.

N.B. This system of allotting stimulus values presumes that $X = 0$ (both stimuli equal) will correspond with 50 per cent correct responses of the subject.

An example is shown in fig. 2. Plus signs and minus signs represent successes and failures of a subject in a psychoacoustical experiment. Zeros correspond with games ending in the yellow position. The difference limen is taken to be the mid point of the transition from plus to minus. Usually one determines a series of consecutive transitions. The average mid point position then serves as the difference limen.

![Diagram](image)

Fig. 2: An unusually long up-and-down chart suggestive of improvement by training during the experiment. The subject had to discriminate a regular filtered pulse train (100 pulses per second) from one in which just one of the impulses was delayed by the amount indicated in the left.

1) This panel was built by Mr. G. Domburg.
It is not unusual to find some indication for improvement of the subject's performance as the number of games increases. However, with thoroughly trained observers this is not so.

The apparatus, as shown in figure 1, contains some extra features, e.g. counters for the number of trials, for the number of incorrect responses in one game and an alarm circuit which tells the subject when he has made three mistakes in succession. This a phenomenon unlikely to occur. Normally the subject is being tested in situations with a probability of an incorrect response of about \(\frac{1}{64}\). Therefore a triple fault \((P = \frac{1}{64})\) is a significant occurrence, probably indicating a lack of attention.

**Rationale of the method**

An up-and-down method is very economical with the total number of responses. It has the disadvantage, however, of changing the stimulus parameter after every response. This usually takes time. It also may confuse the subject. For this reason samples of a reasonable length were thought preferable. Now one would like to avoid taking a large sample at an uninteresting value of the stimulus parameter. This is the idea behind the 'sample graph' that is presented in fig. 3.

![Sample Graph](image)

**Fig. 3**: The sample graph. Arrows indicate a possible path of the subject, leading to a successful ending of a game.

Now suppose a subject has an inherent probability of giving a correct response of 0.50, say. Then his chances for arriving at A, B and C (cf. fig. 3) respectively are 0.031, 0.020 and 0.012, totalling 0.06. So even when the subject is completely unable to perceive any difference between both stimuli, he has a small chance (0.06) of unfairly winning a game. Similarly a subject with an inherent probability of 0.90 has a small chance (0.03) of wrongly losing a game. The risk of making a wrong decision is characteristic of any method operating with small samples.

Figure 4 shows the operating characteristic for the sample graph depicted in figure 3. The solid line represents the chance of winning a game given a certain inherent probability that the subject makes a correct response. The shaded areas correspond to the risk of making a wrong decision, considering the value of the inherent probability

\[ p(x) = 0.71 \]

to be the proper one for the difference limen. This choice will be accounted for later.
Fig. 4: Operating characteristic of the sample graph. The full line represents the chance of the subject for winning a game, given a fixed inherent probability of a correct response. Shaded areas indicate the cumulative risk of deciding wrongly, based on a uniform density of the subject along the abscissa. The broken line depicts the average number of trials in a game (values to be read along the right border).

It may be remarked in passing that the operating characteristic of games consisting of just one trial is the main diagonal of figure 4. Its shaded area (with respect to \( p = 0.75 \)) would be one fourth part of the entire square, which is but slightly more than the shaded area of the present operating characteristic. This may serve to indicate how slowly the discrimination of a sample improves as the length is increased.

In fact, the shaded area is valid only if \( p(x) \) is distributed uniformly along the abscissa. Now this is by no means the case, as it is the very object of the up-and-down method to lead the subject to a position near \( p(x) = 0.71 \).

Let us assume that the inherent probability of giving a correct answer (this is a so-called psychometric function) can be approximated by the function

\[
p(x) = 0.5 + 0.5 \tanh x
\]

If we now accept as a starting value of the stimulus parameter \( x_0 \) such that \( p(x_0) = 0.99 \), then we find \( p(x_1) = 0.90 \), \( p(x_2) = 0.75 \), \( p(x_3) = 0.63 \), \( p(x_4) = 0.57 \), \( p(x_5) = 0.53 \) etc. A brief relaxation computation yields \( p(d) = 0.72 \) where \( d \) is the difference limen, viz. the value of \( x \) determined according to the rules of par. 2.

More realistically, the entire method has been gauged with what might be termed a Monte Carlo technique be it in the simple form of a pencil and a table of random numbers. From ten 'experiments' an average difference limen was found to be

\[
d = 0.70 \pm 0.08
\]

On the average the experiment consisted of 4 games taking somewhat less than 40 trials.

When increasing the average number of games to about 14 the difference limen was found to be

\[
d = 0.70 \pm 0.05
\]
It is very likely that the coarse gradation of the x-axis is responsible for the slowness of the improvement in accuracy.

In practice, it is felt, 40 trials will not do for a good performance of the subject so that usually some 10 and even more games are played. On the other hand it appears reasonable that the best way to get a more accurate difference limen is to spread the experiment over more than one day because it is of little value to have a difference limen statistically dependable within a few units while the subject himself may change from day to day by many more units.

References


D.J.H. Admiraal

A Sound Spectroscope with High Scanning Frequency
In the sound spectroscope in use at the I.P.O. and described in Philips Technical Review, 24, nr. 12, 1959/60, the rectified AC voltages of 79 LC circuits, each with a quality-factor $Q = 16$ and with a ratio between every two resonant frequencies of $1 : 1.06$ are scanned with a repetition frequency of 25 Hz by means of an electronic switch, composed of 79 OR-gates, driven by a ring counter.

The purpose of each rectifier is to store the information present at each circuit because evidently the scanning time is not infinitely small. The rectifier has a memory function. As a scanning-frequency of 25 Hz has been chosen, the scanning time is 40 ms. The RC-time or memory-time of each rectifier also has been chosen 40 ms.

In practice, it has been found that this is not the best solution because the filters for the highest frequencies to be analysed have rise and decay times much shorter than the RC-time of the rectifiers.

At input voltages of the spectroscope of short durations or at voltages of which the spectrum as a function of time changes quickly (speech), the level of the output voltage of a gate of the electronic switch depends strongly on the moment the gate opens. The low scanning frequency can result in a discrepancy of about 60% as a maximum as shown in fig. 1.

![Fig. 1: Rise and decay time of the voltage of a filter circuit. The rise-time $t_r$ and the decay-time $t_d$ decrease according as the resonance-frequency of a filter increases. For the filters tuned to the highest speech frequencies, $t_r$ is very small compared with the cycle-time $T$. The value or the level of the voltage passing through the gate depends on the moment it opens. If this is the moment the AC voltage reaches the top, a good representation of the amplitude is obtained. However, when the gate opens nearly a full cycle-time later, the DC-voltage at the output of the gate can be almost 60% too low. By increasing the scanning frequency this discrepancy can be diminished.](image)

A new spectroscope has been built with a scanning frequency of 250 Hz. The cycle-time here is 4 ms, small indeed in comparison with the rectifier RC-time of 40 ms.

**Block diagram**

Fig. 2 shows the block diagram of this new instrument. Some details will be discussed here.

The output is fed to a large-screen oscilloscope (20 x 30 cm), Airmec, type 279. Since with these dimensions the ratio between the maximum deflection (= height of screen) and the just noticeable deflection is much higher than on a normal but much smaller screen, special attention is paid to using a sensitive rectifier with low threshold.
Fig. 2: Block diagram of Sound Spectroscope with high scanning frequency
Rectification of the voltage at a filter

Each filter F has its own rectifier RF. The circuit of a rectifier is given in fig. 3.

![Diagram of filter, rectifier, OR-gate and a driving flip-flop of the ring counter.]

Fig. 3: Circuit diagram of filter, rectifier, OR-gate and a driving flip-flop of the ring counter.

Via a complementary emitter follower with transistors T1 and T2 the AC voltage of the filter is fed to the rectifier. The coupling capacity \( C_c = 10 \ C_m \) blocks the bias of the conducting transistor T2, so that with no AC input, the output of the rectifier is exactly zero.

With \( C_c \) shortcircuited the strongly temperature-dependent bias of the emitter follower will reach \( C_m \). Owing to transistor properties the bias voltages of all emitter followers differ somewhat, resulting in a strong quantising-noise after being scanned.

We have used an active diode as rectifier, i.e., a transistor having its base and collector shortcircuited. As current-amplification is maintained, the internal resistance is much lower compared with that of a passive diode like the OA5 for instance. This results in a considerable gain in sensitivity which is shown in fig. 4.

The electronic switch

This switch contains 86 OR-gates with 4 additional gates bridging the long fly-back time of the oscilloscope. This point will not be further discussed here.

The number of filters in the new instrument has been added to, as compared with the previous one to get an extension of the wave-band to be analysed (84 - 11,400 Hz), resulting in 86 filters.

All outputs of the gates are connected to a resistor \( R_o \). The gates are controlled by a ring counter of 90 bistable flip-flops FF, driven by a multivibrator of 20 kHz.

Each gate can be built up with diodes as in the older spectroscope. However, this method results in loss in DC amplitude and in a considerable threshold interfering with the great sensitivity of the rectifier.

We have designed a gate without amplitude loss and without threshold by using a transistor as a "floating switch", termed thus because the two points to be connected are free of earth, i.e., floating.
Fig. 4: Comparison of rectifier efficiency for a passive diode like the OA5 and an active diode OC 141. Especially for low AC voltages the latter is much more sensitive.

Fig. 3 shows the circuit diagram. When the base voltage of each gate-transistor reaches +12V, i.e., a level higher than the rectifier can detect, the gate is closed.

When the gate is open a base current flows to the -1V supply. The transistor is now in bottoming and since emitter and collector connections have been inverted (inversed with) the bottoming voltage is very small, viz. about 10mV, with small deviations. Therefore the quantising noise at \( R_0 \) is also small (amplitude a few mV).

When \( V_{CM} \) increases, \( I_e \) increases too. Owing to the fact that \( I_b \) also rises, the transistor remains in bottoming at any value of \( V_{CM} \). After passing through a low-pass filter (\( f_c = 20 \text{ kHz} \) - which suppresses switching peaks - an emitter follower and a mixing circuit, the voltage at \( R_0 \) is fed to the vertical amplifier of the oscilloscope.

An adjustable portion of the filtered signal is DC-amplified and then brought to the Z-input of the oscilloscope. When the amplitude of the signal at \( R_0 \) rises, the intensity or brightness of the spot rises too (Z-modulation).
The display

The new spectroscope has four possibilities of display which can be chosen by using the switch S, fig. 2.

Position 1

One line at the bottom of the screen is traced, since the voltage of the Y saw-tooth generator is shortcircuited.

The X saw-tooth generator gives the horizontal deflection of the spot.

Flip-flop FFa has been locked. The signal is fed to the oscilloscope without attenuation.

This kind of display is useful when maximum sensitivity is required and in all cases where the spectrum of the signal as a function of time is nearly constant.

Position 2

The shortcircuited of the Y saw-tooth voltage is now removed. As a result of the introduction of a scaler of 256, triggered by the reset pulse of the ring counter, a frame of 256 lines is formed on the screen with a frame frequency of 1 Hz.

In this position the signal amplitude is attenuated by a factor 4. This kind of display (according to an idea given by S.J. Campanella, Speech Communication Seminar, Stockholm, 1962. C1: A Spectrographic Study of Formant Tracking) is very useful in analysing signals with spectra which vary quickly as functions of time, e.g. speech.

The peaks are given at different heights or places in the frame, which results in a much better discrimination than is possible in position 1, enlarged as it is by the Z-modulation.

Position 3

As 2, but without signal deflection. Only the Z-modulation is effective and causes strokes of light on the screen which are comparable with the information of the Kay Sonagraph.

Position 4

The signal attenuation is now a factor 4 again. The flip-flop FFa is no longer locked. When the scaler resets to zero position, the flip-flop resets too and at the same moment the Y saw-tooth generator stops. The line is now at the top of the screen in the "waiting" position.

When FFa is triggered by the incoming signal or by an external pulse, the AND gate, controlled by the flip-flop, opens. The scaler starts the vertical saw-tooth generator.

When the scaler resets the Y saw-tooth generator stops and the line returns to its waiting position, etc.

This kind of display is very useful for the analysis of signals with a duration no longer than 1 sec., or of signals with a preceding timing pulse such as used with the segmentator (described elsewhere in this issue).
IPOVOX II: A Speech Synthesizer

L.F. Willems
Introduction

With our previous speech synthesizing machine - the IPOVOX I - we were able to produce simple words by placing in succession elementary segments having the size of a phoneme or part of it. The spectral content of the segments, however, was constant. In designing the IPOVOX II emphasis was laid on having formant transitions and on the ease of handling the machine.

General description

A parallel three-formant synthesizer was chosen for the voiced speech segments and a separate channel for unvoiced consonants. If in a segment the formant locations of the first two formants are different from those in the previous segment, then these new formant locations are reached in a smooth way. The time constant of this formant transition is electronically variable and the formant transition is triggered at the beginning of a segment.

In general, all parameters needed for the synthesis are stored in a digital memory, from which analogue circuits are controlled via digital-to-analogue converters or diode decoding networks. This flip-flop memory contains information for 5 segments and each segment occupies 44 bits. Information for the segments can be put into the memory one after another via a read-in desk.

Block diagram

Fig. 2: Block diagram of IPOVOX II
The output waveform of the buzz source is approximately a sawtooth one. The first formant filter has a constant bandwidth of 50 Hz and the resonant frequency can be set at 16 discrete values between 200 and 1,200 Hz. The second formant filter has approximately a constant Q of 8 and the resonant frequency ranges in 16 steps from 700 to 4,200 Hz. The third formant filter can be switched on or off the circuit and the resonant frequency can be set manually between 1,000 and 5,000 Hz with a constant bandwidth of 150 Hz. The noise formant filters have fixed centre frequencies of 2,200, 2,100, 4,500 and 5,500 Hz. They can be switched into the circuit via a diode decoding network and electronic gates.

Several control voltages are required: for the envelope, the formant transitions and the intonation contours. In general, they are produced by ramp generators. The envelope function is described by three parameters (see fig. 2).

- $t_1$: the rise time can be chosen in 16 steps from 2.5 ms to 150 ms;
- $t_2$: the duration (as we define it) ranges from 10 ms to 200 ms in 16 steps;
- $t_3$: the decay time variable in 8 steps between 3 ms and 150 ms.

**Fig. 2**: Segment envelope defined by 4 time parameters.

**Fig. 3**: Illustration of a formant transition function.

Another time parameter, $t_4$, is important in the operation of the machine. This delay time (ranging from 1 ms to 250 ms) is triggered at the end of $t_2$, and at the end of $t_4$ the following segment is started. At this moment (A in fig. 2) information in the memory is shifted, and thus also all parameter settings.

The formant transition function (see fig. 3) has the same time constant for both formants ($F_1$ and $F_2$) and the values $t_f$ are 10, 20, 30 and 80 ms, values obtained when 80% of the target position has been reached. The formant transition starts at the beginning of a segment (the above mentioned point A). The intonation contour can be built up from a continuously expanding inventory of elementary patterns. At the moment this inventory consists of the following patterns.

- A declining pattern over the whole utterance, i.e. a gradual sloping down, starting at the beginning of the first segment.
- A rising pattern during 100 ms.
- A falling pattern during 67 ms.
- The caesura pattern, which falls for 50 ms and then rises for 90 ms.
- The questioning pattern, which falls for 50 ms and then rises until the end of the utterance is reached.
- The pattern of a syllable with an initial voiceless consonant. This pattern has a sharp rise for 20 ms (which in the synthesis coincides with the voiceless consonant and is, consequently, not heard) followed by a falling contour during 200 ms.
- The pattern of a syllable with an initial voiced consonant. This pattern rises for 100 ms and falls for 150 ms.
All these patterns except the first can be triggered at 8 points to choose in the $t_2 + t_4$ interval. In the block diagram, one block has not yet been referred to, viz. mixing facilities. There are several possibilities as regards the choice of the source and the decision which path in the circuit should be used.

- Periodic source to the formant filters.
- Periodic source to the first formant filter only.
- Noise source to the formant filters ($F_1$, $F_2$ and $F_3$).
- Noise source to the noise formant filters.
- Noise multiplied by the periodic waveform to the formant filters (voiced fricatives).
- No source connected to the synthesizer, but an external signal can be fed into the envelope multiplier.

Circuit details

The resonant frequency of the formant filters is controlled by a voltage. This is accomplished by a multiplier and an inductance in the feedback loop. For the resonant frequency of this circuit one finds

$$fr = \frac{1}{2\pi} \sqrt{\frac{1 + A}{LC}}$$

By connecting two multipliers in cascade one has an almost linear relationship between $fr$ and $A$.

The memory is in fact a large shift register. Information in the uppermost register is not destroyed, when new information is shifted in. But the old information is automatically transferred to the lowest register, so that after a cycle of the five segments the situation is the same as at the start.

One of the five registers can be marked by a special bit in the memory. This mark denotes that the information at that moment present in the read-in desk is being used and not the information in the register. This means that one of the five segments can be varied throughout the experiments. This proved to be a very useful feature.

The Intonator

L.F. Willems

The Intonator, as we have termed this instrument, is a tool for studying the perceptual tolerances of intonation contours. From natural input speech only the intonation is changed according to rules specified by a certain model ("t Hart, 1966).

Use is made of the Vocoder principle, known in analysis-synthesis telephony (Cooper, 1961). Actually, the most difficult part of a vocoder, the pitch extraction, is avoided and at the synthesis side the pitch movements are controlled by a function generator.
The spectrum of the input speech signal is analysed in a set of 25 band-pass filters (simple resonant parallel LRC-circuits), covering a range of 250 to 4,000 Hz. The resonant frequencies of the filters are a full tone apart and the filters have a quality factor Q of 16. The output of each filter is connected to an envelope detector. At the analysis side of the system there is also a voiced-unvoiced detector, using the average number of zero crossings as a criterion (Fant, 1962). At the synthesis side the original spectrum of the input speech is reconstructed. A buzz source or noise source is switched on to the synthesizing filters by the voiced-unvoiced signal. The signals from these filters are multiplied in a set of modulators by the corresponding envelopes. This re-made speech is monotonous. By controlling the periodicity of the buzz source with a special function generator one is able to give the speech output specified pitch movements.
Control of the Fo

The control function of the Fo consists of two elements, called macrointonation and microintonation. They will be described in this order.

In the macrointonation function several elementary patterns can be generated. First there is a gradual sloping down during the whole utterance, and this is called the declination. At a certain moment (governed by a manual setting on a timing device, which is a preset counter) a rising function can be generated. At another moment a falling function can be produced. These elementary patterns, rise and delay, can be generated repeatedly. These ramp waveforms are shaped by integration.

The microintonation is derived from the envelope of the input speech (see fig. 2).

![Fig. 2: Intonation patterns](image)

The concept of microintonation is an old one in our laboratory, and was used originally to give some pitch inflection to synthesized words. It was found that naturalness of synthesized utterances improved greatly by controlling the Fo with a certain fraction of the envelope of the speech segments.

A useful trick is applied to improve the intelligibility of consonants. The high-pass (2,500 Hz) filtered part of the input spectrum is added to the output. The spurious pitch information of the input speech in this part of the spectrum is totally masked at the synthesis side by the new pitch information in the lower part of the spectrum. This was found by R.J. Ritsma in his experiments on pitch perception (oral communication).

References


't Hart, J. 1966 Perceptual Analysis of Dutch Intonation Features. This issue.

D.J.H. Admiraal & M.A. Alewijnse

An Infrared Pupillometer based on the Television Scanning Technique

The response of the pupil of the human eye to light (white or coloured, static or dynamic) is one of the subjects of interest of the visual group of our institute. For instance, one wants to know the contraction and widening as functions of time when the eye is stimulated by a light pulse or light pulses of different durations.

In the course of time many systems have been invented and used to measure the pupil diameter (Van Bussel, 1963). From the literature collected it may appear that the system developed at the I.P.O. is affiliated with techniques applied by Asano et al. (1962) who used an image converter instead of the direct method of the infra-red Plumbicon television camera tube.

On the basis of the scanning technique, used in television, a pupillometer has been developed in our laboratory, which possesses some important and characteristic features. To offer infrared light to the eye, use is made of an infrared Plumbicon television camera tube. This light does not contract the pupil, so the measuring light do not interfere with the stimulus light. An entirely different scanning technique, also with infra-red light, has been developed independently of our work at the Institute for Perception RVO-TNO, Soesterberg (Troelstra, 1965) which is based on a Nipkow disk giving rise to a frame of a hundred lines.

As a result of the great number of scanning lines (312,5) the resolving power is high so that a high accuracy can be reached.

As a result of the great number of frames per second (50) quick changes in the diameter of the pupil can be followed and detected.

Finally, the pupil diameter can be read directly on a decimal counter with printer, while display on a slow-motion oscilloscope or registration on a pen-recorder is possible as a digital to analogue converter has been built in.

The illumination of the eye

Principally, by introducing a semi-transparent mirror, it is possible to offer the stimulus light on an eye which is already illuminated by the infrared light.

In practice, however, it is more convenient to use both eyes, one for the infrared light and the other for the stimulus light. This is allowable because both pupils are coupled in such a way that if light strikes one eye causing its pupil to contract, the pupil of the unilluminated eye also contracts at the same time and in the same ratio.

Before discussing the block diagram of the instrument, something must be said about how the video signal is formed and the corrections of this signal necessary to operate a counter.
Fig. 1: Scanning the iris (shaded) and the pupil (black) of the eye by means of a frame of lines.

The formation of the video signal

In Fig. 1 the black dot indicates the pupil, which is surrounded by the iris (shaded area). Both are scanned by a field or frame of 312,5 lines some of which are shown in the figure.

The momentary value of the output voltage of the camera tube is a measure for the level of the reflected-light intensity at each moment. Each line is a "level line". A level line has a top (black level) as long as the pupil is scanned, (fig. 2).

In this figure the blanking of the line and frame is also given. During these periods the scanning beam in the camera is suppressed and the level is then said to be "blacker than black". Thus these tops are even higher than the tops of the pupil pulses. The latter must be selected from the video signal. So this signal must be "cleaned".

The "whitemaker" circuit

An important operation takes place in the "Whitemaker". This is a kind of gate circuit. When it is open the video signal can pass. When it is closed, the video signal at the output is strongly negative (white), independently of the momentary value of the video signal at the input.

The whitemaker circuit can be controlled by any suitable voltage. When it is controlled by the inverted line and frame pulses available at the synchronisation unit of the closed circuit TV system, the interfering pulses in the video signal are transformed into non-interfering white levels (fig. 3).
The DC-restorer

The DC-component of the video signal thus purified is removed by a capacity and then restored by a shunt diode in such a way that only the maxima (tops) of the pupil pulses can exceed earth potential with an amplitude of about +0.5V. These peaks are amplified up to a standard level of about 6V.

The DC-restorer has been introduced to provide independence of the infrared light source above a certain minimum of intensity.

When this light intensity changes, for instance as a result of variations in the mains voltage, the DC-component of the video signal changes too, but not the amplitude of the positive peaks at the output of the DC-restorer.

Methods of establishing the pupil diameter

There are three methods of defining the pupil diameter from the video signal:

1. The duration of each pupil pulse is compared with that of the preceding one as long as the latter pulse is shorter. As soon as this is no longer the case, the comparison is stopped as now the scanning of the lower half of the pupil begins.

This method has the disadvantage of being complicated and at the same time inaccurate because only one of the lines is used to give the information. An accidental dip or irregularity at the border of the pupil can cause a measurement which is not characteristic of the mean value of the pupil diameter.

2. The number of pupil pulses in each frame is counted. The number is a measure and linearly proportional to the pupil diameter. This method is much more accurate, of course, than the first because it uses the information of a considerable greater number of lines in each frame.

3. Besides the number of pupil pulses per frame, by integration the mean value of the durations of the pupil pulses is also defined. Basically, this method is the most accurate because all the information in one frame is used to the full extent. In fact, one measures the surface of the pupil, so the information is proportional to the square of the pupil diameter.

Method 2 is very simple and gives a purely digital output. Our instrument is based on this system.

The height of the frame and the counting system

If we want to measure the pupil diameter with an accuracy of 1%, the distance between the scanning lines must be 0.01 x 3 = 0.03 mm, since the mean pupil diameter is the half of the minimum (2 mm) and maximum (8 mm) value.
This results in a height of the frame of $312.5 \times 0.05 = 15.625$ mm, high enough to tolerate a certain movement of the eye.

A pupil diameter of say 4 mm results in 80 pulses per frame. Counting 5 frames as a unit, the counter indicates $5 \times 80 = 400$ pulses, i.e. the pupil diameter in hundredths of a mm.

The maximum height of the light-sensitive layer of the "Plumbicon" tube however is only 12.6 mm (with a picture ratio of 4:3). Projection of 1:1 is not possible. A small reduction in the objective is, therefore, necessary.

The illumination of the eye

Direct illumination of the eye results in a picture on the monitor with good contrast and, what is of greater importance, good discrimination between pupil and iris level in the video signal. However, as the eye operates as a spherical mirror the camera tube sees the light source as a very intensive spot and the light sensitive layer will be damaged at that particular point.

![Fig. 4a](image)

**Fig. 4a.** The indirect illumination of the eye by means of a ping pong ball.

a. The ping pong ball.

b. The black strip inside the ball.

c. The pupil, surrounded by part of the iris with the reflection of the black strip thereon, as seen by the camera tube through the hole in the ball.

As the contrast ratio of the infrared "Plumbicon" television camera tube is rather small, it is not possible to decrease the light intensity in order to avoid overloading and at the same time getting sufficient discrimination between pupil and iris in the video signal. So, indirect illumination is necessary.

After many experiments a solution could be found with half of a ping pong ball (fig. 4a). This gives a smooth, transparent illumination of the eye. In the half-ball is a hole of 10 mm diameter.

This kind of illumination, however, diminishes the contrasts in the video signal, because the inside of the half-ball is reflected in the eye. To restore the contrast, a strip 10 mm wide inside the ball has been painted black (fig. 4b). This strip is also reflected in the eye. Fig. 4c shows the picture the camera receives. One sees the contrast-restoring influence of the black strip very clearly.

The purified video signal at the camera, however, is disturbed by the measures just mentioned because the level outside the hole is detected as black, as a result of the sheet of black paper before the ball (fig. 4a).
Fig. 5a. Time pattern of the video signal when the eye is illuminated according to the method of fig. 4a.

b. Time pattern of the output voltage of one-shot flip-flop I.

c. Time pattern of the output voltage of one-shot flip-flop II.

d. Video signal after passing through the whitemaker.

e. The video signal passes through the whitemaker-gate only within an adjustable rectangular figure ABCD. Outside this rectangle the signal is made white.

In this case, too, we can use the whitemaker to overcome the difficulties. Fig. 5 gives an explanation. The time pattern of the video signal is given in Fig. 5a.

On the trailing side of the linepulse a one-shot flip-flop I is activated (fig. 5b). When this mono resets, a second one-shot flip-flop is activated. This mono controls the whitemaker circuit (fig. 5c), so that the video signal can pass only during a restricted lapse of time of each line (fig. 5d). The same limiting process takes place in vertical direction.

By proper adjustments of the flip-flops the video signal passes through the gate only within a rectangular area (fig. 5e). It will be obvious that a separate suppression of the blanking levels of line and frame as explained in fig. 2, is now superfluous.

The block diagram

In the top left corner of fig. 6 is represented the head of the subject with the ping pong ball in front of his right eye. The ball receives the infrared light from two sources, on the right and left of the subject's head. The light stimulus is given on the left eye. The camera, the control unit, the synchronisation unit and the monitor constitute the closed-circuit t.v. system.

The amplified raw video signal is taken from a useful point in the control unit and led to the whitemaker circuit.
Fig. 6: Block diagram of the infrared pupillometer.
Fig. 7: Registrogram showing the contraction and widening of a pupil as a result of pulse-shaped light stimulation. Notice the sharp pulse due to eye winking.
The inverted line and frame pulses are fed to the adjustable one-shot line and frame flip-flops. The whitemaker is controlled via an OR gate.

The purified signal now passes through the DC restorer, the output of which is amplified to a standard amplitude of about 6V.

These normalised pulses are fed to:

1. A binary counter consisting of 8 bistable flip-flops, counting the number of pupil pulses per frame.
   On the front of the frame pulse the position of this counter is conveyed to a register likewise consisting of 8 bistable flip-flops. The first counter resets on the trailing side of the frame pulse.
   The position of the register is converted in the usual manner into an analogue signal by means of a network of resistors.
   The voltage across the output resistor $R_0$ is fed to a penrecorder and an oscilloscope.
   The current through $R_0$ passes through a milliammeter $M$. With the help of the shunt resistor $r$ the deflection of the needle of the meter can be so adjusted that the indication of the meter corresponds to the pupil diameter.

2. A digital counter with printer.

The inverted frame pulses are taken to a scaler of 5, the output of which is the time base of the counter. This means that the pupil pulses are counted during 5 frames. Then the counter stops. After an adjustable display time, the counter resets, and again the pulses of 5 frames are counted etc.

Video-signal-processing part of pupillometer.
Fig. 7 shows a registrogram obtained with a Honeywell "Visicorder" (this is a multichannel instrument with registration by mirror galvanometers on UV sensitive paper).

The block indicates the light stimulus flash. The other line is the time pattern of the pupil diameter. The figure very clearly shows the time delay of the pupil after the light pulse.

References

We would thank Dr. Bouma of this Institute for his valuable advice and suggestions on problems of illumination of the eye.
This device provides a means for allowing smooth running of a tape loop on an ordinary tape recorder. Two main conditions have to be met: tape tension should be approximately the same as in a reel operated mode, whereas no considerable extra friction should result from the fact that a tape loop necessarily shows four angles of about 90° each instead of only two wider angles in reel operation. The latter requirement is satisfactorily met by applying idler rollers with ball bearings. The former led to mounting these rollers on normal tape reels, to the effect that the tape tension depends on the same forces as in reel operation. This is of great importance with non-commercial recorders, in which the tape tension largely determines the contact between the tape and the heads. Generally, in such recorders the tape tension is controlled by the torque of the supply and take up reel motors.
The photograph shows a device by which the histogram of data can be visualized. Each datum is represented by a steel ball and the balls are distributed over 20 histogram classes, each class being represented by a column in the front of the device. The frequency of measurements in each class is shown by the number of balls in each column. The data are obtained electrically and each datum so obtained is assigned to its own column. Technically the process is as follows.

Underneath a rotating supply cylinder, filled with steel balls, is found a transparent panel with 20 parallel columns. In the cylinder are 20 openings corresponding to the 20 columns. Each column is provided with a slide, driven by an electromagnetic device, which prevents balls from falling into it. Immediately upon each measurement one of the electromagnetic devices opens its slide and a ball falls into the corresponding column. The construction of the slides is such that only one ball can fall at each measurement. The piling up of the balls gradually builds up the histogram.
The supply cylinder is rotated by a motor (visible in the illustration). The object of the rotation is to ensure that there is always a ball before its opening. When a histogram of sufficient data has been obtained, the essential values are copied. The balls forming the histogram are removed by shifting the bottom-plate of the columns. All the balls fall into a second supply cylinder, visible at the lower end of the apparatus. The two supply cylinders are interchangeable.

J. Vredenbregt

Suction Cup Electrodes for Electromyography

A pair of suction cup electrodes, as shown above, have been constructed for electromyographic measurements. The electrodes, which are fastened to a flexible strip at a fixed distance from each other, can be quickly placed on and removed from the skin without damaging it, by using vacuum.

Each electrode consists of a silver cup soldered to a bolt with an axial hole. The cup with the bolt is placed through a thin flexible rubber membrane with a sharp edge. By evacuating the space between the skin and the membrane, the cup is pressed against the skin. Because the membrane is sucked on the skin, no damage of the latter occurs in spite of the vacuum being 0.5 atm. The sharp edge of the membrane adapts itself to any irregularities of the skin, so that it is possible to use the electrodes on any part of the skin except where there is abundant growth of hair. The electrodes adhere so tightly to the skin that they can be used in experiments under dynamic conditions.
The necessity of measuring accelerations and decelerations in the frequency range of zero to about 20 Hz has led to the construction of a simple accelerometer, consisting of a cylinder filled with oil, which has a viscosity of 10.5° Engler at 50° celsius, causing critical damping. A small steel plate, at one end provided with a small weight, is attached to the cover of the cylinder and is contained in the oil. The small steel plate is provided with two strain gauges, which detect the bending of the steel plate, due to the acceleration and deceleration of the small weight. The strain gauges are connected to a professional measuring bridge (Philips PR 9300), giving an output voltage, proportional to the acceleration and deceleration in the frequency range from zero to 18 Hz for accelerations from zero to 5 g.

G.J.J. Moonen & C.A. Lammers

Lining Up Information in a Shift Register 1)

When a shift register is used as a buffer between an information source, giving the information at irregular intervals and an information processor that works continuously and which is able to digest the mean speed of the supply of information, then it is important to use the buffer capacity to its full extent. To guarantee that any information fed into the buffer is available at the processor it is necessary that the information is shifted in such a way that possible interspaced idle sections in the shift register must be filled up as soon as possible. As a result there is always a continuous block of information present at the processor end. To realise this it is normal practice to use a second register of the same capacity from which information is fed into the first one in the correct sections; this happens simultaneously. For multichannel shift registers this may be an expensive method. So a cheaper method has been sought.

By cutting the shift pulse circuit in sections each section of the shift register can be operated separately. For the names used see fig. 1.

![Diagram of shift pulse circuit](image)

fig. 1: Illustration of names used
If no information is present the shift pulse is fed to all sections of the register. If, however, information is present in section $S_x$ no shift pulse will be fed to the sections preceding $S_x$ in which the information is lined up. When new information is supplied it will be shifted till it is linked with the information already present. The problem is to cut the shift pulse circuit in a selective way because as soon as information, seen from the supply source, has been shifted so far that no more information is present in section $S_1$ the shift pulse circuit will also be cut at this section. This means that the part in between both blocks of information will not be used. In terms of circuitry this has been solved as follows:

![Circuit Diagram](image)

**fig. 2: Circuit for selective information shifting**

The gates $a$, $b$, and so on are OR-gates whereas the gates $A$, $B - X$ are AND-gates. Via the output of each AND-gate a shift pulse can be fed to the corresponding section. The synchronism of the shift pulses is obtained by feeding the clock pulse to a common input of all AND-gates. One input of each OR-gate is used to record whether or not information is present in its corresponding section. Information is indicated by signal "0"; no information by signal "L".

If no information is present in section $X$ the OR-gate $x$ will have signal "L" at its output. As a consequence all preceding OR-gates have signal "L" independently of the presence of any information in the register. So the pulse present in the common shift circuit will be passed through all the gates $A - X$.

This continues until information arrives in section $S_x$. This will cause the output of OR-gate $x$ to give signal "0" assuming that $S_y$ has "0" too. AND-gate $X$ is blocked immediately and so in gate $C$ if information is present in section $S_3$. If no information is present in $S_3$ then one input of OR-gate $c$ is in position "L" and so all preceding OR-gates are in the "L" position. As a consequence all shift pulses will pass on to the section $C - A$. If information is stored in $S_3$ then OR-gate $c$ will be switched to "0". AND-gate $C$ is cut off and depending on whether or not there is information in $S_2$ the shift pulse circuit will be cut between $B$ and $C$. 
The signal that is fed to Sy determines whether all sections receive a shift pulse or only the part before the cutting. In this way the processor can signal that new information is wanted (all shift) or that the information in the last section still has to be processed (pulse circuit cut).

Note: It is possible to detect the presence of information in a section by connecting an OR-gate input to each channel of the section. Another solution is adding an extra channel in which a signal "0" is written when information is supplied by the source. The outputs of this extra channel can be coupled directly to the inputs of the OR-gates.

It is also possible to use such a code that only a few channels have to be detected to ascertain if information is stored in the subsequent section.

1) Patent pending.

F.F. Leopold & H.S. Fuchs

Conversion of a Typewriter for Typing on Paper Tape

To manufacture typed paper tapes for a reading device (see next page) the carriage of an input-output typewriter is removed. The picture shows a cylinder instead of the carriage. The cylinder is driven by a relay, excited by pulses from contacts underneath the keys.
The paper tape reading device has its application in reading experiments. The apparatus shown gives the possibility of presenting to a subject a printed paper tape behind a screen.

The speed of the tape is continuously adjustable, both in forward and backward direction, between zero and 20 cm per second. The width of the window is adjustable to a maximum of 26 cm.
The following is the description of a simple and handy vowel synthesizer, the design and construction of which was made easy on the basis of our extensive experience with synthetic speech. It is a two-formant parallel synthesizer, and the formant choice can be made by adjusting a stud in a (cartesian) plane having the $F_1$ and $F_2$ frequencies along the axes. The stud can be moved during the production of a vowel and in this way diphthongs can be produced.

**Block diagram**

The waveform of the periodic source is approximately a sawtooth and is fed into the formant filters. The filters are the same as those of the IPOVOX II (described in this report). The $F_1$ filter has a constant bandwidth of 50 Hz and the resonant frequency is variable from 200 to 1,200 Hz. The $F_2$ filter has approximately a constant $Q$ of 10 and the resonance frequency is variable between 700 and 4,200 Hz. The signals from the two formant filters are added and then multiplied by the envelope. The envelope function has a rise time variable between 10 and 100 ms and a decay time variable between 20 and 80 ms.
The total duration of the envelope defined as the time between the beginnings of the rise and decay is variable between 80 and 300 ms. During a vowel the frequency of the periodic source ($F_0$) is not kept constant. A special intonation contour is generated, which has a rising $F_0$ for the first 40% of the duration and after that the $F_0$ decreases. It is also possible to produce continuously a monotonous vowel-like sound.

**Fig. 1: Block diagram of Vowel Synthesizer**
J.P.O. Publications

<table>
<thead>
<tr>
<th>P</th>
<th>Name</th>
<th>Title</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>J.F. Schouten</td>
<td>Nachricht und Signal. v.d. D2(1956), 1-2.</td>
<td>VDE Fachberichte</td>
</tr>
<tr>
<td>Page</td>
<td>Authors</td>
<td>Title</td>
<td>Journal/Conference</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>--------------------</td>
</tr>
<tr>
<td>18</td>
<td>J.F. Schouten</td>
<td>Wat is waar in de waarneming?</td>
<td>Wijsgerig Perspectief</td>
</tr>
<tr>
<td>20</td>
<td>A. Cohen</td>
<td>On the Value of Experimental Phonetics for the Linguist.</td>
<td>Lingua</td>
</tr>
<tr>
<td>24</td>
<td>H.W. Horesman</td>
<td>Perceptieperikelen in het wegverkeer.</td>
<td>7e Politie Verkeersleergang te Utrecht</td>
</tr>
<tr>
<td>29</td>
<td>A. Cohen</td>
<td>Spraakverschijnselen perceptief bekeken.</td>
<td>Spraak. Publicatie no. 1 van het Nederlands Akoestisch Genootschap</td>
</tr>
</tbody>
</table>

A Model of Human Pitch-Extraction Based on Additive Correlation.


Segmentation of the Speech Continuum.


Frequency Discrimination of the Human Ear.

Het Instituut voor Perceptie Onderzoek.


De toonhoogtegebaarwording.

Philips Technisch Tijdschrift, 24(1962), 341-347.

Het fonetisch onderzoek.

Philips Technisch Tijdschrift, 24(1962), 348-353.

Enige onderzoekingen aan het visuele zintuigzysteem.


Op zoek naar een maat voor de perceptieve arbeid.


De DONDERS, een electronisch meetsysteem voor menselijke reacties.


Human Code Transmission.

Ergonomics, 4(1963), 133-143.

Inductive Brightness Depression as Influenced by Configurational Conditions.

Vision Research, 2(1963), 121-190.

Measuring the Force and the Degree and Speed of Shortening in Muscle Contraction.

Acta Physiol. Pharmacol. Neerlandica, 11 (1962), (e.p.).

The Existence Region of the Tonal Residue II.


On Pitch Discrimination of Residue Tones.


Waan, waarneming en werkelijkheid.


Perceptual Tolerances of Isolated Dutch Vowels.

Phonetica, 2(1963), 65-78.

Neural Effects and the Psychophysical Law.


Ein Mehrzweck-Massenspektrometer 1e Teil.

Elektronik, 12(1964), 1-32.


Hearing with One Ear and with Two Ears. International Audiology, 4(1965), 208-209.


Adjusting the Method of Adjustment. J. Acoustical Society of America, 22(1965), 786-793.


Hoe neemt de mens het verkeer waar? Maatschappijbelangen, 8(1965), 395.


<table>
<thead>
<tr>
<th>Page</th>
<th>Authors</th>
<th>Title</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J. 't Hart</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H.E.M. Mélotte</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>J. 't Hart</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A. van Katwijk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P 82</td>
<td>J. Vredenbregt, J.J. Andriessen</td>
<td>Les rapports entre le temps de montage et le jeu relatif des broches pour différents diamètres de trous.</td>
<td>Le Travail Humain, 29(1966), 164-166.</td>
</tr>
<tr>
<td>P 83</td>
<td>W.G. Koster,</td>
<td>La mesure de la charge perceptive, comparaison de quelques méthodes.</td>
<td>Le Travail Humain, 29(1966), 139-141</td>
</tr>
<tr>
<td></td>
<td>G. Taverne</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P 84</td>
<td>H.W. Horeman,</td>
<td>Indication de direction sur feux de circulation.</td>
<td>Le Travail Humain, 29(1966), 135-137</td>
</tr>
<tr>
<td></td>
<td>M.J. Zwart</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>