

The Phonetic Basis of Perceptual Ratings of Running Speech

Netherlands Phonetic Archives

The Netherlands Phonetic Archives (NPA) are modestly priced series of monographs or edited volumes of papers, reporting recent advances in the field of phonetics and experimental phonology. The archives address an audience of phoneticians, phonologists and psycholinguists.

Editors:

Marcel P.R. Van den Broecke

University of Utrecht

Vincent J. van Heuven

University of Leyden

Other books in this series:

- I Nico Willems
English Intonation from a Dutch Point of View
- IIA A. Cohen and M.P.R. Van den Broecke (eds.)
Abstracts of the Tenth International Congress of Phonetic Sciences
- IIB M.P.R. Van den Broecke and A. Cohen (eds.)
Proceedings of the Tenth International Congress of Phonetic Sciences
- III J.R. de Pijper
Modelling British English Intonation

Lou Boves

The Phonetic Basis of Perceptual Ratings of Running Speech



1984

FORIS PUBLICATIONS

Dordrecht - Holland/Cinnaminson - U.S.A.

Published by:
Foris Publications Holland
P.O. Box 509
3300 AM Dordrecht, The Netherlands

Sole distributor for the U.S.A. and Canada:
Foris Publications U.S.A.
P.O. Box C-50
Cinnaminson N.J. 08077
U.S.A.

CIP-data Koninklijke Bibliotheek, Den Haag

Boves, Lou

The Phonetic Basis of Perceptual Ratings of Running Speech / Lou Boves. - Dordrecht [etc.]: Foris Publications. - (Netherlands Phonetic Archives; 4)
ISBN 90-6765-034-X bound
ISBN 90-6765-035-8 paper
SISO 805.2 UDC 801.4
Subject heading: speech; experimental phonetics.

Frontcover illustration taken from: F.M. Helmont [An unabbreviated representation of a true natural Hebrew alphabeth, which simultaneously shows how those born deaf can be thought not only to understand others who speak but even to produce speech themselves], Pieter Rotterdam, Amsterdam 1697.

ISBN 90 6765 034 X (Bound)
ISBN 90 6765 035 8 (Paper)

© 1984 Foris Publications - Dordrecht.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission from the copyright owner.

Printed in the Netherlands by ICG Printing, Dordrecht.

*To the people of the Netherlands
and those abroad who have contributed to our prosperity*

Contents

Preface	XI
1 Research into non-verbal aspects of speech	1
1.0 Introduction: form and content	1
1.1 Previous research into non-verbal aspects of speech	2
1.2 Methodological problems in the description of vocal aspects of speech	3
1.3 Vocal aspects of speech in modern linguistics	4
1.4 Aims of the present research	6
1.5 Operationalization of the research goals	7
1.6 Flavours: methodology versus content	9
2 The construction of a scaling instrument	11
2.0 Introduction	11
2.1 Selection of the objects to be judged	12
2.2 The administration of a semantic differential	14
2.3 The selection of the scales	15
2.3.1 Factorial composition and relevance of the scales	16
2.3.2 Stability of meaning and linearity of scales	16
2.3.3 Choosing from inventories of scales	17
2.3.4 The status of the rating scales	21
2.4 Testing the scales for linearity and relevance	22
2.5 Conclusion	27
3 The evaluation experiment	29
3.0 Introduction	29
3.1 Method	29
3.1.1 The preparation of the stimuli	29
3.1.2 The collection of the ratings	32
3.2 Analysis of the individual scales	33
3.2.1 Introduction of the scaling model	34
3.2.2 Results	38
3.2.2.1 The linearity of the scales	38
3.2.2.2 The goodness-of-fit of the model	40
3.2.3 The reliability of the ratings	41

VIII

3.2.4	Conclusions from the scale analysis	44
3.3	The effect of the texts on the ratings	45
3.3.1	Conclusions	48
3.4	Factor analysis	49
3.4.1	Introductory remarks	49
3.4.2	The dimensionality of the judgement space	51
3.4.2.1	Discussion	53
3.4.3	The construction of an efficient scaling instrument	54
3.4.4	The position of the talkers in the judgement space	55
3.5	Conclusions	58
4	Separation of the contributions of phonation and articulation to the radiated speech signal	61
4.0	Introduction	61
4.1	Registrations of subglottal pressure	68
4.1.1	Techniques for recording subglottal pressure	71
4.2	Making the simultaneous registrations	73
4.3	Establishing criterion instants based on physiological registrations	78
4.3.1	The photoglottogram and the moment of the maximum glottal flow	81
4.3.2	The P_{sg} recordings and the moment of glottal opening	83
4.3.3	Conclusions	86
4.4	Inverse filtering techniques	88
4.5	Inverse filtering of vowels from running speech produced by several speakers	90
4.5.1	Speech material	90
4.5.2	Processing of the inverse filtering results	91
4.5.3	Results	94
4.5.3.1	Speaker characteristics	94
4.5.3.2	Stability of results upon repeated measurements	98
4.5.3.3	Glottal pulse differences between stressed and unstressed syllables	98
4.5.3.4	The influence of supraglottal load	99
4.5.4	Discussion	102
4.6	Conclusions	105
5	Development of acoustic measures to explain perceptual ratings	107
5.0	Introduction	107
5.1	Acoustic measures of tempo	108
5.2	Acoustic measures of pitch	111
5.2.1	Measures of central tendency	112
5.2.2	Measures of variability	113
5.2.3	Pitch perturbation measures	113
5.2.4	The voiced-unvoiced ratio	115

5.2.5	The pitch data	115
5.3	Long-term average spectra	117
5.3.1	The generation of LTAS	118
5.3.2	The representation of LTAS	119
5.3.3	The influence of fundamental frequency on LTAS	121
5.3.4	LTAS as measures to characterize speakers	122
5.3.5	LTAS as representation of the glottal flow spectrum	124
5.4	Measures based on formant analyses	128
5.5	Summary and conclusions	134
6	Relations between acoustic measures and perceptual ratings	135
6.0	Introduction	135
6.1	Simple correlations	136
6.1.1	Voice dynamics	136
6.1.2	Articulation quality	137
6.1.3	The paralinguistic factor in the Nijmegen experiment	141
6.1.4	Potency	143
6.1.5	Voice quality	144
6.1.6	Tempo	148
6.1.7	Discussion and conclusion	149
6.2	Multiple regression analysis	150
6.2.1	Voice dynamics	154
6.2.2	Articulation appreciation	155
6.2.3	The sociolinguistic factor	155
6.2.4	The potency factor	156
6.2.5	Voice quality	156
6.2.6	Tempo	157
6.3	General discussion: method and results	157
6.4	Correlations versus experimental research	160
6.5	Summary and conclusions	162
7	The status quo and perspectives	165
7.0	Introduction	165
7.1	The state-of-the-art in acoustic descriptions of speech signals	167
7.2	The state-of-the-art in perceptual descriptions of running speech	169
7.3	The relations between acoustic and perceptual descriptions of running speech	171
Appendix I Development of an accurate closed glottis interval technique for inverse filtering		173
A1.0	Introduction	173
A1.1	The closed glottis interval approach to inverse filtering	173
A1.2	Determination of the CGI from the speech	174

X

A1.3	Estimating the CGI from the EGG	177
A1.4	The implementation of a closed glottis inverse filtering technique using the EGG	178
A1.5	Accuracy of formant estimates obtained from the CGI	180
References		183
Samenvatting		189
Curriculum vitae		197

Preface

It is well known that phonetics is an interdisciplinary science which draws from linguistics physics, engineering, sociology and psychology, to name just the most obvious disciplines. From the Table of Contents of this book one can see that most, if not all, of the sciences just mentioned have played a role in the work reported here. From this the reader may draw the conclusion that the work cannot have been done by a single person, operating in isolation. This is why the first person plural is used throughout, instead of the singular which might seem to be more appropriate since there is only one author.

Drs. Wil Fagel and drs. Leo van Herpt organized and conducted all rating sessions and saw to it that the listener scores became available in machine readable form. Wil Fagel also took care of a large part of the statistical processing of the rating scores, including the factor analyses.

Drs. Carel Jansen provided essential help in making the recordings of subglottal pressure, photoglottogram and electroglottogram. Nico van Rossum skillfully designed and built all electronics necessary for making the simultaneous recordings of physiological and acoustic signals.

Ir. Bert Cranen was always available and never tired of discussing the intricacies of analog and digital signal processing in general and of inverse filtering in particular. Also, he implemented a considerable part of the programs used to process critical band spectra. He helped in unravelling the mathematical details of Bloms proposal for approaching the problem of the dimensionality of rating scales, discussed in chapter 3. Finally, but perhaps most importantly, he relieved me of all my teaching and service chores.

Ir. Lei Willems, dr. Harm Schutte and dr. Tjeerd de Graaf gave useful comments on a number of practical details of the work. Profs. E. Nuijten, L. Pols and W. Vieregge, the series editors dr. M. Van den Broecke and dr. V. van Heuven, as well as Ir. G. Bloothoof made invaluable comments on earlier drafts of the manuscript.

With so many contributions of so many people to the work reported here it would almost be inevitable that some of its faults come on the account of somebody else than the author. And although it would have been rather unusual, I surely would have put the blame where it belongs, even if it were on somebody else than myself. However, I am sufficiently stubborn to comply with comments when I am convinced they are right, but to disregard

justifiable comments completely if I feel that they interfere with my own set purposes. Thus all shortcomings of the book, including most of the typographical errors, come on the account of the author alone.

Besides people there are also organisations and institutions which contribute to the work of individuals. My work would not have been possible if the money that feed me during three years from all my teaching and administrative duties had not been made available. Thus I should like to acknowledge the indispensable financial contribution, using the standard phrase: This research was supported by the Foundation for Linguistic Research, which is funded by the Netherlands Organisation for the advancement of pure research, Z.W.O.

CHAPTER 1

Research into non-verbal aspects of speech

1.0 INTRODUCTION: FORM AND CONTENT

Speech is often called man's most natural and efficient mode of communication. Whether or not this is true, it can hardly be denied that it is a very treacherous mode, since the speech signal contains a lot of information which is not so much related to the linguistic content of the verbal message but rather to the identity and physical and psychological status of the speaker. For the greater part this non-verbal (para- or extralinguistic) information is conveyed unintentionally and more often than not the speaker is absolutely unaware of its presence and its effects. Also listeners do not normally pay explicit attention to this kind of information which certainly is not to say that they do not normally perceive it and use it to their advantage (or to their detriment). Non-verbal information may be used to the advantage of the listener if he or she recognizes the speaking person correctly or infers the speaker's state and intention correctly from how he or she sounds. Non-verbal information is used to the listener's detriment if the speaker succeeds in making a false impression of friendliness, honesty or trustworthiness.

That the impact of a message does not only depend on its content but also very much on the form in which it is presented has been known for more than 2000 years and most probably much longer. At least since Greek antiquity the art or science of rhetoric exists and has been practiced and studied extensively by many people who aim at making a public career, an undertaking which is aided very much if as many people as possible get a favourable impression of the aspiring person. Training and study of rhetoric has resulted in an extensive list of speech characteristics which are thought to have specific effects on the listeners. A scan of the literature reveals, however, that the rhetoric tradition has concentrated its efforts on the argumentative structuring of speeches. This state of affairs reflects the close relationship of rhetoric and logic. The delivery aspect of a speech has not been neglected by the rhetoric tradition, but here the precepts and advice to prospective career-makers are nowhere nearly as precise and specific as in the domain of argumentation theory. Thus it seems honest to say that classical rhetoric was born a science as far as the linguistic, logic argumentation aspect is concerned and born an art with respect to the delivery aspect.

It should be noted that at least since Aristotle it has been realized that apart from the content of a speech and its structuring, and the way in which a speech is delivered, there is a third aspect which influences the impact of a rendition. This latter aspect is what was called the *ethos* of the speaker by Aristotle. The most general translation of this term is the credibility and honesty of the speaker as the source of the message in question. The *ethos* aspect is, of course, a very complicated one if only because it depends heavily on the situation. In dealing with people one has known and seen behaving for quite some time, prior experience may be much more important in estimating the trustworthiness of a message than the content of the message itself and the way in which it is conveyed. In initial interactions with strangers non-verbal aspects of communication may become extremely important. Thus it is clear that in our present-day mass-communication society a speaker's (perceived) *ethos* is not on the same level as the argumentative structure of his message and its style of delivery, but that the former tends to be one of the products of the latter two. It may safely be assumed that the rhetoric tradition has had a considerable influence on the formation of stereotypical ideas about the relation between the formal characteristics of a message and the personality characteristics of the person who conveys the message. Despite of their vagueness, prescripts concerning delivery may have been more important here than the rules of argumentation.

The single most characteristic feature of the rhetoric tradition is that it is advisory and prescriptive; in no way is it an empirical science. Empirical research into the effects of the verbal and non-verbal features of a message did not start until the rise of the science of mass-communication. Much like rhetoric this science has developed along two quite independent lines, one concentrating on the verbal (content, formulations, structuring) aspects of a message, the other dealing with the non-verbal aspects. The present study is exclusively concerned with the non-verbal aspects of communication, and not necessarily mass-communication. It can be maintained in general that many of the results of the research into the non-verbal aspects of mass-communication hold equally well in person-to-person communication especially in initial interaction between strangers. There have been contributions to the field from such diverse disciplines as social psychology, clinical psychology and psychiatry, personality theory, and - more recently - sociolinguistics.

1.1 PREVIOUS RESEARCH INTO NON-VERBAL ASPECTS OF SPEECH

There is no need for reviewing the research into the non-verbal aspects of communication if only since review papers are beginning to appear in which earlier reviews are reviewed together with more recent original research (Scherer, 1979). At this point it suffices to mention the two most important conclusions. The first conclusion, which one is obliged to draw from a survey

of the literature, is that the research efforts have resulted in little more than a confirmation of the existence of vocal stereotypes, i.e., stereotypical and often incorrect ideas about the relation between non-verbal (= vocal) aspects of a person's speech and her or his personality. The second conclusion, which has been reiterated by Scherer (e.g. Scherer, 1978), is that the first conclusion seems hardly warranted in view of the fundamental methodological shortcomings of the research under review.

Scherer proposes a research model which is meant to alleviate the methodological problems, or at least to point them out explicitly. To honor its originator it is called the Brunswikian Lens model. In this model, the details of which need not concern us here, a strict separation of levels of description is advocated. Specifically it is emphasized that three levels must be distinguished, viz. an 'objective' acoustic description of the speech, a 'perceptual' description of the speech, which is still meant to be quite objective, and a third level concerned with the attribution of personality characteristics. In most respects this proposal is not revolutionary. It parallels the commonly accepted notion of separation of levels in linguistics and phonology. Only in the research practice of the past, the indiscriminate intermixing of levels seems to have been the rule rather than the exception.

1.2 METHODOLOGICAL PROBLEMS IN THE DESCRIPTION OF VOCAL ASPECTS OF SPEECH

The description of the non-verbal aspects of speech is one of the fields where technical and methodological problems abound and solutions are rare. There even seems to be a lack of agreement on the collection of features which rightly belong to the class of non-verbal aspects of speech. In such a situation it is not surprising that terminological clashes abound. Although it is the major aim of this study to contribute to the improvement of the description of the non-verbal aspects of speech, our contribution is concerned with technical problems rather than with the construction of a theoretically sound and comprehensive description system. Therefore we can - and will - afford the luxury of taking a very pragmatic approach to the problem of defining the non-verbal aspects of speech and staking out the field. The pragmatic 'definition' of the non-verbal (or vocal, a term which is considered as synonymous) aspects of speech for the aim of our work is very much facilitated by the fact that it deals exclusively with texts read aloud. In such a situation the non-verbal aspects of a message can be said to be all information which is added to the text in and by the process of reading aloud. This additional information pertains to the identity of the talker (for those who know him or her personally) or to her/his sex, approximate age, approximate regional and social origin and approximate physical and mental states (for the listeners to whom the reader is a stranger). Vocal information is carried by characteristics of the speech signal which are more or less

constantly present, in contrast to the signal features that carry the verbal (linguistic, content) information and which need to vary constantly if any non-trivial message is to be conveyed. The aspects of the speech signal we are aiming at can (informally and without any claim as to completeness and precision) be described as the average speaking pitch, pitch variations, average loudness and loudness variations, average tempo and tempo variations, voice quality and precision of articulation. The majority of those features will be recognized as belonging to what is often called the suprasegmental or paralinguistic properties of speech.

The term 'paralinguistic features' suggests at least in part an explanation why the description of the properties in question suffers from so many problems and perplexities. The mere fact that we have to do with phenomena that are at the same time linguistic in nature (i.e., behave according to rules which are essentially arbitrary and used to convey meanings which are essentially discrete) and non-linguistic in nature (i.e., vary continuously as a function of e.g. the physiological state of arousal of the speaker reflecting conditions which vary continuously rather than discretely) makes their study very difficult and frustrating. Linguistically inclined scientists are hampered by the non-rule based continuous variations of the phenomena, whereas researchers geared towards the empirical correlational methodology of the social sciences have difficulty in coming to grips with the discrete linguistic part of the phenomena.

1.3 VOCAL ASPECTS OF SPEECH IN MODERN LINGUISTICS

Perhaps it is time now to interrupt this fairly general and abstract discussion by the presentation of some research examples which show how vocal aspects of speech play their role in modern linguistic research and at the same time bring to light the problems caused by the difficulties encountered in describing and manipulating non-verbal speech characteristics. The first case chosen for illustration is the much debated topic of the 'inherent norm' versus 'imposed value' hypotheses in sociolinguistics. The case centers around the explanation of the prescientific knowledge and the empirical scientific fact that people tend to display more positive attitudes towards some languages or language varieties than towards others (Nuijten, 1962; Brown & Lambert, 1976; Giles & Powesland, 1975). Brown and Lambert presented recordings of 20 adult male French-Canadians reading a short standard passage to 26 American students, none of whom knowing any French. The listeners were asked (among other things) to rate the social status of the speakers. It appeared that 47% of the variance in the ratings could be accounted for by the real status differences. This result compared favourably with the outcomes of a similar experiment in which the raters were French-Canadians (67% of the rater variance explained by real status differences). Brown and Lambert suggest that the relatively accurate perfor-

mance of the American raters shows the existence of cross-culturally similar voice quality differences between members of the lower and upper class.

Giles and Powesland, on the other hand, describe a series of experiments in which the reactions towards different language varieties, e.g. a dialect and the standard language or a low and a high prestige dialect were investigated using the matched-guise technique. This means that the same bi- or multilingual speakers produced the speech stimuli in all language varieties under study. It was hoped that in this way all properties of the speech signals not related to the language or dialect differences under study were kept constant. In these experiments it was invariably found that the expected differences in the evaluation of the language varieties did not show up if the raters were unable to recognize the languages or dialects and therefore could not identify the high prestige variety as such. This is the case when, for instance, Welsh students are asked to judge talkers using the (high prestige) dialect of Athens and the (low prestige) dialect spoken in the isle of Krete. From these findings it is inferred that there are no real audible differences between languages except those which are a direct result of the linguistic differences and that all differences in evaluative reactions are due to culture specific, learned attitudes which are triggered by the recognition of a certain language or dialect.

Clearly the methodological differences between the studies of Brown and Lambert on the one hand and those of Giles and his associates on the other are more than large enough to preclude any comparison. Also the outcomes of neither suffice to disprove the alternative hypothesis (nor to prove the hypothesis adhered to, for that matter). Brown and coworkers would be much more persuasive if they were able to prove the existence of systematic differences in the paralinguistic properties of the speech of high and low class talkers. Although they have attempted to obtain a description of the French-Canadian stimuli, they do not bring those descriptions to bear. Instead, they limit themselves to making remarks about having the intuitive impression that between-class differences are indeed present on the level of paralinguistic features. Most probably this course of action is taken since Brown et al. did not succeed in obtaining a reliable and accurate description of the paralinguistic properties of their stimuli. This assumption is completely in line with the fact that the group around Brown took refuge to a research strategy in which the stimuli under judgement were synthesized utterances, the prosodic characteristics of which were varied systematically (Brown, Strong & Rencher, 1974).

Giles and his coworkers are in an equally vulnerable position. First of all the lack of a description of the paralinguistic features of the stimuli used in their matched-guise experiments leaves no alternative option than to hope and trust that they were indeed constant over the language varieties. The fact that no effect of the varieties was found on the level of evaluation suggests that they did reach that goal. It may be objected, however, that bi- or tridialectal speakers do not constitute a random and representative sample of the speakers of the language varieties under study. Since we do not have

descriptions of the paralinguistic features of dialects and languages, it is not possible to rule out the possibility that the talkers have transferred those characteristics from one language variety to the others, where they may be less appropriate.

The possibility should seriously be considered that the results of the two research groups are by no means as contradictory as the hypotheses promoted suggest. To reconcile the seeming contradictions or to falsify one (or both) hypotheses would, however, call for experiments in which all relevant properties of the speech stimuli, including the vocal characteristics, are known and controlled systematically. It is quite obvious that this ideal cannot be reached except via the way of developing adequate techniques for dealing with the vocal characteristics of speech.

A quite different example of the importance of the non-verbal characteristics of speech, one which brings us back to the framework of personality perception, is taken from the field of speech technology. An American manufacturer of motor cars decided to equip the luxury models with a speech synthesis module for advising the driver as to the state of his vehicle. Since the synthesis module operates on the basis of Linear Prediction analysis-resynthesis (refer to chapter 4 for an explanation of this technique) a suitable talker had to be chosen. 'Because surveys indicate that luxury-car drivers prefer a male synthesized voice (....) a commanding yet warm male speaking voice' was sought (Finkelstein, 1983). From this example it appears that it may even become of economic importance to know what kind of voice is most appropriate in a given situation and -equally preponderant- to know exactly how that kind of voice can be produced.

1.4 AIMS OF THE PRESENT RESEARCH

It has been said before that in the description of vocal characteristics of speech three levels must be distinguished: acoustic and perceptual descriptions should not be confused and both should be distinguished from the level of attribution of personality characteristics. The first two levels are easily recognized as two subdisciplines which are often distinguished within the phonetic sciences, viz. acoustic and perceptual phonetics. The relation between the levels overlaps with the field of psychoacoustics. The level of personality attribution is less easily interpreted in phonetic or linguistic terms, but it certainly has very much to do with the pragmatic aspects of language behaviour. It has also been said that on all levels of description a large number of technical and methodological problems await solution.

Our study is concerned with the development of techniques for obtaining reliable, valid and cost-effective descriptions of some vocal aspects of speech on all three levels mentioned in the previous paragraph and with establishing links between these levels. At first sight a detailed perceptual description of

speech samples seems to be somewhat superfluous if one has an 'objective' description of the speech on the acoustic level. Also, perceptual descriptions are sometimes considered as very much less reliable than acoustic measurements. Yet, we think that both acoustic and perceptual descriptions are indispensable. Not only will it appear from this study that one can obtain highly reliable perceptual ratings of vocal speech parameters, but it seems also to be impossible to establish the validity of acoustic measurements when bypassing the low-level perceptual processes. This is, of course, not to say that no special precautions have to be taken in order to secure reliable perceptual descriptions of vocal characteristics of speech. Consequently chapters 2 and 3 of the present book are mainly concerned with the description and development of instruments and techniques for collecting and processing perceptual ratings of speech.

1.5 OPERATIONALIZATION OF THE RESEARCH GOALS

Despite of a deliberately pragmatic approach to the problem of the definition of vocal parameters we must, of course, make some decisions as to which features are to be measured both at the acoustic and perceptual level. Similar decisions are also necessary on the level of attribution, but there an independent strategy had to be followed. The choice of features, and therewith the development of measurement instruments, has been guided by two quite different and independent bodies of theory and empirical facts. The first is formed by the outcomes of previous research into the relation between vocal characteristics of speech and person perception. Even if the majority of the older studies has employed questionable methodologies leading to similarly questionable results a review allows one to make an inventory of features which have been found to relate in some way or another to personality characteristics. Since we will explicitly refrain from attempts to link the results of our descriptions with the true personality characteristics of our talker-subjects, the findings from research leading to the disappointing conclusion that everything was based on vocal stereotypes are as good as any other. In passing, it should be noted that the study of vocal stereotypes is as important as research into the relations between true personality and vocal characteristics, because it is those stereotypes which operate when we try to deduce the personality and psychological or physical state of a stranger from his or her voice. It is from an overview of this research that the list of features summed up earlier originated. To be a little more specific, increased pitch and increased pitch variations are known to induce the impression of a person who lacks self-confidence and who is tense. High loudness and a large dynamic range are associated with powerful and extrovert personalities, as is a fast speaking rate. At this stage it is not important that the relations mentioned (and all remaining ones not cited here) may only hold in specific cultures or only for adult males but not females. What counts is that a certain feature has been reported to have an effect on personality judgements.

The second body of theory and knowledge which has guided our work consisted of linguistic and phonetic studies aiming at the construction of a comprehensive framework for defining and describing vocal properties of speech. A recent attempt to integrate existing knowledge into a full-fledged description system is Laver (1980). Although features connected to intonation are lacking, more recent extended versions of his system do, at least in part, alleviate this shortcoming. Laver's work is important for two reasons. Firstly, he brings to bear phonetic theory in order to structure and systematize the rather large collection of vocal features. Secondly, he attempts not to mix different levels of description. On the contrary, while aiming at a description on the perceptual level, he carefully details the acoustic and articulatory manifestations or correlates of the perceptual features. The parallel with Scherer's emphasis on separation of levels will be obvious.

Readily available recent reviews of the literature on the definition and description of vocal features (besides Laver's 1980 book see also Crystal, 1975) and on the relations between various levels of description (physiological, acoustic, perceptual, linguistic) relieves us from the obligation to repeat such a review here. From the data collected in these reviews it appears that we can safely assume that on all these levels vocal phenomena relevant to person perception do exist and that there must be systematic relations between the levels.

Within the fields of physiological, acoustic and perceptual phonetics a further subdivision should be recognized, which is very often made but equally often left implicit, viz. the distinction between phonation and articulation. It is customary in phonetics to concentrate on articulation and to leave phonation to the attention of the more medically (or socially) inclined sciences. This division of labour is obviously adequate if phonetics is conceived of as tightly linked to structural linguistics, since phonation types hardly ever act as distinctive features in the sense of classical phonology. (It should be mentioned, of course, that some authors (e.g. Ladefoged 1973; Ladefoged & Maddieson, 1983) do recognize so called laryngeal features, but the use of these features is limited to a few languages and their status is not unchallenged.) This corresponds with the idea that phonation belongs more to the realm of continuously varying phenomena than to the one of discretely varying objects (or should one rather say: described as discretely varying) which linguistics traditionally deals with.

Within the framework of research into the attribution of personality characteristics based on speech, features related to phonation play a role which is at least as important as that of articulatory features. Since the acceptable range of variation within articulatory features may be much more restricted, due to their linguistic function, phonatory features might even be the more important ones here. Thus there are at least two reasons why much effort will have to be invested into the development of procedures for measuring phonatory features i.e., their importance in forming the speaker's image and their relative neglect in existing phonetic work.