# SINTEF REPORT

**SINTEF Telecom and Informatics**

Address: N-7465 Trondheim
NORWAY

Location Trondheim:
O.S. Bragstads plass, Gløshaugen

Location Oslo:
Forskningsveien 1

Telephone: +47 73 59 30 00
Fax: +47 73 59 43 02

Enterprise No.: NO 948 007 029 MVA

TITLE

# Algorithms for hearing aids

## Listening tests and applications

AUTHOR(S)

Sverre Stensby, Asbjørn Krokstad, Wim van Dommelen

CLIENT(S)

Helse og rehabilitering, prosjekt forskning 1999/2/0144

| REPORT NO. | CLASSIFICATION | CLIENTS REF. | |
|---|---|---|---|
| STF40 A02010 | OPEN | Hørselshemmedes landsforbund v. Steinar Birkeland | |
| **CLASS. THIS PAGE** | **ISBN** | **PROJECT NO.** | **NO. OF PAGES/APPENDICES** |
| OPEN | 82-14-02511-7 | 403006 | 75 |
| **ELECTRONIC FILE CODE** | | **PROJECT MANAGER (NAME, SIGN.)** | **CHECKED BY (NAME, SIGN.)** |
| I: STF40 A02010.doc | | Sverre Stensby | Truls Gjestland |
| **FILE CODE** | **DATE** | **APPROVED BY (NAME, POSITION, SIGN.)** | |
| | 2002-02-05 | Odd Kr. Ø. Pettersen, Research Director | |


ABSTRACT

More than half of the population above 60 years of age suffer from loss of hearing. The most severe consequence of this hearing loss is the reduced ability to understand speech. Hearing aids are being used to compensate for loss of hearing, however, the performance of state-of-the-art hearing aids is far from satisfactory. The reason for this is mainly due to limitations in the analogue signal processing. The main function of traditional hearing aids was to amplify the signal so that the speech spectrum could be detected above the raised threshold of hearing.

The introduction of hearing aids with digital signal processing (DSP) may prove to be of great benefit to the hard of hearing (society).

The goal of the present report is to contribute to the knowledge that is necessary to successfully apply modern DSP in order to improve the performance of next generation hearing aids. The strategy has been to identify which characteristics of the speech signal are lost and thereby cause reduced intelligibility, and then, to design DSP algorithms that can be applied to the speech signal in order to improve the intelligibility.

The main part of the project has been to design and conduct listening tests. The results of these tests show that reduced intelligibility is often caused by errors in the perception of specific speech sounds. The consequence of even small errors in perception by a person suffering from hearing loss may be that the whole content of the speech message is lost. Algorithms that prevent such loss of parts of the speech signal, may be of great help.

Some promising new algorithms have been tested in this project. These are candidates for further development.

The project has been financed by the Norwegian Foundation for Health and Rehabilitation.

| KEYWORDS | ENGLISH | NORWEGIAN |
|---|---|---|
| GROUP 1 | Health | Helse |
| GROUP 2 | Speech | Tale |
| SELECTED BY AUTHOR | Hearing-aid | Høreapparat |
| | Algorithms | Algoritmer |

## Foreword

The project was initiated by Marie Krüger in the Bergen group of "Hørselshemmedes Landsforbund" (The Norwegian Association of the Hard of Hearing People).

The Institute of Telecommunications at NTNU - Norwegian University of Science and Technology has been responsible for the scientific part of the project. Professor Ph.D. Ulf Kristiansen has supervised the project on behalf of the institute. The project has been carried out in close collaboration with the Department of Linguistics at NTNU and SINTEF Telecom and informatics.

The project has so far been presented at two conferences, [ 10 ] and [ 11 ]

The authors of this report are Professor Dr. Wim van Dommelen, the Department of Linguistics at NTNU; Professor Dr.ing. Asbjørn Krokstad, the Institute of Telecommunications at NTNU; and Research Scientist Dr.ing. Sverre Stensby, SINTEF Telecom and Informatics.

Trondheim, 2002-02-05

Wim van Dommelen Asbjørn Krokstad Sverre Stensby

## Acknowledgement

This project has been financed with the aid of EXTRA funds from the Norwegian Foundation for Health and Rehabilitation.

The Norwegian Foundation for Health and Rehabilitation is an institution consisting of 21 nation-wide voluntary health and rehabilitation organisations. The Norwegian Foundation for Health and Rehabilitation is licensed to operate the national TV game EXTRA. Proceeds from the game go towards strengthening the work of voluntary humanitarian organisations in preventive health care, rehabilitation and research in Norway. All voluntary humanitarian/socially dedicated organisations and professional organisations for the functionally disabled in Norway may apply for funds regardless of membership in the foundation.

In 2000 the Norwegian Foundation for Health and Rehabilitation distributed NOK 187 million to 630 different projects. EXTRA funds were first distributed in 1997. By November 2000 the Foundation for Health and Rehabilitation have allocated NOK 720 million to projects in the following areas: Somatic disorders, mental illness, measures for the functionally disabled, the chronically ill and their families, measures to improve living conditions and other areas related to health.

A scientific advisory group has been associated with the project. The members were: Research scientist Olav Kvaløy, SINTEF; Asc. professor Dr. Tormod Njølstad, NTNU; Professor Dr. Torbjørn Svendsen, NTNU; Professor Dr. Peter Svensson, NTNU; Asc. professor Jan Tro, NTNU; and University lecturer Jon Øygarden, HiST. The group has been of great help, especially in the planning phase of the project.

Several students have been engaged in the project: Jørn Almberg, Snefrid Holm, Bjørn Fredrik Meyer, Bård Støfringsdal, Kjell Heggvold Ullestad, and Andreas Vikran.

Credit is given to the persons participating in the recording of the sentences: Tor Martin Antonsen, Anne Berit Fagernes, Petter Haugereid, Katja Elisabeth Holm, Snefrid Holm, Bjørn F. Meyer, Tone Midtgård, and Kjell Heggvold Ullestad.

Credit is also given to the 48 listeners that on a voluntary basis participated in the listening tests. The Trøndelag group of the Norwegian Association of the Hard of Hearing People has been of great help in establishing contacts with these subjects.

## Summary in Norwegian / Sammendrag på norsk

Rapporten (på engelsk) gir en fullstendig dokumentasjon av arbeidet innen prosjektet "Algoritmer for høreapparat", som er gjennomført på oppdrag fra Norske Hørselshemmedes Landsforbund og finansiert av Helse og Rehabilitering.

Prosjektets hoveddel er en oppfattbarhetstest av norsk tale med hørselshemmede lyttere.

Med utgangspunkt i lyttetesten er det arbeidet med utvikling av algoritmer som forbedrer oppfattbarheten, med henblikk på at disse kan implementeres i fremtidige digitale høreapparat.

Det er også utformet en veiledning med råd vedrørende å gjøre seg forstått overfor hørsels-hemmede, [ 17 ]. Denne er spesielt rettet mot grupper som i yrket må gjøre seg godt forstått av personer med hørselssvekkelse, både de som bruker høreapparat og de som ikke gjør det.

Opplegget for oppfattbarhetstesten er at et innspilt (lagret på harddisk) talemateriale ble avspilt over høyttaler i et rom med etterklangstid i samme område som en typisk stue. Innflytelsen av etterklangen ble variert i to trinn ved å bruke en kort og en lang avstand mellom høyttaler og lytter. Bakgrunnsstøyen ved lytteren ble også variert ved å utstråle støy som fra et stort antall personer som snakker samtidig. Det ble brukt tre ulike nivåer av bakgrunnsstøy slik at det totalt ble seks ulike forsøksbetingelser.

Talematerialet er 150 grammatikalsk riktige setninger med kjente norske ord i semantisk uforutsigbare kombinasjoner. Lytteren har da små muligheter for å gjette ord fra sammenhengen. Talematerialet er lest inn av 8 personer med ulik språklig bakgrunn. Flere dialekter er representert i materialet.

Oppfattbarheten ble testet i tre forskjellige tester med ulike vanskelighetsgrader. Testene er benevnt setningstest *(sentence test)*, ordtest *(word test)* og flervalgstest *(forced choice test)*. Det vanskeligste var setningstesten der lytteren ble bedt om å gjenta hele setning. En betydelig enklere test var ordtesten der setningen sto på skjermen, men ett ord var utelatt. Dette skulle lytteren fylle ut. I den enkleste testen, flervalgstesten, skulle lytteren velge mellom fire alternative ord for det som var utelatt. I alle testene skrev forsøksleder svaret på dataskjerm foran lytteren, som så bekreftet at det som sto var hva han hadde hørt.

I oppfattberhetstester er det viktig at lytteren oppfatter noe, men ikke alt, for at testen skal gi informasjon om hva som skaper de største problemene. Det ble derfor variert mellom de seks akustiske betingelsene og vekslet mellom de tre testene i løpet av en økt med lytting. Hvis lytteren ikke fikk noe riktig for hele setninger, gikk forsøksleder over til testen med ett utelatt ord som lytteren skulle fylle inn. Ble det for mange feil også her, gikk en over til testen med fire alternativer.

Til sammen har 48 personer deltatt. De er delt i tre grupper; personer som bruker høreapparat, personer over 55 år som ikke bruker høreapparat, og en referansegruppe med yngre mennesker som er normalthørende (høreterskel innen standardavviket for personer i alder 18-25 år).

Resultatene av de tre typene av lytteprøver foreligger både i form av tabeller som viser forveksling mellom de enkelte språklydene og i sammenfattet form. Følgende observasjoner er av størst interesse i arbeidet med algoritmer for høreapparat:

Setningstest

- Det forekom svært få feil når det gjelder ordgrenser, med andre ord lytterne klarte stort sett å dele opp talestrømmen i enkelte ord. På linje med denne observasjonen ble det ikke registrert noen feil med hensyn til stavelsesgrenser. På dette grunnlaget anbefales det at man i arbeidet med algoritmer først og fremst fokuserer på oppfattbarhet av fonem heller enn ord.

- Oppfattbarhet av ord ble påvirket av ordets posisjon i en ytring. I ytringsinitial posisjon forekom det færrest feil. Dobbelt så mange feil ble det registrert for ytringsfinale ord, mens ytringsmediale ord var enda vanskeligere å oppfatte.
- De kvinnelige talerne var litt bedre å forstå enn de mannlige. Effekten var signifikant, men ikke så stor (3-4% feilratedifferanse mht. ordgjenkjenning).
- Det viste seg at dialekt til innleseren påvirket forståelighet. De færreste feil ble registrert for de sør-østnorske innleserne, uansett dialekt til lytteren. Siden det totalt bare var åtte innlesere involvert i forsøkene, er dette resultatet ikke mer enn en første indikasjon.

Ordtest og flervalgstest

- Begge to testtyper viste liknende resultater, men på grunn av formatet med fire gitte svaralternativer i flervalgstest var forvekslingsmulighetene her begrenset. Konklusjonene baserer derfor hovedsakelig på ordtesten.
- Feil i ordforståelsen skyldtes forvekslinger av både konsonanter og vokaler. Forholdet av gjennomsnittlig feilkvote for vokaler/konsonanter var ca. 0.8
- Feil i vokalpersepsjon var hovedsakelig forårsaket av forvekslinger av trekkene *runding (rundet vs. urundet)* og *artikulasjonssted (framre-bakre)*. Rundete vokaler ble oppfattet som urundet heller enn omvendt. Feil mht. artikulasjonssted forekom spesielt ved *midtre* og i mindre grad ved *bakre* vokaler. *Åpningsgrad* viste seg hovedsakelig å være relevant for *åpen-midtre* vokaler.
- Konsonantforvekslinger skyldtes primært feil når det gjelder *artikulasjonssted*. *Artikulasjonsmåte* spilte en sekundær rolle, mens feil mht. *stemthet* forekom svært sjelden.
- Stort sett forekom de fleste feilene ved nasaler, fulgt av plosiver, frikativer og, til slutt, likvider. Feil ved de første tre lydkategoriene skyldtes mest *artikulasjonssted.* For likvidene /r/ og /l/ derimot var *artikulasjonsmåte* avgjørende; oftest ble /r/ forvekslet med /l/ og omvendt.

# Table of Contents

# 1 Introduction

The goal of the present study is to shed some light on the problems encountered by hard of hearing people in understanding speech. Such information is an urgently needed prerequisite for improving the efficiency of modern hearing aids. Before the introduction of digital signal processing, the functioning of hearing aids has been restricted to bringing the speech spectrum above the users' threshold by frequency- and level-dependent amplification. Modern digital hearing aids, however, allow the implementation of speech enhancement algorithms that treat the speech signal in a far more sophisticated way. Efficient development of such algorithms heavily depends on our knowledge about the main phonetic factors contributing to reduced speech intelligibility.

At the outset of our study, it was decided to focus on the following issues:

- In order to be able to understand a spoken message, the listener has to parse the stream of speech into smaller units (phrases, words, syllables and possibly even single speech sounds). Relatively little is known about the extent to which such parsing is a stumbling-block for hard of hearing people when listening to running speech.
- It might be speculated that word recognition varies with the position of the word in an utterance. In the course of listening to an utterance, a listener receives a steadily increasing amount of information about factors like the speaker's speech rate, speech rhythm and spectral characteristics. Especially, the semantics of an utterance allows the listener to make predictions about the words still to be heard. For example, given a first part of an utterance like "Finally the keeper managed to catch..." the listener will in all probability predict the phrase "the ball" to occur. Thus, one would expect more errors to occur in the recognition of utterance-initial than utterance-medial or –final words. However, since we wanted to focus on purely phonetics aspects of word recognition the use of semantics was eliminated in our study. This was done by developing syntactically correct, but semantically unpredictable sentences as speech material (see Chapter 2). Another advantage of using such sentences is that listeners will not be able to remember them, so the same set of sentences and listeners may be used for example in future research.
- A further issue is to what extent the different classes of speech sounds (phonemes) contribute to recognition errors in Norwegian. First of all, a distinction between consonants and vowels can be made. Within these categories, subclasses can be distinguished (like the four classes plosives/fricatives/nasals/liquids within the main class of consonants). We wanted to investigate which of the (sub)classes of phonemes are most difficult for hard of hearing people to identify.
- Going into even more detail, an important research question is concerned with the role of single phonetic features that constitute the vowel and consonant phonemes in Norwegian. Similar to the investigation of the classes of phonemes, our goal was to reveal the relative importance of the different features for intelligibility. An example of a consonant feature is place of articulation, distinguishing between the Norwegian speech sounds "p" (labial), "t" (alveolar) and "k" (velar). For the vowels, lip rounding can be mentioned as an example, distinguishing "i" (as in "vin") from "y" (as in "lyn"), "e" (as in "sett") from "ø" (as in "søtt"), etc.
- Finally, we wanted to look into the role of dialectal background in speech intelligibility. In Norway, there is a widespread use of local dialects rather than a standard language. Therefore, it can be expected that certain groups of listeners have more difficulties in understanding dialects from other regions compared to their own. Unfortunately, the limited size of our study prevented us from investigating the role of dialect systematically. Nevertheless, through the

inclusion of speakers and listeners with varying dialectal background we hoped to acquire some knowledge in this respect. The use of a group of speakers also reduced the possible influence of speaker idiosyncrasies on speech intelligibility.

An important methodological aspect of this study was concerned with the acoustic room conditions. In daily life, the acoustic quality of spoken language is often deteriorated by concurring background noise and reverberation. So, in order to achieve ecologically valid listening test results both background noise and room acoustics were included as factors (see Chapters 5 and 8). Apart from this function, variation of the acoustic conditions was used instrumentally to adapt the level of speech intelligibility to the individual listeners. Pilot experiments have shown that for listeners with a severe hearing loss a strongly degraded speech signal easily will reduce intelligibility to zero. On the other hand, for listeners with a moderate hearing loss and moderately degraded signal quality intelligibility rates up to 100% will be found. It is obvious that both type of test results are not very conclusive. Therefore, in the present tests a varying degree of deterioration of the speech signals was used to achieve an appropriate amount of perception errors for the different listeners.

## 1.1 This report

The methods used in the project and the results are documented in this report.

In the chapters 2 — 5 the speech material, the speakers, the recording, and the controlled masking noise is documented. The design of the listening test, characterisation of the listeners, and the conditions used in the listening test is described in the chapters 6 — 8. The results from the listening test constitute a major part of the report. These results and conclusions are given in the chapters 9 and 10. The results from the listening test are used in two very different applications, i.e. to design enhancement algorithms for digital hearing aids and in advice regarding talking to hard of hearing people. This is presented in the chapters 11 and 12. The report is concluded with a discussion of future work, appendices and a guide to an accompanying CD-rom.

## 2 Speech material

The speech material was being developed having in mind the following considerations. Ideally, the listening test situation would involve perceiving running speech, for example taken from radio broadcasts. Using such material, the listeners' task could be to indicate parts of speech that were difficult to follow or not understandable at all. It will be clear, of course, that such a design would not allow any systematic investigation. A decisive drawback would be the varying degree of guidance of the listener by the semantics of the utterances to be understood. The extreme alternative to the use of running speech would be to develop systematically varying logatomes (constructed nonsense words), which are being presented in isolation. While having the advantage of allowing specific research questions, the validity of the material as to inferences with regard to everyday-listening situations would be strongly reduced.

Trying to avoid the constraints of the approaches mentioned above, it was decided to develop test sentences having the following characteristics: The sentences were both grammatically and syntactically correct, but were semantically unpredictable (though containing real Norwegian words only). To reduce the possible influence of short-term memory constraints, the sentences contained between (minimally) four and (maximally) seven words. An example is the six-word sentence:

> *Flisen påstår at pillene kunne kyle.*
>
> *('The splinter claims that the pills could fling.')*

Note that the absence of semantic information makes the task of word recognition and, as a consequence, also the recognition of the boundaries between words more difficult. A further advantage of using semantically unpredictable sentences is that listeners will not be able to remember them. Therefore, the listening tests can be repeated with the same listeners, for example after implementation of improved hearing aid algorithms.

In order to be able to investigate the perception of single speech sounds / speech sound categories, minimal pairs like *paver, maver, gaver, haver* (testing the /p-m-g-h/ opposition) were built in into the test sentences as shown in the following example:

| Sentence in Norwegian | Literal translation |
|---|---|
| *Raske paver triller ofte.* | *Fast popes roll often.* |
| *Rare maver griller nå.* | *Strange stomachs grill now.* |
| *Gode gaver traver gjerne.* | *Good gifts trot gladly.* |
| *Kalde haver graver mest.* | *Cold gardens dig mostly.* |

Using only real words rather than logatomes in the test sentences limited the number of minimal pairs to be included. The *forced choice test* described below was designed to focus on the question how the perception of distinctive features is affected by hearing deficiencies.

The whole set of 150 test sentences was built up starting from ten different prototypes. Differences between the prototypes were concerned with sentence structure, sentence type (declarative/question) and number of words (4-7). For each of the ten prototypes, blocks of 15 sentences were developed, thus resulting in a total of 150 sentences. In Appendix B, the sentences are listed together with block number, speaker code and sentence number.

## 3 Speakers

The stimuli were prepared by recording eight different speakers. By involving a group of speakers rather than a single one, both speaker sex and dialect region (East Norwegian, North Norwegian, Trøndelag, West Norwegian) could be introduced as experimental factors (cf. Table 1). Apart from allowing the investigation of these two factors, this design had the advantage of confronting the listeners with different voices, making the listening test situation more realistic. In addition, this design reduced the possible influence of speaker idiosyncrasies on speech intelligibility. The group of speakers consisted mainly of students aged between 20 and 30. None of them had any professional background as a speaker, with the exception of speaker AF, who was a trained radio speaker.

From each of the eight speakers the whole set of 150 sentences was recorded. Following the recordings, for each speaker 18-19 utterances were selected thus dividing the 150 sentences approximately evenly across the eight speakers (see Table 1). Care was taken to choose tokens without any disfluencies or speech errors. Having in mind this quality requirement as a primary criterion, each speaker's 18-19 utterances were divided as evenly as possible across the 10 prototype blocks, i.e. in most cases resulting in 1-2 per block.

**Table 1.** Information on speakers and distribution of the 150 test sentences across the 8 speakers and 10 prototype blocks. For the text of the sentences, see Appendix B.

| | | | Block | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Speaker** | **Sex** | **Dialect** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **Total** |
| TA | male | North Norwegian | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | **19** |
| PH | male | Trøndelag | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | **18** |
| BM | male | East Norwegian | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | **19** |
| KU | male | West Norwegian | 1 | 2 | 1 | 2 | 2 | 2 | 3 | 2 | 2 | 2 | **19** |
| AF | female | Trøndelag | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | **19** |
| KH | female | West Norwegian | 2 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | **18** |
| SH | female | East Norwegian | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | **19** |
| TM | female | Trøndelag | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | **19** |
| | | | **15** | **15** | **15** | **15** | **15** | **15** | **15** | **15** | **15** | **15** | **150** |

# 4 Speech recording and processing

The preparation of the speech material consisted of high quality recording succeeded by sentence selection and speech signal processing.

## 4.1 Speech recording

The recording sessions took place in the studio of the Department of Linguistics at the NTNU. The studio is an acoustically dry room with low background noise and thus well suited for speech recording. The recording chain consisted of the following hardware, software and settings:

- Milab LSR 1000 microphone.
- Fostex D-10 digital recorder with 16 bit resolution and 44.1 kHz sampling rate.
- Soundcard DigiDesign in an Apple Macintosh computer.
- SoundDesigner software.
- Cool Edit Pro, a multitrack digital audio recorder, editor, and mixer program.

The microphone was positioned at a distance of approximately 35 cm from the speakers mouth and directed towards the speaker. Seen from the speaker, the microphone was 30º to the right of the direction of speaking.

The recordings were first saved as .AIF-files (Apple Mac AIFF sound format) and afterwards converted to .WAV-files (Windows Waveform sound)[1].

The speakers read the sentences from a computer screen in front of them. The sentences were shown one after another in a randomised order. The order was different for each of the readers.

The leader of the experiment determined the progress of the recording session. If the rendering of a sentence was judged to be unsatisfactory, the leader could move back and forth in the text file to have the speaker repeat the sentence in question.

The speakers were instructed to read the sentences fluently but at the same time not too fast, comparable to the style of a news broadcast. The first two speakers were assistants involved in the project. Final recordings of their voices were made after some pilot listening tests run to find the most appropriate speaking style. To minimise reading style differences between the eight speakers selected parts of these first accepted recordings were presented for the other readers.

## 4.2 Signal processing

The purpose of the signal processing was to establish the same speech signal level for all the recordings. The process had the following steps:

- Highs pass each of the recorded speech sentences from 100 Hz to remove low frequency noise components. The filtering was done by "Scientific filters" in the program Cool Edit Pro.
- Make a copy of each of the speech sentences filtered with an "A"-filter. The filtering was done by "Scientific filters" in the program Cool Edit Pro.
- Measure the energy of the filtered sentences.
- Correct the amplitude of the original individual sentences to achieve the same speech level. The amplitude correction was done by the "Amplify" function in the program Cool Edit Pro using one the utterances as an arbitrary reference.

---

[1] Acronym definitions from: http://www.ace.net.nz/tech/TechFileFormat.html (2001-12-04).

## 5 Masking noise

The intention of using masking noise in the listening tests was to achieve a typical background noise and systematically control the noise level. The noise level was one of the means used to establish the intended level of difficulty in the tests.

More noise types were considered. Among these were frequency shaped random noise, traffic noise and speech babble. Interfering speech is a common noise that is especially difficult as it by its nature occupies the same frequency domain as the desired speech signal. This is the main reason for choosing this signal.

The speech babble was generated by mixing more independent recordings. The individual recordings are taken from Norwegian radio broadcasting of conversations and other types of speech. All together speech from 6 male and 5 female speakers was included. All the segments had duration of a few minutes. To virtually increase the number of segments they were mixed both direct and in a 5 second delayed version. The final result was speech babble where individual voices and words were almost indistinguishable.

## 6 Listening test design

One of the most crucial parts of the project has been to design an experimental set-up for the listening tests. Pilot studies showed that listeners' ability to understand the speech presented in an experimental situation varied considerably, both between listeners and dependent on experimental conditions (e.g., for speech with concurrent noise). In a listening task where subjects had to understand complete utterances, their performance typically was a matter of "all or nothing". In other words, recognition rates were either close to 100% or almost 0%. As a consequence, with such a design it is impossible to systematically gather information about which parts of the speech signal are difficult to understand in general and most sensitive to noise in particular, and for what reason.

As a solution to the problem described above, three different types of listening tests were designed: *forced choice test, word test* and *sentence test*. While all three tests use the same speech material, the differences between them lie in the listener's task. The varying degree of difficulty of the listening tasks makes the design flexible and apt to cope with variation between listeners and acoustic listening conditions. From an information collecting viewpoint it is important that the tests vary as to the level of detail of investigation.

### 6.1 Forced choice test

With this test type the listener is presented with a sentence written on a computer screen. One of the words in the sentence has been replaced by a horizontal line. The utterance belonging to the sentence is presented via a loudspeaker. The task of the listener is to identify the word that has been deleted in the orthographic representation by choosing one of four alternative words written under the horizontal line in the sentence. The alternatives represent minimal pairs or almost minimal pairs. An example (with phonemic transcription of the alternatives indicated for the present purposes only):

*Bilen vet at barna ville* _____

| | |
|---|---|
| *hale* | [hA:l@] |
| *helle* | [hel@] |
| *holde* | [hOl@] |
| *hyle* | [hy:l@] |

The SAMPA (Speech Assessment Methods Phonetic Alphabet) is used, [ 3 ]. The conventions used in the transcription in this report is given in Appendix D.

A complete list of the 150 test sentences plus the alternatives can be found in Appendix C. The forced choice design makes it possible to elicit responses even from severely hearing-impaired listeners and to investigate possible sound confusions under all conditions. This design allows therefore a systematic investigation of the influence of hearing impairment and noise on perception of single features like consonant voicing or vowel quantity. Table 2 presents an overview of the occurrence of phonemic feature oppositions built in into the answering alternatives belonging to the 150 test sentences. As can be seen from the table, there is a better representation of oppositions for the consonants as compared with the vowels. Also, the number of cases with word pairs distinguished by only one single phonemic feature (like vowel quality in *helle* [hel@] vs. *holde* [hOl@]) is rather restricted compared to the number of cases where two or more features are involved (like vowel quality and vowel quantity in *helle* [hel@] vs. *hale* [hA:l@]. Though the issue of phonemic oppositions has not been studied explicitly for the present investigation, it sees reasonable to assume that the distribution of oppositions as represented in the *forced choice* test material roughly reflects their occurrence in modern Norwegian.

**Table 2.** Occurrence of phonemic feature oppositions in the forced choice test. The category single denotes phoneme pairs differing in only one distinctive feature (e.g., /f-s/ in fanger-sanger; place of articulation) as opposed to one of more (e.g., /f-g/ in fanger-ganger; voicing and place/manner of articulation). Vowel quality comprises both the front-back and the close-open dimension.

| | **Consonants** | | | **Vowels** | | |
|---|---|---|---|---|---|---|
| **Distinctive feature** | **voicing** | **place** | **manner** | **quantity** | **quality** | **rounding** |
| **single** | 20 | 52 | 24 | 1 | 23 | 9 |
| **one of more** | 164 | 175 | 193 | 11 | 36 | 30 |

### 6.2 Word test

This test type bears a strong resemblence to the *forced choice test* and represents the next step on a scale of difficulty for the listener. Also with this test, listeners are being presented with a sentence written on a screen where one word has been replaced by a horizontal line. Here, however, no answering alternatives are being given. See the following example:

*Bilen vet at barna ville* _____

Due to the absence of forced choice alternatives, this test type is more demanding for the listener. At the same time, test evaluation is more complicated because of the increased number of degrees of freedom. The increased potential of sound confusions, auditory epenthesis and deletion is, of course, advantageous from a researcher's viewpoint.

### 6.3 Sentence test

The third test type is the most demanding one for the listener, since the utterances are being presented acoustically only, without any orthographic help. The listener's task is to determine the content of the utterance. Due to the free test format, this test has the potential of discovering all possible kinds of confusion phenomena.

## 7 Listeners

A total of 48 listeners participated in the listening tests.

### 7.1.1 Subgroups

The group was divided into three subgroups according to hearing loss history:

1) Group **AH** (Aided Hearing): 21 hearing-impaired listeners, who used their hearing aids during the listening tests. They were aged between 37 and 87 years (mean= 66.1; standard deviation= 15.4).
2) Group **UAH>55** (UnAided Hearing): Reference group consisting of 20 listeners aged between 56 and 84 (mean= 68.3; standard deviation= 5.5 years). None of them was wearing a hearing aid during the tests, but four participants were in possession of a hearing aid. The other 16 listeners had no reported hearing loss history.
3) Group **REF<55**: Reference group consisting of 7 listeners aged between 22 and 52 (mean= 31.7; standard deviation= 10.7 years). None of the listeners had a reported hearing loss history.

### 7.1.2 Result based verification of subgroup classification

Since it turned out to be too time-consuming to make audiograms of all the subjects, the results of the *word test* were taken as an *a posteriori* justification of the classification into subgroups. This was done in the following way. For each individual listener in the *word test*, for each of the six acoustic conditions (2 loudspeaker distances x 3 noise levels) the number of stimuli and the number of misperceived words (i.e., responses containing one or more misperceived phonemes) were determined. Adding these numbers for all six conditions, a (weighted) average score of misperceived words was calculated. To illustrate this procedure let us consider the results for listener Nr. 1 (group AH):

| Loudspeaker | Noise level | # of misperceived words | # of stimuli |
|---|---|---|---|
| far | low | 3 | 8 |
| | medium | 2 | 6 |
| | high | 2 | 3 |
| close | low | 3 | 8 |
| | medium | 2 | 8 |
| | high | 2 | 6 |
| **total** | | **14** | **39** |

Thus, for this listener in total 14 out of 39 responses contained one or more misperceived phonemes, which corresponds to 35.9%.

Similarly, for each of the resulting 47 subjects individual recognition rates were calculated. The advantage of the solution presented here is that the results are comparable in an absolute sense and therefore reflect the individual listeners' hearing abilities. The results are presented in Figure 1. For individual data, see Appendix E.

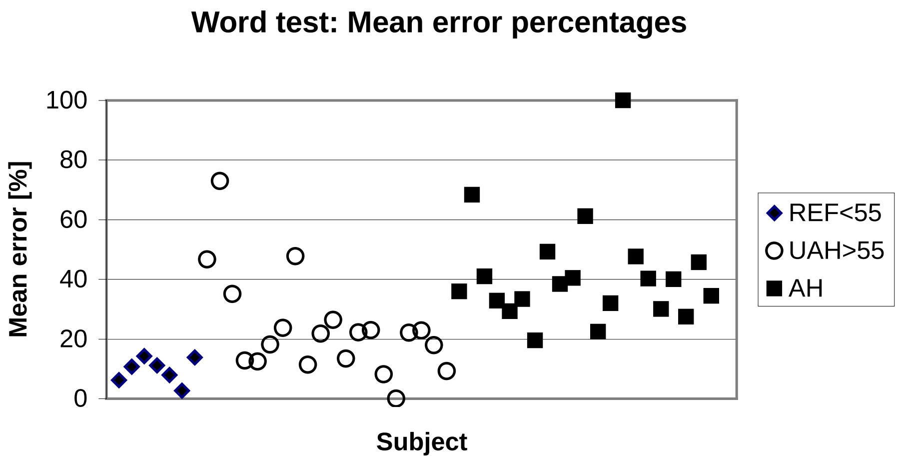


**Figure 1.** Mean error rates in the *word test* for listeners from the three listener groups REF<55, UAH>55 and AHA.

As can be seen from the figure, the results can be said to be in fairly good accordance with the "intuitive" classification from beforehand. There was no overlap between the error percentages for the reference group REF<55 and the group with hearing aid users (AH). As could be expected, the results for the listeners from group UAH>55 showed a strong scatter, with mean percentages varying between 0 % (!) and 73 %. It is important to note that the group AH, in spite of using hearing aids, by and large achieved worse recognition rates than the nearly same aged group UAH>55.

Speech audiometry is in regular use to characterise the hearing ability of persons wearing modern hearing aids. A simplified version of such a test is included in the listening test performed.
A characterisation of the individual listeners based on this method is given in Appendix E.

# 8 Listening test room, equipment and conditions

The listening tests were carried out in a conference room provided with appropriate equipment. The room was sparsely furnished and had the dimensions 6*6*3 meters. The acoustic characteristics of the room are described in section 8.3.

## 8.1 Test equipment and presentation conditions

The primary units were loudspeakers, a computer, and recording equipment. The computer constituted a central part.

There were three signal paths:

1. Speech signal from the computer disk via the computer mixer to one of two active speech loudspeakers in front of the listener.
2. Noise signal from an audio CD-drive via the computer mixer, an adjustable attenuator, and split in one direct and one delayed path to two active noise loudspeakers.
3. The acoustic signal in the room picked up by microphones in an artificial head went via a microphone amplifier and was recorded on DAT and direct to computer disk in parallel.

The speech loudspeakers were situated in front of the listener at two different distances denoted "Near" and "Far". The distances were 1.3 and 4.1 m respectively. In accordance with the selected distance condition, the speech signal was fed to the appropriate loudspeaker.

The noise signal was reproduced by two loudspeakers under the chair on which the listener was situated. Due to the delay in the two-channel reproduction, the origin of the noise signal was hard to identify. It sounded like the noise was diffuse and was almost impossible to locate. Three different noise levels were used. These were denoted "Low", "Medium" and "High".

The purpose of the recording was to have the chance of listening to the presentation and answer for inspection in the future. The artificial head was positioned close to the person participating in the test.

The combination of the two distances to the speech loudspeakers and the three noise levels gives at set of six conditions.

## 8.2 Test sequence

The written information given to the listener during the test was presented on a computer screen in front of the listener and the test conductor.

When the test session was started the test conductor set the appropriate conditions (speech loudspeaker and level of the masking noise) and the relevant information for the selected test type was presented in large font on the screen. The leader of the experiment started the acoustic presentation by pressing the start button initialising the following stages:

1. The masking noise was turned on.
2. The utterance was presented in accordance with the sequence file.
3. The masking noise was turned off.
4. The five-second direct-to-disk recording was started.

The listener dictated his answer immediately after the sentence was presented. The five seconds of direct-to-disk recording was intended for capturing the answer for possible further analysis. The answer may also be inspected later by listening to the DAT recording if it failed to be caught by the five seconds window. The leader of the experiment typed the answer on the keyboard, and the listener read it on the screen for verification.

The leader of the experiment then initiated the next presentation, and the test went on. Breaks were taken when desired.

### 8.3 Characterisation of listening test conditions.

Three physical parameters are affecting the listening test conditions; speech level, noise level and reverberation.

Reverberation may be characterised by the reverberation time. For practical reasons the reverberation time was kept constant during the test. Even the speech level emitted from the loudspeakers was kept constant. The effects of reverberation is varied by using two distances from listeners to the loudspeaker, thus varying both the level of reflected sound relative to the direct sound, and the total level of speech.

All these effects may be characterised numerically by a single parameter; the Speech Transmission Index (STI). The STI was introduced by Houtgast and Steeneken in 1980 and is now standardised [ 2 ].

STI is based on the measurement of the Modulation Transfer Function (MTF), which describes the ability of a system to transfer the variations in the signal envelops. MTF may be described in the time domain, but is most commonly given in the frequency domain. The envelop spectrum of the speech signal has most of the important spectral power in the frequency range below 30 Hz, with 4-7 Hz as the most important range. Reverberation, noise and non-linear distortion influence the MTF.

MTF is found as the Fourier Transform of the squared impulse response, measured by using WinMLS, a PC-based measuring system developed by Morset Sound. The measured STI has to be corrected for Signal to Noise Ratio (SNR). SNR is measured separately by an integrating sound level meter[2] using sequences of recorded speech and recorded noise. All measurements were made in the head position of the listener without the subject being present.

MTF is found, according to Schroeder, as the Fourier Transform of the squared impulse response, for each octave band 125-8000 Hz. STI based on the influence of reverberation alone (the impulse response not corrupted by noise) is called $STI_{imp}$.

The formula below presents the basis of calculations. WinMLS perform these calculations.

***STI based on the impulse response with noise correction:***

$$STI(f) = \frac{\int_{-\infty}^{\infty} h^2(t)\cdot e^{-j2\pi ft}\,dt}{\int_{-\infty}^{\infty} h^2(t)\,dt} \bullet \frac{\overline{p}_s^{\,2}}{\overline{p}_s^{\,2} + \overline{p}_n^{\,2}} = STI_{imp}(f) \bullet \frac{1}{1+10^{\frac{L_n - L_s}{10}}}$$

Legend of the variable:

| | | | |
|---|---|---|---|
| t: | time | $p_s$: | sound pressure of signal |
| f: | frequency | $p_n$: | sound pressure of noise |
| | | $L_s$: | SPL of signal |
| | | $L_n$: | SPL of noise |

The results of the measurements are given in Table 3.

---

[2] Norsonic integrating Sound Level Meter (Norsonic Type 110)

**Table 3.** Calculated STI (Speech Interference Index) for the different listening conditions used in the listening test in "Studio". Speech Levels and Noise Levels are measured by Norsonic 110, STI(imp) by WinMLS.

| Frequency-octaves | 125 | 250 | 500 | 1000 | 2000 | 4000 | 8000 | A-weighted | Correc-tions. | STI (imp) | **STI corr.** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Speech greatest distance ("Far") level dB | 53 | 58 | 59 | 56 | 50 | 43 | 34 | 60 | | | |
| "Near" speech | 66 | 69 | 70 | 61 | 56 | 48 | 35 | 69 | | | |
| Noise low | 42 | 42 | 39 | 32 | 27 | 30 | 33 | 41 | | | |
| Noise medium | 52 | 57 | 53 | 45 | 37 | 35 | 34 | 53 | | | |
| Noise high | 57 | 61 | 58 | 51 | 41 | 39 | 36 | 58 | | | |
| Corrections Far-Low | 0.926 | 0.975 | 0.99 | 0.996 | 0.995 | 0.952 | 0.557 | 0.988 | 0.982 | 0.78 | **0.767** |
| Corrections Far-Mid | 0.557 | 0.557 | 0.799 | 0.926 | 0.952 | 0.863 | 0.5 | 0.834 | 0.820 | 0.78 | **0.639** |
| Correction Far-High | 0.285 | 0.334 | 0.557 | 0.76 | 0.888 | 0.715 | 0.387 | 0.613 | 0.651 | 0.78 | **0.508** |
| Corrections Near-Low | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.000 | 1.000 | 0.84 | **0.840** |
| Corrections Near-Mid | 0.888 | 0.941 | 0.98 | 0.975 | 0.988 | 0.952 | 0.557 | 0.9754 | 0.967 | 0.84 | **0.813** |
| Corrections Near-High | 0.888 | 0.863 | 0.941 | 0.909 | 0.969 | 0.888 | 0.443 | 0.9264 | 0.914 | 0.84 | **0.767** |

Loudspeaker distances, speech levels and noise levels were selected in a pilot test. The intention was an intelligibility in the range 0.3-0.9.

The relation between STI, Articulatin loss of Consonants ($Al_{cons}$) and classification of listening conditions are given in Figure 2.

$Al_{cons}$ is described in [ 9 ].

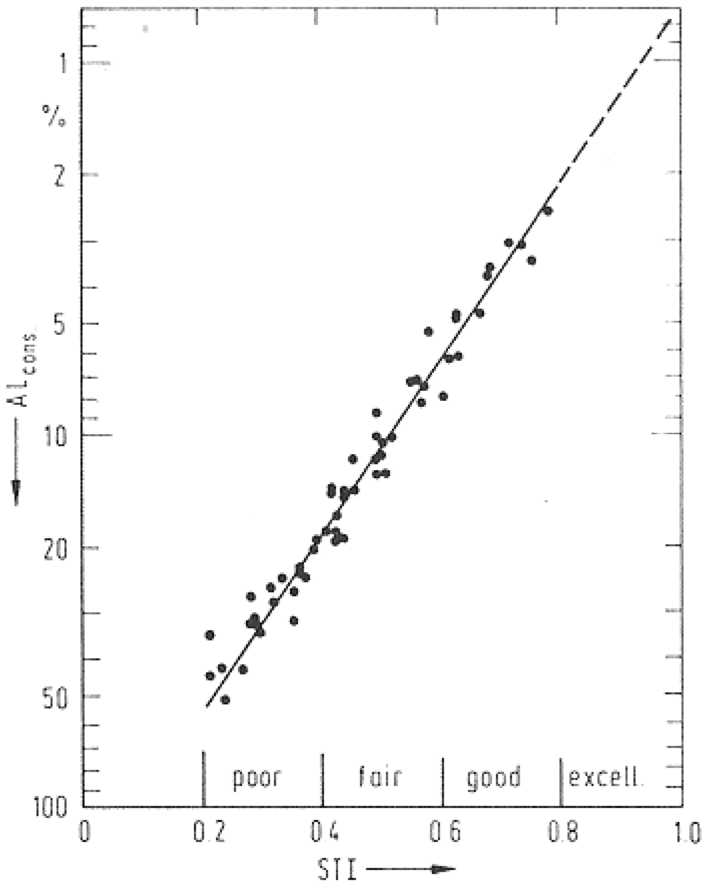


**Figure 2.** Relation between STI, Articulatin loss of Consonants ($Al_{cons}$) and classification of listening condition, [ 1 ] and [ 12 ].

# 9 Listening tests - Results

The listeners' ability to understand the test utterances was investigated applying three different test formats: the *forced choice test*, the *word test* and the *sentence test*. The use of these increasingly difficult test formats was dependent on individual listeners' degree of hearing impairment. From Table 4 it can be seen that both the *forced choice test* and the *word test* were run with all members of the three listener groups (n= 7, 20, and 21, respectively). Not all listeners, however, took part in the more demanding *sentence test*: 16 out of 20 listeners from group UAH>55 and only 13 out of 21 from the aided hearing group (AH).

The differences in hearing between the three groups are also reflected by the distribution of test types to which the listeners were exposed. The younger listeners from group REF<55 judged approximately twice as many stimuli in the *sentence test* (mean= 78.3 stimuli/listener) as in the two other tests (36.0 and 35.7 stimuli/listener). For the older group UAH>55 the distribution is almost level with mean numbers of 48.7, 52.7, and 53.9 stimuli per listener in the *forced choice*, *word*, and *sentence test*, respectively. The hearing-impaired judged most stimuli using the *forced choice* and *word* format (66.8 and 61.8 stimuli/listener), whereas in the *sentence test* the number of stimuli per listener was considerably less (21.2).

**Table 4.** Number of listeners (n) and mean number of stimuli per listener (S/L) used with three listener groups and three different test types. REF<55= reference group <55 years; UAH>55= unaided hearing group >55 years; AH= aided hearing group.

| | | forced choice | | word test | | sentence test | |
|---|---|---|---|---|---|---|---|
| Listener group | n | S/L | n | S/L | n | S/L | n |
| REF<55 | 7 | 36.0 | 7 | 35.7 | 7 | 78.3 | 7 |
| UAH>55 | 20 | 48.7 | 20 | 52.7 | 20 | 53.9 | 16 |
| AH | 21 | 66.8 | 21 | 61.8 | 21 | 21.2 | 13 |

## 9.1 Sentence test

### 9.1.1 Overall performance

Table 5 gives an overall impression of the listeners' performance in the *sentence test*. It appears that group REF<55 performed best, having a mean error rate of 9.6% words per utterance (i.e., statistically one word out of 10 was perceived incorrectly). The error rates for the two other listener groups are approximately twice and three times as high, respectively. In line with this observation, the percentage of utterances containing one or more misperceived words increases going from REF<55 via UAH>55 to AH (31.4%, 52.7%, and 64.4%). It is important to keep in mind that, since the acoustic listening conditions varied across groups and listeners, these results do not reflect absolute abilities to perceive the stimuli correctly. For example, group REF<55 judged 76.8% (421 out of 548) utterances under the less favourable *back* loudspeaker position. While a similar percentage (77.6%; 670 out of 863) is found for UAH>55, the proportion for the AH group was merely 30.9% (85 out of 275). From these results it can be concluded that the hearing differences between the three listeners groups are still larger than is being suggested by the data presented in Table 5.

**Table 5.** Error distribution in the sentence test for three listener groups. Mean percentages of misheard words/utterance, utterances with one or more misheard words/no errors and number of stimuli judged (total and per listener).

| | | % utterances with | | Stimuli | |
|---|---|---|---|---|---|
| **Listener group** | **% misheard words** | **≥1 misheard words** | **no errors** | **n** | **n/listener** |
| **REF<55** | 9.6 | 31.4 | 68.6 | 548 | 78.3 |
| **UAH>55** | 19.9 | 52.7 | 47.3 | 863 | 53.9 |
| **AH** | 29.1 | 64.4 | 35.6 | 275 | 21.2 |

### 9.1.2 Speaker gender and dialect

This section looks into the influence of speaker gender and dialect on speech intelligibility. Prior to discussing the results it should be noted that due to differences in dialectal composition the two speaker groups are not completely similar (cf. Table 1 p4). Only the male group has a North Norwegian speaker, while in the female group there are two Trøndelag speakers. This means that the experimental design is not completely balanced.

From Table 6 it can be seen that the error rates for the male speaker group are generally somewhat higher than for the female speakers. According to an ANOVA with *speaker gender*, *speaker dialect* and *listener group* as factors, the effect of speaker gender is statistically significant (F(1, 1665)= 4.515; p= .034) with no significant interaction between speaker gender and listener group (F(2, 1665)= .457; p= .633).

**Table 6** The influence of speaker gender on intelligibility in the *sentence test*. Mean percentages of misheard words/ utterance and number of stimuli judged pooled across speaker dialect groups.

| Listener group | Speaker | % misheard words | n |
|---|---|---|---|
| **REF<55** | male | 11.7 | 267 |
| | female | 7.7 | 281 |
| **UAH>55** | male | 21.7 | 438 |
| | female | 18.1 | 425 |
| **AH** | male | 30.6 | 131 |
| | female | 27.8 | 144 |

Table 7 depicts the influence of speaker dialect on intelligibility for the three groups of listeners. Lowest error rates were found for East Norwegian, while the rates for North and West Norwegian were substantially higher. Error rates for the Trøndelag dialect were also relatively high, with the exception of the relatively low error rate for UAH>55 (18.2%). The effect of speaker dialect turned out to be statistically highly significant (F(1, 3)= 13.901; p< .001), whereas the interaction of this factor with listener group did not reach significance (F(6, 1665)= 1.625; p= .136).

**Table 7.** The influence of speaker dialect on intelligibility in the *sentence test*. Mean percentages of misheard words/ utterance and number of stimuli judged pooled across speaker gender. The number of speakers was North Norwegian: 1, Trøndelag: 3, and East/West Norwegian: 2 each.

| Listener group | Speaker dialect | % misheard words | n |
|---|---|---|---|
| **REF<55** | North Norwegian | 12.7 | 73 |
| | Trøndelag | 10.2 | 202 |
| | East Norwegian | 5.0 | 140 |
| | West Norwegian | 12.0 | 133 |
| **UAH>55** | North Norwegian | 27.8 | 116 |
| | Trøndelag | 18.5 | 309 |
| | East Norwegian | 13.1 | 224 |
| | West Norwegian | 24.8 | 214 |
| **AH** | North Norwegian | 30.9 | 27 |
| | Trøndelag | 31.4 | 111 |
| | East Norwegian | 19.7 | 66 |
| | West Norwegian | 33.8 | 71 |

### 9.1.3 Listener dialect

In this section we will investigate the interaction between the subjects' ability to understand utterances from varying dialects and their own dialectal background. At the outset of the project it was hypothesised that understanding one's own dialect would be the most easy task, while listeners would have a harder time coping with other dialects. Relevant data from the *sentence test* are presented in Table 8. Though not all the possible combinations of speaker/ listener dialect were represented in the data and some combinations were underrepresented (hearing-impaired listeners with East or West Norwegian background), the picture emerging from the data is interesting. In all three listener groups REF<55, UAH>55, and AH, the error rate distributions were very similar for the different listener dialects. Without exception, error rates were lowest for East Norwegian regardless of the listeners' dialectal background. Beyond that, congruence of speaker dialect and listener dialect (indicated in bold in the table) did reduce error rates only to a certain degree. Taking the results for West Norwegian listener dialect from REF<55 as an example, the error rate for West Norwegian utterances (13.5%) is higher than for East Norwegian (1.1%), while at the same time being comparable with the rates for North Norwegian (16.7%) and Trøndelag (13.0%). A case were listeners seem to have benefited from their dialectal competence is the category Trøndelag listener dialect from group UAH>55. Here, the error rate for Trøndelag utterances amounted to 12.9%, which is lower than for North and West Norwegian (22.8% and 20.5%, respectively) but still somewhat higher than for East Norwegian (11.1%). Similarly, the 21.5% error rate for the combination West Norwegian speaker/listener dialect from UAH>55 is, at the one hand, lower than for North Norwegian (30.4%) and Trøndelag (23.3%) but higher than for East Norwegian at the other hand (15.9%).

**Table 8.** The interaction of speaker dialect and listener dialect in the *sentence test*. Mean percentages of misheard words/ utterance and number of stimuli judged. Congruence of speaker and listener dialect indicated in bold. Percentages in brackets and small italics indicate too low number of observations to allow any conclusions.

| Listener group | Speaker dialect | Listener dialect | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | North | | Trøndelag | | East | | West | |
| **REF<55** | North Norw. | 10.3 | (25) | | | 13.4 | (40) | 16.7 | (8) |
| | Trøndelag | 13.8 | (52) | | | 7.9 | (121) | 13.0 | (29) |
| | East Norw. | 7.2 | (36) | | | 5.1 | (82) | 1.1 | (22) |
| | West Norw. | 12.1 | (39) | | | 11.6 | (75) | 13.5 | (19) |
| **UAH>55** | North Norw. | | | 22.8 | (52) | 32.4 | (46) | 30.4 | (18) |
| | Trøndelag | | | 12.9 | (137) | 22.9 | (107) | 23.3 | (65) |
| | East Norw. | | | 11.1 | (94) | 13.8 | (85) | 15.9 | (45) |
| | West Norw. | | | 20.5 | (86) | 30.8 | (86) | 21.5 | (42) |
| **AH** | North Norw. | | | 28.2 | (26) | | | (100.0) | (1) |
| | Trøndelag | | | 28.5 | (100) | (79.0) | (5) | (39.2) | (6) |
| | East Norw. | | | 17.0 | (56) | (23.3) | (6) | (52.5) | (4) |
| | West Norw. | | | 34.5 | (68) | (16.7) | (2) | (16.7) | (1) |

### 9.1.4 Linguistic factors

The data from the *sentence test* were also used to collect information about the occurrence of misperceptions related to various linguistic factors. These will be presented in the following sections.

#### 9.1.4.1 Sentence-internal word position

The open format of the *sentence test*, where the listeners had to rely on the acoustic signal only, allowed us to investigate the influence of sentence-internal word position on intelligibility. It was hypothesised that due to the absence of any other than acoustic (e.g., visual) indication of the start of an utterance sentence-initial words would be more error-prone than words in sentence-medial or final position. To test this hypothesis, for each of the three listener groups the number of words were counted that had been misperceived, specified for sentence-initial, medial and final position. The medial position is different from the other two word positions in that the number of words in that position varies and is always >1. To compensate for this, both one single misperceived word and two or more misperceived words in sentence-medial position were always counted as one single error for that sentence position.

As can be seen from Table 9, the working hypothesis was falsified by the data. For all three groups of listeners, least word confusions occurred sentence-initially. The highest error rates were found for the words in sentence-medial position, with the rates for words in sentence-final position lying in-between. The ratios of the error numbers for the three categories initial, medial, and final are approximately 1:3:2. According to $\chi^2$-testing, there are no significant differences between the distributions for the three listener groups ($\chi^2(4)= 1.233$; $p= .873$). In contrast, the influence of word position is highly significant (pooled across listener groups: $\chi^2(2)= 254.921$; $p< .0001$).

**Table 9.** *Sentence test*. The occurrence of misheard words as a function of the words' position in the sentence. The category medial indicates the occurrence of one or more words being misheard (annihilating the effect of the varying number of sentence-medial words).

| Listener group | Word position | Misheard words | | Stimuli |
|---|---|---|---|---|
| | | n | % | n |
| | initial | 38 | 6.9 | |
| REF<55 | medial | 123 | 22.4 | 548 |
| | final | 69 | 12.6 | |
| | initial | 110 | 12.7 | |
| UAH>55 | medial | 361 | 41.8 | 863 |
| | final | 184 | 21.3 | |
| | initial | 43 | 15.6 | |
| AH | medial | 143 | 52.0 | 275 |
| | final | 86 | 31.3 | |

### 9.1.4.2 Word and syllable level

This section focuses on the occurrence of errors related to the word and syllable level. It was hypothesised that due to the open format of the *sentence test* misperceptions would not be confined to the speech sound level, but also comprise the larger entities syllable and word. Specifically, it was expected that hearing-impaired listeners would to a relatively high degree experience difficulties in determining boundaries between words and syllables. An example of a word boundary error was found in the utterance *Uten <u>lakk svikter</u> lukten farlig* being misheard as *Uten <u>laks lukter</u> lukten farlig* (relevant words underlined). This example demonstrates that errors can be the result of complicated processes. One might speculate that the subject had problems with understanding all three sentence-medial words, perceived the initial fricative of *svikter* as belonging to *lakk* (resulting in *laks*) and retrieved the word *lukten* from memory to come to a phonotactically correct word for *vikter(?)=> lukter*.

An example of the insertion of an extra syllable due to mishearing was the disyllabic word *barna* in the utterance *Bilen vet at barna skulle hale*, which was rendered as *vinteren*. The opposite phenomenon, the subject missing a syllable, occurred in the utterance *Alltid lapper fela en dram* which was rendered as *Alltid lakker kjeler dram* (*en* missing, which has been counted as both a word and a syllable error).

Inspection of the data revealed that, contrary to our expectations, word boundary errors were rare in the data. As is shown by Table 10, the percentages for this type of errors varied between 1.5% and 4.6%.

**Table 10.** *Sentence test*. Occurrence of word/syllable boundary errors and inserted/missed syllables in percent, and number of stimuli.

| Listener group | Word | Syllable | Syllable | | Stimuli |
|---|---|---|---|---|---|
| | | | inserted | missed | n |
| REF<55 | 1.5 | 0.0 | 1.5 | 2.4 | 548 |
| UAH>55 | 4.6 | 0.0 | 2.1 | 4.5 | 863 |
| AH | 2.5 | 0.0 | 4.0 | 6.2 | 275 |

Also, no cases of errors at all in the perception of syllable boundaries were documented. This is true for all three groups of listeners, even the hearing-impaired. Finally, a certain number of inserted and missed syllables were observed. But also for this error category the numbers are relatively low.

#### 9.1.4.3 Phoneme level

In this section, first of all it should be pointed out that the sentence test was developed primarily for investigating the higher level effects discussed above (like speaker and listener dialect, gender, overall word error rates, etc.). The format of the test was not very well suited to study effects at the most detailed level, i.e. that of the phoneme. To illustrate the difficulties encountered in evaluating the results at this level, the following example is given. The utterance *Når mottar dere sjøer?* had been rendered by a subject as *Du har bodd ... sjøer*. In the evaluation the first three words have been counted as being misperceived (which certainly is correct). No attempt has been made, however, to further categorise the errors contained in the subject's orthographic rendition. The discrepancy between the original utterance and the sentence as perceived by the subject is too large to allow a specification of errors related to word/syllable boundaries, single consonants and vowels, etc. The data presented in Table 11 involve, therefore, more straight-forward cases like the perception of *frukta* as *flukta* (/r/-/l/-confusion) or *ball* misheard as *vann* (confusion of /b/-/v/ and /l/-/n/).

A second caveat with regard to the interpretation of the results presented in Table 11 is concerned with the distribution of error rates for the consonants and vowels. The ratio of numbers of consonants/vowels in the speech material is unknown, but certainly not 1:1. A count of the consonants/vowels contained in ten sentences from the speech material (one utterance from each block) resulted in averages of 12.1 consonant and 7.8 vowel phonemes per utterance, thus yielding a consonant/vowel ratio of 1.6 (or approximately 3:2). As a consequence, error rates could *a priori* not be expected to be the same for consonants and vowels but to show a bias for the former category.

As can be seen from Table 11, the ratios of the error rates for consonants vs. vowels were not dramatically higher than 1.6 (1.9, 2.1, and 1.8 for the respective listener groups). These results seem, therefore, to suggest that the listeners encountered difficulties in identifying both consonant and vowel phonemes correctly. Further, for the REF<55 group inserting or missing a phoneme was relatively rare. For the other two groups, the rates were approximately twice as high.

**Table 11.** *Sentence test*. Occurrence of phoneme (consonant/vowel) errors and inserted/missed phonemes in percent, and number of stimuli. The category consonant indicates the occurrence of one single consonant or a consonant cluster being misheard.

| Listener group | Consonant | Vowel | Phoneme | | Stimuli |
|---|---|---|---|---|---|
| | | | inserted | missed | n |
| **REF<55** | 19.2 | 10.0 | 5.7 | 5.3 | 548 |
| **UAH>55** | 33.7 | 15.8 | 10.2 | 10.7 | 863 |
| **AH** | 39.3 | 21.8 | 9.5 | 10.5 | 275 |

For reasons explained above, the results for phoneme identification should be taken with care. Both the *forced choice test* and the *word test allowed* a more detailed investigation analysis of phoneme identification. The results of these tests will be described below.

## 9.2 Word test

### 9.2.1 Preliminary remarks

It may be recalled that in this test the listeners' task was to "fill in" exactly one word in a (nearly complete) sentence presented on a screen. The goal of this design was twofold. Firstly, this test format was expected to be less demanding than that of the sentence test and, therefore, more suited for listeners with a hearing loss. Secondly, the more stringent format would allow us to extract more detailed information about sound confusions. In accordance with this rationale, the description of the experimental results will focus on the confusions of vowel and consonant phonemes observed for the three listener groups.

In the evaluation of the results, the character "\" (backslash) has been used as a template for empty slots. For example, for the two-phoneme stimulus word *sy* the three-slot transcription "s.\.y:" has been used to accomodate three-phoneme responses like *fri* (transcribed "f.r.i:"). In this way stimulus and response contained the same number of elements and possible categorisation problems in the semi-automatic production of confusion matrices were avoided. In the present example, the third element of the response (i:) will correctly be assigned to the corresponding element of the stimulus (y:), with the insertion of an r-sound. Similarly, the "\" character has been used in transcribing responses (e.g., in the response "l.O.\.t." to the stimulus "l.O.f.t."). Since the empty slot "\" does not give any specific information as to sound confusions, its occurrence was excluded from the interpretation. The "." (full stop ) was included in the one-character phonemes to give all phonemes the same number of symbols, i.e. two.

### 9.2.2 Vowel confusions

Confusion matrices for the vowels are presented in Table 12 to Table 14. First of all it can be seen that the overall error rates varied for the different listener groups, being lowest for REF<55 (3.6%), higher for UAH>55 (10.1%) and highest for AH (15.9%). In this respect, the present results thus fully confirm those from the *sentence test*.

In order to be able to quantify the information about vowel confusions given in the matrices the vowels were classified according to the following phonemic features:

- *place*: front - mid - back (for example /i:/ - /}:/ - / u:/)
- *opening*: close - close-mid - open-mid - open (for example /i:/ - /e:/ - /{:/ - /A:/)
- *rounding*: unrounded - rounded (for example /i:/ - /y:/)
- *quantity*: long - short (for example /i:/ - /i/)

Following this classification, the errors that had occurred in the test were specified and, finally, the occurrences of phonemic feature confusions were counted. Illustrating the procedure with an example, an /i:/ - /y:/ confusion was categorized as a *rounding* error, while an /i:/ - /y/ confusion involved both a *rounding* and a *quantity* error. A threefold error is observed in the case of an /i:/ - /A/ confusion (*place*, *opening* and *quantity*). In the following tables, error rates are expressed as percentages of the number of tokens involved for each phoneme (i.e., "n" in Table 12 to Table 14 and Table 17 to Table 19). The data emerging from this analysis are presented in Table 15. It is clear that due to the small number of errors that occurred for the group REF<55 no information has become available about vowel sounds and phonemic features that could play a role in speech intelligibility for that category of listeners. More informative, however, are the error rates for the groups with moderate to more serious hearing losses. Obviously, both the feature *rounding* (error rates of 5.1% and 7.5% for UAH>55 and AH, resp.) and *place* (error rates of 6.2% and 5.2%) contributed much to vowel confusions. The two other features *opening* and *quantity* were of less importance but still involved in vowel confusions (error rates of 1.8% and 1.1% for UAH>55 and 3.8% and 3.7% for AH, respectively).

**Table 12.** *Word test.* Vowel confusion matrix for listener group REF<55.

| Response | Stimulus | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \. | i: | i. | y: | y. | e: | e. | 2: | 2. | {: | {. | A: | A. | }: | }. | @. | u: | u. | O: | O. |
| \. | | | | | | | | | | | | | | | | | | | | |
| i: | | 18 | | | | | | | | | | | | | | | | | | |
| i. | | | 34 | | 1 | | | | | | | | | | | | | | | |
| y: | | | | 12 | | | | | | | | | | | | | | | | |
| y. | | | | | 5 | | | | | | | | | | | | | | | |
| e: | | | | | | 32 | | | | | | | | | | | | | | |
| e. | | | | | | | 24 | | | | | | | | | | | | | |
| 2: | | | | | | | | 11 | | | | | | | | | | | 3 | |
| 2. | | | | | | | | | 3 | | | | | | | | | | | |
| {: | | | | | | | | | | 9 | | | | | | | | | | |
| {. | | | | | | | | | | | 10 | | | | | | | | | |
| A: | | | | | | | | | | | | 35 | | | | | | | | |
| A. | | | | | | | | | | | | | 39 | | | | | | | |
| }: | | | | | | | | | | | | | | 7 | | | | | | |
| }. | | | | | | | | | | | | | | | 4 | | | | | |
| @. | | | | | | | | | | | | | | | | 178 | | | | |
| u: | | | | | | | | | | | | | | | | | 6 | | | |
| u. | | | | | | | | | | | | | | | | | | 3 | | |
| O: | | | | | | | | | | | | | | | | | | | 3 | |
| O. | | | | | | | | | | | | | 1 | | | | | | | 4 |
| {i | | | | | | | | | | | | | | | | | | | | |
| Ai | | | | | | | | | | | | | | | | | | | | |
| Oy | | | | | | | | | | | | | | | | | | | | |
| A} | | | | | | | | | | | | | | | | | | | | |
| }i | | | | | | | | | | | | | | | | | | | | |
| ui | | | | | | | | | | | | | | | | | | | | |
| 2y | | | | | | | | | | | | | | | | | | | | |
| **n** | | 18 | 34 | 12 | 6 | 32 | 24 | 11 | 3 | 9 | 10 | 35 | 40 | 7 | 4 | 178 | 6 | 3 | 6 | 4 |
| **% error** | | 0.0 | 0.0 | 0.0 | 16.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 50.0 | 0.0 |

**Table 13.** *Word test.* Vowel confusion matrix for listener group UAH>55.

| Response | Stimulus | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \. | i: | i. | y: | y. | e: | e. | 2: | 2. | {: | {. | A: | A. | }: | }. | @. | u: | u. | O: | O. |
| \. | | | 1 | | 1 | | | | | | | | | | | 7 | | | | |
| i: | | 67 | | 7 | 1 | | | | | | | | | 2 | 1 | | | | | |
| i. | | | 135 | | 8 | | | | | | | | | | | 1 | | | | |
| y: | | 4 | 1 | 26 | | | | | | | | | | 1 | | | 1 | | | |
| y. | | | | 2 | 37 | | | | | | | | | | 1 | | | | | |
| e: | | | | | | 159 | 1 | | | | | | | | | | | | | |
| e. | | | | | 1 | 1 | 75 | | 4 | | 1 | | | | | | | | | |
| 2: | | | | | | | | 26 | | | | | | | | | | | 5 | |
| 2. | | | | | | | 3 | | 31 | | | | | | 1 | | | | | |
| {: | | | | | | | | | | 32 | | | | | | | | | | |
| {. | | | | | | | | | | | 27 | | | | | | | | | |
| A: | | | | | | | | | | 1 | | 166 | | | | | | | | |
| A. | 1 | | | | | | | | | | 1 | | 161 | | | 4 | | | | 1 |
| }: | | | | 1 | | | | | | | | | | 19 | 2 | 1 | | | | |
| }. | | | 2 | | 1 | 1 | | | | | | | | | 28 | | | | | |
| @. | 2 | | 1 | | | | | | | | | | | | | 727 | | | | |
| u: | | | 1 | | | | | | | | | | | | | | 31 | | | |
| u. | | | | | | | | | 1 | | | | | | | | | 6 | | |
| O: | | | | | | | | | | | | | | | | | | | 6 | |
| O. | | | | | | | 2 | | | | 1 | | 1 | | | | | | | 9 |
| {i | | | | | 1 | | 1 | | | | | | | | | | | | | |
| Ai | | | | | | | | | | | | | | | | | | | | |
| Oy | | | | | | | | | | | | | | | | | | | | |
| A} | | | | | | | | | | | | | | | | | | | | |
| }i | | | | | | | | | | | | | | | | | | | | |
| ui | | | | | | | | | | | | | | | | | | | | |
| 2y | | | | | | | 1 | | | | | | | | | | | | | |
| **n** | | 71 | 141 | 36 | 50 | 161 | 83 | 26 | 36 | 33 | 30 | 166 | 162 | 22 | 33 | 740 | 32 | 6 | 11 | 10 |
| **% error** | | 5.6 | 4.3 | 27.8 | 26.0 | 1.2 | 9.6 | 0.0 | 13.9 | 3.0 | 10.0 | 0.0 | 0.6 | 13.6 | 15.2 | 1.8 | 3.1 | 0.0 | 45.5 | 10.0 |

**Table 14.** *Word test.* Vowel confusion matrix for listener group AH.

| Response | Stimulus | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \. | i: | i. | y: | y. | e: | e. | 2: | 2. | {: | {. | A: | A. | }: | }. | @. | u: | u. | O: | O. |
| \. | | | | | | | | | | | | | 1 | | | 13 | | | | |
| i: | | 77 | 15 | 11 | 6 | | | | | | | | | | | | | | | |
| i. | 2 | 1 | 126 | | 18 | | | | | | | | | | | | | | | |
| y: | | 9 | 4 | 31 | 1 | | | | | | | | | 4 | | | | | | |
| y. | | | 8 | 2 | 31 | | | | | | 2 | | | | 4 | | | | | |
| e: | | | 1 | | | 168 | 3 | | | 2 | | | | | | | | | 1 | |
| e. | 1 | | | | | 2 | 90 | | 9 | 1 | 1 | | | | 1 | | | | | |
| 2: | | | | | | 2 | | 51 | 1 | | | 1 | | | | | | | 5 | |
| 2. | | | | | 1 | | 1 | | 29 | | | | | | 1 | | | | | |
| {: | | | | 1 | | 2 | | 1 | | 31 | 1 | | | | | | | | | |
| {. | | | | | | | 1 | | | 1 | 30 | 1 | 1 | | | | | | | |
| A: | | | | | | | | | | | | 189 | 5 | | | | | | 1 | |
| A. | | | 1 | | | | 1 | | | | 8 | 2 | 207 | | | 5 | | | | |
| }: | | | 1 | 1 | 2 | | | | | | | | | 20 | | | | | | |
| }. | | | 5 | | | | 1 | | 1 | | | | | 1 | 38 | | | 1 | | |
| @. | 3 | 1 | | | | | | | | | | | 1 | | | 895 | | | | |
| u: | | | | | | | | | | | | | | 1 | | | 43 | | | 1 |
| u. | 1 | | | | | | | | | | | | 1 | | | | 3 | 8 | | |
| O: | | | | | | | | | | | | | | | | | | | 12 | |
| O. | | | 1 | | | | | | | | | | 1 | | | | | | | 17 |
| {i | | | | | | 1 | | | | | | | | | | | | | | |
| Ai | | | | | | | | | | | | | | | | | | | | |
| Oy | | | | | | | | | | | | | | | | | | | | |
| A} | | | | | | | | | | | | | | | | | | | | |
| }i | | | | | | | | | | | | | | | | | | | | |
| tb | | | | | | | | | | | | | | | | | | | | |
| 2y | | | | | | | 1 | | | | | | | | | | | | | |
| **n** | | 88 | 162 | 46 | 59 | 175 | 98 | 52 | 40 | 35 | 42 | 193 | 217 | 26 | 44 | 913 | 46 | 9 | 19 | 18 |
| **% error** | | 12.5 | 22.2 | 32.6 | 47.5 | 4.0 | 8.2 | 1.9 | 27.5 | 11.4 | 28.6 | 2.1 | 4.6 | 23.1 | 13.6 | 2.0 | 6.5 | 11.1 | 36.8 | 5.6 |

**Table 15.** *Word test.* Error rates for phonemic features involved in vowel confusions and number of stimuli. *Place* refers to the front-back dimension.

| Listener group | Rounding | Opening | Place | Quantity | Stimuli (n) |
|---|---|---|---|---|---|
| **REF<55** | 1.0 | 0.1 | 2.6 | 0.0 | 250 |
| **UAH>55** | 5.1 | 1.8 | 5.2 | 1.1 | 1054 |
| **AH** | 7.5 | 3.7 | 6.2 | 3.8 | 1298 |

As a next step in the analysis it was investigated whether the data showed specific tendencies with regard to the direction of the vowel confusions. For example, we have already seen that the feature *rounding* ranked highest. That is to say that rounded vowels rather often have been perceived as being unrounded and vice versa. This does not tell us, however, whether one of the two confusions (*rounded => unrounded* or *unrounded => rounded*) predominated. To be able to answer this question the data were treated in the following way. For each of the vowel phonemes, the occurrences of feature confusions were counted separately for each of the possible directions (i.e., two in the case of *rounding*) and expressed as percent of the total number of vowel tokens for that specific vowel phoneme. As a final step, the percentages were averaged across the vowel phonemes with the relevant feature (in the case of the feature *rounded* the ten vowels /y:  y/, /}:  }/, /u:  u/, /2:  2/, and /O:  O/).

The results of these calculations are presented in Table 16. As to the perceptually most important feature *rounding*, the results for all three groups demonstrate a clear predominance of *rounded => unrounded* confusions. Of course, because of the generally higher error rates for the two listener groups UAH>55 and AH, their data are the most conclusive (7.3% vs. 1.7% and 10.4% vs. 3.9%, respectively).

As to the confusions related to vowel *opening*, there is no clear general tendency (Table 16b). Higher values were found for *open mid* vowels being perceived as either more closed or more open. Note, however, that these data are based on only two vowels (i.e., the only two *mid-open* vowels in Norwegian, (/{:/ and /{/).

The data for the place feature (Table 16c) show that most errors occurred for *mid* and *back* vowels. Like above with the *open-mid* vowels, it should be kept in mind that only two *mid* vowel phonemes are involved, viz. /}:/ and /}/.

Finally, more *short => long* confusions rather than the other way round were found for group AH (4.9% vs. 2.7%). Inspection of Table 14 reveals that it was mainly the *short* /i/ being perceived as *long* /i:/ that contributed to this effect (22 out of 41 *short => long* confusions).

### 9.2.3 Consonant confusions

The results for the consonant confusions are presented in Table 17 to Table 19. The mean overall error rates appeared to be somewhat higher than those for the vowels, but showed otherwise the same differences between the three groups of listeners (REF<55: 4.3%,  UAH>55: 12.3% and AH: 20.3%). Table 20 gives the error rates broken down by phoneme category (or manner of articulation: *plosive*, *fricative*, *nasal*, and *liquid*). As appears from these data, errors varied substantially as a function of manner of articulation. Consistently for all three listener groups, the nasals were strongest prone to confusions. In descending order, the overall error rates for the nasals are followed by those for the plosives, the fricatives and the liquids. The only exception from this picture are the numbers for the REF<55 group's for the fricative (1.7%) and liquid (1.8%) categories. It should be noted, however, that both the error rates and the numbers of observations for this listener group (fricatives: 244; liquids: 258) were relatively small. Therefore, more importance should be attached to the error rate hierarchy fricatives > liquids as found for the groups UAH>55 and AH (fricatives vs. liquids: 9.6% vs. 6.1%, n =1124 and 981; 17.1% vs. 13.2%, n= 1336 and 1289, respectively).

**Table 16.** *Word test.* Percentage of phonemic feature confusions for the vowels (relative to the number of tokens per vowel category) averaged across the number of vowel phonemes involved. Number of vowel phonemes in parentheses.

| **(a) Rounding** | rounded => unrounded | | unrounded => rounded | |
|---|---|---|---|---|
| REF<55 | 1.7 | | 0.3 | |
| UAH>55 | 7.3 | (10) | 2.4 | (8) |
| AH | 10.4 | | 3.9 | |

| **(b) Opening** | close => other | | close-mid => other | | open-mid => other | | open => other | |
|---|---|---|---|---|---|---|---|---|
| REF<55 | 0.0 | | 0.0 | | 0.0 | | 1.3 | |
| UAH>55 | 0.6 | (8) | 2.2 | (6) | 6.5 | (2) | 0.3 | (2) |
| AH | 1.3 | | 3.2 | | 17.4 | | 1.2 | |

| **(c) Place** | front => other | | mid => other | | back => other | |
|---|---|---|---|---|---|---|
| REF<55 | 0.0 | | 0.0 | | 8.3 | |
| UAH>55 | 2.2 | (10) | 11.4 | (2) | 8.1 | (6) |
| AH | 3.4 | | 16.4 | | 7.4 | |

| **(d) Quantity** | long => short | | short => long | |
|---|---|---|---|---|
| REF<55 | 0.0 | | 0.0 | |
| UAH>55 | 0.8 | (9) | 1.5 | (9) |
| AH | 2.7 | | 4.9 | |

Analogous to the procedure followed for the vowels (see Table 15, p.25), the next step in the analysis was to look more closely at the specific features that played a role in the consonant confusions. To that aim the consonants were classified according to the following three features:

- *manner*: plosive - fricative - nasal - lateral (for example /p/ - /f/ - /m/ - /l/)
- *place*: bilabial - labiodental - alveolar - postalveolar - retroflex - palatal - velar - glottal (for example /p/ - /f/ - /t/ - /S/ - /rt/ - /C/ - /k/ - /h/)
- *voicing*: unvoiced - voiced (for example /p/ - /b/)

**Table 17**. *Word test.* Consonant confusion matrix for listener group REF<55.

| Response | Stimulus | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \. | p. | b. | t. | d. | k. | g. | m. | n. | =n | J. | rt | rd | rn | N. | r. | R. | f. | v. | s. | S. | C. | j. | h. | l. |
| \. | 1 | | | | | | 1 | | | | | | | | | | | | | | | | | | 1 |
| p. | | 24 | | 1 | | | | | | | | | | | | | | | | | | | | 1 | |
| b. | | | 18 | | | | | | | | | | | | | | | | | | | | | | |
| t. | 1 | | | 49 | | 1 | | | | | | | | | | | | | | | | | | | |
| d. | | | | | 19 | | | | | | | | | | | | | | | | | | | | |
| k. | | 2 | | | | 37 | 1 | | | | | | | | | | | | | | | | | | |
| g. | | | | | | | 24 | | | | | | | | | | | | | | | | | | |
| m. | | | | | | | | 9 | | | | | | | | | | | | | | | | | |
| n. | | | | | | | | | 33 | | | | | | 1 | | 1 | | | | | | | | 2 |
| =n | | | | | | | | | | 13 | | | | | | | | | | | | | | | |
| J. | | | | | | | | | | | 4 | | | | | | | | | | | | | | |
| rt | | | | | | | | | | | | | | | | | | | | | | | | | |
| rd | | | | | | | | | | | | | | | | | | | | | | | | | |
| rn | | | | | | | | | 1 | | | | | 6 | | | | | | | | | | | |
| N. | | | | | | | | 1 | | | 1 | | | | 12 | | | | | | | | | | |
| r. | | | | | | | | | | | | | | | | 174 | | | | | | | | | |
| R. | | | | | | | | | | | | | | | | | 62 | | | | | | | | |
| f. | | | | | | | | | | | | | | | | | | 18 | | | | | | | |
| v. | | | | | | | 1 | | | | | | | | | | 2 | | 30 | | | | | | |
| s. | | | | | | | | | | | | | | | | | | | | 70 | | | | | |
| S. | | | | | | | | | | | | | | | | | | | | | 22 | | | | |
| C. | | | | | | | | | | | | | | | | | | | | | | 19 | | | |
| j. | | | | | | | | | | | | | | | | | | | | | | | 2 | | |
| h. | | 2 | | | | 2 | | | | | | | | | | | | | | | | | | 15 | |
| l. | | | | | | | | | 1 | | | | | | | | 2 | | | | | | | | 81 |
| **n** | | 28 | 18 | 50 | 19 | 40 | 27 | 10 | 35 | 13 | 5 | 0 | 0 | 6 | 13 | 174 | 67 | 18 | 30 | 70 | 22 | 19 | 2 | 16 | 84 |
| **% error** | | 14.3 | 0.0 | 2.0 | 0.0 | 7.5 | 11.1 | 10.0 | 5.7 | 0.0 | 20.0 | - | - | 0.0 | 7.7 | 0.0 | 7.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 6.3 | 3.6 |

**Table 18.** *Word test.* Consonant confusion matrix for listener group UAH>55

| Response | Stimulus | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \. | p. | b. | t. | d. | k. | g. | m. | n. | =n | J. | rt | rd | rn | N. | r. | R. | f. | v. | s. | S. | C. | j. | h. | l. |
| \. | | 4 | 3 | 6 | | 7 | 5 | | 3 | 1 | | | | 3 | 2 | 17 | 16 | 3 | 7 | 10 | | | | 3 | 4 |
| p. | 1 | 79 | | 5 | 1 | 1 | | | | | | | | | | | | 1 | | | 1 | | | 2 | |
| b. | 1 | | 65 | | 1 | | | | | | | | | | | | | | 2 | | | | | 1 | 1 |
| t. | 3 | 5 | 1 | 199 | 1 | 7 | | | | | | | | | | 1 | 1 | 1 | | 7 | | 1 | | | 1 |
| d. | 1 | | 3 | 1 | 65 | | 5 | | 1 | | | | | | | 1 | | | 1 | | | | | | |
| k. | 4 | 5 | | 5 | | 156 | | | | | | | | | | | | | | | 1 | | | | |
| g. | 2 | | | | 2 | | 103 | | | | | | | | 2 | | 1 | | 1 | | | | | | 1 |
| m. | 1 | | 1 | | | 1 | 1 | 26 | 1 | | 2 | | | | 3 | | 1 | | 3 | | 1 | | | | |
| n. | 5 | | | | | | 1 | 3 | 92 | | 1 | | | 1 | 12 | 3 | | | | 1 | | | | | 2 |
| =n | | | | | | | | | | 60 | | | | | | | | | | | | | | | |
| J. | | | | | | | | | | | 15 | | | | | | | | | | | | | | |
| rt | | | | 1 | | | | | | | | 6 | | | | | | | | | | | | | |
| rd | | | | | | | | | | | | | | | | | | | | | | | | | |
| rn | | | | | | | | | | | | | | 25 | | | | | | | | | | | |
| N. | | | | | | 1 | 1 | 2 | 1 | | 2 | | | | 38 | 1 | | | | | | | | | |
| r. | 8 | | | | 2 | | | | 1 | | | | | 1 | | 592 | | | 1 | 1 | | | | | 9 |
| R. | 1 | | | | | | 1 | | | | | | | | | | 260 | | 1 | 2 | | | | | 1 |
| f. | | | | 3 | | | | | | | | | | | | | 2 | 84 | | 6 | 3 | 2 | | 1 | |
| v. | 2 | | 1 | 1 | 2 | | | | | | | | | | | 3 | 7 | 1 | 109 | | | | | 1 | 3 |
| s. | 6 | 1 | | 2 | | | | | | | | | | | | 2 | 1 | 3 | | 282 | | 1 | | 1 | |
| S. | | | | | | | | | | | | | | | | | | 2 | | 7 | 94 | 5 | | | |
| C. | | | | | | 2 | | | | | | | | | | | | | | 1 | 1 | 80 | | | |
| j. | | | | | | | | | | | | | | | | | | | | | | | 9 | | |
| h. | | | | 2 | | 3 | | 1 | 1 | | | | | | | | 2 | | | 1 | | 1 | | 82 | |
| l. | 6 | | | | 3 | | 3 | 1 | 6 | | | | | | 1 | 9 | 1 | | 3 | | | | | | 330 |
| **n** | | 94 | 74 | 225 | 77 | 178 | 120 | 33 | 106 | 61 | 20 | 6 | 0 | 30 | 58 | 629 | 292 | 95 | 128 | 318 | 101 | 90 | 9 | 91 | 352 |
| **% error** | | 16.0 | 12.2 | 11.6 | 15.6 | 12.4 | 14.2 | 21.2 | 13.2 | 1.6 | 25.0 | 0.0 | - | 16.7 | 34.5 | 5.9 | 11.0 | 11.6 | 14.8 | 11.3 | 6.9 | 11.1 | 0.0 | 9.9 | 6.3 |

**Table 19.** *Word test.* Consonant confusion matrix for listener group AH.

| Response | Stimulus | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \. | p. | b. | t. | d. | k. | g. | m. | n. | =n | J. | rt | rd | rn | N. | r. | R. | f. | v. | s. | S. | C. | j. | h. | l. |
| \. | 7 | 6 | 8 | 9 | | 13 | 8 | 1 | 4 | 3 | | | | | 1 | 31 | 25 | 5 | 9 | 30 | | 2 | | 3 | 14 |
| p. | 2 | 80 | | 8 | 2 | 2 | | | | | | | | | | 1 | 1 | 1 | 2 | 3 | | | | | |
| b. | 6 | 1 | 74 | | 3 | 1 | 3 | 1 | | | | | | | | | 1 | | | 2 | | | | 1 | |
| t. | 14 | 19 | | 212 | 1 | 9 | 1 | | 1 | | 1 | | | | | 1 | | 2 | 1 | 7 | | 1 | | 1 | 4 |
| d. | 1 | | 4 | 4 | 57 | 1 | 4 | | 3 | | | | | | | 2 | | 2 | | | | 1 | | 1 | 3 |
| k. | 5 | 7 | | 14 | 2 | 172 | 4 | | 1 | | | | | | 2 | 1 | | 1 | 1 | 6 | | | | 2 | 1 |
| g. | 1 | | 5 | | 5 | 1 | 126 | | | | | | | | | 1 | | | 5 | | | | | 1 | 2 |
| m. | 3 | | | | | 1 | 1 | 26 | 3 | | 4 | | | | 4 | | 3 | | 2 | | | | | 1 | 3 |
| n. | 9 | | | 1 | 1 | | | 4 | 121 | | 1 | | | 2 | 9 | 5 | | 2 | 2 | 2 | | | | | 11 |
| =n | | | | | | | | | | 65 | | | | | | 1 | | | | | | | | | |
| J. | | | | | | | | | | | 11 | | | | | | | | | | | | | | |
| rt | | | | 1 | | | | | | | | 8 | | | | | | | | 1 | | | | | |
| rd | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| rn | | | | | | | | | | 1 | | | | 25 | | 1 | | 1 | | | | | | | 2 |
| N. | 1 | | | | | | 3 | 5 | 1 | | 9 | | | | 45 | | 1 | | 1 | | | | | | |
| r. | 16 | 2 | | | 4 | 1 | 7 | 1 | 1 | | | | | 1 | | 753 | | | 4 | 1 | | | | | 20 |
| R. | 4 | | | | | 1 | 7 | | | | | | | | | | 282 | | 2 | 1 | | | | 2 | 6 |
| f. | 2 | | 1 | 10 | 2 | 2 | | | | | | | | | | | 2 | 87 | 2 | 10 | 1 | 4 | | 3 | |
| v. | 14 | | 2 | 1 | 4 | | 1 | | 2 | | | | | | | 2 | 7 | | 115 | | | | | 6 | 7 |
| s. | 8 | | | 2 | 1 | 2 | 1 | | | | | | | | | 2 | | 6 | 1 | 331 | | 2 | | | 3 |
| S. | 1 | | | | | | | | | | | | | | | | | 4 | | 8 | 115 | 13 | | 1 | |
| C. | | | | | | 1 | | | | | | | | | | | | 1 | | 3 | 3 | 72 | | 1 | |
| j. | 1 | | | | | | 1 | | 3 | | | | | | | 4 | | | | 1 | | 1 | 8 | | 1 |
| h. | | 3 | | | | 2 | | | 1 | | | | | | | 1 | | 4 | 1 | | | 1 | | 79 | 1 |
| l. | 5 | | | 1 | 4 | | | | 6 | | | | | | 4 | 24 | 10 | | 5 | 1 | | | | 2 | 380 |
| **n** | | 118 | 94 | 263 | 86 | 209 | 167 | 38 | 147 | 69 | 26 | 8 | 0 | 28 | 65 | 830 | 332 | 116 | 153 | 407 | 119 | 97 | 8 | 104 | 459 |
| **% error** | | 32.2 | 21.3 | 19.4 | 33.7 | 17.7 | 24.6 | 31.6 | 17.7 | 5.8 | 57.7 | 0.0 | - | 10.7 | 30.8 | 9.3 | 15.1 | 25.0 | 24.8 | 18.7 | 3.4 | 25.8 | 0.0 | 24.0 | 17.2 |

**Table 20.** *Word test.* Mean error rates in consonant perception plus number of phoneme categories and number of stimuli.

| Listener group | Category | Error rate | Phonemes (n) | Stimuli (n) |
|---|---|---|---|---|
| | plosive | 5.8 | 6 | 182 |
| **REF<55** | fricative | 1.7 | 8 | 244 |
| | nasal | 7.2 | 6 | 82 |
| | liquid | 1.8 | 2 | 258 |
| | plosive | 11.7 | 7 | 774 |
| **UAH>55** | fricative | 9.6 | 8 | 1124 |
| | nasal | 19.2 | 6 | 308 |
| | liquid | 6.1 | 2 | 981 |
| | plosive | 21.3 | 7 | 945 |
| AH | fricative | 17.1 | 8 | 1336 |
| | nasal | 25.7 | 6 | 373 |
| | liquid | 13.2 | 2 | 1289 |

As for the vowels, the number of features involved in consonant confusions were counted. For example, a /p/ - /b/ confusion was being counted as one *voicing* error, while a /p/ - /v/ confusion was registered as a threefold error (*manner*, *place,* and *voicing*). *Voicing* errors were only specified for speech sounds that are phonologically [± voiced]. In consequence, a confusion of a phonologically voiceless /t/ with an (only phonetically, not phonologically voiced) nasal /n/ did not result in a *voice* error but was only counted as a *manner* error. However, the reverse situation was treated differently: The confusion of /n/ with a /t/ was considered to involve both a *manner* and a *voice* error, the reason for this being that /n/ usually is phonetically voiced and the two options for the listener with this *manner* confusion are voiceless /t/ and voiced /d/. The following sounds were classified as unspecified with regard to the feature [± voiced]: the nasals /n/, /N/, /rn/, /J/ and /m/, the liquids /l/ and /r/, the fricatives /R/, /s/, /S/, and /h/.

The results of the analysis are shown in Table 21.

**Table 21.** *Word test.* Error rates for phonemic features involved in consonant confusions and number of stimuli.

| Listener group | Manner | Place | Voicing | Stimuli (n) |
|---|---|---|---|---|
| **REF<55** | 1.3 | 3.4 | 0.2 | 250 |
| **UAH>55** | 3.5 | 8.2 | 0.3 | 1054 |
| **AH** | 6.4 | 14.6 | 1.6 | 1298 |

Clearly, the *place* feature was most often involved in consonant confusions. This is true for all three listener groups, with increasing absolute numbers for REF<55 - UAH>55 - AH. Especially for the latter group the correct identification of consonant *place* has been problematic: The corresponding error rate amounted to 14.6%. Expressed in absolute numbers, with 1298 word identifications a total of 392 errors occurred. If we for the sake of simplicity assume that not more than only one consonant per word was involved, this means that on average almost one third (30.2%) of the words have been identified incorrectly.

Lower but still sizeable error rates were found for consonant *manner*, with scores of 3.5% (UAH>55) and 6.4% (AH), respectively. Apart from the AH group having 1.6% *voicing* errors, this latter feature did not seem to play any important role in consonant confusions.

Further, the data from the confusions matrices (Table 17 to Table 19) show, e.g., that *place* was a very important feature in the perception of nasals, whereas the liquids /l/ and /r/ showed predominantly *manner* confusions (*lateral* /l/ being perceived as *vibrant* /r/ and vice versa). In order to quantify these data the errors occurring in consonant perception were further analyzed by investigating the contributions of features specified for the four consonant types plosives, fricatives, nasals, and liquids. In order to get an easy to grasp overview it was felt appropriate to refrain from specifying errors for the eight *place* features (from *bilabial* to *glottal*) or for *voicing*. Similar to the procedure chosen for the vowels, the errors were expressed as percent of the total number of consonant tokens for each consonant phoneme and averages across the phonemes were calculated.

The results shown in Table 22 indicate that *place* confusions, which contributed most to misperceptions (cf. Table 21), first of all were found for the nasals. For all three listener groups, the corresponding error percentages are by far the highest - not only within the *place* category, but also compared with the *manner* and *voicing* categories (REF<55: 5.5%; UAH>55: 13.9%; AH: 21.6%). Not only for the nasals, but also for the plosives and fricatives the *place* feature contributed more to consonant confusions than the features *manner* and *voicing*. The only exception to the general high contribution of the *place* feature was found for the liquids /l/ and /r/, where for listeners from groups UAH>55 and AH higher percentages were found than for *place* and *voicing*. This outcome can to a large degree be explained by the reciprocal confusion of /l/ and /r/, reflecting their auditory similarity. Finally, the (rather low) error percentages for *voicing* confusions showed no specific pattern and appeared to be of marginal importance.

**Table 22.** *Word test.* Percentage of phonemic feature confusions for the consonants (relative to the number of tokens per consonant category) averaged across the number of consonant phonemes involved. n= number of consonant phonemes.

| Listener group | Category | Manner | Place | Voicing | n |
|---|---|---|---|---|---|
| **REF<55** | plosive | 2.6 | 4.6 | 0.6 | 6 |
| | fricative | 1.3 | 1.7 | 0.0 | 8 |
| | nasal | 0.5 | 5.5 | 0.0 | 6 |
| | liquid | 0.0 | 0.0 | 0.0 | 2 |
| **UAH>55** | plosive | 3.7 | 7.1 | 0.7 | 7 |
| | fricative | 2.8 | 6.5 | 0.1 | 8 |
| | nasal | 3.8 | 13.9 | 0.0 | 6 |
| | liquid | 4.1 | 1.2 | 0.1 | 2 |
| **AH** | plosive | 7.0 | 13.9 | 2.6 | 7 |
| | fricative | 5.7 | 12.7 | 1.1 | 8 |
| | nasal | 5.5 | 21.6 | 1.4 | 6 |
| | liquid | 9.9 | 3.5 | 1.0 | 2 |

## 9.3 Forced choice test

### 9.3.1 Preliminary remarks

In the *forced choice test* the listeners' task was to choose one of four alternative words presented together with the test sentence on a computer screen. As a consequence of this special test format the evaluation has been somewhat different from that for the *sentence test* and the *word test*. Similar to the evaluation procedure followed for the *word test*, comparisons of stimulus words and responses were used to construct confusion matrices. For example, a stimulus word *haver* identified as *paver* (chosen from the four alternatives *gaver*, *haver*, *maver*, *paver*) was counted as a /h/ - /p/ confusion. The difference with the evaluation of the *word test* lies in the calculation of error percentages. In the *word test* all the phonemes of a stimulus word were potentially confusable: *haver* could not only be perceived as *paver* but also, e.g., as *haler*. Therefore, in calculating the total number of potential errors all five phonemes of the stimulus word were entries. In contrast, in the forced choice example given above only the word-initial phoneme was confusable. In order to be able to express the confusions as a relative measure the following steps were taken. For each confusable phoneme, the number of potential confusions was determined. There were, e.g., three potential confusions of /h/ in the case of *gaver*, *haver*, *maver*, *paver*. In the case of the stiumlus word *rager* with the alternatives *rager*, *sager*, *raver*, *raker* there were two possible /g/ confusions (with /v/ or /k/) and one possible /r/ confusion (with /s/). Having established the number of potential confusions for each stimulus word (i.e., each sentence), for each group of listeners the total number of occurrences of each sentence was counted. (Note that the order as well as the number of sentences presented in the *forced choice test* condition varied from listener to listener.) Then, the number of possible phoneme errors was multiplied with this number of occurrences. For example, sentence SH139 containing the *rager*, *sager*, *raver*, *raker* alternatives was presented in total ten times for the AH listener group.This gave a total of 10x2= 20 possible /g/ confusions and 10 possible /r/ confusions. Finally, for each of the confusable phonemes in the *forced choice test* the total of possible phoneme errors was calculated for each of the three listener groups.

### 9.3.2 Vowel confusions

Confusion matrices for vowel perception in the *forced choice test* are presented in Table 23 to Table 25. As can be seen from the tables the overall error rates are relatively low: 0.0% for group REF<55, 4.9% for group UAH<55 and 6.8% for group UAH>55. This is certainly due to test-inherent constraints, viz. having only a selected number of possible phoneme confusions compared to the large number of potential confusions in the *word test*. The general increasing tendency of the error rates found for the three different listener groups, however, is in congruence with the tendencies emerging from both the *sentence test* and the *word test*.

Following the same procedure as for the *word test*, it was investigated to what degree different phonemic features contributed to vowel confusions. The data presented in Table 26 can be compared directly to the corresponding results from the *word test* given in Table 15, p.25. Though of a relative limited value due to the small number of confusions, also these data suggest that at least for the AH group the *rounding* feature contributed most to vowel confusions (error rate = 3.8%) The results for the features *opening, rounding* and *quantity* show tendencies that are different from those for the *word test* but they are not very conclusive.

**Table 23.** F*orced choice test.* Vowel confusion matrix for listener group REF<55. Sum= sum of possible errors (indicated along the diagonal) and actual confusions.

| | **Stimulus** | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Response** | \. | i: | i. | y: | y. | e: | e. | 2: | 2. | {: | {. | A: | A. | }: | }. | @. | u: | u. | O: | O. |
| \. | | | | | | | | | | | | | | | | | | | | |
| i: | | 13 | | | | | | | | | | | | | | | | | | |
| i. | | | 1 | | | | | | | | | | | | | | | | | |
| y: | | | | 12 | | | | | | | | | | | | | | | | |
| y. | | | | | 15 | | | | | | | | | | | | | | | |
| e: | | | | | | 12 | | | | | | | | | | | | | | |
| e. | | | | | | | 16 | | | | | | | | | | | | | |
| 2: | | | | | | | | 10 | | | | | | | | | | | | |
| 2. | | | | | | | | | 6 | | | | | | | | | | | |
| {: | | | | | | | | | | 0 | | | | | | | | | | |
| {. | | | | | | | | | | | 0 | | | | | | | | | |
| A: | | | | | | | | | | | | 18 | | | | | | | | |
| A. | | | | | | | | | | | | | 19 | | | | | | | |
| }: | | | | | | | | | | | | | | 0 | | | | | | |
| }. | | | | | | | | | | | | | | | 12 | | | | | |
| @. | | | | | | | | | | | | | | | | 0 | | | | |
| u: | | | | | | | | | | | | | | | | | 5 | | | |
| u. | | | | | | | | | | | | | | | | | | 8 | | |
| O: | | | | | | | | | | | | | | | | | | | 7 | |
| O. | | | | | | | | | | | | | | | | | | | | 0 |
| **sum** | | 13 | 1 | 12 | 15 | 12 | 16 | 10 | 6 | 0 | 0 | 18 | 19 | 0 | 12 | 0 | 5 | 8 | 7 | 0 |
| **% error** | | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | - | - | 0.0 | 0.0 | - | 0.0 | - | 0.0 | 0.0 | 0.0 | - |

**Table 24.** *Forced choice test.* Vowel confusion matrix for listener group UAH>55. Sum= sum of possible errors (indicated along the diagonal) and actual confusions.

| | **Stimulus** | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Response** | \. | i: | i. | y: | y. | e: | e. | 2: | 2. | {: | {. | A: | A. | }: | }. | @. | u: | u. | O: | O. |
| \. | | | | | | | | | | | | | | | | | | | | |
| i: | | 68 | | 2 | | | | | | | | | | | | | | | | |
| i. | | | 4 | | | | | | | | | | | | | | | | | |
| y: | | 2 | | 63 | | | | | | | | | | | | | | | | |
| y. | | | | | 33 | | | | | | | | | | 1 | | | 1 | | |
| e: | | | | 1 | | 54 | | 1 | | | | | | | | | | | | |
| e. | | | | | | 1 | 34 | | 6 | | | | | | | | | | | |
| 2: | | | | | | | | 16 | | | | | | | | | | | | |
| 2. | | | | | | | | | 21 | | | | | | | | | | | |
| {: | | | | | | | | | | 0 | | | | | | | | | | |
| {. | | | | | | | | | | | 0 | | | | | | | | | |
| A: | | | | | | | | | | | | 80 | | | | | | | | |
| A. | | | | | | | | | | | | | 62 | | | | | | | |
| }: | | 1 | | | | | | | | | | | | 0 | | | | | | |
| }. | | | | | 1 | | | | | | | | | | 21 | | | 1 | | |
| @. | | | | | | | | | | | | | | | | 0 | | | | |
| u: | | | | | | | | | | | | | | | | | 29 | | 2 | |
| u. | | | | | | | | | | | | | | | | | 1 | 35 | | |
| O: | | | | | | | | | | | | | | | | | | | 33 | |
| O. | | | | | | | | | | | | | | | | | | | | 0 |
| **sum** | | 71 | 4 | 66 | 34 | 55 | 34 | 17 | 27 | 0 | 0 | 80 | 62 | 0 | 22 | 0 | 30 | 37 | 35 | 0 |
| **% error** | | 4.4 | 0.0 | 4.8 | 3.0 | 1.9 | 0.0 | 6.3 | 28.6 | - | - | 0.0 | 0.0 | - | 4.8 | - | 3.4 | 5.7 | 6.1 | - |

Table 25. *Forced choice test.* Vowel confusion matrix for listener group AH. Sum = sum of possible errors (indicated along the diagonal) and actual confusions.

| | **Stimulus** | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Response** | \. | i: | i. | y: | y. | e: | e. | 2: | 2. | {: | {. | A: | A. | }: | }. | @. | u: | u. | O: | O. |
| \. | | | | | | | | | | | | | | | | | | | | |
| i: | | 111 | 1 | 2 | | | | | | | | | | | | | 1 | | | |
| i. | | | 10 | | | | | | | | | | | | | | | | | |
| y: | | 3 | | 81 | | | | | | | | | | | | | | | | |
| y. | | | | | 29 | | | | | | | | | | 2 | | | | | |
| e: | | 1 | | 1 | | 76 | | | | | | | | | | | 1 | | | |
| e. | | | | | | | 68 | | 8 | | | | 2 | | 1 | | | | | |
| 2: | | | | | | | | 21 | | | | | | | | | | | | |
| 2. | | | | | | | | | 27 | | | | | | | | | | | |
| {: | | | | | | | | | | 0 | | | | | | | | | | |
| {. | | | | | | | | | | | 0 | | | | | | | | | |
| A: | | | | | | | | | | | | 96 | 3 | | | | | | | |
| A. | | | | | | | | | | | | | 88 | | | | | | | |
| }: | | 2 | | | | | | | | | | | | 0 | | | | | | |
| }. | | | | | | | 3 | | | | | | | | 24 | | | | | |
| @. | | | | | | | | | | | | | | | | 0 | | | | |
| u: | | | | | | | | | | | | | | | | | 36 | | | |
| u. | | | | | | | | | | | | | | | | | 6 | 58 | | |
| O: | | | | | | | | | | | | | | | | | | | 42 | |
| O. | | | | | | | | | | | | 1 | | | | | | | | 0 |
| **sum** | | 117 | 11 | 84 | 29 | 76 | 71 | 21 | 35 | 0 | 0 | 97 | 93 | 0 | 27 | 0 | 44 | 58 | 42 | 0 |
| **% error** | | 5.4 | 10.0 | 3.7 | 0.0 | 0.0 | 4.4 | 0.0 | 29.6 | - | - | 1.0 | 5.7 | - | 12.5 | - | 22.2 | 0.0 | 0.0 | - |

**Table 26.** *Forced choice test.* Error rates for phonemic features involved in vowel confusions and number of stimuli. *Place* refers to the front-back dimension.

| Listener group | Rounding | Opening | Place | Quantity | Stimuli (n) |
|---|---|---|---|---|---|
| **REF<55** | 0.0 | 0.0 | 0.0 | 0.0 | 252 |
| **UAH>55** | 1.1 | 0.5 | 1.1 | 0.4 | 974 |
| **AH** | 3.8 | 1.2 | 1.9 | 2.2 | 1403 |

The results of a further analysis of the direction of vowel feature confusions comparable to those presented in Table 16, p.26 are given in Table 27. Having in mind the same caveats as above (see discussion of the data from Table 16), similar tendencies can be observed. Here, too, there was a dominance of *rounded* => *unrounded* confusions (most conclusive for the AH group with 5.4% for *rounded* => *unrounded* vs. 1.6% for *unrounded* => *rounded*). Further, for the place feature *mid* => *other* confusions dominated for both groups UAH>55 and AH.

**Table 27.** *Forced choice test.* Percentage of phonemic feature confusions for the vowels (relative to the sum of possible errors per vowel category) averaged across the number of vowel phonemes involved. Number of vowel phonemes in parentheses.

| **(a) Rounding** | rounded | | unrounded | | | | | |
|---|---|---|---|---|---|---|---|---|
| | => unrounded | | => rounded | | | | | |
| REF<55 | 0.0 | | 0.0 | | | | | |
| UAH>55 | 1.4 | (8) | 0.7 | (6) | | | | |
| AH | 5.4 | | 1.6 | | | | | |
| **(b) Opening** | close | | close-mid | | open-mid | | open | |
| | => other | | => other | | => other | | => other | |
| REF<55 | 0.0 | | 0.0 | | - | | 0.0 | |
| UAH>55 | 0.2 | (7) | 1.2 | (5) | - | (2) | 0.0 | (2) |
| AH | 1.3 | | 0.8 | | - | | 1.7 | |
| **(c) Place** | front | | mid | | back | | | |
| | => other | | => other | | => other | | | |
| REF<55 | 0.0 | | 0.0 | | 0.0 | | | |
| UAH>55 | 0.6 | (8) | 4.8 | (1) | 1.1 | (5) | | |
| AH | 0.7 | | 12.5 | | 1.6 | | | |
| **(d) Quantity** | long | | short | | | | | |
| | => short | | => long | | | | | |
| REF<55 | 0.0 | | 0.0 | | | | | |
| UAH>55 | 0.8 | (7) | 0.0 | (7) | | | | |
| AH | 2.5 | | 1.9 | | | | | |

Summarising it can be stated that the number of vowel confusions has been too low to allow firm conclusions concerning the role of features. It is reassuring, however, that the main tendencies shown by the *forced choice test* data are in line with those from the *word test*.

### 9.3.3 Consonant confusions

Table 31 to Table 33 present the results for the consonant confusions in the forced choice test. Confirming all the previously discussed results, the three listener groups showed increasing mean overall error rates of 0.6 %, 4.5% and 8.1% for REF<55, UAH>55, and AH, respectively. Note that also these percentages are well below the mean error rates found for consonant confusions in the *word test* (4.3%, 12.3%, and 20.3%).

Comparing the error percentages for the consonant categories (Table 28), for the listener groups UAH>55 and AH the class of the *fricatives* was most prone to errors, followed by *nasals, plosives* and *nasals*. This hierarchy is different from the one that emerged from the *word test* (cf. Table 20, p.30). Since the *word test* offered more degrees of freedom as to phoneme confusions and in fact a larger number of errors occurred with that test format, the results that emerged from that test can be regarded to be more reliable.

**Table 28.** *Forced choice test.* Mean error rates in consonant perception plus number of phoneme categories and number of possible errors.

| Listener group | Category | Error rate | Phonemes (n) | Possible (n) |
|---|---|---|---|---|
| | plosive | 0.2 | 6 | 262 |
| **REF<55** | fricative | 0.9 | 6 | 191 |
| | nasal | 0.0 | 4 | 72 |
| | liquid | 1.7 | 3 | 130 |
| | plosive | 3.9 | 6 | 1069 |
| **UAH>55** | fricative | 5.7 | 6 | 744 |
| | nasal | 4.4 | 4 | 248 |
| | liquid | 3.2 | 3 | 509 |
| | plosive | 6.8 | 6 | 1591 |
| **AH** | fricative | 10.3 | 7 | 1030 |
| | nasal | 8.9 | 4 | 465 |
| | liquid | 4.3 | 3 | 676 |

In contrast to the diverging results for the two different test formats discussed above, the picture found for the contribution of the consonant features in the *forced choice test* confirmed previous results (cf. Table 29 and Table 21,p.30): The *place* feature was of primary importance (cf. the error rate of 6.5% for the AH group) followed by *manner* (4.3%) and *voicing* (1.2%). For the other two listener groups, the order is the same.

**Table 29.** *Forced choice test.* Error rates for phonemic features involved in consonant confusions and number of stimuli.

| Listener group | Manner | Place | Voicing | Stimuli (n) |
|---|---|---|---|---|
| **REF<55** | 0.3 | 0.4 | 0.0 | 252 |
| **UAH>55** | 2.0 | 3.7 | 0.9 | 974 |
| **AH** | 4.3 | 6.5 | 1.2 | 1403 |

Finally, the detailed information on the role of consonant features given in Table 30 differs somewhat from the distribution of errors for the *word test* (cf. Table 22, p.31). It seems to be likely that the less prominent roles of the *place* feature for the nasals and the *manner* feature for the liquids can are due to test-inherent constraints.

**Table 30.** *Forced choice test.* Percentage of phonemic feature confusions for the consonants (relative to the sum of possible errors per consonant category) averaged across the number of consonant phonemes involved. n= number of consonant phonemes.

| Listener group | Category | Manner | Place | Voicing | n |
|---|---|---|---|---|---|
| | plosive | 0.0 | 0.2 | 0.0 | 6 |
| **REF<55** | fricative | 0.0 | 0.9 | 0.0 | 6 |
| | nasal | 0.0 | 0.0 | 0.0 | 4 |
| | liquid | 1.7 | 0.0 | 0.0 | 3 |
| | plosive | 1.0 | 2.8 | 1.7 | 6 |
| **UAH>55** | fricative | 3.1 | 5.4 | 0.7 | 6 |
| | nasal | 1.0 | 4.2 | 0.2 | 4 |
| | liquid | 3.2 | 1.5 | 0.7 | 3 |
| | plosive | 2.5 | 5.1 | 2.5 | 6 |
| **AH** | fricative | 7.0 | 9.4 | 1.1 | 7 |
| | nasal | 2.3 | 7.5 | 0.2 | 4 |
| | liquid | 4.0 | 1.3 | 0.3 | 3 |

**Table 31.** *Forced choice test.* Consonant confusion matrix for listener group REF<55. Sum= sum of possible errors (indicated along the diagonal) and actual confusions.

| | Stimulus | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Response** | \. | p. | b. | t. | d. | k. | g. | m. | n. | =n | J. | rn | N. | r. | f. | v. | s. | S. | C. | j. | h. | l. | L. |
| \. | | | | | | | | | | | | | | | | | | | | | | | |
| p. | | 79 | | | | | | | | | | | | | | | | | | | | | |
| b. | | | 34 | | | | | | | | | | | | | | | | | | | | |
| t. | | | | 22 | | | | | | | | | | | | | | | | | | | |
| d. | | | | | 24 | | 1 | | | | | | | | | | | | | | | | |
| k. | | | | | | 35 | | | | | | | | | | | | | | | | | |
| g. | | | | | | | 68 | | | | | | | | | | | | | | | | |
| m. | | | | | | | | 29 | | | | | | | | | | | | | | | |
| n. | | | | | | | | | 29 | | | | | | | | | | | | | | |
| =n | | | | | | | | | | 0 | | | | | | | | | | | | | |
| J. | | | | | | | | | | | 12 | | | | | | | | | | | | |
| rn | | | | | | | | | | | | 2 | | | | | | | | | | | |
| N. | | | | | | | | | | | | | 0 | | | | | | | | | | |
| r. | | | | | | | | | | | | | | 55 | | | | | | | | 1 | |
| f. | | | | | | | | | | | | | | | 36 | | | 3 | | | | | |
| v. | | | | | | | | | | | | | | | | 6 | | | | | | | |
| s. | | | | | | | | | | | | | | | | | 54 | | | | | | |
| S. | | | | | | | | | | | | | | | | | | 55 | | | | | |
| C. | | | | | | | | | | | | | | | | | | | 22 | | | | |
| j. | | | | | | | | | | | | | | | | | | | | 0 | | | |
| h. | | | | | | | | | | | | | | | | | | | | | 18 | | |
| l. | | | | | | | | | | | | | | 2 | | | | | | | | 71 | |
| L. | | | | | | | | | | | | | | | | | | | | | | | 4 |
| **sum** | | 79 | 34 | 22 | 24 | 35 | 69 | 29 | 29 | 0 | 12 | 2 | 0 | 57 | 36 | 6 | 54 | 58 | 22 | 0 | 18 | 72 | 4 |
| **% error** | | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.5 | 0.0 | 0.0 | - | 0.0 | 0.0 | - | 3.6 | 0.0 | 0.0 | 0.0 | 5.5 | 0.0 | - | 0.0 | 1.4 | 0.0 |

**Table 32.** *Forced choice test.* Consonant confusion matrix for listener group UAH>55. Sum= sum of possible errors (indicated along the diagonal) and actual confusions.

| | **Stimulus** | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Response** | \. | p. | b. | t. | d. | k. | g. | m. | n. | =n | J. | rn | N. | r. | f. | v. | s. | S. | C. | j. | h. | l. | L. |
| \. | | | | 1 | | | | | | | | | | | | | | | | | | | |
| p. | | 307 | 2 | 1 | | 1 | 1 | | | | | | | 2 | 4 | | | | | | 1 | | |
| b. | | 3 | 142 | 1 | | | | | | | | | | | | 1 | 2 | 1 | | | | 3 | |
| t. | | | | 96 | | 3 | | | | | | | | | | | 1 | | | | | 1 | |
| d. | | | | | 166 | 1 | 1 | | | | | | | | 1 | 1 | 1 | | 1 | | | | |
| k. | | 4 | | | | 131 | 1 | | 1 | | | | | 1 | | | 1 | | | | | 1 | |
| g. | | | | | 2 | 2 | 227 | | | | | | | | | | 1 | | | | | | |
| m. | | | | | | | | 67 | | | 2 | | | 2 | | | | | | | | | |
| n. | | 1 | | | | | | 1 | 119 | | | 1 | | 2 | | | 1 | | | | | 3 | |
| =n | | | | | | | | | 1 | 0 | | | | | | | | | | | | | |
| J. | | | | | | | | | | | 42 | | | | | | | | | | | | |
| rn | | | | | | | | | 1 | | | 20 | | | | | | | | | | | |
| N. | | | | | | | | | 1 | | | | 0 | | | | | | | | | | |
| r. | | | | | | | | | | | | | | 237 | | | | | 1 | | | 1 | |
| f. | | 3 | 1 | | | | | | | | | | | | 159 | | 6 | 10 | | | | | |
| v. | | | | 3 | 1 | | 1 | | | | | | | 1 | | 36 | | | | | | | |
| s. | | | | | | 1 | | | | | | | | | | | 198 | 4 | 1 | | | | |
| S. | | | | 1 | | | | | | | | | | | 3 | | | 174 | 1 | | | | |
| C. | | | | | | | | | | | | | | | 1 | | | | 132 | | | 1 | |
| j. | | | | | | | | | | | | | | | | | | | | 0 | | | |
| h. | | 1 | | | | | | 2 | | | | | | | | | | | 1 | | 45 | | |
| l. | | | | | 1 | | | | | | | | | 6 | 1 | | 1 | | 1 | | | 262 | |
| L. | | | | | | | | | | | | | | | | | | | | | | | 10 |
| **sum** | | 319 | 145 | 103 | 170 | 139 | 231 | 70 | 123 | 0 | 44 | 21 | 0 | 251 | 169 | 38 | 212 | 189 | 138 | 0 | 46 | 272 | 10 |
| **% error** | | 3.9 | 2.1 | 7.3 | 2.4 | 6.1 | 1.8 | 4.5 | 3.4 | - | 4.8 | 5.0 | - | 5.9 | 6.3 | 5.6 | 7.1 | 8.6 | 4.5 | - | 2.2 | 3.8 | 0.0 |

**Table 33.** *Forced choice test.* Consonant confusion matrix for listener group AH. Sum= sum of possible errors (indicated along the diagonal) and actual confusions.

| | **Stimulus** | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Response** | \. | p. | b. | t. | d. | k. | g. | m. | n. | =n | J. | rn | N. | r. | f. | v. | s. | S. | C. | j. | h. | l. | L. |
| \. | | | 1 | 2 | | | | | | | | | | 3 | | | | 1 | | | | | |
| p. | | 464 | 2 | 3 | 1 | 1 | | | 1 | | | | | | 8 | | 2 | | | | 2 | | |
| b. | | 8 | 216 | | | | 1 | | 1 | | | | | | 3 | 2 | | | | 2 | | 1 | |
| t. | | | | 130 | 2 | 5 | 1 | | | | | | | | | | | | | | | | |
| d. | | 1 | | 3 | 244 | | 5 | | | | | 1 | | | 1 | 1 | 5 | | 3 | | | | |
| k. | | 9 | 1 | | | 230 | 2 | | 1 | | | | | | | | 6 | | | | | 1 | |
| g. | 3 | 4 | 1 | 3 | 6 | 3 | 307 | | | | | | | | | | 1 | | | | | | |
| m. | | 3 | | | | | 1 | 149 | | | 4 | | | 1 | | | | | | | | | |
| n. | | | | | | 1 | | 3 | 217 | | | 6 | | 1 | | | 1 | 2 | 1 | | | | |
| =n | | | | | | | | | 4 | 0 | | | | | | | | | | | | | |
| J. | | | | | | | | | | | 57 | | | | | | | | | | | | |
| rn | | | | | | | | | | | | 42 | | | | | | | | | | | |
| N. | | | | | | | | | 3 | | | | 0 | | | | | | | | | | |
| r. | 31 | | 1 | | | | | | | | | | | 328 | | | 3 | | | | | 6 | |
| f. | 1 | 9 | 2 | | 4 | | | | | | | | | 1 | 204 | | 7 | 7 | 7 | | | 1 | |
| v. | 2 | | | | 2 | | 1 | | | | | | | 3 | | 63 | | | | | | | |
| s. | | | | | | | | | 1 | | | | | | | | 256 | 6 | | | | | |
| S. | | | | 5 | | | | | | | | | | | 4 | | 5 | 241 | 4 | | | | |
| C. | | | | | | | | | | | | | | | 5 | | | 6 | 167 | | | | |
| j. | | | | | | | | | 2 | | | | | | | | | | | 0 | | 1 | |
| h. | | 7 | 1 | | | | 3 | 1 | | | | | | 4 | | | | | 3 | | 87 | | |
| l. | 2 | | | | | | 2 | | 7 | | | | | 19 | 2 | 2 | 3 | | 2 | | | 328 | |
| L. | | | | | | | | | | | | | | | | | | | | | | | 20 |
| **sum** | | 505 | 225 | 146 | 259 | 240 | 323 | 153 | 237 | 0 | 61 | 49 | 0 | 360 | 227 | 68 | 289 | 263 | 187 | 2 | 89 | 338 | 20 |
| **% error** | | 8.8 | 4.2 | 12.3 | 6.1 | 4.3 | 5.2 | 2.7 | 9.2 | - | 7.0 | 16.7 | - | 9.8 | 11.3 | 7.9 | 12.9 | 9.1 | 12.0 | 16.7 | 2.3 | 3.0 | 0.0 |

# 10 General conclusions from the listening test

## 10.1 Methodology

The final goal of the project reported here has been to shed some light on what kind of problems hearing impaired listeners encounter when listening to spoken Norwegian. The scope of the issues investigated ranged from fairly general questions like word boundary recognition to very specific questions concerning the contribution of distinctive features to phoneme confusions. To the best of our knowledge, the chosen test design using syntactically correct but semantically unpredictable utterances in three different test formats (*sentence*, *word* and *forced choice*) has not been applied before.

As an overall result, the varying degree of participation of the three listener groups in the different test types revealed the actual need for varying degrees of difficulty in the listening tests. For example, for the group with normal hearing, testing exclusively through the *forced choice* format would have excluded us from gaining the rich information now present in the *sentence test*'s results. On the other hand, the *sentence test* posed a problem for the aided hearing group in that intelligibility rapidly could go down to zero. For this listener group, using the *sentence test* format alone would have led to often inconclusive results and implied loosing the aided hearing group as a potential source of information.

It should be kept in mind, however, that besides advantages the flexible listening test design also implied certain drawbacks. Lack of systematic variation of the experimental factors can be mentioned as the probably most obvious disadvantage. Since both the number of observations and the listening conditions regarding noise level and room acoustics varied from group to group and even from listener to listener, absolute comparisons of listening performance were impossible. Further, the number of observations for the speech sounds under scrutiny varied, so that for some speech sounds insufficient data were collected. As a consequence, for a number of cases no firm conclusions could be drawn. The latter seems especially to be true for the *forced choice* test, where a larger amount of misperceived speech sounds would have been desirable. In general, due to the varying number of observations for the experimental conditions, statistic evaluation of the listening test results was often problematic or appeared even to be inappropriate.

## 10.2 Listening test results

The results of the listening tests have been presented and discussed in detail above. Therefore, in the following only a summary of the main results and conclusions will be given.

### 10.2.1 Sentence test

- The listening tests have succeeded in achieving the main goal, viz. to make an inventory of the problems hearing-impaired people encounter when trying to understand spoken Norwegian. Listening to speech in the presence of concurrent babbling noise led to misperceptions not only for hearing-impaired but also for normal-hearing subjects. It is noteworthy that while error rates were lowest for the normal-hearing reference group, similar error patterns were found for all the three listener groups. This allows the conclusion that research directed toward improving hearing aids can to a large extent recruit normal-hearing subjects. In this way experimental testing can be facilitated and speeded up.
- Speaker dialect appeared to have an impact on intelligibility. Lowest error rates were found for utterances recorded from the East Norwegian speakers, irrespective of the dialectal background of the listener. Therefore, it might be speculated that familiarity with a given dialect does not necessarily facilitate intelligibility. Of course, it should be kept in mind that a

rather limited number of speakers were involved in this investigation. This is certainly an issue for future research.

- Investigation of speaker gender showed that utterances spoken by the female speakers were more intelligible than those produced by the males. This effect was statistically significant, though with a difference in error rate of 3-4% not very big.
- As evidenced by remarkably low errors in word boundary recognition, listeners succeeded well in dividing up the stream of speech into word entities. In line with this observation, no syllable boundary errors were registered. Though this performance might in part be due to the relative simple and to a certain degree predictable syllable/word structure, these results seem to justify focussing on phoneme rather than word level in future research.
- Recognition errors varied as a function of position of the word in the utterance. Contrary to what had been postulated, least errors occurred in utterance-initial words. Approximately twice as many errors were registered in utterance-final words, while words in medial position appeared to be hardest to understand. There seem to be obvious explanations for this outcome, neither in connection with linguistic factors like sentence structure and word categories (e.g., content vs. function word) nor with regard to articulation behaviour on the part of the speakers.

### 10.2.2 Word test and forced choice test

A comparison of the results from the *word test* and the *forced choice test* indicated that the former was better suited to elicit all kinds of misperceptions. The more stringent format of the latter test type, where the listeners had restricted degrees of freedom and not all possible phoneme oppositions were included in the response alternatives, implied a reduced potential of errors.

To a certain extent the data from the two tests showed similar tendencies, but having the above-mentioned limitations of the *forced choice test* in mind the following conclusions are mainly based on the *word test* results.

- First of all the data showed that besides consonants also vowels are to a large extent error-prone: The ratio of mean error rates for vowels/consonants was approximately 0.8.
- Vowel confusions were primarily caused by confusions in *rounding* and *place*. Specifically, *rounded* vowels were perceived as being *unrounded* (rather than vice versa). *Place* errors occurred especially with *mid* but to a lesser degree also with *back* vowels. Degree of *opening* appeared to be relevant mainly for open-mid vowels.
- Consonant confusions were primarily due to *place* errors. Consonant *manner* was of secondary importance, while the feature *voicing* played a very limited role.
- Overall, nasals were most prone to error, followed by plosives, fricatives and, finally, liquids. For the first three sound categories, errors were mostly due to misperceived *place* of articulation. Errors in perception of the liquids (/r/ and /l/) were typically due to *manner* confusions, mostly /r/ being perceived as /l/ and vice versa.

# 11 Algorithms for digital hearing aids

A "brain-storming" was organised within the scientific groups in acoustics, signal processing, phonetics and audiology. Only speech enhancement algorithms were discussed. More than 20 proposals were presented and initially discussed. The ideas are classified in 4 groups:

1. Algorithms in the frequency domain
2. Algorithms in the time domain
3. Hybrids between the two above.
4. Binaural algorithms

Promising ideas in each group were implemented, partly by students supervised by members of the project team, or by members of the team.

Unfortunately, algorithms must be implemented in a hearing aid or a real time signal processor simulating a hearing aid for testing. Time and resources did not allow real time implementation. The proposed algorithms were implemented in MatLAB. A very limited speech material was processed, and a simple comparison of processed and unprocessed speech was used for evaluation.

Real time implementation of the algorithms was therefore left as an important future task.

## 11.1 Problems in speech perception converted to problems in signal processing

The listening test confirms that the problems mainly may be characterised as confusion between phonemes. *The main problems thus seems to be rather general, not specific for the Norwegian language. Speech enhancement algorithms may thus be optimised regardless of language.* The consequences of confusion for the intelligibility of running speech may be highly language dependent. As an example, in Norwegian the numbers "ni" (9) and "ti" (10) may easily be exchanged. Other problems may occur in for example English and German.

The reason for phoneme confusions in running speech may be:

- The loudness of the whole phoneme or spectral parts of the phonemes is too low, falling below the threshold of hearing or being masked by background noise or by reverberation.
- The perceived differences are too small to allow a secure discrimination.

Persons with neural hearing losses need generally greater feature differences for an appropriate discrimination compared to persons with a normal hearing.

All hearing aids try to compensate parts of the hearing loss by frequency dependent amplification and by compressing the dynamic range. The fundamental problem with increasing sound levels is that the excitation of the basilar membrane is broadening and the spectral discrimination is reduced.

The signal to noise ratio is principally not affected by amplification. The consequences of information loss by masking are however more serious for people having marginal hearing compared to those having normal hearing.

*Syllabic compression* has to some extent been implemented for reducing the effects of *postmasking.*

Speech enhancement algorithms should in priority address the following class of phonemes:

Consonants:

- plosives
- fricatives
- nasals

- liquids

It is mostly the effects of *place* of articulation which cause the problems.

Also among vowel confusions may be found and the greatest problem is the feature *rounding.*

Further discussion may be found in [ 13 ].

### 11.1.1 Plosives

Plosives are characterised by a closing phase, a closed phase (occlusion), and a release burst.

The important frequency range of the spectra of the bursts of plosives are above 2000 Hz and thus in the range of severe hearing losses by people with presbyacusis or noise induced hearing loss. Normal procedures by fitting hearing aids compensate only half of this loss. Insufficient loudness of the bursts may be one reason for identification problems between plosives.

The occlusion may be masked by background noise or reverberation and thus making it difficult to identify plosives in running speech. Important cues may even be found in the closing phase, where the vocal tract of the previous phoneme is gradually changed to reach the target of the plosive.

Reverberation may mask this phase and even the temporal details in the plosive. Selective amplification or prolongation of the bursts may improve the identification of the plosives.

*Increasing the dynamics of the plosives in the context may be of importance.*

*Enhancement of the spectral characteristics of the burst will improve intelligibility.*

### 11.1.2 Fricatives

Fricatives, like the burst noise of plosives, are characterised by having most of the spectral density above 2000 Hz, and may thus easily fall below the threshold of hearing by neural hearing losses. The spectral structure is characterised by both broad peaks and valeys (poles and zeroes). Zeroes are more easily corrupted by noise and reverberation than poles, so *reshaping the zeroes may be of importance.*

### 11.1.3 Nasals

Nasals are generally of lower sound intensity than oral voiced phonemes. Radiation from the nose openings are less effective than from the mouth.

The spectrum is characterised by a formant at about 1500 Hz. For a given speaker the formant frequency is the same for all nasals, due to the fact that the volume and form of the nasal cavity is not varied by articulatory organs.

The closed oral cavity introduces zeroes in the spectra. The different nasals are mainly identified by the zeroes and are thus more easily corrupted by noise, reverberation and also by hearing loss in the frequency-range above 1000 Hz. Most hearing aids remove the normal ear-canal resonance which is in the frequency range of formant and zeroes in nasals.

*Increasing the relative sound levels and enhancing the zeroes in the spectrum may increase the possibility of ditinguishing between nasals.*

### 11.1.4 Liquids

*The liquids, "r" and "l"*, have the same place of articulation but differ in manner of articulation. Spectral differences are found only at high frequencies (3. formants) and is thus reduced by neural hearing losses.

The identification of "r" may be difficult due to the fact that several different realisations occur in Norwegian dialects: Single flap, rolled, velar, and uvular.

The "roll" modulation, a low frequency modulation of intensity in the range 10 Hz, may be reduced by compression in hearing aids.

*Enhancement of envelope variations in the frequency range above 4 Hz may increase the sound levels of "r", and improve the distinguishing between different liquids.*

## 11.2 Classes of algorithms

The ideas for the algorithms are classified in the four groups: Frequency domain, time domain, hybrids between these two, and binaural algorithms. See also [ 14 ] and [ 15 ].

### 11.2.1 Algorithms in the frequency domain

*Bård Støfringsdal* [ 7 ] has implemented an adaptive filter to increase the levels of formants in fluent speech. The principle is similar to *postfiltering* used in speech coding.

Three sections of 2.order IIR filters [ 8 ] are used for moderate spectral enhancement of peaks in the speech signal spectrum. Maximum 10 dB relative amplification of peak frequencies is allowed. This limitation is due to the problems of artefacts.

Spectral peaks are found by Linear Predictive Coding Analysis of the input signal, determining the optimal inverse FIR-filter. The algorithm is tracking the formants in frames of 20 ms, updated every 5 ms. To reduce artefacts due to quickly varying formant frequencies parameters are changed only at zero crossings in the signal.

The algorithm is implemented in MatLAB. It is tested by comparing processed and original speech using a limited speech material only. The method is promising, but refinements are necessary.

### 11.2.2 Algorithms in the time domain.

*Barbara Resch* [ 6 ], supervised by *A. Krokstad* and *G. Ottesen,* has developed an algorithm for speech enhancement in the time domain. The basic idea is to modify the envelope of the speech signal. Two methods are tested for extracting the signal envelope: Squaring of the signal samples or using the magnitude of the samples.

The Modulation Spectrum (MS) is found as the Fourier transform of the envelope. The modulation spectrum is divided in a Low Frequency (LF) part and a High Frequency (HF) part. The cross over frequency is selectable in the range 1-4 Hz.

The basic idea is that LF is a measure of the speech level, while HF represent important information in the speech signal, especially transients.

Speech transmission index (STI), a method for estimating the properties of a speech transmission system, is based on the modulation spectrum.

After Inverse Transforms, LF is used as basic for speech compression, while HP is used for expansion. Rapid changes are thus enhanced while the mean speech level is tried to be kept constant. Ordinary compressors used in hearing aids are trying to reduce all envelope variations.

The algorithms are tested on some German sentences. Optimising selectable parameters is found to be very important. A basic problem is to avoid delay in the processing of the envelope.

A Japanese speech-processing group at ARAI Laboratories simultaneously has been working on a very similar approach, [ 16 ]. Preliminary results are very promising.

### 11.2.3 Hybrid methods

*Anna Kim* [ 4 ], supervised by *Jarle Svean,* has worked on algorithms where frames of 20 ms speech are classified in three classes, voiced, unvoiced and "not speech", and each class is processed differently. The classification of each frame is based on Cepstrum analysis.

Signal to noise ratio is improved by suppressing less important parts of the spectrum for voiced and unvoiced frames, and of course suppressing the whole frame when it is classified as "not speech".

An intelligibility test using nonsense syllables was performed. Persons with both normal hearing and with hearing loss were tested. Significant improvements of intelligibility were found.

### 11.2.4 Binaural methods

Based on an idea of Wim van Dommelen, Hans-Petter Vadset [ 5 ] supervised by *Peter Svensson* has tried to verify whether different processing methods should be implemented for the left and the right ear. The rationale is that the left and right hemisphere of the brain extract different types of signal information.

The hypothesis is that the right hemisphere processes more long term global information (e.g. intonation) while the left one is specialised for processing the quick variations (e.g. the sequence of sounds).

The speech signal was divided into a sequence of pure spectrum patterns, and a pure envelope signal. The two signals were fed to left and right ear by headphones. Differences in intelligibility by shifting the two headphones were indicated.

The intelligibility was a little higher when the time sequence exited the right ear and the spectral pattern the left ear. Due to crossing of the nerves (the right ear feeding mostly the left hemisphere of the brain), this is as expected. The difference was small, just significant, so it is doubtful if differences in processing will represent an improvement when localising is taken into account. The test, however, confirms the importance of both the time sequence (the rhythmical pattern) and the spectral patterns.

## 12 Talking to hard of hearing people

A guide in Norwegian, [ 17 ], especially addresses health personal who professionally communicates with elderly people. The guidelines should be of general interest, however. The guide is formulated as *advises* with comments and some discussion, and is cited in English below:

1. *Never exaggerate.*
   The hard of hearing people don't like to announce their handicap.
2. *Speak a little louder than you usually do, but never shout.*
   The hearing loss has to be partly compensated by increasing the loudness of the speech. But it is important not to over-compensate. The possibility of discriminating between phonemes is greatest when all important features of the sounds are suitable above threshold of hearing. When speaking, try to test as soon as possible if the speaking level is adequate.
3. *Speak a little slower than usual.*
   Speaking slower improves both the quality of articulation and the listeners' possibilities for processing the information.
4. *Speak face to face or at least with visual contact.*
   Lip reading is of great help. Critical high frequency sound as "t", "p", "k", "s", and "f" are radiated strongest directly ahead.
   A small distance between speaker and listener will reduce the influence of room reverberation and will contribute to achieving a higher speech level.
   Do not try to communicate with hard of hearing people from one room to another. While people with normal hearing may tolerate the lack of direct sound, people with hearing losses may not. (It is a frequent complain that the husband in the living room does not understand his wife speaking from the kitchen).
5. *Distinct articulation is of outmost importance.*
   Speaking very clearly, carefully underlining differences between similar phonemes and words is of great benefit to compensate for hearing losses.
   Increasing loudness and duration of the fricatives ("f", "s", "sh", and "kj") and the burst of plosives ("p", "t", and "k") is more important than a general increase in loudness.
   Differences between nasal sounds ("m", "n", "ng" and "rn") should be underlined by careful pronunciation.
6. *Speak at your normal low pitch.*
   Avoid increasing the pitch du to the louder voice. It is better to use low pitch.
7. *Be aware of confusable words.*
   Confusion is more frequent between words of the same word class than of different classes. For example are the are the place names "Moss"-"Voss" and the numbers "ni"-"ti" and " førti"-"søtti" (9-10 and 40-70 in Norwegian) a greater problem than the words "kam"-"kan" (noun "comb" and verb "can").
8. *The start of an information unit is important.*
   We usually do not need to hear all words in a sentence as some may be guessed from the context. Hard of hearing people has to rely heavily on guessing.
   The beginning is important for guessing the rest and people who have "lost the track" will often stop trying to understand at all.
9. *Numbers and names can not be guessed.*
10. *Using dialects or slang may represent a greater problem for hard of hearing people than by people with normal hearing.*
    The main problem in understanding people using dialects is not the occurrences of uncommon

words but the frequent increase of speaking rate and less careful pronunciation. The listening test suggested that people with hearing loss may have greater problems understanding even their own dialect than a normalised language, especially a language commonly used in Radio and TV broadcast.
In marginal listening conditions especially hard of hearing people may experience great differences between dialects in their ability to comprehend speech.

11. *Avoid speaking in environments with a high background noise level.*
We automatically compensate for background noise by speaking louder but it is a limited range of background noise levels that may be compensated for. This range is smaller for hard of hearing people.
Background noise (e.g. traffic noise) is frequently dominated by low frequencies. The information density in speech is moderate in this frequency range. The residual information here may however be of outmost importance to people having hearing loss at high frequencies.

12. *Especially avoid "Speech Noise" from several simultaneous speakers in the same room.*
The ability to localise a certain speaker and to follow his of speech among several is deteriorated by hearing loss and is not successfully restored by hearing aids.
Reduced reverberation time in rooms reduces mixing of several conversations and may ease the comprehension of speech.
"The Cocktail Party" problem in acoustics describes the situation in a party where everyone tries to compensate for a high speech noise level by speaking louder. This in turn causes a further increase in the overall noise level. The problem may be remedied by dividing the party into small and geometrically concentrated groups.

# 13 Conclusion

The type of speech material used in the listening tests described in the preceding chapters is known as semantically unpredictable sentences (SUS). The main innovative idea in the design of the listening tests was the use of the same speech material with three different formats, the *sentence test*, the *word test*, and the *forced choice test*. The rationale behind this flexible design was to maximise the amount of information to be extracted from listeners with varying degrees of hearing loss.

The results of the tests indicated that in general the application of our listening test design has been successful. Especially through the data from the *word test* we have been able to collect detailed and potentially useful information about the phonetics of speech intelligibility for hard of hearing people. In the *forced choice test* only a limited set of alternatives was available. Due to these test-inherent limitations the results from this test were somewhat less informative. With future use of the present test design it may, therefore, seem advisable to refine and extend the inventory of response alternatives in this test.

Beyond its flexibility, the test design had the advantage of supplying us with data related to different levels in speech perception. On the one hand, the phoneme confusion matrices based on the *word* and *forced choice test* data allowed detailed conclusions with regard to the role of single phonemes and phonetic features. On the other hand, the *sentence test* data shed light on more global issues in connection with parsing. In conclusion it can be stated that the total of information emerging from the listening tests forms a solid basis for the development and further improvement of digital hearing aid algorithms. During the course of such work, the listening test material can be reused to monitor progress and to facilitate the prioritisation of research and development issues.

Though the listening tests series has supplied a considerable amount of valuable information, a number of issues deserve to be investigated in future research. One such issue is concerned with the role of dialect in speech intelligibility. The present results suggested that the utterances spoken by South-East Norwegian speakers were easier to understand than those from other dialects, even for other than South-East Norwegian listeners. It should be kept in mind, however, that the number of speakers and different dialects in our investigation was rather limited. Therefore, it might be interesting to pursue the issue of dialectal differences by enlarging the number of speakers and, possibly, dialects represented. Results based on a more systematic investigation of this issue will potentially allow us to draw more firm conclusions.

Both production and perception are important issues in the process of understanding speech. One of the topics of the project was thus how to make oneself understood by the hard of hearing people. Obviously, careful pronunciation is of great importance. A chapter with guidelines is included in the report.

A set of promising processing algorithms for digital hearing aids has evolved during the project. Together with other algorithms these are candidates for further development, implementation in real time processing hardware for functional testing, and finally implementation in digital hearing aids.

## References

**[ 1 ]** Houtgast and Steeneken: "A Multi-Language Evaluation of the RASTI-Method for Estimating Speech Intelligibility in Auditoria". Acustica, Vol. 54 (1984), pp185-199.

**[ 2 ]** IEC International Standard 268 Sound Systems, Part 16 The objective rating of speech intelligibility in auditoria by the "RASTI" method.

**[ 3 ]** SAMPA (Speech Assessment Methods Phonetic Alphabet) is a machine-readable phonetic alphabet. Homepage URL: http://www.phon.ucl.ac.uk/home/sampa/home.htm
URL for Norwegian: http://www.phon.ucl.ac.uk/home/sampa/norweg.htm

**[ 4 ]** Kim, Anna Na: "Speech Enhancement Algorithm for Digital Hearing Aid". Master thesis NTNU Inst for teleteknikk May 2000.

**[ 5 ]** Vadset, Hans-Petter: "Binaural Signal Processing of Speech Sounds". Master thesis NTNU Inst for teleteknikk, 20.12.2000.

**[ 6 ]** Resch, Barbara: "Algorithm for digital hearing aid with temporal expansion". Project thesis NTNU Inst, for teleteknikk/Acoustics, 2000/2001.

**[ 7 ]** Støfringsdal, Bård: "Algoritmer for høreapparat. Spesifikasjon MATLAB-rutiner for forbetring av taleoppfattbarheit". Prject memo (this project), August 2001. (In Norwegian).

**[ 8 ]** Regalia, P.A. and Mitra, S.K.: "Tunable Digital Frequency Response Equalization Filters". IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-35, No. 1, 1987, pp118.

**[ 9 ]** Peutz, V. M. A.: "Articulation Loss of Consonants as a Criteria for Speech Transmission in a Room," J. Audio Eng. Soc., Vol. 19 (Dec 1971).

**[ 10 ]** Stensby, Sverre; Krokstad, Asbjørn; og van Dommelen, Wim A.: "Algoritmer for høreapparet". Presentation at NAS (Norsk Akustisk Selskap. The Norwegian Acoustical Society). høstmøte 27 oktober 2001 in Trondheim (in Norwegian). URL (2002-01-15): http://www.akustisk-selskap.com/hoestmoetet_2001/html_hostmotet/sverre_stensby.htm

**[ 11 ]** van Dommelen, Wim A.; Stensby, Sverre; og Krokstad, Asbjørn: "Algoritmer for høreapparat". Presentation at "Etterutdanningskurs for Audiofysikere, Audioingeniører og Audiografer, 2001-11-15 - 2001-11-17" at Hell. (Included on the CD).

**[ 12 ]** American National Standard Methods for the Calculation of the Speech Intelligibility Index (Draft).

**[ 13 ]** Fant, Gunnar: "Speech Sounds and Features". MIT Press 1973.

**[ 14 ]** Parsons, Thomas W.: "Voice and Speech Processing". McGraw-Hill 1987.

**[ 15 ]** Furui, Sadaoki: "Future Directions in Speech Information Processing". Proceedings of The International Congress of Acoustics Seattle 1998.

**[ 16 ]** Kusumoto et al: "Modulation Enhancement of Speech as a Preprocessing for Reverberant Chambers with Hearing-impaired". Proc. of ICASSP 2000.

**[ 17 ]** Krokstad, Asbjørn; van Dommelen, Wim A.; and Stensby, Sverre: "Gode råd om å snakke med folk som er hørselshemmet". Notat 420202, Institutt for teleteknikk, NTNU. 01.02.2002. (In Norwegian)

# Appendices

## A. CD-rom

A CD-rom containing the following items is attached to the report:

- Material used in developing the listening test.
- The speech material used in the test.
- Results and evaluation of the listening test.
- This report, STF40 A02010. (File code: STF40 A02010.doc)
- Krokstad, Asbjørn; van Dommelen, Wim A.; and Stensby, Sverre: "Gode råd om å snakke med folk som er hørselshemmet". Notat 420202, Institutt for teleteknikk, NTNU. 01.02.2002. (In Norwegian). The reference [ 17 ]. (File code: Teleteknikk420202.doc)
- van Dommelen, Wim A.; Stensby, Sverre; og Krokstad, Asbjørn: "Algoritmer for høreapparat". Presentation at "Etterutdanningskurs for Audiofysikere, Audioingeniører og Audiografer, 2001-11-15 - 2001-11-17" at Hell. The reference [11] (PowerPoint presentation). (File code: HELLmøte.ppt)

During the project more terms have been used. In the table below the alternative terms are connected to the preferred English term:

| Preferred term | Norwegian terms or alternative terms |
|---|---|
| Sentence test | Setningstest |
| Word test | Ordtest   Hulltest. |
| Forced choice test | Flervalgstest  Rimtest. |
| (REF<55) | (REF<55) |
| REF>55 | UAH>55 |
| AH | TH   HI |

## B. Speech material used in the listening tests

The speech material used in the listening tests is presented in this appendix. The 150 sentences are presented together with the speaker code and the assigned block number.

Words in *italics* were left out from the orthographic representation in the *word test*.

In the column *Dialect forms*, actual deviant pronunciation is indicated using normal orthography.

Speaker codes (dialect region):

- males: TA (North Norwegian); PH (Trøndelag); BM (East Norwegian); KU (West Norwegian)
- females: AF (Trøndelag); KH (West Norwegian); SH (East Norwegian); TM (Trøndelag)

| Block | Speaker | Number | Sentence | Dialect forms |
|---|---|---|---|---|
| 1 | BM | 001 | Bilen vet at *barna* ville hale. | |
| 1 | BM | 002 | Boka tror at billene vil *helle*. | |
| 1 | TA | 003 | De tenker at *broren* tør holde. | |
| 1 | SH | 004 | Dere skriker at fillene torde *hulle*. | |
| 1 | SH | 005 | Du sier at flyet skulle *hyle*. | |
| 1 | AF | 006 | Faren ser at frukten skal *hylle*. | frukta |
| 1 | KU | 007 | Fniset røper at heftet måtte *kile*. | |
| 1 | KH | 008 | Han roper at hodet må *kjæle*. | |
| 1 | PH | 009 | Flisen påstår at pillene kunne *kyle*. | |
| 1 | PH | 010 | Hunden mener at *rillene* kan lære. | |
| 1 | TA | 011 | Jeg lærer at selen kan *se*. | Æ lær |
| 1 | KH | 012 | Loven hvisker at *sjefen* får skjære. | |
| 1 | TM | 013 | Mora husker at sjelen fikk *sy*. | |
| 1 | AF | 014 | Posten *frykter* at tanten bør så. | |
| 1 | TM | 015 | Vi forteller at veien burde *vanne*. | |
| 2 | BM | 016 | Når *truer* avisen feler? | |
| 2 | BM | 017 | Hvorfor treffer bilen *skyer*? | |
| 2 | TA | 018 | Når spiser blomstene *skjeer*? | spis blomstan skjee |
| 2 | TA | 019 | Hvor ser de *ski*? | |
| 2 | SH | 020 | Når mottar dere *sjøer*? | |
| 2 | KU | 021 | Hvorfor maler du *sju*? | |
| 2 | KU | 022 | Hvor *løfter* han sjeler? | |
| 2 | KH | 023 | Når lurer hun *sanger*? | |
| 2 | KH | 024 | Hvor *krever* hunden paver? | |
| 2 | AF | 025 | Når kjører huset *land*? | |
| 2 | PH | 026 | Hvorfor gjetter jeg *lam*? | |
| 2 | PH | 027 | Hvordan *føler* mormor kjeler? | |
| 2 | TM | 028 | Hvordan *frykter* snøen gaver? | |
| 2 | KH | 029 | Hvordan forteller tankene *deler*? | |
| 2 | TM | 030 | Når elsker vi *barn*? | |
| 3 | BM | 031 | Aldri synger faren en *rille*. | |
| 3 | BM | 032 | Alltid *spiser* fela en pille. | |
| 3 | SH | 033 | Denne uka spiller fienden et *pass*. | |
| 3 | SH | 034 | Helst *skriver* gaven en måke. | |

| | | | | |
|---|---|---|---|---|
| 3 | TA | 035 | I dag ser havet en *masse*. | |
| 3 | AF | 036 | I morgen leser kisten en *liste*. | morra læs kista ei |
| 3 | KU | 037 | Neste uke kjører *kjolen* en labb. | |
| 3 | TA | 038 | Ofte hører kjæresten en *kål*. | |
| 3 | PH | 039 | Om høsten hilser listen en *kiste*. | |
| 3 | KH | 040 | Om sommeren har læreren en *kasse*. | |
| 3 | AF | 041 | Om vinteren får maven en *kappe*. | magen |
| 3 | PH | 042 | Om våren *føler* paven en handel. | |
| 3 | TM | 043 | Sjelden fanger *sjelen* en fille. | |
| 3 | TM | 044 | Til frokost elsker *tjæren* en bille. | |
| 3 | KH | 045 | Til middag bærer *været* en ball. | |
| 4 | BM | 046 | Den faste pilen *synger* aldri. | |
| 4 | BM | 047 | Det feige skjerfet *skyller* alltid. | |
| 4 | AF | 048 | Det gale skjegget *raser* denne uka. | |
| 4 | SH | 049 | Det ferske loddet *roser* best. | |
| 4 | SH | 050 | Den fete *bilen* reiser i dag. | |
| 4 | KH | 051 | Den fine øya *raker* i morgen. | |
| 4 | KU | 052 | Den fjerde lyden *rager* neste uke. | |
| 4 | TA | 053 | Det fjerne *skinnet* raller ofte. | ralle |
| 4 | KH | 054 | Den flate snøen *rekker* om høsten. | |
| 4 | KU | 055 | Det flaue skipet *renner* om sommeren. | |
| 4 | TA | 056 | Den fremre tanken *revner* om vinteren. | revne |
| 4 | TM | 057 | Det frie *preget* rir om våren. | |
| 4 | PH | 058 | Den friske fuglen *rår* sjelden. | |
| 4 | PH | 059 | Den fulle tavla *rører* til frokost. | |
| 4 | AF | 060 | Den første dokken *rygger* til middag. | |
| 5 | BM | 061 | Blide tanker *firer* ukentlig. | |
| 5 | BM | 062 | Ekte sjeler *fisker* tregt. | |
| 5 | TA | 063 | Falske sjefer *flirer* snart. | sjefe |
| 5 | SH | 064 | Glemske *seler* føler sjelden. | |
| 5 | SH | 065 | Gode piller *grubler* overalt. | |
| 5 | KU | 066 | Kalde paver *husker* ofte. | |
| 5 | KU | 067 | Mørke maver *hvisker* nå. | |
| 5 | KH | 068 | Pene *kjeler* hører nødig. | |
| 5 | TM | 069 | Rare *kjeder* jubler minst. | |
| 5 | PH | 070 | Raske *haver* kjører mest. | |
| 5 | AF | 071 | Slemme gaver *snakker* gjerne. | gava |
| 5 | KH | 072 | Store filler *snubler* bestandig. | |
| 5 | PH | 073 | Svarte feler *sparker* best. | |
| 5 | AF | 074 | Vakre *deler* tror alltid. | dela |
| 5 | TM | 075 | Våkne biller *vokser* aldri. | |
| 6 | AF | 076 | Til middag rugger været en *tram*. | |
| 6 | BM | 077 | Til frokost lugger tjæren *nitten* riller. | |
| 6 | SH | 078 | Sjelden rekker sjelen *sju* prammer. | |
| 6 | SH | 079 | Om våren skriver paven *tjue* piller. | |
| 6 | BM | 080 | Om vinteren lekker maven *førti* pass. | |
| 6 | TA | 081 | Om sommeren leser læreren *søtti* måker. | læs |
| 6 | KU | 082 | Om høsten *roser* listen en masse. | |
| 6 | TM | 083 | Ofte loser kjæresten et *løft*. | |

| | | | | |
|---|---|---|---|---|
| 6 | KU | 084 | Neste uke lager kjolen et *loft*. | |
| 6 | PH | 085 | I morgen *sager* kisten et lodd. | |
| 6 | KH | 086 | I dag nager havet *nitti* lister. | |
| 6 | PH | 087 | Helst lønner *gaven* et ledd. | |
| 6 | TA | 088 | Denne uka skjønner fienden en *gran*. | skjønne ei |
| 6 | TM | 089 | Alltid lapper fela en *dram*. | |
| 6 | AF | 090 | Aldri lader faren en *brann*. | lade |
| 7 | TA | 091 | Våkne briller *brer* ukentlig. | brilla |
| 7 | BM | 092 | Vakre sjeler *trer* tregt. | |
| 7 | SH | 093 | Svarte sjefer *grer* snart. | |
| 7 | SH | 094 | Store seler *driller* sjelden. | |
| 7 | BM | 095 | Slemme piller *brister* overalt. | |
| 7 | KU | 096 | Raske paver *triller* ofte. | |
| 7 | KU | 097 | Rare maver *griller* nå. | |
| 7 | AF | 098 | Pene *kjeler* gror nødig. | kjela |
| 7 | KH | 099 | Mørke kjeder *tror* minst. | |
| 7 | KU | 100 | Kalde haver *graver* mest. | |
| 7 | AF | 101 | Gode gaver *traver* gjerne. | gava trave |
| 7 | PH | 102 | Glemske filler *skriker* bestandig. | |
| 7 | TM | 103 | Falske feler *spriker* best. | |
| 7 | TA | 104 | Ekte deler *fletter* alltid. | dela flette |
| 7 | TM | 105 | Blide biller *sletter* aldri. | |
| 8 | BM | 106 | Blide kjeler *sprekker* aldri. | |
| 8 | TA | 107 | Ekte kjeder *spreller* alltid. | kjeda sprelle |
| 8 | SH | 108 | Falske haver *spretter* best. | |
| 8 | SH | 109 | Glemske *gaver* skvetter bestandig. | |
| 8 | BM | 110 | Gode filler *sprader* fort. | |
| 8 | KU | 111 | Kalde feler *spraker* høflig. | |
| 8 | KU | 112 | Mørke *deler* skraper i dag. | |
| 8 | AF | 113 | Pene *biller* skryter minst. | billa skryt |
| 8 | KH | 114 | Rare briller *skriver* månedlig. | |
| 8 | AF | 115 | Raske *sjeler* skratter overalt. | sjela skratte |
| 8 | TM | 116 | Slemme *sjefer* skrubber plutselig. | |
| 8 | PH | 117 | Store *seler* skreller raskt. | |
| 8 | PH | 118 | Svarte piller *skritter* sakte. | |
| 8 | TA | 119 | Vakre paver *nøler* sikkert. | pava nøle |
| 8 | TM | 120 | Våkne *maver* brøler umiddelbart. | |
| 9 | TA | 121 | Om høsten *lader* pakken et skip. | lade |
| 9 | BM | 122 | Helst *leder* fetteren en hånd. | hand |
| 9 | SH | 123 | Om kvelden lider sykkelen en *stjerne*. | |
| 9 | SH | 124 | Om natten *rugger* tjeneren en skralle. | |
| 9 | AF | 125 | Alltid rydder *høsten* en skramme. | rydde |
| 9 | KU | 126 | Ofte skyver *hesten* en sjel. | |
| 9 | KU | 127 | Fort skyter *hosten* en kjede. | |
| 9 | AF | 128 | Snart spisser *hasten* en kjele. | spisse |
| 9 | KH | 129 | Sikkert vrikker hasen en *tang*. | |
| 9 | KH | 130 | Trolig vraker haren en *gang*. | |
| 9 | PH | 131 | Flittig *vrenger* hanen et tak. | |
| 9 | PH | 132 | Ivrig *sprenger* haken en prikk. | |

| | | | | |
|---|---|---|---|---|
| 9 | TM | 133 | Tregt *slenger* hunden en purk. | |
| 9 | TM | 134 | Begeistret *trenger* huden en prest. | |
| 9 | TA | 135 | Nesten stikker hånden en *klage*. | stikk ei |
| 10 | AF | 136 | Med *land* skjenner du fort. | |
| 10 | BM | 137 | Uten *tann* skinner luften sjelden. | |
| 10 | SH | 138 | Med lokk *klipper* kona bedre. | |
| 10 | SH | 139 | På sand *rager* helten dårlig. | |
| 10 | BM | 140 | Uten *rand* glipper lykten minst. | |
| 10 | AF | 141 | Uten lakk *svikter* lukten farlig. | |
| 10 | TA | 142 | På jorda *lyder* ferskenen lykkelig. | lyd |
| 10 | PH | 143 | Den halve pillen *rager* ofte. | |
| 10 | KU | 144 | Den kalde nesen *brenner* presist. | |
| 10 | KH | 145 | Mindre haker *fletter* trolig. | |
| 10 | KH | 146 | Sjenerte haner *sletter* farlig. | |
| 10 | KU | 147 | Spisse *harer* brister dristig. | |
| 10 | TA | 148 | Flate haser *glipper* ivrig. | hase glipp |
| 10 | TM | 149 | Hvorfor *kjøler* bilen sprekker? | |
| 10 | TM | 150 | Hvordan søler gutten *glass*? | |

## C. Forced choice test

Sentence material plus answering alternatives as presented in orthographic form in the *forced choice test*. For reasons of space, only the first sentence (BM 001) is shown as presented on the computer screen during the test. The spoken utterances in this test were the same as those used in the *sentence test* and the *word test* listed in Appendix A.

BM 001 Bilen vet at barna ville _____

hale
helle
holde
hyle

| | | |
|---|---|---|
| BM 002 | Boka tror at billene vil _____ | \helle\hulle\hyle\hylle |
| TA 003 | De tenker at _____ tør holde | \baren\broen\broren\troen |
| SH 004 | Dere skriker at fillene torde _____ | \helle\holde\hulle\hylle |
| SH 005 | Du sier at flyet skulle _____ | \holde\hulle\hyle\hylle |
| AF 006 | Faren ser at frukten skal _____ | \helle\holde\hulle\hylle |
| KU 007 | Fniset røper at heftet måtte _____ | \kile\hyle\kyle\kjæle |
| KH 008 | Han roper at hodet må _____ | \bære\kjæle\lære\skjære |
| PH 009 | Flisen påstår at pillene kunne _____ | \hyle\kile\kyle\kjæle |
| PH 010 | Hunden mener at rillene kan _____ | \lære\kjæle\skjære\bære |
| TA 011 | Jeg lærer at selen kan _____ | \be\se\sy\så |
| KH 012 | Loven hvisker at sjefen får _____ | \lære\kjæle\skjære\bære |
| TM 013 | Mora husker at sjelen fikk _____ | \by\se\sy\så |
| AF 014 | Posten frykter at tanten bør _____ | \gå\se\sy\så |
| TM 015 | Vi forteller at veien burde _____ | \banne\lande\vinne\vanne |
| BM 016 | Når truer avisen _____ | \deler\feler\kjeler\sjeler |
| BM 017 | Hvorfor treffer bilen _____ | \byer\sjøer\skjeer\skyer |
| TA 018 | Når spiser blomstene _____ | \feer\sjøer\skjeer\skyer |
| TA 019 | Hvor ser de _____ | \ni\sju\ski\ti |
| SH 020 | Når mottar dere _____ | \sjøer\skjeer\skyer\løer |
| KU 021 | Hvorfor _____ du sju | \haler\maler\mater\saler |
| KU 022 | Hvor løfter han _____ | \deler\feler\kjeler\sjeler |
| KH 023 | Når lurer hun _____ | \fanger\ganger\sanger\senger |
| KH 024 | Hvor krever hunden _____ | \gaver\haver\maver\paver |
| AF 025 | Når kjører huset _____ | \lakk\lam\land\lass |
| PH 026 | Hvorfor gjetter jeg _____ | \lakk\lam\land\lass |
| PH 027 | Hvordan føler mormor _____ | \deler\feler\kjeler\sjeler |
| TM 028 | Hvordan frykter snøen _____ | \gaver\paver\haver\maver |
| KH 029 | Hvordan forteller tankene _____ | \deler\feler\kjeler\sjeler |
| TM 030 | Når elsker vi _____ | \bad\bar\barn\garn |
| BM 031 | Aldri synger faren en _____ | \bille\fille\pille\rille |
| BM 032 | Alltid spiser fela en _____ | \bille\fille\pille\rille |
| SH 033 | Denne uka spiller fienden et _____ | \glass\lass\ras\pass |
| SH 034 | Helst skriver gaven en _____ | \måke\måne\måte\tåke |
| TA 035 | Idag ser havet en _____ | \kasse\mappe\masse\matte |
| AF 036 | I morgen leser kisten en _____ | \kiste\lisse\liste\niste |
| KU 037 | Neste uke kjører kjolen en _____ | \labb\lapp\lakk\last |

| | | |
|---|---|---|
| TA 038 | Ofte hører kjæresten en _____ | \kul\kål\nål\ål |
| PH 039 | Om høsten hilser listen en _____ | \kiste\liste\niste\skisse |
| KH 040 | Om sommeren har læreren en _____ | \kanne\kappe\kasse\masse |
| AF 041 | Om vinteren får magen en _____ | \kagge\kanne\kappe\kasse |
| PH 042 | Om våren føler _____ en handel | \gaven\haven\maven\paven |
| TM 043 | Sjelden fanger sjelen en _____ | \bille\fille\pille\rille |
| TM 044 | Til frokost elsker tjæren en _____ | \bille\fille\pille\rille |
| KH 045 | Til middag bærer været en _____ | \ball\pall\hall\stall |
| BM 046 | Den faste _____ synger aldri | \bilen\filen\pilen\silen |
| BM 047 | Det feige skjerfet _____ alltid | \fyller\hyller\skjeller\skyller |
| AF 048 | Det gale skjegget _____ denne uka | \raser\reiser\riser\roser |
| SH 049 | Det ferske loddet _____ best | \raser\reiser\riser\roser |
| SH 050 | Den fete _____ reiser idag | \bilen\filen\pilen\silen |
| KH 051 | Den fine øya _____ i morgen | \rager\raker\raller\raser |
| KU 052 | Den fjerde lyden _____ neste uke | \rager\raker\raller\raser |
| TA 053 | Det fjerne skinnet _____ ofte | \faller\raker\raller\raser |
| KH 054 | Den flate snøen _____ om høsten | \lekker\rekker\renner\revner |
| KU 055 | Det flaue skipet _____ om sommeren | \hekker\rekker\renner\revner |
| TA 056 | Den fremre tanken_____om vinteren | \rekker\renner\retter\revner |
| TM 057 | Det frie preget _____ om våren | \frir\glir\rir\rår |
| PH 058 | Den friske fuglen _____ sjelden | \rir\rår\ror\trår |
| PH 059 | Den fulle tavla _____ til frokost | \hører\kjører\røper\rører |
| AF 060 | Den første dokken _____ til middag | \bygger\rigger\rygger\tygger |
| BM 061 | Blide tanker _____ ukentlig | \firer\flirer\forer\fyrer |
| BM 062 | Ekte _____ fisker tregt | \feler\kjeler\seler\sjeler |
| TA 063 | Falske sjefer _____ snart | \firer\flirer\forer\fyrer |
| SH 064 | Glemske _____ føler sjelden | \deler\feler\seler\sjeler |
| SH 065 | Gode _____ grubler overalt | \biller\filler\piller\riller |
| KU 066 | Kalde _____ husker ofte | \gaver\haver\maver\paver |
| KU 067 | Mørke _____ hvisker nå | \gaver\haver\maver\paver |
| KH 068 | Pene _____ hører nødig | \deler\feler\kjeder\kjeler |
| TM 069 | Rare _____ jubler minst | \kjeder\kjeler\kjever\reder |
| PH 070 | Raske _____ kjører mest | \gaver\haver\maver \paver |
| AF 071 | Slemme _____ snakker gjerne | \gaver\haver\maver\paver |
| KH 072 | Store _____ snubler bestandig | \biller\filler\piller\riller |
| PH 073 | Svarte _____ sparker best | \deler\feler\seler\sjeler |
| AF 074 | Vakre _____ tror alltid | \deler\feer\seler\sjeler |
| TM 075 | Våkne _____ vokser aldri | \biller\filler\piller\riller |
| AF 076 | Til middag rugger været en _____ | \dram\pram\tram\trapp |
| BM 077 | Til frokost _____ tjæren nitten riller | \lugger\pugger\rugger\vugger |
| SH 078 | Sjelden rekker sjelen sju _____ | \drammer\prammer\trammer\skrammer |
| SH 079 | Om våren skriver paven tjue _____ | \biller\filler\piller\riller |
| BM 080 | Om vinteren _____ maven førti pass | \lekker\rekker\brekker\trekker |
| TA 081 | Om sommeren leser læreren søtti _____ | \måker\måner\måter\sokker |
| KU 082 | Om høsten _____ listen en masse | \koser\loser\moser\roser |
| TM 083 | Ofte _____ kjæresten et løft | \koser\loser\moser\roser |
| KU 084 | Neste uke _____ kjolen et loft | \jager\lager\nager\sager |
| PH 085 | I morgen _____ kisten et lodd | \jager\lager\nager\sager |
| KH 086 | I dag _____ havet nitti lister | \jager\lager\nager\sager |
| PH 087 | Helst lønner _____ et ledd | \gaven \haven\maven\paven |

| | | |
|---|---|---|
| TA 088 | Denne uka skjønner fienden en ____ | \gran\klan\kran\plan |
| TM 089 | Alltid _____ fela en dram | \lader\lager\lakker\lapper |
| AF 090 | Aldri _____ faren en brann | \lader\lager\nager\sager |
| TA 091 | Våkne briller _____ ukentlig | \brer\grer\rer\trer |
| BM 092 | Vakre sjeler _____ tregt | \brer\grer\rer\trer |
| SH 093 | Svarte sjefer _____ snart | \brer\grer\rer\trer |
| SH 094 | Store seler _____ sjelden | \diller\driller\griller\triller |
| BM 095 | Slemme _____ brister overalt | \biller\filler\piller\riller |
| KU 096 | Raske _____ triller ofte | \gaver\haver\maver\paver |
| KU 097 | Rare maver _____ nå | \diller\driller\griller\triller |
| AF 098 | Pene _____ gror nødig | \deler\feler\kjeler\sjeler |
| KH 099 | Mørke kjeder _____ minst | \gror\ror\trer\tror |
| KU 100 | Kalde _____ graver mest | \gaver\haver\maver\paver |
| AF 101 | Gode _____ traver gjerne | \gaver\haver\maver\paver |
| PH 102 | Glemske _____ skriker bestandig | \biller\filler\piller\riller |
| TM 103 | Falske _____ spriker best | \deler\feler\kjeler\sjeler |
| TA 104 | Ekte _____ fletter alltid | \deler\feler\kjeler\sjeler |
| TM 105 | Blide biller _____ aldri | \fletter\letter\sletter\slutter |
| BM 106 | Blide kjeler _____ aldri | \skvetter\sprekker\spreller\spretter |
| TA 107 | Ekte kjeder _____ alltid | \skvetter\sprekker\spretter\spreller |
| SH 108 | Falske _____ spretter best | \gaver\haver\maver\paver |
| SH 109 | Glemske _____ skvetter bestandig | \gaver\haver\maver\paver |
| BM 110 | Gode filler _____ fort | \skraper\skratter\sprader\spraker |
| KU 111 | Kalde feler _____ høflig | \spraker\sprader\skraper\skryter |
| KU 112 | Mørke _____ skraper i dag | \deler\feler\seler\sjeler |
| AF 113 | Pene _____ skryter minst | \biller\filler\piller\riller |
| KH 114 | Rare briller _____ månedlig | \river\skrider\skriker\skriver |
| AF 115 | Raske _____ skratter overalt | \feler\seler\sjefer\sjeler |
| TM 116 | Slemme _____ skrubber plutselig | \kjeler\seler\sjefer\sjeler |
| PH 117 | Store _____ skreller raskt | \deler \feler\seler\sjeler |
| PH 118 | Svarte _____ skritter sakte | \biller\filler\piller\riller |
| TA 119 | Vakre paver _____ sikkert | \brøler\føler\nøler\søler |
| TM 120 | Våkne _____ brøler umiddelbart | \gaver\haver\maver\paver |
| TA 121 | Om høsten _____ pakken et skip | \lader\lager\leder\lider |
| BM 122 | Helst _____ fetteren en hånd | \lader\leder\leger\lider |
| SH 123 | Om kvelden _____ sykkelen en stjerne | \lader\leder\lider\liker |
| SH 124 | Om natten _____ tjeneren en skralle | \lugger\pugger\rugger\ruller |
| AF 125 | Alltid rydder _____ en skramme | \hasten\hesten\hosten\høsten |
| KU 126 | Ofte skyver _____ en sjel | \hasten\hesten\hosten\høsten |
| KU 127 | Fort skyter _____ en kjede | \hasten\hesten\hosten\høsten |
| AF 128 | Snart spisser _____ en kjele | \hasten\hesten\hosten\høsten |
| KH 129 | Sikkert vrikker _____ en tang | \haken\hanen\haren\hasen |
| KH 130 | Trolig vraker _____ en gang | \haken\hanen\haren\hasen |
| PH 131 | Flittig vrenger _____ et tak | \haken\hanen\haren\hasen |
| PH 132 | Ivrig sprenger _____ en prikk | \haken\hanen\haren\hasen |
| TM 133 | Tregt _____ hunden en purk | \slenger\slynger\sprenger\trenger |
| TM 134 | Begeistret _____ huden en prest | \slenger\sprenger\trenger\vrenger |
| TA 135 | Nesten stikker hånden en _____ | \klage\klase\klasse\plage |
| AF 136 | Med _____ skjenner du fort | \lam\land\rand\tann |
| BM 137 | Uten _____ skinner luften sjelden | \land\rand\tang\tann |

| | | | |
|---|---|---|---|
| SH | 138 | Med lokk _____ kona bedre | \glipper\klapper\klipper\slipper |
| SH | 139 | På sand _____ helten dårlig | \rager\sager\raver\raker |
| BM | 140 | Uten _____ glipper lykten minst | \land\rand\rang\tann |
| AF | 141 | Uten lakk svikter _____ farlig | \bukten\fukten\lukten\lykten |
| TA | 142 | På _____ lyder ferskenen lykkelig | \bordet\gjerdet\hjula\jorda |
| PH | 143 | Den halve pillen _____ ofte | \rager\raker\raver\sager |
| KU | 144 | Den kalde _____ brenner presist | \nesen\nepen\neven\nisen |
| KH | 145 | Mindre haker _____ trolig | \fletter\gjetter\letter\sletter |
| KH | 146 | Sjenerte _____ sletter farlig | \haker\haner\harer\haser |
| KU | 147 | Spisse _____ brister dristig | \haker\haner\harer\haser |
| TA | 148 | Flate _____ glipper ivrig | \haker\haner\harer\haser |
| TM | 149 | Hvorfor _____ bilen sprekker | \føler\kjøler\kjøper\søler |
| TM | 150 | Hvordan _____ gutten glass | \føler\kjøler\siler\søler |

## D. Conventions used in the phonetic transcription

The SAMPA (Speech Assessment Methods Phonetic Alphabet) is phonetic alphabet is used [ 1 ].

### Classification of vowels and diphthongs

| phoneme | place | opening | rounding | quantity | example |
|---|---|---|---|---|---|
| i: | front | close | unrounded | long | vin |
| i. | front | close | unrounded | short | vind |
| y: | front | close | rounded | long | lyn |
| y. | front | close | rounded | short | lynne |
| e: | front | close-mid | unrounded | long | sen |
| e. | front | close-mid | unrounded | short | send |
| 2: | front | close-mid | rounded | long | søt |
| 2. | front | close-mid | rounded | short | søtt |
| {: | front | open-mid | unrounded | long | vær |
| {. | front | open-mid | unrounded | short | vært |
| A: | back | open | unrounded | long | hat |
| A. | back | open | unrounded | short | hatt |
| }: | mid | close | rounded | long | lun |
| }. | mid | close | rounded | short | bukk |
| u: | back | close | rounded | long | bok |
| u. | back | close | rounded | short | bukk |
| O: | back | close-mid | rounded | long | våt |
| O. | back | close-mid | rounded | short | vått |
| @. | the neural vowel "schwa" | | | | (kjel)e |
| {i | diphthong | | | | vei |
| Ai | diphthong | | | | kai |
| Oy | diphthong | | | | konvoy |
| A} | diphthong | | | | sau |
| }i | diphthong | | | | hui |
| ui | diphthong | | | | hoi |
| 2y | diphthong | | | | høy |

### Classification of consonants

| phoneme | manner | place | voicing | example |
|---|---|---|---|---|
| p. | plosive | bilabial | unvoiced | hopp |
| b. | plosive | bilabial | voiced | labb |
| t. | plosive | alveolar | unvoiced | lat |
| d. | plosive | alveolar | voiced | ladd |
| k. | plosive | velar | unvoiced | takk |
| g. | plosive | velar | voiced | tagg |
| m. | nasal | bilabial | | lam |
| n. | nasal | alveolar | | vann |
| =n | nasal | alveolar | | hesten |
| J. | nasal | palatal | | brann[3] |
| rt | plosive | retroflex | unvoiced | hardt |
| rd | plosive | retroflex | voiced | verdi |

[3] Typical in Trøndelag.

| | | | | |
|---|---|---|---|---|
| rn | nasal | retroflex | | garn |
| N. | nasal | velar | | sang |
| r. | liquid | alveolar | | prøv |
| R. | fricative | velar | | graver[4] |
| f. | fricative | labiodental | unvoiced | fin |
| v. | fricative | labiodental | voiced | vin |
| s. | fricative | alveolar | | lass |
| S. | fricative | postalveolar | | skyt |
| C. | fricative | palatal | unvoiced | kino |
| j. | fricative | palatal | voiced | gi |
| h. | fricative | glottal | | ha |
| l. | liquid | alveolar | | fall |
| L. | liquid | palatal | | ball[5] |

**Transcription conventions**

| **Phoneme** | **Explanation** | **Example** | **Transcription** | **Typical area** |
|---|---|---|---|---|
| =n | Syllabic n. | hesten | [h.e.s.t.=n] | |
| J. | Palatal n. | brann | [b.r.A.J.] | Trøndelag. |
| rn | Retroflex nasal. | barna | [b.A.rnA.] | |
| N. | Velar nasal. | tang | t.A.N. | |
| S. | Postalveolar frictative | sjel | S.e:l. | |
| C. | palatal fricative | kjele | C.e:l.@. | |
| L. | palatal l | ball | b.A.L. | Trøndelag |
| R. | Velar/uvular friktativ (r) | graver | g.R.A:v.@.r. | Bergen |
| @. | schwa<br>Neural vowel | kjele | C.e:l.@. | |

In the evaluation of the results, the character "\" (backslash) has been used as a template for empty slots and a "**.**" (period sign) was included in the one-character phonemes to give all phonemes the same number of symbols, i.e. two. This was done to ease the semi-automatic production of the confusion matrices.

[4] Typical in Bergen.
[5] Typical in Trøndelag.

## E. Characterisation of the hearing ability of the subjects

The characterisation of the hearing ability of the individual subjects is based on the mean error rates in the *word test*. In the table below this is shown for the subjects divided into the tree three listener groups REF<55, UAH>55 and AHA (see Chapter 7, p9). The results allow absolute comparisons and can therefore be used as "speech audiograms" for individual listeners.

| REF<55 | | | UAH>55 | | | AH | | |
|---|---|---|---|---|---|---|---|---|
| **Subject** | **Error (%)** | **n** | **Subject** | **Error (%)** | **n** | **Subject** | **Error (%)** | **n** |
| **4** | 6,3 | 32 | **23** | 46,7 | 75 | **1** | 35,9 | 39 |
| **11** | 10,7 | 28 | **27** | 73,0 | 37 | **2** | 68,3 | 60 |
| **12** | 14,3 | 42 | **29** | 35,1 | 77 | **3** | 41,0 | 61 |
| **21** | 11,1 | 36 | **31** | 12,8 | 47 | **5** | 32,8 | 64 |
| **22** | 7,9 | 38 | **32** | 12,5 | 40 | **6** | 29,3 | 75 |
| **24** | 2,6 | 38 | **33** | 18,2 | 33 | **7** | 33,3 | 54 |
| **25** | 13,9 | 36 | **34** | 23,7 | 38 | **8** | 19,6 | 46 |
| | | | **35** | 47,8 | 67 | **9** | 49,3 | 71 |
| | | | **36** | 11,3 | 53 | **10** | 38,5 | 52 |
| | | | **37** | 21,7 | 46 | **13** | 40,5 | 84 |
| | | | **38** | 26,4 | 53 | **14** | 61,1 | 54 |
| | | | **39** | 13,5 | 52 | **15** | 22,4 | 58 |
| | | | **40** | 22,2 | 54 | **16** | 31,9 | 72 |
| | | | **42** | 22,9 | 48 | **17** | 100,0 | 7 |
| | | | **43** | 8,2 | 49 | **18** | 47,7 | 65 |
| | | | **44** | 0,0 | 39 | **19** | 40,3 | 77 |
| | | | **46** | 22,1 | 68 | **20** | 30,0 | 70 |
| | | | **47** | 22,8 | 57 | **26** | 40,0 | 70 |
| | | | **48** | 17,9 | 67 | **28** | 27,4 | 62 |
| | | | **49** | 9,3 | 54 | **30** | 45,7 | 70 |
| | | | | | | **45** | 34,5 | 87 |

In addition to this data, for some of the subjects either a pure tone audiometer or a speech audiogram recorded by an audiologist, or both, were available. These are included below.

The hearing threshold found by pure tone audiometry without hearing aid:

| Subject | Ear L/R | Ferquency, Hz | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 250 | 500 | 1000 | 2000 | 3000 | 4000 | 6000 | 8000 |
| 2 | L | 65 | 80 | 85 | 100 | 90 | 95 | 95 | 105 |
| | R | 50 | 65 | 85 | 80 | 80 | 85 | 95 | 100 |
| 5 | L | 70 | 95 | 100 | 80 | 105 | 110 | 115 | — |
| | R | 30 | 30 | 50 | 50 | 55 | 75 | 80 | 85 |
| 13 | L | 45 | 70 | 85 | 95 | 95 | 90 | 110 | 100 |
| | R | 40 | 70 | 70 | 75 | 85 | 80 | 110 | 105 |
| 15 | L | 45 | 45 | 50 | 50 | 35 | 45 | 55 | 55 |
| | R | 20 | 40 | 25 | 35 | 50 | 55 | 40 | 75 |
| 16 | L | 45 | 50 | 90 | 95 | 95 | 110 | 105 | 85 |
| | R | 55 | 55 | 50 | 50 | 65 | 70 | 85 | 100 |
| 34 | L | 20 | 30 | 40 | 40 | 55 | 85 | 90 | 85 |
| | R | 25 | 30 | 40 | 20 | 30 | 50 | 85 | 75 |
| 36 | L | 5 | 5 | 5 | 5 | 5 | 20 | 50 | 70 |
| | R | 5 | 0 | 0 | 10 | 5 | 10 | 35 | 65 |
| 37 | L | 35 | 35 | 30 | 35 | 35 | 35 | 60 | 70 |
| | R | 50 | 55 | 50 | 30 | 20 | 25 | 35 | 50 |
| 38 | L | 0 | 0 | 10 | 10 | 45 | 55 | 65 | 70 |
| | R | 0 | 5 | 5 | 10 | 15 | 50 | 75 | 75 |
| 39 | L | 10 | 10 | 10 | 45 | 65 | 55 | 60 | 65 |
| | R | 10 | 5 | 0 | 20 | 5 | 5 | 30 | 50 |
| 40 | L | 25 | 10 | 35 | 40 | 50 | 45 | 60 | 75 |
| | R | 5 | 15 | 25 | 35 | 35 | 30 | 40 | 55 |
| 42 | L | 5 | 5 | 0 | 0 | 30 | 25 | 50 | 65 |
| | R | 0 | 0 | 0 | 5 | 10 | 15 | 35 | 70 |
| 43 | L | 5 | 0 | 0 | 0 | 5 | 15 | 10 | 35 |
| | R | 0 | 5 | 5 | 20 | 20 | 30 | 20 | 60 |
| 44 | L | 15 | 15 | 5 | 30 | 30 | 40 | 35 | 70 |
| | R | 20 | 25 | 15 | 15 | 25 | 20 | 30 | 35 |
| 45 | L | 25 | 20 | 10 | 35 | 65 | 65 | 75 | 75 |
| | R | 5 | 10 | 5 | 0 | 55 | 60 | 60 | 65 |
| 46 | L | 5 | 20 | 5 | 0 | 10 | 5 | 30 | 20 |
| | R | 5 | 5 | 0 | 5 | 5 | 0 | 25 | 10 |
| 47 | L | 10 | 5 | 5 | 25 | 40 | 40 | 45 | 80 |
| | R | 0 | 0 | 10 | 10 | 40 | 40 | 40 | 65 |
| 48 | L | 0 | 5 | 10 | 5 | 10 | 50 | 70 | 75 |
| | R | 15 | 10 | 0 | 10 | 15 | 40 | 80 | 80 |
| 49 | L | 5 | 5 | 5 | 10 | 25 | 35 | 45 | 60 |
| | R | 10 | 5 | 5 | 5 | 5 | 10 | 30 | 60 |

The speech level for 50 % recognition in speech audiometry without hearing aid.

| Subject | Speech level in dB SPL | | Subject | Speech level in dB SPL | |
|---|---|---|---|---|---|
| | Right ear | Left ear | | Right ear | Left ear |
| 2 | 85 | — | 34 | 60 | 80 |
| 15 | 35 | 40 | 37 | 40 | 30 |

## F. Presentation at NAS October 27th, 2001

The paper included below was presented at NAS (Norsk Akustisk Selskap. The Norwegian Acoustical Society). høstmøte 27 oktober 2001 in Trondheim (in Norwegian), [ 10 ].

# Algoritmer for høreapparat

FOREDRAG PÅ NAS HØSTMØTE 27 OKTOBER 2001

Sverre Stensby, SINTEF. Asbjørn Krokstad og Wim van Dommelen, NTNU

## Innledning

I dette foredraget presenteres prosjektet "Algoritmer for høreapparat".

Funksjonen til tradisjonelle høreapparat er å forsterke lyd slik at talespektret kommer over høreterskelen til den hørselshemmede. Det er vel kjent at dette alene ikke gir så god oppfattbarhet av tale som ønsket. Spesielt er problemet stort hos personer med nevrogene hørselstap.

Den store prosesseringskapasiteten i dagens og fremtidens digitale høreapparat gir mulighet for avansert signalprosessering. Dette åpner for forbedret individuell tilpassing av talesignalet til den hørselsrest som disse personene har. Detaljene i aktuelle algoritmer er imidlertid ikke tilstrekkelig kjent.

## Prosjektformål

Hensikten med prosjektet er å bidra til utviklingen av det faglige grunnlaget for konstruksjon av effektive algoritmer for bedret taleforståelighet i digitale høreapparater.

Prosjektet bearbeides langs to linjer, lytteprøver og utvikling av algoritmer. Hensikten med lytteprøvene er å diagnostisere hvilke fonetiske trekk i tale som er vanskeligst å oppfatte. Ut fra denne diagnosen kan det så utvikles algoritmer som modifiserer talesignalet slik at det vil bli lettere å forstå for hørselshemmede. I denne presentasjonen legges hovedvekten på lytteprøvene.

Et sekundært mål med prosjektet er at erfaringene skal komme til nytte i taleopplæring av helsepersonell, nyhetsopplesere, foredragsholdere og andre.

## Lytteprøver

Hensikten med lytteprøvene er å diagnostisere hvilke fonetiske trekk i tale som er vanskeligst å oppfatte i dagligdagse situasjoner. Med "dagligdagse situasjoner" menes lytting til radio og TV, samtale, foredrag mm. Metodisk er det vanskelig å kombinere "dagligdagse" situasjoner med et systematisk opplegg som kan gi kvantitative resultater. For å ivareta disse til dels motstridende hensyn ble det utviklet et opplegg med spesielt oppbygde setninger som ble presentert via høyttaler.

## Setningsmateriale

Setninger har den fordelen at de er mer "naturlige" enn ord sagt isolert. Men vanlige setninger har den ulempen at de inneholder redundans slik at enkelte av ordene kan gjettes ut fra sammenhengen. Det er derfor utviklet et sett med setninger med en semantikk som gjør det vanskelig å gjette hvilke ord som forekommer, men der den syntaktiske strukturen er korrekt. Alle ordene er vanlige norske ord. Setningen "*Flisen påstår at pillene kunne kyle.*" fra materialet illustrerer prinsippet. Metoden benyttes ofte i lytteprøver og benevnes "Semantically Unpredictable Sentences" (SUS).

Språklyder kan grupperes i minimale par. I setningene er det også lagt vekt på å avdekke feil i slike tilfelle. Det er derfor valgt ord fra grupper av minimale par. I setningene "*Raske paver triller* *ofte.", "Rare maver griller nå.", "Gode gaver traver gjerne."* og *"Kalde haver graver mest."* kan det testes på opposisjonen mellom */p/, /m/, /g/* og */h/* i *"paver", "maver", "gaver"* og *"haver"*.

Det ble utviklet 150 setninger. Disse ble lest av fire menn og fire kvinner med østnorsk, vestnorsk, trøndersk, eller nordnorsk dialektbakgrunn.

## Typer av lytteprøver

Lytteprøver har best diskriminerende effekt når andelen av riktig og galt er i nærheten av 50-50. Et av virkemidlene for å oppnå ønsket andel av feil var å benytte tre ulike typer av lytteprøver dynamisk gjennom testen. Prøvetypene ble benevnt setningstest, ordtest og rimordtest.

Rimordtesten var den letteste testformen. Her fikk lytteren presentert det meste av setningen skriftlig. Et av ordene manglet, og det ble gitt fire svaralternativer for dette ordet. Et eksempel er

*Bilen vet at barna ville* ____ . */hale/helle/holde/hyle/*

Her skulle _____ erstattes med det svaralternativet som passet best med det som ble hørt.

Ordtesten var mer krevende. Også her fikk lytteren presentert det meste av setningen skriftlig og ett av ordene manglet. Men det ble ikke gitt svaralternativ.

Setningstesten var den mest krevende. Her skulle hele setningen oppfattes når den ble presentert. Det ble ikke gitt noen skriftlig støtte.

## Lyttere

Det har deltatt 48 lyttere i testen. Den største gruppen brukte høreapparat, men det deltok også normalthørende og personer som hadde hørselstap, men som ikke nyttet høreapparat. Lokallag av Hørselshemmedes Landsforbund i Sør-Trøndelag har vært behjelpelig med å sette oss i kontakt med aktuelle lyttere.

## Lytteprøver

Lytteprøvene ble foretatt i et konferanserom med dimensjonene 6x6x3 m. Det ble systematisk benyttet to avstander til høyttaleren som presenterte setningsmaterialet, henholdsvis 1.3 og 4.1 m.

Kontrollert støy ble presentert over andre høyttalere for å maskere talen. Som støy ble det benyttet såkalt "talekor" med tre ulike nivåer. Dette gir til sammen 6 ulike betingelser for virkningen av støy og rom. Disse betingelsene ble variert i løpet av testen sammen med testtype for å oppnå ønsket andel av feil for hver lytter.

Lytter satt sammen med forsøksleder. Foran seg hadde de en dataskjerm. Detaljene i presentasjonen ble styrt av en PC. Hver presentasjon hadde følgende faser:

- Forsøksleder initierer presentasjonen av neste setning. Presentasjonen styres så av PC:
- – Informasjon relevant for testtypen (setnings-, ord- eller rimordtest) vises på skjermen.
- – Støyen slås på.
- – Setningen presenteres.
- – Støyen slås av.
- Lytter dikterer svaret til forsøksleder.
- Forsøksleder taster inn svaret på tastaturet.
- Lytter stadfester at forsøksleder har skrevet svaret som ønsket.

Lyden fra hele sesjonen tas opp på bånd via kunsthode for senere analyse.

## Resultat av lytteprøven

Testmetoden og setningsmaterialet var utviklet spesielt for dette prosjektet. Et resultat er derfor at dette fungerte etter forutsetningene.

Mønstret i typer av feil var de samme hos normalthørende og hos de med hørselstap uavhengig om de i den siste gruppen brukte høreapparat eller ikke. Mange av problemtypene ser derfor ut til å være universelle, men de skaper størst problem hos de med hørselstap.

Innleserne med østnorsk dialekt ble forstått best. Dette var uavhengig av lytternes egen dialektbakgrunn. Vi har foreløpig ikke funnet noen endelig forklaring på dette. En mulighet er at de som leste østnorsk naturlig lå tettest opp mot skriftbildet og benyttet mest distinkt uttale.

Grensen mellom ordene ble stort sett oppfattet. Dette var et av resultatene fra setningstesten.

Konsonantlydene hadde større andel feil enn vokallydene, men det var også betydelig andel feil i vokallydene. Algoritmene bør derfor gi positive bidrag til oppfattelsen av alle kategorier av talelyder. Nasaler hadde størst andel av feil etterfulgt av plosiver, frikativer og likvider.

## Algoritmer

Ulike algoritmer har vært undersøkt i studentoppgaver. Algoritmer som tar utgangspunkt i henholdsvis tidsfunksjon, transformasjoner og binaural hørsel har vært behandlet. Algoritmene er lovende for videre utvikling.

## Taleopplæring

Resultatene fra prosjektet støtter opp om mange av de anerkjente anvisningene vedrørende det å tale til hørselshemmede, og de gir grunnlag for noen nye. Det viktigste er å bruke naturlig stemme og snakke litt sterkt, litt langsomt, og med øyekontakt rett til samtalepartneren. Lyder og ord som lett forveksles bør tillegges litt ekstra tydelighet.

## Bidragsytere

Prosjektet "Algoritmer for høreapparat" er finansiert med Extra-midler fra Helse og Rehabilitering.

Helse og Rehabilitering er en norsk stiftelse bestående av 21 landsomfattende helse- og rehabiliteringsorganisasjoner, som har konsesjon til å drive det landsomfattende TV-spillet EXTRA. Overskuddet fra spillet skal bidra til å øke de frivillige humanitære organisasjonenes innsats innenfor forebygging, rehabilitering og forskning i Norge. Alle frivillige humanitære/samfunnsnyttige organisasjoner og funksjonshemmedes interesseorganisasjoner i Norge kan søke om midler uavhengig av medlemskap i stiftelsen.

Søkerorganisasjon er Hørselshemmedes Landsforbund. Spesielt har lokallaget i Bergen engasjert seg i problematikken og vært pådriver for å få i gang dette prosjektet.

Prosjektet utføres ved Institutt for Teleteknikk, Fakultet for elektronikk og telekommunikasjon, NTNU. Arbeidet gjøres i nært samarbeid med Lingvistisk institutt ved NTNU og SINTEF Tele og data. En rekke studenter har gjort en stor innsats i prosjektet. Takk rettes også de som har lest inn setningsmaterialet, de som har deltatt som lyttere og til prosjektets rådgivende ekspertgruppe.

## Kontakt

Foredragsholderene har følgende elektroniske postadresser:
Asbjorn.Krokstad@tele.ntnu.no, Wim.van.Dommelen@hf.ntnu.no og
Sverre.Stensby@informatics.sintef.no