

Dept. for Speech, Music and Hearing
**Quarterly Progress and
Status Report**

**Perceptual analysis of
dysarthric speech in the
ENABL project**

Rosengren, E.

journal: TMH-QPSR
volume: 41
number: 1
year: 2000
pages: 013-018



**KTH Computer Science
and Communication**

<http://www.speech.kth.se/qpsr>

Perceptual analysis of dysarthric speech in the ENABL project

Elisabet Rosengren

Abstract

This paper presents the perceptual analysis of dysarthric speech recorded for use in the ENABL project. Twelve dysarthric speakers were tested with a Swedish dysarthria test that evaluates several speech functions; attention was focused on articulation and intelligibility. Articulation was tested with sentences representing different articulation types, and articulation deviation rated for each type. Word intelligibility was measured both in isolated words and in sentence context. The severity of the dysarthria was found to vary from very mild to moderate/severe. The results of the perceptual analysis were later put to use in other project tasks for investigating relationships between these perceptual findings, acoustic-phonetic data and results on speech recognition.

Introduction

The ENABL project has provided access by voice, via speech recognition, to vocational software for custom-driven product design and configuration (Bickley & Hunnicutt, this volume). By using such a speech user interface, persons with motoric disabilities can be enabled to enter or remain in a technical work environment. Because such persons may also have a speech impairment (dysarthria), one part of the project dealt with speech recognition of dysarthric speech. To this end, a number of persons with dysarthria participated in the project and this paper describes the perceptual evaluation of their speech.

Locating speakers

The subjects originally sought were twelve speakers with dysarthria of six types (flaccid, spastic, mixed, ataxic, hypokinetic and hyperkinetic) and two levels of severity, mild and moderate. When it was understood that the speech characteristics that needed to be considered for speech recognition for these types of dysarthria overlapped substantially, it was decided that a reclassification was in order. After discussions with other speech pathologists experienced in the dysarthria field, it was decided that a good alternative would be to find subjects with varying speech abnormalities. The population sample of dysarthric subjects could then be selected and characterised by distinctive deficits, or features, as opposed to categories of dysarthria. This manner of subject selection was considered more accurate as well as being able to ensure the procurement of the desired

spectrum of speech deficits with which to test the speech recognition. The following speech characteristics affecting speech recognition were discussed: frequency of pauses (inter and intra word), slow rate of speech, non-speech (involuntary) sounds, irregular breakdowns (inconsistency), low intensity and severity of articulation difficulties. A categorisation based on such speech characteristics was chosen when all the dysarthric speakers had been tested. This was regarded to better represent possible users of speech recognition.

More detailed rankings of severity of dysarthria than the two categories mild and moderate were considered and decided on in the project. A 7-point severity scale served as a basis for the ranking; the scale steps were: none, minimal, mild, moderate, moderate/severe, severe, profound. Having a more detailed classification of severity rather than just two discrete levels was of benefit when correlating the severity of dysarthria to the results on the speech recogniser.

Test persons

Of the twelve dysarthric speakers, five were men and seven women (Table 1). The age varied from 15 to 58 years. In three cases the dysarthria was caused by stroke, four persons had cerebral palsy, two had the diagnosis MS, one had hereditary ataxia, one had ataxia caused by a tumour and one had a congenital cerebellum damage. The severity of the dysarthria was found to vary from very mild to moderate/severe.

Table 1

| Test person | Sex | Age | Degree of dysarthria | Diagnosis |
|-------------|--------|-----|----------------------|-----------------|
| M1 | male | 54 | mild | stroke |
| F1 | female | 52 | very mild | MS |
| M2 | male | 45 | mild | hered ataxia |
| F2 | female | 32 | moderate | CP |
| F3 | female | 39 | mild | MS |
| M3 | male | 30 | mild | tumour/ataxia |
| F4 | female | 58 | moderate | stroke |
| M4 | male | 31 | mild/moderate | spastic CP |
| M5 | male | 51 | mild/moderate | stroke/ataxia |
| F5 | female | 15 | moderate | CP |
| F6 | female | 15 | moderate | cerebell damage |
| F7 | female | 15 | moderate/severe | CP |

Description of test persons

M1 was a 54-year-old male who had mild dysarthria after a stroke. He had hemiplegia and used a walker. He was an engineer who had worked with CAD/CAM systems and was very interested in our project. His respiratory and phonatory functions were good. His problems became apparent in the articulation of consonants that require precision. His spontaneous speech sounded somewhat slurred, but he was able to compensate and spoke more clearly with a slower rate. He was the only speaker who reached 100% intelligibility on the isolated words test.

F1 was a 52-year-old woman with MS and very mild dysarthria. Her ability to sustain the sounds /s/ and /a/ was reduced, which reveals a weakness of the respiratory muscles. Her voice was weak, breathy and hypernasal with hints of hyponasality. The speech rate was normal, but with frequent pauses and in short phrases. The articulation was flaccid, but she was able to pronounce all phonemes, including consonant clusters.

M2 was 45 years old and had a progressive hereditary form of ataxia. His voice had a certain resonance but was rather weak, and weakened fast during syllable repetition. The articulation was flaccid, especially when he spoke fast, and phonemes then were more indistinct or entirely omitted. He could not fully control the voice-voiceless distinction in plosives and his speech was somewhat nasal.

F2 was a 32-year-old woman with CP of an athetosis type and moderate dysarthria. In spite of her limited muscular capacity, her phonation in speech was functional and her voice had rather low pitch. Her articulation was slow and

laboured; she was the second slowest speaker. Although many types of consonants sounded distorted due to her difficulties with the articulatory muscles, they were seldom confused with another consonant, so the results on the intelligibility test were very high. However, she commented that she put in extra effort and usually did not speak that clearly.

F3 was 39 years old and had MS. She used a walker and had reduced hand function and mild dysarthria. Her spontaneous speech often sounded a little slurred; plosives could become fricatives, phonemes could be skipped over and unstressed front vowels could become neutral. She sometimes took a deep breath in preparation to get initial consonant clusters clear. She had met with situations where she had been accused of being drunk.

Test person M3, 30 years old, was diagnosed to have ataxic type of dysarthria of a mild degree since he was 15 years old. He had a slightly exaggerated articulation with very few articulation errors, although some phonemes sounded a little distorted and prolonged and his speech was slow due to difficulties in timing and rhythm. At times he showed somewhat deviant prosody.

F4 was a 58-year-old woman who had had a stroke several years ago. She used a wheelchair for mobility and had moderate dysarthria. Her articulation was slow and effortful. Her voice was rather resonant with adequate intensity. Her speech was slurred, especially the consonant clusters, but she often corrected herself; the speech then became even slower. Because of this awareness and ability to make the articulation more precise, the intelligibility was good in spite of the rather severely affected

articulatory muscles. In faster speech, phonemes and syllables could be lost.

Test person M4, a 31-year-old male, had cerebral palsy with mild to moderate degree of spastic dysarthria. He used a wheelchair and could not use his hands. His respiratory function was reduced, so he spoke in short phrases with frequent pauses. He had difficulties in articulating some sounds, especially consonant clusters. His speech was characterised by tenseness and imprecise articulation requiring effort.

Test person M5 was a 51-year-old male who had had a stroke in the cerebellum some months earlier. He could walk a few steps and was in a rehabilitation clinic. He had relatively mild dysarthria of an ataxic type. His ability to sustain /s/ and /a/ was within the normal range and he had no problems with the vocal intensity. However, he had some difficulties with the coordination of phonation, velum and articulators. This sometimes led to the perception of nasals as stops, or voiceless stops as voiced. The intelligibility was slightly reduced.

F5 was 15 years old, had CP and moderate dysarthria. Her voice was weak but she was able to increase the intensity. She had a high fundamental frequency and her voice was rather tense with frequent instances of vocal fry. Her speech was rather slow, but she sometimes rushed towards the end of a phrase due to lack of air. The speech could be somewhat inconsistent due to varying speed and clarity.

F6 was a 15-year-old girl with a cerebellar damage and a moderate hearing impairment. Her speech was slow and rather weak with a high pitch and with a "childish" prosody. She had difficulties in reading, which influenced her speech and intelligibility. The reading of the "hinge script" (see below) was reduced to six sentences.

F7 was a 15-year-old girl with spastic CP and dysarthria of a moderate/severe degree. Her voice was weak and the pitch was very high. She had difficulty with controlling and coordinating breathing, phonation and articulation because of her spasticity, which led to sudden changes in intensity and speed, short phrases and many pauses, sometimes even within words. According to her speech pathologist, F7 was able to use a deeper voice (lower pitch) when she was more relaxed, and this made her speech more intelligible. The speech also varied from day to day. Because of her rather severe speech impairment, the complicated sentences of the "hinge script" (see below) were excluded from the test material.

The Dysarthria Test

The twelve persons have been tested with the Swedish Dysarthria Test (Hartelius & Svensson, 1990). The test consists of six sections that evaluate respiration, phonation, oral motor function, articulation, prosody and intelligibility. To test intelligibility, a new beta version of the intelligibility section was used (Allemark & Lillvik, 1996). The new version consists of lists of 68 one- and two-syllable words that are randomly generated from a program that contains a lexicon with 1500 minimal pairs. The program also generates ten grammatically correct sentences with a nonsense content.

The test also includes reading of a short passage (92 words). In addition, the test persons also read ten sentences (called the "hinge script") that were relevant to this project. Finally there were a few minutes of spontaneous speech.

The subjects' performance on the dysarthria test was recorded on digital audio tape. The perceptual analysis of all speakers was carried out by investigators at KTH.

Attention was focused on the articulation and intelligibility sections, but a few tasks from other sections were also carried out.

Perceptual analysis

Respiration, phonation and oral motor function

The following tasks from the dysarthria test were selected in order to assess respiration, phonation and oral motor function; these tasks correlate best with the test in general and with the articulation and intelligibility sections.

The ability to sustain the fricative /s/ was the chosen measure in judging specific functions of the respiratory muscles that are important for speech. Normal values are between 10 and 40 seconds. The values for the test persons varied from normal to very reduced (<5 seconds).

Phonation was tested by asking the person to sustain the vowel /a/ as long as possible. The values for normal speakers vary between 6 and 30 seconds. Some dysarthric speakers had normal values while others had very low values (1-2 seconds).

The tasks that were chosen from the section that evaluates oral motor function were fast syllable repetitions. The test persons were asked to repeat pa, ta and ka at an even pace as fast as possible. Almost all test persons had lower values (3-4 per second) than normal (6-7 per second). Differences between the three syllables were also seen, depending on where the person

had the worst muscular problem; in the lips, the front of the tongue or the back of the tongue.

Articulation

In the articulation test, the test person reads seven sentences that are constructed to represent different articulation types; bilabials, labio-dentals, dentals, palatals, velars, consonant clusters and vowels. The deviation of the speech for each category is rated from 0 (normal) to 4 (very deviant or lacking phonemes). The highest score representing the most deviant speech (sum for the 7 types with maximum value 28) reached among the twelve speakers was 9.5 and the lowest, 0.5.

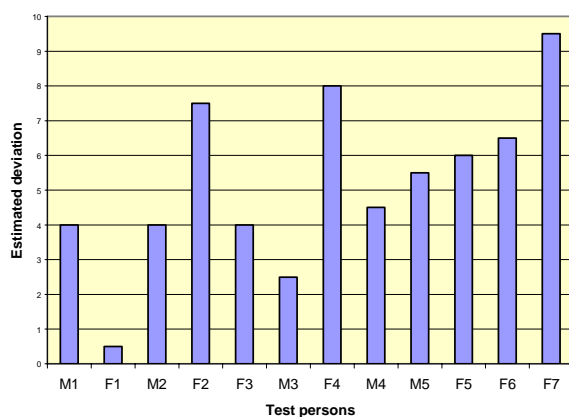


Figure 1. Articulation deviation scores for the 12 dysarthric speakers.

Intelligibility

Isolated words

In the first part of the intelligibility test, the test persons read a list of 68 one- and two-syllable isolated words. Each person had a different list that had been randomly generated by a computer program. The listener selected the perceived word out of five alternatives. The results varied from 100% intelligibility for one test person to 56% for the person with the most severe dysarthria (Figure 2). Seven persons had more than 90% intelligibility. Single word tests are considered to be insensitive regarding estimation of the intelligibility of mildly dysarthric speakers (Yorkston et al., 1988). Besides, for most dysarthric speakers, the pronunciation of single words is less demanding, enabling the speaker to compensate optimally. Most of our test persons were able to compensate considerably for their difficulties in the isolated word test and reached high values, while their spontaneous speech was quite different. The results of the isolated word test can therefore not

be seen as a measure of the “overall” intelligibility of the test persons, but is of value in revealing certain error types. A word that was incorrectly perceived by the listener had been assigned a code that represented a certain error type due to a certain articulation difficulty (see below).

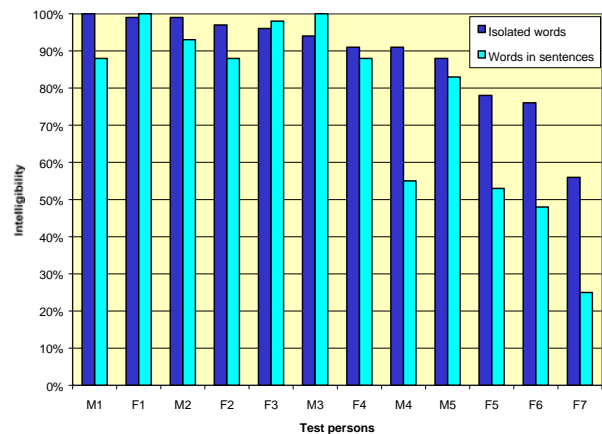


Figure 2. Word intelligibility in isolated words and in nonsense sentences for the 12 dysarthric speakers. The test persons are presented from left to right in order of their intelligibility.

Words in sentences

The other part of the intelligibility test measured word intelligibility in a sentence context. The test persons read 10 grammatically correct sentences with a nonsense content, e.g., “en hetsig alg tvättar en dans” (a passionate alga washes a dance). The listener transcribed orthographically the perceived words. The results in word intelligibility in a sentence context varied from 100% for two test persons to 25% intelligibility for the person with the most severe dysarthria (Figure 2).

Intelligibility in sentences was lower than for isolated words for most test persons. This could partially be explained by the more spontaneous and indistinct pronunciation of the words in the sentence context. The transcription task in the sentence test was also more difficult for the listener than the multiple-choice task in the isolated word test. The three persons with the highest intelligibility in the sentence part, however, had somewhat lower scores on the isolated words test.

A few test persons had some reading or vision problems. Additional examination of the intelligibility test was necessary in order to discover obvious reading mistakes. Such words were excluded from the list so that false listening results would not incorrectly reduce the intelligibility percentage. In spite of this it was

apparent that the vision and reading problems of a few test persons affected their reading in such a way that their dysarthria was worsened and the intelligibility thereby reduced.

Articulation errors: Results of speakers with lowest intelligibility

The following show some examples of the articulation errors that were made by the four test persons with scores under 90% (i.e. more than 6 errors) on the word intelligibility test.

The most common error types for the person with the most severe dysarthria (F7, 56% intelligibility) were the following: /r/ became /l/ (“rida” was perceived as “lida”), her inability to round the lips caused vowel errors (“rusa” became “rysa”), and consonant clusters lost one phoneme (“strå” became “stå”).

Test person F6 reached 76% intelligibility. Typical errors were of three kinds with three or four instances of each. /r/ became /j/ because of lack of precision and control of the tongue (bröd --> bjöd) and initial /h/ was not present (or the opposite case) due to lack of timing control of phonation (Haga --> aga, äng --> häng). There were also mistakes in voiced/voiceless stops due to deficient coordination of phonation and articulation (spädda --> spätta, para --> bara).

F5 had 78% intelligibility. The most common error (6 occasions) was that /r/ was perceived as /l/, which, as in the case of person F7, shows a lack of precision and control of the tongue.

Test person M5 had had a stroke in the cerebellum. He had 88% intelligibility. Typical errors were that initial /h/ was not perceived at all (“hatt” --> “att”) and that nasals were perceived as plosives (“ring” --> “rigg”). His errors indicate problems with coordination of the velum, phonatory and articulatory muscles.

Reading rate

The test persons also read a piece of text containing 92 words and their reading rate was calculated. The rate varied between 158 words per minute for the fastest speaker and 53 words per minute for the slowest. The test persons with dysarthria acquired in adulthood tended to read faster. The five speakers who had a reading rate over 100 words per minute had all acquired dysarthria. Their values were within the normal range or slightly below (values for normal speakers: 121 - 221).

The value for test person F6 was based only on the first half of the text, as she had to repeat the text after the speech pathologist because of her reading problems.

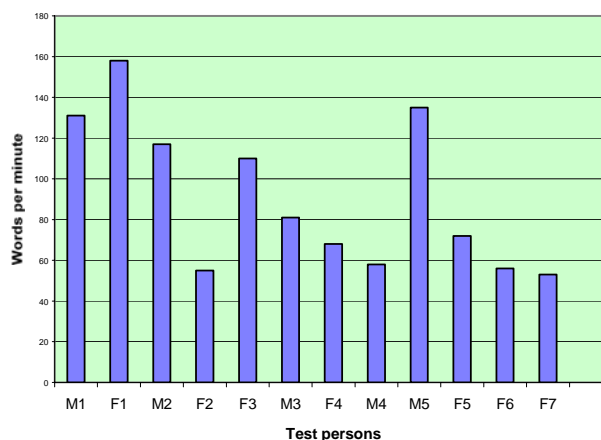


Figure 3. Reading rate for the 12 dysarthric speakers.

Frequency of pauses

The six slowest speakers (except for F6, see below), with a reading rate between 53 and 81 words per minute, were examined in order to determine the frequency of breathing pauses in their reading. The first two sentences (44 words) of the passage “Ett svårt fall” were chosen and the perceived pauses were calculated for each person. However, F6’s pauses could not be calculated in a reliable way, as she was repeating the text after the speech pathologist. Three persons with a rather continuous flow of the speech had 6, 7 and 10 pauses respectively. The three other speakers had 14, 15 and 18 pauses, respectively. They used another speech strategy to compensate for their insufficient respiratory function.

Spontaneous speech

In order to get a sample of spontaneous speech, the test persons were asked to describe their trip to KTH. Most of the test persons had no difficulty talking for a few minutes about that topic or about something else that was of interest to them. The spontaneous speech was taken into account when rating the overall intelligibility and describing the test persons’ speech.

Classification

The dysarthric speakers have been classified with reference to different speech characteristics as a result of the perceptual analysis.

Severity of dysarthria

The test persons have been ranked in order of severity of dysarthria, from the mildest to the most severe. The grades of dysarthria that are represented in the test group are very mild, mild, mild/moderate, moderate and moderate/severe. The test persons can also be arranged into two equally large groups; one group with moderate degrees of dysarthria (articulation deviation score above 5) and the other with milder dysarthria (articulation deviation score below 5). None of the test persons had severe or profound dysarthria. The ranking is mainly based on the deviation scores, but other speech material is also taken into account.

Speech rate

The speakers have also been ranked according to their reading rate, from the fastest to the slowest. They fall into two groups; relatively fast speakers (>100 words/minute) and slow speakers (<100 words/minute). There are five speakers belonging to the “fast” group; it is interesting to note that they have all acquired dysarthria.

Frequency of pauses

Six of the slow speakers were further examined with regard to the number of breathing pauses in their reading. Test person F6 could not be measured because of reading difficulties that made it necessary for her to repeat the passage in phrases after the speech pathologist. The

speakers have been divided into two groups; one group with relatively few pauses in the specified passage (three speakers), and the other group with many pauses.

Consequences for the project

These categorisations and rankings were used to investigate relationships between the perceptual findings, the acoustic-phonetic data and results on speech recognition (Magnuson & Blomberg, this volume; Talbot, this volume). Intelligibility results were also correlated with recognition results. In addition, the influence of factors such as involuntary sounds and inconsistencies were also assessed, although not many such occurrences were found among these speakers.

References

- Allemark E & Lillvik M (1996). Förståelighetstestning av dysartriska talare (Intelligibility testing of dysarthric speakers). *Senior thesis in speech-language pathology, University of Gothenburg*.
- Bickley C & Hunnicutt S (2000). ENABL – Enabler for Engineering Software using Language and Speech. *TMH-QPSR, KTH, 1/2000: 1-11* (this volume).
- Hartelius L & Svensson P (1990). *Dysartritest (Dysarthria Test)*, Psykologiförlaget AB, Hägersten.
- Magnuson T (2000). Acoustic analysis of dysarthric speech with some implications for automatic speech recognition. *TMH-QPSR, KTH, 1/2000: 19-30* (this volume).
- Talbot N (2000). Improving the Speech Recognition in the ENABL project. *TMH-QPSR, KTH, 1/2000: 31-38* (this volume).
- Yorkston KM, Beukelman DR, & Bell KR (1988). *Clinical Management of Dysarthric Speakers*. Boston: College-Hill Press.