Text-to-Speech Technology: A Survey of German Speech Synthesis Systems

Nicolas Léwy

Thomas Hornstein

UBILAB Technical Report 94.10.2

UBILAB (UBS Information Technology Laboratory)

UBILAB is the corporate information technology innovation center of Union Bank of Switzerland. It pursues a small number of attractive, highly competitive research projects with the aim of maintaining the status of a recognized research institution.
It is the task of UBILAB to actively assist UBS with its goal of becoming a leader in the mastery of modern IT equipment, techniques, and methods. This task is pursued by means of intimate involvement in application-oriented research and advanced development projects. Furthermore, UBS contributes to fostering the interaction between IT research and application, i.e. between academia in its search for new methods, and business in its application of them. To this purpose UBILAB cooperates with universities and other research institutions on a worldwide basis. The goal is to expand the scope of the Laboratory by carrying out common projects, thus guaranteeing – provided top-notch partners are found – the quality of the research. The impact of UBILAB is therefore focussed on both the IT departments of UBS Group and the IT research community at large.

For further information about UBILAB, staff, projects, or publications see our World Wide Web (WWW) node.

UBILAB offices and laboratories are situated in two locations. Please contact the staff personally via e-mail or write to the main office located at Lab Flur.

Lab Flur:

Enzianweg 4
8048 Zürich

Mailing address:

Union Bank of Switzerland
UBILAB
Bahnhofstrasse 45
CH-8021 Zurich, Switzerland

Phone: ++41 1 236 57 14
Fax: ++41 1 236 46 71
E-mail: *lastname*@ubilab.ubs.ch
WWW: http://www.ubilab.ubs.ch

Lab Universitätsstrasse:

Universitätsstrasse 84
8033 Zürich

Mailing address:

Union Bank of Switzerland
UBILAB
P.O. Box 2336
CH-8033 Zurich, Switzerland

Phone: ++41 1 234 21 47
Fax: ++41 1 234 65 18
E-mail: *lastname*@ubilab.ubs.ch
WWW: http://www.ubilab.ubs.ch

1 Introduction

The work described in this report is part of the QuickPhone project at UBILAB. The QuickPhone research project started in 1992. Its aim is to investigate new technologies for building state-of-the-art telephone information services.

So far, a prototype of a phone banking service has been built. This service incorporates the experiences gained in the domain of interactive voice response (IVR) systems during the last three years. Different aspects of the service are of interest and some of them can be classified as novel in comparison with existing ones. One important aspect is the user interface which makes use of an application-independent, list-based navigation technique. This is in contrast to the traditional hierarchical tree-based menu selection technique. The user interface also supports both voice and key input. Voice recognition vocabularies, including the digits and a set of application independent navigation commands, were trained in the three languages German, French and Italian on a homogeneous distributed Swiss population. A detailed description of the user interface aspects in the QuickPhone project can be found in [Hor94]. It has been shown that the voice input technology is ready for building easy-to-use and efficient telephone voice interfaces.

In order to apply IVR systems successfully in current and future banking environments additional aspects should be considered as well. The first is to gain more flexibility by using text-to-speech (TTS) as the voice output technology instead of, or in addition to, playing pre-recorded phrases, words and numbers as in most current systems. The second is to build more secure systems by using speaker identification and verification techniques for applications with restricted access, such as account information or transaction services. Both technologies are to be investigated in the QuickPhone project. The goal, again, is to show the usability and reliability of the technology for operational services at UBS.

As a first step toward TTS integration we launched a 13-week student project with the aim of investigating the state-of-the-art technology of the TTS systems currently available. Its results are described in this report.

In Europe, and especially in Switzerland, IVR systems must support multiple languages. Many TTS systems on the market are designed for different languages and different hardware and software environments. All of them have their merits and drawbacks in terms of speech quality, software and hardware requirements, and costs. Only a few of them, however, support multiple languages. For this project only systems supporting multiple languages were considered.

The goal of the project was to judge and compare different TTS systems according to the following criteria:

- synthesis quality (pronunciation rules, names, subjective evaluation)
- platform independence
- applicability for real world IVR systems

This report describes the results of the study of synthesis quality. Due to the restricted time of 13 weeks the number of TTS systems investigated had to be limited to four. During the project it was decided to focus the work on the synthesis quality in German. The procedures and set-ups for these tests could be applied to French and Italian as a part of a subsequent project.

TTS systems from Berkeley Speech Technologies (BeST), Elan Informatique, Infovox, and Lernout & Hauspie Speech Products were tested. The BeST system was purchased from Dialogic Corporation in Belgium. Despite the generous support of Dialogic, the synthesis quality could not be tested due to problems with the driver software. This will not prevent further investigations of the BeST system in the subsequent work.

In addition to the importance of the scope of the QuickPhone project this report also serves as feedback and is distributed to all four companies. The results presented do not claim to be complete nor are they absolute. The report should, however, provide an opportunity for eradicating some of the synthesis errors found during the tests. Not all aspects could be investigated in this short time. There are still a lot of questions to be answered and many tests to be performed. Nevertheless, the report provides a representative overview of where the technology for German TTS systems is today.

1.1 Organization of the Report

The practical evaluation of the synthesizers consists of three parts: Chapter 2, 3, and 4. *Chapter 2* discusses the evaluation setup with a description of the evaluated systems and a definition of the three evaluation criteria. *Chapter 3* deals with the objective evaluation which consists of testing the pronunciation of names and the general correctness of pronunciation. *Chapter 4* is the exposition of a subjective evaluation of the synthesizers' quality. The glossary in *Chapter 6* contains terms of the linguistic vocabulary used in the report. Most of them are well known to experts in the field and are added for completeness for inexperienced readers. In addition to the references made in the text, *Chapter 7* also contains references which serve as an overview of speech synthesis [Kel94], [Lin85]. Finally, the complete text material and the details of the test results are listed in appendices A, B and C respectively.

1.2 Acknowledgments

We would like to thank Elan Informatique, Infovox, and Lernout & Hauspie for lending us their material. They all supplied an evaluation kit including all necessary software and hardware. All of the companies were very open and helpful and offered invaluable support.

Dr. Karl Huber and Christof Traber clarified many questions about the theory and methods of speech synthesis and were kind enough to provide copies of good papers on the domain.

An example of how to evaluate speech synthesizers was kindly given by Prof. François Grosjean from the University of Neuchâtel. Help in statistical questions came from Delphine Guillelmon.

Béatrice Léwy and José Clarinval organized the 42 test persons for the subjective evaluation.

Last but not least, we would like to express our appreciation to all the test persons themselves who were willing to participate in the telephone tests.

2 Evaluation Setup

2.1 Evaluated Systems

Four different TTS systems for German were evaluated which are precisely defined in the subsequent sections. All systems also have the capability to synthesize languages other than German. We use our own abbreviations for the four systems throughout the report in order to keep the tables and figures concise. These abbreviations are given in parentheses in the company's title.

Lernout & Hauspie (LHS)

Manufacturer:	Lernout & Hauspie Speech Products, Brussels, Belgium
Platform:	Microsoft Windows
Hardware:	Sound Blaster or compatible interface
Software:	Text-To-Speech Evaluator, Version 1.00
Language:	German

Infovox (IVX)

Manufacturer:	Telia Promotor Infovox AB, Solna, Sweden
Platform:	MS-DOS
Hardware:	Infovox 500 (half-length PC board) with PROM
Software:	System software for Text-to-Speech Converter, PC-Version 3.71
Language:	German

Elan Informatique (Elan)

Manufacturer:	ELAN Informatique, Toulouse, France
Platform:	MS-DOS
Hardware:	Board TELEVOX PSOLA-8M
Software:	ALLVOC, Version 1.11
Language:	German

Berkeley Speech (BeST)

Manufacturer:	Berkeley Speech Technologies Inc., Berkeley, California
Platform:	SCO-UNIX
Hardware:	Dialogic board TTS/20[1]
Software:	Dialogic PEB T-T-S German Firmware, Version 7.9[2]
Language:	German

***Remark**: The BeST system could not be evaluated due to problems with the driver software, and therefore it will not be discussed further in the subsequent text.*

[1] in combination with the Dialogic boards: D121 and LSI-80/CH.

[2] This is the language firmware for the board. The driver software was "Base Development Package for Unix SVR 3.2 System Release 4.1".

2.2 Evaluation Criteria and General Method

In order to investigate the possibilities of integrating TTS technology into the QuickPhone project, and to judge the linguistic quality of the synthesizers, we formulated three different evaluation criteria which we regard as most important. These three criteria (see Figure 1) guided the whole evaluation process.

Criterion 1:
We want a TTS system which is subjectively accepted by the general public.

Criterion 2:
We want a TTS system which correctly pronounces general German text.

Criterion 3:
We want a TTS system which pronounces the specialized text portions which are of the greatest importance for the banking environment.

Figure 1. The three evaluation criteria.

In order to determine the acceptability of TTS systems in general and to show possible differences in the acceptance between the systems, a subjective evaluation by the general public was carried out. The test was performed using an existing text section from a banking services advertisement written for customers.

The correct pronunciation of general German text was tested by compiling a representative collection of several hundred isolated German words. This test set and its application to the three synthesizer systems, are presented in Chapter 3.2.

The specialized text portions which we consider important for the banking environment are (proper) names and numbers. In order to test the ability of the three systems to pronounce names, a separate name evaluation was carried out which is described in detail in Chapter 3.1. For the problem of pronouncing numbers, a number dictation test (with a sufficiently large group of test persons) would be interesting and helpful, but it had to be skipped due to lack of time.

3 Objective Evaluation

3.1 Names

A TTS system which pronounces every name correctly would be extremely useful for a phone banking service since it allows to address each single customer of the bank in a polite way. Pronouncing names is needed not only for addressing people but for other telephone applications, such as automatic reverse directory services[3], or for "Who's calling" facilities[4]. Apart from the pronunciation of proper names, one can think of applications which deal with product names, e.g. in a phone-based ordering system.

To test the feasibility of using the synthesizers to pronounce names, a name evaluation was undertaken. For this purpose, we used a telephone directory available in electronically readable form, consisting of more than 13'500 names. For convenience we first filtered out all entries which did not designate persons (but e.g. institutions). In a second step we formatted the entries in order to get a usable file of names.[5] Furthermore, we preceded each name with a greeting formulation ("Guten Tag, Herr/Frau/Frl. ...") which simulated the suggested usage and enabled comfortable hearing sessions.

From the large file obtained, we constructed three different test files. The first two were built automatically by taking every 100th from the total of 13'500 names (starting with name No. 50). A further separation into those entries with a female and those with a male first name resulted in two useful test files and a representative coverage of the Swiss population. The third file was built manually by choosing examples of difficult names, primarily foreign names of many different origins, and names with other specialties such as double first names, names of aristocratic origin, or names with titles. The three test files are printed in full in Appendixes A.1 and A.2.

The evaluation was performed by the authors. For the judgments, we defined and used our own rating system which is shown in Figure 2. The ratings were attributed from the point of view of potential future users. That means, a pronunciation of a name is either acceptable or not (>=2, <=1). Acceptable names may contain minor errors or are pronounced entirely correctly. Unacceptable names may contain major errors or are incomprehensible. Names marked with '-' did not produce any sound.

---

[3] A reverse telephone directory service provides the name of the person (and maybe his or her address) associated to a given telephone number.

[4] A "Who's calling" service identifies acoustically the name of the person calling the receiver of the telephone call. It is a means of preventing anonymous telephone calls and an alternative to the visual display information provided by some digital telephones today.

[5] The UNIX tools `awk` and `sed` were of great help in this process.

3 = correct
2 = acceptable
1 = unacceptable
0 = incomprehensible
- = system does not accept input

Figure 2. Rating system for the evaluation of name pronunciation.

The overall impression is that when pronouncing names none of the three systems provide satisfactory output. LHS even refused any names with letters specific to foreign languages such as "é", "à" etc. Therefore, none of the systems is suitable for applications where name pronunciation is crucial.

The details of the results of the evaluations are given in the Appendix A.[6] It is important to note that fairly common abbreviations, such as "Frl.", "Dr.", and "Prof.", were not treated properly by all systems. Some were not recognized at all. Others resulted in unnecessary pauses, probably due to incorrect sentence terminations.

To do justice to the systems evaluated, we must emphasize that they had to perform a task for which they were not designed. It is shown in the literature that systems with large lexicons of names are rarely used [Chu85]. Specialized name synthesizers are necessary, and such modules are currently being developed, at least for English [Spi90]. See also [Bas93] and [Gol93] for comprehensive surveys.

The methods used in anglophone systems are, however, probably not directly usable for German name pronunciation systems. In English, all names are pronounced anglicized, at least to a certain degree, whereas in German, names are usually pronounced as they are in the original language.

We believe that neither general purpose synthesizers with a large name dictionary nor specialized name synthesizers as stand-alone systems are suitable for real world applications such as automatic ordering systems. Our vision[7] of a flexible and open solution consists of first determining the language were a name comes from, and then, pronouncing it according to the rules of that language. This can be done by a "language detector" followed by a combination of a general purpose synthesizer for ordinary text and a name synthesizer for names for every language. The determination of the language of origin of names is still an interesting research challenge.

3.2 Pronunciation Errors

In order to test the general correctness of pronunciation, a large test collection of German words was developed from scratch. For this purpose, we used a standard

---

[6] Nicolas Léwy evaluated the test file containing special cases of names and Thomas Hornstein evaluated the files containing ordinary names. Therefore, the results from one evaluation cannot be compared directly to the those from the other.

[7] The idea of guessing the underlying language of a name is already proposed in [Lib79]. One solution involving frequency statistics of trigrams was suggested in [Chu85].

reference for correct German pronunciation, the *Siebs* [Boo69].[8] The International Phonetic Alphabet is used for the phonetic writing.

Our test collection consists of three parts: vowels and diphthongs, consonants, and - a factor which is not often considered - stress. As evidenced in Appendix B, the three parts are further divided into groups which each concern a specific sound or sound group. The words in a sound group are ordered according to phonemes which belong to that sound group. Only one sound per word is tested. It is checked whether or not the requested sound is chosen by the synthesizer. It is important to emphasize that we only tested the correctness of the particular sound in a word and not the correctness of the complete word.

It was necessary to skip four test items of a preliminary draft of the test collection because their pronunciation given by the Siebs did not correspond to today's usage. They are listed in Table B1 at the end of Appendix B.4.

Furthermore, we did not check the correct pronunciation of words written with <ss> instead of <ß>. In Switzerland, they are written with <ss>. A TTS system optimized to Swiss usage has to take this particularity into account.[9]

The results of the pronunciation test is too extensive to be discussed in detail. The complete test results are given in Appendix B. The following paragraphs give a rough overview of the linguistic difficulties of the systems.

An sample of words tested which none of the synthesizers pronounced correctly is given in tables 2 & 3. Table 2 contains words with errors but which are considered to be understandable and, therefore, acceptable. Most of them originate in a foreign language but they are German words nonetheless. Table 3 lists words pronounced with errors which we consider severe. In both tables the underlined characters mark the part (or sound) of the word which was tested. Words with an apostrophe are words where stress was checked. Most of them are derived words and a few are simple lexemes or compounds.

Melodram	Ökumene	Clown
Memoiren	Tweed	Couch
loyal	Seldwyla	Ha-iti
Gabriel	Milieu	Alle-uten
Benefiz	Etude	Krelusa
Etui	Garage	brillant
Bazar	Jury	Billard
Chinchilla	Giro	Herme'lin
Boccia	Arpeggien	

Table 2. Words incorrectly pronounced by all synthesizers but acceptable

[8] Volume 6 of DUDEN would have been another possibility.

[9] E.g. frassen, Spässe, bloss, Geschoss, Grösse, Schössling, Musse, Genuss, süss, etc. should be pronounced exactly as their counterparts with <ß> are.

Bratsche	Juwel	Behörde
Magnet	Segment	flugs
hätscheln	Türkis	Konsul
weg	Meteor	Rüsche
Ungar	Frauchen	Obmann
Frevel	Unbe'rechenbarkeit	vor'züglich
Lebe'wohl	Un'heilbarkeit	Wiederinbe'sitznahme
durch'queren	um'fluten	vor'trefflich
voll'bringen	um'kreisen	

Table 3. Words incorrectly pronounced by all synthesizers with severe errors

Table 4 lists the number of errors for each synthesizer for the different sounds and sound groups (or stress problems) which we distinguish in the test collection. An error is an incorrectly pronounced word. The last column in the table indicates the number of words tested, or in other words, the number of possible errors.
All systems have problems with vowels. It makes little difference whether these vowels are in borrowings or non-borrowings. IVX is the best system for the non-borrowings, and Elan for the borrowings. Both are better than LHS for vowels in general. Diphthongs and voice onsets are produced incorrectly in more than one third of the cases.

For consonants, IVX gives the best overall results. Difficulties are more likely to arise for some fricatives and for liquids and nasals than for plosives. Fricatives such as <‘>, <t‘>, <“>, and <d“> (e.g. Böschung, Boccia, Garage and Gin) are rarely produced correctly. Also the <v> is often treated incorrectly.

Stress, however, seems to be difficult to automate. Maybe it has been underestimated and tackled with too simplistic rules, such as single stress patterns for a given prefix, combined with simple exceptions lists. Elan has the best stress results, whereas LHS again has the worst.

	LHS	**IVX**	**Elan**	**max.**
(1) Vowels and diphthongs				
(a) a-sounds	15	10	14	31
(b) e-sounds	21	13	18	47
(c) i-sounds	12	8	11	26
(d) o-sounds	11	13	9	29
(e) ö-sounds	9	3	5	16
(f) u-sounds	14	11	6	29
(g) ü-sounds	10	7	4	23
(h) Diphthongs and voice onset	10	10	12	28
Non-Borrowings Total	41	23	31	103
Borrowings Total	51	42	36	98
Diphthongs etc. Total	10	10	12	28
Vowels & Diphthongs Total	**102**	**75**	**79**	229

(2) Consonants				
(a) Liquids and nasals	7	4	5	18
(b) Fricatives: [h]	3	0	3	13
[v, ph, pf, ps, qu]	7	4	6	22
[z, s, ts, ks]	6	5	6	28
[st, sp]	2	1	3	13
[ʃ, tʃ, ʒ, dʒ]	11	11	11	13
[ç, x, j, lj]	4	6	6	19
(c) Plosives	2	1	1	13
CONSONANTS TOTAL	**42**	**32**	**41**	139
(3) Stress				
(a) Simple lexemes and compounds	3	3	3	7
(b) Derived items	35	29	24	88
STRESS TOTAL	**38**	**32**	**27**	95
OVERALL TOTAL	**182**	**139**	**147**	463

Table 4. Number of pronunciation errors on the test collection

4 Subjective Evaluation

The purpose of the subjective evaluation is to measure the speech quality and the acceptance rating of a TTS system. Persons who have not previously used a synthesized speech system or have had no exposure to synthesizers at all, give quality judgments which are directly comparable with the reactions of future users to an application which integrates this synthesizer. In addition, the judgments given for different synthesizers may be compared in order to reveal the advantages and drawbacks of each of them.

Quality evaluation can be achieved in various ways. Methods reported in the literature ([Kla87], [All87], [Bai92]) range from intelligibility of isolated words via reading comprehension tests, to analytic speech assessment techniques. Because our interest in TTS systems is related primarily to the development of computer applications, we did not investigate methods which reveal detailed behavior; rather we chose a method which shows the overall characteristics. Therefore, subjective opinion rating tests were conducted, based on the reading of a coherent and self-contained text.

4.1 Text Material

A section from a banking services advertisement (see Figure 3), written for the general public, was chosen as text material for three reasons. First, the text is simple enough to be understood on the first hearing; second, it is not too trivial in its content and vocabulary (compared to the children's stories which are used sometimes); and third, one can imagine a similar text in a real telephone application. [10]

---

10 A fourth reason is its availability in French and Italian. Consequently an evaluation of more than one language reveals possible differences in the quality and usability of a given system. A system designed primarily for one language, and subsequently adapted to other languages, might show such differences.

Informieren Sie sich rund um die Uhr über Ihr Geld. Mit der Liberty-Line!

Die Liberty-Line ist ein exklusiver Telefonservice nur für Liberty-Kunden. Damit haben Sie den direkten Draht zu all Ihren Liberty-Konten. Und können so überall und rund um die Uhr den Kontostand und die letzten fünf Buchungen abfragen. Bequem, sicher, einfach und günstig. Denn Sie bezahlen nur 33,3 Rappen PTT-Gebühren pro Minute. Stand Juni 94.

So einfach geht das: Wählen Sie die Nummer der Liberty-Line. 157, 03, 45, 2. Nochmals: 157 03 45 2. Dann Kundennummer und PIN-Code eingeben. Schon werden Sie automatisch durch den Service der Liberty-Line geführt.

Falls Sie kein Telefon mit Tonwahl haben, benötigen Sie ein Zusatzgerät, den Liberty-Link. Sie erhalten ihn als Schlüsselanhänger gratis in jeder SBG.

Figure 3. The German text for subjective evaluations.

A well-defined number of pronunciation errors was corrected in the text. This rather unconventional measure was taken to avoid any possible distraction caused by incorrect pronunciation which might therefore interfere with the assessment of speech quality. It is generally easier to test and explain one single factor (here the speech quality) than many different factors simultaneously. Pronunciation errors of the synthesizers were dealt with in Chapter 3.2 already. We also assume that for real world systems some pronunciation correction must be done. This is usually implemented by using an extended system dictionary which contains error-prone words of the vocabulary of the language (such as the "bequem" of our text) and a user dictionary where foreign and domain-specific words like "Liberty-Line", "SBG", and "PTT" can be entered. Table 5 shows the pronunciation corrections which were applied to the three systems. They were implemented with the tools provided (phonetic transcriptions, and orthographic rewriting).

Original -> Correction	**LHS**	**IVX**	**Elan**
foreign words:			
"Liberty" -> "Libberti"	no	yes	yes
"Line" -> "Lein"	yes	yes	yes
"service" -> "sörvis"	yes	yes	yes
"Konto" -> "Konnto"	no	yes	yes
"PIN-Code" -> "Pinn-Code"	yes	yes	no
"Code" -> "Cood" (or "Cod" resp.)	yes	yes	yes
acronyms:			
"SBG" -> "es beh geh" (or "S B geh" resp.)	yes	no	yes
"PTT" -> "peh teh teh" (or "P teh T" resp.)	yes	no	yes
simple words:			

The translations of the text into French and Italian are given in Appendix C.1.

"bequem" -> "bequeem"	yes	no	no
"informieren" -> is stressed on the third syllable	no	no	yes
composite words:			
"Kundennummer" -> "Kunden-Nummer"	yes	no	no
"Kontostand" -> "Konto-Stand" (or "Konto Stand" resp.)	yes	yes	yes
"Schlüsselanhänger" -> "Schlüssel-Anhänger"	yes	yes	yes

Table 5. Pronunciation corrections (yes = "needed", no = "not needed").

The vocabulary and grammar of the text are appropriate for our requirements. The vocabulary contains a range of nontrivial, banking and technical, terms. The grammatical structures found are diverse, including emphasis ("... *nur* für Liberty-Kunden ..."), incomplete sentences (such as the adjective sequence), and also the exclamation "Mit der Liberty-Line!". By means of prosody judgments, structures of this kind allow us to test the systems' knowledge of German grammar.

4.2 Speech Production

Each of the three synthesizers received the text and the necessary pronunciation corrections as input and had to generate on-line speech output. All options of each synthesizer were set to their default values, i.e. to the settings recommended by the manufacturers.[11] This was done for the following parameters (if adjustable): pitch level, pitch dynamics, aspiration, rate, and pause durations.[12]

The synthesizers' output was recorded with a digital audio tape-recorder (DAT) Sony TCD-D7, with a sampling frequency of 48 kHz. A separate tape was created for each synthesizer.[13] The recorded text is repeated four times on the tape, each slice is separated by 4 sec of silence. The duration of the spoken texts was of comparable length (LHS: 71 sec; IVX: 69 sec; Elan: 80 sec). Only the Elan synthesizer seems to have a slower articulation rate as a default.[14]

4.3 Subjects

A total number of 42 subjects ran the evaluation, 14 (7 women and 7 men) for each of the 3 synthesizers. They were selected from the professional and personal environment of the authors. Care was taken to have a representative sample of the population of potential future users. Table 6 lists the properties which were considered.

---

[11] It is generally a good idea to use the default values of the manufacturer because they were determined through extensive opinion tests.

[12] As described below, the volume was controlled independently of the synthesiser systems by selecting similar output levels when playing the DAT recorder.

[13] The recording procedures were different for each synthesiser because of differing technical difficulties; the exact handling and configuration are described in Appendix C.2.

[14] See the chapter Results & Discussion as to whether this fact had an impact on the judgements of rates.

	LHS	IVX	Elan
(a) Age groups			
≤ 20	2	1	1
20-30	4	4	4
30-40	2	2	5
40-50	3	4	2
50-60	1	0	1
60-70	2	2	0
≥ 70	0	1	1
(b) Technical interest			
very technical	1	5	3
somewhat technical	1	2	1
not at all	12	7	10
(c) Previous exposure			
considerable	2	3	0
little	5	5	7
none	7	6	7

Table 6. Subject properties

The first criterion is the age of the subjects. The subjects were only asked to which of the seven age groups they belong.

The second criterion, technical interest, shows the level of computer and other technical experience or interest. The subjects were asked how technical their job and activities were. This was expected to have an influence on the attitude of the subjects towards synthetic speech, as it might have on the attitude towards any technical system.

As a third variable, previous exposure to synthesized speech was checked. People were grouped into three categories, depending whether they had considerable, little or no experience with synthesizers. It is possible that some of the subjects confused computer-generated speech with recorded human voices played back by a computer. This was not regarded as an interpretation error because it was the user's impression which we were interested in.

The distribution by categories differs slightly from synthesizer to synthesizer because the categories into which the participants fell were checked only after the tests had been run. However, we found the test results discussed below to be reliable overall.

4.4 Procedure

Each evaluation with a subject consists of three telephone calls. In a first informal call, they are asked whether they are ready to assist in a telephone test. They are told that a test of three different speech synthesizers for German is being undertaken, and that they will hear only one of them. They are also told that they will hear the computer-generated text of a banking advertisement. The quality of the produced text is to be judged according to 10 criteria with ratings from 1 (very bad) to 6 (very good). None of the criteria are to do with the content of the text. Subjects are also told that the text will be played several times and that they can hang up as soon as they are ready to give their judgments.

In the second call, the subjects hear the recorded synthesized text. The recording is played straight-away without any previous introductory text or sound and without any interaction from the experimenter. The output from the DAT was coupled via a loudspeaker to the microphone of the telephone's handset. Care was taken that test conditions remained identical throughout the tests.

In the third call, the questions for the ten criteria are posed. Each question covered a specific part of the quality of synthesized speech. Table 7 lists the 10 questions. An English translation of the original is given for each question. The questions are written explicitly on a prepared questionnaire used by the experimenter in order to achieve identical conditions for each run. The subjects are asked to give ratings between 1 and 6 with half points (e.g. 3.5) being allowed. This method is borrowed from the Swiss school rating system, which is familiar to most of the subjects.

After the 10 questions, the subjects are asked, as explained in the previous section, in which age category they are, whether they have technical interests, and/or previous exposure to synthesized speech. In addition, subjects can make special observations or state opinions. All but four of the tests are conducted in Swiss German, the conversational language in Switzerland, the exceptions being persons who preferred standard German.[15]

Q1	**Listening effort** *Wie anstrengend, wie mühsam haben Sie es empfunden, den Sinn des Textes zu verstehen?* How difficult was it for you to understand the sense of the text?
Q2	**Comprehensibility** *Wie gut verständlich ist der Sinn, der Inhalt, des Textes?* How well could you understand the meaning, the content of the text?
Q3	**Articulation** *Wie deutlich ist die Aussprache?* How clear is the articulation?
Q4	**Rate** *Wie gut beurteilen Sie die Geschwindigkeit der Sprache?* How do you rate the speed of the voice?
Q5	**Prosody** *Wie gut beurteilen Sie den Satzrhythmus (die Betonung und Pausen)?* How do you rate the prosody (the stress and the pauses)?
Q6	**Naturalness** *Ist die Stimme natürlich ? Wie stark natürlich?* Does the voice seem natural? To what extend?
Q7	**Pleasantness** *Ist die Stimme angenehm, sympathisch?* Is the voice pleasant?
Q8	**Foreign accent** *Wie stark finden Sie hat die Stimme einen fremdsprachigen Akzent?* Does the voice have a foreign accent, and if so, how strong?

[15] Subjects No. 8, 10, 23, and 29.

Q9	**System acceptability** *Wie fest wären Sie bereit, ein automatisches Telefonsystem zu benützen, das genau diese Stimme (und nicht eine andere) verwendet?* How willing would you be to use an automatic telephone system which uses exactly this voice (and not another synthetic one)?
Q10	**General acceptability** *Wie fest wären Sie bereit, ein Telefonsystem mit einer künstlichen Stimme zu benützen (verglichen mit einer menschlichen, z. B. von einer Tonbandaufnahme)?* How willing would you be to use a telephone system which uses a synthetic voice (compared to a human voice, e. g. from a tape recording)?

Table 7. The 10 criteria and the associated questions.

4.5 Results & Discussion

The complete answer data is given in full detail in the table in Appendix C.3. The table lists the type of TTS system, the characteristics of the subjects, and the answers to the questions for each subject. Appendix C.4 shows the distribution of answers for each of the ten questions in separate charts.

Unacceptable ambient noise was established during the test of one of the 42 subjects; so, this test was repeated with another person.[16] Furthermore, in three cases, subjects did not want to give an answer to a particular question for various reasons; theses cases are marked in Table C3 in the appendix as "noa" (no answer).

The discussion of the results is divided into four parts. The first part shows the results of the examination of the 10 criteria. The second part discusses the relationship between the three properties of Table 6 and some of the criteria. The third part illustrates the influence of the sex of the subjects on their preferences. Finally, the influence of the age of the subjects on the general acceptability was analyzed.

Figure 4 summarizes the answers of the subjects. It shows the mean values of the answers versus the criteria for each synthesizer.[17] A legend of the symbols is given in Figure 5.

[16] For completeness, the answer set of the substituted subject is given here:

-	IVX	M	40-50	T	M	3.5	5	3	5	1	1	2	5	3	6

[17] An analysis of the significance of the differences between systems has not been carried out. It can be done by means of an analysis of the variance.

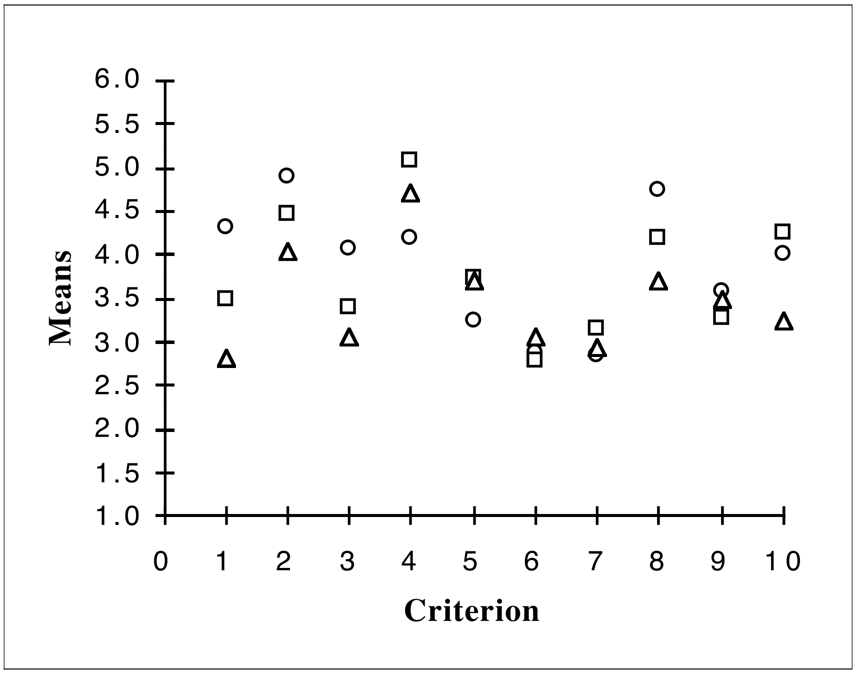

Figure 4. Mean values of answers for each criterion and synthesizer

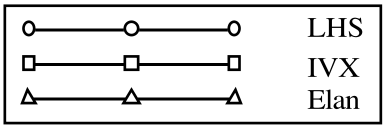

Figure 5. Legend to Figure 4.

In a first task, the answers to the 10 criteria are discussed and compared between the synthesizers:

- The clearness of the articulation (Q3) is best for LHS, followed by IVX and Elan. This might be interpreted as having an influence on comprehensibility (Q2) and listening effort (Q1). The results for both of these criteria show a similar distribution to the result from the articulation criterion. The mean values for comprehensibility are significantly higher than for listening effort and articulation. This means that people may understand most of the text but they probably don't feel very comfortable. It is interesting to see that for all three systems the rating for listening effort shows the most marked divergence.
- The results for the default speaking rate (Q4) follow another pattern: the results for LHS suggest an unsatisfactory rate, whereas IVX and also Elan score well or fairly well. The good results here are not surprising because adopting an optimal speaking rate is an easy task compared to the other criteria.
- Prosody (Q5), naturalness (Q6), and pleasantness (Q7) have similar ratings for all systems. It cannot be clearly said which one scores best. However, the mean values for the prosody give an incorrect impression of the real situation. The distribution chart of Figure C9 in Appendix C.4 shows a very different judgment among subjects. It could be argued that what constitutes good and bad prosody is

not clear to the subjects. The criteria naturalness and pleasantness probably also had an influence on this judgment. Care should be taken in subsequent investigations that the evaluation of the prosody is well isolated from other criteria.

- The criterion of foreign accent (Q8) may give an indication as to the language which the system was originally designed for. Low ranking means bad quality or strong foreign accent. Elan seems to produce impressions of a French or English accent. Some subjects supposed English or Dutch influence in IVX. LHS scored best and demonstrated least evidence of a foreign language influence: in a few cases traces of a French accent were detected.
- The most interesting question probably concerns the system acceptability (Q9). All systems have nearly the same mediocre mean value. The distribution of the answers results in two groups; one with a high and one with a low acceptance. This is accurate for all three systems. It is interesting to note that the general acceptability (Q10) has the same characteristic. The latter is given for completeness and cannot be compared between the systems.

In a second task, we compared the general acceptability of TTS versus the technical interest and the previous exposure of the subjects.

Figure 6 illustrates that people who are interested (very interested or somewhat interested) in technical matters are also ready and willing to use a speech synthesis system. People who are not interested at all in technical matters, however, can be divided into two equally important groups: one is also willing to use the speech synthesis system, to almost the same extent as the people with technical interest. The others are clearly against TTS systems.

The influence of previous exposure is shown in Figure 7. In the case where the subjects had no previous exposure, they will either strongly refuse the use of TTS systems or they will accept it quite readily, but nothing in between. As soon as some experience with TTS systems has been gained, people show a tendency towards greater acceptance, i.e. people who are in some way accustomed to the technology will accept it more easily. As TTS systems are more widely used in future applications, people will become more accustomed to them. Eventually, this will lead to greater acceptance.

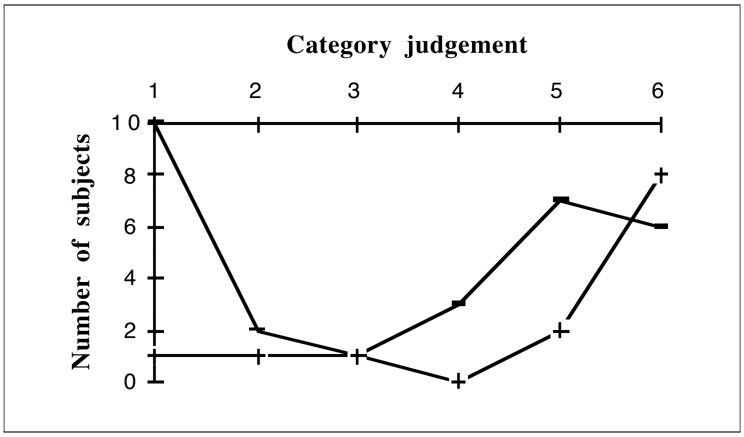

Figure 6. Acceptability of TTS depending on technical interest
(+ "very technical," - "somewhat or not technical").

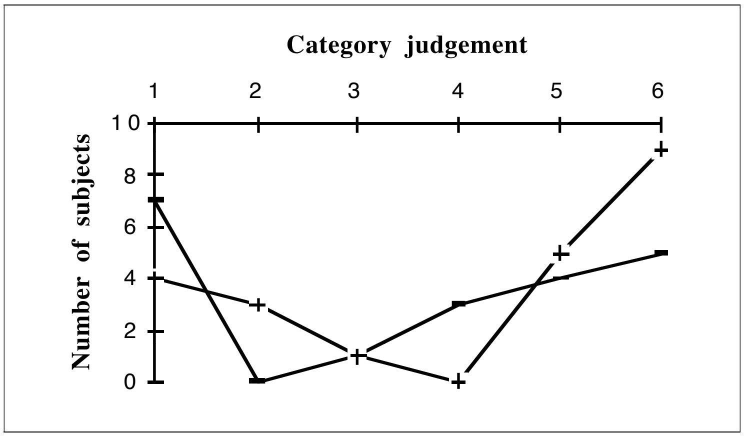

Figure 7. Acceptability of TTS depending on previous exposure
(+ "many voices", - "some or no voices").

It is possible, that this positive result is only due to correlation with the criterion of technical interest, in other words, that the subjects who are interested in technical matters are those with greater opportunity for previous exposure. In order to test whether this is the case or not, we did the same test but only for those subjects who have no technical interests. Figure 8 shows the results. The characteristics of Figure 7 are slightly weakened but could be confirmed.

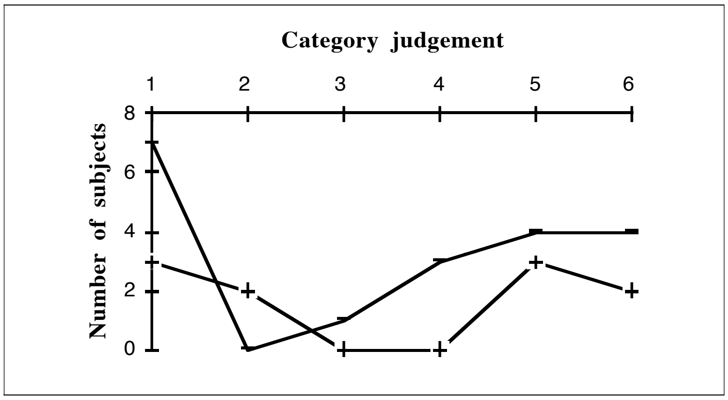

Figure 8. Acceptability of TTS, depending on previous exposure of subjects without technical interest
(+ "many voices", - "some or no voices").

As a third task, the judgment in terms of naturalness and pleasantness depending on the sex of the subjects are investigated. A frequent claim is that women prefer a male voice for telephone applications, and vice versa. This statement was confirmed for LHS, and at least partly for Elan as far as naturalness was concerned. For the criterion of pleasantness we did not find any significant differences between sexes. Naturalness and pleasantness are without a doubt two completely different facets of quality. At this point, it is important to repeat that the default voice of LHS is female, whereas the two others are male.

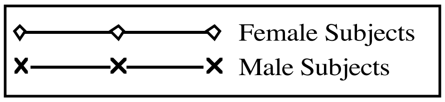

Figure 9. Legend for Figure 7-9.

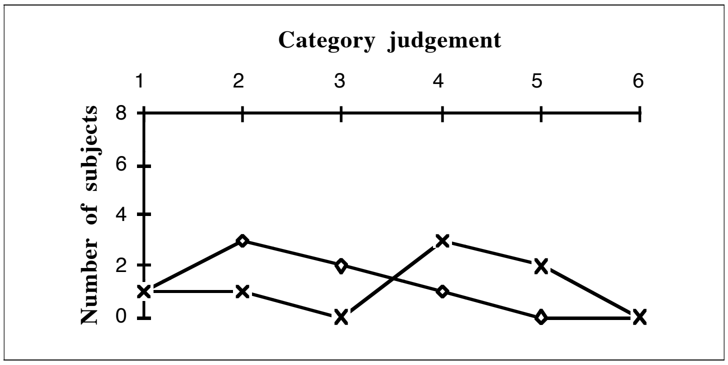

Figure 10. Naturalness of LHS's default female voice.

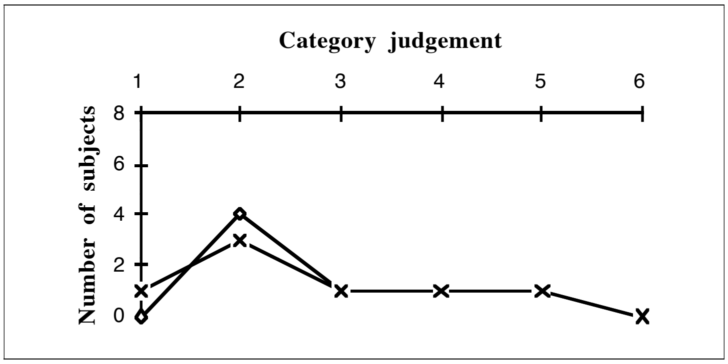

Figure 11. Naturalness of IVX's default male voice.

Both men and women judged the naturalness of IVX to be poor, there is no difference between the sexes at all. But for the two other systems there are differences between the sexes: on the one hand, male subjects clearly judged LHS's female voice as being more natural than their female counterparts did, and to some extent, female subjects rated Elan's male voice more natural than their male counterparts did. Further evidence for the previously stated assertion was found, therefore with regard to the naturalness of the tested systems we think that men and women do seem to like certain voices more than others. To some degree this depends on the sex of the voice. However, other characteristics may also play a role, as the IVX result shows. Such factors remain to be investigated in depth in a further study.

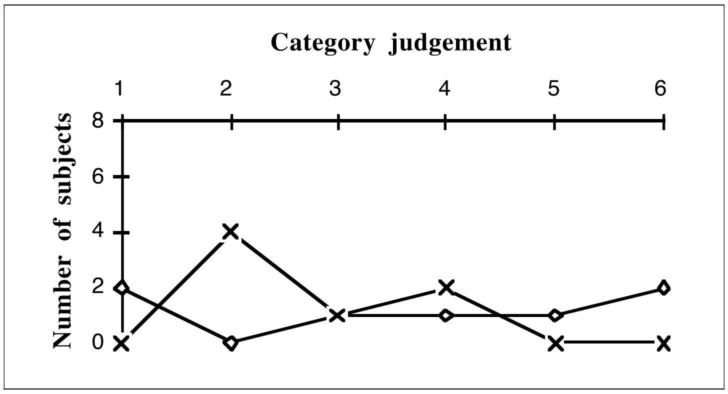

Figure 12. Naturalness of Elan's male default voice.

Finally, we analyzed the influence of the age of the subjects on their judgment - whether they would accept and use systems with synthetic voice or not. We expected that the younger subjects would accept TTS more easily than the older subjects would. This prejudice was proved completely wrong. For this analysis the 7 age classes checked in the experiment were grouped into three larger classes: ≤ 30, 30-50, and ≥ 50. Table 8 shows the mean value for each age group; distributive data is not provided because the groups are not of the same size.

age group	**≤ 30**	**30-50**	**≥ 50**
mean rate	3.3	3.8	5.1
group size	16	18	8

Table 8. General acceptability across age groups.

The results show that the older the subjects are, the more they accept TTS. We interpret this unexpected result in the following way. People over 50 (at least in Switzerland), did not grow up with the same level of technological progress as the younger ones did. They may, therefore, be rather surprised by the current possibilities, whereas the younger people are more critical in this regard.

5 Conclusion

No miracles can be expected from the German TTS systems currently available on the market. The voices of these systems still sound artificial, i.e. users immediately recognize that the voice is generated synthetically. Despite this fact, TTS systems have made progress in the last years. With respect to our three evaluation criteria we can emphasize the following conclusions:

The acceptance of the systems by the general public can be judged positive. Broader deployment of services using TTS will result in better acceptance of TTS.

Correct pronunciation of general German text is achieved without serious shortcomings. Progress has been made mainly in improving the stress and the rhythm of the sentences.

For the pronunciation of specialized text portions such as names or foreign words none of the systems functions properly. Exception dictionaries can be set up but for a large number of exceptions these quickly assume proportions which make them impractical to use and to maintain. Solutions to this problem can be found only by adopting entirely new approaches. So far, TTS systems are not suitable for an operation a banking environment with occasional users. The correct pronunciation of names is an important issue in such an environment.

The advantages of TTS systems compared to playing prerecorded words and phrases become fully apparent when used in applications for a limited or trained number of users. In the scope of prototyping IVR services using TTS, development cycles may be shortened significantly. This enables an exploratory development of the dialogues which helps to improve the user interface of a service in an iterative way.

6 Glossary

Borrowing
A word which belongs to the vocabulary of a particular language (in our case German) but where the whole word or part(s) of it etymologically came from another language (often from Latin or Romance languages).

Compound
In the context of this report, a compound is a word which consists of two or more lexemes. The expression compound is also often used in the context of sentence grammar. A compound sentence consists of two or more clauses.

Diphthong
A diphthong is a combination of two vowels (e.g. <au>, <ei>, <eu> etc.).

Fricative
Fricatives (also called constrictives) are formed when the stream of air must flow through a very narrow opening in the vocal tract, so that turbulence (friction) is created. Examples are the <s> and <f> of "singen" and "fast".

Lexeme
A lexeme is a vocabulary word.

Liquid
Liquids are sounds where the tip of the tongue approaches the alveolar ridge. Liquid consonants are <r> and <l>, found in the initial position of the words "rot" and "leicht" respectively.

Morpheme
A morpheme is seen as the basic unit of a grammatical structure. Traditionally, the different building blocks that make up a word are called morphemes.

Nasal
Nasals are speech sounds where the passage into the nasal cavity is kept open. Air can escape through the nose at the same time as it also emerges through the mouth. Examples are <m> or <n>.

Phoneme
A phoneme is the smallest abstract sound unit of a language. The phonetics of a language is highly structured and consists of a phonetic alphabet. The phonetic alphabet used in this report is defined by the IPA (International Phonetic Association).

Plosive
Plosives are consonants produced with a sharp release of breath. Examples in German are <p>, <t>, <k> (all unvoiced), and <b>, <d>, <g> (all voiced).

Stress
Stress is the emphasis that you put on a word or a part of a word when you pronounce it, with the result that it sounds slightly louder.

7 Literature

[Akm79] Akmanjian A, Demers R A, Harnish R M: *Linguistics: An Introduction to Language and Communication*. The MIT Press, Cambridge Massachusetts, 1979

[All87] Allen J, Hunnicutt M S, Klatt D: *From text to speech: The MITalk system,* Cambridge University Press. 1987

[Bai92] Bailly G, Benoit C, eds: *Talking Machines: Theories, Models, and Designs*. Amsterdam: North-Holland, 1992

[Bas93] Basson S. et. al: Comparing synthesizers for name and address provision. *Proc. Am. Voice I/O Soc. (AVIOS-93)*, 1993

[Boo69] De Boor H. et al., eds: *Siebs, Deutsche Aussprache; Reine und gemäßigte Hochlautung mit Aussprachewörterbuch*. 19th edition. Berlin: de Gruyter 1969

[Chu85] Curch K. W: Stress Assignment in Letter-to-Sound Rules for Speech Synthesis. *Proc. 23rd Meeting Assoc. Comp. Ling.*, pp. 246-53, 1985

[Dud90] Duden 6: *Das Ausprachewörterbuch*. 3. Auflage, Bibliographisches Institut & F. A. Brockhaus AG, 1990

[Gol93] Golding A, Rosenbloom P: A comparison of Anapron with seven other name-pronunciation systems. *J. Am. Voice I/O Soc.* 14, pp. 1-21, 1993

[Hor94] Hornstein T: Telephone Voice Interfaces on the Cheap. In Bischofberger W. & Frei H.-P., eds: *Computer Science Research at UBILAB, Strategy and Projects; Proc. UBILAB Conf.-94*. Konstanz: Universitätsverlag Konstanz, 1994

[Lin85] Linggard R: *Electronic Synthesis of Speech*. Cambridge University Press, Cambridge UK, 1985

[Kel94] Keller E. ed: *Fundamentals of Speech Synthesis and Speech Recognition*: Basic Concepts, State of the Art and Future Challenges. John Wiley, Chichester UK, 1994

[Klu93] Klaus H. et al: An evaluation system for ascertaining the quality of synthetic speech based on subjective category rating tests. *Proc. European Conference on Speech Communication and Technology (Eurospeech-93)*, 1993

[Kla87] Klatt D H: Review of text-to-speech conversion for English. *Journal of the Acoustical Society of America,* Vol. 82 No. 3, pp. 737-793, 1987

[Lio81] Lyons J: *Language and Linguistics*. Cambridge University Press, Cambridge UK, 1981

[Spi90] Spiegel M, Macchi M: Synthesis of names by a demisyllable-based speech synthesizer (Orator). *J. Am. Voice I/O Soc.* 7, pp. 1-10, 1990

8 The Authors

Nicolas Léwy

Nicolas Léwy has been studying computer science at the Swiss Federal Institute of Technology (ETH) in Zurich and is a part-time research assistant at the Laboratoire de traitement du langage et de la parole (LTLP), University of Neuchâtel. He spent 13 weeks at UBILAB undertaking a practical training which is part of his studies.

Thomas Hornstein

Thomas Hornstein received a B.Sc. degree in Electronics from the Basle Institute of Technology in 1987 and a diploma in SW Engineering (NDS Informatik) from the Softeware-Schule Schweiz in Berne in 1991. Since the autumn of 1991, he has been working as a research staff member at UBILAB. He headed the QuickPhone project and his research interests are concentrated on HCI (Human Computer Interaction) aspects of interactive voice response systems.

9 Appendices

A Material for the Study of Names

A.1 Ordinary names

	LHS	IVX	Elan
A.1.1 Names with female first names			
Guten Tag, Frau Sylviane Andrey.	2	2	1
Guten Tag, Frau Catherine Bastian.	2	2	2
Guten Tag, Frau Gisela Bigler.	1	2	1
Guten Tag, Frau Franciska Bosshard.	3	3	3
Guten Tag, Frau Beatrice Burkhardt.	1	1	1
Guten Tag, Frau Christiane Ciaccio.	1	1	2
Guten Tag, Frau Micheline Derber.	1	1	1
Guten Tag, Frau Verena Eichinger.	3	3	2
Guten Tag, Frau Chantal Folly.	1	1	1
Guten Tag, Frau Elisabeth Gattiker.	3	2	3
Guten Tag, Frl. Anja Grand.	2	2	3
Guten Tag, Frau Ingrid Haltenberger.	2	2	2
Guten Tag, Frau Katharina Hofer.	2	1	2
Guten Tag, Frau Sybille Häusermann.	2	3	2
Guten Tag, Frau Elisabeth Kaiser.	3	2	3
Guten Tag, Frau Matilde Kohler.	3	3	3
Guten Tag, Frl. Leopoldine Lachmayer.	3	2	2
Guten Tag, Frau Katica Lovricic.	1	1	1
Guten Tag, Frau Verena Matzenauer.	2	3	2
Guten Tag, Frau Monica Moledo.	2	1	2
Guten Tag, Frau Ann Nelson.	1	1	1
Guten Tag, Frau Isabelle Parlier.	1	1	1
Guten Tag, Frau Diana Preid.	1	1	1
Guten Tag, Frau Marie-Claire Rochat.	1	1	1
Guten Tag, Frau Caterina Savi.	2	1	2
Guten Tag, Frl. Regula Schmucker.	3	2	2
Guten Tag, Frau Suzanne Seyfried.	1	2	2
Guten Tag, Frau Doris Steudler.	3	3	3
Guten Tag, Frau Marianne Thyssen.	3	3	3
Guten Tag, Frau Evangelia Virvilis.	1	2	2
Guten Tag, Frau Annemarie Widmer.	3	1	1
Guten Tag, Frau Monika Zehnder.	3	2	2
A.1.2 Names with male first names			
Guten Tag, Herr Erich Aeschlimann.	3	3	1
Guten Tag, Herr Josef Aranya.	1	2	2
Guten Tag, Herr Matthew Baglay.	1	1	1
Guten Tag, Herr Josef Baumgartner.	3	3	3
Guten Tag, Herr Imre Berkovics.	1	2	1
Guten Tag, Herr Karl Binzegger.	2	1	3
Guten Tag, Herr Harald Bohne.	2	3	3
Guten Tag, Herr Claude Bovy.	1	1	1
Guten Tag, Herr Walter Brupbacher.	1	2	2
Guten Tag, Herr Mikael Busch.	2	2	2
Guten Tag, Herr Peter Böni.	3	3	2

	LHS	IVX	Elan
Guten Tag, Herr Michal Cellar.	1	2	1
Guten Tag, Herr José Cortes.	-	1	1
Guten Tag, Herr Benedikt Deiss.	2	1	1
Guten Tag, Herr Gilbert Dreyfuss.	1	2	2
Guten Tag, Herr Guido Egeter.	2	1	2
Guten Tag, Herr René Erne.	-	1	1
Guten Tag, Herr Roger Ferner.	1	1	1
Guten Tag, Herr Christian Fraipont.	1	1	2
Guten Tag, Herr Ralph Fütterer.	1	2	1
Guten Tag, Herr Tryphon Georgallides.	1	2	1
Guten Tag, Herr Leopold Goldner.	3	3	3
Guten Tag, Herr Reto Grubenmann.	3	3	3
Guten Tag, Herr Hans-Peter Gämperli.	2	3	2
Guten Tag, Herr Michel Hartmann.	3	3	3
Guten Tag, Herr Herbert Helbling.	3	3	2
Guten Tag, Herr André Hirzel.	-	2	2
Guten Tag, Herr Franz Huber.	3	3	3
Guten Tag, Herr Beat Häfliger.	2	2	1
Guten Tag, Herr Laurence Jaccard.	1	1	1
Guten Tag, Herr Franz Jäggi.	2	2	2
Guten Tag, Herr Jürg Keller.	3	2	3
Guten Tag, Herr Dr. Erich Klopfenstein.	1	2	2
Guten Tag, Herr Daniel Kornmann.	2	1	1
Guten Tag, Herr Herbert Kupferschmied.	3	3	2
Guten Tag, Herr Rolf Landis.	3	3	2
Guten Tag, Herr André Lenweiter.	-	2	2
Guten Tag, Herr Andrew Lorenc.	1	1	1
Guten Tag, Herr Hubert Magne.	2	3	3
Guten Tag, Herr Fritz Mathys.	1	1	1
Guten Tag, Herr Jürg Meili.	3	3	3
Guten Tag, Herr Harald Minnig.	2	2	2
Guten Tag, Herr Heinz Mühlebach.	3	3	3
Guten Tag, Herr Dr. Dulal Nanda Nandi.	2	2	1
Guten Tag, Herr André Nydegger.	-	1	1
Guten Tag, Herr Renato Pajarola.	2	1	1
Guten Tag, Herr Pierre Peyer.	0	1	1
Guten Tag, Herr Kurt Probst.	3	2	3
Guten Tag, Herr Pius Regli.	2	1	3
Guten Tag, Herr Martin Rindlisbacher.	3	2	3
Guten Tag, Herr Dr. Jörg Rossow.	2	3	2
Guten Tag, Herr Max Rüegger.	1	2	2
Guten Tag, Herr Thierry Schafflützel.	1	1	2
Guten Tag, Herr Rainer Schlotter.	3	3	3
Guten Tag, Herr Arnold Schnyder.	1	1	1
Guten Tag, Herr Ernst Schüpbach.	1	1	3
Guten Tag, Herr Georg Sgier.	1	2	1
Guten Tag, Herr Massimo Sonego.	2	1	2
Guten Tag, Herr Paul Stehrenberger.	3	3	3
Guten Tag, Herr Ulrich Strickler.	1	2	1
Guten Tag, Herr Istvan Szabo.	2	2	2
Guten Tag, Herr Jeffrey Tolmie.	1	1	1
Guten Tag, Herr Franz Unternährer.	2	2	2
Guten Tag, Herr Gerd Voswinkel.	2	2	1
Guten Tag, Herr Joe Weber.	1	1	1

Guten Tag, Herr René Weyermann…………………………	-	1	1
Guten Tag, Herr Kurt Witschi………………………………	3	2	2
Guten Tag, Herr Marco Zannini.……………………………	1	2	2
Guten Tag, Herr Anton Zubler. ……………………………	1	3	3

A.1.3 Summary of the results

SUM OF THE RATINGS FOR WOMEN	**62**	**58**	**60**	**(96)**
SUM OF THE RATINGS FOR MEN	**117**	**133**	**128**	**(207)**
RATINGS TOTAL	**179**	**191**	**188**	**(303)**
MEAN RATING FOR WOMEN	**1.94**	**1.81**	**1.88**	
MEAN RATING FOR MEN	**1.70**	**1.93**	**1.86**	
MEAN RATING TOTAL	**1.77**	**1.89**	**1.86**	
N. OF CORRECT OR ACCEPTABLE, FOR WOMEN	**20**	**19**	**20**	**(32)**
N. OF CORRECT OR ACCEPTABLE, FOR MEN	**37**	**45**	**41**	**(69)**

A.2 Special cases of names (with both male and female first names)

A.2.1 German

Guten Tag, Herr Rolf Haeny.	1	0	0
Guten Tag, Frau Rose Dietrich.	3	3	3
Guten Tag, Herr Michael Zbinden.	3	3	1
Guten Tag, Herr Christoph Frey.	1	3	3
Guten Tag, Herr Hans Gasteyger.	1	2	3
Guten Tag, Herr Daniel Furtwängler.	2	1	2
Guten Tag, Frau Ruth Bernhardsgrütter.	3	3	2
Guten Tag, Herr Jakob Waespe.	3	2	0
Guten Tag, Frau Ursula Feierabend.	2	3	3
Guten Tag, Frau Rita Busch.	3	3	3
Guten Tag, Frau Ruth Bärtschi.	3	3	3

A.2.2 French

Guten Tag, Frau Christiane Ansermet.	1	1	1
Guten Tag, Frl. Corinne L'Eplattenier.	0	1	1
Guten Tag, Herr Pierre Borgognon.	1	0	1
Guten Tag, Frau Huguette Banderet.	1	1	1
Guten Tag, Frau Janine Borloz.	1	1	2
Guten Tag, Frau Marie-Thérèse Chetélat.	-	0	0
Guten Tag, Frau Nathalie Corminboeuf.	0	0	0

A.2.3 Italian

Guten Tag, Frau Mirella Anniballo.	2	2	2
Guten Tag, Herr Biagio Zoccolillo.	1	1	1
Guten Tag, Herr Egidio Alfieri.	1	1	1
Guten Tag, Herr Giovanni Berizzi.	1	1	1
Guten Tag, Frau Cristina Capodifoglia.	1	1	1
Guten Tag, Frau Grazia Bresciani.	1	1	1
Guten Tag, Herr Vittorio Bonapace.	1	2	1
Guten Tag, Frl. Ines Giaquinto.	0	0	0
Guten Tag, Herr Renato Bernasconi.	3	2	2

A.2.4 English

Guten Tag, Frau Rosemary Brown.	2	1	1
Guten Tag, Frau Carole McFarlane.	0	0	0
Guten Tag, Herr Philip Goodman.	2	1	2
Guten Tag, Herr Mike Bellhouse.	0	0	1
Guten Tag, Frau Lin Edwards.	1	1	2
Guten Tag, Herr Simon Wheatley.	0	0	0
Guten Tag, Herr James Whittlestone.	1	0	0

A.2.5 Spanish

Guten Tag, Frl. Concepcion Vazquez.	0	0	1
Guten Tag, Herr Joaquim Fortunato.	1	1	1
Guten Tag, Herr Juan Carlos Bailén.	-	0	1
Guten Tag, Herr Javier Alonso-Perez.	1	0	0
Guten Tag, Herr José Antonio Blanco.	-	1	2
Guten Tag, Frl. Georgina Gonçalves.	0	0	0

A.2.6 Greek

Guten Tag, Frau Charula Angelopoulos.	1	2	2
Guten Tag, Herr Leandros Yannakopoulos.	2	1	1
Guten Tag, Frau Gisela Papastergios.	1	2	3
Guten Tag, Herr Dimitrios Andreou.	1	2	3
Guten Tag, Frau Angelika Simantirakis.	1	1	1
Guten Tag, Frau Irini Papazoglou.	2	1	2

A.2.7 Slavic

Guten Tag, Frau Jana Blazek.	2	2	2
Guten Tag, Frau Stefica Gajic.	1	1	1
Guten Tag, Herr Branislav Dimovski.	2	3	2
Guten Tag, Frau Helga Morczinek.	1	1	1
Guten Tag, Frau Danja Krampikowski.	3	2	2
Guten Tag, Frau Katica Lovricic.	1	1	1
Guten Tag, Frau Tanja Maljkovic.	0	1	1
Guten Tag, Frau Teresa Majewski.	0	1	3
Guten Tag, Frau Libusa Kucera.	1	1	1

A.2.8 Hungarian

Guten Tag, Herr Bela Borföi.	1	2	2
Guten Tag, Frau Magdalena Magyar.	1	0	1
Guten Tag, Frau Szilvia Orszagh.	1	1	1

A.2.9 Turkish

Guten Tag, Frau Nuriye Oguz.	1	1	0
Guten Tag, Herr Tarkan Oezdemir.	2	2	1
Guten Tag, Herr Bülent Oezkaynak.	2	1	1
Guten Tag, Frau Gülnur Yücel.	1	1	1

A.2.10 Arabic

Guten Tag, Herr Elmontasser Abdelraouf.	2	1	2
Guten Tag, Herr Moheb Zaki.	1	2	2
Guten Tag, Herr Ahmet Avsar.	1	1	1
Guten Tag, Herr Mohammed Bishtawi.	1	1	2
Guten Tag, Herr Esmail Aghdami.	2	2	2

A.2.11 Finnish

Guten Tag, Frau Kirsi Vättö.	3	2	2

A.2.12 Scandinavian

Guten Tag, Herr Olav Kvaale.	3	3	1
Guten Tag, Herr Björn Bengtsson.	-	3	3
Guten Tag, Herr Jens Bjorheim.	-	3	3
Guten Tag, Herr Sören Bjönness.	-	3	2

A.2.13 Chinese

Guten Tag, Frau Hwei-Ying Wang.	2	1	2
Guten Tag, Frau Mai Phuong.	2	3	1
Guten Tag, Frau Tsering Yangkyi Sesung.	2	1	1

A.2.14 Japanese

Guten Tag, Herr Takashi Harazawa.	2	1	2
Guten Tag, Herr Yoichi Nagakura.	1	0	1

A.2.15 Latin

Guten Tag, Herr Dr. Aleidus Bosman.	3	2	2
Guten Tag, Herr Carl Castritius.	2	3	1
Guten Tag, Herr Pius Nigg.	0	1	2

A.2.16 Compound first names

Guten Tag, Herr Pierre-Yves Bruchez.	0	0	0
Guten Tag, Herr Jean-Gilles Fauquex.	0	0	0
Guten Tag, Herr Daniel-Marc Brunner.	1	1	1
Guten Tag, Herr Hans-Jörg Demuth.	2	2	2
Guten Tag, Herr Jean-Marie Fornerod.	0	0	0
Guten Tag, Herr Hans-Beat Deuber.	2	2	3
Guten Tag, Herr Karl-Heinz Bär.	3	3	3
Guten Tag, Herr Heinz-Dieter Druschke.	3	3	3

A.2.17 Names of aristocratic origin

Guten Tag, Herr Klaus 'von Horsten.	2	2	2
Guten Tag, Frau Vera 'von Planta.	1	1	1
Guten Tag, Herr Jürg von 'Wartburg.	3	3	3
Guten Tag, Frau Rita von 'Holzen.	3	3	3
Guten Tag, Herr Bertrand de Saint Remye.	0	0	0
Guten Tag, Herr Philippe de Weck.	1	1	1
Guten Tag, Herr Gian-Michele a Marca.	0	0	0
Guten Tag, Frau Julia ten 'Elsen.	2	2	2
Guten Tag, Frau Elly van 'Ham.	2	3	1
Guten Tag, Herr Frans van den 'Brul.	2	3	2
Guten Tag, Frau Lesley van 'Rijn.	1	0	0
Guten Tag, Herr Stefan Van der 'Sluijs.	0	0	1
Guten Tag, Herr Patrik 'Auf der Mauer.	1	1	2
Guten Tag, Herr Alex 'Auf der Maur.	2	2	2

A.2.18 Names with titles

Guten Tag, Herr Dr. Niklaus Ammann.	2	2	3
Guten Tag, Herr Dr. Olivier Cuendet.	0	0	1
Guten Tag, Herr Dr. von Schwarzenberg.	3	3	3
Guten Tag, Herr Dr. Marco Finardi.	3	3	2
Guten Tag, Herr Dr. Hans-Peter Bauer.	3	3	3
Guten Tag, Herr Dr. Heinz Boppart.	3	3	2
Guten Tag, Herr Prof. Hans-Peter Frei.	3	3	3
Guten Tag, Herr Prof. Dr. Hans-Peter Frei.	3	3	3

German	**25**	**26**	**23**	**(33)**
French	**4**	**4**	**6**	**(21)**
Italian	**11**	**11**	**10**	**(27)**
English	**6**	**3**	**6**	**(21)**
Spanish	**2**	**2**	**5**	**(18)**
Greek	**8**	**9**	**12**	**(18)**
Slavic	**11**	**13**	**14**	**(27)**
Hungarian	**3**	**3**	**4**	**(9)**
Turkish	**6**	**5**	**3**	**(12)**
Arabic	**7**	**7**	**9**	**(15)**
Finnish	**3**	**2**	**2**	**(3)**
Scandinavian	**3**	**12**	**9**	**(12)**
Chinese	**6**	**5**	**4**	**(9)**
Japanese	**3**	**1**	**3**	**(6)**
Latin	**5**	**6**	**5**	**(9)**
Compound First Names	**11**	**11**	**12**	**(24)**
Names of aristocratic origin	**20**	**21**	**11**	**(42)**
Names with Titles	**20**	**20**	**20**	**(24)**
Ratings Total	**154**	**158**	**167**	**(330)**
Mean Rating Total	**1.40**	**1.44**	**1.52**	
N. of Correct or Acceptable	**47**	**48**	**34**	**(110)**

B Material to the Pronunciation Error Test

B.1 Vowels and diphthongs[18]

A "√" means a correct pronunciation of the sound tested and an "x" means an incorrect pronunciation.

B.1.1 a-sounds

		LHS	IVX	Elan	
[ɑː]	Staat	√	√	√	
	fraßen	×	√	×	
	Lebertran	√	√	×	
	Grab	√	√	√	
	Mühsal	√	√	×	
	tags	√	√	√	
	Gas	×	√	√	
	brachliegen	√	×	√	
	Jagd	√	√	×	
	Papst	√	√	×	
[ɑ]	Amt	√	√	√	
	Scharte	√	√	×	
	Walfisch	√	√	√	
	Bräutigam	×	√	×	
borrowings with [ɑː]	Natrium	×	√	√	
	Dramatik	×	√	√	
	Melodram	×	×	×	
	Roman	√	√	√	
	Pirat	×	√	√	
	Bratsche	×	×	×	
	Spanien	×	√	×	
	Memoiren	×	×	×	
borrowings with [ɑ]	Aroma	√	√	√	
	Karneval	×	×	√	
	Biwak	×	×	√	
	Atlas	√	×	√	
	Januar	√	√	×	
	Kap	√	√	×	
	Tram	×	×	√	
	Toilette	×	×	√	
	loyal	×	×	×	
	NON-BORROWINGS	**3**	**1**	**7**	**(14)**
	BORROWINGS	**12**	**9**	**7**	**(17)**
	TOTAL	**15**	**10**	**14**	**(31)**

[18] For each vowel, there are separate entries for the borrowings so that the underlying regularities can be detected more easily.

B.1.2 e-sounds

		LHS	IVX	Elan
[eː]	Teer	√	√	×
	Efeu	√	√	√
	Erde	×	√	√
	Pferd	×	√	×
	Krebs	√	√	√
	Lebkuchen	√	√	√
	Magnet	×	×	×
	nebst	√	√	√
	stets	√	√	√
[ɛː][19, 20]	Ähre	×	√	×
	schräg	√	√	√
	Späße	×	√	√
	Rätsel	×	√	√
	Gemälde	×	√	√
	Städte	×	×	√
	nämlich	√	√	×
	hätscheln	×	×	×
	zärtlich	√	×	√
	erklärt	√	√	√
[ɛ][21]	Held	√	√	√
	Herzog	×	×	×
	Esche	√	√	√
	häßlich	√	√	√
	fertig	×	√	×
	älter	√	√	√
	weg	×	×	×
	Verse	×	√	√
[ə]	lebe	√	√	√
	wahren	√	√	√
	Atem	√	√	√
borrowings with [ɛː]	Ära	×	√	×
	Sphäre	×	√	×
borrowings with [eː]	Allee	√	√	√
	Juwel	×	×	×
	Prophet	√	√	×
	Problem	√	√	√
	Bibliothek	√	√	√
	Phänomen	√	√	√
borrowings with [ɛ]	Seg men t	×	×	×
	Echo	√	√	√
	Gabriel	×	×	×
	Hotel	√	√	×
	Requiem	×	×	√

[19] The [ɛː] of Elan is generally too open.

[20] LHS does not make a sensible difference between [ɛː] and [eː], and therefore there is less potential for errors.

[21] IVX distinguishes [ɛ] from [e] in the phonetic trace although [e] is not part of the German vowel system. However, as the realizations of the two phonemes are identical, there was no need to count errors of this type.

borrowings with [e]	Benefiz	×	×	×	
	Ökumene	×	×	×	
borrowings with [ə]	Saurier	√	×	√	
	Spanier	√	√	√	
	NON-BORROWINGS	**13**	**6**	**9**	**(30)**
	BORROWINGS	**8**	**7**	**9**	**(17)**
	TOTAL	**21**	**13**	**18**	**(47)**

B.1.3 i-sounds

[iː]	Biber	×	√	√	
	Igel	×	√	×	
	Nische	×	×	×	
[ɪ]	selig	√	√	√	
	Kürbis	√	√	√	
	Himbeere	√	×	√	
	Rettich	√	√	√	
	Viertel	×	√	√	
[i]	vielleicht	×	√	√	
borrowings with [iː]	studieren	×	√	×	
	Liter	×	√	√	
	Agronomie	√	√	√	
	Akkusativ	√	√	√	
	Scharnier	√	√	√	
	Fabrik	×	×	×	
	massiv	√	√	√	
	Türkis	×	×	×	
	Tweed	×	×	×	
borrowings with [ɪ]	Christ	√	√	√	
	Sellerie	×	×	×	
	Ethik	√	√	√	
	Fazit	√	√	×	
	Ysop	√	×	×	
	Seldwyla	×	×	×	
borrowings with [i]	Linie	√	√	×	
	Italien	√	√	√	
	NON-BORROWINGS	**5**	**2**	**2**	**(9)**
	BORROWINGS	**7**	**6**	**9**	**(17)**
	TOTAL	**12**	**8**	**11**	**(26)**

B.1.4 o-sounds

[oː]	Moor	√	×	√
	bloß	√	√	×
	Ostern	√	×	×
	Propst	×	√	×
	Lotse	×	×	√
	vorwärts	√	√	√
	Herzog	√	√	√
	Vogt	×	√	√
	Dom	×	√	√

		LHS	IVX	Elan	
[ɔ]	Geschoß	√	√	√	
	Most	√	√	√	
	Brombeere	√	×	×	
	Hochzeit	×	√	×	
borrowings with [oː]	Domino	√	√	√	
	Jod	√	√	√	
	Symptom	√	√	√	
	Apostroph	√	×	×	
	reziprok	√	×	√	
	Baron	√	√	×	
	Meteor	×	×	×	
	grandios	√	×	×	
borrowings with [o]	Obelisk	×	×	√	
	Protest	√	√	√	
	Silo	√	√	√	
borrowings with [ɔ]	Kanon	×	×	√	
	Marmor	√	×	√	
	Pathos	×	√	√	
	Lektor	×	×	√	
	Grog	×	×	√	
	NON-BORROWINGS	**5**	**4**	**5**	**(13)**
	BORROWINGS	**6**	**9**	**4**	**(16)**
	TOTAL	**11**	**13**	**9**	**(29)**

B.1.5 ö-sounds

		LHS	IVX	Elan	
[øː]	öde	√	√	√	
	Erlös	√	√	√	
	Größe	√	√	√	
	höchst	√	√	√	
	Behörde	×	×	×	
	Gehöft	×	√	×	
[œ]	öfters	√	√	√	
	Schößling	×	×	×	
	röcheln	×	√	√	
	östlich	√	√	√	
	Böschung	√	√	√	
borrowings with [øː]	Milieu	×	×	×	
borrowings with [œ]	Zölibat	×	√	×	
	Ökumene	×	√	√	
borrowings with [œː]	Deserteur	×	√	√	
	Kontrolleur	×	√	√	
	NON-BORROWINGS	**4**	**2**	**3**	**(11)**
	BORROWINGS	**5**	**1**	**2**	**(5)**
	TOTAL	**9**	**3**	**5**	**(16)**

B.1.6 u-sounds

		LHS	IVX	Elan	
[uː]	duzen	×	√	√	
	Spuk	√	√	√	
	Armut	√	√	√	
	Muße	×	×	√	
	ruchlos	√	×	√	
	Schuster	×	×	√	
	Geburtstag	×	×	√	
	Wuchs	×	√	×	
	Ursprung	√	×	√	
[ʊ]	Genuß	√	√	√	
	Spruch	×	√	√	
	Furt	√	√	√	
	flugs	×	×	×	
	Urteil	√	√	×	
borrowings with [uː]	Gnu	√	√	×	
	Eunuch	×	√	√	
	immun	×	×	√	
	Matur	√	√	√	
	Rekrut	√	√	√	
	Konsum	×	√	√	
	Route	×	×	√	
borrowings with [u]	Statue	×	×	√	
borrowings with [ʊ]	Sirup	×	√	√	
	minus	√	√	×	
	Konsul	×	×	×	
	Justiz	√	√	√	
	Klub	√	×	√	
	Rum	√	√	√	
	Radium	√	√	√	
	NON-BORROWINGS	**7**	**6**	**3**	**(14)**
	BORROWINGS	**7**	**5**	**3**	**(15)**
	TOTAL	**14**	**11**	**6**	**(29)**

B.1.7 ü-sounds

		LHS	IVX	Elan
[yː]	üben	√	√	√
	schwül	√	√	√
	Gemüt	√	√	√
	hochmütig	√	√	√
	süß	×	√	√
	Küchlein	√	×	×
	düster	×	√	√
	Rüsche	×	×	×
[ʏ]	mürbe	√	√	√
	Büsche	√	√	√
	rüsten	√	√	√
	Gelübde	×	√	√
borrowings with [yː]	Asyl	√	×	√
	Zypern	√	√	√
	Etude	×	×	×

		LHS	IVX	Elan	
borrowings with [y]	Büro	×	√	√	
	kostümieren	×	√	√	
	Labyrinth	×	√	√	
	Zypresse	×	×	√	
	Nuance	×	×	√	
	Rhythmus	√	√	×	
	lynchen	√	×	√	
	Ypsilon	√	√	√	
	NON-BORROWINGS	**4**	**2**	**2**	**(12)**
	BORROWINGS	**6**	**5**	**2**	**(11)**
	TOTAL	**10**	**7**	**4**	**(23)**

B.1.8 Diphthongs and voice onset

		LHS	IVX	Elan	
[ɑe, ɑɪ]	bleiben	√	√	√	
	Hai	√	√	√	
	Laich	√	√	√	
	Tokayer	√	√	×	
	Nylon	×	×	√	
[ɑo, ɑʊ]	rauh	√	√	√	
	Clown	×	×	×	
	Couch	×	×	×	
[ɔø, ɔɪ]	Heu	√	√	√	
	läuten	√	√	√	
	Pseudonym	√	√	√	
	Boykott	√	×	×	
no diphthongs	archa-isch	×	√	√	
	Ha-iti	×	×	×	
	Alle-uten	×	×	×	
	Tede-um	√	×	√	
	Jubilä-um	√	√	×	
hiatus	O|ase	×	×	×	
	Ka|in	×	√	×	
	Kre|usa	×	×	×	
new voice onset	Ver|ein	√	√	√	
	über|all	√	√	√	
	ver|alten	√	√	√	
no new voice onset	hinab	√	√	×	
	vollends	√	×	√	
	voraus	√	√	√	
	einander	√	√	×	
	Initialen	×	√	√	
	DIPHTHONGS	**3**	**4**	**4**	**(12)**
	NO DIPTHONGS, HIATUS, VOICE ONSET	**7**	**6**	**8**	**(16)**
	TOTAL	**10**	**10**	**12**	**(28)**
	NON-BORROWINGS TOTAL	**41**	**23**	**31**	**(103)**
	BORROWINGS TOTAL	**51**	**42**	**36**	**(98)**
	DIPHTHONGS ETC. TOTAL	**10**	**10**	**12**	**(28)**
	VOWELS & DIPHTHONGS TOTAL	**102**	**75**	**79**	**(229)**

B.2 Consonants

B.2.1 Liquids and nasals

		LHS	IVX	Elan	
[ʀ]	Rhythmus	√	√	√	
	Katarrh	×	√	√	
<gm, gn>	Pigment	×	√	√	
	Magnat	√	√	√	
	Signal	√	√	√	
[ŋ]	langsam	√	√	√	
	Angst	√	√	√	
	fängt	√	√	√	
[ŋk, ŋg, ŋks]	Dank	√	√	√	
	Ungar	×	×	×	
	Mangan	×	√	×	
	Sphinx	√	√	×	
<ng, nk> without [ŋ]	konkav	×	√	×	
	inkognito	√	×	√	
	unklar	√	√	√	
	Anklang	√	√	×	
[ɲ]	Kognak	×	×	√	
	Kastagnetten	×	×	√	
	LIQUIDS AND NASALS TOTAL	**7**	**4**	**5**	**(18)**

B.2.2 Fricatives

		LHS	IVX	Elan
[h]	heiser	√	√	×
	zugehörig	√	√	×
	Ahorn	√	√	×
	Hoheit	√	√	√
	Alkohol	√	√	√
	Hausherr	√	√	√
	Hotel	√	√	√
mute <h>	nahe	×	√	√
	ruhig	×	√	√
	Ehe	√	√	√
	stehen	√	√	√
	Hypot<u>h</u>ese	√	√	√
	Walther	×	√	√
<v> as [f]	Vater	√	√	√
	Vesper	×	×	√
	Vögel	√	√	×
	Veilchen	×	×	√
	Frevel	×	×	×
	Nerven	√	√	√
	brav	√	√	√
	aktiv	√	√	√
<v> as [v]	Vasall	√	√	√
	Verben	×	√	√
	Sklave	×	√	√
	bravo	×	√	√
	mißverstehen	√	√	×
	nervig	√	√	×

		LHS	IVX	Elan
<ph> as [f]	Philosophie	√	√	√
	Aphorismus	√	√	√
<pf>:	Pferd	√	√	√
	Dampf	√	√	√
<ps>:	Psi	√	√	×
<u> after cons as [v]	Requiem	√	√	√
	Qual	√	√	√
	Etui	×	×	×
[z]	Gesang	√	√	√
	lose	√	√	√
	Amsel	√	√	√
	Salat	√	√	√
	Symbiose	√	√	√
	System	√	√	√
	Prosa	√	√	√
	Bazar	×	×	×
[s]	Hals	√	√	√
	raspeln	√	√	√
	Röslein	√	√	√
	Smaragd	√	√	√
	Skizze	√	√	√
	präzis	√	√	√
	Nuance	√	×	×
<s + -chen>	Häuschen	×	√	×
	Röschen	×	×	×
[ts]	Kranz	√	√	√
	Skizze	×	√	√
[sts]	Szene	√	√	√
	obszön	√	√	√
<t> before <i> as [ts]	Nation	√	√	√
	partiell	×	√	√
[ks]	Axt	√	√	√
	weckst	√	√	√
	klügste	√	√	√
	Dachse	√	×	×
	Deichsel	×	×	×
<st> as [ʃt], <sp> as [ʃp]	Gestein	√	√	√
	Schachspiel	×	√	√
	verspielen	√	√	√
	Frühstück	√	√	×
	Ursprung	√	×	√
	Stadion	√	√	√
	spekulieren	√	√	√
	spezifisch	√	√	√
	Bevölkerungsstatistik	√	√	√
<st> as [st], <sp> as [sp]	Korrespondent	√	√	√
	Abstinenz	×	√	√
	Perspektive	√	√	×
	Apostroph	√	√	×
[ʃ]	Ski	×	×	×
	Chiffon	×	×	×
	Recherchen	×	×	×
	Böschung	√	√	×

		LHS	IVX	Elan	
[tʃ]	Chinchilla	×	×	×	
	Boccia	×	×	×	
[ʒ]	Garage	×	×	×	
	Genie	×	×	√	
	Regie	√	√	×	
	Jury	×	×	×	
	Giro	×	×	×	
[dʒ]	Gin	×	×	√	
	Arpeggien	×	×	×	
[ç]	nächste	×	√	√	
	Köcher	√	√	√	
	mancher	√	√	√	
	Molch	√	√	√	
	Storch	√	√	√	
	Frauchen	×	×	×	
	ewig	√	√	√	
	freudigste	√	√	√	
	Chemie	√	√	√	
	Chinese	√	√	√	
	Melancholie	√	√	√	
[x]	wachst	√	×	×	
	Epoche	√	√	√	
[j]	Joch	√	√	√	
	Juli	√	√	√	
	York	√	×	×	
	Yacht	√	×	×	
[lj]	brillant	×	×	×	
	Billard	×	×	×	
	<h>	**3**	**0**	**3**	**(13)**
	<V, PH, PF, PS, QU>	**7**	**4**	**6**	**(22)**
	[z, s, ts, ks]	**6**	**5**	**6**	**(28)**
	<ST, SP>	**2**	**1**	**3**	**(13)**
	[ʃ, tʃ, ʒ, dʒ]	**11**	**11**	**11**	**(13)**
	[ç, x, j, lj]	**4**	**6**	**6**	**(19)**
	FRICATIVES TOTAL	**33**	**27**	**35**	**(108)**

B.2.3 Plosives

loss of voice, final pos.	grob	√	√	√
	Rad	√	√	√
	Obmann	×	×	×
	Kindheit	√	√	√
	Schlagzeile	√	√	√
	Weisheit	√	√	√
	gibt	√	√	√
voiced	Gegend	√	√	√
	übler	√	√	√
	regnet	√	√	√

<-ig> as [-ɪg] or [-ɪk]	königlich	×	√	√	
	ewiges	√	√	√	
	Vierziger	√	√	√	
	PLOSIVES TOTAL	**2**	**1**	**1**	**(13)**
	CONSONANTS TOTAL	**42**	**32**	**41**	**(139)**

B.3 Stress[22]

B.3.1 Simple lexemes and compounds

		LHS	IVX	Elan	
simple lexemes	le'bendig	√	√	×	
	Fo'relle	√	√	√	
	Ho'lunder	√	√	√	
	Wa'cholder	√	√	√	
	Herme'lin	×	×	×	
compounds	will'kommen	×	×	√	
	Lebe'wohl	×	×	×	
	TOTAL	**3**	**3**	**3**	**(7)**

B.3.2 Derived items

		LHS	IVX	Elan
derived with <da->	'dazumal	×	×	×
	'Dasein	×	√	×
	da'hinter	√	√	√
	da'rinnen	√	×	√
derived with <durch->	'Durchblick	×	√	√
	'Durchmesser	×	√	√
	'durchhalten	×	√	√
	'durchzählen	×	√	√
	Durch'leuchtung	√	×	×
	durch'trieben	√	√	×
	durch'queren	×	×	×
derived with <her->	'herstellen	√	√	√
	her'über	√	√	√
	her'bei	√	√	√
derived with <hin->	'hinlänglich	√	√	√
	Hinfahrt	√	√	√
	hi'nab	×	√	√
derived with <hinter->	'Hintergebäude	×	√	√
	'hinterrücks	×	×	√
	hinter'gehen	√	×	√
derived with <in->	'Inhaber	√	√	√
	'Inland	√	√	√
	In'kraftsetzung	×	√	√
	In'anspruchnahme	√	√	√
derived with <miß->	'Mißernte	√	√	√
	'Mißverhältnis	√	√	×
	miß'achten	×	√	×
	miß'gönnen	×	×	√
	'mißverstehen	√	√	×
derived with <über->	'Überbleibsel	×	×	√
	'Übereifer	×	×	√
	'Überschall	×	×	√
	'überglücklich	×	×	√
	'überschießen	×	×	√
	über'lebendsgroß	√	√	√
	über'völkert	√	√	×

22 Stress correctness in borrowings was not tested.

	über'hitzen	√	√	×	
	über'wältigen	√	√	√	
derived with <um->	'Umweg	√	√	√	
	'Umriß	√	√	√	
	'umwerfen	√	√	×	
	'umändern	√	√	√	
	Um'gebung	√	√	√	
	um'fluten	×	×	×	
	um'kreisen	×	×	×	
derived with <un->	'Ungeziefer	√	√	√	
	'Unwetter	√	√	√	
	'uneins	√	√	√	
	'unangebracht	√	√	√	
	'ungebräuchlich	√	√	√	
	Unbe'rechenbarkeit	×	×	×	
	Un'heilbarkeit	×	×	×	
derived with <unter->	'Unterbegriff	×	√	√	
	'Unterbewußtsein	×	√	√	
	'untergründig	×	×	√	
	'unterbelichten	×	×	√	
	Unter'drückung	√	√	√	
	Unter'malung	√	√	×	
	unter'haltsam	√	√	×	
	unter'suchen	√	√	√	
derived with <voll->	'Volldampf	√	√	√	
	'Vollmilch	√	√	√	
	'volljährig	√	√	√	
	voll'bringen	×	×	×	
	'vollends	√	×	√	
derived with <vor->	'Voralpen	√	√	√	
	'Vorfrühling	√	√	√	
	'voreingenommen	√	×	√	
	'vorgestern	√	√	√	
	vor'trefflich	×	×	×	
	vor'züglich	×	×	×	
derived with <wider->	'Widerhall	√	×	√	
	'Widerlager	√	×	√	
	'Widerwille	√	×	√	
	'widerborstig	√	×	√	
	wider'fahren	×	√	×	
	wider'streben	×	√	√	
	wider'legen	×	√	×	
derived with <wieder->	'Wiederwahl	√	√	√	
	Wiederinbe'sitznahme	×	×	×	
	'wiedervereinigen	√	√	√	
derived with <zu->	'Zufahrt	√	√	√	
	'Zunahme	√	√	√	
	'zugehörig	√	√	√	
	'zureden	√	√	√	
	Zu'hause	×	√	×	
	zu'hinterst	×	×	√	
	zu'mindest	√	√	√	
	TOTAL	**35**	**29**	**24**	**(88)**

STRESS TOTAL	38	32	27	(95)
PRONUNCIATION TEST TOTAL	**182**	**139**	**147**	**(463)**

B.4 Eliminated Test Items

<Fährte>: today [ɛː] instead of [eː]
<Rebhuhn>: [e] instead of [ɛ]
<Kongreß>: today [ŋg] instead of [ng]
<Gotthard>: today with [h] instead of without

Table B1. Items which were eliminated from the original test set.

C Material for the Subjective Evaluation

C.1 Translations of the text used for the evaluation

The translations are adaptations of the original banking advertisement leaflets, reflecting the changes which were appropriate for the German text in order to be used in the subjective evaluation.

Vous pouvez connaître l'état de vos finances à toute heure. Avec la Liberty-Line.

Un service téléphonique réservé aux clients Liberty. Cette ligne directe vous permet de connaître à toute heure le solde de vos comptes Liberty et les cinq dernières opérations effectuées sur chacun d'eux. Simple, commode, sûr et avantageux. 33,3 centimes la minute seulement. Tarif juin 94.

Rien de plus facile. Pour appeler Liberty-Line, il suffit de composer le 157, 03, 45, 1. Encore une fois : 157 03 45 1. Puis composer votre numéro client et votre code confidentiel CIP. A partir de là, Liberty-Line est tout à votre service.

Si vous n'avez pas de téléphone à fréquences acoustiques, il vous faut un sélecteur: le Liberty-Link. L'UBS vous le fournit gratuitement sous forme de porte-clés.

Figure D1. French translation.

Potete informarvi 24 ore su 24 sul vostro denaro. Colla Liberty-Line !

La Liberty-Line è un servizio telefonico esclusivo a disposizione solo dei clienti Liberty. In tal modo siete in collegamento diretto con tutti i vostri conti Liberty e potete conoscere ovunque e in ogni momento il saldo disponibile e gli ultimi cinque movimenti registrati sui conti. Un sistema comodo, sicuro, semplice e conveniente. Infatti pagate solo la tariffa PTT: 33,3 centesimi al minuto. Dati del giugno 94.

E' molto facile usarla: basta fare il numero della Liberty-Line 157, 03, 45, 3. Ancora: 157 03 45 3. Quindi comporre il numero personale di cliente e il codice PIN. E la Liberty-Line provedde a tutto il resto automaticamente.

Se il vostro telefono non fosse provvisto di selezione a frequenza vocale, è sufficiente avere un piccolo strumento, Liberty-Link, che potete ottenere gratuitamente in forma di portachieve presso qualsiasi sportello UBS.

Figure D2. Italian translation.

C.2 Recording configuration and procedures

C.2.1 Lernout & Hauspie (1 step)

- Text: C:\tts\texts\g_lhs.txt
- LHS: default values, but volume to max, freq = 11.025 kHz
- PC output: upper
- Volume PC: exactly medium, left/right = balanced
- DAT input: mic, mic sens = L, rec mode = manual, rec level = max
- DAT quality: SP
- DAT playback volume for telephone test: 15

C.2.2 Infovox 2 (steps)

Step 1 (board to PC)

- Text: C:\tts\texts\g_ivx.txt
- Board output: line
- voxtalk: default values (speaker, pitch, vol, rate)
- PC input: mic, lower
- Quick Recorder: quality CD (44 kHz, PCM 16 bit), version 2.0, mono
- Recording control: mic = 3 steps from bottom

Step 2 (PC to DAT)

- PC output: upper
- Volume PC: exactly medium, lef/right = balanced
- DAT input: mic, mic sens = H, rec mode = manual, rec level = 4
- DAT quality: SP
- DAT playing volume for telephone test: 14

C.2.3 Elan (2 steps)

Step 1 (board to first DAT)

- Text: written in 2 portions into the demo text windows (window too small)
- Board: volume as shipped
- Board out: jack 3.5 mm
- ALLVOC demo options: vol = 128, pitch = 100, time = 100 (defaults)
- DAT input: mic, mic sens = L, rec mode = manual, rec level = 6
- DAT quality: SP

Step 2 (first DAT to second DAT using Sony RMR-D3 digital I/O adapter)

- Volume: 1:1
- DAT quality: SP
- DAT playing volume for telephone test: 16

C.3 Complete answer data

	TTS system	Sex	Age class	Techn. interest	Prev. exposure	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1	LHS	F	≤ 20	N	S	2	4	4	4	1	2	1	4	2	1
2	LHS	F	20-30	N	N	3.5	4.5	3	5	2	2	3	5	5	5
3	LHS	F	20-30	N	N	5	4	4	6	2	1	1	5	1	1
4	LHS	F	30-40	N	N	5	6	4.5	4	3	3	4	4	5	4
5	LHS	F	40-50	N	N	1	3	3	3	3	4	3	3	1	1
6	LHS	F	40-50	N	N	3	4	3.5	2	2	2	2	6	1	5
7	LHS	F	60-70	N	N	5	5.5	4	4	5.5	3	4	4.5	noa	5.5
8	LHS	M	≤ 20	S	M	6	5.5	4.5	5	4	4.5	5	6	5	6
9	LHS	M	20-30	N	S	5	6	4	4	5	4.5	4	5	5	5
10	LHS	M	20-30	N	S	5	4.5	5	4	1	1	1	noa	3.5	1
11	LHS	M	30-40	N	S	4.5	5	3	4.5	4.5	2	2	4	5	5
12	LHS	M	40-50	N	S	5	5	5	6	3.5	3.5	3	5	4	5.5
13	LHS	M	50-60	T	M	5.5	5.5	5.5	3	5	4	4	6	6	6
14	LHS	M	60-70	N	N	5	6	4	4	4	4	3	4	3	5
15	IVX	F	20-30	N	S	4	5	4	5	3	3	3	5	6	5.5
16	IVX	F	20-30	N	N	4	5.5	5	3	1	2	2	6	4	3.5
17	IVX	F	20-30	N	S	2	1	2	5	2	2	5	1	1	2
18	IVX	F	30-40	N	N	6	6	4	6	5	5	6	5	6	6
19	IVX	F	40-50	N	N	4	6	4	6	5	2	4	3.5	1	1
20	IVX	F	60-70	N	N	2	3	2	4	4	2	1	2	1	5
21	IVX	F	≥ 70	N	N	2	3	3	6	6	4	1	6	1	1
22	IVX	M	≤ 20	S	S	4	4	3	5	4	3	4	4.5	4	5
23	IVX	M	20-30	T	M	2	4	2	6	3.5	1	1	2.5	1	6
24	IVX	M	30-40	T	M	4	5	3	5	2	2	1	3	4	2
25	IVX	M	40-50	S	S	4	5	4.5	6	noa	4	6	6	6	6
26	IVX	M	40-50	T	S	3	5	2	3	3	2	2	2	2	5
27	IVX	M	40-50	T	M	3	6	4	5	4	2	2	6	3	5.5
28	IVX	M	60-70	T	N	5	4	5	6	6	5	6	6	6	6
29	Elan	F	≤ 20	N	S	4	5	3	4	4.5	3.5	4	6	3.5	5
30	Elan	F	20-30	N	N	2	4	2	5	1	1	1	1	1	1
31	Elan	F	30-40	N	N	1.5	4	2	5	5	2.5	2	2	1	1
32	Elan	F	40-50	N	N	5	5	5	5	5.5	5.5	5	6	6	3
33	Elan	F	40-50	N	S	2	6	1	6	1	1	2	3	6	1
34	Elan	F	50-60	N	N	1	1	4	3.5	5	5.5	5.5	3	6	6
35	Elan	F	≥ 70	N	N	5	6	4	4	6	5	6	4	5	6
36	Elan	M	20-30	N	N	2	4	3	5	3	4	3	4	5	1
37	Elan	M	20-30	S	S	4	4	4	6	4	3	2.5	6	1	1
38	Elan	M	20-30	N	N	2.5	2.5	2	3.5	2.5	2	1	2	2	3.5
39	Elan	M	30-40	N	S	2	4	3	5	4.5	4	3	4	2	2
40	Elan	M	30-40	T	S	3	5	3	5	4	2	3	4	5	6
41	Elan	M	30-40	T	S	2.5	3	4	4	3	2	1	5	3	6
42	Elan	M	30-40	T	S	3	3	3	5	3	2	2	2	2.5	3

Table C3. Complete answer data
(Technical interest: N = "not at all", S = "somewhat", T = "very technical";
Previous exposure: N = "no", S = "some", M = "many synthesizers";
Questions: noa = "no answer").

C.4 Distributions of the answer data

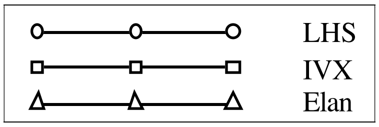

Table C4. Legend.

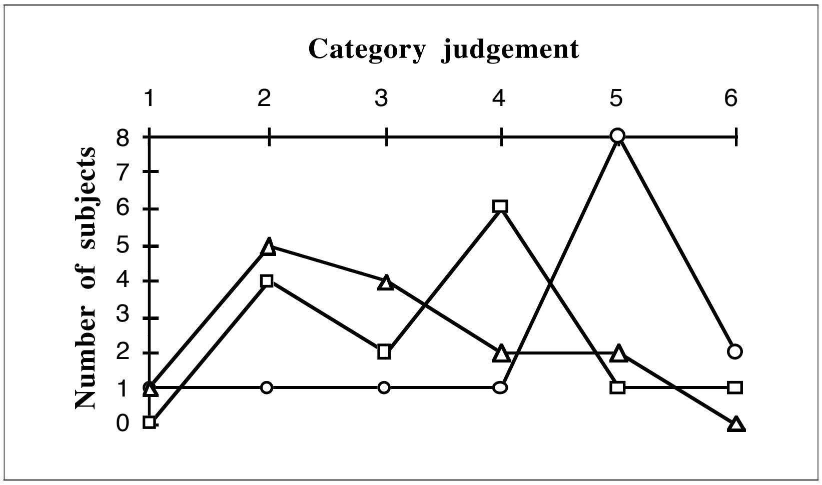

Figure C5. Listening effort (Q1).

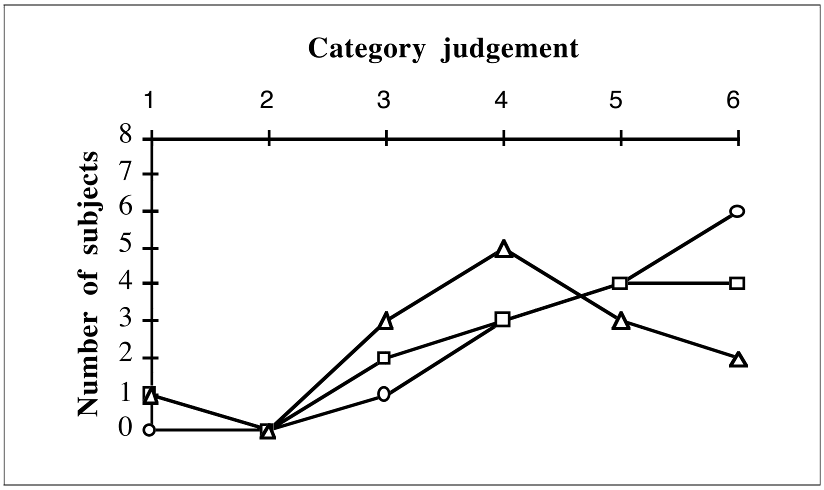

Figure C6. Comprehensibility (Q2).

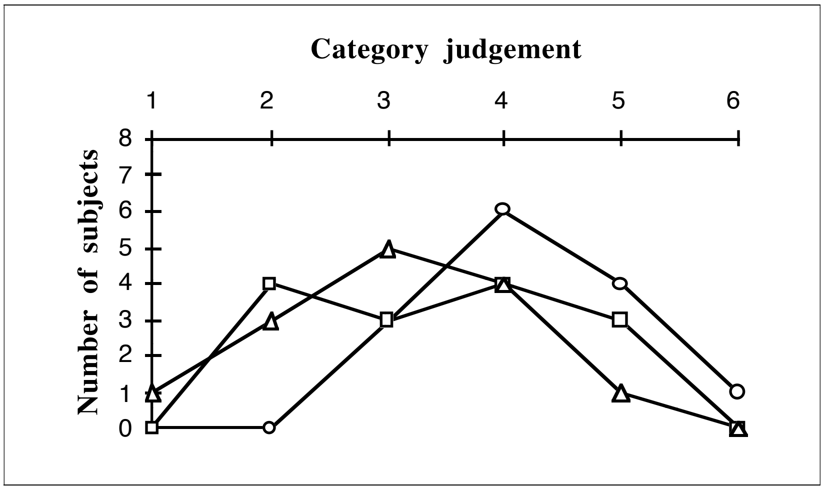

Figure C7. Articulation (Q3).

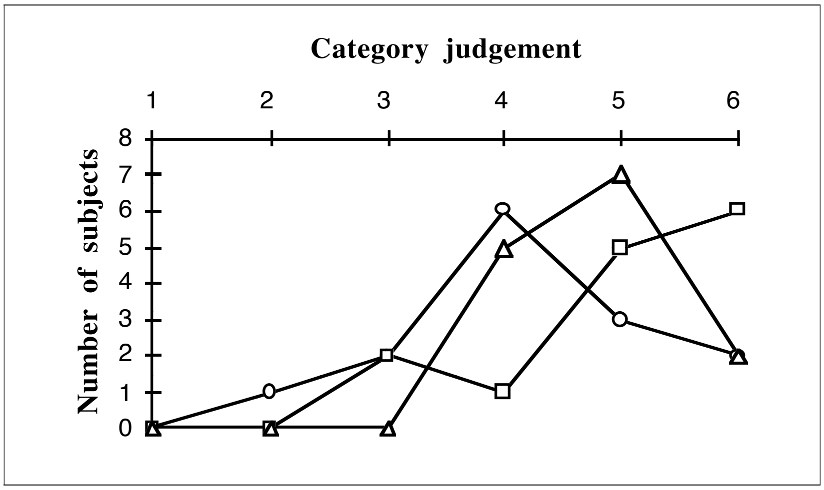

Figure C8. Rate (Q4).

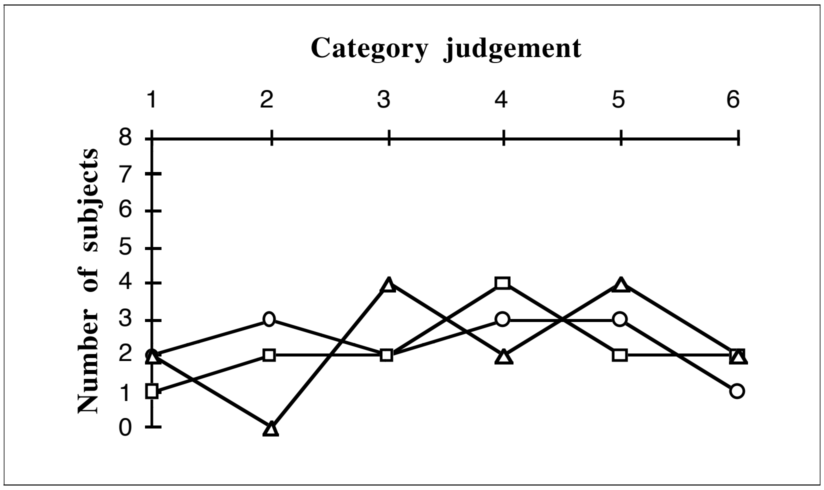

Figure C9. Prosody (Q5).

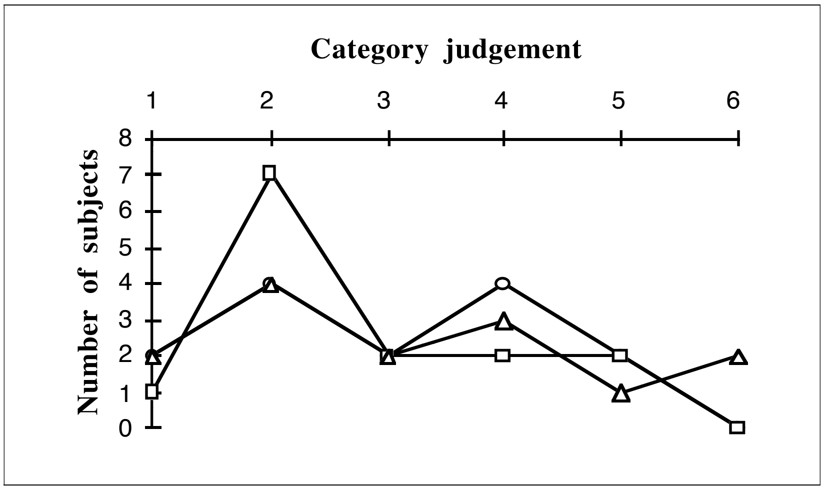

Figure C10. Naturalness (Q6).

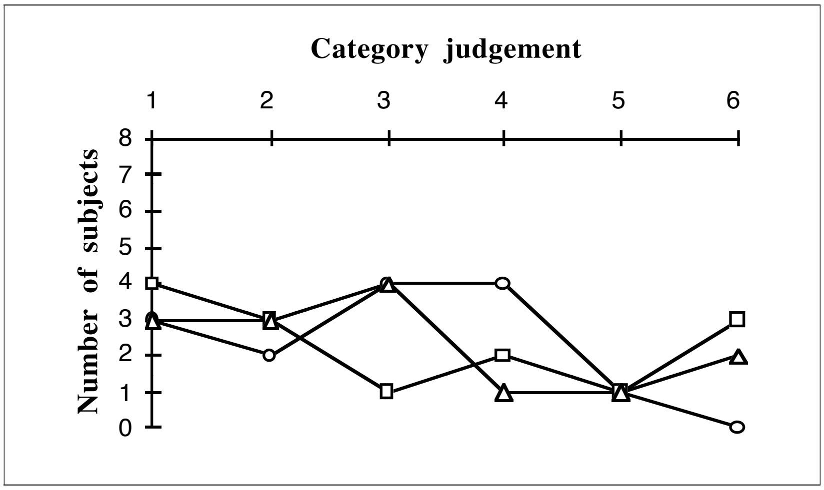

Figure C11. Pleasantness (Q7).

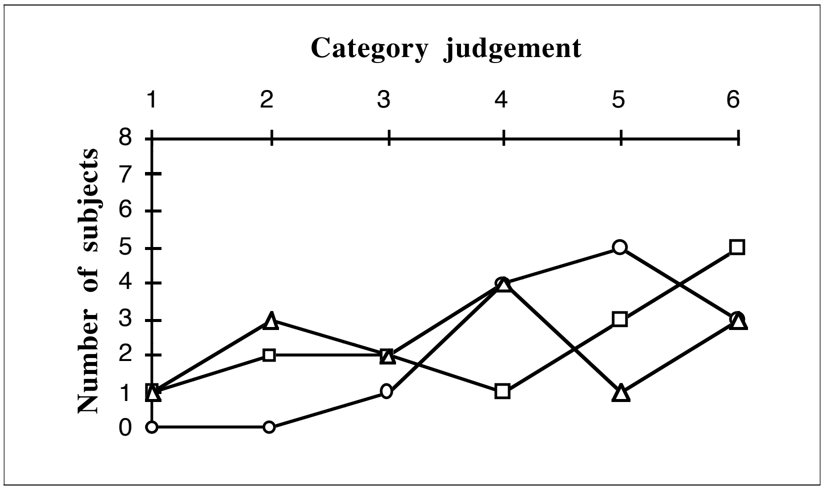

Figure C12. Foreign accent (Q8).

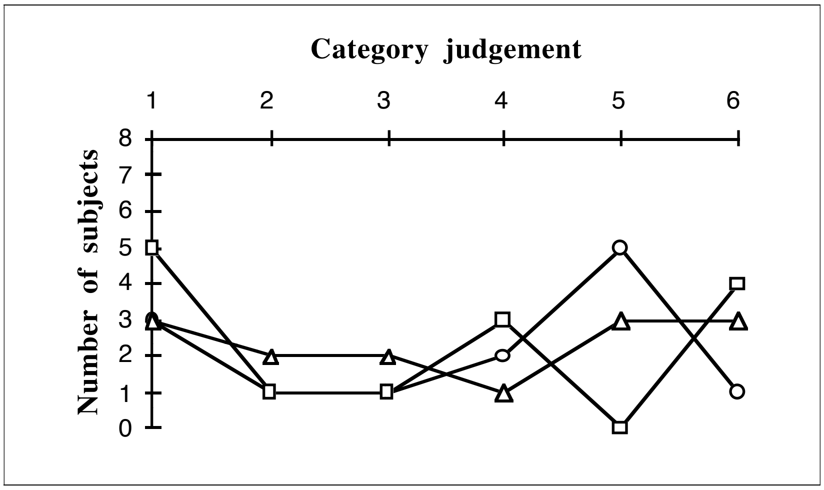

Figure C13. System acceptability (Q9).

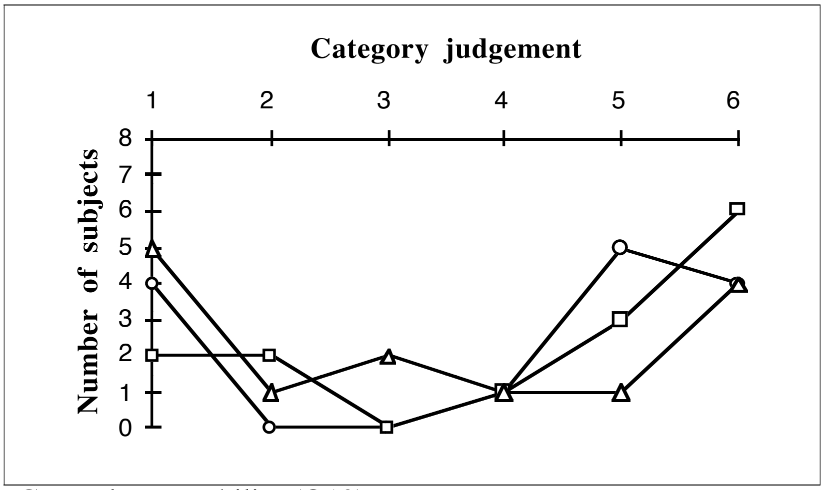

Figure C14. General acceptability (Q10)