

Perzeptive Evaluierung standardisierter und nicht-standardisierter Diphone in synthetisierten Wörtern und Sätzen

Citation for published version (APA):

Rump, W. (1989). *Perzeptive Evaluierung standardisierter und nicht-standardisierter Diphone in synthetisierten Wörtern und Sätzen*. (IPO-Rapport; Vol. 706). Instituut voor Perceptie Onderzoek (IPO).

Document status and date:

Gepubliceerd: 10/07/1989

Document Version:

Uitgevers PDF, ook bekend als Version of Record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Rapport no. 706

Perzeptive Evaluierung standardisierter
und nicht-standardisierter Diphone in
synthetisierten Wörtern und Sätzen

W. Rump

Perzeptive Evaluierung standardisierter und
nicht-standardisierter Diphone in synthetisierten Wörtern
und Sätzen*

Willem Rump

30. Juni 1989

*Bericht eines ersten Experiments im Rahmen meiner Tätigkeit als wissenschaftlicher Praktikant am IPO, 1989.

Abstract

Mittels zweier kurzer Perzeptionstests habe ich festgestellt, ob die Standardisierung der Dauern deutscher Diphone einen Einfluß auf die Qualität der Sprachsynthese hat.

Die Diphone waren so standardisiert, daß die Dauer jedes Phonemanteils des Diphons einen vorher festgelegten Wert bekam. Die Standardisierung der Diphondauern führte im vorliegenden Diphonmaterial zu kürzeren Diphonen. Dadurch wurde segmentelle Information aus dem Diphon entfernt. Wenn ein Diphon aus dem ursprünglichen Diphonmaterial zu kurz war, wurde dafür ein neues Diphon segmentiert mit standardisierten Dauerwerten. Auf diese Weise erzeugte die Standardisierung ein längeres Diphon, mit anderer segmenteller Information. Die Standardisierung führte bei der Synthese in einigen Fällen zu einem perzeptiv relevanten, veränderten Intonationsverlauf.

Aus den Tests ging hervor, daß die Standardisierung der Dauern der deutschen Diphone keinen signifikanten Einfluß auf die Qualität der Sprachsynthese hat, denn die Versuchspersonen fanden manchmal die standardisierte Version besser, und manchmal die nicht-standardisierte Version.

Teil 1 dieses Berichtes beschreibt die Problemstellung, Teil 2 gibt eine kurze Einführung in die angewandte Sprachsynthese, Teil 3 die Beschreibung des Experiments und Teil 4 die Schlußfolgerungen.

Vorwort

Im Rahmen meiner Tätigkeit als Praktikant am Instituut voor Perceptieonderzoek (IPO) in Eindhoven habe ich in den Monaten März, April und Mai 1989 ein erstes Experiment durchgeführt. Ich machte zwei Perzeptionstests mit deutscher synthetischer (Computer-) Sprache.

Ich möchte Herrn Prof. Dr. W. Hess in Bonn für die gebotene Gelegenheit danken, die Tests an seinem Institut (IKP) durchzuführen. Auch Herrn Dr. D. Stock und Herrn Dr. W. Sendlmeier möchte ich herzlich danken für den sehr netten Empfang und für ihre Unterstützung. Sie haben dafür gesorgt, daß es einen geeigneten Testraum für die Durchführung der Tests gab, und daß die benötigte Apparatur bereitstand. Außerdem danke ich den Versuchspersonen für ihre bereitwillige Mitarbeit.

Mein Dank geht nicht zuletzt an Herrn Drs. Wim Peeters, OTS, Rijksuniversiteit Utrecht, der es mir ermöglicht hat, diese Arbeit am IPO durchzuführen, und an Frau Drs. Ursula Adriaens-Porzig, die mir mit Tat und gutem Rat am IPO sehr fachfraulich und geduldig zur Seite gestanden hat.

Contents

I	Problemstellung	5
II	Sprachsynthese	5
1	SPICOS	5
2	Erstellung des 1. Diphonlexikons	5
3	Erstellung des 2. Diphonlexikons	6
III	Das Experiment: Test 1, Test 2	7
4	Testdatum, Testort und Versuchspersonen	7
5	Methode	7
6	Test 1: Der Vergleichstest	7
6.1	Stimuli	7
6.2	Synthese der Wörter	7
6.3	Testband	8
6.4	Testantwortbogen	8
6.5	Ergebnisse	8
6.6	Auswertung	8
7	Test 2: Der Akzeptabilitätstest	10
7.1	Stimuli	10
7.2	Synthese der Sätze	10
7.3	Testband	11
7.4	Testantwortbogen	11
7.5	Ergebnisse	11
7.6	Auswertung	11
IV	Schlußfolgerungen	15

I Problemstellung

Meine Aufgabe war es zu untersuchen, ob die Standardisierung von Diphondauern einen Einfluß auf die Sprachqualität der Diphonsynthese hat. Zu diesem Zweck habe ich Diphonsprache synthetisiert, einmal mit nicht-standardisierten und einmal mit standardisierten Diphonen. Die Verständlichkeit und Natürlichkeit der Synthesen habe ich von deutschsprachigen Versuchspersonen in zwei Tests beurteilen lassen.

II Sprachsynthese

1 SPICOS

Im SPICOS-Projekt wird die Kommunikation mittels Sprache zwischen Mensch und Computer für das Deutsche untersucht. Das Forschungsprojekt ist eine Zusammenarbeit des IPO mit dem Philips Forschungslabor Hamburg und Siemens München (SPICOS heißt: Siemens Philips IPO Continuous Speech). Das Projekt besteht aus drei Teilen: 1. Automatische Spracherkennung, 2. Automatische Interpretation des Gesprochenen und Antwortgenerierung, und 3. Sprachsynthese, Van Hemert et al. (1987). Mit einem Teil der Sprachsynthese habe ich mich in diesem Experiment befaßt.

Es gibt verschiedene Arten, Sprache mit Hilfe eines Computers zu synthetisieren: 1. Synthese nach Regeln (Vollsynthese), und 2. Synthese mit vorher abgespeicherten Sprachsegmenten. Im SPICOS-Projekt werden für die Synthese Sprachsegmente, nämlich Diphone benutzt. Diphone werden gebildet, indem man aus natürlicher Sprache aus zwei aufeinanderfolgenden Phonemen die zweite Hälfte des ersten Phonems und die erste Hälfte des zweiten Phonems segmentiert. Dabei bleibt die Transition zwischen diesen beiden Phonemhälften erhalten und wird mit im Diphon abgespeichert, De Pijper (1988).

Im Deutschen gibt es 44 Phoneme ¹ (Tabelle im Anhang A). Da jedes Diphon aus einer Kombination von zwei Phonemen besteht, gibt es etwa 1600 Diphone. Im Anhang B werden die Vorkommensfrequenzen der deutschen Diphone (aus den sogenannten Sotscheksätzen) aufgelistet.

In meinem Experiment habe ich deutschen Versuchspersonen (Muttersprachler) zwei Versionen der deutschen Diphonsynthese zum Vergleich dargeboten.

2 Erstellung des 1. Diphonlexikons

Die Diphone wurden aus natürlicher Sprache, aus Logatomen (Nonsensewörtern) segmentiert. Die Segmentation erfolgte automatisch mit dem Programm AUTOSEG, Van Hemert (1985). Danach wurden die Diphone in LPC-analyzierter Form, Vogten (1981) abgespeichert. Für die Synthese konkateniert man die entsprechenden Diphone und gibt sie, nachdem man eine geeignete Intonationskontur hinzugefügt hat, über einen Lautsprecher aus. Die Diphone aus dem 1. Lexikon werde ich im weiteren die 'alten' Diphone nennen.

¹Es wäre besser zu sagen, daß 44 Zeichen gewählt sind, die die deutschen Phoneme darstellen.

3 Erstellung des 2. Diphonlexikons

Nach der automatischen Segmentierung wurden manuell Korrekturen durchgeführt. Die Korrekturen umfaßten die Standardisierung der Diphondauern und die Korrektur aller Segmentgrenzen, Adriaens-Porzig (1989). Die Matrix im Anhang C zeigt die Zahl der Frames im korrigierten und standardisierten (neuen) Diphon minus die Zahl der Frames im alten Diphon. Ein Frame dauert 10 Millisekunden. Die stilisierten Diphone bilden ein neues Lexikon. Die Diphone aus diesem Lexikon werde ich weiterhin die 'neuen' Diphone nennen.

Die beiden zu vergleichenden Lexika für die Diphonsynthese waren also:

- Lexikon 1: automatisch segmentierte Diphone mit automatisch gesetzten Phonemgrenzen
- Lexikon 2: manuell segmentierte Diphone mit manuell gesetzten Phonemgrenzen mit standardisierten verkürzten Diphondauern.

III Das Experiment: Test 1, Test 2

4 Testdatum, Testort und Versuchspersonen

Das Experiment habe ich am 24. Mai 1989 an der Universität Bonn im Institut für Kommunikationsforschung und Phonetik (IKP) durchgeführt. Beim ersten Test haben 15 Versuchspersonen (Vp) mitgemacht, beim zweiten Test 18. Die ersten 15 Vp haben an beiden Tests teilgenommen. Die Vp waren Universitätsmitarbeiter und Phonetikstudenten des Instituts. Sie hatten alle als Muttersprache Deutsch.

5 Methode

In dem Experiment sollte getestet werden, ob es bezüglich der Verständlichkeit und der Natürlichkeit der beiden Synthesen Unterschiede gibt. Dazu habe ich zwei verschiedene Tests benutzt: Der erste Test war ein Vergleichstest, der zweite ein Akzeptabilitätstest. Im ersten Test sollten die Vp entscheiden, welche der beiden Syntheseversionen sie bevorzugten. Im zweiten Test sollten die Vp die Qualität der jeweiligen Version mit einer Note zwischen 1 und 6 beurteilen.

6 Test 1: Der Vergleichstest

6.1 Stimuli

Die Stimuli für den ersten Test waren verschiedene Versionen einzelner Wörter. Jedes Wort war je einmal mit den alten und einmal mit den neuen Diphonen synthetisiert worden. Die Vp sollten die beiden Versionen des gleichen Wortes miteinander vergleichen. Die vierzig mehrsilbigen Wörter waren so gewählt, daß möglichst viele Phonemkombinationen enthalten waren (Matrix im Anhang D).

6.2 Synthese der Wörter

Die Wörter wurden mit dem LVS-Programm DS synthetisiert. Zuerst wurden die Wörter in Phonemschreibweise (Anhang A) in den Graphem-Phonem-Konverter (Nr. 2) eingegeben. Das erste Mal habe ich die Wörter mit den alten Diphonen mit einer durchschnittlichen Framedauer von 80% (Mean Frame Duration=80), also ein wenig verkürzt, synthetisiert. Das zweite Mal habe ich die Wörter mit den neuen Diphonen mit Mean Frame Duration=100 synthetisiert. Da die neuen Diphone kürzer sind, war das Sprechtempo auf diese Art ungefähr wieder gleich geworden. (Der Gebrauch der unverkürzten alten Diphone hätte ein zu langsames Tempo ergeben). Die entstandenen AP-files habe ich mit dem LVS-Programm SYN zu N-files synthetisiert. Mit Hilfe des LVS-Programms RYN wurden die N-files der zwei Syntheseversionen (Alt und Neu) eines Wortes zu zwei Stimuluspaaren verbunden, einmal in der Kombination Alt/Neu und einmal in der Kombination Neu/Alt. Innerhalb eines Stimuluspaares gab es eine Pause von 1200 Millisekunden.

Table 1: Übersicht über die von den Vp getroffenen Entscheidungen (80 Stimuluspaare).

VP	A	Alt	Neu	Alt - Neu
1	31	37	43	-6
2	42	42	38	4
3	40	46	34	12
4	45	47	33	14
5	47	47	33	14
6	34	48	32	16
7	39	41	39	2
8	35	35	45	-10
9	34	46	34	12
10	43	43	37	6
11	35	37	43	-6
12	41	39	41	-2
13	55	39	41	-2
14	49	39	41	-2
15	36	40	40	0
Insgesamt	626	574	52	

6.3 Testband

Von den N-files der Stimuli habe ich eine SPL-Liste für das LVS-Sprachausgabeprogramm SPL gemacht. Die Stimuli habe ich in randomisierter Reihenfolge, vgl. Moses und Oakford (1963), auf ein Tonband (BASF), Tonbandgerät Revox A77, und auf eine Tonbandkassette (Philips LH), Kassettenrekorder Philips FC153, aufgenommen. Die Pausen zwischen den Stimuluspaaren waren 3600 Millisekunden. Nach jedem Block von zehn Stimuluspaaren gab es einen kurzen Pfeifton, um den Vp Gelegenheit zu bieten, sich zu orientieren. Es gab im ganzen achtzig Stimuluspaare, die den Vp über Lautsprecher vorgespielt wurden.

6.4 Testantwortbogen

Vor Beginn des Testes bekamen die Vp schriftliche Instruktionen und den Testantwortbogen für den ersten Test (Anhang H). Die Instruktionen habe ich mit den Vp durchgesprochen. Die Vp wurden gebeten, die Reihe A anzukreuzen, wenn sie das erste Wort des Stimuluspaares bevorzugten, oder die Reihe B, wenn sie das zweite Wort bevorzugten.

6.5 Ergebnisse

Die fünfzehn Vp sollten achtzig Mal entscheiden, ob sie die alte Version oder die neue Version verständlicher oder natürlicher fanden. Das ergibt eine Anzahl von 1200 Entscheidungen. Davon waren 626 Entscheidungen für die alte Version und 574 für die neue Version. In der Tabelle 1 stehen die Ergebnisse aufgelistet nach den Vp, in der Tabelle 2 nach den Wörtern .

6.6 Auswertung

Die Vp.

Mit einem Chi-Quadrat (χ^2 -) Test habe ich festgestellt, daß die Gesamtzahlen der Entscheidungen für die alte oder für die neue Version nicht signifikant (z-Score= 1.39) von den erwarteten Gesamtzahlen (für beide Versionen: 600) abweichen. Die Signifikanzgrenze habe ich auf 5 Prozent gelegt (Grenzen 566 und 634). Ebenfalls mit einem χ^2 -Test (N=80, p=0.5) habe ich festgestellt, daß keine der Vp eine der beiden Versionen signifikant öfter

Table 2: Übersicht über die von den Vp gewählten Versionen (15 Vp) (In Klammern die Wortpaarnummer im Test (Alt/Neu, Neu/Alt)).

WORT	VERSION		WORT	VERSION	
	Alt	Neu		Alt	Neu
Stimulus (61,12).	21	9	Computer (19,59).	1	29
Lautkombination (2,27).	24	6	Verständlichkeit (64,55).	20	10
Perzeption (60,33).	19	11	Kontext (28,53).	13	17
Versuchsperson (16,5).	16	14	Anforderung (45,80).	12	18
Diphone (9,78).	4	26	Schwierigkeit (14,20).	14	16
Vergleichstest (52,56).	18	12	Kommunikation (67,71).	12	18
Fernseher (79,38).	12	18	Nachrichtensendung (72,43).	22	8
Sprachsynthese (30,6).	18	12	Information (3,22).	17	13
Bildschirm (7,47).	22	8	Ausserung (11,47).	12	18
Forschung (41,49).	6	24	Wissenschaftler (10,4).	7	23
Entwicklung (46,39).	23	7	Akustik (76,62).	5	25
Automatisch (21,73).	10	20	Lexikon (70,63).	26	4
Identifikation (57,18).	19	11	Phonologie (66,75).	15	15
Glottisschlag (77,37).	16	14	Vokalübergänge (24,36).	9	21
Herausgestellt (34,65).	18	12	Erfahrung (17,69).	28	8
Wettervorhersage (42,13).	18	12	Staumeldung (32,51).	18	12
Reihenfolge (31,23).	18	12	Frequenzbereich (15,1).	26	4
Tonbandgerät (35,8).	23	7	Gebrauchsmöglichkeit(50,44).	13	17
Darbietung (48,25).	12	18	Sinuston (40,26).	21	8
Deutschsprachig (58,29).	11	19	Testteilnehmer (54,68).	6	24

Table 3: Wörter, wobei die ALTE Version bedeutend besser beurteilt wurde.

Lautkombination	24	6
Nachrichtensendung	22	8
Bildschirm	22	8
Entwicklung	23	7
Lexikon	26	4
Erfahrung	28	2
Frequenzbereich	26	4
Tonbandgerät	23	7
Sinuston	22	8

gewählt hat (Grenzen 30 und 50). Eine Vp (Nr. 13) hat signifikant öfter A angekreuzt (55 mal), aber dies hatte keinen Einfluß auf ihre Scoreverteilung über Alt und Neu.

Die Wörter.

Aus der Scoreverteilung für die Wörter zeigen sich für die einzelnen Wörter große Unterschiede. Die Position des Stimulus im Test hatte keinen Einfluß auf die Ergebnisse. Ein χ^2 -Test (N=30, p=0.5) ergab, daß für 15 Wörter die Unterschiede zwischen den beiden Versionen signifikant waren (Grenzen 21 und 9). Es fällt dabei auf, daß für einige Wörter die alte Version öfter bevorzugt wurde (Tabelle 3), und für andere die neue Version (Tabelle 4). Für 25 Wörter gab es also keinen bedeutenden Unterschied zwischen der alten und der neuen Version.

Bei einigen Wörtern, wobei die alte Version bedeutend bevorzugt wurde, war infolge der Korrekturen eine Intonationsänderung aufgetreten. Dies gilt vor allem für das Wort 'Tonbandgerät'. Es fiel mir auf, daß nicht die normale, sondern die auffälligere (falsch

Table 4: Wörter, wobei die NEUE Version bedeutend besser beurteilt wurde.

Computer	1	29
Diphone	4	26
Forschung	6	24
Wissenschaftler	7	23
Akustik	5	25
Testteilnehmer	6	24

betonte) Version öfter angekreuzt wurde. In 'Lautkombination' war das au-t Diphon korrigiert worden. Die erste Silbe ist deshalb in der neuen Version länger und mehr betont. In dem Wort 'Nachrichtensendung' ist die segmentelle Information in den Silben -ich und -ten durch die Standardisierung zu gering geworden.

Bei anderen Wörtern ist mir nach wiederholtem Abhören nicht ganz klar, weshalb die Vp die alte Version so stark bevorzugt haben. Dies gilt besonders für das Wort 'Erfahrung', wobei nach meiner Meinung die letzte Silbe in der alten Version schlechter war. Bei dem Wort 'Forschung' wird dahingegen die neue Version von den Vp stark bevorzugt.

Beidrei Wörtern, von denen die neue Version bevorzugt wurde, ist zu hören, daß die Korrektur des Enddiphons er-si der wichtigste Faktor ist (Computer, Wissenschaftler, Testteilnehmer). Bei dem Wort 'Diphone' ist das e-si Diphon in der alten Version zu kurz. Außerdem war das oh-n Diphon korrigiert. In dem Wort 'Akustik' waren die Diphone ah-k und i-k korrigiert. Vor allem das erste k ist deutlicher.

7 Test 2: Der Akzeptabilitätstest

7.1 Stimuli

Für den zweiten Test habe ich vollständige Sätze verwendet. Die zwölf verwendeten Sätze stammen aus dem SPICOS-Inventar und waren so gewählt, daß sie im LVS-DS-Programm vom SPICOS-Graphem-Phonem-Konverter (Nr.4) richtig synthetisiert wurden und die richtige Intonationskontur bekamen. Außerdem sollten sie möglichst viele Phonemkombinationen enthalten (Matrix im Anhang E). Jeder Satz (mit alten oder mit neuen Diphonen synthetisiert) bildete einen Stimulus.

7.2 Synthese der Sätze

Die Synthese der Sätze wurde auf ähnliche Weise wie die der Wörter gemacht. Der einzige Unterschied war der Gebrauch des SPICOS-Graphem-Phonem-Konverters (Nr. 4 statt Nr. 2). In diesem Konverter (Nr. 4) wird die Akzentverteilung in einem Satz automatisch generiert (im Konverter Nr. 2 werden Akzente manuell eingegeben (Anhang A)). Außerdem ist sein Lexikon nur für bestimmtes Vokabular zugelassen. Die Tempi der verschiedenen Satzversionen waren durch die Verkürzung der alten Diphone auf 80% (Mean Frameduration=80) etwa gleich.

Table 5: Übersicht über die von den Vp vergebenen Noten (Mittelwert über 12 Sätze).

Vp Nr	Version	
	Alt	Neu
1	4.083	4.167
2	4.250	4.500
3	4.667	4.417
4	4.333	4.333
5	4.000	4.667
6	3.667	4.000
7	4.583	4.667
8	4.417	4.833
9	4.417	4.417
10	3.083	3.333
11	3.167	3.333
12	3.917	3.833
13	3.333	4.000
14	4.167	4.417
15	3.000	3.250
16	3.833	4.083
17	3.750	4.167
18	4.500	4.167

7.3 Testband

Auch von den Sätzen habe ich eine SPL-Liste für das Sprachausgabeprogramm SPL gemacht. Die verschiedenen Versionen der Sätze habe ich in randomisierter Reihenfolge, Moses und Oakford (1963), aufgenommen (mit demselben Material wie bei den Wörtern). Nach jedem Stimulussatz gab es jedesmal 2400 ms Pause, dann einen Pfeifton (500 ms) und dann nochmal 1200 ms Pause.

7.4 Testantwortbogen

Vor Beginn des zweiten Testes bekamen alle Vp die Instruktionen für den zweiten Test und einen Testantwortbogen (Anhang I). Die Vp wurden gebeten, jeden Satz einzeln zu beurteilen. Sie sollten den Satz mit einer Note von 1 bis 6 bewerten: die Note 1 für einen sehr guten Satz, die Note 6 für einen sehr schlechten. Ich habe die Instruktionen durchgesprochen, und die Vp dabei gebeten, die ganze Skala von 1 bis 6 zu benutzen.

7.5 Ergebnisse

Die Vp haben jeden der 24 Stimuli mit einer Note von 1 (sehr gut) bis 6 (sehr schlecht) bewertet (Im Anhang F ist die Scoreverteilung über die Sätze zu sehen). Über alle 18 Vp, 2 Versionen und 12 Sätzen gerechnet ergab diese Bewertung die Note 4.05. Dieses und folgende Ergebnisse sind mit dem Statistikprogramm ALICE auf dem IPO-Vax errechnet. Die Noten nach Vp und nach Satz geordnet, stehen in den Tabellen 5 und 6.

7.6 Auswertung

Als erstes habe ich mit den rohen Daten eine Varianzanalyse durchgeführt. Das Ergebnis steht in Tabelle 1 im Anhang G. Es zeigte sich, daß es verschiedene signifikante Effekte gab. Signifikante Effekte gab es 1. für die Vp, 2. für die Sätze und 3. für die Versionen. Der Vp-Effekt deutet daraufhin, daß die Streuung und der Mittelwert der Scores pro Vp sehr unterschiedlich waren. Diesen Vp-Effekt habe ich ausgeglichen, indem ich alle

Table 6: Übersicht über die Bewertung der Sätze (18 Vp). In Klammern die Stimulusnummer (alt, neu)

Satz Nr.	rohe Scoredaten		standardisierte Scoredaten	
	Version		Version	
	Alt	Neu	Alt	Neu
1	4.889	3.333	0.648	-0.503
2	4.556	5.500	0.365	1.108
3	3.556	4.222	-0.364	0.119
4	4.000	5.833	-0.072	1.356
5	5.056	3.444	0.759	-0.420
6	3.944	3.389	-0.054	-0.514
7	5.611	5.833	1.197	1.350
8	4.056	4.500	0.007	0.311
9	2.611	3.722	-1.088	-0.257
10	4.222	3.722	0.126	-0.224
11	2.333	2.944	-1.291	-0.885
12	2.611	3.278	-1.090	-0.576

Bewertungen zu z-Scoren standardisiert habe. Je höher die z-Score ist, desto schlechter war die Bewertung.

Das Ergebnis einer erneuten Varianzanalyse, aber jetzt mit standardisierten Scorewerten, steht in Tabelle 2 im Anhang G. Sie zeigt 1. einen signifikanten Satzeffekt, 2. einen signifikanten Versionseffekt, aber 3. leider auch einen signifikanten Interaktionseffekt von Sätzen mit Versionen. Dieser Interaktionseffekt deutet daraufhin, daß die verschiedenen Wörter in den Sätzen die Bewertungen beeinflusst haben. Der Satzeffekt bedeutet, daß die Sätze sehr unterschiedlich beurteilt wurden. Tabelle 3 im Anhang G zeigt, daß man den Versionseffekt über alle Sätze als nicht-signifikant ansehen muß. Der Effekt der einzelnen Sätze ist nämlich so groß, daß der Versionseffekt wegfällt. Deshalb habe ich für jeden einzelnen Satz mit Hilfe eines F-Tests beurteilt, ob es einen signifikanten Versionseffekt gibt. Dabei zeigte es sich, daß nur drei Sätze einen signifikanten Versionseffekt aufwiesen (Die Nummern 1, 4 und 5). Bei den Sätzen 1 und 5 wurden die neuen Diphone besser bewertet, bei Satz 4 die alten Diphone.

Meinen ursprünglichen Testplan habe ich der Umstände halber nicht ausführen können. Ich mußte die Reihenfolge der Tests ändern, sodaß zuerst die Sätze getestet wurden. Am Scoreverlauf (Fig. 1) sieht man, daß die Vp im Laufe des Tests die Sätze immer höher bewerten. Den signifikanten Unterschied in der Bewertung der beiden Versionen von Satz 1 muß man also der Gewöhnung der Vp an die Diphonsprache zuschreiben. Diesen Effekt will ich wegen der geringen Zahl der Sätze aber nicht berücksichtigen.

Die Sätze 2,4 und 7 wurden besonders schlecht beurteilt (Fig. 2). Die Versionen von Satz 2 wurden beide in der ersten Hälfte des Testes dargeboten. In der neuen Version des Satzes war von 'Zeige' die erste Silbe undeutlicher als in der alten Version, trotzdem wurde die alte Version nicht viel besser beurteilt.

Auch von Satz 4 wurde zuerst die neue Version am Anfang des Testes dargeboten. Es fiel

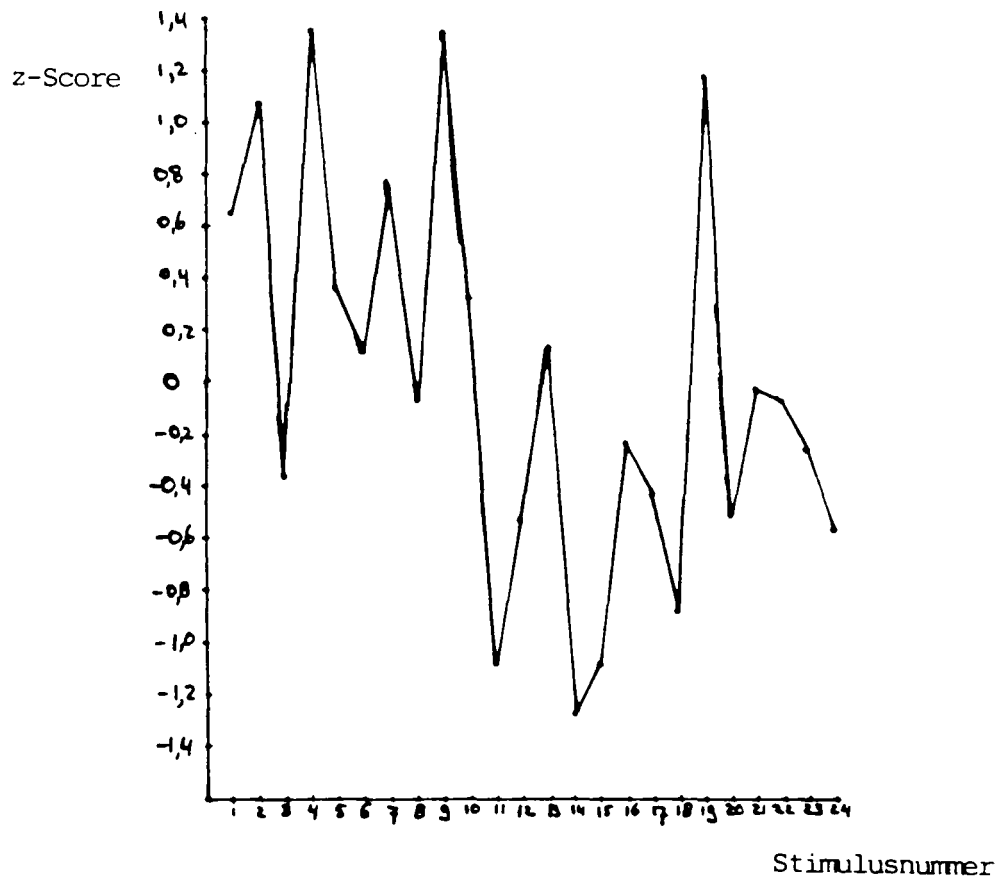


Figure 1: Verlauf der Scores im Test. (Je höher die z-Score, desto schlechter die Bewertung).

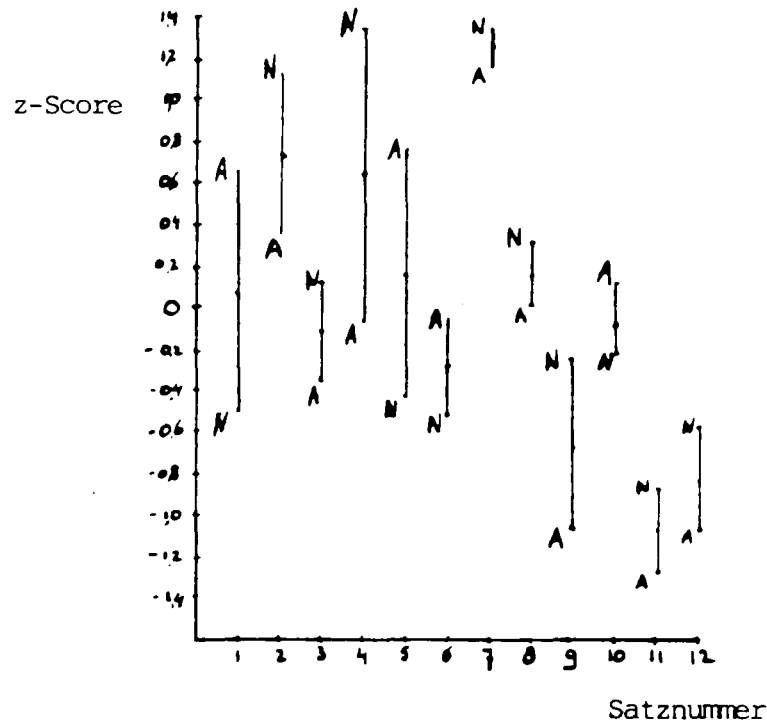


Figure 2: Scoreverteilung pro Satz. A = Alt, N = Neu. Unterschied zwischen den Sätzen: Satzefekt. Unterschied pro Satz: Versionseffekt.

auf, daß der Satz nicht glatt verlief, sondern holperte. Vor allem das Wort 'Aufgaben' klang zu schnell. Die Diphone au-f und f-g waren sehr stark verkürzt, wodurch die segmentelle Information zu gering geworden war. In der alten Version war das Enddiphon e-si in 'Netze' der störende Faktor (Vergl. bei den Wörtern 'Diphone').

In Satz 7 waren es wahrscheinlich die Wörter 'Abstracts' und 'Aufsätze', die den Vp nicht gefielen. Das Wort 'Abstracts' war weniger bekannt, und deshalb wurde eine hohe Anforderung an die Verständlichkeit dieses Wortes gestellt. In dem Wort 'Aufsätze' hatte das Intonationsmodul einen Unterschied zwischen beiden Versionen verursacht (Vergl. bei den Wörtern 'Tonbandgerät').

IV Schlußfolgerungen

Aus dem Akzeptabilitätstest ist hervorgegangen, daß die Natürlichkeit und die Verständlichkeit der Diphonsprache von den Vp im Mittel mit der Note 4 bewertet wurde. Wenn man das Mittel für die zweite Hälfte des Testes nimmt, um den Gewöhnungseffekt ein wenig auszugleichen, wurde die (höhere) Note 3.5 vergeben.

Es zeigte sich, daß die Korrekturen und die Standardisierung bei den einzelnen Wörtern einen Effekt hatten, der das eine Mal positiv, das andere Mal aber negativ bewertet wurde. Ein positiver Einfluß ging von einigen Korrekturen aus, vor allem am Ende eines Wortes oder eines Satzes. Einen negativen Einfluß hatten die Veränderungen der Intonation, die durch die Standardisierung und die Korrekturen verursacht wurden.

Die alten Diphone wurden in verkürzter Form verwendet. Der Unterschied zu den standardisierten Diphonen war ziemlich klein, da die Standardisierung eines Diphons meist ein kürzeres Diphon erzeugte.

Man kann sagen, daß der negative Effekt der Standardisierung so klein ist, daß die Sprachqualität mit standardisierten Diphonen nicht signifikant schlechter ist als die mit nicht-standardisierten Diphonen. Es zeigte sich, daß kleine Korrekturen in der Segmentierung an bestimmten Stellen eine große Verbesserung der Verständlichkeit und Natürlichkeit der Diphonsprache bewirken können. Eine Anpassung des Intonationsmoduls an die standardisierten Diphone würde meines Erachtens eine erhebliche Qualitätsverbesserung der synthetischen Sprache bewirken. Eine Zeitstrukturverbesserung der Diphonsprache wäre außerdem ein weiterer Schritt zur größeren Natürlichkeit und Verständlichkeit der aus Diphonen synthetisierten Sprache.

Literaturverzeichnis

Adriaens-Porzig, U. (1989): Interner Rapport IPO, (in Vorbereitung).

Van Hemert, J.P. (1985): 'Automatic diphone preparation', IPO Annual Progress Report 20, S. 23-32.

Van Hemert, J.P.; Adriaens-Porzig, U.M.; Adriaens, L.M.H. (1987): 'Speech synthesis in the SPICOS-project', in: Tillmann, H.G.; Willee, G. (Hg.) Analyse und Synthese gesprochener Sprache, Georg Olms Verlag, S. 34-39.

Moses, L.E.; Oakford, R.V. (1963): 'Tables of random permutations', George Allen & Unwin Ltd., London.

De Pijper, J.R. (1988): 'Van tekst naar spraak met behulp van difoonsynthese', IPO Report, Ms 664.

Vogten, L.L.M. (1981): 'Spraak-analyse, resynthese en enkele toepassingen', IPO Report, Ms 402.

LISTE DER DEUTSCHEN PHONEME

KLASSEN: 1-18 VOKALE, 16-18 DIPHTONGE, 19-24 PLOSIVE, 25-30 FRICATIVE, 31-33 NASALE, 34-36 HALBVOKALE, 37-38 LIQUIDA, 39-40 PAUSE, GLOTTISSCHLAG UND 41-44 SILBISCHE KONSONANTEN.

NR	PHONEM- SCHREIBWEISE	BEISPIEL	NR	PHONEM- SCHREIBWEISE	BEISPIEL	NR	PHONEM- SCHREIBWEISE	BEISPIEL
1	I	ICH	19	P	PATER	34	W	WAHR
2	IE	IHRE	20	T	TAT	35	Y	JA
3	AE	ELF	21	K	KATER	36	H	HOFFEN
4	EH	EBEN	22	B	BAD			
5	A	APRIL	23	D	DA	37	L	LANG
6	AH	ABER	24	G	GABEL	38	R	ROT
7	E	STUNDE						
8	O	ONKEL	25	F	FAHRRAD	39	SI	<PAUSE>
9	OH	ODER	26	S	DAS	40	GS	<GLOTTISSCHLAG>
10	U	UND	27	X	ACH			
11	UH	UHR	28	CH	ICH	41	EL	EDEL
12	UE	BUERO	29	SCH	SCHADE	42	EM	JEDEM
13	UEH	UEBEL	30	Z	SEHEN	43	EN	JEDEN
14	OE	OEFFENTLICH				44	ER	JEDER
15	OEH	OEFEN	31	M	MUTTER			
			32	N	NATUR			
16	AI	EIFRIG	33	NQ	JUNGE			
17	AU	AUCH						
18	EU	EURE						

LISTE DER IN DEM TEST VERWENDETEN WÖRTER (PHONEMSCHREIBWEISE)

scht'iemulus.	kompy'uhter.
l'autkombinatsyohn.	fae rscht'aentlichkait.
paertsaeptysy'ohn.	k'ontaekst.
fae rz'uhsxae rzohn.	@'anforde runq.
d'ief'ohne.	schw'ierichkait.
fae rgl'aichstaest.	komuhniekatsy'ohn.
f'aernzeher.	n'ahxrichtenzaendung.
schpr'ahxzuentehze.	informatsy'ohn.
b'iltschirm.	@'euser runq.
f'orschunq.	w'isenschaftler.
@aentw'i klunq.	@ahk'ustik.
@autohm'ahtisch.	l'aeksikon.
@iedaentiefiekatsy'ohn.	fohnohlohg'ie.
gl'otisschlahk.	wohk'ahl@uehbergaenge.
her'ausgeschtaelt.	@aerf'ahrung.
w'aeterfohrzehrzahge.	scht'au maeldunq.
r'aienfolge.	frehkw'aentsbe raich.
t'ohnbantge reht.	gebr'auxsmoehklichkait.
d'ahrbietunq.	z'ienustohn.
d'eutschschprahxich.	t'aesttailnehmer.

Ein "'" gibt die betonte Silbe im Wort an.

"e r" ist nicht dasselbe wie "er" (vergl. Nr. 44 der Phoneme).

Ein "@" steht für den Glottisschlag (=gs).

Ein " " bewirkt eine Pause (=si).

VORKOMMENSFREQUENZ DER DEUTSCHEN DIPHONE (AUS DEN SOGENANTEN SOTSCHESKÄTZEN)

F S SCHCH X H W Z Y R P T K B D G M N NO L SI GS																	IE UE EH AH UH OH OE HI UE E ⁶ AE A U O OE EL EM EN ER AI EU AU																																		
F	--	--	--	--	--	1	--	5	--	6	2	1	4	--	--	3	--	--	F	2	3	3	8	1	2	--	1	--	2	10	1	1	3	--	1	--	5	--	2	--	--										
S	1	--	3	1	--	1	3	--	--	21	2	1	3	3	4	4	--	1	--	S	6	--	--	6	2	--	3	1	4	1	--	9	1	--	--	--	7	2	3	--	1										
SCH	--	--	--	--	--	1	--	--	--	3	18	--	--	1	--	3	--	3	--	SCH	--	--	--	1	2	2	--	2	1	2	1	--	--	--	--	--	1	1	--	1	--	--									
CH	--	--	1	--	--	1	1	1	--	--	20	2	--	2	1	1	2	--	--	CH	--	--	--	--	--	--	--	1	--	--	--	--	--	--	--	--	4	--	--	--	--										
X	--	--	1	--	--	1	--	--	--	--	9	--	--	1	--	--	--	--	1	X	--	--	--	--	--	--	1	--	--	--	--	--	--	--	--	4	--	1	--	--	--										
H	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	H	4	--	2	--	2	3	2	--	1	3	1	--	--	--	--	--	4	--	1	2	2	--									
W	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	W	14	--	7	3	--	3	--	7	--	1	2	6	1	2	--	--	--	1	--	2	--	--	--								
Z	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	Z	3	1	--	3	1	--	--	10	--	3	1	1	1	3	--	--	--	5	--	2	--	1	--	--							
Y	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	Y	--	--	--	--	--	--	--	--	3	1	1	--	--	--	--	--	--	--	--	--	--	--	--	--							
R	3	3	1	2	--	4	5	1	3	--	2	15	6	8	3	6	7	7	--	R	1	3	2	2	--	2	1	8	2	2	5	2	4	1	--	--	--	4	1	2	--	4	--	--							
P	5	2	--	--	--	--	--	1	--	3	--	--	--	--	--	2	--	--	1	P	1	1	--	--	--	--	--	--	1	--	--	--	--	--	--	1	--	--	1	--	--	--	--	--							
T	7	37	6	1	--	2	7	3	--	5	--	--	2	20	4	2	2	--	1	T	1	1	6	3	1	1	--	7	2	22	2	5	2	3	--	1	--	9	13	3	--	1	--	--							
K	--	2	--	--	--	3	1	--	--	--	9	--	1	--	1	--	1	--	2	K	--	--	--	1	1	--	--	1	2	--	6	1	5	1	--	--	2	2	2	--	2	--	--								
B	--	--	--	--	--	--	--	5	--	--	--	--	--	--	--	--	--	--	2	B	--	--	--	4	--	--	2	--	5	--	3	2	--	--	1	--	5	5	5	1	1	--	--								
D	--	--	--	--	--	--	--	4	--	--	--	--	--	--	--	--	--	--	--	D	27	--	24	2	2	--	2	--	6	3	11	1	2	--	--	--	4	3	1	--	--	--	--	--							
G	--	--	--	--	--	--	--	2	--	--	--	--	--	--	--	--	--	7	--	G	4	--	4	2	3	--	--	1	11	4	2	--	--	--	--	6	1	1	--	--	--	--	--	--							
M	2	--	3	--	--	1	1	--	1	1	5	--	3	1	1	--	--	--	1	M	2	--	1	1	--	1	--	10	2	3	3	4	5	1	3	1	--	1	1	1	--	--	--	--	--						
N	2	8	5	1	--	2	3	5	--	1	31	1	3	18	2	1	--	--	2	N	--	--	2	4	2	--	9	--	13	3	3	1	5	--	--	2	6	3	--	--	--	--	--	--							
NO	--	1	--	--	--	--	--	--	--	--	1	2	--	1	--	--	--	--	--	NO	--	--	--	--	--	--	--	1	--	--	--	--	--	--	--	1	2	--	--	--	--	--	--	--							
L	2	--	1	--	--	1	--	1	9	1	--	2	1	--	2	--	--	3	L	4	1	1	2	1	--	1	4	1	4	2	4	2	--	--	--	3	2	9	4	1	--	--	--	--							
SI	5	--	1	--	--	4	10	3	2	6	--	11	1	1	26	5	6	5	--	SI	--	--	1	--	--	--	3	--	1	2	--	--	--	--	--	--	--	--	2	--	2	--	--	--	--						
GS	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	1	--	--	--	GS	2	3	1	3	1	1	1	14	--	13	16	7	2	--	--	--	--	--	17	--	9	--	--	--	--						
IE	6	--	3	1	--	2	1	--	18	4	5	7	4	1	2	--	4	--	5	IE	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	1	3	--	--	--	--	--						
UEH	--	1	1	--	--	--	--	4	--	--	4	--	--	--	--	1	--	1	--	UEH	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	5	2	--	--	--	--	--	--	--					
EH	--	--	--	--	--	1	--	20	--	6	1	2	1	1	4	7	--	--	--	EH	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--					
AH	1	1	--	--	3	--	1	--	11	--	7	4	4	--	1	1	4	--	2	AH	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--				
UH	--	1	--	--	2	--	1	--	4	--	3	4	--	--	--	--	--	--	--	UH	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--				
OH	4	1	--	--	--	1	--	--	1	1	--	--	1	--	5	--	2	--	1	OH	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	2	--	--	--	--	--	--	--	--				
OEH	--	1	--	--	--	--	--	3	--	1	--	--	--	--	--	2	--	1	--	OEH	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--			
I	--	4	6	22	--	--	--	2	--	11	1	--	4	6	20	2	3	--	--	I	--	--	--	2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--			
UE	--	2	--	1	--	--	--	1	--	--	4	--	--	--	--	3	--	--	--	UE	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--		
E	4	4	5	--	--	2	5	4	--	8	1	7	2	2	3	3	2	3	--	E	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--		
AE	1	13	--	3	--	--	--	11	--	6	3	--	--	2	5	2	7	--	12	AE	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	2	--	--	--	--	--	--	--	--			
A	5	7	1	--	9	1	--	7	2	4	1	2	1	1	8	18	--	6	--	A	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--		
U	2	5	--	--	2	2	--	5	--	3	1	--	--	4	8	6	1	--	--	U	--	--	--	--	--	--	--	1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--		
O	1	--	--	--	7	--	--	6	3	--	--	--	--	--	1	5	--	5	--	O	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--		
OE	--	--	3	--	--	--	--	--	--	--	--	--	--	--	--	1	--	--	--	OE	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--		
EL	--	--	--	--	--	--	--	--	--	1	--	--	--	--	--	1	--	--	1	EL	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	
EM	--	--	--	--	--	--	--	--	--	--	1	--	--	--	--	--	--	--	--	EM	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	
EN	4	3	4	--	--	3	5	1	--	1	3	5	2	4	6	2	--	--	1	EN	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	
ER	3	1	4	--	--	1	3	5	--	1	3	1	1	1	2	3	3	--	2	ER	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
AI	2	2	--	2	--	1	--	1	--	6	1	--	1	1	3	23	--	3	--	AI	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	
FU	1	1	--	1	--	--	--	--	--	3	--	--	--	--	1	1	--	--	--	FU	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	
AU	8	5	--	3	--	2	--	1	--	1	--	--	--	--	1	1	--	--	--	AU	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	

VORKOMMENSFREQUENZ DER IN DEN WÖRTERN VERWENDETEN DIPHONE

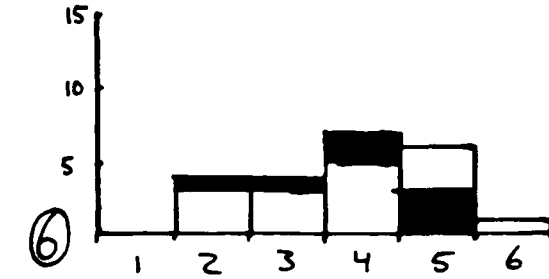
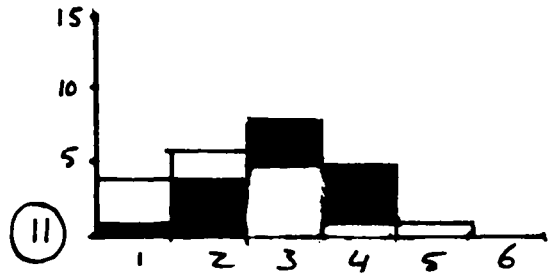
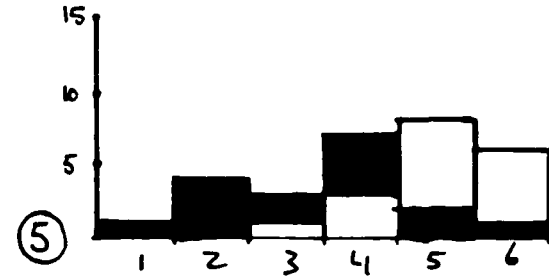
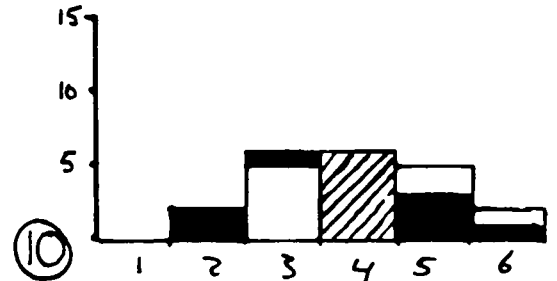
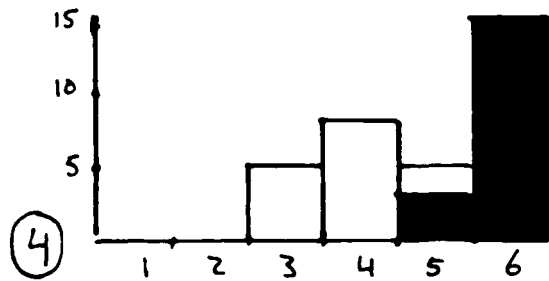
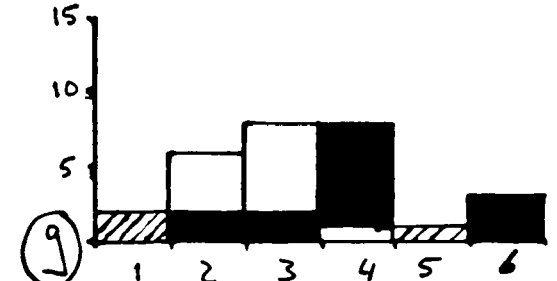
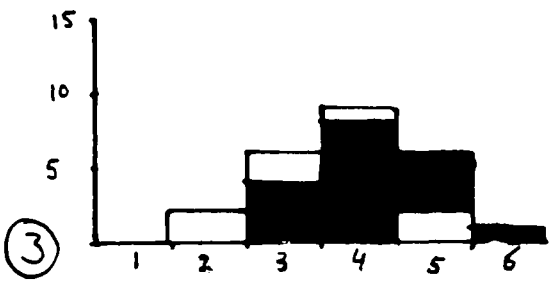
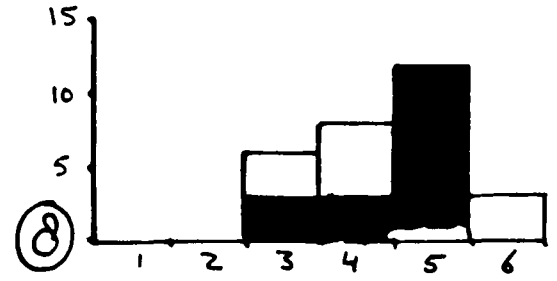
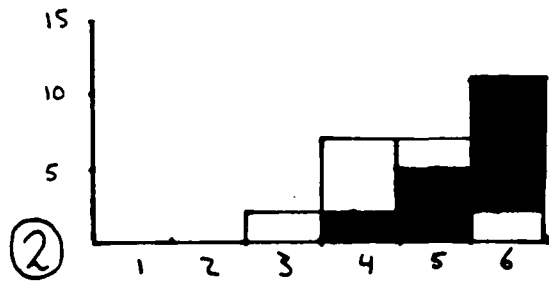
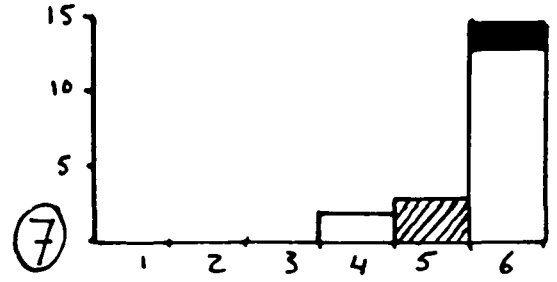
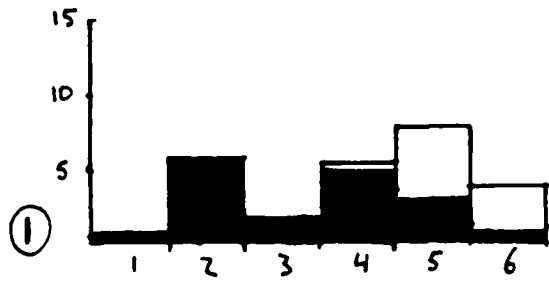
Diphone	Vorkommensfrequenz																							
	P	T	B	D	G	F	S	CH	SCH	M	Z	Y	H	U	NS	R	L	M	EL	PH	EM	ER	SI	GS
P	1	1				2		2	1							1	2							5
T	1																2							
B			1																					
D				1																				
G					1																			
F						2																		
S							1																	
CH								1																
SCH									1															
M										1														
Z											1													
Y												1												
H													1											
U														1										
NS															1									
R																1								
L																	1							
M																		1						
EL																			1					
PH																				1				
EM																					1			
ER																						1		
SI																							1	
GS																								1

VORKOMMENSFREQUENZ DER IN DEN SÄTZEN VERWENDETEN DIPHONE

	P	T	S	B	D	G	F	B	CH	SCH	M	Z	V	F	N	NB	R	L	M	EL	PH	EH	EP	SI	CS	
PT	1	3																								
TS		2		1	1	1																				
SB			1																							
BD				1																						
DG					1																					
FB						1																				
BC							1																			
BS								1																		
CH									1																	
SCH										1																
M											1															
Z												1														
V													1													
F														1												
N															1											
NB																1										
R																	1									
L																		1								
M																			1							
EL																				1						
PH																					1					
EH																						1				
EP																							1			
SI																								1		
CS																									1	

	A	I	U	E	O	AE	UE	OE	AN	IE	UN	EN	ON	LN	LM	AI	AU	EU		
A	1																			
I		4																		
U			1																	
E				4																
O					2															
AE						1														
UE							2													
OE								1												
AN									1											
IE										1										
UN											1									
EN												1								
ON													1							
LN														1						
LM															1					
AI																1				
AU																	1			
EU																		1		

ÜBERSICHT ÜBER DIE VERGEBENEN NOTEN (1: SEHR GUT, 6: SEHR SCHLECHT) ANHANG F



□ = ALT

■ = NEU

▨ = Beide

○ = SATZNR.

HORIZONTAL : Note

VERTIKAL : SUMME DER ANTWORTEN PRO NOTE

SOURCE	DF1/	DF2	F	P VALUE	MEAN SQUARE	SUM OF SQ
VV	1/	11	0.3916	0.5441	3.8912	3.8912
PP	17/	187	6.4135	0.0000	5.4963	93.4375
VV PP	17/	187	0.6264	0.8684	0.4353	7.4005
SA	11				28.7077	315.7647
SA VV	11				9.9316	109.2477
SA PP	187				0.8570	160.2569
SA VV PP	187				0.6950	129.9606

SOURCE	DF1/	DF2	F	P VALUE	MEAN SQUARE	SUM OF SQ
VV	1/	17	9.4172	0.0070	2.2069	2.2069
SA	11/	187	34.4117	0.0000	16.5775	182.3530
VV SA	11/	187	14.2967	0.0000	5.6217	61.8389
PP	17				0.0000	0.0000
PP VV	17				0.2343	3.9839
PP SA	187				0.4817	90.0857
PP VV SA	187				0.3932	73.5316

SOURCE	DF1/	DF2	F	P VALUE	MEAN SQUARE	SUM OF SQ
VV	1/	11	0.3926	0.5437	2.2069	2.2069
PP	17/	187	0.0000	1.0000	0.0000	0.0000
VV PP	17/	187	0.5960	0.8925	0.2343	3.9839
SA	11				16.5775	182.3530
SA VV	11				5.6217	61.8389
SA PP	187				0.4817	90.0857
SA VV PP	187				0.3932	73.5316

Instruktionen zum Perzeptionstest (1).

Willkommen zu diesem Test.

In diesem Experiment soll die Qualität von zwei verschiedenen Sprachsynthesen verglichen werden.

Im ersten Test wird die Synthese einzelner Wörter getestet. Sie hören achtzig Stimuli. Sie bestehen aus je einer Zweiergruppe von verschieden synthetisierten Wörtern (Wort A und Wort B). Bitte entscheiden Sie, welches Wort Ihnen besser gefällt (welches natürlicher oder verständlicher ist). Dieses Wort kreuzen Sie bitte in Ihrem Fragebogen an.

Zum Beispiel: Version B von Stimulus 1 finden Sie besser:

Stimulus	A	B
-----		-----
1		X

Vor Beginn des Testes werde ich Ihnen zwei Beispiele vorspielen, damit Sie sich vorstellen können, wie synthetische Sprache klingt. Im Test gibt es nach je zehn Stimuli eine kurze Pause von etwa drei Sekunden und danach einen Piep. Wenn Sie mal aus dem Rhythmus kommen, ist es am Besten auf die nächste Pause zu warten, um dort weiterzumachen. Der Test wird etwa zehn Minuten dauern.

Viel Erfolg und viel Vergnügen!

Fragebogen für den ersten Teil.

Name der Versuchsperson:

Datum:

Stimulusnummer	A	B	Stimulusnummer	A	B
1			21		
2			22		
3			23		
4			24		
5			25		
6			26		
7			27		
8			28		
9			29		
10			30		
11			31		
12			32		
13			33		
14			34		
15			35		
16			36		
17			37		
18			38		
19			39		
20			40		

Fragebogen für den ersten Teil.

Stimulusnummer	A	B	Stimulusnummer	A	B
41			61		
42			62		
43			63		
44			64		
45			65		
46			66		
47			67		
48			68		
49			69		
50			70		
51			71		
52			72		
53			73		
54			74		
55			75		
56			76		
57			77		
58			78		
59			79		
60			80		

Experiment 1, zweiter Teil: Sätze

In der zweiten Hälfte des Experiments wird die Synthese ganzer Sätze getestet. Sie werden insgesamt 24 Stimuli hören. Die Stimuli bestehen diesmal aus je einem Satz, den Sie beurteilen sollen.

Bitte entscheiden Sie, wie gut Sie den Satz finden. Sie können für jeden Satz die Noten 1 bis 6 vergeben. Ist ein Satz natürlich und verständlich, dann bekommt er die Note 1. Finden Sie die Qualität eines Satzes sehr schlecht, dann bekommt er die Note 6.
Ein Beispiel:

Sie finden Satz 7 gut, deshalb bekommt er von Ihnen die Note 2.

Satznummer	Note					
7	1	2	3	4	5	6

Vor Beginn des Testes werde ich Ihnen zwei Beispielsätze vorspielen. Nach jedem Stimulus gibt es eine Pause von zwei Sekunden, in der Sie Ihre Wahl treffen. Vor jedem weiteren Stimulus hören Sie einen Piep. Der Test wird etwa fünf Minuten dauern. Viel Spaß und viel Erfolg!

Name des Testteilnehmers:.....

Datum:.....

Satznummer	Note					
1	1	2	3	4	5	6
2	1	2	3	4	5	6
3	1	2	3	4	5	6
4	1	2	3	4	5	6
5	1	2	3	4	5	6
6	1	2	3	4	5	6
7	1	2	3	4	5	6
8	1	2	3	4	5	6
9	1	2	3	4	5	6
10	1	2	3	4	5	6
11	1	2	3	4	5	6
12	1	2	3	4	5	6
13	1	2	3	4	5	6
14	1	2	3	4	5	6
15	1	2	3	4	5	6
16	1	2	3	4	5	6
17	1	2	3	4	5	6
18	1	2	3	4	5	6
19	1	2	3	4	5	6
20	1	2	3	4	5	6
21	1	2	3	4	5	6
22	1	2	3	4	5	6
23	1	2	3	4	5	6
24	1	2	3	4	5	6