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

Rule controlled data base search

Carlson, R. and Granström, B.

journal: STL-QPSR

volume: 26 number: 4

year: 1985 pages: 029-042



II. SPEECH SYNTHESIS

A. RULE CONTROLLED DATA BASE SEARCH*
Rolf Carlson and Björn Granström

Abstract

In current acoustic-phonetic research, there is a need for huge data bases. There are considerable problems in administering such data bases, both to transcribe and segment the speech and to easily access stored material. We have created a speech analysis system to attempt to alleviate these problems. Speech data are stored in sentence sized files. These files are segmented and transcribed semi-automatically given a phonetic transcription of the utterance. This transcription is generated by the text-to-phonetic component of our synthesis system. The same rule structure, similar to the notation used in generative phonology, is used for accessing the data. By a brief rule statement, speech segment meeting the specified contextual conditions can be identified. Durational data can be collected directly during the data base search. Spectral analysis programs operating with a variety of spectral representations have also been created that display the result, typically as a mean/SD spectrum or as a contour histogram spectrum.

Introduction

To develop detailed acoustic-phonetic knowledge we want to access large speech data bases according to linguistically meaningful principles. We have developed a method in which generative rules form an integral part of the data base management. There are considerable problems in administering such data bases, both to transcribe and segment the speech and to easily access stored material. We have created a speech analysis system to attempt to alleviate some of these problems. In this paper we will present the main function of this system. The data presented should be regarded merely as examples. Both the data base and the programs for handling it is still in continuous development.

A block diagram is shown in Fig. 1 that illustrates the method. The approach can be divided into three parts:

- * how the data base is created.
- * how the data base is searched.
- * how the data base can be analyzed.

^{*)} Summary of a paper presented at the Acoustical Society of America fall meeting 1985

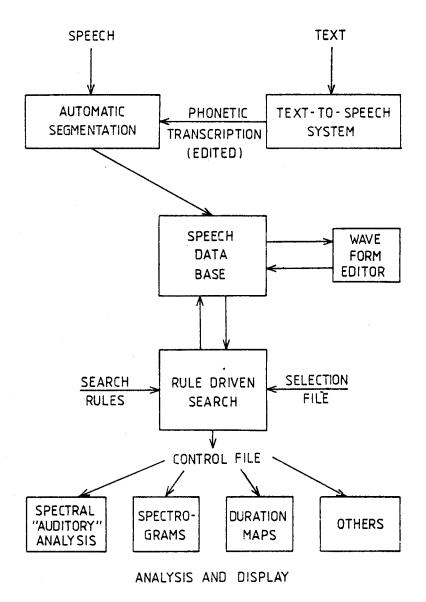


Fig. 1. Block diagram of the rule controlled data base search.

TEXT: HULLER OM BULLER.

PHONETIC: H'ULA4R 'AM+ B'ULA4R.

96 0	£H	.0599 sec
1809	Ú	.1130 sec
2911	L	.1819 sec
5048	Ä4	.3154 sec
5909	R	.3693 sec
6333	ÆÅ	.3958 sec
7798	M+	.4873 sec
8509	£B	.5318 sec
9342	'υ	•5838 sec
11472	L	.7169 se c
12122	Ä4	.7576 sec
14040	R	.8774 sec
16000	£.	1.0000 sec

Figure 2. Example of a label file.

orani sakosti sastata ka

We will in the following describe these parts individually and give some illustrative examples. The test data base in our example consists of 150 Swedish sentences, containing about 5000 phonemes, read by one male speaker. The same material has earlier been used in a redundancy study (Hunnicutt, 1985).

Creating the data base

The first step to create the data base is to record and label speech. Speech data is stored in sentence sized files. Our text-to-speech system is used to create a phonetic transcription of the utterances (Carlson, Granström, & Hunnicutt, 1982). This transcription is edited to match the pronunciation as good as possible. A phoneme can be transcribed by one or two characters. Stress level is marked by special signs. Additional markers like syntactic boundaries can be added if needed.

The phonetic transcription is used by an automatic segmentation program (Blomberg & Elenius, 1985) to distribute the phonetic labels along the wave form. The segmentation program gives an estimate of the time positions of each phoneme. Segmentation of speech in phone-sized parts in an unambiguous way is a classical problem, possibly without a solution. When several persons are contributing to the data base, it is important that the same criteria are used throughout. For that reason it is attractive to leave the segmentation to a self-consistent algorithm. The accuracy of the present program is, however, not enough.

When a detailed analysis should be done, the labels have to be checked and corrected. This is done with by means of a wave form editor program, which is a general purpose program for labeling and editing sampled files. It is able to handle multiple channels and has been used for labeling EMG signals beside speech and music wave forms. The editor is screen oriented and has two different displays. The lower half of the screen shows a condensed representation of the signal. A cursor marks the position of the detailed plot at the top of the screen, in which each sample can be seen. The editor is controlled by a joystick in addition to keyboard commands. Selected parts of the signal could be replayed during the editing process.

By means of the joystick samples can be labeled or labels can be changed. These can be used to edit the wave form, to include wave forms from other recordings, to delete parts of it, and to do scalings of amplitudes. The labels are stored in label files which are used by all following programs to be described. An example of a label file is shown in Fig. 2. At the top the text is shown. The rule generated phonetic transcription is displayed below the text and is followed by the sample number associated to each label.

During the editing the program can suggest good positions for labels. This is done by an automatic procedure that places the cursor at zero crossings or at the closing time of the glottal source. These features make the program fast, interactive, and user friendly.

Rule-driven search

The data base is accessed by means of rules. By a brief rule statement, speech segments meeting the specified contextual conditions can be identified. The rule structure is similar to the notation used in generative phonology and is also used in the text-to-speech system (Carlson & Granström, 1975). A definition file specifies the phonetic inventory and the features that are associated to each phoneme, Fig. 3.

The rules are used to insert a '*' symbol in front of the phoneme to be analyzed and to give it a set of parameter values. These parameters can be used to specify the time position for each phoneme, the duration of the phoneme, the stress level, or any information that can be derived from the phonetic transcription or the durational information in the label file. Fig. 4 gives an example of a simple rule system to find all representations of the tense vowel /a:/. The midpoint of the vowel is stored in parameter A and the duration in parameter B.

A special feature of the system is that the notation itself is a powerful tool to describe a model such as a text-to-speech system. The model prediction can, thus, be part of the data bank search and the difference between the model, and the actual data can be derived. Some examples will be given in this paper.

A selection file specifies which label files that form the data base. The label files are read into the program one by one, processed by the rules and the result is stored in a control file. The control file contains the phonetic transcription with the star positions marked. Each star is then decoded into a control line including the wave form file name, the specified parameters and the name of the phonetic symbol that follows the star. This control file, Fig. 5, is used by all analyzing programs described below.

Spectral analysis

As a first example we will make a spectral analysis of the 384 occurrences of the /a:/ phoneme in the data base. The control file produced by the rule system in Fig. 4 was used to find the midpoint of the vowels. At each point an FFT was made and the mean and standard deviations were calculated. The result is shown in Fig. 6a. It is interesting to note that the formant positions are well marked even after such a heavy pooling of data. One reason for this consistency is that this vowel mainly occurs in stressed position and that reduction phenomena have a small effect.

Fig. 6b shows a more detailed representation. A histogram is calculated for the energy distribution at each point along the frequency axis. These histograms are connected into a contour spectrum, each line representing 10 percent of the distribution. The spectrum is more condensed at low compared to high frequency indicating that the glottal source has a steeper slope for low effort whereas the amplitude of the voice is relatively independent on effort. If we use a broader category

- Marine

```
01.00 DEF A::=<-STRESS,-ISTRESS,-GRAV,+HSTR,-MSTR,-LSTR,+VOC,
- DIPH,-CONS,+TENSE,-HIGH,+LOW,+BACK,-ROUND,-OBST,+CONT,
- NAS,+VOICE,+SEG,-SYNTAX,-DIG>
```

- 01.01 DEF A:=<-STRESS,-ISTRESS,-GRAV,-HSTR,-MSTR,-LSTR,+VOC,
 DIPH,-CONS,-TENSE,-HIGH,+LOW,+BACK,-ROUND,-OBST,+CONT,
 NAS,+VOICE,+SEG,-SYNTAX,-DIG>
- 01.02 DEF E::=<-STRESS,-1STRESS,-GRAV,+HSTR,-MSTR,-LSTR,+VOC,
 +DIPH,-CONS,+TENSE,-HIGH,-LOW,-BACK,-ROUND,-OBST,+CONT,
 NAS,+VOICE,+SEG,-SYNTAX,-DIG>
- 01.03 DEF E:=<-STRESS,-1STRESS,-GRAV,-HSTR,-MSTR,-LSTR,+VOC,
 DIPH,-CONS,-TENSE,-HIGH,-LOW,-BACK,-ROUND,-OBST,+CONT,
 NAS,+VOICE,+SEG,-SYNTAX,-DIG>
- 02.00 DEF B:=<-TENSE,-STRESS,-VOC,+CONS,+OBST,-CONT,-NAS,+VOICE,
 +ANT,-COR,+SEG,-SYNTAX,-DIG>
- 02.01 DEF P:=<-TENSE,-STRESS,-VOC,+CONS,+OBST,-CONT,-NAS,-VOICE,
 +ANT,-COR,+SEG,-SYNTAX,-DIG>
- 02.13 DEF S:=<-TENSE,-STRESS,-VOC,+CONS,+OBST,+CONT,-NAS,-VOICE,
 +ANT,+COR,FRIC,+SEG,-SYNTAX,-DIG>
- 02.14 DEF N:=<-TENSE,-STRESS,-VOC,+CONS,-OBST,-CONT,+NAS,+VOICE,
 +ANT,+COR,+SEG,-SYNTAX,-DIG>

Figure 3. Example of a definition file.

insert * in front of a: 01.00: ^ * / & A:

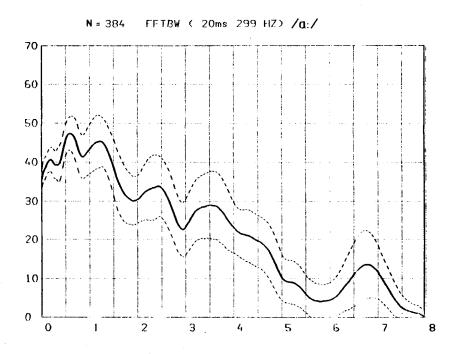
save the position and the duration of the vowel in the *
02.00: * ^ <POS:=X,DUR=Y> / & <VOC,SEG,X:=POS,Y:=DUR>

Figure 4. Example of a rule system to find /a:/

P*Å:*U:*T*F*L*Y*K*T*EO*N*G*R*I*L*A*D*EO*B*A:*2N*EO*N*G*L*A*T*K*Å*R*V*EO*N.

WD9001 13 -1 5 U: 13 1: WD9001 25 -2 2 T 9 WD9001 34 0 2 F 6 WD9001 41 1 2 L 4 6 WD9001 44 0 4 Y 6 6 WD9001 50 -3 2 K 10 10 10 WD9001 60 -2 2 T 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	WD9001	6	0	3	A:	7	7
WD9001 25 -2 2 T 9 WD9001 34 0 2 F 6 WD9001 41 1 2 L 4 WD9001 50 -3 2 K 10 WD9001 60 -2 2 T 9 WD9001 68 -1 3 EO 8 WD9001 86 0 2 G 5 WD9001 96 0 6 I 10 1 WD9001 107 5 2 L 3 8 WD9001 107 5 2 L 3 8 WD9001 110 0 4 A 8 9 WD9001 121 0 3 EO 6 6 WD9001 128 0 2 B 6 6 WD9001 153 1 2 2N 4 6 WD9001 162 0 2 N 6							12
WD9001 34 0 2 F 6 WD9001 41 1 2 L 4 WD9001 50 -3 2 K 10 WD9001 60 -2 2 T 9 WD9001 68 -1 3 EO 8 WD9001 86 0 2 G 5 WD9001 96 0 6 I 10 1 WD9001 107 5 2 L 3 6 WD9001 107 5 2 L 3 6 WD9001 110 0 4 A 8 9 WD9001 118 2 2 D 4 6 WD9001 128 0 2 B 6 6 WD9001 153 1 2 2N 4 6 WD9001 162 0 2 N 6 6 WD9001 168 -1 2 G							7
WD9001 41 1 2 L 4 6 WD9001 50 -3 2 K 10 WD9001 60 -2 2 T 9 WD9001 68 -1 3 EO 8 WD9001 76 -2 2 N 9 WD9001 86 0 2 G 5 WD9001 90 -1 2 R 6 WD9001 96 0 6 I 10 1 WD9001 107 5 2 L 3 9 WD9001 110 0 4 A 8 9 WD9001 118 2 2 D 4 6 WD9001 128 0 2 B 6 6 WD9001 153 1 2 2N 4 6 WD9001 162 0 2 N 6 6 WD9001 168 -1 2 G							7
WD9001 44 0 4 Y 6 WD9001 50 -3 2 K 10 WD9001 60 -2 2 T 9 WD9001 68 -1 3 EO 8 WD9001 76 -2 2 N 9 WD9001 86 0 2 G 5 WD9001 90 -1 2 R 6 WD9001 96 0 6 I 10 1. WD9001 107 5 2 L 3 9 WD9001 110 0 4 A 8 9 WD9001 118 2 2 D 4 6 WD9001 128 0 2 B 6 6 WD9001 153 1 2 2N 4 6 WD9001 162 0 2 N 6 6 WD9001 168 -1 2 G 7							6
WD9001 50 -3 2 K 10 MD9001 60 -2 2 T 9 MD9001 68 -1 3 EO 8 MD9001 76 -2 2 N 9 MD9001 9 -1 2 R 6 9 MD9001 9 -1 2 R 6 9 9 -1 2 8 6 9 9 -1 2 8 9 9							7
WD9001 60 -2 2 T 9 1<			-3	2			7
WD9001 76 -2 2 N 9 WD9001 86 0 2 G 5 WD9001 90 -1 2 R 6 WD9001 107 5 2 L 3 WD9001 110 0 4 A 8 WD9001 118 2 2 D 4 WD9001 121 0 3 EO 6 WD9001 128 0 2 B 6 WD9001 153 1 2 2N 4 WD9001 157 0 3 EO 5 WD9001 162 0 2 N 6 WD9001 168 -1 2 G 7 WD9001 182 0 6 A 10 1 WD9001 192 -1 2 T 12 1 WD9001 203 -2 2 K 11 9 WD9001 214			-2				7
WD9001 76 -2 2 N 9 WD9001 86 0 2 G 5 WD9001 90 -1 2 R 6 WD9001 107 5 2 L 3 WD9001 110 0 4 A 8 WD9001 118 2 2 D 4 WD9001 121 0 3 EO 6 WD9001 128 0 2 B 6 WD9001 153 1 2 2N 4 WD9001 157 0 3 EO 5 WD9001 162 0 2 N 6 WD9001 168 -1 2 G 7 WD9001 182 0 6 A 10 1 WD9001 192 -1 2 T 12 1 WD9001 203 -2 2 K 11 9 WD9001 214			-1				7
WD9001 86 0 2 G 5 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	WD9001		-2		N	9	7
WD9001 90 -1 2 R 6 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9		86			G	5	5
WD9001 96 0 6 I 10 13	WD9001	90	-1		R		5
WD9001 110 0 4 A 8 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	WD9001	96	0	6	I	10	11
WD9001 110 0 4 A 8 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9		107	5	2	L	3	9
WD9001 121 0 3 EO 6 6 WD9001 128 0 2 B 6 6 WD9001 134 -5 5 A: 19 14 WD9001 153 1 2 2N 4 6 WD9001 157 0 3 EO 5 6 WD9001 162 0 2 N 6 6 WD9001 168 -1 2 G 7 6 WD9001 175 -1 2 L 7 6 WD9001 182 0 6 A 10 1 WD9001 192 -1 2 T 12 1 WD9001 214 3 6 A 8 12 WD9001 222 -8 2 R 13 5 WD9001 235 -1 2 V 6 6 WD9001 241 0 3 EO 6 6 <td></td> <td>110</td> <td>0</td> <td>4</td> <td>Α</td> <td>8</td> <td>9</td>		110	0	4	Α	8	9
WD9001 121 0 3 EO 6 6 WD9001 128 0 2 B 6 6 WD9001 134 -5 5 A: 19 14 WD9001 153 1 2 2N 4 6 WD9001 157 0 3 EO 5 6 WD9001 162 0 2 N 6 6 WD9001 168 -1 2 G 7 6 WD9001 175 -1 2 L 7 6 WD9001 182 0 6 A 10 1 WD9001 203 -2 2 K 11 9 WD9001 214 3 6 Å 8 12 WD9001 235 -1 2 V 6 6 WD9001 241 0 3 EO 6 6	WD9001	118	2	2	D	4	6
WD9001 134 -5 5 A: 19 14 WD9001 153 1 2 2N 4 6 WD9001 157 0 3 EO 5 6 WD9001 162 0 2 N 6 WD9001 168 -1 2 G 7 WD9001 175 -1 2 L 7 WD9001 182 0 6 A 10 13 WD9001 192 -1 2 T 12 13 WD9001 203 -2 2 K 11 9 WD9001 214 3 6 Å 8 12 WD9001 222 -8 2 R 13 WD9001 235 -1 2 V 6 WD9001 241 0 3 EO 6	WD9001	121	0		EO	6	6
WD9001 134 -5 5 A: 19 14 WD9001 153 1 2 2N 4 6 WD9001 157 0 3 EO 5 6 WD9001 162 0 2 N 6 6 WD9001 168 -1 2 G 7 6 WD9001 175 -1 2 L 7 6 WD9001 182 0 6 A 10 13 WD9001 192 -1 2 T 12 13 WD9001 203 -2 2 K 11 9 WD9001 214 3 6 Å 8 12 WD9001 222 -8 2 R 13 9 WD9001 235 -1 2 V 6 9 WD9001 241 0 3 EO 6	WD9001	128	0	2	В	6	7
WD9001 153 1 2 2N 4 6 6 6 6 WD9001 162 0 2 N 6 6 6 WD9001 168 -1 2 G 7 6 WD9001 175 -1 2 L 7 6 WD9001 182 0 6 A 10 13 WD9001 192 -1 2 T 12 13 WD9001 203 -2 2 K 11 9 WD9001 214 3 6 Å 8 12 WD9001 222 -8 2 R 13 9 WD9001 235 -1 2 V 6 9 WD9001 241 0 3 EO 6	WD9001	134	 5		A:	19	14
WD9001 162 0 2 N 6 6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	WD9001	153			2N	4	6
WD9001 168 -1 2 G 7 6 WD9001 175 -1 2 L 7 6 WD9001 182 0 6 A 10 1	WD9001	157	0	3	EO	5	6
WD9001 175 -1 2 L 7 6 WD9001 182 0 6 A 10 13 WD9001 192 -1 2 T 12 13 WD9001 203 -2 2 K 11 9 WD9001 214 3 6 Å 8 13 WD9001 222 -8 2 R 13 9 WD9001 235 -1 2 V 6 9 WD9001 241 0 3 E0 6 6		162	0		N	6	6
WD9001 182 0 6 A 10 13 WD9001 192 -1 2 T 12 13 WD9001 203 -2 2 K 11 9 WD9001 214 3 6 Å 8 13 WD9001 222 -8 2 R 13 9 WD9001 235 -1 2 V 6 9 WD9001 241 0 3 E0 6				2	G	7	6
WD9001 192 -1 2 T 12 13 WD9001 203 -2 2 K 11 9 WD9001 214 3 6 Å 8 12 WD9001 222 -8 2 R 13 WD9001 235 -1 2 V 6 WD9001 241 0 3 EO 6					L	7	6
WD9001 203 -2 2 K 11 9 WD9001 214 3 6 Å 8 12 WD9001 222 -8 2 R 13 9 WD9001 235 -1 2 V 6 9 WD9001 241 0 3 EO 6		182	0		Α		11
WD9001 214 3 6 Å 8 12 WD9001 222 -8 2 R 13 5 WD9001 235 -1 2 V 6 5 WD9001 241 0 3 EO 6					${f T}$	12	11
WD9001 222 -8 2 R 13 5 WD9001 235 -1 2 V 6 5 WD9001 241 0 3 EO 6						11	9
WD9001 235 -1 2 V 6 9 WD9001 241 0 3 EO 6		214	3		Å	8	12
WD9001 241 0 3 EO 6 6					R	13	5
					V	6	5
WD9001 247 -1 2 N 7 6							6
	WD9001	247	-1	2	N	7	6

Figure 5. Example of control file.



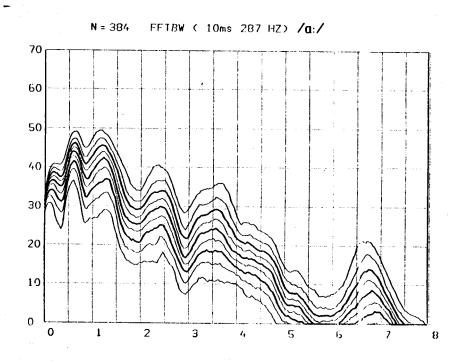


Fig. 6. Spectral analysis of 384 occurrences of /a:/ with mean/SD representation (a) and contour histogram representation (b).

and analyze all 1948 vowels in the data base, we will get similar results (Fig. 7). Fig. 8 shows examples of consonant spectra derived with the same method. By studying these kinds of representations we can see how consistent each phoneme is realized, and coarticulation effects can be searched for. It could be seen that the /s/ sometimes is voiced even if such a voiced fricative is not regarded to be part of the Swedish phonetic system.

The spectral analysis can be done with a variety of spectral representations including auditory modeling (Carlson & Granström, 1982). Fig. 9 shows two examples where the frequency axes are transformed into a Bark scale and the energy is measured in dB or phon.

Duration analysis

Already during the data base search, the duration can be measured for each phoneme. In the rule system, shown in Fig. 10, a classifier is used to group duration data on vowels depending on the stress level, phonological length, and whether the vowel is followed by an unvoiced stop or not. The result is shown in Fig. 11. As expected, vowels get shortened in front of the stops, and unstressed vowels are shorter than stressed. The example is chosen to show the power of the rule based approach. More sophisticated relations can be studied by more complex rule systems.

Duration relations and prediction of duration

Normally, when duration is studied, we are more interested in relations between phonemes than the absolute duration. Fig. 12 shows a duration map of such an analysis. The purpose was to examine if a consonant vowel sequence could be regarded as a prosodic unit. Along the x-axes the duration is of a phonologically short vowel and along the y-axes the duration is of a preceding consonant. The consonant is not part of a cluster and is not a semi-vowel. The data has been subjected to a two-dimensional smoothing. Unfiltered data can be seen in the scatter plot in Fig. 13. The inherent duration of each consonant category plays an important role, and the separation of fricatives and nasals give a more coherent picture. It is possible to formulate the relations between fricative/vowel and nasal/vowel in two rules, and these rules can be tested against our data. The result is shown in Fig. 14. The actual duration measured in the data base is plotted along the x-axes and the prediction along the y-axes. The duration of the following vowel is used as parameter in this prediction.

Testing of rule systems against the data base

Given the power of rule-formulated models and rule-formulated data base search we are able to test complete durational models against actual data in a fast and effective manner. The difference between N=1948 FFTBW (20ms 299 HZ)

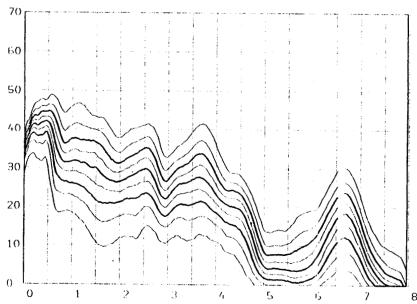
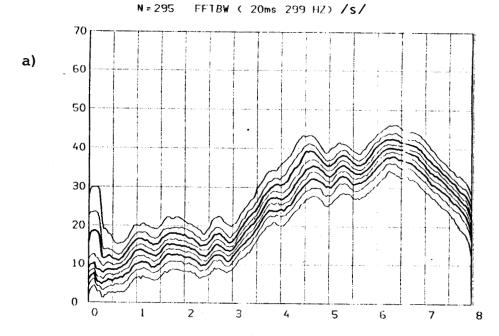
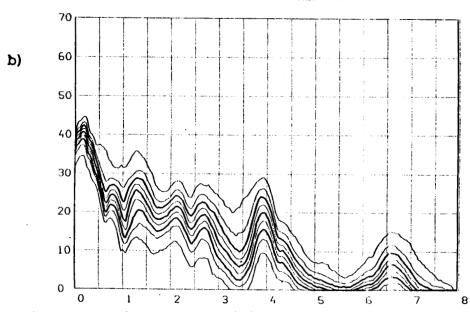
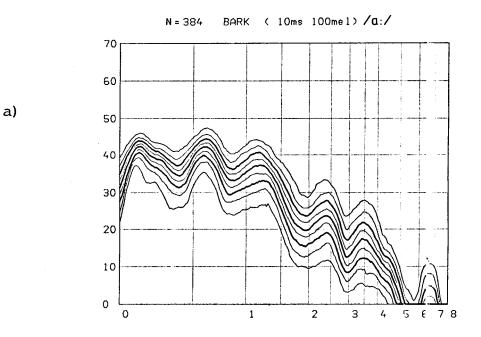


Fig. 7. Contour histograms for all 1948 vowels in the data base.



N = 477 FFTBW (20ms 299 HZ) /n/





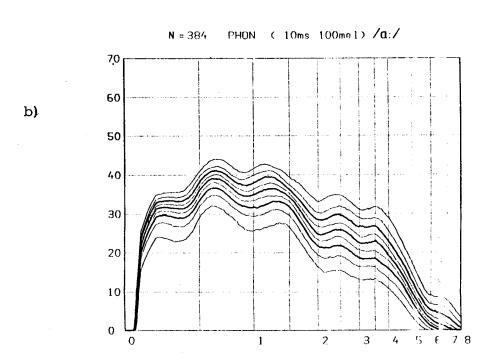


Fig. 9. Contour histograms with a dB/mel (a) and a phon/mel (b) representation.

insert * in front of vowels 01.00: ^ * / & <VOC,SEG>

save the duration of the vowel in the * 02.00: * ^ <DUR:=Y> / & <VOC,SEG,Y:=DUR>

give all vowels class 1
03.00: * ^ <CLASS:=1> / & <VOC,SEG>

give class 2 to short vowels with primary stress 04.00: * ^ <CLASS:=2> / & <VOC,SEG,STRESS,1STRESS,-TENSE>

give class 3 to short vowels with secondary stress 05.00: * ^ <CLASS:=3> / & <VOC,SEG,STRESS,-ISTRESS,-TENSE>

give class 4 to long stressed vowels
06.00: * ^ <CLASS:=4> / & <VOC,SEG,STRESS,TENSE>

add 4 to the class if the vowel is before of voiceless stop

07.00: * ^ <CLASS=CLASS+4> /
- & <VOC,SEG> <CONS,SEG,-CONT,-NAS,-VOICE>

Figure 10. Rule system to find and classify vowels.

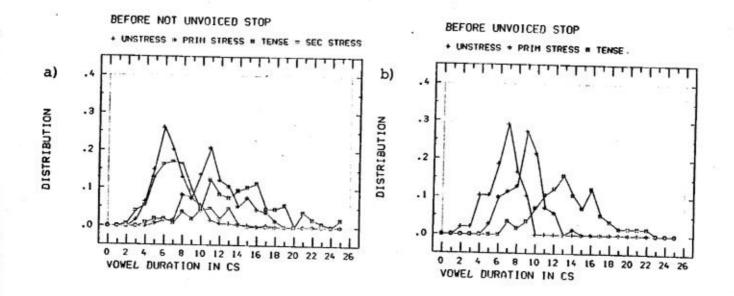


Fig. 11. Result of duration analysis of vowels in front of non unvoiced stops (a) and in front of unvoiced stops (b).

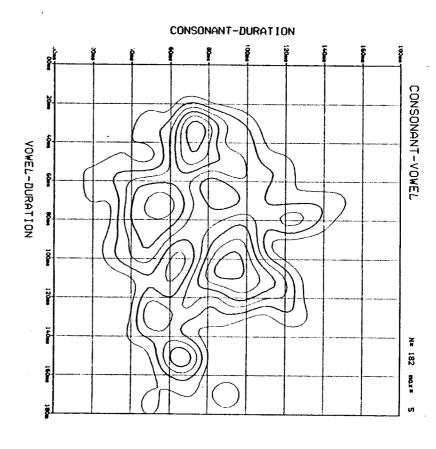


Fig. 12. Duration map of consonant-vowel duration relation.

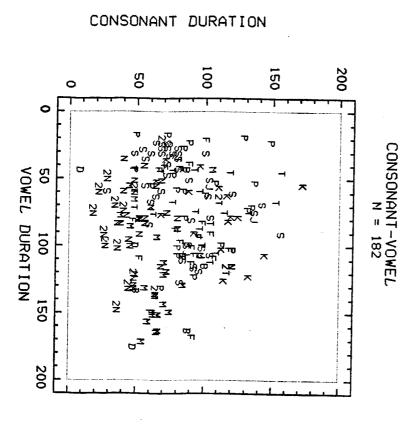


Fig. 13. Scatter plot of consonant-vowel duration relation.

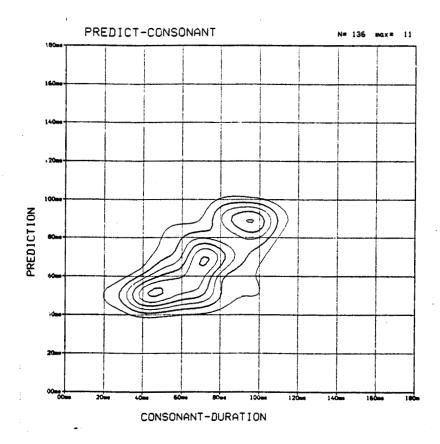


Fig. 14. Predicted and measured consonant duration.

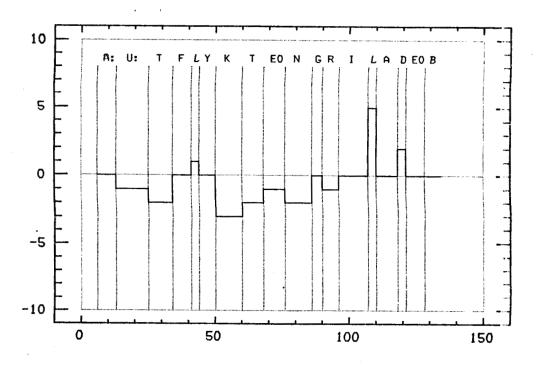


Fig. 15. Error analysis of predicted/measured phoneme duration.

predicted and measured data can be plotted as in Fig. 15 to study systematic errors. The errors can be plotted in scatter plots, three-dimensional duration maps, or two-dimensional curves.

Conclusion

We have developed a system to access a data base in an effective manner by means of rules. These rules can also be used to describe models that can be tested against the data. It has not been the goal in this paper to present scientifically important results. Rather, it has been our intention to illustrate the method and to show the power of the approach. Many components will be added to present the data in more informative manners but the current system enables us to test hypotheses and to transform the gained knowledge to our text-to-speech system or speech recognition system in a fast and effective manner.

References

Blomberg, M. & Elenius, K. (1985): "Automatic time alignment of speech with a phonetic transcription", STL-QPSR 1/1985, pp. 37-45.

Carlson, R. & Granström, B. (1975): "A phonetically oriented programming language for rule description of speech", pp. 245-253 in Speech Communication, Vol. 2 (ed. G. Fant), Almqvist & Wiksell, Stockholm.

Carlson, R. & Granström, B. (1982): "Towards an auditory spectrograph", pp. 109-115 in The Representation of Speech-in the Peripheral Auditory System, (eds. R. Carlson & B. Granström), Elsevier/North-Holland Biomedical Press, Amsterdam.

Carlson, R. & Granström, B. (1985): "Rule-controlled data base search", J. Acoust. Soc. Am. 78, S54(A).

Carlson, R. & Granström, B. (1986): "51.4", Conference Record, IEEE-ICASSP, Tokyo.

Carlson, R., Granström, B., & Hunnicutt, S. (1982): "A multi-language text-to-speech module", Conference Record, IEEE-ICASSP, Paris.

Hunnicutt, S. (1985): "Intelligibility versus redundancy - conditions of dependency", pp 47-56, Language and Speech 28: 1, pp. 47-56.

But the first of t