

# Speechdat(II) GERMAN Database

for the fixed Telephone network

Author(s):	Christoph Draxler
------------	-------------------

Institute:	Department of Phonetics, University of Munich
------------	---

Address:	Schellingstr. 3, D 80799 Munich
----------	---------------------------------

email:	draxler@phonetik.uni-muenchen.de
--------	----------------------------------

Date:	16.9.97
-------	---------

Version:	1.1 updated according to validation report of August 1997 1.0
----------	--

# CONTENTS

<b>1. INTRODUCTION.....</b>	<b>3</b>
1.1 SPEECH FILE FORMAT .....	3
1.2 FILE NOMENCLATURE .....	3
1.3 DIRECTORY STRUCTURE.....	5
1.4 LABEL FILES.....	6
<b>2. DATABASE DESIGN AND COLLECTION.....</b>	<b>7</b>
2.1 RECORDING SITE AND PLATFORM .....	7
2.2 SPEAKER RECRUITMENT .....	7
2.3 DESIGN OF PROMPTING AND PROMPT-SHEET .....	7
2.4 TRANSCRIPTION.....	7
<b>3. DATABASE CONTENTS DEFINITION.....</b>	<b>8</b>
3.1 APPLICATION WORDS .....	9
3.2 ISOLATED DIGITS .....	10
3.3 CONNECTED DIGITS.....	10
3.4 DATES .....	11
3.5 EMBEDDED APPLICATION WORD.....	12
3.6 SPELLED NAMES/WORDS .....	16
3.7 MONEY AMOUNT .....	17
3.8 NATURAL NUMBER .....	17
3.9 DIRECTORY ASSISTANCE NAMES .....	17
3.10 YES/NO QUESTIONS .....	17
3.11 PHONETICALLY RICH SENTENCES.....	18
3.12 TIMES .....	18
3.13 PHONETICALLY RICH WORDS .....	20
3.14 ANY OTHER ADDITIONAL MATERIAL.....	20
3.15 LINKS TO OTHER DATABASES .....	21
<b>4. SPEAKER DEMOGRAPHIC INFORMATION.....</b>	<b>21</b>
4.1 ACCENT/REGIONS .....	21
4.2 SPEAKER CHARACTERISTICS.....	21
<b>5. THE LEXICON.....</b>	<b>22</b>
<b>6. RECORDING CONDITIONS.....</b>	<b>22</b>
6.1 ENVIRONMENTS.....	22
<b>7. DEVIATIONS FROM SPEECHDAT SPECIFICATIONS .....</b>	<b>23</b>
<b>8. SAMPLE PROMPT SHEETS.....</b>	<b>23</b>
8.1 SAMPLE INSTRUCTION SHEET AND PROMPT SHEET .....	23
8.2 SAMPLE INSTRUCTION SHEET AND PROMPT SHEET (ENGLISH TRANSLATION).....	23

## 1. Introduction

SpeechDat(II) is a telephone speech database collection sponsored by the EU under contract number LE2-4001. This document describes the first 1000 recordings of a total of 4000 German Recordings in SpeechDat(II).

The German Fixed Network 1000 database collection is carried out by the Department of Phonetics and Speech Communication of the University of Munich under a subcontract with SIEMENS AG, Munich.

The German Fixed Network 1000 database consists of 4 CD-ROMs:

CD-ROM	Begin Block	End Block
FIXED1DE_01	BLOCK10	BLOCK12
FIXED1DE_02	BLOCK13	BLOCK15
FIXED1DE_03	BLOCK16	BLOCK18
FIXED1DE_04	BLOCK19	BLOCK21

**Table 1 – CD-ROM contents**

### 1.1 Speech file format

The speech file format is 8bit, 8kHz, A-law speech files, uncompressed.

Each speech file has an accompanying ASCII SAM label file with the same file name but a different file extension in the same directory as the speech file.

### 1.2 File nomenclature

File names follow the ISO 9660 file name conventions (8 plus 3 characters) according to the main CD-ROM standard. The following template is used:

DD NNNN CC. LL F

where:

DD	Database identification code (00-ZZ) For SpeechDat: A1=fixed network recordings; B1=mobile network recordings; C1=speaker verification database
NNNN	Recording session progressive number (0000-9999)
CC	Corpus code (A1-Z9) obtained by collating the corpus and the item identifiers
LL	Two letter ISO 639 language code
F	File type code O=Orthographic label file, A=A-law coded speech file

**Table 2 - SpeechDat filename convention**

Thus, for the German Fixed Network Database, the file names are

A1 NNNN CC. DE A and A1 NNNN CC. DE O

respectively.

As it is useful for users to clearly identify the speech file contents by looking at the filename we have specified the corpus code to support one letter corpus identifier and one letter item identifier, by the following table. All items are read, unless marked as spontaneous.

Number	Item Type	Specification	Mnemonic
2	isolated digit	isolated digit	I1
		sequence of 10 isolated digits	B1
4	digit/number	prompt sheet number 5 digits	C1
		telephone number (9-11 digits)	C2
		credit card number (15-16 digits)	C3
		PIN code (6 digits)	C4
1	natural number		N1
1	money amount	currency amount, mixed size and units	M1
2	yes/no questions	spontaneous predominantly yes	Q1
		spontaneous predominantly no	Q2
3	dates	spontaneous date, e.g. birthdate	D1
		prompted text form	D2
		relative and general date form	D3
2	times	time of day, spontaneous	T1
		prompted, mixed analogue/digital	T2
3	application words	keywords/keyphrases	A1-3
1	word spotting phrase	using embedded application words	E1
5	directory assistance	names spontaneous (e.g. forename)	O1
		city of birth/growing up (spontaneous)	O2
		set of 150 SDB full names	O7
		most frequent cities	O3
		most frequent , companies, agencies	O5
3	spellings	spontaneous spelling, e.g. forename	L1
		directory city name	L2
		real/artificial word	L3
4	isolated words		W1-4
9	phonetically rich sentences		S1-9

**Table 3 - SpeechDat filename convention**

The proposed format uses mnemonic values. It permits selection of all files belonging to one of the twelve corpora by using one command (e.g. in DOS “dir /s/b ??????C\*”, in UNIX “find . -name “??????C\*” -print”).

### 1.3 Directory structure

The directory structure uses a shallow directory nesting with contiguous numbers to identify the individual sub-directories and call directories. The following three-levels directory structure is defined:

\FIXED1DE\

Where:

FIXED1DE	Defined as: <name><#><language code> i.e. DE Where: <name> is FIXED indicating a fixed network database <#> is 1 for SpeechDat <LL> is the ISO 2-letters code DE for German
<block >	Defined as: BLOCK<nn> where <nn> is a progressive number from 00 to max. 99 These numbers are the same as the first 2 digits used in <nnnn> described below.
<session >	Defined as: SES<nnnn> Where <nnnn> is a progressive number in the range 0000 to max. 9999, being the numeric call identification number also encoded in each filename. As there are no more than 50 utterances per call, the total number of speech files and associated transcription files does not exceed the CD-ROM recommended limit of approximately 100 files in a directory.

**Table 4 - SpeechDat directory structure**

Both signal files and label files are put in the same directory.

All sessions have complete recordings for all prompted items with the following exceptions (see the file SUMMARY.TXT in directory DOC on the CD-ROMs).

In addition to the previous structure the following directories are used to store some other files:

\FIXED1DE\DOC	documentation files, including subword occurrence files
\FIXED1DE\TABLE	speaker and lexicon tables
\FIXED1DE\INDEX	index files - contents file
\FIXED1DE\PROMPT	prompt sheet tables
\FIXED1DE\SOURCE	source code for SAMLIB DOS/Unix file access routines

**Table 5 - Non-speech related directory structure**

Finally the root directory contains three files:

- a “README.TXT” ASCII file describing all files in the CD-ROM; signal and label files are reported by specifying their templates;
- a “DISK.ID” ASCII file containing the volume name (11 characters long); it supplies the volume label to UNIX systems that are unable to read the physical volume label, e.g. “FIXED1DE\_01”.

- a “COPYRIGH.TXT” ASCII file to protect the authors rights.

All these support files are duplicated in each CD-ROM.

#### 1.4 Label files

The labels files use US-7bit ASCII SAM labels. The meaning of the labels is as defined in SD1.3.1. A sample file is given below.

```
LHD:SAM, 5.10
DBN:SpeechDat_German_Fixed_Network
VOL:FIXED1DE
SES:0003
DIR:\FIXED1DE\BLOCK00\SES0003
SRC:A10003A1.DEA
CCD:A1
SHT:4294-8
BEG:
END:24000
REP:Dept. of Phonetics, University of Munich, Germany
RED:14/Dec/1997
RET:18:14:00
SAM:8000
SNB:1
SSB:8
QNT:A-LAW
SCD:UNKNOWN
SEX:F
AGE:30
ACC:BY
REG:UNKNOWN
ENV:OFFICE
NET:PSTN
PHM:TOUCH-TONE
LBD:
LBR:0,24000,,,Nachricht
LBO:0,12000,24000,Nachricht
ELF:
```

## **2. Database design and collection**

### **2.1 Recording site and platform**

The recording platform is a 32 MB 486-PC running SCO UNIX. The machine is equipped with a 1 GB internal and a 4 GB external hard disk, a Dialogic AG30 and an NMS ISDN primary rate interface board.

The software for recording was developed in house from the sample software provided with the boards. It has been used for the German SpeechDat(M) recordings already. The software allows a maximum of 6 parallel recordings; if more than 6 calls are received in parallel, all calls in excess of 6 receive a busy signal and are not answered.

Each call is given a unique ID-number, and no attempt is made to identify a caller. Hence, if a call is aborted, it will be incomplete on the hard disk.

### **2.2 Speaker recruitment**

Speakers are recruited in two ways:

- 1) letters are sent to Siemens department managers who are asked to persuade 10 employees to participate in the recordings
- 2) each caller receives a telephone card as an immediate reward and further sheets to distribute to family member, friends, and colleagues.

A third manner of recruitment is currently being tested: the Siemens internal magazine „SiemensWelt“ publishes calls for participation; people can request prompt sheets via phone, fax, or use sheets distributed via the Internet.

As a further incentive besides the telephone card, a prize of DM 1000.- is offered.

### **2.3 Design of prompting and prompt-sheet**

The prompting was taken over from SpeechDat(M) with a minor modification: the prompt sheet is divided into sections; these section numbers are spoken by the speech server and serve as an orientation for the speaker. Within the sections, there is no further explicit guidance by the speech server. However, since some prompts are read by the speech server, this helps speakers to stay synchronized with the speech server. Each recording is preceded by a short beep.

The limiting factor for the number of prompt sheets is 5 (phonetically rich words may be repeated by 5 speakers). Hence, each prompt sheet can be used up to 5 times if no phonetically rich word occurs on more than one prompt sheet. The individual items of the prompt sheet are spread out over the sheet in such a way that list effects are minimized, and that they fit in the layout. Most of the partner specific material is recorded last, so that not answering them does not affect the SpeechDat material.

Recordings are monitored closely to detect possible deviations from the demographic balance early. However, some oversampling is necessary to compensate for the time delay with which new prompt sheets can be sent out, and for unusable calls.

### **2.4 Transcription**

The transcription software used was developed in house. It is based on the WWW and thus allows transcriptions on any machine with a modern WWW browser. The signal is presented auditively only and can be output as often as requested. For ease of editing, some special conversion buttons have been implemented, e.g. digit to number string, spelling to letter names (A -> Anton, B -> Berta) etc.

Transcriptions are saved to the final position immediately.

For quality assurance, 5% of all transcriptions are checked by two transcribers (they do not see who transcribed the signal) and the result is recorded in a log file.

The character set used for the transcription is ISO 8859-1 (ISO Latin). A sample ISO 8859-1 table can be found in [ISO8859 file]

Transcriptions follow the usual German rules for upper and lower case, i.e. nouns, names and nominalized verbs are written with capital letters, everything else with lower case letters.

Punctuation marks are removed from the transcription.

The following symbols are used to denote word truncations, mispronunciations, non-understandable speech, non-speech acoustic events; other symbols:

word truncations	~abc or abc~		at signal begin or end only
mispronunciations	*abc		
non-understandable speech	**		separated by blank from rest of text
non-speech acoustic events	[fil]	filled pause	at correct location between words
	[spk]	speaker noise	at correct location between words
	[int]	intermittent noise	at correct location between words
	[sta]	stationary noise	at correct location between words

**Table 6 - Noise and Mispronunciation markers**

### 3. Database contents definition

The final specification for the German recordings is as follows:

Number	Item Type	Specification	Mnemonic
2	isolated digit	isolated digit	I1
		sequence of 10 isolated digits	B1
4	digit/number	prompt sheet number 5 digits	C1
		telephone number (9-11 digits)	C2
		credit card number (15-16 digits)	C3
		PIN code (6 digits)	C4
1	natural number		N1
1	money amount	currency amount, mixed size and units	M1
2	yes/no questions	spontaneous predominantly yes	Q1
		spontaneous predominantly no	Q2
3	dates	spontaneous date, e.g. birthdate	D1
		prompted text form	D2

		relative and general date form	D3
2	times	time of day, spontaneous	T1
		prompted, mixed analogue/digital	T2
3	application words	keywords/keyphrases	A1-3
1	word spotting phrase	using embedded application words	E1
5	directory assistance	names spontaneous (e.g. forename)	O1
		city of birth/growing up (spontaneous)	O2
		set of 150 SDB full names	O7
		most frequent cities	O3
		most frequent , companies, agencies	O5
3	spellings	spontaneous spelling, e.g. forename	L1
		directory city name	L2
		real/artificial word	L3
4	isolated words		W1-4
9	phonetically rich sentences		S1-9
4	partner specific material	speaker gender question	Y1
		birthdate request	Y3
		speaker region question	Y4
		today's date	Y5

**Table 7 - Database Contents**

40 items are mandatory for SpeechDat, and 4 are additional but included as part of the database.

### 3.1 Application words

<b>German</b>	<b>English</b>	<b>Meaning</b>
Vermittlung	operator	user requests transfer to operator
wiederholen	repeat	repeat the last directory information given
wählen	call/dial	place the call using the given number
deutsch	German	keyword for switching to German
Wahlwiederholung	redial	redial last Number or Name
Abbruch	cancel	cancel editing or dialling operation
Hilfe	help	request help information for current dialogue mode
programmieren	program	enter program mode to list or edit user-defined settings or preferences
Verzeichnis	directory	enter directory mode (or omit – simply having one

		menu level)
Menü	menu	return to main menu (or omit – simply having one menu level)
Liste	list	list directory entries
Stop	stop	stop directory listing
nächste	next	list or change next directory entry
vorherige	previous	list or change previous directory entry
neu	add	add name to directory
löschen	delete	delete name from directory
ändern	change	change name or number details in directory
Ende	terminate	quit the application
weiter	continue	continue playout
erste	rewind	rewind to start of list
Ausgabe	play	play current message
Versand	send	send current message
Aufnahme	record	record a greeting message
Sprache	language	select language of interaction

**Table 8 - Application words**

## **3.2 Isolated digits**

### **3.2.1 Single digit**

Random selection from 0 to 9. Note that there is no compensation for „2“ spoken as „zwei“ or „zwo“.

### **3.2.2 Digit string**

Random selection of 10 digit strings, each digit being different. No compensation for „2“ being spoken as „zwei“ or „zwo“. In most recordings, star and hash were used as begin and end symbols to encourage the pronunciation as single digits.

## **3.3 Connected digits**

### **3.3.1 Sheet number**

NNN MMMM-X digit number. NNN is a series number, MMMM the sheet number proper, and X a check digit for MMMM (in the 100 and 200 series, the check digit is incorrect due to a programming error in the computation).

### **3.3.2 Telephone number**

Random selection of digit strings with some adaptation to phone number format. Area codes always begin with a „0“; there are three, four, and five digit area codes in Germany; the short area codes have long telephone numbers (6 to 7 digits), the long area codes have short telephone numbers (4 to 5 digits). The telephone numbers may not begin with a „0“.

Some service identifying numbers, e.g. 0130, 0180, 0800, etc. have been included.

Mobile phone numbers begin with the area codes 017X, with X the identifying number of the network provider. Mobile phone area codes are included, but only with 7 digit telephone numbers.

### 3.3.3 Credit card number

The credit card number is given in the NNNN NNNN NNNN NNNN format. No checksum was used. The credit card numbers are taken from the common SpeechDat list of 150 numbers.

### 3.3.4 PIN code

The 6-digit PIN code was given as a single digit string NNNNNN. No checksum is used. The PIN codes are taken from the common SpeechDat list of 150 PIN codes.

## 3.4 Dates

### 3.4.1 Spontaneous date

One spontaneous date expression was elicited by the following prompt:

„Bitte nennen Sie ein beliebiges Datum“

(„please say some date“).

### 3.4.2 Prompted date

The format of date expressions to read is as follows:

```
{day_name ", " } {determiner} num_day "." [month_name |
num_month] "." [{"19"}num_year | "2000" | "2001" |
"2002" | "2003" | "2004" | "2005"]
```

with

day_name	Montag, Dienstag, Mittwoch, Donnerstag, Freitag, Samstag, Sonnabend, Sonntag
determiner	der, den
month_name	Januar, Februar, März, April, Mai, Juni, Juli, August, September, Oktober, November, Dezember

**Table 9 - Date expression vocabulary**

### 3.4.3 Relative and general date expression

The list of relative and general date expressions is as follows

```
{relative_expression ", " } {determiner} num_day "."
num_month "." [{"19"}num_year | "2000" | "2001" | "2002"
| "2003" | "2004" | "2005"]
```

with

relative_expression	heute, gestern, morgen, übermorgen, vorgestern, letzte Woche, letzter Monat, letztes Jahr, nächste Woche, nächster Monat, nächstes Jahr
---------------------	---

## Table 10 - Date expression vocabulary

### 3.5 Embedded application word

The list of application sentences is as follows:

Vermittlung	die Vermittlung bitte
	zur Vermittlung bitte
	bitte rasch die Vermittlung
	sofort Vermittlung zuschalten
	ich hätte gerne die Vermittlung bitte
Wiederholung	die Wiederholung bitte
	bitte keine Wiederholung mehr
	Wiederholung!
	die Nachricht wiederholen
	bitte wiederholen Sie
Wahl	die Wahl bitte
	geben Sie die Wahl ein
	jetzt die Wahl eingeben
	bitte wählen Sie
	wählen!
deutsch	alles in deutsch bitte
	Sprache deutsch
	ab sofort weiter in deutsch
	deutsch bitte
	ich spreche deutsch.
Wahlwiederholung	die Wahlwiederholung bitte
	Wahlwiederholung!
	Hilfe zur Wahlwiederholung
	Ende der Wahlwiederholung
	nochmal Wahlwiederholung
Abbruch	sofort Abbruch bitte
	Abbruch, schnell
	Abbruch!
	abbrechen!
	jetzt abbrechen
Hilfe	die Hilfe bitte
	sofort Hilfe

	Hilfe!
	Ende der Hilfe jetzt
	Hilfe Faxe löschen
Programm	das Programm bitte
	den Modus "Programm" einschalten
	Programm!
	ich möchte zum Programm
	jetzt programmieren
Verzeichnis	das Verzeichnis bitte
	zum nächsten Verzeichnis
	zeige mir das Verzeichnis
	schliesse das Verzeichnis sofort
	das steht im Verzeichnis drin
Menü	das Menü bitte
	aus dem Menü auswählen
	zeige mir das Menü
	Menü!
	gehe zum nächsten Menü
Liste	die Liste bitte
	alle Einträge der Liste zeigen
	Einträge auflisten
	aus der Liste löschen
	in die Fax-Liste aufnehmen
Stop	sofort Stop bitte
	Stop!
	stoppen!
	danach dann Stop!
	stoppe die Ausgabe
nächste	der nächste Eintrag
	zum nächsten Eintrag
	die nächsten Einträge zeigen
	nächster Eintrag
	Eintrag in die Liste
vorherige	der vorherige Eintrag
	zum vorherigen Eintrag
	vorheriger Eintrag

	den vorherigen Eintrag ändern
	lösche Eintrag
neu	die neue Nachricht
	zur neuen Nachricht
	keine neue Nachricht
	Nachrichten ausgeben
	Nachricht Stop!
löschen	den Eintrag löschen bitte
	lösche alle Einträge
	löschen!
	die Nachricht von gestern löschen
	alle Faxe löschen bitte
ändern	den Eintrag ändern bitte
	ändern!
	ändere den Empfänger der Nachricht
	das Fax ändern
	die Sprachnachricht ändern und verschicken
Ende	sofort Ende bitte
	Ende!
	beenden!
	danach bitte zum Ende
	beenden des Programms
weiter	zur Nachricht weiter
	weiter!
	die weiteren Faxe anzeigen
	keine weitere Eingabe
	mach weiter mit der Liste
erste	die erste Nachricht
	zur ersten Nachricht
	erster Eintrag
	zur ersten
	erster Eintrag
Ausgabe	die Ausgabe der Nachricht
	alle Nachrichten ausgeben
	Ausgabe!
	Ausgabe aller Nachrichten

	die Ausgabe umleiten
Versand	der Versand der Nachricht
	bitte das Fax versenden
	Status des Versands anzeigen
	Fax zum Versand weiterleiten
	Hilfe zum Versand ausdrucken
Aufnahme	die Aufnahme der Nachricht
	das Gespräch aufnehmen
	Aufnahme!
	aufnehmen!
	Aufnahme abspielen
Sprache	das Menü Sprache bitte
	Sprache wechseln
	wechsle zur Sprache deutsch
	zeige die Liste der Sprachen
	das Menü "Sprache" ausgeben
Bestätigung	die Bestätigung der Nachricht
	bestätige!
	Bestätigung!
	Bestätigung der Eingabe
	die Eingabe bestätigen
an	Aufnahme an, bitte
	an!
	schalte die Ausgabe an
	anschalten!
	Eingabe: an
aus	Aufnahme aus, bitte
	aus!
	schalte die Ausgabe aus
	ausschalten!
	Eingabe: aus
Eingabe	die Eingabebitte
	Eingabe!
	zurück zur Eingabe
	weiter zur nächsten Eingabe
	Eingabe der Faxnummer

Nachricht	die Nachricht bitte
	Nachricht ausgeben
	dritte nicht dringende Nachricht vorspielen
	letzte Nachricht anzeigen
	Nachrichten von Florian Murr von gestern vorspielen
Beantworter	den Beantworter bitte
	schalte den Beantworter ein
	der Beantworter ist eingeschaltet
	Beantworter neu programmieren
	Beantworter Stop!
Raute	die Raute bitte
	Raute eingeben
	als nächstes die Taste mit der Raute drücken
	Taste Raute drücken
	drücken Sie die Raute
Stern	den Stern bitte
	Stern eingeben
	als nächstes die Taste mit dem Stern drücken
	Taste Stern drücken
	drücken Sie den Stern
Quadrat	das Quadrat bitte
	Quadrat eingeben
	als nächstes die Taste mit dem Quadrat drücken
	Taste Quadrat drücken
	drücken Sie das Quadrat

**Table 11 - Application words and phrases**

### **3.6 Spelled names/words**

#### **3.6.1 Spontaneous name**

The speaker is asked to spell his or her first name:

„Bitte buchstabieren Sie Ihren Vornamen“

For privacy reasons, the surname was not asked for.

#### **3.6.2 Prompted name/word**

The names to spell are randomly selected items from the list of words, or the geographic and company names.

### **3.6.3 Artificial name**

The list of artificial spellings consists of 10 letters (selected randomly; both lower and upper case letters are used including the German umlauts Ä, Ö, Ü, ä, ö, ü, ß, accented characters é, è, ê, à, ç, and some common symbols, e.g. -, \*, ' ). The spelling items are given with a blank between them.

### **3.7 Money amount**

The format is DM NNNN.MM or DM NNNN,MM with NNNN a four digit number, MM a number between „01“ and „99“ with leading zeroes for amounts smaller than 10 Pfennig, or „--“ for „00“.

### **3.8 Natural number**

The number is given as a 3 to 5 digit number.

### **3.9 Directory assistance names**

#### **3.9.1 Spontaneous forename**

Speakers are asked for their first name:

„Bitte nennen Sie Ihren Vornamen.“

#### **3.9.2 Spontaneous city name**

Speakers are asked from where they are calling:

„Aus welcher Stadt rufen Sie an?“

#### **3.9.3 City name (set of 500)**

The list of geographical names is taken from the list of German telephone directory names, European capitals and countries, and city and country names taken from newspaper texts.

#### **3.9.4 Company/agency name (set of 500)**

Company and agency names are taken from the Fortune 500 list of the largest national and international companies, common German political office names (e.g. Bürgermeister, Bundestag, Parlament, etc.), and from the economics pages of newspapers.

#### **3.9.5 Forename & surname (set of 150)**

The list of forenames and surnames was taken from a random selection of the telephone directory of Offenbach with a total of ~55.000 entries (available in electronic form from a CD-ROM based telephone directory). A few distinctly foreign names were included in the list: English, Scandinavian, Turkish, Serbo-Croatian, Greek, Italian, Spanish, etc.

### **3.10 Yes/No questions**

The two yes/no questions asked were

„Waren Sie schon einmal im Kino?“

„Benutzen Sie ein schnurloses Telefon?“

„Have you ever been to a cinema?“ (yes expected) and „Are you calling with a cordless phone?“ (no expected).

For the FDB 1000, the following results were achieved:

yes question	934
no question	690

**Table 12 - Yes/no question summary**

### 3.11 Phonetically rich sentences

The following tables contains the phoneme count for each word in the transcription:

Phoneme	Count	Phoneme	Count	Phoneme	Count
O~	2	h	5660	k	13264
dZ	16	N	5816	e:	13957
dS	22	p	6698	m	14503
Z	59	C	7230	f	15609
a~	94	o:	7383	E	16535
tS	612	O	7694	s	17632
pf	1530	a:	7972	l	19972
2:	2089	S	8612	a	20274
E:	2106	v	8983	l	20384
OY	2202	z	9149	d	23860
9	2223	ts	9641	r	36261
Y	2393	6	10872	t	37500
y:	2825	b	11111	@	40695
j	3152	aI	11534	n	52613
x	3653	U	11576		
u:	4325	g	11978		
aU	5164	i:	12874		

**Table 13 - Phoneme count**

The phonetically rich sentences were selected to contain each phoneme at least twice on a prompt sheet. However, not all speakers were able to read the prompt sheets without errors. This results in some phonemes being uttered only once per prompt sheet and speaker.

### 3.12 Times

For the time phrases, three main formats were used: numerical, e.g. „10:12“, alphanumerical, e.g. „23 Uhr 03“, and in a format using only words, e.g. „kurz nach fünf nachmittags“. Time expressions, e.g. to indicate relative times, were also used.

For the spontaneous time expression the speaker was asked to speak the current time of day:

„Wie spät ist es jetzt?“

(„What time is it now?“).

The list of words used in time expressions and their frequency is given in the following table:

Time word	Count	Time word	Count
der	2	vier	45
Früh	2	zehn	50
in	2	abends	51
Mitternacht	6	fünf	54
Mittag	6	über	58
nachts	12	halb	67
mittags	23	nachmittags	72
morgens	23	dreiviertel	80
vorgestern	26	so	85
acht	30	gegen	85
heute	30	fast	87
drei	31	um	89
zwölf	34	kurz	98
eins	35	ungefähr	98
elf	35	vor	123
vormittags	36	nach	132
übermorgen	36	morgen	165
neun	37	gestern	202
sieben	37	viertel	215
zwei	39	Uhr	301
sechs	42	genau	327

**Table 14 - Time expression word count**

### 3.13 Phonetically rich words

The phonetically rich words were selected from the original lexicon in such a way that words with rare phonemes had a higher priority than more normal words. The following table contains the frequencies for each phoneme of the phonetically rich word transcriptions:

Phoneme	Count	Phoneme	Count
O~	2	v	679
a~	5	p	698
dZ	7	o:	791
Z	13	a:	794
E:	37	6	805
tS	49	S	833
pf	59	i:	844
9	72	U	908
Y	145	m	924
OY	161	g	942
y:	178	b	1066
2:	181	I	1120
h	365	ts	1230
aU	383	f	1241
C	434	k	1278
j	513	s	1375
u:	527	E	1541
aI	547	a	1638
N	549	l	1652
z	568	t	2870
e:	569	r	2882
O	583	@	3217
d	654	n	3533
x	676		

**Table 15 - Phoneme count**

### 3.14 Any other additional material

Empty

### 3.15 Links to other databases

The links to other DBs are the use of the standard sets of 150 credit card and PIN numbers.

## 4. Speaker demographic information

### 4.1 Accent/Regions

The federal states of Germany were chosen as accent regions for two reasons: this information can easily be provided by the speaker, and it provides 16 region classes.

The question asked to the speaker was

„In welchem Bundesland haben Sie die Grundschule besucht?“

(„in which federal state did you enter school?“).

Note that the federal states do not coincide exactly with dialect regions, and that the size of the states vary considerably.

Item	Name	Count	Target	Percent	Target Percent
A	Austria	8	0	0.81	0.00
BB	Brandenburg	4	31	0.40	3.20
BE	Berlin	33	43	3.34	4.40
BW	Baden-Württemberg	91	120	9.21	12.20
BY	Bayern	405	142	40.99	14.40
CH	Switzerland	1	0	0.10	0.00
HB	Bremen	2	9	0.20	1.00
HE	Hessen	74	72	7.49	7.30
HH	Hamburg	12	19	1.21	2.00
MV	Mecklenburg-Vorpommern	1	23	0.10	2.40
NI	Niedersachsen	85	91	8.60	9.30
NW	Nordrhein-Westfalen	120	217	12.15	22.00
OTHER		21	0	2.13	0.00
RP	Rheinland-Pfalz	39	48	3.95	4.90
SH	Schleswig Holstein	15	31	1.52	3.20
SL	Saarland	3	13	0.30	1.40
SN	Sachsen	35	56	3.54	5.70
ST	Sachsen-Anhalt	12	34	1.21	3.50
TH	Thüringen	17	30	1.72	3.10
UNKNOWN		10	0	1.01	0.00

**Table 16 - Distribution of speakers over regions**

## 4.2 Speaker characteristics

Age groups	Number of male speakers	Number of female speakers	Total count	Percentage of total
under 16	13	24	37	3
16-30	32	31	63	6
31-45	163	190	353	35
46-60	164	164	328	34
over 60	129	77	206	21
unknown			1	
total	501	486	988	

**Table 17 - Distribution of speakers over age groups and sexes**

## 5. The lexicon

The German SpeechDat(II) lexicon was mainly provided by Siemens AG. It covers all texts on the prompt sheets. The transcription of the speaker's utterances were not completely covered by this lexicon. For these missing words, a two-step procedure was used to generate the lexicon entries: 1) a proprietary grapheme-to-phoneme converter was used to generate a first version of the SAM-PA pronunciation; 2) this reference pronunciation was then checked manually (~ 15% error rate).

The SAM-PA symbols used are those of the German SAM-PA inventory with two additions: the ~ was used to indicate nasalization of /O/ and /a/ which occur in words of French origin (e.g. „Restaurant“). Not all diphthongs allowed in the SAM-PA table were used.

The lexicon entries are case sensitive and follow the German rules for case, i.e. nouns and names are written with the first letter a capital letter, all other words are written in lower case. Note that the polite form of personal pronomina, e.g. „Du“, „Sie“ are written in lower case too; nominalized verbs, e.g. „das Laufen“ are treated as nouns and hence are written with a capital first letter.

A lexicon entry may contain hypens, apostrophes, or slashes. Idiomatic expressions are split into their word constituents.

Only one spelling of a word is allowed. This spelling is determined by the Duden or the source from which the original word was taken (e.g newspaper text for words not in the Duden).

The entries are ordered by their ISO8859-1 encoding.

## 6. Recording conditions

### 6.1 Environments

Speakers are asked for the environment from which they called the speech server. This information is contained on a data sheet which speakers are asked to return to the Phonetics department (rate of return: 972 data sheets for 1200 recordings, i.e. 81%). On this data sheet, speakers could mark the environment; the options were home, office, public phone booth.

The following table contains the environment, network and handset data obtained from the data sheets:

HOME	528	53.44
OFFICE	231	23.33
BOOTH	5	0.51
UNKNOWN	224	22.67

**Table 18 - Environment**

**Network**

PSTN	971	0.98
MOBILE	17	0.02

**Table 19 - Network**

CORDLESS	207	20.95
MOBILE	17	1.72
ROTARY	23	2.33
TOUCH-TONE	599	60.63
UNKNOWN	142	14.37

**Table 20 – Phone model**

**7. Deviations from SpeechDat specifications**

Empty

**8. Sample Prompt sheets**

**8.1 Sample instruction sheet and prompt sheet**

A sample instruction and prompt sheet is contained on the CD-ROM in the directory PROMPT.

**8.2 Sample instruction sheet and prompt sheet (English translation)**

An english translation of the instruction and a prompt sheet is contained on the CD-ROM in the directory PROMPT.