SPEECH DATABASE DEVELOPMENT: DESIGN AND ANALYSIS OF THE ACOUSTIC-PHONETIC CORPUS*

Lori F. Lamel, Robert H. Kassel, and Stephanie Seneff

Department of Electrical Engineering and Computer Science, and
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT

The need for a comprehensive, standardised speech database is threefold: first, to acquire acoustic-phonetic knowledge for phonetic recognition; second, to provide speech for training recognisers; and third, to provide a common test base for the evaluation of recognisers. There are many factors to consider in corpus design, making it impossible to provide a complete database for all potential users. It is possible, however, to provide an acceptable database that can be extended to meet future needs. After much discussion among several sites, a consensus was reached that the initial acoustic-phonetic corpus should consist of calibration sentences, a set of phonetically compact sentences, and a large number of randomly selected sentences to provide contextual variation. The database design has been a joint effort including MIT, SRI, and TI. This paper describes MIT's role in corpus development and analyses of the phonetic coverage of the complete database. We also include a description of the phonetic transcription and alignment procedure.

INTRODUCTION

The development of a common speech database is of primary importance for continuous speech recognition efforts. Such a database is needed in order to acquire acoustic-phonetic knowledge, develop acoustic-phonetic classification algorithms, and train and evaluate speech recognisers. The acoustic realisation of phonetic segments results from a multitude of factors, including the canonical characteristics of the phoneme, contextual dependencies, and syntactic and extralinguistic factors. A large database will make it possible to examine in detail many of these factors, with the hope of eventually understanding acoustic variability well enough to design robust speech recognisers. A complete database should include different styles of speech, such as isolated words, sentences and paragraphs read aloud, and conversational speech. The speech samples should be gathered from many speakers (at least several hundred) of varying ages, both male and female, with a good representation of the major regional dialects of American English.

*This research was supported by DARPA under contract N00039-85-C-0341, monitored through Naval Electronic Systems Command.

DESIGN CONSIDERATIONS

There are many factors to consider in designing a large corpus for speech analysis. Unfortunately, some of the goals are limited by practical considerations. Ideally we would like to include multiple samples of all phonemes in all contexts, a goal that is clearly impractical for a manageable database.

At the last DARPA review meeting it was decided that an initial acoustic-phonetic database would be designed to have good phonetic coverage of American English. It was agreed that the initial acoustic-phonetic corpus would include calibration sentences (spoken by every talker), a small set of phonetically compact sentences (each spoken by several talkers) and a large number of sentences (each to be spoken by a single talker). This combination was chosen to balance the conflicting desires for compact phonetic coverage, contextual diversity, and speaker variability. Another requirement of the corpus was that the sentences should be reasonably short and easy to say.

The database design is a joint effort between MIT, SRI, and TI. The speaker *calibration sentences*, provided by SRI, were designed to incorporate phonemes in contexts where significant dialectical differences are anticipated. They will be spoken by all talkers. The second set of sentences, the *phonetically compact* sentences, was hand-designed by MIT with emphasis on as complete a coverage of phonetic pairs as is practical. Each of these sentences will be spoken by several talkers, in order to provide a feeling for speaker variation. Since it is extremely time-consuming and difficult to create sentences that are both phonetically compact and complete, a third set of *randomly selected* sentences, chosen by TI, provides alternate contexts and multiple occurrences of the same phonetic sequence in different word sequences.

A breakdown of the actual sentence corpus is shown in Table 1. This arrangement was chosen to balance the conflicting desires for capturing inter-speaker variability and providing contextual diversity. Since the calibration

	No. Talkers	No. Sentences	Total
Calibration (SRI)	640	2	1280
Compact (MIT)	7	450	3150
Random (TI)	1	1890	1890
Total	—	—	6320

Table 1: Breakdown of Frequencies of Occurrence of Sentences in Corpus

sentences are spoken by all of the speakers, they should be useful for defining dialectical differences. For multiple instances of the exact same phonetic environments, but with a much richer acoustic-phonetic content than in the calibration sentences, the MIT set would be appropriate. The TI sentences, to be spoken by one talker per sentence, should provide data for phoneme sequences not covered by the MIT database.

DESIGN OF THE COMPACT ACOUSTIC-PHONETIC SENTENCES

A set of 450 sentences was hand-designed at MIT, using an iterative procedure, to be both compact and comprehensive. We made no attempt to phonetically balance the sentences. We used *ALexis* and the Merriam-Webster Pocket Dictionary (Pocket) to interactively create sentences and analyze the resulting corpus. We began with the "summer" corpus created for the MIT speech spectrogram reading course to include basic phonetic coverage and interesting phonetic environments. We initially augmented these sentences by looking at pairs of phonemes, trying to have at least one occurrence of each phoneme pair sequence. *ALexis* was used to search the Pocket dictionary for words having sequences that were not represented and for words beginning or ending with a specific phoneme. We then created sentences using the new words and added them to the corpus. Certain difficult sequences were emphasized, such as vowel-vowel and stop-stop sequences. Some phoneme pairs are impossible; others are extremely rare and occur only across word boundaries. For example, /w/ and /y/ never close a syllable, except as an off-glide to a vowel, so many /w/-phoneme pairs are impossible. After filling some of the gaps in coverage, we reanalyzed the sentences with regard to phoneme pair coverage, consonant sequence coverage, and the potential for applying phonological rules both within words and across word boundaries. In a final pass through the sentence set, we modified and enriched sentences where simple substitutions could introduce variety or generate an instance of a rare phoneme pair.

ANALYSIS OF PHONETIC COVERAGE

This section discusses the phonetic coverage of the compact sentence set developed at MIT and the resulting corpus consisting of the combined MIT and TI sentences. This analysis does not include the calibration sentences as we consider their use to be of a different nature.

	POCKET	HL	MIT-450	APDB
# sentences		720	450	5040
# unique words	19,837	1894	1792	5107
# words	19,837	5745	3403	41,161
ave # words/sent		7.9	7.6	8.2
min # words/sent		5	4	2
max # words/sent		12	13	19
ave # syls/word	1.38*	1.1	1.58	1.54
ave # phones/word	3.34*	2.97	4.0	3.89

* The ave # syls/word and ave # phones/word have been weighted by Brown Corpus[1] word frequencies.

Table 2: Description of Databases

Table 2 compares some of the distributional properties of the Pocket Lexicon (Pocket), the Harvard List (HL)[2], the MIT-selected sentences (MIT-450), and the Acoustic-Phonetic Database selected sentences (APDB). The APDB includes seven copies of each MIT-450 sentence, to account for the number of talkers per sentence, and a single copy of each randomly selected sentence (TI-1890). Since we were given only the orthographies for the TI-1890 sentences, we generated phonemic transcriptions by dictionary lookup, by rule-based expansion of the dictionary entries, and, as a last resort, by a text-to-speech synthesiser. We expect that there are pronunciation variations between the dictionary and the text-to-speech synthesizer, particularly with respect to vowel color. There may also be some pronunciation errors, but we think these will be statistically insignificant.

The proportion of unique words relative to the total number of words is substantially larger in the MIT-450 than the APDB, probably due to the selection procedure. We tried to use new words in sentences and to avoid duplication when at all possible. Roughly 50% of the MIT-450 words are unique, as compared to only 25% of the APDB words. The TI-1890 sentences are, on the average, slightly longer than those in the MIT-450. The 10 most frequently occurring words for all of the corpora are function words or pronouns. In both the MIT-450 and the APDB corpora, the most common word is "the," accounting for roughly 7%of all words.

The average numbers of syllables and phones per word are longer for the MIT-450 and the APDB than for the HL. This is presumably due to the higher percentage of polysyllabic words.

Figure 1 shows the distribution of the number of syllables per word for the two corpora. The distributions are quite similar, with the majority of the words being mono- or bi-syllabic. The MIT-450 corpus has a slightly higher percentage of polysyllabic words than does the combined

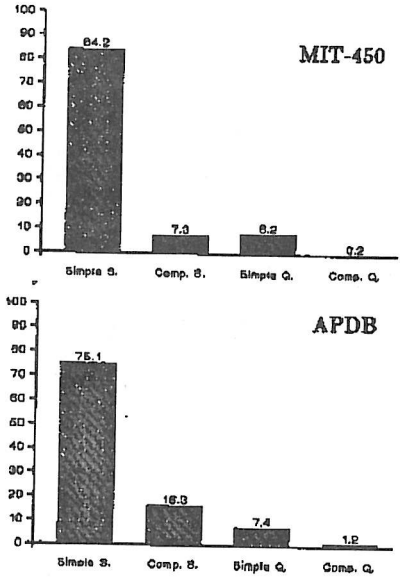

Figure 4: Histogram of sentence types.

10% of both corpora.

Figure 5 shows counts of environments where major phonological rules may apply. We chose to gather information on the following possibilities:

- gemination (GEM)
- vowel-vowel sequences (VVS)
- vowel-schwa sequences (VSS)
- schwa-vowel sequences (SVS)
- flapping of /t/,/d/, and /n/ (FLAP)
- homorganic stop insertion (HSI)
- schwa devoicing (S-DVC)
- fricative devoicing (F-DVC)
- /s/-/š/ and /z/-/ž/ palatalisation (PAL)
- y-palatalization: /dy/→/ǰ/ (DY-Jh)
- y-palatalisation: /ty/→/č/ (TY-Ch)
- y-palatalization: /sy/→/š/ (SY-Sh)

The histograms show that both corpora have many potential environments for flapping and homorganic stop insertion. The vowel-vowel environments are also well covered. The analysis for phonological rule application is difficult, because of the difficulties in predicting what different speakers will say.

RECORDING, LABELING, AND ALIGNMENT

The recording of the sentences is currently under way at TI. Speech is recorded digitally at 20 kHz, simultaneously on a pressure-sensitive microphone and on a Sennheiser close-talking microphone. Digital tapes are shipped to NBS, where they are filtered and downsampled to 16 kHz. The resampled tapes are then shipped to MIT where the orthographic and phonetic transcriptions are generated.

Transcriptions are generated using the *Spire* facility, in conjunction with the automatic alignment system pro-

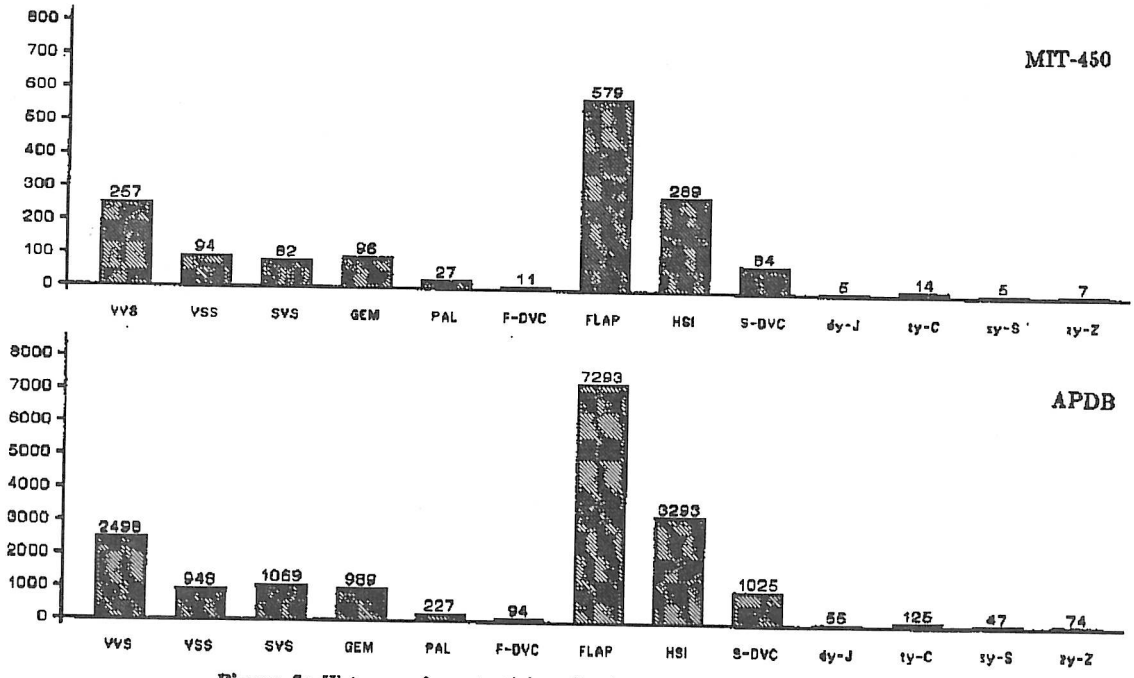

Figure 5: Histogram for potential application of phonological rules.

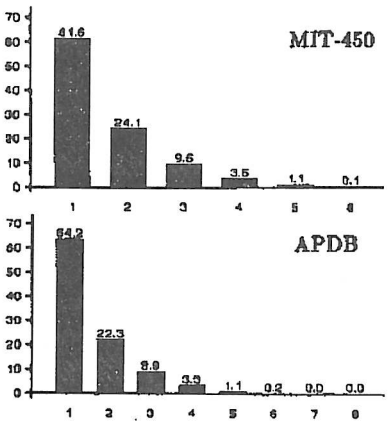

Figure 1: Histograms of the number of syllables per word.

corpus. We specifically tried to include polysyllabic words in the sentences, since these are likely to be spoken with greater variability.

Distributions of the number of phonemes per word are shown in Figure 2. The 10 most common phonemes and their frequency of occurrence are given in Figure 3.

Table 3 shows the distribution of within-word consonant sequences for the four databases. The MIT-450 sentence set covers most of the consonant sequences occurring within words. The APDB has more complete coverage, particularly for the word-final and word-medial sequences. We examined a list of all of the word-initial and word-final clusters in the sentence list, and compared these with the occurrences in Pocket. We verified that essentially every initial cluster that occurred more than once in the Pocket lexicon was included at least once in the APDB, and that most of the final clusters were covered. Often, if a word-final cluster did not occur in word-final position in the APDB, the sequence did occur within a word or across a word boundary. Generally, the sequences occurring in Pocket that are not covered by APDB are from borrowed words such "moire" and "svelte."

The APDB includes many word-final consonant sequences that were not present in MIT-450. In fact, there are more word-final consonant sequences in the APDB than actually occur in Pocket. The reason is that the Pocket lexicon does not include suffixes.

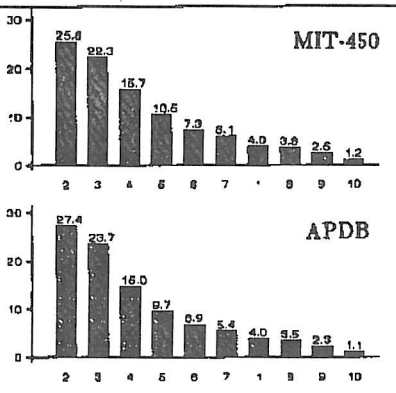

Figure 2: Histograms of the number of phonemes per word.

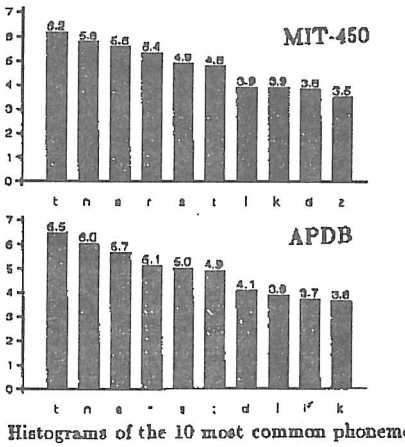

Figure 3: Histograms of the 10 most common phonemes.

	POCKET	HL	MIT-450	APDB
# unique words	19,837	1894	1792	6103
# WI	75	59	64	68
# WF	129	105	102	146
# WM	608	123	228	388
# boundaries		4305	2953	36,121
# WB		976	805	1639

Table 3: Distribution of Consonant Sequences

A more detailed phonetic analysis of all *phoneme pairs* is included in Appendix 1 in tabular form. The tables are broken down into phoneme subsets, and data are included for both the MIT-450 and the APDB. Some of the gaps in the MIT-450 table have been filled in by sentences in the TI-1890 corpus (e.g., the syllabic /l/ column of the vowel-sonorant pairs table and the /y/ column of the vowel-sonorant pairs table). Note also that some gaps occur in both tables. Such gaps are expected, since some phoneme sequences are impossible or quite rare. For example, the lax vowels (excluding schwa) are never found in syllable-final position in English. As a consequence, table entries requiring lax vowels as the first member of a pair have many gaps (see for example, the vowel-vowel entries in the pair tables.)

Figure 4 compares histograms of the sentence types for the MIT-450 and the APDB. Simple sentences (Simple S.) and questions (Simple Q.) have no major syntactic markers. Complex sentences (Complex S.) and questions (Complex Q.) are expected to have a major syntactic boundary when read. As can be seen, the APDB has a wider variety of sentence types, with 75% being simple declarative sentences. In the MIT-450, almost 85% of the sentences are of the simple declarative form. Questions form about

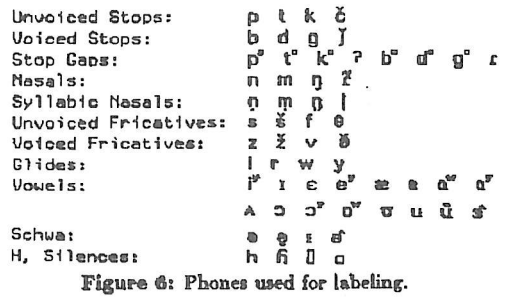

Figure 6: Phones used for labeling.

vided by Leung [3]. The transcription process involves three steps:

1. A "Phonetic Sequence," which consists of a list of the phones of the utterance in correct temporal order but with no boundaries marked in time, is entered.

2. The utterance is run through an automatic system to generate an alignment for the sequence.

3. The automatically generated alignment is hand-corrected.

Only the data recorded through the pressure microphone are transcribed. Transcriptions for the close-talking version are generated by duplicating the results for the pressure microphone.

The phones used in the labeling are shown in Figure 6. In many cases, it is not possible to define a boundary between two phones, such as /ɔr/, because features appropriate for both phones often occur simultaneously in time. When no obvious positioning of the boundary is apparent, arbitrary rules, such as an automatic 2/3:1/3 split, are invoked. There are also some cases in which none of our standard phones are appropriate for a given portion of the speech, primarily because of severe coarticulation effects. In such cases, the segment is labeled as the nearest phone equivalent, according to the transcriber's judgment. There are other difficult cases, such as syllable-initial /pl/, where the /l/ is devoiced at onset. Should the portion before voicing begins be thought of as part of the aspiration of the /p/, or as part of the /l/? We have decided, somewhat arbitrarily, to define the onset time of the phone following an unvoiced stop as coincident with the onset of voicing. These remarks serve simply as examples of some of the difficulties that arise in transcribing continuous speech. We are mainly interested in using consistent methods of transcribing in situations where ambiguity exists. Currently the transcription rate is 100 sentences per week.

SUMMARY

We have described various components of the preliminary acoustic-phonetic database and discussed some of the issues in its design. Evaluating the phonetic coverage of the database is difficult primarily because no standard for comparison exists. We have chosen to compare the phonetic coverage of the database to two well-known sources, the Merriam-Webster Pocket Dictionary of 1964 and the Harvard List sentences. The dictionary does not reflect spoken English very well, and can only guide us in judging the possible phonemic sequences within words. The Harvard List sentences, while phonemically balanced, consist primarily of very simplistic sentences and monosyllabic words. In addition, they are balanced for phoneme occurrences, whereas we tried to account for occurrences of phoneme pairs.

We believe that we have adequate coverage of most phonemes and phoneme pairs. In cases where the phoneme pairs are scarce, there are often other phoneme pairs that will provide similar information. For example, the class sequence **[alveolar consonant] [back vowel]** is more general than /t/ /ɔ/, and has a higher frequency of occurrence.

We hope that the APDB database will provide guidelines for the development of future databases. An analysis of the spoken corpus will enable us to judge our phonetic analysis procedure. In particular, we will be able to evaluate the relationship between our phonological rule predictions and the frequency with which a phonological modification actually occurred.

REFERENCES

[1] Kucera, H. and W.N. Francis (1967) *Computational Analysis of Present-Day American English*, Brown University Press, Providence, R.I.

[2] Egan,J. (1944) "Articulation testing methods II," OSRD Report No. 3802, U.S. Dept. of Commerce Report PB 22848, November.

[3] Leung, H. C. and V.W. Zue (1984) "A Procedure for Automatic Alignment of Phonetic Transcriptions with Continuous Speech," *Proc. ICASSP-84*, 2.7.1-2.7.4.

Appendix 1

MIT

	iʸ	eʸ	ɔ	oʷ	u	aʸ	ɔʸ	aʷ	ɝ
iʸ	3	7	8	9		3	1	3	1
eʸ	1	1	4	4		1		1	1
ɔ	1	1	1	1				1	
oʷ	2	2	3	1		1	1	3	
u	1	5	5	3	1	2	1	2	1
aʸ	4	4	6	2	1	3	1		1
ɔʸ	1	1	1	1					
aʷ	1	2	1	1		1	1		
ɝ	7	1	1	2					1
ɪ									
ɛ									
æ									
ɑ	1								
ʌ									
ʊ									
ɚ	4	7	4	1		6	1	3	1
ɔ	3	1		7		4	1	1	1
ɨ	1								

APDB

	iʸ	eʸ	ɔ	oʷ	u	aʸ	ɔʸ	aʷ	ɝ
iʸ	28	70	75	88		27	7	26	22
eʸ	10	7	38	31		10		7	9
ɔ	7	7	7	7				7	
oʷ	14	14	26	10		9	7	21	4
u	12	48	43	26	7	19	7	16	10
aʸ	28	30	47	16	8	21	7	2	8
ɔʸ	10	7	7	7					
aʷ	7	14	9	7		7	7		
ɝ	62	10	12	18		3		2	7
ɪ									
ɛ				2					3
æ									
ɑ	8								
ʌ									
ʊ									
ɚ	64	73	37	13		52	10	25	7
ɔ	22	8	7	60		33	7	11	10
ɨ	7								

	iʸ	eʸ	ɔ	oʷ	u	aʸ	ɔʸ	aʷ	ɝ
n	38	5	8	11	10	14	4	3	
m	17	24	10	18	9	24	2	1	4
ŋ	1	1	3					1	1
m̩			1						
n̩		1							
l̩	2		2			1		1	
l	88	23	13	25	9	22	4	12	7
r	81	28	16	28	24	24	3	5	
w	20	20	20	6	1	17	1	1	10
y			1	1	88				1

	iʸ	eʸ	ɔ	oʷ	u	aʸ	ɔʸ	aʷ	ɝ
n	422	79	88	252	117	133	32	95	21
m	246	318	100	227	86	248	18	22	48
ŋ	9	7	32	8		1	1	10	7
m̩			7						
n̩		7							
l̩	55	5	17	5		13		7	2
l	1001	287	168	286	115	308	34	110	60
r	936	312	169	303	260	272	29	71	6
w	294	212	199	67	8	210	7	9	134
y	3		9	7	933			5	10

	iʸ	eʸ	ɔ	oʷ	u	aʸ	ɔʸ	aʷ	ɝ
b	27	2	9	10	1	28	4	3	8
d	18	22	8	10	14	7		9	2
ɡ	1	3	1	13	4	1		1	1
p	14	14	2	11	2	9	8	3	12
t	44	19	11	13	86	18	1		4
k	9	9	19	15	7	3	4	8	11
ʧ	6	9	1	2	4	3	1		3
ʤ	7	2	3	1	6	2	5		4
s	22	5	12	5	9	13	2	2	10
z	14	7	10	3	2	2		1	4
ʃ	11	2	5	5	3	1		1	1
ʒ					1				
f	9	5	17	46	2	8		3	12
v	10	6	2	5		3	2	1	4
θ	6		2					3	5
ð	3	9		5					
h	8		7	4	5	10		12	18

	iʸ	eʸ	ɔ	oʷ	u	aʸ	ɔʸ	aʷ	ɝ
b	364	35	78	118	15	282	48	54	79
d	270	212	95	133	165	80	1	107	44
ɡ	7	51	18	137	29	15	1	7	28
p	149	141	32	142	21	81	75	38	137
t	501	234	157	175	1060	220	9	21	99
k	99	137	191	146	56	31	32	75	106
ʧ	54	73	10	18	56	28	9	1	26
ʤ	61	16	25	12	55	17	40		36
s	275	83	137	176	85	181	20	26	123
z	130	63	109	58	18	33		12	37
ʃ	198	33	42	54	35	9		11	15
ʒ					22				
f	120	72	172	506	25	117	4	32	120
v	92	68	28	44		60	28	15	63
θ	56	2	31	3	3	3		37	53
ð	65	192		59				2	3
h	381	18	58	79	54	117		137	213

MIT

	ɪ	ɛ	æ	ɑ	ʌ	ʊ	ɚ	ə	i
ɪˠ	11	4	16	11			5	28	1
ɛˠ	4	3	3	1	1		1	3	
ɔ	1	1	1	2					1
oʷ	4	1	2	3			1		3
u	6	7	4	6	1		1	4	
ɑʸ	5	2	7	3	2		4	11	6
ɔʸ	3	1	1				2	1	2
ɑʷ		1	1		1		3	3	
ɝ	2		1		1			1	1
ɪ									
ɛ									
æ									
ɑ									
ʌ									
ʊ									
ɚ	6	4	7	6	2		1	7	4
ə	5	6	8	2	2			3	
i									

	ɪ	ɛ	æ	ɑ	ʌ	ʊ	ɚ	ə	i
n	38	16	13	19	10		16	29	14
m	17	15	13	7	23		9	44	8
ŋ	12		3	1				2	2
m̩									
n̩	2	3	1	1				1	1
l̩	10	2	2	1				7	3
l	37	22	29	20	11	5	12	33	24
r	69	39	33	26	38	1	2	49	16
w	60	25	3	7	7	12	6	30	
y	4	4		3	6	38	2	1	1

	ɪ	ɛ	æ	ɑ	ʌ	ʊ	ɚ	ə	i
b	25	12	16	15	11	3	5	12	5
d	56	9	17	6	7	1	16	56	22
g	9	10	9	13	7	4	9	3	5
p	12	13	14	20	2	5	23	16	3
t	63	21	30	16	8	2	44	47	45
k	10	11	42	25	23	6	6	33	14
ʧ	6	4	2	7	1		10	3	6
ʃ	12	12	4	4	6		5	2	25
s	44	15	11	8	19	1	7	39	20
z	37	7	17	13	4		4	36	9
ʒ	5	5	3	2		9	1	3	42
ǯ						1	4		3
f	23	5	6	5	5	2	2	15	2
v	12	12	10	3			33	16	3
θ	14	1	1	1			4	3	
ð	15	14	21		1		12	219	
h	24	7	31	15	3	3			

APDB

	ɪ	ɛ	æ	ɑ	ʌ	ʊ	ɚ	ə	i
ɪˠ	174	47	182	102	7		46	399	27
ɛˠ	38	22	26	21	9		8	38	7
ɔ	9	7	9	14	2			3	8
oʷ	54	14	26	25	3		28	28	34
u	75	58	48	54	11		8	79	13
ɑʸ	51	18	61	23	14		30	107	61
ɔʸ	21	8	8				14	7	15
ɑʷ	5	17	9	1	8		32	30	
ɝ	27	8	13	1	9			30	16
ɪ									
ɛ								1	
æ									
ɑ									
ʌ									
ʊ									
ɚ	89	38	73	48	21		8	106	52
ə	57	52	82	19	23			42	2
i									

	ɪ	ɛ	æ	ɑ	ʌ	ʊ	ɚ	ə	i
n	434	185	185	260	148	3	144	388	214
m	231	223	219	93	240		91	516	87
ŋ	125	6	51	10	11			43	14
m̩	1		3					1	
n̩	18	24	11	9			2	11	10
l̩	93	29	37	10	2			96	43
l	397	281	330	191	127	67	119	342	279
r	808	415	364	271	393	10	17	562	211
w	659	293	31	125	150	145	144	376	5
y	44	55	3	29	46	353	44	15	9

	ɪ	ɛ	æ	ɑ	ʌ	ʊ	ɚ	ə	i
b	271	123	185	145	204	31	69	185	41
d	646	130	209	82	121	10	174	595	245
g	99	132	97	118	67	57	87	30	54
p	160	150	166	208	41	66	222	223	44
t	764	288	326	160	105	35	456	690	481
k	127	97	415	266	251	81	51	404	178
ʧ	61	41	29	56	14	4	101	47	70
ʃ	110	120	35	39	68	1	50	29	230
s	523	349	166	88	268	8	86	488	270
z	386	69	228	157	52		36	437	112
ʒ	56	48	40	32	14	112	13	84	472
ǯ			1			7	39	3	34
f	268	87	96	74	46	26	39	172	31
v	152	145	124	40	11		345	230	52
θ	171	16	18	7	3		34	50	19
ð	238	196	323		14		187	2230	
h	455	118	452	137	46	27	9	2	4

MIT

	b	d	g	p	t	k	č	ǰ
iʸ	7	27	9	27	32	25	17	1
eʸ	8	13	2	9	47	17	2	7
ɔ	2	5	5	1	13	4	3	1
oʷ	5	2	2	9	15	12	3	
u	10	14	5	15	17	10	4	7
aʸ	7	17	2	2	24	14	1	1
ɔʸ	2	22	1	2				
aʷ		2	1	1	22	1		
ɝ	8	5	1	5	10	10	6	5
ɪ	5	13	28	16	21	87	8	12
ɛ	2	16	8	5	20	37	1	8
æ	10	10	10	9	37	27	4	2
ɑ	13	7	3	19	31	19	2	10
ʌ	10	3	5	15	12	4	8	3
ʊ	1	24	1	1	4	13	1	
ɚ	6	14	2	3	10	8	4	
ə	48	46	28	45	43	46	2	14
ɨ	8	13	2	3	27	9		8

	b	d	g	p	t	k	č	ǰ
n	13	119	7	11	120	17	5	16
m	15	3	3	33	6	6	1	1
ŋ	3	4	9	3	4	10	1	1
ɱ								
ɳ	1	3			11		1	
l̩			5	4	5	8	4	1
l	7	29	1	6	7	8	1	1
r	16	36	9	5	22	15	5	15
w								
y								

	b	d	g	p	t	k	č	ǰ
b	3	5	1	2	4	2	1	3
d	25	5	3	5	14	8	2	3
g	4	3	3	2	3	1	1	1
p	4	1	2	4	11	2	2	1
t	18	14	6	11	18	13	1	3
k	2	6	3	7	45	6	1	1
č	2	2	3	1	4	3	2	1
ǰ	1	8	1	3	4	1	2	2
s	9	7	4	46	158	56	7	1
z	15	21	9	11	16	17	2	4
š	1	1	1	1	4	1	1	
ž		1						
f	1	1	1	2	12	3	1	
v	1	3	1	7	3	5	1	1
θ		5		2	1	2		1
ð	2	1		1		1	1	
h								

APDB

	b	d	g	p	t	k	č	ǰ
iʸ	116	357	102	292	387	270	155	13
eʸ	113	188	20	84	503	228	18	62
ɔ	16	41	48	8	130	52	23	7
oʷ	68	61	30	104	157	124	27	
u	121	173	58	152	196	120	28	57
aʸ	60	249	19	31	292	166	7	8
ɔʸ	15	22	7	14	1	3		
aʷ	3	42	9	8	262	8		3
ɝ	69	75	11	53	121	112	52	46
ɪ	48	186	266	186	501	873	100	145
ɛ	16	224	76	83	269	438	8	71
æ	95	192	99	118	519	357	43	26
ɑ	133	91	33	195	408	183	16	87
ʌ	101	54	60	171	202	64	104	23
ʊ	8	314	10	8	48	149	7	
ɚ	63	203	23	43	101	83	34	5
ə	537	616	307	477	517	514	31	116
ɨ	79	168	28	37	387	137		75

	b	d	g	p	t	k	č	ǰ
n	144	1482	72	118	1343	204	76	155
m	181	63	26	369	80	55	9	9
ŋ	39	37	125	41	77	146	7	10
ɱ		1						
ɳ	8	28		1	141	2	7	
l̩	24	75	40	55	84	47	10	1
l	71	348	15	74	129	79	14	15
r	156	359	93	86	277	173	41	121
w								
y								

	b	d	g	p	t	k	č	ǰ
b	24	41	8	14	30	14	7	27
d	283	82	51	95	196	105	16	30
g	28	30	21	17	21	8	7	7
p	33	11	16	32	138	18	15	7
t	210	161	75	119	214	154	15	24
k	26	48	31	58	513	59	26	7
č	16	17	22	9	46	26	14	8
ǰ	12	67	8	25	31	9	14	14
s	98	77	39	503	1778	523	69	12
z	174	221	85	124	189	175	19	40
š	8	8	7	8	44	7	7	
ž		7						
f	13	17	8	15	155	30	7	
v	36	77	21	67	50	61	8	9
θ	7	38	3	16	14	24		7
ð	14	8		7		7	7	
h								

MIT

	n	m	ŋ	ɱ	ɳ	ɻ	l	ɾ	w	y
iʸ	23	15				1	15	11	15	2
eʸ	29	11					8	1	2	1
ɔ	29	1	10				39	61		
oʷ	24	6					20	77	3	
u	15	23				1	9	5	7	5
ɑʸ	16	15				1	15	12	2	2
ɔʸ	9	1					4			
ɑʷ	20	2					2	10	2	
ɝ	8	7					4	1	5	
ɪ	112	32	46				49	24	2	
ɛ	56	17	3				48	26		
æ	125	24	8				23	31		
ɑ	17	15	1				16	100		
ʌ	46	34	11				11			
ʊ		2					16	30		
ɚ	7	5				5	6	2	13	2
ə	104	62					50	10	10	4
ɨ	94	1	60							

APDB

	n	m	ŋ	ɱ	ɳ	ɻ	l	ɾ	w	y
iʸ	276	224				41	213	143	219	31
eʸ	293	157				6	109	32	43	10
ɔ	335	10	127				480	574		2
oʷ	293	92				1	258	987	38	3
u	161	247				60	115	72	78	57
ɑʸ	220	171				14	161	151	22	17
ɔʸ	83	7				5	55		1	2
ɑʷ	241	22				2	21	122	20	2
ɝ	104	79				6	63	19	43	6
ɪ	1372	423	500				509	330	16	
ɛ	795	213	27				540	340		
æ	1544	233	90			1	231	350		
ɑ	201	160	8				183	1022		
ʌ	522	437	108				121			
ʊ		28					163	299		
ɚ	74	58				63	62	35	110	20
ə	1324	680	6				537	162	152	37
ɨ	1013	37	758				5			

	n	m	ŋ	ɱ	ɳ	ɻ	l	ɾ	w	y
n	5	15				8	9	3	8	8
m	3	10				2	5	2	3	10
ŋ	2	1					2	5	5	4
ɱ							1			
ɳ						2			1	
ɻ	2	1					4	3	2	
l	2	8					3	3	12	8
ɾ	16	23			1	2	18	5	13	6
w						1				
y					1					

	n	m	ŋ	ɱ	ɳ	ɻ	l	ɾ	w	y
n	59	169				88	129	50	113	108
m	33	88				33	49	24	53	90
ŋ	23	28					33	48	50	31
ɱ	1						7		2	
ɳ	1	1				20	7	1	9	
ɻ	24	20					85	31	24	2
l	27	85		2			29	41	115	81
ɾ	160	262			9	34	167	55	142	58
w						8		1		
y					7	1				

	n	m	ŋ	ɱ	ɳ	ɻ	l	ɾ	w	y
b	1	1				21	21	27	2	9
d	8	5			11	4	7	32	10	5
ɡ	3	2				4	8	36	6	5
p	4	3				10	41	57	4	10
t	6	10		5	5	7	18	63	16	14
k	2	3			1	25	36	30	26	18
ʧ	1	1			1	1	2	1	1	2
ʤ	1	2					1	1	1	1
s	5	18			6	2	16	4	18	5
z	9	14		2	7		4	9	14	7
ʃ	1	3			7	5	2	2		1
ʒ		1								
f	1	1				6	12	35	1	6
v	8	5			1	6	3	12	3	5
θ	1	1					1	12	1	
ð				1						1
h									18	5

	n	m	ŋ	ɱ	ɳ	ɻ	l	ɾ	w	y
b	8	9				236	228	274	17	82
d	114	106			116	43	123	330	148	55
ɡ	46	19				43	98	357	47	56
p	30	33				102	425	633	35	85
t	99	152		39	60	102	222	728	250	125
k	36	37			7	251	376	324	287	185
ʧ	9	12			7	15	20	10	13	14
ʤ	8	16				4	10	7	9	11
s	66	184			70	31	197	54	213	47
z	154	163		29	100	8	71	98	175	74
ʃ	10	24			49	55	17	22	1	7
ʒ		7							1	
f	10	11				98	149	330	16	71
v	68	71			8	74	44	129	37	54
θ	9	14		1		1	16	128	15	1
ð		1		7						7
h						1			315	44

MIT

	s	z	ʃ	ʒ	f	v	θ	ð	h
iʸ	28	43	3	1	11	18	2	6	6
eʸ	14	12	20	2	2	5	1		1
ɔ	9	6	2		12		5		2
oʷ	14	18	5	2	2	16	4	1	2
u	18	25	3		8	9	5	10	8
aʸ	15	17	1		2	8		4	2
ɔʸ	7	5						1	
aʷ	6	1							
ɝ	10	7	1		5	6	4	2	1
ɪ	50	64	32	2	21	22	7	24	
ɛ	27	3	7	2	4	25		3	
æ	17	15	6	1	17	15	3	1	
ɑ	11	1	2	5	2	2			
ʌ	16	2	3		4	7	2	7	
ʊ									
ɚ	7	36	2		9	4	1	7	8
ə	67	54	3		28	55	6		14
ɨ	43	17			2	2	1		

APDB

	s	z	ʃ	ʒ	f	v	θ	ð	h
iʸ	374	435	52	9	170	194	56	102	158
eʸ	194	121	248	15	36	72	13	10	22
ɔ	105	55	19		147	1	45	5	14
oʷ	174	184	59	17	24	165	51	29	33
u	193	216	38	19	88	91	53	131	109
aʸ	154	211	11		53	94	3	42	23
ɔʸ	67	45			2			7	
aʷ	55	22	2		2	5	7	7	7
ɝ	129	97	17		51	68	44	23	13
ɪ	647	901	290	21	252	251	152	168	5
ɛ	332	43	67	23	47	286	12	33	1
æ	206	263	67	9	166	180	28	14	1
ɑ	119	11	16	36	18	19	2	8	1
ʌ	238	52	30		58	83	25	111	
ʊ		2	5		1				
ɚ	100	346	23		86	40	10	83	95
ə	797	639	66		325	825	66	2	174
ɨ	452	162	12		34	46	7		1

	s	z	ʃ	ʒ	f	v	θ	ð	h
n	60	45	7		12	1	3	34	11
m	6	20	1		8	1	2	11	4
ŋ	10	5	2		5	2	2	4	8
ɱ		2							
ɲ	2	4			1	1		1	
l̩	6	17			6	1	1	1	
l	7	13	1		11	3	5	5	4
r	15	16	2		8	3	2	9	1
w									
y									

	s	z	ʃ	ʒ	f	v	θ	ð	h
n	718	451	83		165	43	37	367	180
m	93	213	9		81	9	20	97	58
ŋ	94	61	17		48	18	23	54	82
ɱ		18	1					1	
ɲ	25	36			9	7	1	10	2
l̩	64	169	4		67	12	10	11	24
l	123	152	16		140	52	39	57	58
r	191	157	25	1	101	30	27	110	60
w									5
y									

	s	z	ʃ	ʒ	f	v	θ	ð	h
b	2	4	1		1	1	1	1	1
d	15	17	1		15	3	4	15	11
ɡ	1	11	1	1	1		1	1	1
p	13	1			2	1	1	3	1
t	77	1	2		12	2	1	21	12
k	53	1	4		6	1	1	4	2
ʔ	1		1		1		1	2	1
s	7	1	7		10	1	2	7	7
z	19		9		19	2	2	5	16
ʃ	3	1	1		3	1	1		
ʒ	1								1
f	6				1	1	3	1	
v	4	9	1		6	1		11	5
θ	3	1			2	1	1		
ð	1				2	1		4	1
h									

	s	z	ʃ	ʒ	f	v	θ	ð	h
b	32	36	7		8	12	7	7	8
d	191	233	25		174	54	43	191	214
ɡ	10	131	8	7	10		8	8	10
p	152	7	11		23	7	7	35	14
t	910	7	41		158	24	18	272	233
k	604	7	81		66	11	8	49	47
ɾ	32	7	7		16	7	7	14	14
ʔ	11		7		9		7	16	11
s	92	7	61		127	14	21	95	107
z	227	1	80		185	26	22	129	204
ʃ	22	7	7		22	7	9	1	3
ʒ	7							1	7
f	62		7		10	7	21	17	14
v	72	98	10		58	13	2	152	71
θ	35	7	2		20	9	8	16	15
ð	7	3			15	7		30	7
h									