Internal Use Only（非公開）

TR－SLT－0012

Robust Speech Recognition for Non-Native Speech Based on Phoneme Lattice Processing

Norbert Binder　　Rainer Gruhn

March 29th, 2002

In this report, the recent research for robust recognition of non-native speech is analyzed and a new approach is introduced. Target of this method is to eliminate typical variations in non-native speech on phoneme level.

During training, the phoneme substitutions and identities are extracted in a data driven way. Variations with a low occurrence frequency are rejected and all other accepted as rules. In the recognition process, a phoneme lattice is generated. By applying the previously generated rules to this lattice, new variations are added. The resulting modified lattice is then transferred to word level. Task are English conversations on hotel reservation spoken by Japanese, which were in part recorded for this thesis. An English acoustic model (AM) was trained on the Wall Street Journal speech database, and a Japanese model on the ATR TRA database. By merging, a mixed AM is generated, which allows the recognition of the phonemes of both languages for rule generation and the recognition-process itself.

（株）国際電気通信基礎技術研究所
音声言語コミュニケーション研究所
〒619－0288 京都府相楽郡精華町光台二丁目２番地２ TEL：0774－95－1301

Advanced Telecommunication Research Institute International
Spoken Language Translation Research Laboratories
2-2-2 Hikaridai Seika-cho Soraku-gun Kyoto 619-0288,Japan
Telephone:+81-774-95-1301
Fax　:+81-774-95-1308

Contents

1 Introduction

The importance of Communication with people all over the world increases, and speech recognition systems become more and more popular in everyday live. One area, where these two factors occur very often is the travel domain. This includes e. g. hotel reservation, tourist information systems or other voice controlled devices like navigation systems in rental cars. One of the main topics in ATR-SLT is the robust recognition of speech for large vocabulary in the above mentioned task. A sub-project emphasizes on multilingual and non-native speech recognition systems.

The biggest problems appearing in those tasks are accents of people not speaking in their native language, but the language the system is designed for. depending on the proficiency of a speaker, the influences of this non-native speech can effect the recognition performance very strong. Human listeners can adapt quickly to pronunciation variations while listening to a person who is not speaking in his or her mother tongue [Com01]. A recognition system is restricted to phoneme sequences given in the dictionary. In case of deviations, the recognition performance decreases rapidly. If it is possible to compensate those typical non-native disturbances a higher performance in recognition can be achieved.

In this approach, phoneme recognition is performed and typical variations of non-native speech, which were extracted before from the training-set, are applied to the resulting phoneme lattice. The recognition performance of either a native English or a mixed English-Japanese acoustic model for both, rule generation and recognition were compared.

2 Dynamic Lattice Processing

In this thesis, the robust recognition of non-native speech by elimination of typical variations on phoneme level by phoneme-lattice processing is proposed. Speech data from the WSJ- and TRA-databases, explained in Subsection 3, and the corresponding transcriptions and dictionaries are given. A native English and a mixed Japanese-English acoustic model (AM) are trained on this data. Additional speech for a database of English spoken by Japanese is collected for rule generation and evaluation.

The approach is based on automatically-derived rules which define the most frequent substitutions for a phoneme. Information about the pronunciation-variations is extracted by aligning the recognition result and the transcription on phoneme-level. Two different AMs are used for the experiments, a native English and a mixed English Japanese. As only the English transcription is used for generation, these rules also contain the mapping from Japanese to English phonemes. Unlike in [WY99], a single phoneme is not substituted by a single, but a group of the most likely phonemes. This group can also contain the original phoneme if this was from the English set, whereas in case the mixed acoustic model is applied to recognition, the resulting lattice will contain phonemes of the English phoneme-set only. The number of rules is selected according to the number of appearances of a phoneme in the training data and the number of occurrence of a single variation.

The method allows the recognition of all English phonemes, if the English AM is used, or the phonemes of both languages in case of the mixed AM and gene/rates a phoneme-lattice. The rules are applied to this lattice to map Japanese phonemes on English ones and to add likely variations of all phonemes to the lattice. During transformation to a word lattice with a pronunciation dictionary it is decided which variation is correct. This way as much acoustical information as possible can be kept while the dictionary size can stay the same, and especially a native English dictionary can be used without modification. For the first series of experiments, the variation were extracted with the native phoneme-set and acoustic model to prove the performance of the method without distortion caused by the mixed acoustic model.

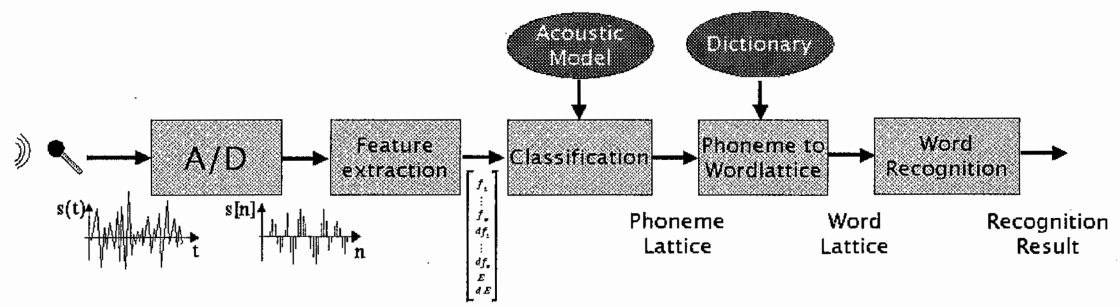

Figure 1: The recognition system introduced in this approach.

In [Tom00] a similar approach on word-lattice level was introduced and achieved good results. A disadvantage of this method is, that the initial recognition uses a native acoustic and language model and non-native variations are treated in the word-lattice and has therefore only influence the rescoring with the language model. Information about the original speech can be lost e. g. because of the dictionary and the acoustic model. This missing data can hardly be restored by processing the word-lattice.

Adding non-native pronunciation variations to the dictionary is another common way to achieve more robustness [EW00][][]. The size of the dictionary increases though only frequent variations are selected. This can lead to rising misrecognition rates.

2.1 Rule Generation

The proposed method is based on rules derived from frequent mispronunciations in non-native speech. For the creation of these rules, a careful analysis of the speech data is necessary. In a first step, the speech data is recognized on phoneme-level. The result and the given transcription are compared to extract correct recognitions, insertions, deletions and substitutions. After collecting these information from all of the training data (see Sec. 3.3), rules are generated to remove substitution errors from the lattice.

(2.1.1) Pronunciation Variation Analysis

Two monophone models, a native English and a native Japanese one, are trained (see Section (4.1.3)). The training data was taken from the WSJ- respective the TRA-database, described in Section 3. A third model is generated by combining the two native models.

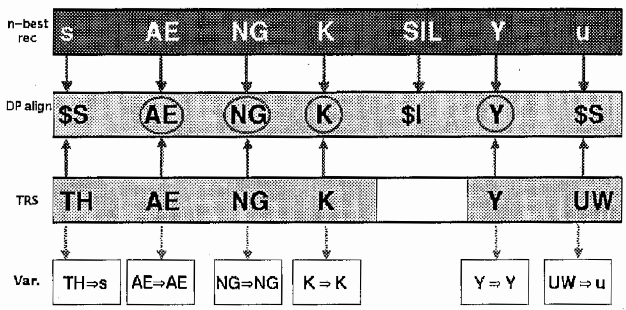

Figure 2: Example for rule derivation from n-best result and transcription with DP-alignment for "thank you". *$S* marks substitutions, *$I* insertions and *$D* deletions. Japanese phonemes are written in lower case and English phonemes in capital letters.

Monophone recognition was selected, as the influences of the context in bi- and triphones are considered to blur the non-native effect. Additionally the amount of non-native training data for rule generation is too small to produce a sufficient number of variations. During 1-best phoneme recognition of the non-native training data, no restrictions are given (similar to [Tom00]), which means that all phonemes defined in the acoustic model can be recognized. The recognition produces two results: 1-best and net-best (Fig. 3). The first one is the result with the highest acoustic score, latter shows the existence of the correct word sequence in the lattice if the transcription is given. To find the pronunciation variations in a speech file, both results are aligned by dynamic programming (DP) with the transcription [AKS00], which is defined to be the correct result. As the transcription is derived from the pronunciation lexicon, it can only represent one possible correct version. Alternatives may exist, but are not considered in this work. The results of this process is a string containing the information about correctly recognized phonemes, insertions, deletions and substitutions. As shown in Figure 2, the variations can be extracted by string-comparison.

The number of variations generated by this method is high (1298 in case of the native AM) and 95% have an appearance-frequency under 10%. This indicates a high number of random variations, which have to be discarded. To keep a high flexibility in selection, all of these variations are kept until the end of rule generation. Methods for limitation will be introduced later when the rules are applied to the lattice. The ideal number of rules has to be found during the evaluation.

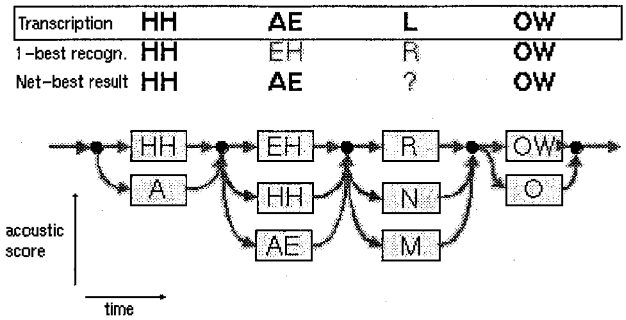

Figure 3: Simple example for 1-best and net-best evaluation for the word *hello*.

When using the net-best alignment, in case of misrecognition it is not possible to extract one single *most possible* variation, as all phoneme at that position are possible candidates. A possible way to solve this is to define the correct recognized phonemes as states of a *finite state automaton* and set the states for misrecognition to *any phoneme*. This way the start and end nodes for these places are fixed and it is possible to find the phoneme or phoneme sequence with the highest acoustic score in between (see Fig. 4).

Due to the size of the lattice, the calculation costs are high and time consuming. Thus only the first method was evaluated in this thesis. As the second approach is considered to produce more robust variations, it has to be subject of future research.

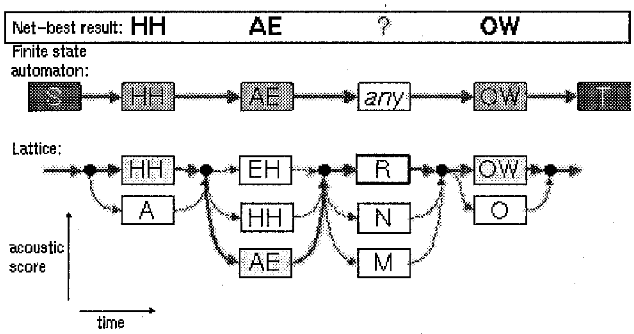

Figure 4: Example for net-best rule generation with finite state automaton. In this case R will be selected for the missing phone and creates the variation $L \rightarrow R$

(2.1.2) Data Driven Rule Generation

The rules for lattice processing need to fulfill the following requirements:

- to remove all Japanese phonemes from the lattice (in case a mixed AM is used) and
- add alternative variations which are likely to be correct.

To transform the phonemes in the lattice, rules have to be given, which show which phonemes ph_{corr} are likely to be correct, when the phoneme ph_{rec} is recognized. Not all variations have the same appearance probability $P(ph_{i,corr}|ph_{j,rec})$. With the number of occurrences $N_{ph_{i,rec}}$ of the recognized phoneme $ph_{i,rec}$ and $N_{ph_{rec}\rightarrow ph_{corr}}$ for the variation given, this probability can be calculated with

$$P_{sub} = P(ph_{i,corr}|ph_{j,rec}) = \frac{N_{ph_{i,rec}\rightarrow ph_{j,corr}}}{N_{ph_{i,rec}}}. \tag{1}$$

For the first experiments (Section (4.2.2)) this method of rule-generation was used.

So far, only the relation between $N_{ph_{i,rec}}$ and $N_{ph_{rec}\rightarrow ph_{corr}}$ is taken in account. If phoneme $ph_{j,corr}$ appears once in the training data and is recognized as $ph_{i,rec}$, this will lead to the probability of 1.0 for the rule $P(ph_{i,rec}|ph_{j,corr})$. Another phoneme $ph_{a,corr}$ appears 100 times and is 75 times recognized as $ph_{b,rec}$. Therefore the probability is 0.75. The first case might also be random while the second does not suffer to such an extent from insufficient training data. Nevertheless, according to the above calculation latter case has a lower probability. Therefore, the occurrence of $ph_{i,rec}$ relative to the number of all phonemes N_{all} must be taken into consideration. To allow a comparison, $N_{ph_{i,rec}}$ is normalized.

$$\overline{N}_{ph_i} = \frac{N_{ph_{i,rec}}}{N_{all}} \tag{2}$$

This value is calculated for each single rule and can then together with P_{sub} be used later on to suppress rules that are generated by rare appearing phonemes [AKS00].

For the final decision whether a rule is used or not, two thresholds are introduced. T_{sub} defines the minimum value for the substitution probability P_{sub} and $T_{\overline{N}_{ph}}$ is the limit for $\overline{N}_{ph_i}$.The decision is then implemented as

$$P_{sub} > T_{sub} \bigwedge \overline{N}_{ph_i} > T_{\overline{N}_{ph}}, \tag{3}$$

If the result is *true*, the rule is selected, otherwise it is dismissed.

A different method to include pronunciation variations into the dictionary found in [EW00] keeps only the most likely pronunciation of each word and then adds $n\%$ of its variations with the highest appearance frequency. Besides of using words instead of phonemes as units, this idea could also be transferred to this here introduced approach. The selection of rules is more depending on the number of occurrence of a phoneme in the test-data, so that frequently appearing phonemes have a higher number of variations than others. The number of variations is assumed to depend more on the phoneme itself than on the number of its occurrence, so that this method was not evaluated within this thesis and remains subject of future research.

Especially for Japanese accented English speech another method should be mentioned which was introduced in [Tom00] by Laura Mayfield Tomokiyo. The Japanese language contains many English words which are written in the in Katakana, a syllable character system. *Restaurant* is then written as *re-su-to-ra-n-to* or *Washington* as *wa-shi-n-to-n*. Influences of this writing can be recognized in the spoken language. In her approach every word in the pronunciation dictionary was also represented in syllable writing. For the approach of this thesis a transformation from syllable- to "normal" writing would be necessary, which is not so straightforward than the other way round.

2.2 Lattice Processing

Similar to the phoneme-recognition for the extraction of the variations (Section (2.1.1)), an initial phoneme lattice is created with an unrestricted recognition run. This lattice is then parsed arc by arc following the algorithm shown in Figure 5. For the phoneme ID in each arc a rule lookup is performed and if one or more rules are found, the corresponding alternatives are included between the same nodes. As already pointed out in Section (2.1.2), only rules that have been selected with Equation 3 were taken into account. This ensures, that rare and therefore insignificant variations will not corrupt the recognition.

The original lattice contains the likelihood $L(ph_{orig})$ for each source phoneme. Together with the probability of a substitution $P(ph_{i,corr}|ph_{j,rec})$ from Equation 1, the likelihood of the new arc is then

$$L(ph_{new}) = L(ph_{orig}) + log\left(P(ph_{corr}|ph_{rec})\right). \tag{4}$$

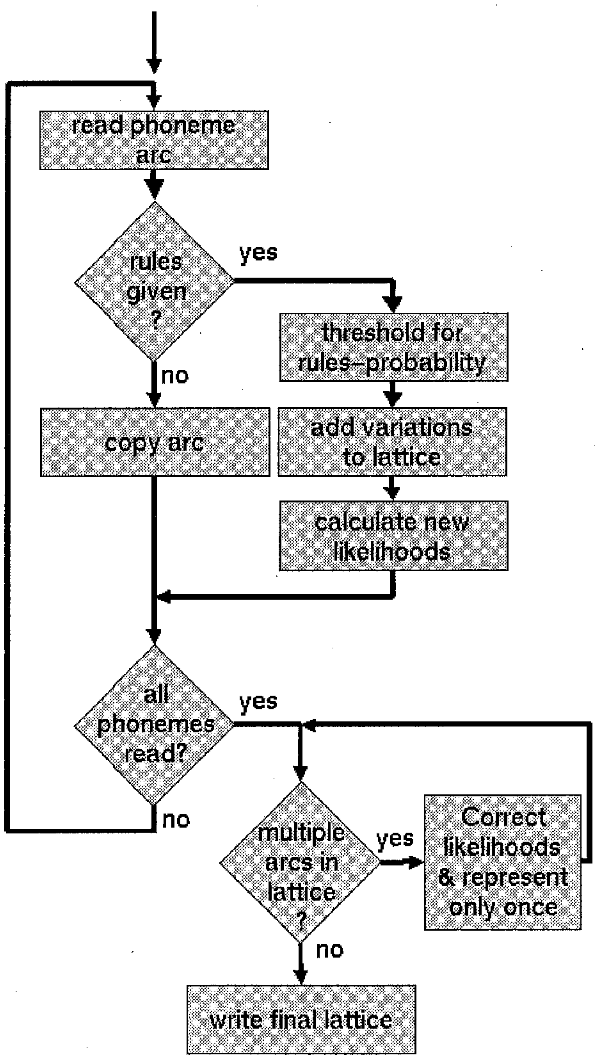

Figure 5: Algorithm for the lattice processing.

Figure 6 shows a comparison between a lattice before and after processing. The so generated new lattice has the same number of nodes as the original one, but contains more arcs. In case of a recognition with the mixed acoustic model, the lattice contains phonemes of both languages, but after processing only phonemes of the English set are remaining.
The original lattice already contains frequently more than one link between two nodes. If

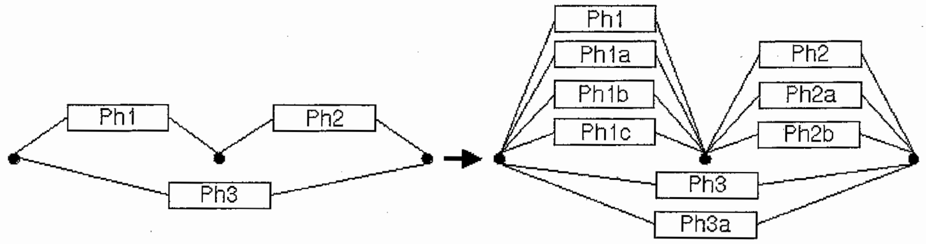

Figure 6: An example for a lattice before (left) and after processing (right).

additional variations are included, it can happen that one phoneme i appears j times between the same nodes. In this case, those multiple transitions have to be merged to one single arc. To keep the correct log-likelihood score $L(ph_i)$, the likelihoods $l(ph_{i,j}) = exp(L(ph_{i,j})$ of the arcs $j = 1 \ldots n$ are summed up.

$$L(ph_i) = log\left(\sum_{j=1}^{n} exp\left(L(ph_{i,j})\right)\right) \tag{5}$$

2.3 Word Recognition

All steps until now were limited to the phoneme level. To apply this method to continuous speech recognition, it has to be extended to word level. Therefore the modified phoneme-lattice has to be transformed to a word lattice and then rescored using a language model to get a final recognition result.

(2.3.1) Lattice Transformation

The entries in the pronunciation dictionary are defined as phoneme sequences. For the conversion, these sequences have to be extracted from all possible phoneme combinations in the lattice. The calculation costs for this process depend much on the sizes of the dictionary and the lattice. As the phoneme lattices tend to have up to $2 * 10^3$ nodes and the dimension of arcs is 10^6 this problem is not solvable in a reasonable time. For the calculation of the *word correct* rate only the number of correct words as defined in the transcription are necessary. Thus, the dictionary-lookup was restricted to those entries. The lattice in Figure 7 also contains the word *low (L+OW)*, which cannot be extracted with this method. The whole process gets faster, though it has an influence on the calculation of the recognition rate (see Sec. (2.3.2)).

Another way to solve this problem is to perform the word search not only with a dictionary, but in combination with a language model. This way the search area can be limited and the calculation gets fast, too. The used decoder does not support rescoring of an input lattice with an LM and therefore this method could not be tested.

(2.3.2) Word Recognition Rate

The word lattice is DP-aligned with the transcription and the recognition rate is calculated. Independent of the method of transformation from phoneme to word lattice, there are always some phonemes left in the word lattice which will cause insertion errors. Especially for the

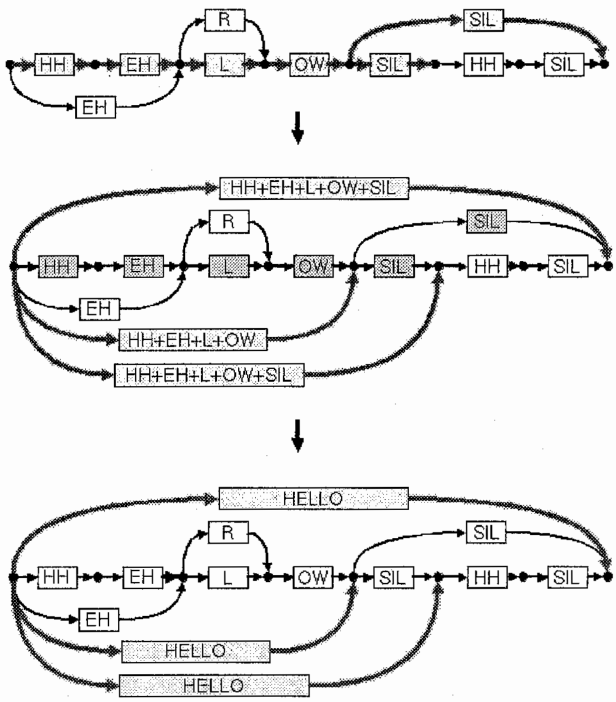

Figure 7: Example for conversion of a phoneme- into a word-lattice.

in Section (2.3.1) discussed method with the minimized dictionary, the number of words represented by these insertions is not known. There is also no information about which word was not recognized (deletion) and which one was recognized wrong (substitution). As an approximation, a sequence of misrecognized phonemes is assumed to represent one word and therefore one single mistake. This can be justified by the shortness of the utterances with an average of five words. The recognition sequence

$I $I $I $I $I $I $I $I 10052 10094 $S

is shortened to:

$I 10052 10094 $S.

The approximated word accuracy (WA) is then calculated as:

$$Errors = N_{rec.words} - N_{correctwords} \tag{6}$$

$$WA = \frac{N_{wordsintrans} - Errors}{N_{wordsintrans}} \tag{7}$$

The word correct recognition rate can be calculated as usual:

$$WC = \frac{N_{correctwords}}{N_{wordsintrans}} \tag{8}$$

2.4 Transfer on Other Languages

Considering the location of ATR and the importance of English speech recognition in international conversation, the signification of Japanese people speaking English as a research topic is quite obvious. But the recognition of non-native speech is important for other languages as well. Also the different dialects of one language cause a decrease in recognition accuracy because of its variations in pronunciation.

The proposed method can be adapted fast to new environments, if a sufficient amount of data for the rule generation is available. Once those rules exist, this approach should be able to cope with the new conditions.

ATR-SLT's non-native database also contains the speech of four Germans speaking English. This is a rather small amount and insufficient for statistically reliable rule-generation and evaluation, but it allows a first look at the possibilities of a transfer of this approach to another language.

3 Database

3.1 Wall-Street Journal (WSJ)

The *Continuous Speech Recognition Wall Street Journal Phase I (CSR-WSJ0) Corpus* [ARP94] consists of news from the WSJ, recorded simultaneously with close talking and a distant microphone. The native English acoustic model was trained with the SI- training set, offering 7200 utterances and a total of 12 hours of read speech.

3.2 ATR-TRA Database

This native Japanese database was originally introduced by ATR in 1996 as *SDB-L* [NMS+96] as one part of three databases of native Japanese speech (see Table 1). Because of the specific design for the travel- and hotel-reservation domain, it is also known as ATR-TRA. Main application of this collection is research in large vocabulary continuous speech recognition and more robustness for spontaneous speech. The semi-spontaneous speech of 215 Japanese persons with a total of 655 conversations was recorded. The recording conditions can be seen in Table 2. For the conversation content guideline basic information was provided to the subjects. The size of vocabulary is in the range of 10^4 words. There was no selection of speaking skill, but people with experience in hotel reservation were taking the role of the hotel clerks to achieve a more realistic conversation.

	SDB-L	SDB-S	SDB-I
speaking style	dialogue	dialogue	read and dialogue
vocabulary size	$\approx 10^4$	$\approx 10^3$	$\approx 10^5$ & $\approx 10^3$
spontaneity	middle	high	middle
# of speakers	$\approx 10^2$	$\approx 10^2$	$\approx 10^4$

Table 1: Features of the database (adapted from [NMS+96]).

microphone:	uni-direcitonal (Sanken MU-2C, Sony C-355)
sampling frequency:	48 kHz
quantization:	16 bit linear
noise environment:	quiet
facial content:	none
dialoge topology:	1 to 1
language:	Japanese monolingual

Table 2: Recording conditions (adapted from [NMS+96]).

3.3 Non-Native English Database

In this database read English speech of different native and non-native speakers is collected. Topic is - like in TRA - hotel reservation. Additionally, phoneme balanced sentences and sequences of digits similar to telephone- or credit-card numbers are included. The detailed information is given in Table 4

The number of Japanese speakers was originally limited to two male and nine female speakers. For derivation of the variances and evaluation, additional data-collection was necessary. The hardware for the recording was the same than that of the given data, but the room was different. The recording of the new data was carried out in an anechoic chamber with a Sennheiser HMD 410 close talking microphone headset and a Sony DTC-2000ES DAT device. The data was transfered via DAT-link to a computer outside the room where it was stored.

The sentences were displayed on a computer screen. The speakers were allowed to read the text before recording to reduce nervosity effects. Afterwards, each utterance was checked and re-recorded in case of reading mistakes such as missing words or stuttering to minimize those errors, while those of non-nativeness were kept. A positive side effect is that the data sounds more natural, as after test-reading a sentence, the speakers also have increased confidence.

The final number of speakers was increased to eight male and nine female speakers. Due to longterm stays in English speaking countries, the speaking proficiency of all female speakers was better than that of the male. From every speaker a small part of utterances were taken for evaluation, the rest for rule generation (see Table 4). The amount of data

	training	evaluation	total
hotel reservation	612	244	856
digits	250	—	250
phoneme balanced	480	—	480
total	1342	244	1586

Table 3: Number and type of utterances for training and evaluation.

	male	female	total
for rule generation	94	84	178
evaluation	14	12	26
total	108	96	204
total speakers	8	9	17

Table 4: Database overview on the amount of non-native data in minutes.

may be insufficient to completely rule out speaker specific effects. A bigger fundament could be achieved by adding similar data collected at CMU [Tom00], but that data is not open to the research community at the time of this thesis.

3.4 Native English Database

The data of three native English speakers is evaluated as a native English baseline. The recorded data was the same as for the non-native database in Section 3.3.

3.5 Non-Native German Database

Four German male members of the ATR-SLT research staff were recorded along with the original and the new data of the non-native database. The contents of the recording is the same as for the non-native speakers.

3.6 Phoneme Sets

(3.6.1) English Phoneme Set

Most research on the WSJ data bases on a similar phoneme set, which is also standard at ATR. The notation is in capital letters and it contains the symbol *SIL* for silence. In combination with HTK, additionally the short-pause symbol *SP* is included.

AA AE AH AO AW AX AXR AY B CH D DH DX EH ER EY F G HH
IH IX IY JH K L M N NG OW P R S SH T TH UH UW V W Y Z SIL

Table 5: The English 42 model phoneme set.

(3.6.2) Japanese Phoneme Set

ATR-SLT's standard phoneme set for Japanese speech was also chosen for this work. For easier distinction from English phonemes in the mixed phoneme set (Sec.: (3.6.3)), this set uses lowercase letters and the "-" character for silence and if needed, e.g. for the HTK-toolkit, *sp* for short-pause. At ATR all transcriptions and the dictionary were designed for this set.

a i k j o zh z u d m g ch ng r sh ts s e b q t w n p h f -

Table 6: The Japanese phoneme set with 27 models.

(3.6.3) Mixed Phoneme Set

This set was especially created for these experiments. The English and Japanese phoneme set were merged and phonemes which shared the same SAMPA symbol were only represented once in the English notation. Also only the English symbol *SIL* for silence was included. All recognition is performed with ATR-SPREC, which does not support a short pause. The different notations with capital letters for English and lowercase for Japanese in one set was chosen on purpose. It allows to check for leftover Japanese phonemes in the processed phoneme lattice.

AA AE AH AO AW AX AXR AY B CH D DH DX EH ER EY F G HH
IH IX IY JH K L M N NG OW P R S SH T TH UH UW V W Y Z SIL
a i k o zh u r sh e w n h f

Table 7: Phonemes of the mixed phoneme set with 55 models.

4 Experiments

4.1 Preparation

(4.1.1) Data Collection

As already pointed out in Section 3.3, additional recording of non-native speech was necessary to get more data for a training- and a test-set. ATR-SLT has its own anechoic experimental and recording room where the collection took place. Six male Japanese and one male German speakers were recorded for the experiments. All of them were recruited among ATR staff and had certain experience in speech recognition. Each recording lasted about two hours and was performed under supervision to ensure good results.

(4.1.2) Feature Calculation

The feature vectors were generated with the HCopy module from the HTK Toolkit [Y^{+}99]. Each vector contains 26 components, consisting of 12 MEL-frequency cepstral coefficients (MFCC), their first time-derivates, energy and delta-energy. A Hamming-window with a size of 20ms and a frame-shift of 10ms were set as parameters for the framing. A python-script was implemented which reads a list of audio files, calculates the features with HCopy and stores the results depending on language and training- or test-data. The generated MFCCs contain a 12bit HTK header which has to be removed for the usage with other software.

(4.1.3) Acoustic Modeling

WSJ and TRA have fixed pre-defined sets for training and evaluation. Two separate monophone AMs were trained on the native English respective the native Japanese training-data. An initial AM was created from a prototype model and the training data. This model was reestimated in three iterations Additional to the silence model, HTK has a short-pause model for the recognition of short silence parts between. SPREC uses an optional silence at the end of each word in the pronunciation dictionary instead. The short-pause model was included but not used during recognition. The Model with short-pause was reestimated again in ten iterations and a copy of each iteration level was kept for a comparison experiment of the AM performance (Sec.: (4.2.1)).

The silence conditions for TRA and the non-native database are different from those of WSJ. To increase recognition performance, the silence model in the English AM was substituted with that of the Japanese AM.

The target of this task is to recognize non-native speech. Therefore, for one approach both models are combined to one mixed AM to be able to recognize all phonemes of the mixed phoneme set. The HMMs of all phonemes defined in the mixed phoneme set were included.

(4.1.4) Rule Generation

All utterances of the non-native training-set were recognized and then DP-aligned with their transcriptions to derive the variations as described in Sections (2.1.1) and (2.1.2). The algorithm for deriving the variations had to be implemented additionally. Also the function to generate rules from those variations is part of this work.

Four sets of rules were generated:

- nnE: non-native English, trained from non-native speech with the native English AM

- mixed: mixed Rules, trained from non-native English and the mixed Japanese-English AM
- German: German Rules, trained from the small amount of German data available in the database

4.2 Evaluation

A first series of experiments should show the performance of the method. Therefore, only native phoneme-set and acoustic models were used for the rule generation and recognition of the non-native speech. The main reason was to avoid influences of the mixed acoustic models on the recognition result. The mixed acoustic model and phoneme-set were evaluated in further experiments.

(4.2.1) Acoustic Model Evaluation

In a first experiment the performance of the proposed method was evaluated, comparing acoustic models with a different number of reestimations. As baseline the native test-set (see Section 3.3) is evaluated. A set of 99 rules and the non-native test-set are the corpus for these experiments. The net-best accuracy for each of the results is calculated and shown in Table 8 and Figure 8.

	native 2.	nn 2	nn 4	nn 6	nn 8	nn 10
unprocessed	72.07%	60.87%	61.41%	61.65%	61.62%	55.7%
processed	—	67.44%	67.39%	67.71%	67.76%	61.1%

Table 8: Comparison of net-accuracy for unprocessed and processed lattices and AMs with a different number of iterations for reestimation.

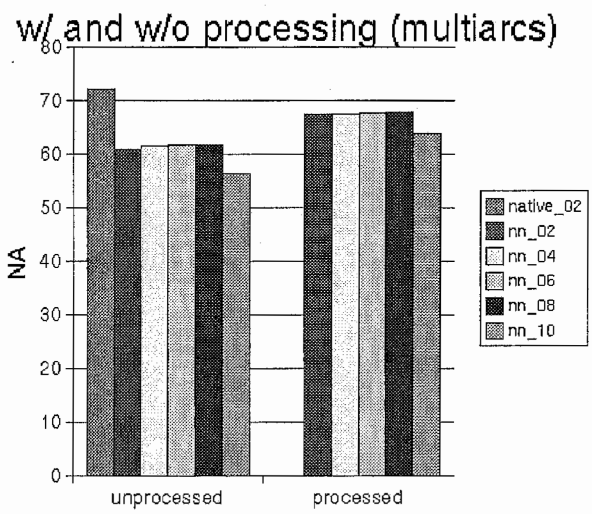

Figure 8: Comparison of the performance of acoustic models with a different number of iterations for reestimation.

(4.2.2) Non-Native Speech and Native Acoustic Model

In this Experiment, the performance of this approach depending on the number of rules was evaluated. The threshold for the occurrence-frequency of each variation varies from 3% to 11%. All rules with a lower frequency were discarded. The total number of rules in the rule-set was 1298. The Threshold for the phoneme appearance frequency calculated with Equation 2 in Section (2.1.2) did not have much influence. In case of a limit of 0.005,which is kept throughout the whole evaluation, the number of rules sorted out is 31. As can be

	unpr	0.03	0.07	0.08	0.10	0.11	nE
PNA	57.3%	64.7%	65.9%	64.4%	62.0%	60.9%	72.1%
PNC	100%	100%	100%	100%	100%	99.6%	100%
PBC	34.5%	30.1%	31.5%	31.5%	32.5%	32.3%	41.6%
# rules	—	236	80	65	50	44	—

Table 9: Phoneme net correct for the recognition of Japanese accented English with a native English acoustic model.

seen in Table 9 and Figure 9, the phoneme net-accuracy has a maximum for a threshold of 7% (80 rules). If more rules are applied, the recognition suffers from confusion, for fewer rules, the recognition result converges to the recognition rate of the unprocessed lattice. The overall phoneme best correct rate for the processed lattice is lower than for the unprocessed lattice. As net-best and net-correct results indicate, the correct phonemes exist in the lattice. One explanation is, that the calculation of the score during variation application has to be improved. Another hint for this suggestion is the word recognition rate from Section (4.2.4).

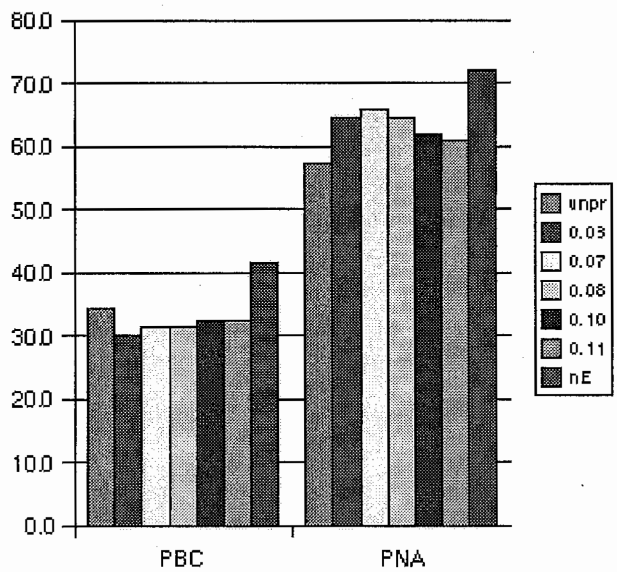

Figure 9: Results for the recognition of non-native speech with the native English acoustic model.

(4.2.3) Non-Native Speech and Mixed Acoustic Model

Similar to the previous experiment, the recognition performance depending on the threshold was calculated. Here, the mixed AM was applied to the decoding system. The total number of rules for this set is 1838 of which 3 are sorted out because of the phoneme occurrence probability (Sec.: (2.1.2)). Again, the phoneme best correct for the processed lattice is

	unpr	0.03	0.07	0.08	0.10	0.11
PNA	57.3%	77.1%	67.9%	64.4%	60.9%	69.4%
PNC	100%	100%	100%	100%	100%	100%
PBC	34.5%	19.3%	31.3%	31.5%	32.8%	24.6%
rules	—	658	83	46	13	8

Table 10: Phoneme net correct for mixed acoustic model.

smaller than for the unprocessed, but it shows clear that the recognition rates decrease for both, a high and a low number of rules.

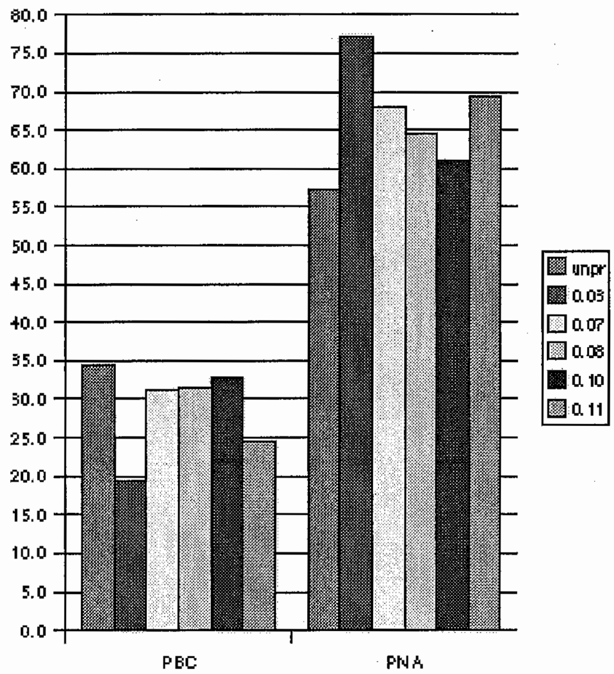

Figure 10: Results for the recognition of non-native speech with the mixed acoustic model.

(4.2.4) Word Recognition

The phoneme lattices of the previous experiments were transferred to word lattices and then evaluated with the methods introduced in Section 2.3.

For both, the native and the mixed acoustic model, the performance could be improved by applying additional variations to the lattice (See Tables 11 and 12). In case of the unprocessed lattices, the approach with the mixed acoustic model performs best. The reason is the ability to recognize Japanese phonemes, which occur in the non-native speech.

	unprocessed	processed
WNA	34.9%	47.0%
WNC	73.0%	82.3%

Table 11: Estimated word net accuracy and word net correct for the native English acoustic model.

	unprocessed	processed
WNA	39.2%	47.0%
WNC	74.0%	79.2%

Table 12: Estimated word net accuracy and word net correct for the mixed English-Japanese acoustic model.

(4.2.5) Application of Rules to Pronunciation Dictionary

To estimate the performance of former experiments a conventional recognition run is performed. For an additional experiment the previous generated rules were applied directly to the dictionary. To allow a comparison, the same acoustic models, test-set and rules were used. Two methods were evaluated which differ in the way the rules were applied to the single word entries:

- Method one added the pronunciation variations as alternatives to the original phoneme. With the same variations, the results in recognition should be the same as those of the lattice processing approach. Thus, this is a good basement for comparison. *Hello* is e. g. defined as $h\ \{a|e\}\ \{l|r\}\ o$. All phonemes in brackets are possible alternatives of which one has to be selected. So all combinations of this variations are possible (Table 13). The size of the dictionary increases and calculation costs are high.
- the second method applies only one modification at a time per word entry (Table 14). Thus less variations are applied to the dictionary and the recognition is faster compared to the previous version. Although, not all possible variations can be recognized.

For the word *hello* these two methods produce the following results:

hello	→	halo
		haro
		helo
		hero

Table 13: Possible variations for *hello* created with the first version.

The comparison shows, that the recognition rates are higher for the first method. This indicates, that more than one variation per word appear which have to be considered. The word net-correct is significant lower than that of the lattice processing methods while the word net-accuracy is higher. This shows, that the number of insertions for the dictionary method is lower. An effect that can be caused by the restricted recognition in combination with the pronunciation dictionary. The word recognition of this approach has to be compared with those results to evaluate the improvement or decrease of recognition performance.

hello	→	helo
		halo
		hero

Table 14: Possible variations for *hello* created with the second version.

	WBA	WBC	WNA	WNC
with LM				
unprocessed	24.3%	38.7%	45.3%	58.25%
processed 1	26.9%	40.6%	50.1%	61.1%
processed 2	25.8%	37.9%	49.1%	59.9%

Table 15: Recognition results for the standard unprocessed pronunciation dictionary and a dictionary including the new pronunciation variations.

(4.2.6) Extension to Other Languages: Non-Native German Speech

This approach can be applied on other languages, too, as already pointed out in Section 2.4. The German database (Sec.: 3.5) consists of only four speakers. Thus the rules for this evaluation are speaker dependent and will not represent all variations which are typical for Germans speaking English. The purpose of this experiment is to show up the possibility for a transfer on other languages.

	unprocessed	processed
WBC	4.4%	4.8%
WNA	31.1%	35.4%
WNC	33.5%	37.7%

Table 16: Phoneme net correct for recognition of German accented English with a native English acoustic model.

The phoneme recognition rates in Table 16 show the intended tendency but due to the small amount of data, the results are lower than those for Japanese speakers. For a bigger database, more reliable rules can be generated, which will increase the performance.

5 Discussion

5.1 Restricted Phoneme Recognition

The recognition experiments were run in sequential order:

- phoneme recognition
- processing of phoneme lattice
- conversion of phoneme- to word lattice
- calculation of recognition rate from word lattice

Thus no limitation in search space for phonemes during phoneme recognition and for words during the lattice transformation is applied. Therefore, the amount of data that has to be processed is huge and calculation costs are high. As a consequence, e. g. a real transformation from phoneme to word lattice was not possible. Also the number of possibly correct solutions gets bigger from step to step and getting the correct result will become difficult.

In conventional recognition, the pronunciation dictionary and language model are used during the recognition process, to discard phoneme sequences with a low acoustic- and language model score. This parallel use of information from the training data helps to decrease calculation costs and get a better performance in recognition. According to a discussion with other researchers at the laboratory, this approach is judged to achieve higher recognition rates than the direct application of rules to a dictionary as the likelihood scores during recognition can be included an the variations are applied more flexible. Research for this approach will be continued at ATR.

5.2 Improvement of Score Calculation

The results of the phoneme recognition indicate that the correct phonemes are applied to the lattice but for calculation of the recognition results, their score is too low to be chosen. Discussions about this problem resulted in two main ideas for improvement:

- The original phoneme always has the lowest likelihood score if it does not appear in its own rules as an identity. As the original phoneme is not wrong in general, this discriminates its probability of selection. By calculation of the average probability of all variations of one phoneme and subtract it from the each probability, some variations also will get a lower score than the original phoneme.
- A bonus score should be applied for phonemes that were recognized correctly during rule generation. This will lower the probability for randomly appearing variations to get high scores and thus are treated as best result. On the other hand, this might also decrease the chance for necessary variations. Which of these two effects is bigger has to be evaluated.

5.3 Biphone Models

The recognition performance for monophones is in general lower than that with a n-phoneme context dependency. The reason why it was selected for this approach is to extract variations in speech as precise as possible and avoid a blurring effect by caused by ngrams. Evaluation showed, that both, too many and too few rules perform worse. The correct selection of the right threshold is difficult but important for the results. The usage of bigrams can improve the recognition of the phonemes, especially as an influence of the syllable structure of the Japanese language is given [HZ01].

5.4 Bigger Non-Native Database

For the experiments of this thesis a non-native database of 17 Japanese speakers was used. In case of a bigger training-set for rule generation, typical pronunciation variations will become more definite, while the frequency for randomly occurring variations will decrease. The so created rules will be more precise and reliable and thus will increase the recognition performance.

5.5 Future Experiments

The next evaluation steps for this project will deal with the aspects already mentioned in the previous discussion. Additional to that, some other applications have to be mentioned.

In Section (4.2.6) already a first evaluation was introduced. As the size of the database was small, further experiments with a bigger amount of data and also data of other languages have to be performed.

Also the regional dialects cause problems in recognition. Evaluations have to show if the performance of a recognition system can be increased if this approach is applied to native but dialected speech.

A final introduced experiment will be a test in a standard recognition system which already achieves good results. In case of systematic recognition errors caused by e. g. the acoustic model, these also can be eliminated and increase recognition rates.

6 Summary and Conclusion

In this thesis, a brief introduction to speech recognition is given. Publications of recent research at conferences or in journals were analyzed and different methods for robust speech recognition and rule generation were discussed on the background of available databases. Especially the problems of robust recognition of non-native speech is treated. After comparing the different possibilities, the method of phoneme lattice processing was proposed. It is based on rules, extracted in a data driven way from the phoneme-recognition result of the non-native training set. These rules are applied to the phoneme lattice during the recognition, offering new pronunciation variations with a high probability to be correct. The modified lattice is converted to a word lattice. The recognition performance is evaluated on phoneme- and word-level.

In order to get enough speech data for rule generation and evaluation of the system, the existing non-native database on English hotel reservation task was enlarged by recording six Japanese for the main experiment, and one German for the portation on other languages. The recording conditions and contents of the conversations were the same as those of the original data (Section 3).

An English and a Japanese acoustic model were trained separately on the databases of WSJ and ATR-TRA. A mixed model was generated by merging these two models to one (Section 3).

The algorithms for the extraction of pronunciation variations, rule generation and modification of the lattice were formulated as introduced in Section 2 and implemented, including existing modules of the decoder.

For the evaluation (Section 4), two sets of rules were generated separately. One with the native English, and one with the mixed acoustic model. The recognition is performed with the native English AM and rule set, or the mixed AM and rules. The recognition results for these two combinations are calculated with and without lattice processing. The recognition rates could be improved relative to the unprocessed baseline for the native English and the mixed acoustic model on both, the phoneme- and word-level.

The rules were also applied directly to the native English pronunciation dictionary to allow a comparison of the recognition performance with a conventional system. Two different methods of application were tested and both of them achieved better recognition performance than the recognition with the original native dictionary. The dictionary size grows and calculation costs increase. Another disadvantage becomes obvious in case of changes in the rule set or the language. The whole dictionary has to be created new for this method, while for the lattice processing method only a new set of rules is needed. Nevertheless, this evaluation is necessary to estimate, which performance is possible with the acoustic models and test-data of this experiments.

A small database of German speakers was used to test the approach on a second language. Rules for German pronunciation variations were generated and applied to the phoneme lattice in the same way than for the Japanese accented data. An increase in recognition performance could be achieved which will become bigger in case of a larger training set for rule generation. But also this first evaluation with the small database showed the right tendency and thus proved the possibility of portation to other accents.

The results of the evaluation showed the right tendency of recognition rates. Nevertheless, the analysis of the experiments also indicated some problems which were discussed in Section 5. The unrestricted phoneme recognition was identified to decrease the performance and caused big lattices with high calculation costs. The score calculation during phoneme recognition was often not in favor of the correct phoneme in the lattice and thus decreased the recognition accuracy.

Possible solutions are a context-dependent phoneme recognition which will create more accurate rules and a better phoneme lattice. The variations should be applied during the phoneme recognition. Unlikely sequences of phonemes should be discarded during the lattice creation using a native dictionary and language model. Thus, the lattice size will be smaller which will decreases calculation costs. The calculation of the acoustic score during rule application has to be changed according to the suggestions in Section 5 to enable the selection of the correct phoneme from the lattice.

A bigger non-native database like that from CMU [Tom00] would allow to create more precise and speaker independent rules. Also other methods of rule generation which require more training data e. g. with left and/or right phoneme-context dependency could be evaluated.

The approach introduced in this thesis showed an improvement in recognition performance for non-native speech, but further improvements are necessary to achieve better results. Papers about this research were submitted to the *Spring Meeting of the Acoustic Society of Japan* [BGN02] and *ICSLP 2002* (to appear).

A Software

A.1 Software Manuals

This chapter contains the short-manuals to the software created for the dynamic phoneme lattice processing. As all of the scripts despite the AM training create big lattices with high calculation costs. Thus, it is recommended to start these scripts with *nohup* and write a logfile (> &! logfile) in case the computer crashes. Another big problem is the size of the lattices. It has to be made sure, that enough disc space is available!

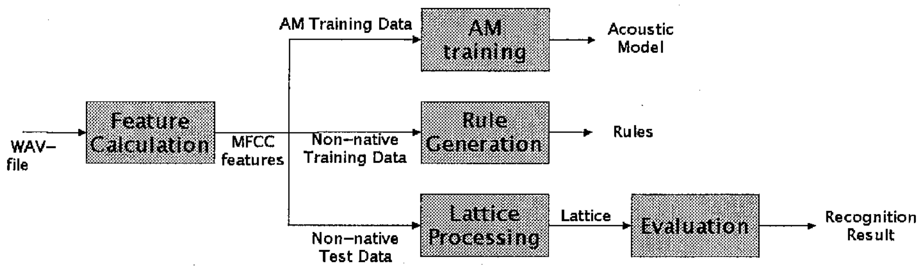

Figure 11: Survey over software.

The types of data used during the evaluation are defined as follows

.16k	audio data
.mfc	MFCC files (HTK)
.lat	phoneme lattices
.final.lat	processed phoneme lattices
.word.lat	word lattices
.final.word.lat	processed word lattices
.nb	n-best results
.var	files containing the variations
.rul	rules
.rescore	recognition results of the processed lattices
.compare	recognition results of the unprocessed lattices

Table 17: Types of data.

The options for each skript are given for each file in the following sections. In case *int* or *string* follows an option, it has to be separated with a space from the option. E. g.: `-h 12` for starting setting the headerbytes to 12. All other options have to be used alone e. g. `-nAM` is a switch to use the native AM instead of the default mixed AM.

(A.1.1) AM_train.py and definitions.py

This script follows the steps given in the HTK-book [Y+99] to train an acoustic model. The language specifications have to be defined in the "language configuration" part of the software. Additionally, *definitions.py* contains some basic definitions for the AM.

- silence module

- short pause module
- script to start HHEd

These are imported by *AM_train.py*.

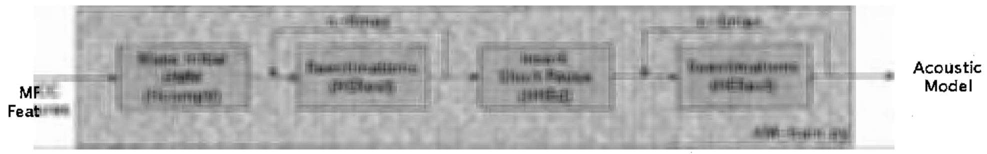

Figure 12: Acoustic model training.

The options are shown in Table 18.

[1|2]	training-part
-l en|jap	language "en|jap"
-i	additional information
-help	help message

Table 18: options for AM_train.py.

There is a problem starting HHEd from the python script. Thus after part 1 is finished, the command line which has to be executed to start HEReest is printed on the screen and has to be executed manually. After that, part 2 has to be started to finish the rest of the reestimations.

(A.1.2) generate_rules.py

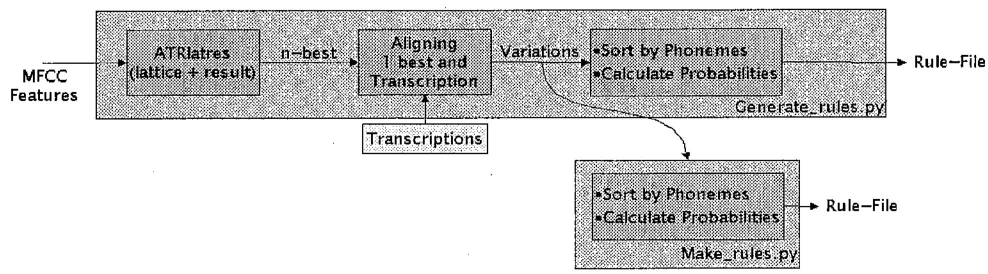

Figure 13: Rule generation.

Variations are extracted from an alignment of transcription and the here calculated 1-best recognition result. For each input-file a separate output-file with the ending **.var* is generated. At the end, all variation files are read in and are combined combined to one rule file (*.rul).

In the case that the script crashes, the already generated variations can be converted to rules with the *make_rules.py* script (Section (A.1.3)).
usage: generate_rules.py *id* [options]

-nE	using native English speech data (else: non-native)
-nAM	using native English AM (else mixed J-E-AM)
-f	define own file-list for mfcc input files
-r *int*	define reestimation level of AM (default: 9)
-l all|rules	make all (=default) or generate only rules
-n *int*	n-best
-h *int*	headerbytes
-local	use own SPREC (pre-release-version) in /data/tetsu1/xnbinder/SPREC/
-help	help message

Table 19: options for generate_rules.py.

If the option *local* is set, a local version of ATRlatres is used (defined in the settings) which is independent from the version of ATR-SPREC in $ATRSPREC. If a newer version of the combined ATRlattice and result should be used, this can be defined in the source-code.

(A.1.3) make_rules.py

This is a stand alone version of the rule generation part of the software in the previous section. It reads a filelist of *.var files and generates one rule-file *.rul (Section 13).

Two versions exist:

- make_rules.py generates rules with probabilities needed for rule-application to the dictionary ($P\,(correct \rightarrow variation)$).
- make_rules_V2.py generates rules with probabilities needed for rule-application to the lattice ($P\,(variation \rightarrow correct)$). This version was used in the Evaluatins for the paper

Usage is: make_rules.py variation-filelist rule-file [file_id]

variation-filelist	list containing the needed *.var files
rule-file	outputfile (including *.rul*)
file_id	ID for identification e. g. nnE_mixAM

Table 20: options for make_rules.py.

(A.1.4) add_rules_V5.py

A list of lattice files and a rule-file are read and the lattices are processed according to the rules (Section (2.1.2)). Various options are given to define the input language, the AM and so on. The values for these options are defined in *config_languages.py*

Please note that the 3gram AM was included for experimental reasons only and is not yet tested.

(A.1.5) phoneme2word

This script transforms a set of phoneme-lattices into a word lattice. To reduce calculation costs, not the whole dictionary is used, but only the words defined in the transcription

-nE	native English (default non-native E)
-nAM	native AM (default JE mixed AM)
-3gram	native English 3-gram AM (SPREC)
-nbest *int*	nbestvalue
-d *string*	dictionary
-c	calculate recognition-rates from original lattice
-h *int*	headerbytes
-id *string*	define identification for test-series
-dir *string*	write results into subdir in your working dir
-compare	don't process lattice, just rescore
-local	use local version of SPREC
-info	version information

Table 21: options for add_rules.py.

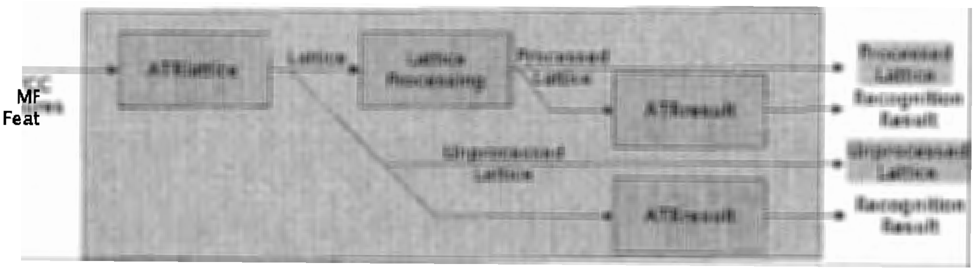

Figure 14: Lattice processing.

of the corresponding task. This reduces the calculation costs, but permitts to calculate the correct *word accuracy* or *phoneme-accuracy*. The wordlattice is piped through ATR result to calculate the recognition rates. These rates treat single each single not transferred phoneme as an insertion error.
The script is started with:
`phoneme2word_lattice.py flist_mix_final_lat2 hrt.1000_nr_compound.lex mix_final_p2w`

The program *find_words* for finding the phoneme sequences in the lattice was implemented by Yamamoto and can be found in the same directory as the other scripts.

(A.1.6) parse_word_lattice.py

The created recognition results of the previous script is parsed and a series of phoneme errors is merged to one. And the rates are re-calculated. As the calculation contained a bug during the evaluation, the final recognition rates were calculated with the following script which parses the logfile. Therefore, a logfile has to be written when running this script!

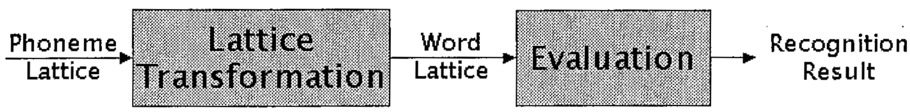

Figure 15: Lattice evaluation.

(A.1.7) eval_phoneme_lattice.py

This file parses the logfile and calculates the *word correct* and the *estimated word accuracy.*

B Rules

B.1 non-native English Speech and native English AM

In this part, the different sets of rules, used for the evaluation are shown in detail. The data is taken from the logfile of the lattice processing software, which included the final selection of the rules. An explanation about the different entries is given in table 22.

column	content
1	recognized phoneme
2	additional variation
3	P_{sub}
4	occurrence of this variation
5	occurrence of recognized phoneme
6	number of all phonemes

Table 22: Information about the rule set

(B.1.1) Rules for an Occurrence Probability > 3%

AA AA 0.151515151515 30.0 198.0 15585.0
AA AH 0.136363636364 27.0 198.0 15585.0
AA AX 0.0757575757576 15.0 198.0 15585.0
AA ER 0.0353535353535 7.0 198.0 15585.0
AA F 0.0353535353535 7.0 198.0 15585.0
AA N 0.106060606061 21.0 198.0 15585.0
AA R 0.0707070707071 14.0 198.0 15585.0
AA W 0.0454545454545 9.0 198.0 15585.0
AE AE 0.26938239159 205.0 761.0 15585.0
AE AH 0.030223390276 23.0 761.0 15585.0
AE AY 0.0604467805519 46.0 761.0 15585.0
AE EH 0.101182654402 77.0 761.0 15585.0
AE N 0.109067017083 83.0 761.0 15585.0
AE R 0.0341655716163 26.0 761.0 15585.0
AH AE 0.037147102526 25.0 673.0 15585.0
AH AH 0.187221396731 126.0 673.0 15585.0
AH AX 0.0490341753343 33.0 673.0 15585.0
AH IX 0.038632986627 26.0 673.0 15585.0
AH L 0.0430906389302 29.0 673.0 15585.0
AH N 0.0579494799406 39.0 673.0 15585.0
AH OW 0.0490341753343 33.0 673.0 15585.0
AH R 0.0668647845468 45.0 673.0 15585.0
AH UW 0.0475482912333 32.0 673.0 15585.0
AO AA 0.0332103321033 9.0 271.0 15585.0
AO AH 0.0627306273063 17.0 271.0 15585.0
AO AO 0.110701107011 30.0 271.0 15585.0
AO AX 0.0405904059041 11.0 271.0 15585.0
AO F 0.129151291513 35.0 271.0 15585.0
AO L 0.0332103321033 9.0 271.0 15585.0
AO N 0.0442804428044 12.0 271.0 15585.0
AO OW 0.09963099631 27.0 271.0 15585.0
AO R 0.0774907749077 21.0 271.0 15585.0
AO W 0.09963099631 27.0 271.0 15585.0
AW AA 0.0384615384615 6.0 156.0 15585.0
AW AE 0.0512820512821 8.0 156.0 15585.0
AW AH 0.0512820512821 8.0 156.0 15585.0
AW AW 0.115384615385 18.0 156.0 15585.0
AW AX 0.0576923076923 9.0 156.0 15585.0
AW EH 0.0705128205128 11.0 156.0 15585.0
AW IY 0.0384615384615 6.0 156.0 15585.0
AW L 0.0384615384615 6.0 156.0 15585.0
AW N 0.0897435897436 14.0 156.0 15585.0
AW OW 0.0705128205128 11.0 156.0 15585.0

AW R 0.0705128205128 11.0 156.0 15585.0
AW V 0.0448717948718 7.0 156.0 15585.0
AW Z 0.0320512820513 5.0 156.0 15585.0
AX AH 0.0497237569061 9.0 181.0 15585.0
AX AX 0.14364640884 26.0 181.0 15585.0
AX DH 0.0331491712707 6.0 181.0 15585.0
AX EH 0.110497237569 20.0 181.0 15585.0
AX IX 0.0441988950276 8.0 181.0 15585.0
AX N 0.0828729281768 15.0 181.0 15585.0
AX R 0.0497237569061 9.0 181.0 15585.0
AX S 0.0386740331492 7.0 181.0 15585.0
AX T 0.0331491712707 6.0 181.0 15585.0
AX UW 0.0331491712707 6.0 181.0 15585.0
AX V 0.0441988950276 8.0 181.0 15585.0
AY AH 0.0452830188679 24.0 530.0 15585.0
AY AY 0.637735849057 338.0 530.0 15585.0
AY N 0.0396226415094 21.0 530.0 15585.0
AY R 0.0358490566038 19.0 530.0 15585.0
D B 0.0405982905983 19.0 468.0 15585.0
D D 0.262820512821 123.0 468.0 15585.0
D N 0.108974358974 51.0 468.0 15585.0
D R 0.034188034188 16.0 468.0 15585.0
D T 0.0448717948718 21.0 468.0 15585.0
D UW 0.034188034188 16.0 468.0 15585.0
D V 0.0320512820513 15.0 468.0 15585.0
D Z 0.0384615384615 18.0 468.0 15585.0
DH AH 0.0392749244713 13.0 331.0 15585.0
DH B 0.0876132930514 29.0 331.0 15585.0
DH D 0.0422960725076 14.0 331.0 15585.0
DH DH 0.138972809668 46.0 331.0 15585.0
DH EH 0.036253776435 12.0 331.0 15585.0
DH M 0.0302114803625 10.0 331.0 15585.0
DH N 0.0876132930514 29.0 331.0 15585.0
DH OW 0.0453172205438 15.0 331.0 15585.0
DH P 0.0332326283988 11.0 331.0 15585.0
DH R 0.0604229607251 20.0 331.0 15585.0
DX AH 0.032 8.0 250.0 15585.0
DX D 0.06 15.0 250.0 15585.0
DX DH 0.036 9.0 250.0 15585.0
DX EH 0.04 10.0 250.0 15585.0
DX K 0.036 9.0 250.0 15585.0
DX L 0.032 8.0 250.0 15585.0
DX M 0.032 8.0 250.0 15585.0
DX N 0.128 32.0 250.0 15585.0
DX R 0.072 18.0 250.0 15585.0
DX T 0.052 13.0 250.0 15585.0
DX V 0.1 25.0 250.0 15585.0
DX W 0.036 9.0 250.0 15585.0
EH AH 0.0380952380952 16.0 420.0 15585.0
EH AX 0.0357142857143 15.0 420.0 15585.0
EH EH 0.283333333333 119.0 420.0 15585.0
EH ER 0.0309523809524 13.0 420.0 15585.0
EH IX 0.0357142857143 15.0 420.0 15585.0
EH L 0.047619047619 20.0 420.0 15585.0
EH N 0.0833333333333 35.0 420.0 15585.0
EH R 0.0714285714286 30.0 420.0 15585.0
EH UW 0.0357142857143 15.0 420.0 15585.0
EY EH 0.051638530288 52.0 1007.0 15585.0
EY EY 0.334657398213 337.0 1007.0 15585.0
EY IH 0.0933465739821 94.0 1007.0 15585.0
EY IX 0.0397219463754 40.0 1007.0 15585.0
EY IY 0.0506454816286 51.0 1007.0 15585.0
EY L 0.0327706057597 33.0 1007.0 15585.0
EY N 0.0764647467726 77.0 1007.0 15585.0
EY R 0.0327706057597 33.0 1007.0 15585.0
F F 0.592274678112 276.0 466.0 15585.0
F OW 0.0321888412017 15.0 466.0 15585.0
F R 0.0343347639485 16.0 466.0 15585.0
HH AX 0.0373443983402 9.0 241.0 15585.0
HH D 0.0373443983402 9.0 241.0 15585.0
HH F 0.0414937759336 10.0 241.0 15585.0
HH HH 0.253112033195 61.0 241.0 15585.0
HH IY 0.0331950207469 8.0 241.0 15585.0
HH K 0.0373443983402 9.0 241.0 15585.0

HH N 0.0705394190871 17.0 241.0 15585.0
HH OW 0.0580912863071 14.0 241.0 15585.0
HH R 0.045643153527 11.0 241.0 15585.0
HH W 0.045643153527 11.0 241.0 15585.0
IH AX 0.037422037422 18.0 481.0 15585.0
IH D 0.037422037422 18.0 481.0 15585.0
IH EH 0.0706860706861 34.0 481.0 15585.0
IH IH 0.189189189189 91.0 481.0 15585.0
IH IX 0.037422037422 18.0 481.0 15585.0
IH IY 0.035343035343 17.0 481.0 15585.0
IH L 0.031185031185 15.0 481.0 15585.0
IH N 0.0602910602911 29.0 481.0 15585.0
IH R 0.101871101871 49.0 481.0 15585.0
IH T 0.033264033264 16.0 481.0 15585.0
IH UW 0.0852390852391 41.0 481.0 15585.0
IH V 0.033264033264 16.0 481.0 15585.0
IX AH 0.0561797752809 20.0 356.0 15585.0
IX AX 0.0505617977528 18.0 356.0 15585.0
IX AY 0.0393258426966 14.0 356.0 15585.0
IX IX 0.188202247191 67.0 356.0 15585.0
IX M 0.0365168539326 13.0 356.0 15585.0
IX N 0.129213483146 46.0 356.0 15585.0
IX OW 0.0421348314607 15.0 356.0 15585.0
IX R 0.0674157303371 24.0 356.0 15585.0
IX UW 0.0308988764045 11.0 356.0 15585.0
IY IH 0.134955752212 122.0 904.0 15585.0
IY IY 0.429203539823 388.0 904.0 15585.0
IY N 0.0818584070796 74.0 904.0 15585.0
K K 0.566153846154 184.0 325.0 15585.0
K N 0.0430769230769 14.0 325.0 15585.0
K R 0.04 13.0 325.0 15585.0
K T 0.0338461538462 11.0 325.0 15585.0
L AH 0.0326797385621 15.0 459.0 15585.0
L F 0.0305010893246 14.0 459.0 15585.0
L L 0.244008714597 112.0 459.0 15585.0
L N 0.037037037037 17.0 459.0 15585.0
L OW 0.135076252723 62.0 459.0 15585.0
L R 0.0871459694989 40.0 459.0 15585.0
L W 0.0762527233115 35.0 459.0 15585.0
M B 0.0585106382979 11.0 188.0 15585.0
M M 0.372340425532 70.0 188.0 15585.0
M N 0.175531914894 33.0 188.0 15585.0
M R 0.0425531914894 8.0 188.0 15585.0
M V 0.031914893617 6.0 188.0 15585.0
N M 0.051724137931 18.0 348.0 15585.0
N N 0.735632183908 256.0 348.0 15585.0
N R 0.0316091954023 11.0 348.0 15585.0
OW AA 0.03125 13.0 416.0 15585.0
OW IX 0.0504807692308 21.0 416.0 15585.0
OW L 0.0384615384615 16.0 416.0 15585.0
OW N 0.0841346153846 35.0 416.0 15585.0
OW OW 0.389423076923 162.0 416.0 15585.0
OW R 0.0456730769231 19.0 416.0 15585.0
OW W 0.0360576923077 15.0 416.0 15585.0
P D 0.0380434782609 7.0 184.0 15585.0
P IH 0.0380434782609 7.0 184.0 15585.0
P IX 0.0326086956522 6.0 184.0 15585.0
P K 0.0326086956522 6.0 184.0 15585.0
P L 0.0380434782609 7.0 184.0 15585.0
P N 0.070652173913 13.0 184.0 15585.0
P OW 0.0380434782609 7.0 184.0 15585.0
P P 0.157608695652 29.0 184.0 15585.0
P R 0.0978260869565 18.0 184.0 15585.0
P UW 0.0326086956522 6.0 184.0 15585.0
P W 0.0434782608696 8.0 184.0 15585.0
R F 0.0346820809249 6.0 173.0 15585.0
R L 0.121387283237 21.0 173.0 15585.0
R R 0.514450867052 89.0 173.0 15585.0
R W 0.0462427745665 8.0 173.0 15585.0
S R 0.0314049586777 38.0 1210.0 15585.0
S S 0.604132231405 731.0 1210.0 15585.0
S Z 0.0462809917355 56.0 1210.0 15585.0
T K 0.0911256700417 153.0 1679.0 15585.0
T N 0.0428826682549 72.0 1679.0 15585.0

T P 0.0333531864205 56.0 1679.0 15585.0
T R 0.0506253722454 85.0 1679.0 15585.0
T T 0.415128052412 697.0 1679.0 15585.0
TH F 0.0790774299835 48.0 607.0 15585.0
TH K 0.0329489291598 20.0 607.0 15585.0
TH L 0.0329489291598 20.0 607.0 15585.0
TH N 0.0774299835255 47.0 607.0 15585.0
TH OW 0.0395387149918 24.0 607.0 15585.0
TH P 0.0345963756178 21.0 607.0 15585.0
TH R 0.0626029654036 38.0 607.0 15585.0
TH T 0.0362438220758 22.0 607.0 15585.0
TH TH 0.130148270181 79.0 607.0 15585.0
TH UW 0.0313014827018 19.0 607.0 15585.0
TH Z 0.0313014827018 19.0 607.0 15585.0
UH IH 0.0336134453782 8.0 238.0 15585.0
UH L 0.134453781513 32.0 238.0 15585.0
UH M 0.0504201680672 12.0 238.0 15585.0
UH N 0.046218487395 11.0 238.0 15585.0
UH R 0.142857142857 34.0 238.0 15585.0
UH UH 0.0966386554622 23.0 238.0 15585.0
UH UW 0.0420168067227 10.0 238.0 15585.0
UH W 0.0546218487395 13.0 238.0 15585.0
UW IX 0.0344036697248 15.0 436.0 15585.0
UW M 0.0596330275229 26.0 436.0 15585.0
UW N 0.126146788991 55.0 436.0 15585.0
UW R 0.0550458715596 24.0 436.0 15585.0
UW UW 0.394495412844 172.0 436.0 15585.0
UW W 0.0366972477064 16.0 436.0 15585.0
V L 0.0393258426966 21.0 534.0 15585.0
V N 0.0692883895131 37.0 534.0 15585.0
V R 0.0486891385768 26.0 534.0 15585.0
V V 0.432584269663 231.0 534.0 15585.0
W W 0.836842105263 159.0 190.0 15585.0
Y IY 0.111764705882 19.0 170.0 15585.0
Y K 0.0352941176471 6.0 170.0 15585.0
Y N 0.0529411764706 9.0 170.0 15585.0
Y R 0.0588235294118 10.0 170.0 15585.0
Y Y 0.505882352941 86.0 170.0 15585.0
Y Z 0.0352941176471 6.0 170.0 15585.0
Z DH 0.0642458100559 23.0 358.0 15585.0
Z N 0.0586592178771 21.0 358.0 15585.0
Z R 0.0530726256983 19.0 358.0 15585.0
Z S 0.0307262569832 11.0 358.0 15585.0
Z T 0.0307262569832 11.0 358.0 15585.0
Z Z 0.416201117318 149.0 358.0 15585.0

RULES

SP: 0 (0.000000)

SIL: 0 (0.000000)

P(ph) < 0.005000: 79 (0.060863)

P(var) < 0.030000: 983 (0.757319)

rules: 1298

accepted rules: 236 (0.000000)

added jap. rules: 0 (0.000000)

(B.1.2) Selection for an Occurance Probability > 7%

AA AA 0.151515151515 30.0 198.0 15585.0
AA AH 0.136363636364 27.0 198.0 15585.0
AA AX 0.0757575757576 15.0 198.0 15585.0
AA N 0.106060606061 21.0 198.0 15585.0
AA R 0.0707070707071 14.0 198.0 15585.0
AE AE 0.26938239159 205.0 761.0 15585.0
AE EH 0.101182654402 77.0 761.0 15585.0
AE N 0.109067017083 83.0 761.0 15585.0
AH AH 0.187221396731 126.0 673.0 15585.0
AO AO 0.110701107011 30.0 271.0 15585.0

AO F 0.129151291513 35.0 271.0 15585.0
AO OW 0.09963099631 27.0 271.0 15585.0
AO R 0.0774907749077 21.0 271.0 15585.0
AO W 0.09963099631 27.0 271.0 15585.0
AW AW 0.115384615385 18.0 156.0 15585.0
AW EH 0.0705128205128 11.0 156.0 15585.0
AW N 0.0897435897436 14.0 156.0 15585.0
AW OW 0.0705128205128 11.0 156.0 15585.0
AW R 0.0705128205128 11.0 156.0 15585.0
AX AX 0.14364640884 26.0 181.0 15585.0
AX EH 0.110497237569 20.0 181.0 15585.0
AX N 0.0828729281768 15.0 181.0 15585.0
AY AY 0.637735849057 338.0 530.0 15585.0
D D 0.262820512821 123.0 468.0 15585.0
D N 0.108974358974 51.0 468.0 15585.0
DH B 0.0876132930514 29.0 331.0 15585.0
DH DH 0.138972809668 46.0 331.0 15585.0
DH N 0.0876132930514 29.0 331.0 15585.0
DX N 0.128 32.0 250.0 15585.0
DX R 0.072 18.0 250.0 15585.0
DX V 0.1 25.0 250.0 15585.0
EH EH 0.283333333333 119.0 420.0 15585.0
EH N 0.0833333333333 35.0 420.0 15585.0
EH R 0.0714285714286 30.0 420.0 15585.0
EY EY 0.334657398213 337.0 1007.0 15585.0
EY IH 0.0933465739821 94.0 1007.0 15585.0
EY N 0.0764647467726 77.0 1007.0 15585.0
F F 0.592274678112 276.0 466.0 15585.0
HH HH 0.253112033195 61.0 241.0 15585.0
HH N 0.0705394190871 17.0 241.0 15585.0
IH EH 0.0706860706861 34.0 481.0 15585.0
IH IH 0.189189189189 91.0 481.0 15585.0
IH R 0.101871101871 49.0 481.0 15585.0
IH UW 0.0852390852391 41.0 481.0 15585.0
IX IX 0.188202247191 67.0 356.0 15585.0
IX N 0.129213483146 46.0 356.0 15585.0
IY IH 0.134955752212 122.0 904.0 15585.0
IY IY 0.429203539823 388.0 904.0 15585.0
IY N 0.0818584070796 74.0 904.0 15585.0
K K 0.566153846154 184.0 325.0 15585.0
L L 0.244008714597 112.0 459.0 15585.0
L OW 0.135076252723 62.0 459.0 15585.0
L R 0.0871459694989 40.0 459.0 15585.0
L W 0.0762527233115 35.0 459.0 15585.0
M M 0.372340425532 70.0 188.0 15585.0
M N 0.175531914894 33.0 188.0 15585.0
N N 0.735632183908 256.0 348.0 15585.0
OW N 0.0841346153846 35.0 416.0 15585.0
OW OW 0.389423076923 162.0 416.0 15585.0
P N 0.070652173913 13.0 184.0 15585.0
P P 0.157608695652 29.0 184.0 15585.0
P R 0.0978260869565 18.0 184.0 15585.0
R L 0.121387283237 21.0 173.0 15585.0
R R 0.514450867052 89.0 173.0 15585.0
S S 0.604132231405 731.0 1210.0 15585.0
T K 0.0911256700417 153.0 1679.0 15585.0
T T 0.415128052412 697.0 1679.0 15585.0
TH F 0.0790774299835 48.0 607.0 15585.0
TH N 0.0774299835255 47.0 607.0 15585.0
TH TH 0.130148270181 79.0 607.0 15585.0
UH L 0.134453781513 32.0 238.0 15585.0
UH R 0.142857142857 34.0 238.0 15585.0
UH UH 0.0966386554622 23.0 238.0 15585.0
UW N 0.126146788991 55.0 436.0 15585.0
UW UW 0.394495412844 172.0 436.0 15585.0
V V 0.432584269663 231.0 534.0 15585.0
W W 0.836842105263 159.0 190.0 15585.0
Y IY 0.111764705882 19.0 170.0 15585.0
Y Y 0.505882352941 86.0 170.0 15585.0
Z Z 0.416201117318 149.0 358.0 15585.0

RULES
SP: 0 (0.000000)
SIL: 0 (0.000000)

P(ph) < 0.005000: 31 (0.023883)
P(var) < 0.070000: 1187 (0.914484)

rules: 1298

accepted rules: 80 (0.000000)

added jap. rules: 0 (0.000000)

(B.1.3) Rules for an Occurance Probability > 8%

AA AA 0.151515151515 30.0 198.0 15585.0
AA AH 0.136363636364 27.0 198.0 15585.0
AA N 0.106060606061 21.0 198.0 15585.0
AE AE 0.26938239159 205.0 761.0 15585.0
AE EH 0.101182654402 77.0 761.0 15585.0
AE N 0.109067017083 83.0 761.0 15585.0
AH AH 0.187221396731 126.0 673.0 15585.0
AO AO 0.110701107011 30.0 271.0 15585.0
AO F 0.129151291513 35.0 271.0 15585.0
AO OW 0.09963099631 27.0 271.0 15585.0
AO W 0.09963099631 27.0 271.0 15585.0
AW AW 0.115384615385 18.0 156.0 15585.0
AW N 0.0897435897436 14.0 156.0 15585.0
AX AX 0.14364640884 26.0 181.0 15585.0
AX EH 0.110497237569 20.0 181.0 15585.0
AX N 0.0828729281768 15.0 181.0 15585.0
AY AY 0.637735849057 338.0 530.0 15585.0
D D 0.262820512821 123.0 468.0 15585.0
D N 0.108974358974 51.0 468.0 15585.0
DH B 0.0876132930514 29.0 331.0 15585.0
DH DH 0.138972809668 46.0 331.0 15585.0
DH N 0.0876132930514 29.0 331.0 15585.0
DX N 0.128 32.0 250.0 15585.0
DX V 0.1 25.0 250.0 15585.0
EH EH 0.283333333333 119.0 420.0 15585.0
EH N 0.0833333333333 35.0 420.0 15585.0
EY EY 0.334657398213 337.0 1007.0 15585.0
EY IH 0.0933465739821 94.0 1007.0 15585.0
F F 0.592274678112 276.0 466.0 15585.0
HH HH 0.253112033195 61.0 241.0 15585.0
IH IH 0.189189189189 91.0 481.0 15585.0
IH R 0.101871101871 49.0 481.0 15585.0
IH UW 0.0852390852391 41.0 481.0 15585.0
IX IX 0.188202247191 67.0 356.0 15585.0
IX N 0.129213483146 46.0 356.0 15585.0
IY IH 0.134955752212 122.0 904.0 15585.0
IY IY 0.429203539823 388.0 904.0 15585.0
IY N 0.0818584070796 74.0 904.0 15585.0
K K 0.566153846154 184.0 325.0 15585.0
L L 0.244008714597 112.0 459.0 15585.0
L OW 0.135076252723 62.0 459.0 15585.0
L R 0.0871459694989 40.0 459.0 15585.0
M M 0.372340425532 70.0 188.0 15585.0
M N 0.175531914894 33.0 188.0 15585.0
N N 0.735632183908 256.0 348.0 15585.0
OW N 0.0841346153846 35.0 416.0 15585.0
OW OW 0.389423076923 162.0 416.0 15585.0
P P 0.157608695652 29.0 184.0 15585.0
P R 0.0978260869565 18.0 184.0 15585.0
R L 0.121387283237 21.0 173.0 15585.0
R R 0.514450867052 89.0 173.0 15585.0
S S 0.604132231405 731.0 1210.0 15585.0
T K 0.0911256700417 153.0 1679.0 15585.0
T T 0.415128052412 697.0 1679.0 15585.0
TH TH 0.130148270181 79.0 607.0 15585.0
UH L 0.134453781513 32.0 238.0 15585.0
UH R 0.142857142857 34.0 238.0 15585.0
UH UH 0.0966386554622 23.0 238.0 15585.0
UW N 0.126146788991 55.0 436.0 15585.0
UW UW 0.394495412844 172.0 436.0 15585.0
V V 0.432584269663 231.0 534.0 15585.0

W W 0.836842105263 159.0 190.0 15585.0
Y IY 0.111764705882 19.0 170.0 15585.0
Y Y 0.505882352941 86.0 170.0 15585.0
Z Z 0.416201117318 149.0 358.0 15585.0

RULES

SP: 0 (0.000000)

SIL: 0 (0.000000)

P(ph) < 0.005000: 23 (0.017720)

P(var) < 0.080000: 1210 (0.932203)

rules: 1298

accepted rules: 65 (0.000000)

added jap. rules: 0 (0.000000)

(B.1.4) Rules for an Occurance Probability > 10%

AA AA 0.151515151515 30.0 198.0 15585.0
AA AH 0.136363636364 27.0 198.0 15585.0
AA N 0.106060606061 21.0 198.0 15585.0
AE AE 0.26938239159 205.0 761.0 15585.0
AE EH 0.101182654402 77.0 761.0 15585.0
AE N 0.109067017083 83.0 761.0 15585.0
AH AH 0.187221396731 126.0 673.0 15585.0
AO AO 0.110701107011 30.0 271.0 15585.0
AO F 0.129151291513 35.0 271.0 15585.0
AW AW 0.115384615385 18.0 156.0 15585.0
AX AX 0.14364640884 26.0 181.0 15585.0
AX EH 0.110497237569 20.0 181.0 15585.0
AY AY 0.637735849057 338.0 530.0 15585.0
D D 0.262820512821 123.0 468.0 15585.0
D N 0.108974358974 51.0 468.0 15585.0
DH DH 0.138972809668 46.0 331.0 15585.0
DX N 0.128 32.0 250.0 15585.0
DX V 0.1 25.0 250.0 15585.0
EH EH 0.283333333333 119.0 420.0 15585.0
EY EY 0.334657398213 337.0 1007.0 15585.0
F F 0.592274678112 276.0 466.0 15585.0
HH HH 0.253112033195 61.0 241.0 15585.0
IH IH 0.189189189189 91.0 481.0 15585.0
IH R 0.101871101871 49.0 481.0 15585.0
IX IX 0.188202247191 67.0 356.0 15585.0
IX N 0.129213483146 46.0 356.0 15585.0
IY IH 0.134955752212 122.0 904.0 15585.0
IY IY 0.429203539823 388.0 904.0 15585.0
K K 0.566153846154 184.0 325.0 15585.0
L L 0.244008714597 112.0 459.0 15585.0
L OW 0.135076252723 62.0 459.0 15585.0
M M 0.372340425532 70.0 188.0 15585.0
M N 0.175531914894 33.0 188.0 15585.0
N N 0.735632183908 256.0 348.0 15585.0
OW OW 0.389423076923 162.0 416.0 15585.0
P P 0.157608695652 29.0 184.0 15585.0
R L 0.121387283237 21.0 173.0 15585.0
R R 0.514450867052 89.0 173.0 15585.0
S S 0.604132231405 731.0 1210.0 15585.0
T T 0.415128052412 697.0 1679.0 15585.0
TH TH 0.130148270181 79.0 607.0 15585.0
UH L 0.134453781513 32.0 238.0 15585.0
UH R 0.142857142857 34.0 238.0 15585.0
UW N 0.126146788991 55.0 436.0 15585.0
UW UW 0.394495412844 172.0 436.0 15585.0
V V 0.432584269663 231.0 534.0 15585.0
W W 0.836842105263 159.0 190.0 15585.0
Y IY 0.111764705882 19.0 170.0 15585.0
Y Y 0.505882352941 86.0 170.0 15585.0
Z Z 0.416201117318 149.0 358.0 15585.0

RULES

SP: 0 (0.000000)

SIL: 0 (0.000000)

P(ph) < 0.005000: 14 (0.010786)

P(var) < 0.100000: 1234 (0.950693)

rules: 1298

accepted rules: 50 (0.000000)

added jap. rules: 0 (0.000000)

(B.1.5) Rules for an Occurance Probability > 11%

AA AA 0.151515151515 30.0 198.0 15585.0
AA AH 0.136363636364 27.0 198.0 15585.0
AE AE 0.26938239159 205.0 761.0 15585.0
AH AH 0.187221396731 126.0 673.0 15585.0
AO AO 0.110701107011 30.0 271.0 15585.0
AO F 0.129151291513 35.0 271.0 15585.0
AW AW 0.115384615385 18.0 156.0 15585.0
AX AX 0.14364640884 26.0 181.0 15585.0
AX EH 0.110497237569 20.0 181.0 15585.0
AY AY 0.637735849057 338.0 530.0 15585.0
D D 0.262820512821 123.0 468.0 15585.0
DH DH 0.138972809668 46.0 331.0 15585.0
DX N 0.128 32.0 250.0 15585.0
EH EH 0.283333333333 119.0 420.0 15585.0
EY EY 0.334657398213 337.0 1007.0 15585.0
F F 0.592274678112 276.0 466.0 15585.0
HH HH 0.253112033195 61.0 241.0 15585.0
IH IH 0.189189189189 91.0 481.0 15585.0
IX IX 0.188202247191 67.0 356.0 15585.0
IX N 0.129213483146 46.0 356.0 15585.0
IY IH 0.134955752212 122.0 904.0 15585.0
IY IY 0.429203539823 388.0 904.0 15585.0
K K 0.566153846154 184.0 325.0 15585.0
L L 0.244008714597 112.0 459.0 15585.0
L OW 0.135076252723 62.0 459.0 15585.0
M M 0.372340425532 70.0 188.0 15585.0
M N 0.175531914894 33.0 188.0 15585.0
N N 0.735632183908 256.0 348.0 15585.0
OW OW 0.389423076923 162.0 416.0 15585.0
P P 0.157608695652 29.0 184.0 15585.0
R L 0.121387283237 21.0 173.0 15585.0
R R 0.514450867052 89.0 173.0 15585.0
S S 0.604132231405 731.0 1210.0 15585.0
T T 0.415128052412 697.0 1679.0 15585.0
TH TH 0.130148270181 79.0 607.0 15585.0
UH L 0.134453781513 32.0 238.0 15585.0
UH R 0.142857142857 34.0 238.0 15585.0
UW N 0.126146788991 55.0 436.0 15585.0
UW UW 0.394495412844 172.0 436.0 15585.0
V V 0.432584269663 231.0 534.0 15585.0
W W 0.836842105263 159.0 190.0 15585.0
Y IY 0.111764705882 19.0 170.0 15585.0
Y Y 0.505882352941 86.0 170.0 15585.0
Z Z 0.416201117318 149.0 358.0 15585.0

RULES

SP: 0 (0.000000)

SIL: 0 (0.000000)

P(ph) < 0.005000: 13 (0.010015)

P(var) < 0.110000: 1241 (0.956086)

rules: 1298

accepted rules: 44 (0.000000)

added jap. rules: 0 (0.000000)

B.2 Non-Native Speech and Mixed AM

In this section, the rule sets created with the mixed acoustic model are shown. The structure is the same as for the rules of the previous sections.

The difference to the rules created with the native model is, that all Japanese phonemes have to appear in the rule-set as they have to be mapped on a English phoneme. If a Japanese leaves in the lattice, it causes an insertion error, as the pronunciation dictionary is designed with the English phoneme set only. In case, all variations of a japanese phoneme were rejected, the variation with the highest occurance frequency was inserted.

(B.2.1) Rules for an Occurance Probability > 3%

AA EY 0.0344827586207 5.0 145.0 10786.0
AA IH 0.0620689655172 9.0 145.0 10786.0
AA IX 0.0344827586207 5.0 145.0 10786.0
AA K 0.0413793103448 6.0 145.0 10786.0
AA N 0.0896551724138 13.0 145.0 10786.0
AA OW 0.0344827586207 5.0 145.0 10786.0
AA R 0.0344827586207 5.0 145.0 10786.0
AA S 0.0758620689655 11.0 145.0 10786.0
AA T 0.0551724137931 8.0 145.0 10786.0
AA UW 0.0620689655172 9.0 145.0 10786.0
AA W 0.048275862069 7.0 145.0 10786.0
AE AE 0.0405405405405 3.0 74.0 10786.0
AE AH 0.0405405405405 3.0 74.0 10786.0
AE AY 0.0675675675676 5.0 74.0 10786.0
AE EH 0.0540540540541 4.0 74.0 10786.0
AE HH 0.0405405405405 3.0 74.0 10786.0
AE IH 0.0540540540541 4.0 74.0 10786.0
AE M 0.0540540540541 4.0 74.0 10786.0
AE OW 0.0540540540541 4.0 74.0 10786.0
AE R 0.0810810810811 6.0 74.0 10786.0
AE S 0.0675675675676 5.0 74.0 10786.0
AE T 0.0405405405405 3.0 74.0 10786.0
AH AE 0.0416666666667 6.0 144.0 10786.0
AH AH 0.0486111111111 7.0 144.0 10786.0
AH AX 0.0347222222222 5.0 144.0 10786.0
AH AY 0.0416666666667 6.0 144.0 10786.0
AH D 0.0625 9.0 144.0 10786.0
AH IY 0.0416666666667 6.0 144.0 10786.0
AH K 0.0347222222222 5.0 144.0 10786.0
AH L 0.0416666666667 6.0 144.0 10786.0
AH N 0.0833333333333 12.0 144.0 10786.0
AH S 0.0694444444444 10.0 144.0 10786.0
AH T 0.0416666666667 6.0 144.0 10786.0
AH V 0.0416666666667 6.0 144.0 10786.0
AH Z 0.0486111111111 7.0 144.0 10786.0
AO AH 0.0406504065041 5.0 123.0 10786.0
AO AX 0.0406504065041 5.0 123.0 10786.0
AO AY 0.0569105691057 7.0 123.0 10786.0
AO EH 0.0406504065041 5.0 123.0 10786.0
AO F 0.0569105691057 7.0 123.0 10786.0
AO IH 0.0406504065041 5.0 123.0 10786.0
AO IY 0.0325203252033 4.0 123.0 10786.0
AO N 0.0731707317073 9.0 123.0 10786.0
AO OW 0.0487804878049 6.0 123.0 10786.0
AO R 0.0650406504065 8.0 123.0 10786.0
AO S 0.0894308943089 11.0 123.0 10786.0
AO T 0.0650406504065 8.0 123.0 10786.0
AO V 0.0731707317073 9.0 123.0 10786.0
AO W 0.0325203252033 4.0 123.0 10786.0
AO Z 0.0325203252033 4.0 123.0 10786.0
AW AA 0.04 3.0 75.0 10786.0
AW AE 0.04 3.0 75.0 10786.0
AW AX 0.0533333333333 4.0 75.0 10786.0
AW D 0.04 3.0 75.0 10786.0
AW DH 0.04 3.0 75.0 10786.0
AW EY 0.0533333333333 4.0 75.0 10786.0
AW IH 0.04 3.0 75.0 10786.0
AW K 0.04 3.0 75.0 10786.0

AW N 0.106666666667 8.0 75.0 10786.0
AW OW 0.0666666666667 5.0 75.0 10786.0
AW R 0.04 3.0 75.0 10786.0
AW S 0.0666666666667 5.0 75.0 10786.0
AW V 0.04 3.0 75.0 10786.0
AW Y 0.04 3.0 75.0 10786.0
AX AX 0.0424242424242 7.0 165.0 10786.0
AX AY 0.0484848484848 8.0 165.0 10786.0
AX DH 0.030303030303 5.0 165.0 10786.0
AX EH 0.0545454545455 9.0 165.0 10786.0
AX EY 0.0484848484848 8.0 165.0 10786.0
AX F 0.0363636363636 6.0 165.0 10786.0
AX IX 0.030303030303 5.0 165.0 10786.0
AX K 0.030303030303 5.0 165.0 10786.0
AX N 0.0545454545455 9.0 165.0 10786.0
AX R 0.0727272727273 12.0 165.0 10786.0
AX S 0.0545454545455 9.0 165.0 10786.0
AX T 0.0666666666667 11.0 165.0 10786.0
AX UW 0.0424242424242 7.0 165.0 10786.0
AX V 0.0363636363636 6.0 165.0 10786.0
AXR AH 0.064 8.0 125.0 10786.0
AXR AY 0.048 6.0 125.0 10786.0
AXR F 0.048 6.0 125.0 10786.0
AXR IX 0.032 4.0 125.0 10786.0
AXR IY 0.088 11.0 125.0 10786.0
AXR N 0.136 17.0 125.0 10786.0
AXR OW 0.032 4.0 125.0 10786.0
AXR R 0.08 10.0 125.0 10786.0
AXR S 0.064 8.0 125.0 10786.0
AXR T 0.056 7.0 125.0 10786.0
AXR UW 0.032 4.0 125.0 10786.0
AXR V 0.04 5.0 125.0 10786.0
AXR Y 0.032 4.0 125.0 10786.0
AXR Z 0.032 4.0 125.0 10786.0
AY AH 0.046357615894 7.0 151.0 10786.0
AY AY 0.0596026490066 9.0 151.0 10786.0
AY EH 0.0529801324503 8.0 151.0 10786.0
AY EY 0.0397350993377 6.0 151.0 10786.0
AY F 0.0860927152318 13.0 151.0 10786.0
AY IY 0.0662251655629 10.0 151.0 10786.0
AY N 0.112582781457 17.0 151.0 10786.0
AY OW 0.0662251655629 10.0 151.0 10786.0
AY R 0.046357615894 7.0 151.0 10786.0
AY S 0.0662251655629 10.0 151.0 10786.0
AY T 0.0596026490066 9.0 151.0 10786.0
AY UW 0.0331125827815 5.0 151.0 10786.0
B AE 0.0392156862745 4.0 102.0 10786.0
B AY 0.0686274509804 7.0 102.0 10786.0
B EH 0.0588235294118 6.0 102.0 10786.0
B EY 0.0490196078431 5.0 102.0 10786.0
B IH 0.0392156862745 4.0 102.0 10786.0
B IY 0.0490196078431 5.0 102.0 10786.0
B M 0.0392156862745 4.0 102.0 10786.0
B N 0.078431372549 8.0 102.0 10786.0
B R 0.078431372549 8.0 102.0 10786.0
B S 0.0588235294118 6.0 102.0 10786.0
B T 0.0686274509804 7.0 102.0 10786.0
B UW 0.0392156862745 4.0 102.0 10786.0
D AE 0.047817047817 23.0 481.0 10786.0
D AY 0.033264033264 16.0 481.0 10786.0
D D 0.04158004158 20.0 481.0 10786.0
D F 0.045738045738 22.0 481.0 10786.0
D IY 0.035343035343 17.0 481.0 10786.0
D K 0.031185031185 15.0 481.0 10786.0
D L 0.035343035343 17.0 481.0 10786.0
D N 0.0790020790021 38.0 481.0 10786.0
D OW 0.047817047817 23.0 481.0 10786.0
D R 0.039501039501 19.0 481.0 10786.0
D S 0.049896049896 24.0 481.0 10786.0
D T 0.047817047817 23.0 481.0 10786.0
D UW 0.033264033264 16.0 481.0 10786.0
D W 0.031185031185 15.0 481.0 10786.0
DH AY 0.0322580645161 9.0 279.0 10786.0
DH D 0.0394265232975 11.0 279.0 10786.0

DH EH 0.0465949820789 13.0 279.0 10786.0
DH EY 0.0430107526882 12.0 279.0 10786.0
DH IH 0.0358422939068 10.0 279.0 10786.0
DH IY 0.0430107526882 12.0 279.0 10786.0
DH L 0.0537634408602 15.0 279.0 10786.0
DH N 0.0860215053763 24.0 279.0 10786.0
DH OW 0.0501792114695 14.0 279.0 10786.0
DH R 0.0681003584229 19.0 279.0 10786.0
DH S 0.0358422939068 10.0 279.0 10786.0
DH T 0.0681003584229 19.0 279.0 10786.0
DH UW 0.0322580645161 9.0 279.0 10786.0
DH V 0.0322580645161 9.0 279.0 10786.0
DX AE 0.0428571428571 3.0 70.0 10786.0
DX AH 0.0428571428571 3.0 70.0 10786.0
DX DH 0.0428571428571 3.0 70.0 10786.0
DX F 0.0428571428571 3.0 70.0 10786.0
DX IY 0.0714285714286 5.0 70.0 10786.0
DX N 0.0428571428571 3.0 70.0 10786.0
DX OW 0.0428571428571 3.0 70.0 10786.0
DX R 0.0714285714286 5.0 70.0 10786.0
DX S 0.0571428571429 4.0 70.0 10786.0
DX T 0.157142857143 11.0 70.0 10786.0
DX Y 0.0428571428571 3.0 70.0 10786.0
EH AH 0.0387596899225 5.0 129.0 10786.0
EH AY 0.031007751938 4.0 129.0 10786.0
EH D 0.093023255814 12.0 129.0 10786.0
EH DH 0.031007751938 4.0 129.0 10786.0
EH F 0.0387596899225 5.0 129.0 10786.0
EH IH 0.031007751938 4.0 129.0 10786.0
EH IY 0.062015503876 8.0 129.0 10786.0
EH K 0.031007751938 4.0 129.0 10786.0
EH L 0.0387596899225 5.0 129.0 10786.0
EH N 0.100775193798 13.0 129.0 10786.0
EH R 0.0542635658915 7.0 129.0 10786.0
EH S 0.0542635658915 7.0 129.0 10786.0
EH T 0.046511627907 6.0 129.0 10786.0
EH UW 0.031007751938 4.0 129.0 10786.0
EH W 0.031007751938 4.0 129.0 10786.0
EY AH 0.0337837837838 5.0 148.0 10786.0
EY AY 0.0608108108108 9.0 148.0 10786.0
EY EH 0.0405405405405 6.0 148.0 10786.0
EY EY 0.0472972972973 7.0 148.0 10786.0
EY F 0.0405405405405 6.0 148.0 10786.0
EY IY 0.0337837837838 5.0 148.0 10786.0
EY K 0.0608108108108 9.0 148.0 10786.0
EY M 0.0472972972973 7.0 148.0 10786.0
EY N 0.0810810810811 12.0 148.0 10786.0
EY OW 0.0337837837838 5.0 148.0 10786.0
EY R 0.0608108108108 9.0 148.0 10786.0
EY S 0.0608108108108 9.0 148.0 10786.0
EY T 0.0540540540541 8.0 148.0 10786.0
EY V 0.0405405405405 6.0 148.0 10786.0
EY W 0.0337837837838 5.0 148.0 10786.0
F AY 0.0439189189189 13.0 296.0 10786.0
F D 0.0337837837838 10.0 296.0 10786.0
F EH 0.0337837837838 10.0 296.0 10786.0
F EY 0.0337837837838 10.0 296.0 10786.0
F F 0.0540540540541 16.0 296.0 10786.0
F IY 0.0337837837838 10.0 296.0 10786.0
F K 0.0506756756757 15.0 296.0 10786.0
F N 0.0675675675676 20.0 296.0 10786.0
F OW 0.0337837837838 10.0 296.0 10786.0
F R 0.0675675675676 20.0 296.0 10786.0
F S 0.0574324324324 17.0 296.0 10786.0
F V 0.0506756756757 15.0 296.0 10786.0
F Z 0.0439189189189 13.0 296.0 10786.0
G AE 0.0666666666667 6.0 90.0 10786.0
G EY 0.0444444444444 4.0 90.0 10786.0
G IH 0.0777777777778 7.0 90.0 10786.0
G IY 0.111111111111 10.0 90.0 10786.0
G K 0.0444444444444 4.0 90.0 10786.0
G L 0.0333333333333 3.0 90.0 10786.0
G N 0.0555555555556 5.0 90.0 10786.0
G OW 0.0666666666667 6.0 90.0 10786.0

G R 0.0555555555556 5.0 90.0 10786.0
G S 0.0888888888889 8.0 90.0 10786.0
G T 0.0444444444444 4.0 90.0 10786.0
G V 0.0555555555556 5.0 90.0 10786.0
HH AX 0.0402684563758 6.0 149.0 10786.0
HH D 0.0335570469799 5.0 149.0 10786.0
HH EH 0.0402684563758 6.0 149.0 10786.0
HH IH 0.0402684563758 6.0 149.0 10786.0
HH L 0.0469798657718 7.0 149.0 10786.0
HH N 0.0939597315436 14.0 149.0 10786.0
HH OW 0.0469798657718 7.0 149.0 10786.0
HH R 0.0402684563758 6.0 149.0 10786.0
HH S 0.0671140939597 10.0 149.0 10786.0
HH T 0.0604026845638 9.0 149.0 10786.0
HH UW 0.0335570469799 5.0 149.0 10786.0
HH W 0.0335570469799 5.0 149.0 10786.0
HH Z 0.0469798657718 7.0 149.0 10786.0
IH AH 0.0425531914894 4.0 94.0 10786.0
IH AY 0.031914893617 3.0 94.0 10786.0
IH DH 0.031914893617 3.0 94.0 10786.0
IH F 0.063829787234 6.0 94.0 10786.0
IH IX 0.031914893617 3.0 94.0 10786.0
IH N 0.0957446808511 9.0 94.0 10786.0
IH NG 0.031914893617 3.0 94.0 10786.0
IH P 0.031914893617 3.0 94.0 10786.0
IH R 0.0851063829787 8.0 94.0 10786.0
IH S 0.0744680851064 7.0 94.0 10786.0
IH T 0.0531914893617 5.0 94.0 10786.0
IH UW 0.031914893617 3.0 94.0 10786.0
IH V 0.031914893617 3.0 94.0 10786.0
IH W 0.0425531914894 4.0 94.0 10786.0
IX DH 0.0363636363636 4.0 110.0 10786.0
IX F 0.0363636363636 4.0 110.0 10786.0
IX IH 0.0545454545455 6.0 110.0 10786.0
IX IX 0.0454545454545 5.0 110.0 10786.0
IX K 0.0454545454545 5.0 110.0 10786.0
IX N 0.0727272727273 8.0 110.0 10786.0
IX R 0.0909090909091 10.0 110.0 10786.0
IX S 0.0818181818182 9.0 110.0 10786.0
IX T 0.0909090909091 10.0 110.0 10786.0
IY AH 0.0392857142857 11.0 280.0 10786.0
IY AX 0.0392857142857 11.0 280.0 10786.0
IY AY 0.0321428571429 9.0 280.0 10786.0
IY EY 0.0428571428571 12.0 280.0 10786.0
IY F 0.0464285714286 13.0 280.0 10786.0
IY IH 0.0357142857143 10.0 280.0 10786.0
IY IY 0.0428571428571 12.0 280.0 10786.0
IY L 0.0571428571429 16.0 280.0 10786.0
IY N 0.0678571428571 19.0 280.0 10786.0
IY OW 0.0321428571429 9.0 280.0 10786.0
IY R 0.075 21.0 280.0 10786.0
IY S 0.0571428571429 16.0 280.0 10786.0
IY T 0.0392857142857 11.0 280.0 10786.0
IY V 0.0357142857143 10.0 280.0 10786.0
K AH 0.0359712230216 10.0 278.0 10786.0
K AX 0.0431654676259 12.0 278.0 10786.0
K AY 0.0467625899281 13.0 278.0 10786.0
K EY 0.0323741007194 9.0 278.0 10786.0
K IH 0.0359712230216 10.0 278.0 10786.0
K IY 0.0467625899281 13.0 278.0 10786.0
K K 0.068345323741 19.0 278.0 10786.0
K L 0.0323741007194 9.0 278.0 10786.0
K M 0.0359712230216 10.0 278.0 10786.0
K N 0.0827338129496 23.0 278.0 10786.0
K R 0.0539568345324 15.0 278.0 10786.0
K S 0.0431654676259 12.0 278.0 10786.0
K T 0.0755395683453 21.0 278.0 10786.0
K UW 0.0395683453237 11.0 278.0 10786.0
K W 0.0323741007194 9.0 278.0 10786.0
K Z 0.0323741007194 9.0 278.0 10786.0
L EH 0.0325732899023 10.0 307.0 10786.0
L F 0.0521172638436 16.0 307.0 10786.0
L IY 0.0553745928339 17.0 307.0 10786.0
L L 0.0325732899023 10.0 307.0 10786.0

L N 0.0879478827362 27.0 307.0 10786.0
L OW 0.0586319218241 18.0 307.0 10786.0
L R 0.0846905537459 26.0 307.0 10786.0
L T 0.0618892508143 19.0 307.0 10786.0
L UW 0.0488599348534 15.0 307.0 10786.0
L V 0.0325732899023 10.0 307.0 10786.0
L W 0.0358306188925 11.0 307.0 10786.0
L Z 0.0358306188925 11.0 307.0 10786.0
M AA 0.0441767068273 11.0 249.0 10786.0
M AE 0.0361445783133 9.0 249.0 10786.0
M AY 0.0321285140562 8.0 249.0 10786.0
M F 0.0441767068273 11.0 249.0 10786.0
M IY 0.0441767068273 11.0 249.0 10786.0
M K 0.0441767068273 11.0 249.0 10786.0
M L 0.0321285140562 8.0 249.0 10786.0
M M 0.0361445783133 9.0 249.0 10786.0
M N 0.0722891566265 18.0 249.0 10786.0
M OW 0.0401606425703 10.0 249.0 10786.0
M R 0.0401606425703 10.0 249.0 10786.0
M S 0.0522088353414 13.0 249.0 10786.0
M T 0.0562248995984 14.0 249.0 10786.0
M W 0.0321285140562 8.0 249.0 10786.0
N AE 0.0408805031447 13.0 318.0 10786.0
N AY 0.0503144654088 16.0 318.0 10786.0
N EH 0.0408805031447 13.0 318.0 10786.0
N EY 0.0440251572327 14.0 318.0 10786.0
N IY 0.0534591194969 17.0 318.0 10786.0
N K 0.0314465408805 10.0 318.0 10786.0
N L 0.0345911949686 11.0 318.0 10786.0
N N 0.0723270440252 23.0 318.0 10786.0
N R 0.0566037735849 18.0 318.0 10786.0
N T 0.0691823899371 22.0 318.0 10786.0
N V 0.0503144654088 16.0 318.0 10786.0
N W 0.0314465408805 10.0 318.0 10786.0
NG AY 0.04 3.0 75.0 10786.0
NG F 0.0533333333333 4.0 75.0 10786.0
NG IH 0.0533333333333 4.0 75.0 10786.0
NG IY 0.0533333333333 4.0 75.0 10786.0
NG N 0.146666666667 11.0 75.0 10786.0
NG R 0.08 6.0 75.0 10786.0
NG T 0.0666666666667 5.0 75.0 10786.0
NG TH 0.04 3.0 75.0 10786.0
NG V 0.0533333333333 4.0 75.0 10786.0
NG W 0.0533333333333 4.0 75.0 10786.0
OW AY 0.030303030303 4.0 132.0 10786.0
OW B 0.030303030303 4.0 132.0 10786.0
OW EH 0.0454545454545 6.0 132.0 10786.0
OW EY 0.030303030303 4.0 132.0 10786.0
OW F 0.030303030303 4.0 132.0 10786.0
OW IH 0.0378787878788 5.0 132.0 10786.0
OW IY 0.0530303030303 7.0 132.0 10786.0
OW K 0.0530303030303 7.0 132.0 10786.0
OW L 0.0378787878788 5.0 132.0 10786.0
OW N 0.0757575757576 10.0 132.0 10786.0
OW OW 0.0530303030303 7.0 132.0 10786.0
OW R 0.0454545454545 6.0 132.0 10786.0
OW S 0.0606060606061 8.0 132.0 10786.0
OW T 0.0681818181818 9.0 132.0 10786.0
OW UW 0.0378787878788 5.0 132.0 10786.0
OW V 0.0530303030303 7.0 132.0 10786.0
OW W 0.0530303030303 7.0 132.0 10786.0
P EH 0.0310734463277 11.0 354.0 10786.0
P EY 0.0310734463277 11.0 354.0 10786.0
P F 0.0423728813559 15.0 354.0 10786.0
P IH 0.0480225988701 17.0 354.0 10786.0
P IY 0.0508474576271 18.0 354.0 10786.0
P K 0.0395480225989 14.0 354.0 10786.0
P N 0.0988700564972 35.0 354.0 10786.0
P OW 0.0508474576271 18.0 354.0 10786.0
P R 0.0649717514124 23.0 354.0 10786.0
P S 0.0734463276836 26.0 354.0 10786.0
P T 0.0932203389831 33.0 354.0 10786.0
P V 0.0423728813559 15.0 354.0 10786.0
P W 0.0338983050847 12.0 354.0 10786.0

R AA 0.0441176470588 6.0 136.0 10786.0
R AH 0.0367647058824 5.0 136.0 10786.0
R AY 0.0441176470588 6.0 136.0 10786.0
R D 0.0441176470588 6.0 136.0 10786.0
R IH 0.0661764705882 9.0 136.0 10786.0
R K 0.0514705882353 7.0 136.0 10786.0
R L 0.0367647058824 5.0 136.0 10786.0
R N 0.102941176471 14.0 136.0 10786.0
R OW 0.0367647058824 5.0 136.0 10786.0
R W 0.0735294117647 10.0 136.0 10786.0
S AH 0.0324074074074 14.0 432.0 10786.0
S AY 0.0324074074074 14.0 432.0 10786.0
S F 0.037037037037 16.0 432.0 10786.0
S IH 0.0462962962963 20.0 432.0 10786.0
S IY 0.0578703703704 25.0 432.0 10786.0
S K 0.037037037037 16.0 432.0 10786.0
S L 0.0462962962963 20.0 432.0 10786.0
S N 0.0787037037037 34.0 432.0 10786.0
S R 0.0509259259259 22.0 432.0 10786.0
S S 0.0578703703704 25.0 432.0 10786.0
S T 0.0671296296296 29.0 432.0 10786.0
S TH 0.0324074074074 14.0 432.0 10786.0
S V 0.0324074074074 14.0 432.0 10786.0
T AY 0.034155597723 18.0 527.0 10786.0
T F 0.0303605313093 16.0 527.0 10786.0
T IX 0.0436432637571 23.0 527.0 10786.0
T IY 0.034155597723 18.0 527.0 10786.0
T K 0.0303605313093 16.0 527.0 10786.0
T N 0.0891840607211 47.0 527.0 10786.0
T OW 0.0417457305503 22.0 527.0 10786.0
T R 0.0588235294118 31.0 527.0 10786.0
T S 0.0645161290323 34.0 527.0 10786.0
T T 0.0664136622391 35.0 527.0 10786.0
T UW 0.0322580645161 17.0 527.0 10786.0
T W 0.0303605313093 16.0 527.0 10786.0
T Z 0.034155597723 18.0 527.0 10786.0
TH AE 0.044776119403 9.0 201.0 10786.0
TH AH 0.0398009950249 8.0 201.0 10786.0
TH EH 0.0398009950249 8.0 201.0 10786.0
TH EY 0.044776119403 9.0 201.0 10786.0
TH IY 0.044776119403 9.0 201.0 10786.0
TH L 0.044776119403 9.0 201.0 10786.0
TH N 0.0746268656716 15.0 201.0 10786.0
TH R 0.0497512437811 10.0 201.0 10786.0
TH T 0.0547263681592 11.0 201.0 10786.0
TH UW 0.0497512437811 10.0 201.0 10786.0
UH ER 0.0416666666667 3.0 72.0 10786.0
UH IY 0.0555555555556 4.0 72.0 10786.0
UH L 0.0416666666667 3.0 72.0 10786.0
UH N 0.111111111111 8.0 72.0 10786.0
UH R 0.0694444444444 5.0 72.0 10786.0
UH S 0.0833333333333 6.0 72.0 10786.0
UH T 0.0972222222222 7.0 72.0 10786.0
UH TH 0.0416666666667 3.0 72.0 10786.0
UW AY 0.0502092050209 12.0 239.0 10786.0
UW EH 0.0334728033473 8.0 239.0 10786.0
UW EY 0.0376569037657 9.0 239.0 10786.0
UW IY 0.0334728033473 8.0 239.0 10786.0
UW K 0.0627615062762 15.0 239.0 10786.0
UW L 0.0502092050209 12.0 239.0 10786.0
UW N 0.129707112971 31.0 239.0 10786.0
UW OW 0.0460251046025 11.0 239.0 10786.0
UW R 0.0502092050209 12.0 239.0 10786.0
UW S 0.0502092050209 12.0 239.0 10786.0
UW T 0.0543933054393 13,0 239.0 10786.0
UW UW 0.0376569037657 9.0 239.0 10786.0
UW W 0.0334728033473 8.0 239.0 10786.0
UW Z 0.0334728033473 8.0 239.0 10786.0
V AY 0.0433734939759 18.0 415.0 10786.0
V F 0.0457831325301 19.0 415.0 10786.0
V IH 0.033734939759 14.0 415.0 10786.0
V IX 0.0313253012048 13.0 415.0 10786.0
V K 0.0457831325301 19.0 415.0 10786.0
V L 0.0385542168675 16.0 415.0 10786.0

V N 0.0819277108434 34.0 415.0 10786.0
V OW 0.0361445783133 15.0 415.0 10786.0
V R 0.0409638554217 17.0 415.0 10786.0
V S 0.0771084337349 32.0 415.0 10786.0
V T 0.0506024096386 21.0 415.0 10786.0
V V 0.0433734939759 18.0 415.0 10786.0
V W 0.0385542168675 16.0 415.0 10786.0
W AH 0.0431034482759 5.0 116.0 10786.0
W AX 0.0344827586207 4.0 116.0 10786.0
W AY 0.0689655172414 8.0 116.0 10786.0
W EH 0.0344827586207 4.0 116.0 10786.0
W ER 0.0431034482759 5.0 116.0 10786.0
W EY 0.051724137931 6.0 116.0 10786.0
W IY 0.0344827586207 4.0 116.0 10786.0
W M 0.0344827586207 4.0 116.0 10786.0
W N 0.051724137931 6.0 116.0 10786.0
W OW 0.0344827586207 4.0 116.0 10786.0
W R 0.0775862068966 9.0 116.0 10786.0
W S 0.0689655172414 8.0 116.0 10786.0
W T 0.0948275862069 11.0 116.0 10786.0
W UW 0.0344827586207 4.0 116.0 10786.0
W W 0.0344827586207 4.0 116.0 10786.0
Y AY 0.0545454545455 6.0 110.0 10786.0
Y D 0.0363636363636 4.0 110.0 10786.0
Y EH 0.0454545454545 5.0 110.0 10786.0
Y F 0.0363636363636 4.0 110.0 10786.0
Y IY 0.0545454545455 6.0 110.0 10786.0
Y M 0.0545454545455 6.0 110.0 10786.0
Y N 0.109090909091 12.0 110.0 10786.0
Y R 0.0727272727273 8.0 110.0 10786.0
Y S 0.0363636363636 4.0 110.0 10786.0
Y UW 0.0454545454545 5.0 110.0 10786.0
Z AXR 0.0336134453782 4.0 119.0 10786.0
Z IY 0.0588235294118 7.0 119.0 10786.0
Z L 0.0504201680672 6.0 119.0 10786.0
Z N 0.0840336134454 10.0 119.0 10786.0
Z OW 0.0420168067227 5.0 119.0 10786.0
Z P 0.0336134453782 4.0 119.0 10786.0
Z R 0.0588235294118 7.0 119.0 10786.0
Z S 0.0420168067227 5.0 119.0 10786.0
Z T 0.0504201680672 6.0 119.0 10786.0
Z UW 0.0420168067227 5.0 119.0 10786.0
Z V 0.0336134453782 4.0 119.0 10786.0
Z W 0.0504201680672 6.0 119.0 10786.0
Z Z 0.0336134453782 4.0 119.0 10786.0
a D 0.0649350649351 25.0 385.0 10786.0
a EY 0.0415584415584 16.0 385.0 10786.0
a IH 0.0363636363636 14.0 385.0 10786.0
a IY 0.0415584415584 16.0 385.0 10786.0
a K 0.0519480519481 20.0 385.0 10786.0
a L 0.0519480519481 20.0 385.0 10786.0
a M 0.0311688311688 12.0 385.0 10786.0
a N 0.0753246753247 29.0 385.0 10786.0
a R 0.0311688311688 12.0 385.0 10786.0
a S 0.0571428571429 22.0 385.0 10786.0
a T 0.0649350649351 25.0 385.0 10786.0
e AE 0.0342105263158 13.0 380.0 10786.0
e AX 0.0342105263158 13.0 380.0 10786.0
e D 0.0394736842105 15.0 380.0 10786.0
e EH 0.0342105263158 13.0 380.0 10786.0
e EY 0.0473684210526 18.0 380.0 10786.0
e IY 0.0368421052632 14.0 380.0 10786.0
e K 0.0447368421053 17.0 380.0 10786.0
e L 0.0315789473684 12.0 380.0 10786.0
e N 0.0736842105263 28.0 380.0 10786.0
e R 0.0315789473684 12.0 380.0 10786.0
e S 0.0552631578947 21.0 380.0 10786.0
e T 0.0631578947368 24.0 380.0 10786.0
e UW 0.0421052631579 16.0 380.0 10786.0
f AH 0.0714285714286 5.0 70.0 10786.0
f EH 0.0571428571429 4.0 70.0 10786.0
f IY 0.0571428571429 4.0 70.0 10786.0
f L 0.0714285714286 5.0 70.0 10786.0
f N 0.114285714286 8.0 70.0 10786.0

f R 0.0714285714286 5.0 70.0 10786.0
f S 0.0571428571429 4.0 70.0 10786.0
f T 0.0714285714286 5.0 70.0 10786.0
f TH 0.0428571428571 3.0 70.0 10786.0
f V 0.0571428571429 4.0 70.0 10786.0
f Z 0.0428571428571 3.0 70.0 10786.0
h DH 0.0326530612245 8.0 245.0 10786.0
h IY 0.0530612244898 13.0 245.0 10786.0
h K 0.0367346938776 9.0 245.0 10786.0
h L 0.0448979591837 11.0 245.0 10786.0
h M 0.0489795918367 12.0 245.0 10786.0
h N 0.0775510204082 19.0 245.0 10786.0
h OW 0.0408163265306 10.0 245.0 10786.0
h R 0.0530612244898 13.0 245.0 10786.0
h S 0.0408163265306 10.0 245.0 10786.0
h T 0.065306122449 16.0 245.0 10786.0
h UW 0.065306122449 16.0 245.0 10786.0
h W 0.0408163265306 10.0 245.0 10786.0
i AE 0.030303030303 8.0 264.0 10786.0
i AX 0.030303030303 8.0 264.0 10786.0
i D 0.0530303030303 14.0 264.0 10786.0
i EH 0.0340909090909 9.0 264.0 10786.0
i F 0.0378787878788 10.0 264.0 10786.0
i IH 0.0340909090909 9.0 264.0 10786.0
i IY 0.0416666666667 11.0 264.0 10786.0
i K 0.0340909090909 9.0 264.0 10786.0
i L 0.0378787878788 10.0 264.0 10786.0
i M 0.030303030303 8.0 264.0 10786.0
i N 0.102272727273 27.0 264.0 10786.0
i R 0.0416666666667 11.0 264.0 10786.0
i S 0.0378787878788 10.0 264.0 10786.0
i T 0.0719696969697 19.0 264.0 10786.0
i UW 0.0378787878788 10.0 264.0 10786.0
i Y 0.0340909090909 9.0 264.0 10786.0
i Z 0.0530303030303 14.0 264.0 10786.0
k AE 0.03 6.0 200.0 10786.0
k AH 0.05 10.0 200.0 10786.0
k D 0.035 7.0 200.0 10786.0
k EY 0.045 9.0 200.0 10786.0
k F 0.06 12.0 200.0 10786.0
k IX 0.035 7.0 200.0 10786.0
k IY 0.08 16.0 200.0 10786.0
k K 0.06 12.0 200.0 10786.0
k L 0.035 7.0 200.0 10786.0
k N 0.03 6.0 200.0 10786.0
k R 0.055 11.0 200.0 10786.0
k S 0.04 8.0 200.0 10786.0
k T 0.07 14.0 200.0 10786.0
k UW 0.04 8.0 200.0 10786.0
k W 0.035 7.0 200.0 10786.0
k Z 0.04 8.0 200.0 10786.0
n D 0.0462633451957 13.0 281.0 10786.0
n EY 0.0533807829181 15.0 281.0 10786.0
n IH 0.0498220640569 14.0 281.0 10786.0
n K 0.067615658363 19.0 281.0 10786.0
n L 0.0355871886121 10.0 281.0 10786.0
n N 0.0640569395018 18.0 281.0 10786.0
n OW 0.0462633451957 13.0 281.0 10786.0
n R 0.0355871886121 10.0 281.0 10786.0
n S 0.0355871886121 10.0 281.0 10786.0
n T 0.0747330960854 21.0 281.0 10786.0
n UW 0.0462633451957 13.0 281.0 10786.0
n W 0.0391459074733 11.0 281.0 10786.0
n Z 0.0391459074733 11.0 281.0 10786.0
o AH 0.030985915493 11.0 355.0 10786.0
o D 0.030985915493 11.0 355.0 10786.0
o EH 0.0422535211268 15.0 355.0 10786.0
o EY 0.030985915493 11.0 355.0 10786.0
o IY 0.0422535211268 15.0 355.0 10786.0
o K 0.0394366197183 14.0 355.0 10786.0
o L 0.0535211267606 19.0 355.0 10786.0
o N 0.0676056338028 24.0 355.0 10786.0
o OW 0.0422535211268 15.0 355.0 10786.0
o R 0.0676056338028 24.0 355.0 10786.0

o S 0.056338028169 20.0 355.0 10786.0
o T 0.0507042253521 18.0 355.0 10786.0
o UW 0.0394366197183 14.0 355.0 10786.0
o W 0.0338028169014 12.0 355.0 10786.0
r AE 0.0303951367781 10.0 329.0 10786.0
r AXR 0.0334346504559 11.0 329.0 10786.0
r AY 0.0303951367781 10.0 329.0 10786.0
r EH 0.0395136778116 13.0 329.0 10786.0
r EY 0.048632218845 16.0 329.0 10786.0
r IH 0.0334346504559 11.0 329.0 10786.0
r IY 0.0425531914894 14.0 329.0 10786.0
r K 0.0547112462006 18.0 329.0 10786.0
r L 0.0334346504559 11.0 329.0 10786.0
r N 0.063829787234 21.0 329.0 10786.0
r R 0.048632218845 16.0 329.0 10786.0
r S 0.0607902735562 20.0 329.0 10786.0
r T 0.063829787234 21.0 329.0 10786.0
r UW 0.0334346504559 11.0 329.0 10786.0
sh AE 0.03 3.0 100.0 10786.0
sh AH 0.04 4.0 100.0 10786.0
sh AX 0.03 3.0 100.0 10786.0
sh ER 0.03 3.0 100.0 10786.0
sh EY 0.04 4.0 100.0 10786.0
sh IH 0.08 8.0 100.0 10786.0
sh IX 0.04 4.0 100.0 10786.0
sh K 0.05 5.0 100.0 10786.0
sh L 0.03 3.0 100.0 10786.0
sh N 0.08 8.0 100.0 10786.0
sh NG 0.04 4.0 100.0 10786.0
sh OW 0.05 5.0 100.0 10786.0
sh R 0.06 6.0 100.0 10786.0
sh S 0.04 4.0 100.0 10786.0
sh T 0.03 3.0 100.0 10786.0
sh UW 0.04 4.0 100.0 10786.0
sh V 0.03 3.0 100.0 10786.0
sh W 0.05 5.0 100.0 10786.0
u AE 0.0357142857143 14.0 392.0 10786.0
u AH 0.0357142857143 14.0 392.0 10786.0
u D 0.030612244898 12.0 392.0 10786.0
u EH 0.0382653061224 15.0 392.0 10786.0
u ER 0.0331632653061 13.0 392.0 10786.0
u EY 0.030612244898 12.0 392.0 10786.0
u IH 0.0331632653061 13.0 392.0 10786.0
u K 0.0382653061224 15.0 392.0 10786.0
u L 0.0459183673469 18.0 392.0 10786.0
u M 0.0408163265306 16.0 392.0 10786.0
u N 0.0892857142857 35.0 392.0 10786.0
u OW 0.0433673469388 17.0 392.0 10786.0
u R 0.0433673469388 17.0 392.0 10786.0
u S 0.0382653061224 15.0 392.0 10786.0
u T 0.0408163265306 16.0 392.0 10786.0
u UW 0.0459183673469 18.0 392.0 10786.0
u W 0.0331632653061 13.0 392.0 10786.0
w D 0.0441988950276 8.0 181.0 10786.0
w IY 0.0386740331492 7.0 181.0 10786.0
w K 0.060773480663 11.0 181.0 10786.0
w L 0.0331491712707 6.0 181.0 10786.0
w M 0.0386740331492 7.0 181.0 10786.0
w N 0.0773480662983 14.0 181.0 10786.0
w OW 0.0331491712707 6.0 181.0 10786.0
w P 0.0331491712707 6.0 181.0 10786.0
w R 0.060773480663 11.0 181.0 10786.0
w S 0.0497237569061 9.0 181.0 10786.0
w T 0.0773480662983 14.0 181.0 10786.0
w UW 0.0331491712707 6.0 181.0 10786.0
w W 0.0386740331492 7.0 181.0 10786.0
w Y 0.0331491712707 6.0 181.0 10786.0
zh AX 0.0444444444444 8.0 180.0 10786.0
zh EY 0.0333333333333 6.0 180.0 10786.0
zh F 0.0611111111111 11.0 180.0 10786.0
zh IH 0.0333333333333 6.0 180.0 10786.0
zh IY 0.0444444444444 8.0 180.0 10786.0
zh K 0.0333333333333 6.0 180.0 10786.0
zh L 0.0444444444444 8.0 180.0 10786.0

zh N 0.0555555555556 10.0 180.0 10786.0
zh OW 0.0388888888889 7.0 180.0 10786.0
zh R 0.0777777777778 14.0 180.0 10786.0
zh S 0.0444444444444 8.0 180.0 10786.0
zh T 0.0555555555556 10.0 180.0 10786.0
zh W 0.0611111111111 11.0 180.0 10786.0
zh Y 0.0333333333333 6.0 180.0 10786.0
zh Z 0.0333333333333 6.0 180.0 10786.0
phonemes:
OW w UH u r n o k h i f IH e a Z AX Y V W T EY R S P N AO L M K AH IY F G D AE B IX AA TH AY UW NG AW EH sh DX HH DH zh AXR

RULES

SP: 0 (0.000000)

SIL: 0 (0.000000)

P(ph) < 0.005000: 68 (0.036997)

P(var) < 0.030000: 1112 (0.605005)

rules: 1838

accepted rules: 658 (0.000000)

added jap. rules: 0 (0.000000)

(B.2.2) Rules for an Occurance Probability > 7%

AA N 0.0896551724138 13.0 145.0 10786.0
AA S 0.0758620689655 11.0 145.0 10786.0
AE R 0.0810810810811 6.0 74.0 10786.0
AH N 0.0833333333333 12.0 144.0 10786.0
AO N 0.0731707317073 9.0 123.0 10786.0
AO S 0.0894308943089 11.0 123.0 10786.0
AO V 0.0731707317073 9.0 123.0 10786.0
AW N 0.106666666667 8.0 75.0 10786.0
AX R 0.0727272727273 12.0 165.0 10786.0
AXR IY 0.088 11.0 125.0 10786.0
AXR N 0.136 17.0 125.0 10786.0
AXR R 0.08 10.0 125.0 10786.0
AY F 0.0860927152318 13.0 151.0 10786.0
AY N 0.112582781457 17.0 151.0 10786.0
B N 0.078431372549 8.0 102.0 10786.0
B R 0.078431372549 8.0 102.0 10786.0
D N 0.0790020790021 38.0 481.0 10786.0
DH N 0.0860215053763 24.0 279.0 10786.0
DX IY 0.0714285714286 5.0 70.0 10786.0
DX R 0.0714285714286 5.0 70.0 10786.0
DX T 0.157142857143 11.0 70.0 10786.0
EH D 0.093023255814 12.0 129.0 10786.0
EH N 0.100775193798 13.0 129.0 10786.0
EY N 0.0810810810811 12.0 148.0 10786.0
G IH 0.0777777777778 7.0 90.0 10786.0
G IY 0.111111111111 10.0 90.0 10786.0
G S 0.0888888888889 8.0 90.0 10786.0
HH N 0.0939597315436 14.0 149.0 10786.0
IH N 0.0957446808511 9.0 94.0 10786.0
IH R 0.0851063829787 8.0 94.0 10786.0
IH S 0.0744680851064 7.0 94.0 10786.0
IX N 0.0727272727273 8.0 110.0 10786.0
IX R 0.0909090909091 10.0 110.0 10786.0
IX S 0.0818181818182 9.0 110.0 10786.0
IX T 0.0909090909091 10.0 110.0 10786.0
IY R 0.075 21.0 280.0 10786.0
K N 0.0827338129496 23.0 278.0 10786.0
K T 0.0755395683453 21.0 278.0 10786.0
L N 0.0879478827362 27.0 307.0 10786.0
L R 0.0846905537459 26.0 307.0 10786.0
M N 0.0722891566265 18.0 249.0 10786.0
N N 0.0723270440252 23.0 318.0 10786.0
NG N 0.146666666667 11.0 75.0 10786.0
NG R 0.08 6.0 75.0 10786.0

OW N 0.0757575757576 10.0 132.0 10786.0
P N 0.0988700564972 35.0 354.0 10786.0
P S 0.0734463276836 26.0 354.0 10786.0
P T 0.0932203389831 33.0 354.0 10786.0
R N 0.102941176471 14.0 136.0 10786.0
R W 0.0735294117647 10.0 136.0 10786.0
S N 0.0787037037037 34.0 432.0 10786.0
T N 0.0891840607211 47.0 527.0 10786.0
TH N 0.0746268656716 15.0 201.0 10786.0
UH N 0.111111111111 8.0 72.0 10786.0
UH S 0.0833333333333 6.0 72.0 10786.0
UH T 0.0972222222222 7.0 72.0 10786.0
UW N 0.129707112971 31.0 239.0 10786.0
V N 0.0819277108434 34.0 415.0 10786.0
V S 0.0771084337349 32.0 415.0 10786.0
W R 0.0775862068966 9.0 116.0 10786.0
W T 0.0948275862069 11.0 116.0 10786.0
Y N 0.109090909091 12.0 110.0 10786.0
Y R 0.0727272727273 8.0 110.0 10786.0
Z N 0.0840336134454 10.0 119.0 10786.0
a N 0.0753246753247 29.0 385.0 10786.0
e N 0.0736842105263 28.0 380.0 10786.0
f AH 0.0714285714286 5.0 70.0 10786.0
f L 0.0714285714286 5.0 70.0 10786.0
f N 0.114285714286 8.0 70.0 10786.0
f R 0.0714285714286 5.0 70.0 10786.0
f T 0.0714285714286 5.0 70.0 10786.0
h N 0.0775510204082 19.0 245.0 10786.0
i N 0.102272727273 27.0 264.0 10786.0
i T 0.0719696969697 19.0 264.0 10786.0
k IY 0.08 16.0 200.0 10786.0
k T 0.07 14.0 200.0 10786.0
n T 0.0747330960854 21.0 281.0 10786.0
sh IH 0.08 8.0 100.0 10786.0
sh N 0.08 8.0 100.0 10786.0
u N 0.0892857142857 35.0 392.0 10786.0
w N 0.0773480662983 14.0 181.0 10786.0
w T 0.0773480662983 14.0 181.0 10786.0
zh R 0.0777777777778 14.0 180.0 10786.0
phonemes:
OW w UH u sh n k h i f IH e a Z AX Y V W T EY R S P N AO L M K AH IY G D AE B IX AA TH AY UW NG AW EH DX HH DH zh AXR
added o N 0.0676056338028 24.0 355.0 10786.0
added r N 0.063829787234 21.0 329.0 10786.0

RULES

SP: 0 (0.000000)

SIL: 0 (0.000000)

P(ph) < 0.005000: 12 (0.006529)

P(var) < 0.070000: 1743 (0.948313)

rules: 1838

accepted rules: 83 (0.000000)

added jap. rules: 2 (0.001088)

(B.2.3) Rules for an Occurance Probability > 8%

AA N 0.0896551724138 13.0 145.0 10786.0
AE R 0.0810810810811 6.0 74.0 10786.0
AH N 0.0833333333333 12.0 144.0 10786.0
AO S 0.0894308943089 11.0 123.0 10786.0
AW N 0.106666666667 8.0 75.0 10786.0
AXR IY 0.088 11.0 125.0 10786.0
AXR N 0.136 17.0 125.0 10786.0
AXR R 0.08 10.0 125.0 10786.0
AY F 0.0860927152318 13.0 151.0 10786.0
AY N 0.112582781457 17.0 151.0 10786.0
DH N 0.0860215053763 24.0 279.0 10786.0

DX T 0.157142857143 11.0 70.0 10786.0
EH D 0.093023255814 12.0 129.0 10786.0
EH N 0.100775193798 13.0 129.0 10786.0
EY N 0.0810810810811 12.0 148.0 10786.0
G IY 0.111111111111 10.0 90.0 10786.0
G S 0.0888888888889 8.0 90.0 10786.0
HH N 0.0939597315436 14.0 149.0 10786.0
IH N 0.0957446808511 9.0 94.0 10786.0
IH R 0.0851063829787 8.0 94.0 10786.0
IX R 0.0909090909091 10.0 110.0 10786.0
IX S 0.0818181818182 9.0 110.0 10786.0
IX T 0.0909090909091 10.0 110.0 10786.0
K N 0.0827338129496 23.0 278.0 10786.0
L N 0.0879478827362 27.0 307.0 10786.0
L R 0.0846905537459 26.0 307.0 10786.0
NG N 0.146666666667 11.0 75.0 10786.0
NG R 0.08 6.0 75.0 10786.0
P N 0.0988700564972 35.0 354.0 10786.0
P T 0.0932203389831 33.0 354.0 10786.0
R N 0.102941176471 14.0 136.0 10786.0
T N 0.0891840607211 47.0 527.0 10786.0
UH N 0.111111111111 8.0 72.0 10786.0
UH S 0.0833333333333 6.0 72.0 10786.0
UH T 0.0972222222222 7.0 72.0 10786.0
UW N 0.129707112971 31.0 239.0 10786.0
V N 0.0819277108434 34.0 415.0 10786.0
W T 0.0948275862069 11.0 116.0 10786.0
Y N 0.109090909091 12.0 110.0 10786.0
Z N 0.0840336134454 10.0 119.0 10786.0
f N 0.114285714286 8.0 70.0 10786.0
i N 0.102272727273 27.0 264.0 10786.0
k IY 0.08 16.0 200.0 10786.0
sh IH 0.08 8.0 100.0 10786.0
sh N 0.08 8.0 100.0 10786.0
u N 0.0892857142857 35.0 392.0 10786.0
phonemes:
W UH u k DX Y NG UW i f IH Z sh AY V AW T EY R HH P DH AO L K AH AA G EH AE IX AXR
added a N 0.0753246753247 29.0 385.0 10786.0
added o N 0.0676056338028 24.0 355.0 10786.0
added zh R 0.0777777777778 14.0 180.0 10786.0
added r N 0.063829787234 21.0 329.0 10786.0
added e N 0.0736842105263 28.0 380.0 10786.0
added w N 0.0773480662983 14.0 181.0 10786.0
added n T 0.0747330960854 21.0 281.0 10786.0
added h N 0.0775510204082 19.0 245.0 10786.0

RULES

SP: 0 (0.000000)

SIL: 0 (0.000000)

P(ph) < 0.005000: 12 (0.006529)

P(var) < 0.080000: 1780 (0.968444)

rules: 1838

accepted rules: 46 (0.000000)

added jap. rules: 8 (0.004353)

(B.2.4) Rules for an Occurance Probability > 10%

AA N 0.0896551724138 13.0 145.0 10786.0
AE R 0.0810810810811 6.0 74.0 10786.0
AH N 0.0833333333333 12.0 144.0 10786.0
AO S 0.0894308943089 11.0 123.0 10786.0
AW N 0.106666666667 8.0 75.0 10786.0
AXR IY 0.088 11.0 125.0 10786.0
AXR N 0.136 17.0 125.0 10786.0
AXR R 0.08 10.0 125.0 10786.0
AY F 0.0860927152318 13.0 151.0 10786.0
AY N 0.112582781457 17.0 151.0 10786.0

DH N 0.0860215053763 24.0 279.0 10786.0
DX T 0.157142857143 11.0 70.0 10786.0
EH D 0.093023255814 12.0 129.0 10786.0
EH N 0.100775193798 13.0 129.0 10786.0
EY N 0.0810810810811 12.0 148.0 10786.0
G IY 0.111111111111 10.0 90.0 10786.0
G S 0.0888888888889 8.0 90.0 10786.0
HH N 0.0939597315436 14.0 149.0 10786.0
IH N 0.0957446808511 9.0 94.0 10786.0
IH R 0.0851063829787 8.0 94.0 10786.0
IX R 0.0909090909091 10.0 110.0 10786.0
IX S 0.0818181818182 9.0 110.0 10786.0
IX T 0.0909090909091 10.0 110.0 10786.0
K N 0.0827338129496 23.0 278.0 10786.0
L N 0.0879478827362 27.0 307.0 10786.0
L R 0.0846905537459 26.0 307.0 10786.0
NG N 0.146666666667 11.0 75.0 10786.0
NG R 0.08 6.0 75.0 10786.0
P N 0.0988700564972 35.0 354.0 10786.0
P T 0.0932203389831 33.0 354.0 10786.0
R N 0.102941176471 14.0 136.0 10786.0
T N 0.0891840607211 47.0 527.0 10786.0
UH N 0.111111111111 8.0 72.0 10786.0
UH S 0.0833333333333 6.0 72.0 10786.0
UH T 0.0972222222222 7.0 72.0 10786.0
UW N 0.129707112971 31.0 239.0 10786.0
V N 0.0819277108434 34.0 415.0 10786.0
W T 0.0948275862069 11.0 116.0 10786.0
Y N 0.109090909091 12.0 110.0 10786.0
Z N 0.0840336134454 10.0 119.0 10786.0
f N 0.114285714286 8.0 70.0 10786.0
i N 0.102272727273 27.0 264.0 10786.0
k IY 0.08 16.0 200.0 10786.0
sh IH 0.08 8.0 100.0 10786.0
sh N 0.08 8.0 100.0 10786.0
u N 0.0892857142857 35.0 392.0 10786.0
phonemes:
W UH u k DX Y NG UW i f IH Z sh AY V AW T EY R HH P DH AO L K AH AA G EH AE IX AXR
added a N 0.0753246753247 29.0 385.0 10786.0
added o N 0.0676056338028 24.0 355.0 10786.0
added zh R 0.0777777777778 14.0 180.0 10786.0
added r N 0.063829787234 21.0 329.0 10786.0
added e N 0.0736842105263 28.0 380.0 10786.0
added w N 0.0773480662983 14.0 181.0 10786.0
added n T 0.0747330960854 21.0 281.0 10786.0
added h N 0.0775510204082 19.0 245.0 10786.0

RULES

SP: 0 (0.000000)

SIL: 0 (0.000000)

P(ph) < 0.005000: 12 (0.006529)

P(var) < 0.080000: 1780 (0.968444)

rules: 1838

accepted rules: 46 (0.000000)

added jap. rules: 8 (0.004353)

(B.2.5) Rules for an Occurance Probability > 11%

AXR N 0.136 17.0 125.0 10786.0
AY N 0.112582781457 17.0 151.0 10786.0
DX T 0.157142857143 11.0 70.0 10786.0
G IY 0.111111111111 10.0 90.0 10786.0
NG N 0.146666666667 11.0 75.0 10786.0
UH N 0.111111111111 8.0 72.0 10786.0
UW N 0.129707112971 31.0 239.0 10786.0
f N 0.114285714286 8.0 70.0 10786.0
phonemes:

NG UW AY f G UH DX AXR
added a N 0.0753246753247 29.0 385.0 10786.0
added i N 0.102272727273 27.0 264.0 10786.0
added k IY 0.08 16.0 200.0 10786.0
added o N 0.0676056338028 24.0 355.0 10786.0
added zh R 0.0777777777778 14.0 180.0 10786.0
added u N 0.0892857142857 35.0 392.0 10786.0
added r N 0.063829787234 21.0 329.0 10786.0
added sh IH 0.08 8.0 100.0 10786.0
added e N 0.0736842105263 28.0 380.0 10786.0
added w N 0.0773480662983 14.0 181.0 10786.0
added n T 0.0747330960854 21.0 281.0 10786.0
added h N 0.0775510204082 19.0 245.0 10786.0

RULES

SP: 0 (0.000000)

SIL: 0 (0.000000)

P(ph) < 0.005000: 3 (0.001632)

P(var) < 0.110000: 1827 (0.994015)

rules: 1838

accepted rules: 8 (0.000000)

added jap. rules: 12 (0.006529)

List of Figures

List of Tables

References

[AKS00] Ingunn Amdal, Filipp Korkmazskiy, and Arun C Surendran. Joint Pronunciation Modeling of Non-Native Speakers Using Data-Driven Methods. *Proc. ICSLP*, pages 622–625, 2000.

[ARP94] ARPA. November 1998 ARPA Continuous Speech Recognition Benchmark Tests Corpora and Instructions. Internet: http://www.ldc.upenn.edu/Catalog/readme_files/csr1/c Jun 1994.

[BGN02] Norbert Binder, Rainer Gruhn, and Satoshi Nakamura. Recognition of Non-Native Speech Using Dynamic Phoneme Lattice Processing. *Proc. Acoust. Soc. Jap.*, page 203f, March 2002.

[Com01] D. V. Compernolle. Recognizing Speech of Goats, Wolves, Sheep and ... Non-Natives. *Speech Communication*, 35:71–79, 2001.

[EW00] Matthias Eichner and Matthias Wolff. Data-Driven Generation of Pronunciation Dictionaries in the German Verbmobil Project - Discussion of Experimental Results. *Proc. ICASSP*, 2000. Istanbul, Turkey.

[HZ01] X. He and Y. Zhao. Model complexity optimization for nonnative english speakers. *Proc. EuroSpeech*, pages 1461–1464, 2001. Aalborg, Denmark.

[NMS+96] A. Nakamura, S. Matsunaga, T. Shimizu, M. Tonomura, and Y. Sagisaka. Japanese Speech Databases for Robust Speech Recognition. *Proc. ICSLP*, pages 2199–2202, 1996. Philadelphia, USA.

[Tom00] Laura Mayfield Tomokiyo. Lexical And Acoustic Modeling of Non Native Speech in LVCSR. *Proc. ICSLP*, 2000. Beijing, China, Paper Nr.: 1619.

[WY99] S. Witt and S. Young. Off-Line Acoustic Modeling of Non-Native Accents. *Proc. EuroSpeech*, pages 1367–1370, 1999. Budapest, Hungary.

[Y+99] S Young et al. *The HTK Book*. Entropic Ltd, 1999.