

TR-IT-0016

音素配列構造の制約を用いた連続音声認識  
Phonotactic Constraints for Continuous  
Speech Recognition

大脇 浩      シンガー ハラルド      鷹見 淳一  
Hiroshi Ohwaki   Harald Singer   Jun-ichi Takami

1993.8

## 概要

従来のCFG、PCFG、n-gramなどのような文法を用いた連続音声認識手法では、認識できる語彙を制限することによって、高い音声認識性能が実現されていた。しかしその一方で、これらの方法では文法で表現されていない語については、認識できないという点が問題となる。本実習では、認識語彙に関する制限を設けない音素タイプライタ型の音声認識の実現をめざして、日本語の音素配列構造 (phonotactics) という、緩い言語的制約のみを用いた連続音声認識方式の検討を行なった。言語制約の弱さに起因する認識率の低下を防ぐため、音素モデルとしては、SSSによって得られた3308個の異音モデルを用いた。文節発声データを用いて発話様式適応したモデルを用いて、279文節発声データ (SB3) に対して認識実験を行なったところ、文節正解率 51.3%、音素正解率 92.5% という、良好な認識性能が得られた。

## Abstract

In general, high recognition accuracy can be obtained by using CFG, PCFG, n-grams etc. and thus constraining the vocabulary and grammatical constructs. However, a problem exists in recognizing words which have not been previously recorded in these grammars. In this research, we were researching the possibility of using only weak phonotactic constraints apparent in Japanese, e.g. the sequence p/t/a is not "valid" Japanese. To maintain high recognition accuracy in spite of these very weak constraints, 3278 allophone models which were trained by the Successive State Splitting (SSS) algorithm were then adapted to phrase utterance type using the VFS algorithm. Phrase recognition accuracy for 279 phrases (SB3) was 51.3 %, phoneme accuracy was 92.5 %, a surprisingly good result in view of the fact that only a standard One-Pass DP was used.

©ATR 音声翻訳通信研究所

©ATR Interpreting Telecommunications Research Labs.

# 目次

1	はじめに	2
2	音素タイプライタの実現	3
2.1	音響モデル (隠れマルコフ網)	3
2.2	言語的制約 (音素配列構造)	3
2.3	探索手法 (One-Pass DP 法)	5
2.4	システム構成	7
3	実験と結果	7
3.1	音声資料	7
3.2	実験内容	8
3.3	HMM モデル学習方法	9
3.4	結果	9
3.5	認識過程における所要計算時間について	10
3.6	エラー解析	12
4	考察	13
5	まとめ	13
A	HTK ver1.2 を使うノウハウ	17
B	音素配列構造の制約	18
B.1	文脈自由文法による表現	18
B.2	正規文法による表現	20
C	認識誤り例	21

# 1 はじめに

現在の連続音声認識手法では、音響的特徴量に対するパターン識別法の適用に加えて、何らかの言語情報を併用する方法が主流となっている。通常、音声認識の過程で利用される言語情報は、人為的に作成された有限状態オートマトンや文脈自由文法、あるいは大量の言語データの統計処理により獲得された確率文法 [3] などの形で与えられ、これらによって、認識の対象となる単語の種類 (語彙) や、それらの間の順序関係 (構文) などに対する制約が加えられる。その結果、許された単語と正しい構文を用いて表現された言語音声に対しては、音響パターンに多少の変動が生じている場合でも、高い認識精度が実現されている。

しかし、このような手法には、予め設定された言語的制約から外れた表現を扱うことが本質的に不可能であるといった大きな問題がある。そのため、これまでの手法で広範囲に渡る言語表現を扱おうとする場合には、人的作業による文法規模の拡張や、確率文法獲得のための言語データベースの大規模化などによって、文法で扱うことのできる言語現象の範囲を広げる必要がある。しかしいずれの場合でもその実行には限界があり、多様な音声現象を全て覆い尽くす文法を構築することは不可能である。

このような問題は、音響レベルでの識別結果のみを手がかりとして音素ラベル系列の探索を行う「音素タイプライタ」を実現することによって解決することができる。音素タイプライタの構築は、音声認識システムを考えた場合の一つの究極の目標であり、既にその実現を試みた研究も報告されている。ただしこの例では、離散分布型の音素 HMM を用いた音響レベルでの識別能力だけでは十分に高い性能が得られないため、音節間の連鎖確率のような低レベルの言語情報や、継続時間のような韻律情報が併用されている。

しかし最近では、前後の音素の違いにより生じる音響パターンの変動を積極適に考慮した高精度な音素コンテキスト依存 HMM についての研究も進められ、その自動生成法についても提案されている。このような音素コンテキスト依存モデルを用いることによって音響レベルでの識別性能を向上させることができるため、音響レベル以外の側面からの制約条件を大幅に緩和できる可能性が高い。

そこで今回我々は、高精度な音素コンテキスト依存モデルと、日本語としての最低限の言語的制約である音素配列構造 (phonotactics) のみに基づいて探索範囲に制限を加えた One-Pass DP による探索法を用いた音素タイプライタ型の連続音声認識手法を提案する。音素配列構造に基づく言語的制約は、日本語として起こり得ない音素並びを探索範囲から除外するためだけに使用され、連鎖確率のように特定の音素並びに対する重み付けを行うものではないため、ある特定の音素並びが認識され難くなったり、認識不可能になるといった弊害が生じることはない。従って、音素タイプライタで使用する言語的制約としては適切なものであると考えられる。

本報告では、提案する音声タイプライタを構成する上で重要な役割を持つ音響モデル、言語的制約、および探索手法について述べ、特定話者の連続音声に対する認識実験を通してその性能を明らかにする。

## 2 音素タイプライタの実現

### 2.1 音響モデル (隠れマルコフ網)

本手法では、認識語彙を制限しない音素タイプライター型の音声認識をめざしている。この場合、言語的な制約が非常に弱いため、音響レベルでの音素識別性能を高める必要がある。そこで本手法では、前後の音素の違いによる音響パターンの変動を積極的に考慮した、精密な音素コンテキスト依存モデル (triphone) を使用する。この triphone は、逐次状態分割 (SSS)[4] によって自動生成した隠れマルコフ網 (HMnet) を使用した。HMnet は、複数の triphone モデル間での状態共有を許すことによって、多くの種類の triphone モデルを、比較的少量のモデルパラメータで表現した表現効率の高いモデルとなっている。HMnet の構造の概観を図1に示す。

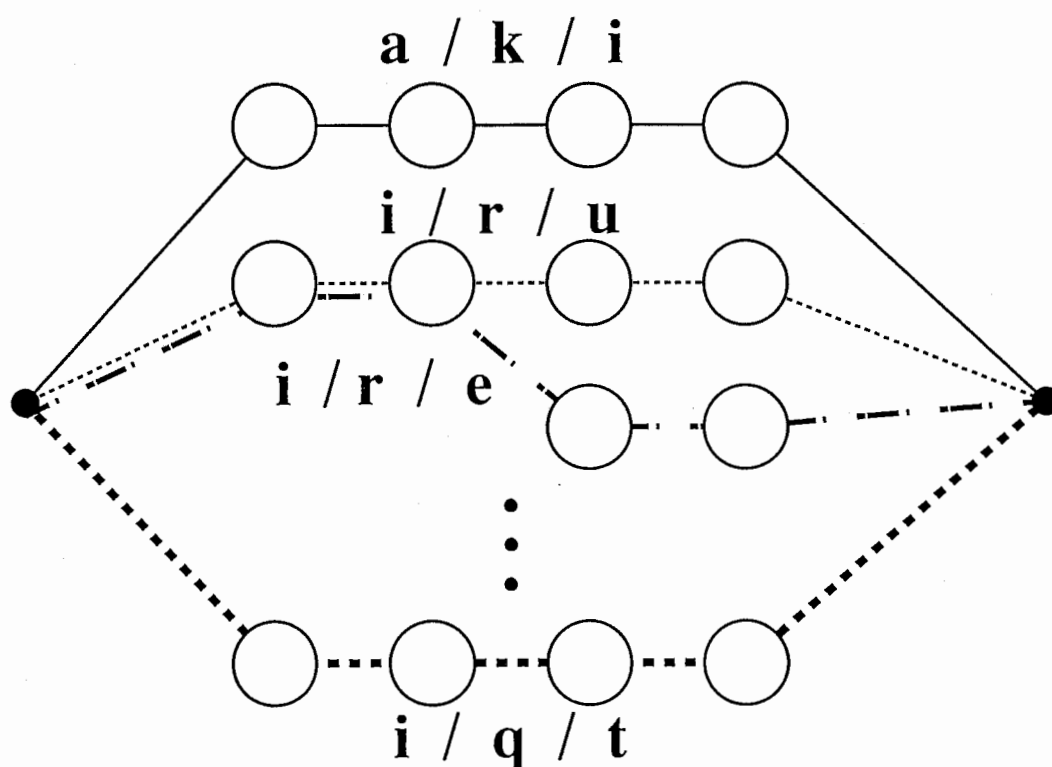


図 1: HMnet 構造

### 2.2 言語的制約 (音素配列構造)

本実験においては、認識語彙の制約の全くない音素タイプライター型音素列認識を目指しているため、文法などによって認識語彙を制限することはできない。しかし、このままでは、あまりに認識語の探索範囲が広過ぎる。そこで最低限、日本語中に存在しないような仮定は、認識語彙の制限がないことを条件に、認識語の探索範囲から外しても良いのではないかと考えた。この目的に適したものに、日本語の音素配列構造制約 (phonotactic constraints) [2] がある。これは、日本語中に存在する音素並びの約束を記述したものである。つまり、日本語中には存在しない音素並びがある。この約束によって、日本語中に存在する音素並びと、存在しない音素並びを識別することができる。この日本語の音素配列構造制約は有限状態オートマトンの形で表現できる。図

2は、日本語中に存在する「音節並び」を表している。

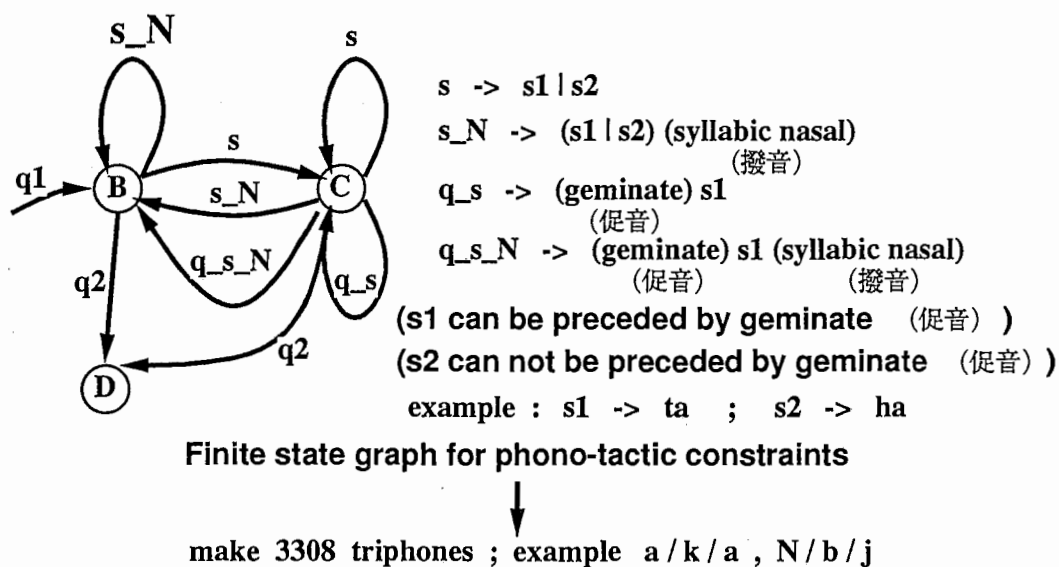


図 2: 日本語中に存在する音節並びを表す有限状態オートマトン

ここに示されている B,C,D は有限状態オートマトンの各状態に相当する。そして、各記号の意味は、

- s1 : 前に促音が来ることができる音節
- s2 : 前に促音が来ることができない音節
- q1 : 語の最初に来る無音
- q2 : 語の最後に来る無音

- s : s1 | s2
- s\_N : (s1 | s2) 撥音
- q\_s : 促音 s1
- q\_s\_N : 促音 s1 撥音

s1 の例 : 「た (ta)」

s2 の例 : 「は (ha)」

である。

本実験では日本語中の音節並びではなく、音素並びの約束が欲しいため、この図 2 から還元して音素並びの約束を抽出した。そして、この日本語中に存在しない音素並びを認識候補の探索仮定から除く。この約束事を実現するには、本来なら、それを表す文法を記述しなければならない。そしてプログラムによって「日本語中に存在する音素並びを表す有限状態オートマトン構造をメモリ上に構築しなければならないが、それをいなくても良い方法がある。それは音素を triphone とし、日本語 phonotactic constraints を満たす triphone だけを認識に使用することである。これによって日本語音素配列構造制約の音素つながり関係が自動的に表現されている。その様子を図 3 に示す。

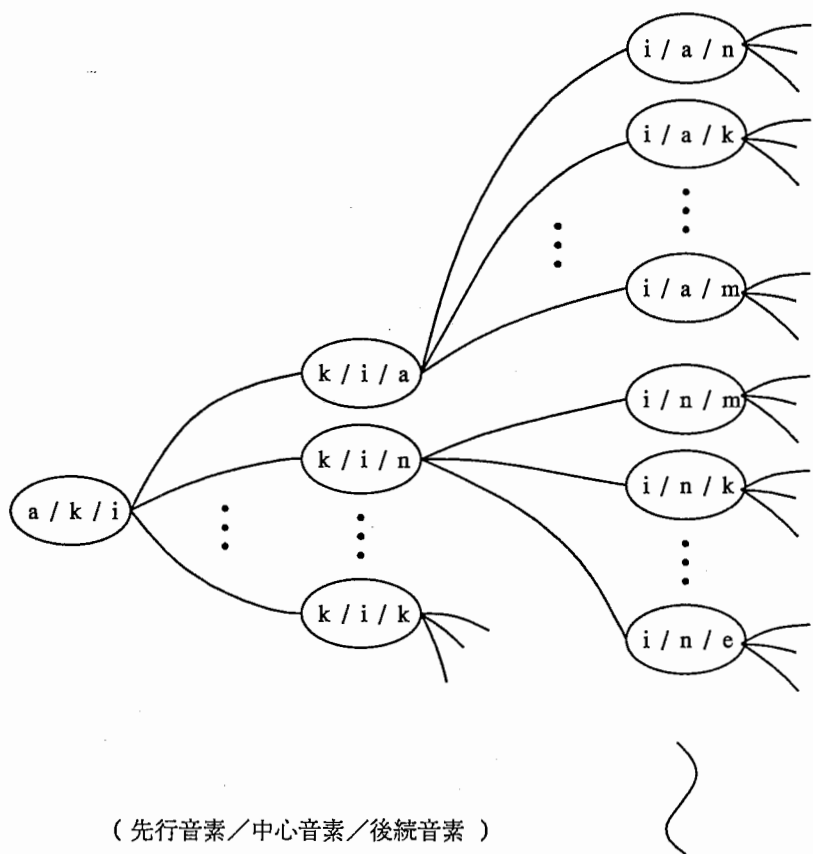


図 3: 使用する triphone の制限によって自動的に構築される日本語音素配列構造制約

このように認識に使用する triphone の種類を日本語 phonotactic constraints を満たすものだけに絞ると、例えば日本語 26 音素ラベルを用いた場合、triphone の数は、 $26^3=17576$  通りから 3308 通り (外来語の音素並びを除く) にまで削減される。なお、これらの triphone は、図 2 に示した有限状態オートマトンと等価な文脈自由文法 (CFG) で表現できる長さ 8 の音素列の中から、異なる 3 つ組音素列を抜き出すことによって求めた。付録 B.1 に図 2 を表現する CFG を示す。

このようにして生成された 3308 個の triphone によって、認識語彙を制限しないという条件を保ったまま、認識時の探索範囲を狭めることができる。ここで日本語中に存在する 3 つ組の音素並びの例と、存在しない 3 つ組の音素並びの例を示す。

- 存在する音素並び : a / k / a , N / b / j
- 存在しない音素並び : p / p / k , j / h / d

### 2.3 探索手法 (One-Pass DP 法)

本実験では、連続音声認識における探索方法として、One Pass DP 法を用いた。One Pass DP 法は

1. フレーム同期での探索が可能
2. ビームサーチを併用する際に、スコアの正規化が不用。(今回は、フルサーチを用いている。)

という利点がある。この One-Pass DP 法を用いた時の認識候補の探索範囲を、音素配列構造制約を満たす音素並びだけに絞った。このことは、前節で述べたように認識に使用する triphone を音素配列構造制約を満たすものだけに制限することによって暗示的に実現される。この様子を図4に示す。

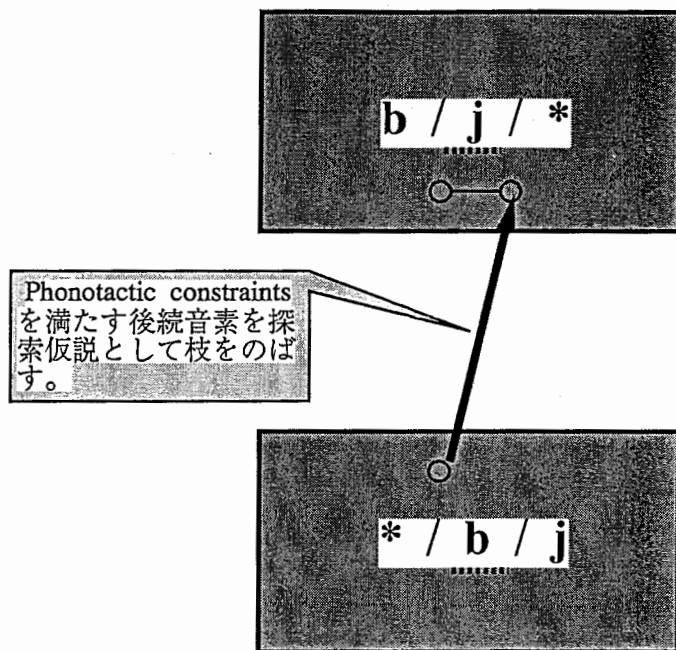


図 4: 音素 phonotactic constrained One Pass DP

この図4のように One-Pass DP 法において後続音素の探索仮説の枝を伸ばす時、音素配列構造制約を満たす音素仮説のみを候補として伸ばす。これによって探索範囲を絞り込むことができ、計算量についても減らすことができる。

## 2.4 システム構成

ここに本システムのブロック図を図5に示す。

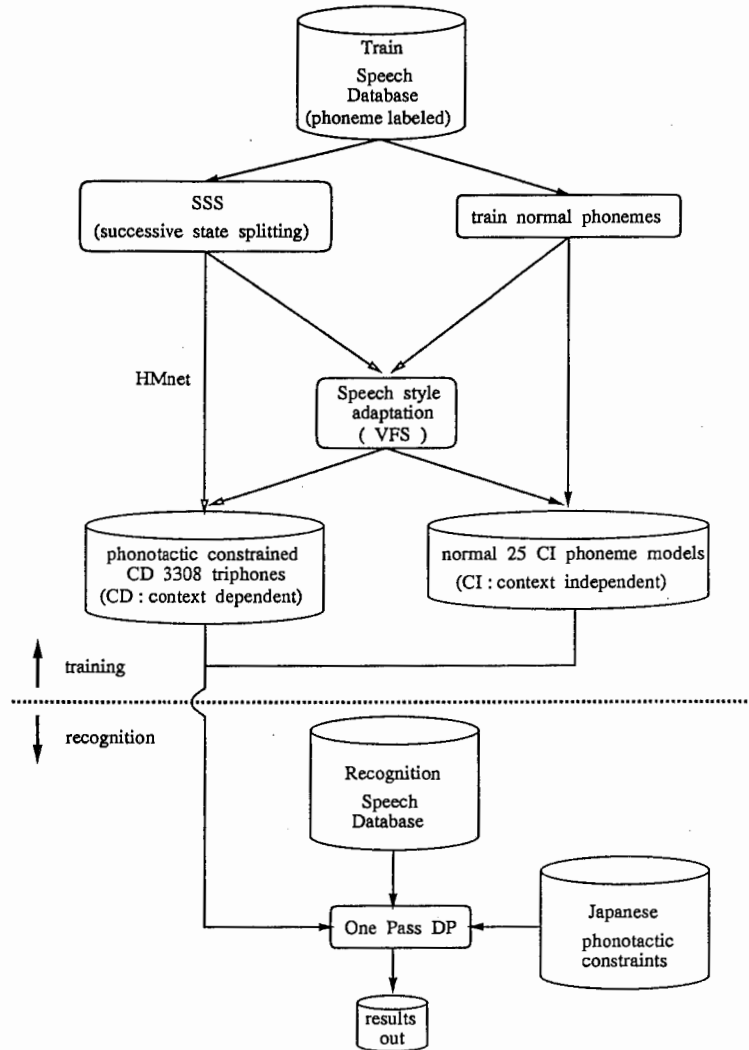


図5: 本システムのブロック図

## 3 実験と結果

### 3.1 音声資料

日本人男性話者 (MHT) 1 名  
サンプリング周波数 : 20kHz  
量子化ビット数 : 16bit  
分析周期 : 5ms  
窓幅 : 20ms  
窓種類 : ハミング窓  
データ構成 : 16 次 LPC ケプストラム  
log パワー



16次 $\Delta$  LPC ケプストラム

$\Delta$  log パワー

本実験で用いた実験用データの発話様式の種類は、以下の通りである。

[モデルの学習]

- 孤立発声された重要語 5240 単語の偶数番目の単語 (2620 単語) 平均 5.7mora/sec (記号 (1-even) と略す。)

[認識実験]

- 孤立発声された重要語 5240 単語の奇数番目の単語 (2620 単語) の中からランダムに取り出した 100 単語平均 5.7mora/sec (記号 (1-odd) と略す。)
- 音素連鎖バランス単語 (216 単語) 平均 5.7mora/sec (記号 (B) と略す。)
- 国際会議参加問い合わせに関する会話文を、複合語を許さない文節単位で区切って発声したもの (279 単語) 平均 7.1mora/sec (記号 (SB3) と略す。)

### 3.2 実験内容

本実験では、音響モデルとして音素コンテキスト依存モデル (triphone) を、また、言語的制約として日本語音素配列構造を使用して、音素タイプライタ型の連続音声認識を行なっている。この音素コンテキスト依存モデルや、日本語音素配列構造制約の有効性を調べるために、さらに音素コンテキスト非依存のモデル (普通の音素モデル) や、日本語音素配列構造による制約を用いない場合の比較実験についても行なった。また、本実験では学習データと認識データの間で発話様式 (発話速度) の異なる場合の認識を試みたため、VFS (Vector Field Smoothing) を用いた発話様式適応の効果についても合わせて検討を行なった。これらの実験条件の組み合わせを表 1 に示す。

表 1: 実験条件の組み合わせ

音響モデル		音素配列構造		発話様式適応
context dependent	→	考慮する	→	あり
	↘		↘	
context independent	→	考慮しない	→	なし

なお、この「発話様式適応 (speaking style adaptation)」は、重要語 5240 単語 (偶数番目) (1-even) によって学習したモデルを、同じ話者の文節発声データ (SA1, 2, 4 及び SB1, 2, 4) を用いた VFS (Vector Field Smoothing) による再学習により行なった。

また、context dependent、音素配列構造考慮なし、の条件の実験項目は cpu time を多大に消費するため、文章発声データ (SB3) のみを評価した。

比較実験において、前語音素に対する context independent の音素モデルを使用するときは、日本語音素配列構造を正規文法により明示することによって実現した。この正規文法を、付録 B.2 に示す。

### 3.3 HMM モデル学習方法

音素コンテキスト依存モデルは逐次状態分割 [4] によって学習した。また、比較実験のための音素モデルは通常の Baum-Welch 法により学習した。どちらの場合も、学習用サンプルには、ラベル情報に従って切り出した音素を使用した。

### 3.4 結果

認識率を、表 2 に示す。

表 2: 実験結果

Speaking Style Adaptation	Rec. Data	Context Dependent		Context Independent	
		YES Phonotactics	NO Phonotactics	YES Phonotactics	NO Phonotactics
NO	1	73.00 / 95.45%		61.00 / 93.55%	41.00 / 90.27%
	B	43.52 / 90.19%		32.41 / 87.63%	21.30 / 83.61%
	SB3	27.24 / 85.18%	20.43 / 82.56%	19.00 / 84.93%	11.11 / 79.55%
YES	1	49.00 / 91.15%		55.00 / 92.54%	37.00 / 88.87%
	B	33.33 / 88.62%		30.56 / 86.80%	19.91 / 82.97%
	SB3	51.25 / 92.47%	29.03 / 87.73%	39.07 / 90.41%	24.73 / 86.84%

51.25 / 95.54% = word recognition rate (%) / phoneme recognition rate (%)

ここで、単語認識率 (word recognition rate)(%) は、

$$\text{単語認識率}(\%) = \frac{N_{\text{words}} - N_{\text{miss}}}{N_{\text{words}}}$$

ただし、

$N_{\text{words}}$  : 評価データの中の全単語(文節)の数

$N_{\text{miss}}$  : 誤り単語(文節)の数

また、音素認識率 (phoneme recognition rate)(%) は、

$$\text{音素認識率(\%)} = \frac{N_{\text{phones}} - N_{\text{subst}} - N_{\text{ins}} - N_{\text{del}}}{N_{\text{phones}}}$$

ただし

$N_{\text{phones}}$  : 評価データの中の全音素数

$N_{\text{subst}}$  : 置換誤りの数

$N_{\text{ins}}$  : 挿入誤りの数

$N_{\text{del}}$  : 脱落誤りの数

とする。

表2から認識率を検討すると、単語認識率、音素認識率に共通した傾向があることがわかる。

まず、モデルのコンテキスト依存、非依存による認識率の良さの傾向は、

コンテキスト依存モデル > コンテキスト非依存モデル

となっている。

また、音素配列構造制限の適用、未適用の違いによる認識率の良さの傾向は

音素配列構造 適用 > 音素配列構造 未適用

となっている。

そして、認識データの発話様式の違いによる認識率の良さの傾向は

(1) > (B) > (SB3)

である。発話様式適応を行なうと (SB3) の認識率は上がり、(1) や (B) の認識率は下がる。

この結果から、コンテキスト依存モデルの音響的精密さや、音素配列構造制約は認識性能向上に貢献していることがわかる。

### 3.5 認識過程における所要計算時間について

認識に要する計算時間 (計算マシン依存) を計測したのでそれを付録として示す。

表 3: 各実験のおおよその cpu time (10 単語の認識時間; machine:HP9000/730; SPECmark=76.8; 76MIPS)

phonotactic constraints (PT)	CD3308	CI26
with	662 sec	273 sec
without	35724 sec	25 sec

ここで

CD3308 = コンテキスト依存の 3308 個の triphone を用いた時

CI26 = コンテキスト非依存の 26 個の音素を用いた時

である。それぞれの所要計算時間に対する説明を加えると、

- A) (CD3308 + without PT) : が最も計算時間がかかるのは、モデル数が 3308 個ある上に、音素配列構造制約を用いていないために、音素のグリッド空間を移る時の探索仮説数が 3308 個になってしまうためである。
- B) (CD3308 + with PT) : が A) よりも計算時間が少ないのは、音素のグリッド空間を移る時の探索仮説数が音素配列構造制約によって 3308 個よりかなり減らせるためである。
- C) (CI26 + without PT) : が最も高速なのは、用いる音素数が 26 個であり、音素のグリッド空間を移る時の探索仮説数が 26 個というのは大した数ではないからである。
- D) (CI26 + with PT) : が C) より計算時間がかかるのは C) と違ってコンテキスト非依存モデルにて音素配列構造制約を取り入れるために、メモリ上に音素配列構造制約を実現する有限オートマトン構造を構築する必要があるからである。

### 3.6 エラー解析

今回の例では、継続時間長制御を用いていないため、母音の長母音化などの音素継続時間に関するエラーが多かった。そこで実際にどのようなエラーがどの程度起こっているか調べるためにエラー解析を行なった。

その結果を表4に示す。

表4: エラー解析結果 (単語認識率 (%))

Rec. Data	P T	SSA	Context	Original	a ↔ a a	t ↔ q t	n ↔ ng n	a ↔ a a t ↔ q t	All
1	Yes	No	CD	73.00	82.00	73.00	74.00	82.00	83.00
			CI	61.00	75.00	63.00	62.00	80.00	81.00
		Yes	CD	49.00	65.00	51.00	52.00	67.00	71.00
			CI	55.00	67.00	56.00	56.00	68.00	69.00
SB3	Yes	No	CD	27.24	38.71	28.67	28.32	40.50	41.94
			CI	19.00	39.07	19.35	20.07	39.78	41.94
		Yes	CD	51.25	67.03	51.97	52.69	68.10	69.53
			CI	39.07	69.18	40.50	39.78	70.97	73.12

CD=Context Dependent  
 CI=Context Independent  
 PT=Phonotactics  
 SSA=Speaking Style Adaptation  
 Rec. Data=Recognition Data

ここで表の中の記述について説明する。

- 1) original: 実験により得られた単語 (文節) 認識率。
- 2) a ↔ a a: a, e, i, o, u の5つの母音が長母音化または短母音化するエラーをエラー対象から除外して考えた時の単語認識率を示す。
- 3) t ↔ q t: / q / ↔ / q k /, / t / ↔ / q t /, / ts / ↔ / q ts / といいた3通りのエラーを、エラー対象から除外して考えた時の単語認識率を示す。
- 4) n ↔ ng n: / n / ↔ / ng n /, / m / ↔ / ng m /, / g / ↔ / ng g / といいた3通りのエラーを、エラー対象から除外して考えた時の単語認識率を示す。
- 5) a ↔ a a, t ↔ q t: 2) と 3) に属するエラーを、エラーの対象から除外して考えた時の単語認識率を示す。

6) All: 2) から 4) までに属する全てのエラーを、エラーの対象から除外して考えた時の単語認識率を示す。

この表 4 から、/ a / が / a a / のように誤認識される母音の長母音化のエラーの頻度が非常に高いことを示している。また、/ t / が / q t / や、/ n / が / ng n / と誤認識されるのも、音素の継続時間にまつわるエラーと考えてさしつかえないので、このようなエラーは継続時間長に関する制御を試みることによって、かなりの改善が見込まれる。

付録 C に、音声データ (SB3) を、文節発声の (SB3) に VFS(Vector Field Smoothing) によって発話様式適応した音素 context dependent モデルによって、認識を行なった時のエラーリストを示す。このエラーリストの中で、LAB は認識音声データにあらかじめ付けられたラベルを示し、REC は実際に認識された音素列を示す。

## 4 考察

本手法と同じような音素列認識 (音素タイプライタ) を試みた過去の例に、89 年に ATR 自動翻訳電話研究所の川端らによって行なわれたもの [1] がある。それと、今回の結果を比較してみると、

表 5: 89 年に川端らによって行なわれた実験結果との比較

	HMM	duration control	constraints	search	phoneme recognition rate
Kawabata	Discrete	Yes	syllable n-gram	LR + beam(170)	88.9%
This work	Continuous	No	phonotactics	One-Pass DP	92.5%

となる。川端らの例と本方式との間で、認識時に与えている制約条件の強さを比較してみる。まず、川端らの用いた syllable n-gram よりも、今回用いた音素配列構造による制約の方が、多様な言語表現に対する柔軟性という点で有理である。それは、音素配列構造による制約では、あらゆる音素列を同じ公平さで認識できるのに対し、syllable n-gram は、音節の出現確率を設定しているため、n-gram の学習用データ中で出現頻度の低かった (あるいは全く現れなかった) 音節連鎖には低い重み付けが行なわれてしまうためである。さらに、川端らの例では、HMM の状態についての継続時間長制御が行なわれているが、本方式ではそれについても行なっていない。このように、今回の実験では言語的・韻律的な制約が緩いにも関わらず、ほぼ同等の認識性能を得ることができた。

## 5 まとめ

本報告では、音素配列構造の制約を用いた音素タイプライタ型の連続音声認識手法について述べた。

本手法を用いて行なった連続音声認識実験では、文節発声データ (SB3) に対して音素認識率 92.5%、文節認識率 51.3% を得ることができた。また、音響モデルの学習に使用したデータ (1-even) に最

も類似した単語発声データ (1-odd) に対しては、音素認識率 95.5%、単語認識率 73.00% という、高い認識性能が得られた。また、エラー解析の結果より、音素の継続時間長を考慮することにより、さらなる認識性能の向上を期待できることが確認された。

今後の課題としては、そのような韻律的側面からの制御による音素列誤り訂正法の検討、および、n-best までの複数候補の獲得法の検討などが挙げられる。

なお、この研究は、大脇浩 (電通大) が学外実習として ATR 音声翻訳通信研究所に滞在した、1993 年 7 月 12 日～8 月 31 日の期間中に行なわれたものである。

## 謝 辞

本実習の機会を与えて下さった電気通信大学の樽松明教授、ATR 音声翻訳通信研究所の山崎泰弘社長、並びに匂坂芳典室長に深く感謝致します。また、日頃から時間を割いて熱心にご指導頂いた鷹見淳一研究員、Harald.Singer 研究員を始めとする音声翻訳通信研究所の皆様感謝します。

## 参考文献

- [1] T. Kawabata, T. Hanazawa, K. Itoh, and K. Shikano. Japanese phonetic typewriter using HMM phone recognition and stochastic phone-sequence modeling. *Transactions of the Institute of Electronics, Information and Communication Engineers*, E74(7):1783–1787, July 1991.
- [2] K. Kita. *A Study on Language Modeling for Speech Recognition*. PhD thesis, Waseda University, October 1992.
- [3] K. Kita. LR parser applications to spoken language processing. Technical Report SP93-11, IEICE, 1993. (in Japanese).
- [4] J. Takami and S. Sagayama. A successive state splitting algorithm for efficient allophone modeling. In *Proc. ICASSP*, volume 1, pages 573–576, San Francisco, March 1992.
- [5] S. Young, N. Russell, and J. Thornton. The use of syntax and multiple alternatives in the VODIS voice operated database inquiry system. *Computer Speech and Language*, 5:65–80, 1991.



# 付 録

## A HTK ver1.2 を使うノウハウ

今回、HTK ver1.2[5]において triphone を用いるときに少々問題があった。この解決法を、後に HTK を使用する方のために書き残す。

HTK ver1.2では、認識過程で音素コンテキストの整合性を保ったまま始端、終端モデルを指定(例えば「無音モデル」など)にすることができない。このため、これに起因する認識誤りが続出した。その誤りの例は

( sil/sil/a → k/i/b → i/b/a → ... → k/e/e → i/sil/sil)

のように、語頭や語尾で音素コンテキストに不整合が生じる。そこで図6のように、データとモデルに少し工夫を施した。

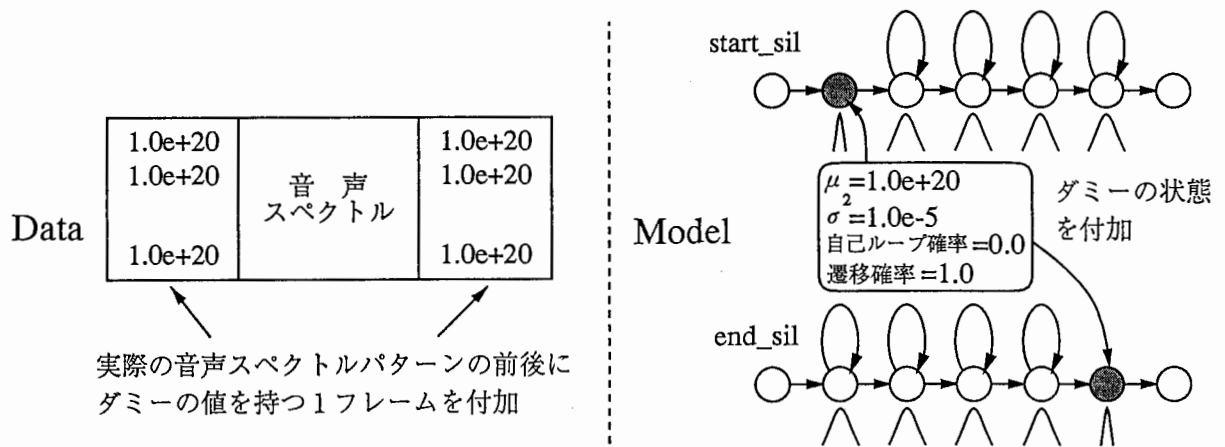


図 6: HTK ver1.2 におけるデータとモデルの工夫

このようにデータには、実際の音声スペクトルパターンの前後に特殊な値を付け加えた。また、無音モデルには、データに付け加えた特殊な値にマッチングする特別な状態(図では、灰色の状態)を付け加えた。

## B 音素配列構造の制約

### B.1 文脈自由文法による表現

音素配列構造の制約 (LR パーザの文脈自由文法にて記述)

---

```
;;; phon.gra
(<start> <--> (<_start>))
(<_start> <--> (<D>))
(<B> <--> (q1))
(<B> <--> (<B> <s_N>))
(<C> <--> (<B> <s>))
(<C> <--> (<C> <s>))
(<C> <--> (<C> <q_s>))
(<B> <--> (<C> <s_N>))
(<B> <--> (<C> <q_s_N>))
(<D> <--> (<B> q2))
(<D> <--> (<C> q2))
(<s> <--> (<s1>))
(<s> <--> (<s2>))
(<s_N> <--> (<s1> <nasal>))
(<s_N> <--> (<s2> <nasal>))
(<q_s_N> <--> (<sokuon> <s1> <nasal>))
(<q_s> <--> (<sokuon> <s1>))
(<s1> <--> (k a))
(<s1> <--> (k i))
(<s1> <--> (k u))
(<s1> <--> (k e))
(<s1> <--> (k o))
(<s1> <--> (k j a))
(<s1> <--> (k j u))
(<s1> <--> (k j o))
(<s1> <--> (p a))
(<s1> <--> (p i))
(<s1> <--> (p u))
(<s1> <--> (p e))
(<s1> <--> (p o))
(<s1> <--> (p j a))
(<s1> <--> (p j u))
(<s1> <--> (p j o))
(<s1> <--> (t a))
(<s1> <--> (ch i))
(<s1> <--> (ts u))
(<s1> <--> (t e))
(<s1> <--> (t o))
(<s1> <--> (ch j a))
(<s1> <--> (ch j u))
(<s1> <--> (ch j o))
(<s1> <--> (s a))
(<s1> <--> (sh i))
(<s1> <--> (s u))
(<s1> <--> (s e))
(<s1> <--> (s o))
(<s1> <--> (sh j a))
(<s1> <--> (sh j u))
(<s1> <--> (sh j o))
(<s2> <--> (i))
(<s2> <--> (u))
(<s2> <--> (e))
(<s2> <--> (o))
(<s2> <--> (m a))
(<s2> <--> (m i))
(<s2> <--> (m u))
(<s2> <--> (m e))
(<s2> <--> (m o))
(<s2> <--> (m j a))
(<s2> <--> (m j u))
(<s2> <--> (m j o))
(<s2> <--> (n a))
(<s2> <--> (n i))
(<s2> <--> (n u))
(<s2> <--> (n e))
(<s2> <--> (n o))
(<s2> <--> (n j a))
(<s2> <--> (n j u))
(<s2> <--> (n j o))
(<s2> <--> (b a))
(<s2> <--> (b i))
(<s2> <--> (b u))
(<s2> <--> (b e))
(<s2> <--> (b o))
(<s2> <--> (b j a))
(<s2> <--> (b j u))
(<s2> <--> (b j o))
(<s2> <--> (r a))
(<s2> <--> (r i))
(<s2> <--> (r u))
(<s2> <--> (r e))
(<s2> <--> (r o))
(<s2> <--> (r j a))
(<s2> <--> (r j u))
(<s2> <--> (r j o))
(<s2> <--> (z a))
(<s2> <--> (zh i))
(<s2> <--> (z u))
(<s2> <--> (z e))
(<s2> <--> (z o))
(<s2> <--> (zh j a))
(<s2> <--> (zh j u))
(<s2> <--> (zh j o))
(<s2> <--> (j a))
(<s2> <--> (j u))
(<s2> <--> (j o))
(<s2> <--> (w a))
(<s2> <--> (h a))
(<s2> <--> (h i))
(<s2> <--> (h u))
```

(<s2> <--> (h e))  
(<s2> <--> (h o))  
(<s2> <--> (h j a))  
(<s2> <--> (h j u))  
(<s2> <--> (h j o))  
(<s2> <--> (d a))  
(<s2> <--> (d u))  
(<s2> <--> (d e))  
(<s2> <--> (d o))  
(<s2> <--> (g a))  
(<s2> <--> (g i))  
(<s2> <--> (g u))  
(<s2> <--> (g e))  
(<s2> <--> (g o))  
(<s2> <--> (g j a))  
(<s2> <--> (g j u))  
(<s2> <--> (g j o))  
(<sokuon> <--> (q))  
(<nasal> <--> (N))  
;;; EOF

---

## B.2 正規文法による表現

音素配列構造の制約 (正規文法にて記述)

---

```
$s1 = k a | k i | k u | k e | k o |
k j a | k j u | k j o |
p a | p i | p u | p e | p o |
p j a | p j u | p j o |
t a | c h i | t s u | t e | t o |
c h j a | c h j u | c h j o |
s a | s h i | s u | s e | s o |
s h j a | s h j u | s h j o ;

$s2 = a | i | u | e | o |
m a | m i | m u | m e | m o |
m j a | m j u | m j o |
n a | n i | n u | n e | n o |
n j a | n j u | n j o |
b a | b i | b u | b e | b o |
b j a | b j u | b j o |
r a | r i | r u | r e | r o |
r j a | r j u | r j o |
z a | z h i | z u | z e | z o |
z h j a | z h j u | z h j o |
j a | j u | j o |
w a |
h a | h i | h u | h e | h o |
h j a | h j u | h j o |
d a | d u | d e | d o |
g a | g i | g u | g e | g o |
g j a | g j u | g j o ;

$s = $s1 | $s2;
$s_N = $s ng;
$q_s = q $s1;
$q_s_N = q $s1 ng;
$ssqs = $s { { $s } { $q_s } };

(start_sil (
{ { $s_N } [ $ssqs ( $s_N | $q_s_N ) ] } |
{ { $s_N } [ $ssqs ] }
) end_sil
)
```

---

## C 認識誤り例

音声データ (SB3) を、文節発声の (SB3) に VFS(Vector Field Smoothing) によって発話様式適応した 音素 context dependent モデルによって、認識を行なった時のエラーリスト

---

```
Aligned transcription: ../L_DSB/MHT_SB3_01_01.lbl vs MHT_SB3_01_01.rec
LAB: sil d a i i q k a i sil
REC: sil d a i i q k a i i sil
Aligned transcription: ../L_DSB/MHT_SB3_01_04.lbl vs MHT_SB3_01_04.rec
LAB: sil k o k u s a i sil
REC: sil k o k u s a i i sil
Aligned transcription: ../L_DSB/MHT_SB3_01_08.lbl vs MHT_SB3_01_08.rec
LAB: sil g o k i b o o s a r e r u sil
REC: sil p o o k i b o o s a r e r u sil
Aligned transcription: ../L_DSB/MHT_SB3_01_10.lbl vs MHT_SB3_01_10.rec
LAB: sil sh j o t e e n o sil
REC: sil sh j o t e n o sil
Aligned transcription: ../L_DSB/MHT_SB3_01_15.lbl vs MHT_SB3_01_15.rec
LAB: sil h a q p j o o sil
REC: sil h a q k j o o sil
Aligned transcription: ../L_DSB/MHT_SB3_01_17.lbl vs MHT_SB3_01_17.rec
LAB: sil b e t s u o sil
REC: sil b e t s u n o sil
Aligned transcription: ../L_DSB/MHT_SB3_01_19.lbl vs MHT_SB3_01_19.rec
LAB: sil k o k u s a i sil
REC: sil k o k u s a i i sil
Aligned transcription: ../L_DSB/MHT_SB3_01_20.lbl vs MHT_SB3_01_20.rec
LAB: sil k a i g i sil
REC: sil k a i i sil
Aligned transcription: ../L_DSB/MHT_SB3_01_22.lbl vs MHT_SB3_01_22.rec
LAB: sil o m o o s h i k o m i sil
REC: sil o m o s h i k u w a m i sil
Aligned transcription: ../L_DSB/MHT_SB3_02_03.lbl vs MHT_SB3_02_03.rec
LAB: sil d a i i q k a i sil
REC: sil d a i i q k a i i sil
Aligned transcription: ../L_DSB/MHT_SB3_02_06.lbl vs MHT_SB3_02_06.rec
LAB: sil k o k u s a i sil
REC: sil k o k u s a i i sil
Aligned transcription: ../L_DSB/MHT_SB3_03_04.lbl vs MHT_SB3_03_04.rec
LAB: sil k o k u s a i sil
REC: sil k o k u s a i i sil
Aligned transcription: ../L_DSB/MHT_SB3_03_06.lbl vs MHT_SB3_03_06.rec
LAB: sil s a n g k a o sil
REC: sil s a n g k a o o sil
Aligned transcription: ../L_DSB/MHT_SB3_03_07.lbl vs MHT_SB3_03_07.rec
LAB: sil m o o s h i k o m i t a i n o d e s u sil
REC: sil m o s h i k o m i t a i n o d e s u sil
Aligned transcription: ../L_DSB/MHT_SB3_03_08.lbl vs MHT_SB3_03_08.rec
LAB: sil k e r e d o m o sil
REC: sil k e r e d a m o sil
Aligned transcription: ../L_DSB/MHT_SB3_03_11.lbl vs MHT_SB3_03_11.rec
LAB: sil t e t s u z u k i o sil
REC: sil t e t s u z u k i j o sil
Aligned transcription: ../L_DSB/MHT_SB3_03_13.lbl vs MHT_SB3_03_13.rec
LAB: sil j o r o s h i i n o d e s h j o o k a sil
REC: sil j o r o s h i i r o d e s h j o k a sil
Aligned transcription: ../L_DSB/MHT_SB3_04_03.lbl vs MHT_SB3_04_03.rec
LAB: sil k o k u s a i sil
REC: sil k o k u s a i i sil
```

Aligned transcription: ../L\_DSB/MHT\_SB3\_04\_07.lbl vs MHT\_SB3\_04\_07.rec  
 LAB: sil sh j o t e e n o sil  
 REC: sil sh j o t e n o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_04\_09.lbl vs MHT\_SB3\_04\_09.rec  
 LAB: sil j o o sh i o sil  
 REC: sil i j o o sh i i j o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_04\_11.lbl vs MHT\_SB3\_04\_11.rec  
 LAB: sil s a ng k a sil  
 REC: sil s a ng k a a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_04\_14.lbl vs MHT\_SB3\_04\_14.rec  
 LAB: sil h i ts u j o o d e s u sil  
 REC: sil k i q ch i j o d e s u sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_05\_02.lbl vs MHT\_SB3\_05\_02.rec  
 LAB: sil h a q p j o o s u r u n o d e w a sil  
 REC: sil h a q k j o o s u r u n o d a w a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_05\_07.lbl vs MHT\_SB3\_05\_07.rec  
 LAB: sil k a k a r i m a s u k a sil  
 REC: sil t a k a r i m a s u k a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_06\_01.lbl vs MHT\_SB3\_06\_01.rec  
 LAB: sil g o h a q p j o o o sil  
 REC: sil b o h a q p j o o o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_06\_03.lbl vs MHT\_SB3\_06\_03.rec  
 LAB: sil b a a i n i w a sil  
 REC: sil b a i n i w a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_06\_04.lbl vs MHT\_SB3\_06\_04.rec  
 LAB: sil j o k o o sh j u d a i sil  
 REC: sil k j o k o o sh j u d a i sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_06\_05.lbl vs MHT\_SB3\_06\_05.rec  
 LAB: sil t o o r o k u r j o o o sil  
 REC: sil t o o r o k u r j o o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_06\_07.lbl vs MHT\_SB3\_06\_07.rec  
 LAB: sil s a ng k a sil  
 REC: sil s a ng k a a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_06\_09.lbl vs MHT\_SB3\_06\_09.rec  
 LAB: sil j o ng m a ng e ng d e s u sil  
 REC: sil j o m a i i e ng d e s u sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_07\_02.lbl vs MHT\_SB3\_07\_02.rec  
 LAB: sil b a a i w a sil  
 REC: sil b a i w a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_07\_04.lbl vs MHT\_SB3\_07\_04.rec  
 LAB: sil u k e ts u k e m o sil  
 REC: sil u k e ts u k e m o o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_07\_06.lbl vs MHT\_SB3\_07\_06.rec  
 LAB: sil j o k o o sh j u d a i o sil  
 REC: sil j o k o o sh j u d a i j o o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_07\_07.lbl vs MHT\_SB3\_07\_07.rec  
 LAB: sil h u k u m e t a sil  
 REC: sil h u k u m e q t a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_07\_09.lbl vs MHT\_SB3\_07\_09.rec  
 LAB: sil s a ng m a ng sil  
 REC: sil s a m a ng sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_07\_10.lbl vs MHT\_SB3\_07\_10.rec  
 LAB: sil g o s e ng e ng sil  
 REC: sil g o s e ng g e ng sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_08\_03.lbl vs MHT\_SB3\_08\_03.rec  
 LAB: sil m o o sh i k o m i sil  
 REC: sil m o sh i k u w a m i sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_08\_08.lbl vs MHT\_SB3\_08\_08.rec  
 LAB: sil t e n i sil

REC: sil t e n g n i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_08\_09.lbl vs MHT\_SB3\_08\_09.rec  
LAB: sil i r e r e b a sil  
REC: sil i r e e r e b a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_08\_10.lbl vs MHT\_SB3\_08\_10.rec  
LAB: sil j o r o s h i i n o d e s h j o o k a sil  
REC: sil j o r o s h i i m a d e s h j o k a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_09\_02.lbl vs MHT\_SB3\_09\_02.rec  
LAB: sil g o z h j u u s h j o o sil  
REC: sil g o z h j u s h j o o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_09\_05.lbl vs MHT\_SB3\_09\_05.rec  
LAB: sil k o k u s a i sil  
REC: sil k o k u s a i i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_09\_07.lbl vs MHT\_SB3\_09\_07.rec  
LAB: sil z h i m u k j o k u k a r a sil  
REC: sil z h i m u k j o h u k a r a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_09\_08.lbl vs MHT\_SB3\_09\_08.rec  
LAB: sil s h j o t e e n o sil  
REC: sil s h j o t e n o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_09\_09.lbl vs MHT\_SB3\_09\_09.rec  
LAB: sil j o o s h i o sil  
REC: sil j o o s h i j o o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_09\_11.lbl vs MHT\_SB3\_09\_11.rec  
LAB: sil i t a d a k i m a s u sil  
REC: sil i t a d a k i b a s u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_10\_01.lbl vs MHT\_SB3\_10\_01.rec  
LAB: sil z h j u u s h j o w a sil  
REC: sil z h j u s h j o w a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_10\_06.lbl vs MHT\_SB3\_10\_06.rec  
LAB: sil n i n o sil  
REC: sil r i n o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_10\_07.lbl vs MHT\_SB3\_10\_07.rec  
LAB: sil r o k u z h j u u i c h i sil  
REC: sil r o k u z h j u i c h i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_10\_08.lbl vs MHT\_SB3\_10\_08.rec  
LAB: sil e e t i i a a r u sil  
REC: sil e e k i i j a a r u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_10\_09.lbl vs MHT\_SB3\_10\_09.rec  
LAB: sil z h i d o o sil  
REC: sil z h i d o o o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_10\_10.lbl vs MHT\_SB3\_10\_10.rec  
LAB: sil h o n g j a k u sil  
REC: sil h o j a k u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_10\_12.lbl vs MHT\_SB3\_10\_12.rec  
LAB: sil k e n g k j u u z h j o d e sil  
REC: sil k e n g k j u z h j a d e sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_10\_15.lbl vs MHT\_SB3\_10\_15.rec  
LAB: sil i c h i t a r o o d e s u sil  
REC: sil i c h i t a r o d e s u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_11\_01.lbl vs MHT\_SB3\_11\_01.rec  
LAB: sil e e t i i a a r u sil  
REC: sil e e k i i j a a r u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_11\_03.lbl vs MHT\_SB3\_11\_03.rec  
LAB: sil h o n g j a k u sil  
REC: sil h o i j a k u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_11\_05.lbl vs MHT\_SB3\_11\_05.rec  
LAB: sil k e n g k j u u z h j o sil  
REC: sil p e n g k i j u z h j o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_11\_06.lbl vs MHT\_SB3\_11\_06.rec



LAB: sil t a k a r a s a m a d e s u n e sil  
REC: sil t a k a r a s a m a d e s u p e e sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_11\_07.lbl vs MHT\_SB3\_11\_07.rec  
LAB: sil d e w a sil  
REC: sil d e o a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_11\_09.lbl vs MHT\_SB3\_11\_09.rec  
LAB: sil m o o s h i k o m i sil  
REC: sil m o s h i k o m i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_11\_10.lbl vs MHT\_SB3\_11\_10.rec  
LAB: sil j o o s h i o sil  
REC: sil j o o s h i j o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_11\_12.lbl vs MHT\_SB3\_11\_12.rec  
LAB: sil i t a d a k i m a s u sil  
REC: sil i t a r a q k i p a s u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_12\_01.lbl vs MHT\_SB3\_12\_01.rec  
LAB: sil k i z h i t s u t o sil  
REC: sil p i z h i t s u t o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_12\_03.lbl vs MHT\_SB3\_12\_03.rec  
LAB: sil k a k u n i n g s a s e t e sil  
REC: sil k a k u g i n g s a s a t e sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_12\_04.lbl vs MHT\_SB3\_12\_04.rec  
LAB: sil i t a d a k i t a i n o d e s u g a sil  
REC: sil i t a d a k i t a i m a d e s u g a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_12\_06.lbl vs MHT\_SB3\_12\_06.rec  
LAB: sil s e n g k j u u h j a k u k j u u z h j u n a n a n e n g n o sil  
REC: sil s e n g k j u h j a k u q k j u z h j u n a n a n a n g r o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_12\_07.lbl vs MHT\_SB3\_12\_07.rec  
LAB: sil h a c h i g a t s u sil  
REC: sil h a c h i g a t s u u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_12\_12.lbl vs MHT\_SB3\_12\_12.rec  
LAB: sil k o k u s a i sil  
REC: sil k o k u s a i i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_12\_14.lbl vs MHT\_SB3\_12\_14.rec  
LAB: sil j u u sil  
REC: sil i j u u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_12\_15.lbl vs MHT\_SB3\_12\_15.rec  
LAB: sil k o t o d e sil  
REC: sil k o t o o d e sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_12\_16.lbl vs MHT\_SB3\_12\_16.rec  
LAB: sil j o r o s h i i d e s h j o o k a sil  
REC: sil j o r o s h i i d e s h j o k a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_14\_09.lbl vs MHT\_SB3\_14\_09.rec  
LAB: sil s h j u d a n g g a sil  
REC: sil k e d a g a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_15\_02.lbl vs MHT\_SB3\_15\_02.rec  
LAB: sil k i t a o o z h i e k i m a d e sil  
REC: sil k i t a o u z h i e k i m a d e sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_15\_04.lbl vs MHT\_SB3\_15\_04.rec  
LAB: sil s o k o k a r a sil  
REC: sil s o k o k u r a w a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_15\_07.lbl vs MHT\_SB3\_15\_07.rec  
LAB: sil b a s u k a sil  
REC: sil b a s u t a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_15\_09.lbl vs MHT\_SB3\_15\_09.rec  
LAB: sil t a k u s h i i o sil  
REC: sil t a k u s h i i j o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_16\_01.lbl vs MHT\_SB3\_16\_01.rec  
LAB: sil k j o o t o e k i k a r a sil  
REC: sil k j o o t o e h j a r a sil

Aligned transcription: ../L\_DSB/MHT\_SB3\_16\_03.lbl vs MHT\_SB3\_16\_03.rec  
 LAB: sil k o k u s a i sil  
 REC: sil k o k u s a i i sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_16\_05.lbl vs MHT\_SB3\_16\_05.rec  
 LAB: sil zh i k a ng w a sil  
 REC: sil zh i k a ng g u w a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_16\_06.lbl vs MHT\_SB3\_16\_06.rec  
 LAB: sil d o n o sil  
 REC: sil d o o n o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_16\_08.lbl vs MHT\_SB3\_16\_08.rec  
 LAB: sil k a k a r i m a s u k a sil  
 REC: sil t a k a r i m a s u k a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_17\_02.lbl vs MHT\_SB3\_17\_02.rec  
 LAB: sil b a s u o sil  
 REC: sil b a s u n o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_17\_04.lbl vs MHT\_SB3\_17\_04.rec  
 LAB: sil b a a i sil  
 REC: sil b a i sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_17\_05.lbl vs MHT\_SB3\_17\_05.rec  
 LAB: sil m a ch i zh i k a ng o sil  
 REC: sil m a ch i zh i k a ng g o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_17\_06.lbl vs MHT\_SB3\_17\_06.rec  
 LAB: sil i r e t e sil  
 REC: sil i d e t e sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_17\_08.lbl vs MHT\_SB3\_17\_08.rec  
 LAB: sil t e e d o sil  
 REC: sil t e d o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_18\_01.lbl vs MHT\_SB3\_18\_01.rec  
 LAB: sil k a i zh j o o sil  
 REC: sil k a i zh j o o o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_18\_02.lbl vs MHT\_SB3\_18\_02.rec  
 LAB: sil h u k i ng n o sil  
 REC: sil z u k i n o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_18\_03.lbl vs MHT\_SB3\_18\_03.rec  
 LAB: sil ch i z u j a sil  
 REC: sil ch j u z u i a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_18\_06.lbl vs MHT\_SB3\_18\_06.rec  
 LAB: sil zh i k o k u h j o o w a sil  
 REC: sil zh i k u o k a h j o o w a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_18\_08.lbl vs MHT\_SB3\_18\_08.rec  
 LAB: sil j o o i s a r e t e sil  
 REC: sil j o o i i s a n a t e sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_18\_09.lbl vs MHT\_SB3\_18\_09.rec  
 LAB: sil i r u d e sh j o o k a sil  
 REC: sil i r u d e sh j o k a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_19\_01.lbl vs MHT\_SB3\_19\_01.rec  
 LAB: sil s a ng k a sil  
 REC: sil s a ng k a a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_19\_08.lbl vs MHT\_SB3\_19\_08.rec  
 LAB: sil k o k u s a i sil  
 REC: sil k o k u s a i i sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_19\_10.lbl vs MHT\_SB3\_19\_10.rec  
 LAB: sil k o o t s u u r o n o sil  
 REC: sil k o o t s u u r o o n o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_19\_11.lbl vs MHT\_SB3\_19\_11.rec  
 LAB: sil r j a k u z u j a sil  
 REC: sil r j a k u z u i j a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_19\_12.lbl vs MHT\_SB3\_19\_12.rec  
 LAB: sil zh i k o k u h j o o g a sil

REC: sil zh i k o q k j o o g a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_19\_13.lbl vs MHT\_SB3\_19\_13.rec  
 LAB: sil h u k u m a r e t e sil  
 REC: sil h u k u m a r a t e sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_20\_02.lbl vs MHT\_SB3\_20\_02.rec  
 LAB: sil k o k u s a i k a i g i n i sil  
 REC: sil k o k u s a i i q k a i g i n i sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_20\_03.lbl vs MHT\_SB3\_20\_03.rec  
 LAB: sil r o n g b u n g o sil  
 REC: sil r o n g b u n g n o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_20\_04.lbl vs MHT\_SB3\_20\_04.rec  
 LAB: sil h a q p j o o s h i t a i n g d e s u g a sil  
 REC: sil h a q k j o s h i t a i n g d e s u g a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_20\_05.lbl vs MHT\_SB3\_20\_05.rec  
 LAB: sil k e n g k j u u sil  
 REC: sil k e n g k i j u sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_20\_06.lbl vs MHT\_SB3\_20\_06.rec  
 LAB: sil b u n g j a w a sil  
 REC: sil b u n g i a b a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_20\_07.lbl vs MHT\_SB3\_20\_07.rec  
 LAB: sil t s u u j a k u sil  
 REC: sil t s u j a k u sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_20\_08.lbl vs MHT\_SB3\_20\_08.rec  
 LAB: sil d e n g w a n i sil  
 REC: sil d e n g w a r i sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_20\_09.lbl vs MHT\_SB3\_20\_09.rec  
 LAB: sil k a g i r a r e r u n o d e s h j o o k a sil  
 REC: sil k a i r a r e r u n o d e s h j o k a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_21\_01.lbl vs MHT\_SB3\_21\_01.rec  
 LAB: sil i i e sil  
 REC: sil i i e n g sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_21\_07.lbl vs MHT\_SB3\_21\_07.rec  
 LAB: sil o n g s e e sil  
 REC: sil p o n g s u e e sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_21\_09.lbl vs MHT\_SB3\_21\_09.rec  
 LAB: sil g e n g g o sil  
 REC: sil g e n g n o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_21\_11.lbl vs MHT\_SB3\_21\_11.rec  
 LAB: sil k i k a i sil  
 REC: sil k i k a i i t s u sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_21\_12.lbl vs MHT\_SB3\_21\_12.rec  
 LAB: sil h o n g j a k u n a d o n o sil  
 REC: sil h o u n g j a k u n a d o n o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_21\_14.lbl vs MHT\_SB3\_21\_14.rec  
 LAB: sil b u n g j a o sil  
 REC: sil b u n g j o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_21\_15.lbl vs MHT\_SB3\_21\_15.rec  
 LAB: sil h u k u n g d e sil  
 REC: sil h u k o u n g d e sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_22\_04.lbl vs MHT\_SB3\_22\_04.rec  
 LAB: sil n i h o n g g o t o sil  
 REC: sil h j u h o n g g o t o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_22\_05.lbl vs MHT\_SB3\_22\_05.rec  
 LAB: sil e e g o t o sil  
 REC: sil e e n g m o t o sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_22\_07.lbl vs MHT\_SB3\_22\_07.rec  
 LAB: sil k o t o d e s u g a sil  
 REC: sil k o t o o d a s u g a sil  
 Aligned transcription: ../L\_DSB/MHT\_SB3\_22\_09.lbl vs MHT\_SB3\_22\_09.rec

LAB: sil d o o sil  
REC: sil d o o o sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_22\_11.lbl vs MHT\_SB3\_22\_11.rec  
LAB: sil i m i d e sh j o o k a sil  
REC: sil i m i d e sh j o k a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_23\_02.lbl vs MHT\_SB3\_23\_02.rec  
LAB: sil h a q p j o o w a sil  
REC: sil h a q k j o o w a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_23\_03.lbl vs MHT\_SB3\_23\_03.rec  
LAB: sil n i h o ng g o n i sil  
REC: sil n i h o ng g o r i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_23\_07.lbl vs MHT\_SB3\_23\_07.rec  
LAB: sil n a q t e sil  
REC: sil n a t e sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_23\_11.lbl vs MHT\_SB3\_23\_11.rec  
LAB: sil h a q p j o o w a sil  
REC: sil h a q p i j o o w a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_23\_12.lbl vs MHT\_SB3\_23\_12.rec  
LAB: sil e e g o n i sil  
REC: sil e e g o r i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_23\_13.lbl vs MHT\_SB3\_23\_13.rec  
LAB: sil h o ng j a k u s a r e m a s u sil  
REC: sil h o i j a q s a r e m a s u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_24\_02.lbl vs MHT\_SB3\_24\_02.rec  
LAB: sil j o k o o sh j u u n i sil  
REC: sil j o k o o sh j u n i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_24\_03.lbl vs MHT\_SB3\_24\_03.rec  
LAB: sil t s u i t e w a sil  
REC: sil t s u i t e b a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_24\_05.lbl vs MHT\_SB3\_24\_05.rec  
LAB: sil e e g o d e sil  
REC: sil e e ng g o d e sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_24\_06.lbl vs MHT\_SB3\_24\_06.rec  
LAB: sil sh i q p i t s u sh i t e sil  
REC: sil sh i q t s u e sh i t e sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_24\_07.lbl vs MHT\_SB3\_24\_07.rec  
LAB: sil i t a d a k i m a s u sil  
REC: sil i t a d a q p a s u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_25\_01.lbl vs MHT\_SB3\_25\_01.rec  
LAB: sil s o o d e s u k a sil  
REC: sil s o d a s u k a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_25\_03.lbl vs MHT\_SB3\_25\_03.rec  
LAB: sil d e w a sil  
REC: sil d e a w a sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_25\_04.lbl vs MHT\_SB3\_25\_04.rec  
LAB: sil m o o sh i k o m i sil  
REC: sil m o sh i k u w a m i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_25\_07.lbl vs MHT\_SB3\_25\_07.rec  
LAB: sil j o r o sh i k u sil  
REC: sil j o r a sh i k u sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_25\_08.lbl vs MHT\_SB3\_25\_08.rec  
LAB: sil o n e g a i sil  
REC: sil o n a i g a i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_25\_09.lbl vs MHT\_SB3\_25\_09.rec  
LAB: sil m o o sh i sil  
REC: sil m o sh i sil  
Aligned transcription: ../L\_DSB/MHT\_SB3\_25\_10.lbl vs MHT\_SB3\_25\_10.rec  
LAB: sil a g e m a s u sil  
REC: sil a i g e m a s u sil

----- Overall Results -----

PHRASE: %Correct=51.25 [H=143, S=136, N=279]

PHONE: %Corr=95.54, Acc=92.47 [H=2701, D=54, S=72, I=87, N=2827]

---