

低相異度単語リストの構成と音声認識性能の研究*

5Q-7

小杉好人 江袋林蔵 鈴木眞一郎†
足利工業大学工学部‡

1. はじめに

単語間の音素数の相異総数を相異度指標 I といい、 $I \leq 2$ の範囲に誤りのマジョリティーが存在する。ある実験では、一単語リストの $I \leq 2$ の範囲の単語対が 2% 程度の分布であっても誤りの 80% 以上がここから得られていた。これを更に確認するために、新たに低相異度単語対のみの単語リストを作成し実験し新たな知見を得たので報告する。

2. 相異度法

2.1 相異度法

[原理]相異度法とは、一对の単語対の音素相異数で単語対の相異の程度を定量化する方法の一つである。これまでの多くの実験から、この相異度と誤りとの間に相関が見いだされ、これにより音声認識や人の受聴特性などの分析評価が行われてきている。

一对の単語の相異の大きさ Dw は、一方にのみ音素が存在する空音素対数 J と、相異する音素対の数 K とで定義し、次のように表す。

$$Dw = [J, K] \quad (1)$$

相異度指標 I は一単語対における相異音素対の総数で定義し、しき (1) より次のように表される。

$$I = J + K \quad (2)$$

相異度法の基本的な計算例を次に示す。

やまじ Y A M A Z I

かわ K A W A * *

↓ ↓ ↓ ↓ J:空音素対

K K J J K:異音素対

式 (1) および (2) より、

$$Dw = [J, K] = [2, 2], I = J + K = 4.$$

*A study on the voice input performances regarding the low disparity word list

† Yoshito Kosugi, Rinzou Ebukuro, Shinichiro Suzuki

‡ Ashikaga Institute of Technology

(相異度自動計算のためのソフト DISPA が開発済み
以下の計算は全てこれによって行われている)

2.2 相異度指標と誤りの関係

認識誤りの相異度指標分布は、図 1 に示すように、機械認識で繰り返し誤る単語対や、人の受聴誤りにおいては、これが、相異度指標 $I \leq 2$ の範囲におよそ分布する傾向があり、このような領域を低相異度領域といい、また $I \geq 6$ の高相異度領域には機械認識の誤りが曲線の傾斜に応じて分布する。その中間の領域を相異度中間領域と言っている。

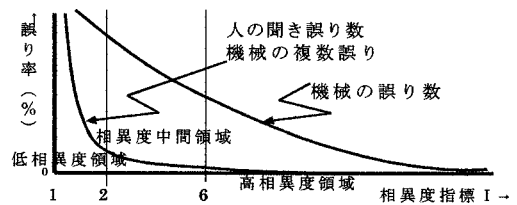


図 1 相異度指標と誤り率の関係

[相異度指標分布]図 1 に示ように、誤りのマジョリティーは、図中 $I \leq 2$ の低相異度領域から生成されている。しかし、一般の単語リストの単語対相異度指標分布は、これまでに調べた範囲では、およそ $I=5$ 前後をピークにしており、低相異度領域に分布するものは数%程度に過ぎない。これまでに、低相異度単語対を意図的に挿入した実験でも、それを除去した単語リストより大きな誤り率であることが確認されていてこれにより、また、単語リストの相異度仕様が、音声認識性能に大きく関わる事が考えられている。

[低相異度単語対リスト]ここでは、これをさらに確認するために、表 1 に示すように、全ての単語の組み合わせが $I \leq 2$ であるような単語リスト作り、これにより音声認識を用いて、表 1 のような単語リストにより認識実験を行った。

3 調査

3.1 調査項目

低相異度単語対リストによる音声入力性能調査

表1 低相異度領域のみ単語リスト (LD・WL)

No	単語	ヨミガナ	No	単語	ヨミガナ
1	体躯	タイク	11	大気	タイキ
2	胎児	タイジ	12	太守	タイシュ
3	大儀	タイギ	13	大愚	タイグ
4	大社	タイシャ	14	対語	タイゴ
5	大拳	タイキョ	15	対座	タイザ
6	大尉	タイイ	16	太祖	タイソ
7	大火	タイカ	17	大河	タイガ
8	太古	タイコ	18	大旨	タイシ
9	大佐	タイサ	19	大家	タイケ
10	急情	タイダ	20	大所	タイショ

3.2 調査方法

表1の単語リストを、5人の被験者が単語1~20まで1人1読み上げパスでDATにて録音しそれを再生して認識させた。これを1人5回ずつ繰り返した。使用した音声認識装置は、A社製大語彙不特定話者認識装置であったが、これを、表1:20単語に限定し、元来の機能どおりのエンロールメントなしで認識させた。

4 調査結果

(1) 単語リストの検証

DISPAにより相異度指標Iを計算した結果、すべての単語対は相異度指標 $I \leq 2$ であった。

(2) 音声入力テスト

発声者別誤り率を図2に示す。

5. 検討

全単語の平均誤り率は、発声者1~5でそれぞれ、27%、47%、53%、40%、および、49%で、別に5人の被験者により確認のために行った、上述2.2[相異度指標分布]に述べた通常の相異度指

標分布を持つ32単語による実験結果の誤認識率0.63%に比較して格段に大きいことがわかる。発声者別には、1と2、3などの間に有意の差が見られるが、2と3との間には見られない。最大で26%の開きがあった。

5人が発声した全単語の平均は43.2%である。単語別にみると、20単語中、すべての発声者が1度も誤らない単語はなかった。

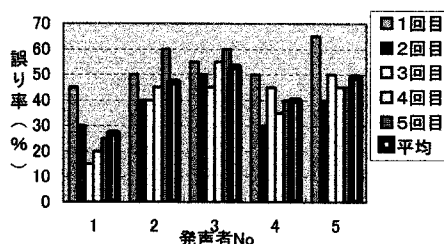


図2 発声者別誤り率

6. 考察

【単語リスト】今回作成した単語リストは、意味のある3モーラ単語のみ、しかも3モーラ目のみを変更したが、1モーラ目、2モーラ目の音素を変更し、意味の分かりにくいような低相異度領域単語リストを作成し実験を行う必要がある。

【音声入力テスト】

(1) 予想通り、音声認識性能は、単語相異度に強く依存することがわかった。

(2) 発声者により、誤り率に26%もの差があるが、通常の数%程度の誤りには当然見られない現象である。この個人差が統計的なものか、比較的固定的なものかは、今後の調査に待つ必要がある。

(3) すべての単語が誤りとなったことは、これまでの実験では見られなかったことである。これによっても、単語リストが誤りに関与することが考えられる。

7. 結論

低相異度単語リストでは、誤り率が格段に大きい、個人差が大きく見られる、すべての単語が少なくとも1回は誤る、ということがみられた。