

二者択一型日本語音声了解度試験方法の検討*

近藤和弘^{*1} 泉 良^{*1,†} 藤森雅也^{*1,††}
加賀 類^{*1} 中川清司^{*1}

【要旨】 未熟な被験者でも安定した単語了解度を得ることができる試験方法 DRT の日本語への適用を検討した。DRT は語頭 1 音素のみ異なる単語対の候補の内一方を被験者に聴取させ、単語対の内一方を選ばせる了解度試験方法である。まず日本語音素を英語 DRT の特徴分類にあわせて分類を試み、語頭の音素が各特徴の内 1 つのみ異なり、他の音素は同じである 2 モーラ単語対リストを提案した。単語親密度の影響を調べたところ、単語対より正答を選ばせる DRT でも、自由に単語を記述させる従来の了解度試験方法と同様に、厳密には親密度の影響があることが分かった。各種雑音を混入させたサンプルを用いて評価を行ったところ、英語 DRT と良く似た傾向が確認された。

キーワード 主観音質評価, 了解度, 音素特徴, ミニマルペア単語試験, 親密度

1. はじめに

近年の無線通信システムの発展や普及等により、今まででは考えられなかった環境で音声通信が行われるようになった。また高エネルギー音声符号化方式や、各種ノイズ・キャンセラ、エコー・キャンセラの導入により今まででは存在しなかった人工的な妨害音の混入が見られるようになり、ますます普遍的な音質の評価が困難となっている。主観音質の評価には音の「自然さ」等受聴品質を総合的に計測する Mean Opinion Score (MOS) などに代表される評価と、受聴音の分かりやすさを評価する明瞭度・了解度試験があるが、ここでは後者を扱う。

従来通信における音質劣化要因は帯域幅、ノイズ等比較的単純なものであり、その評価も比較的単純な手順で行うことができた。一般にはランダムに配置した 1 音節、あるいは無意味な 2 ないしは 3 連音節を被験者に聴取させ、聞こえたと思われる音を答えさせることが行われた [1]。しかしながら日本語には 100 種の音節が存在する。よって、特に妨害の多い環境下では、単音節明瞭度評価でも 100 のうち 1 つを選ぶ必要があり、訓練された受聴者にとってさえも困難な作業である。2 連音、3 連音になると更に 1 万、100 万の可能な

音節組み合わせから 1 つを選ぶ必要があり、ますます困難な作業となる。このため、音節明瞭度試験は一般に被験者の訓練を必要とし、安定性が比較的悪く、また劣悪環境下では必ずしも受聴音の明瞭度を反映していないとされている。無意味音節ではなく単語リストを用いる単語了解度試験もあるが [2]、単語に対する親密度が結果に影響することが知られている [3]。

英語の明瞭度試験でも同様の傾向が報告されている。これに対し、House 等は語頭、ないしは語尾の 1 音素のみ異なるミニマルペア 6 単語を 1 セット、計 50 セットよりなる単語リストを用意し、被験者にセットの中から 1 単語を聞かせ、セット内 6 単語のうちから選ばれる Modified Rhyme Test (MRT) を提案した [4]。6 単語のうちから選ぶため、比較的簡単な作業であり、訓練が不要とされる。また単語の親密度もそれほど影響しないとされている。さらに Voiers は語頭のみ異なるミニマルペア 2 単語を 1 セット、計 96 セットより成る Diagnostic Rhyme Test (DRT) を提案した [5,6]。DRT では各セットにおいて対比させる語頭の音素が音素特徴空間内の特定の要素のみ異なるように吟味されている。よって音素特徴別の了解度が評価できる。MRT, DRT 共に米国では標準化され、広く用いられている [7]。

本論文ではまず日本語音素の特徴分類を試み、これを元に特徴毎に先頭音素のみ異なるミニマルペア単語対より構成される日本語 DRT 単語対リストを提案する。ついで DRT のような被験者に二者択一を行わせる試験方法に対する親密度の影響について調べるため、簡単な実験を行った。その結果、選択肢が二者択一の様

* On a Two-to-one Selection Based Japanese Speech Intelligibility Test,
by Kazuhiro Kondo, Ryo Izumi, Masaya Fujimori,
Rui Kaga and Kiyoshi Nakagawa.

^{*1} 山形大学工学部

[†] (株) YCC 情報システム

^{††} (株) トヨタコミュニケーションシステム

(問合せ: 近藤 和弘 kondo@yz.yamagata-u.ac.jp)
(2006年5月30日受付, 2006年11月11日採録決定)

に極端に少ない場合でも、厳密には親密度の影響が見

られることが判明した。更に白色雑音、マルチトーカーノイズ、および擬似音声雑音の3種類の雑音を異なったレベルで混入させたサンプルを用いてDRTを行ったところ、白色雑音では音素特徴によりノイズの正答率への影響に差があるのに対し、他の雑音では音素特徴間に有意差が見られなくなり、英語DRTと似た傾向を示すことが判明した。以下、提案日本語DRTについて簡単に述べ、次に親密度の影響について実験を通して考察する。次いで提案試験単語リストを各種ノイズ混入時のサンプルで評価し、最後にまとめと今後の課題について述べる。

2. 日本語用 Diagnostic Rhyme Test の概要

2.1 Diagnostic Rhyme Test

Diagnostic Rhyme Test (DRT) はいくつかの仮定と簡略化を行うことにより、訓練されていない被験者を用いても効率よく理解度を測定できることを目的として開発された。DRTでは以下を仮定している。

- 1) 加算雑音や伝送歪等の妨害の影響は比較的エネルギーの小さい子音に限定でき、母音はほとんど影響を受けない。また、子音が言語的情報の大部分を伝達し、よって音声通信ではより重要な役割を担う。これはFairbanksらのRhyme Testの根拠にもなっている[8]。
- 2) 語頭1音素のみを用いて明瞭度を評価しても、他の位置の音素明瞭度も計ることができる。すなわち語頭音素の明瞭度と他の位置の明瞭度は妨害に対し同等の傾向を示す。たとえば[9]では、鈴木らは第1モーラと第2モーラの正答率に相関があることを示している。
- 3) 被験者に選ばせる候補数を極端に限定することにより単語の親密度の影響(3.1参照)、コンテキスト(音素、文脈)の影響[6, 10]を排除できる。

以上の仮定に基づいて、DRTでは語頭の1音素のみ異なるミニマルペアより一つを選ばせる構成になっている。候補内には必ず正答が含まれている。また効率的・系統的な評価を行うために、語頭の1音素にはJacobson, Fant, Halleによる音素特徴分類[11]のうち、1特徴のみ異なる音素対を用いている。

2.2 日本語子音の特徴分類

表-1に英語音素分類に倣った日本語音素分類表を示す。この分類法はJacobson, Fant, Halle(以下JFH)による英語音素の特徴分類に基づいている。表内では、各音素について特定音素特徴のあることを“+”, ないことを“-”, どちらもないことを“0”で示している。7つの音素特徴としては以下が用いられている。

- 1) Voicing: JFH分類ではvocalic-novocalicに相当する。有声-無声分類のことであり、比較的明解な特徴分類である。
- 2) Nasality: JFH分類ではnasal-oralに相当する。鼻韻性のことであり、やはり比較的明解である。
- 3) Sustention: JFH分類ではcontinuant - interruptedである。持続性のある音と、それ以外(破裂音、破擦音)に相当し、明解な分類である。ただし両鼻韻性子音[m][n]が“-”と分類されるのは意外であるが、口腔内の空気の流れが阻害される(鼻腔に抜ける)と言う意味では持続性がないとされるようである。
- 4) Sibilant: JFH分類ではstrident-mellowに相当する。波形の不規則性に関する分類であり、不規則な音素をstridentとして“+”と分類し、規則的な音素をmellowとして“-”と分類する。破裂音がすべてmellow(“-”)であるのに破擦音がstrident(“+”)であること、[s]がstridentで[θ]がmellowであると分類されることなど意外な点が多く、明確な分類ではない。
- 5) Graveness: JFH分類ではgrave-acuteであり、抑音と鋭音に相当する。スペクトル上のエネルギーが低周波に集中すれば前者、高周波に集中すれば後者とされる。また発声時、口腔内の体積が大きければ前者、口腔内が舌により細かく分割されていれば後者とされる。
- 6) Compactness: JFH分類ではcompact-diffuseに相当する。スペクトル上のエネルギーが一つのフォルマント周波数に集中すれば前者、分散すれば後者とされる。また発声時、口腔内前部の体積が狭鎖部を挟んだ後部より大きければ前者、逆なら後者とされる。
- 7) Vowel-like: この特徴は実際には分類には使われないが、glideとそれ以外の子音を区別するのに用いられる。明解な特徴である。

なお、表中、以下の記述に注意されたい。

- 「ラ」行の子音は正確には[r]であるが、ここでは敢えて[r]と表記する。
- [ɸ]は「フ」の子音部分である。
- [ç]は「ホ」の子音部分であるが、ここでは敢えて[h]と区別しないものとする。よってすべての特徴において同一の分類を与えている。

大部分において英語の音素と類似のものは同等の分類を与えた。しかし以下の点は異なる。

- 「ガ」行は語中では鼻韻化し易いが、ここでは特に語頭での特徴を問題にしているため、鼻韻化され難いとしてnasalityは“-”とした。

表-1 日本語音素の特徴分類

子音特徴分類	m	n	z	j	b	d	g	w	r	j	ɸ	s	š	č	p	t	k	h	N	ts	ç
Voicing (vocalic-nonvocalic)	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	+	-	-
Nasality (nasal-oral)	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-
Sustention(continuant-interrupted)	-	-	+	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+	-	-	+
Sibilantion (strident-mellow)	-	-	+	+	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	+	-
Graveness (grave-acute)	+	-	-	0	+	-	0	+	-	0	+	-	0	0	+	-	0	0	0	-	0
Compactness (compact-diffuse)	-	-	-	+	-	-	+	-	-	+	-	-	+	+	-	-	+	+	-	-	+
Vowel-like (glide-nonglide)	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-

- [ŋ] 等の異音 (allophone) は区別していない。

2.3 日本語 DRT 用単語対リスト

まず英語版に倣い、1 特徴につき 16 の単語対候補を挙げた [12]。これを付録の表-A.1 に示す。英文では 8 種の後続母音環境を網羅しているが、日本語では 5 種しかない。よって 10 単語対でも十分であるが、まずは対等に比較できるように英語単語リストと同様に、1 特徴あたり 16 候補を挙げておく。リスト作成に当たっては以下の方針を採った。

- 2 モーラの単語で統一した。これは 2 モーラの単語数が多いため、音素特徴や後続母音環境などの制約を満たす候補数が多いためである。今後必要に応じて 3 モーラ以上も検討する。
- 外来語は極力避けた。ただし「パ」行等は外来語以外の候補が極端に少ないため、その限りではない。
- 単語対内のアクセント型は統一した。これは本試験方法の目的である先頭音素を用いて単語対を識別するのではなく、アクセントを用いて識別されること避けるためである。
- 極力普通名詞を選んだ。固有名詞は極力避けた。これは被験者の背景により親密度等が偏ることを避けるためである。
- 拗音、促音をなるべく避けた。

また、日本語では見られない、あるいはまれな特徴と後続音素の組み合わせもあるので、これは他のもので可能な限り代用している。

予備検討の後、俗語あるいは専門用語と判断された単語を置き換え、必要最低限の 10 単語対に最適化した単語対リストも付録の表-A.2 にあわせて示した [13]。この単語リストの文字音声親密度 [15] (3.1 参照) は平均 5.55、標準偏差 0.72 で、全体的に高い親密度を持つ単語を用いている。アクセント型は 77% が 1 型で、23% が 0 型であった。各音素特徴リストに少なくとも一対の 0 型単語対が含まれている。なお、アクセント型については [15] と [16] を併用して参照し、複数候補がある場合は単語対内で統一が取れる型を採用した。

2.4 DRT の運用と評価

通常複数の話者より録音したサンプルを用いて、8 名以上の被験者に聴取させ、前述のように単語対の内一方を選択させる。呈示順序は全くランダムでも良いし、後続母音が順序良く一巡するように構成しても良い (すなわち 5 種の母音に応じて 5 単語周期)。被験者の正答率を 6 種の音素特徴別、あるいは全回答数の平均で評価する。正答率は以下の調整により偶然性を排除する。

$$S = \frac{100(R - W)}{T} \quad (1)$$

ここに、 S : 調整後正答率 [%], R : 正答数, W : 不正解数, T : 全試行回数である。すなわち、全くでたらめに回答しても二者択一なので $R \approx W$ となり、この時 $S \approx 0$ と成るように原点を調整している。

3. 日本語 DRT の予備検討

3.1 親密度の影響

一般に単語了解度試験においては、単語親密度が大きく影響すると言われている [14]。単語親密度は単語に対する「なじみの程度」を主観的に評価した値であり、天野、近藤は 8 万語について 32 名の被験者が 7 段階で親密度を評価した平均値をデータベース化している [15]。一方、DRT や 6 単語候補グループを用いる MRT のような少数の候補の中から聞いたと思われる単語を選ばせる、いわゆる closed-set 試験では、「単語リストの習熟」の影響が見られないとされている [4, 17]。しかし、厳密な親密度の影響が示されていない上、日本語のデータは存在しない。そこで天野等の親密度データを用いて DRT のような closed-set 試験への親密度の影響を定量的に評価する [18]。なお、親密度には音声、文字およびその組み合わせに対する親密度が各々定義されているが、DRT においては音声および文字が呈示されるので、文字音声単語親密度を親密度として用いる。

3.1.1 実験方法

まず 2 種類の親密度グループを定義した。高親密度

グループを親密度 6 以上, 低親密度グループを親密度 4 以下の単語とした。親密度 4 から 6 の単語は, 高親密度と低親密度を明確に分けるため敢えてはずした。以下の 2 種類の単語リストを用意した。いずれも DRT と同様に先頭子音のみ異なる単語によりグループを作り, グループ内の 1 単語により刺激し, 同一グループ内の全単語を呈示し選ばせる closed-set 試験を想定している。

- a) 単語対リスト: DRT と同様に 1 グループ 2 単語, 31 グループよりなるリストであるが, グループ内の 1 単語は低親密度, 他方は高親密度グループに属する。全単語 2 モーラである。低親密度グループの平均親密度と標準偏差は 3.39 および 1.00, 高密度グループの平均と標準偏差は 6.23 および 0.23, 全体では平均 4.81 および標準偏差 1.60 となった。
- b) 4 単語グループリスト: 1 グループ 4 単語, 9 グループよりなるリストであるが, 2 単語は低親密度, 2 単語は高親密度グループに属する。該当する親密度の単語が不足したため, 2 グループは 3 モーラ, 残り 7 グループは 2 モーラの単語よりなる。低親密度グループの平均親密度と標準偏差は 2.74 および 0.81, 高密度グループの平均と標準偏差は 6.26 および 0.28, 全体では平均 4.50 および標準偏差 1.88 となった。

上記リストの単語音声日本語を母国語とする非訓練話者男性 1 名 (20 代) 女性 1 名 (50 代) を用いて収録した。この時, 単語文字を呈示するとともに読みもあわせて呈示し, 標準語で読むように指示した。このときはアクセントの指示は行っていない。収録後指定アクセントであることを検聴し, 異なる場合のみアクセントを指示して再収録を行なった。収録音声に白色雑音を SNR が -15, -10, 0, 10[dB] となるように加えた。このサンプルを用いて以下の 3 種類の了解度試験を行った。

- 1) 2 単語 RT(2-word RT): 通常の DRT と同様にグループ内単語音声で刺激する。音声再生を終了直後に図-1 に示すようなダイアログボックスをディスプレイ上に表示し, その単語が属するグループの 2 単語を呈示して, 正答をマウスを用いて選ばせる。選択後, “Next” ボタンを押すことで次の音声刺激の再生を開始する。この試験では, 前記 a) の単語対リストを用いる。
- 2) 4 単語 RT(4-word RT): 前記 b) の 4 単語グループリストを用いて, グループ内単語音声で刺激し, グループ内全 4 単語を呈示して正答を選ばせる。図-2 に示すダイアログボックスを表示し,



図-1 2-word RT の応答入力ダイアログボックス例

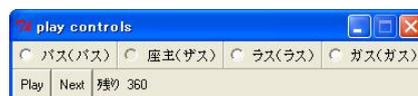


図-2 4-word RT の応答入力ダイアログボックス例



図-3 Intell. Test の応答入力ダイアログボックス例

応答を選択する。やはり選択後, “Next” ボタンを押すことで次の音声刺激の再生を開始する。

- 3) 単語了解度試験 (Intell. Test): 前記 a) の単語対リスト内の単語を用いて従来型 (open-set) の単語了解度試験を行う。すなわち, リスト内単語音声で刺激し, 正答を自由に記入させる。図-3 に示すダイアログボックスのテキストボックスに全角カタカナで入力する。テキストはカナモードに固定して, ローマ字入力する。入力後, “Next” ボタンを押すことで次の音声刺激の再生を開始する。

聴取試験者は日本語を母国語とする 20 代男子学生 10 名とした。各被験者は 3 種類の試験方法を全て行なっているが, その順番はランダムである。また, 各試験内の試験音声呈示順も被験者毎にランダムに並べ替えた。また, 2-word RT, 4-word RT については選択枝の呈示順も呈示毎にランダムに並べ替えた。各試験とも一つの試験音に対する応答が完了した時点で次の試験音を呈示するようにした。表-A.3, 表-A.4 にここで用いた親密度別単語グループリストの一部を示した。収録は全てヘッドマウント型エレクトレットマイクで行い, 再生はヘッドホンよりダイオティック刺激として両耳に呈示した。呈示音圧は被験者毎に試験音が十分可聴となるレベルに調整させた。なお, 一般的には音圧は了解度に影響するとされるが, 了解度に影響を与える受音量は被験者の聴力差が問題となる極端に低い音量であることが知られている [19]。よってこの試験での呈示音圧は了解度に影響を与えない範囲であると考え。提案試験方法が使用者が音量を調整できる機器の了解度を測ることを目的としていること

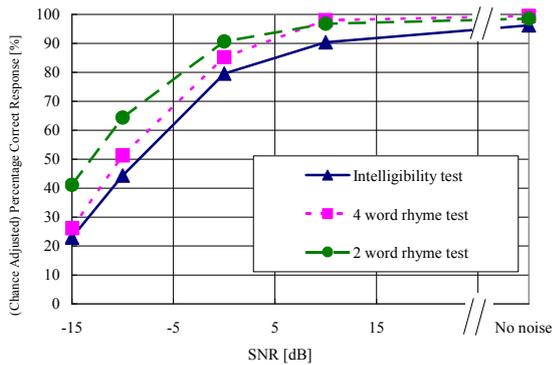


図-4 3種の試験方法による正答率

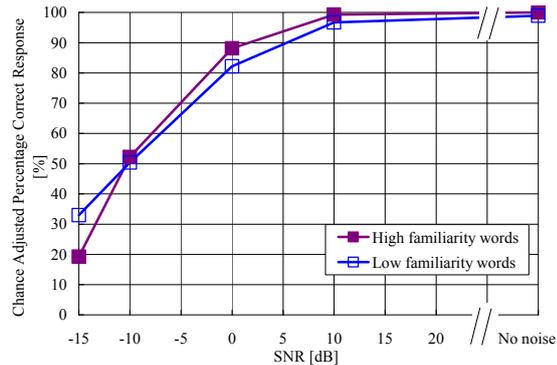


図-6 4-word RT 親密度グループ別正答率

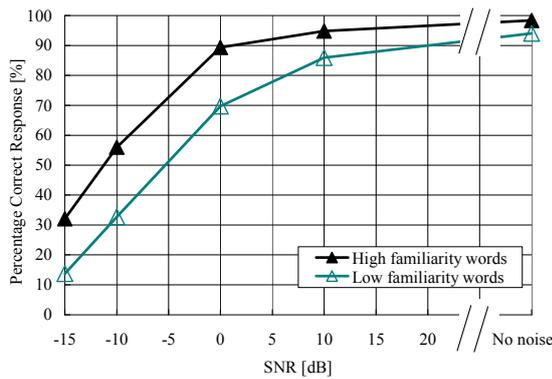


図-5 Intell. Test 親密度グループ別正答率

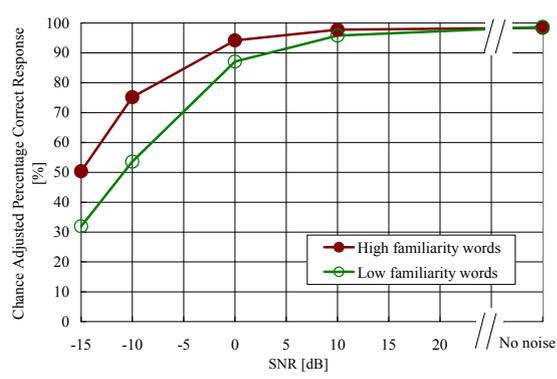


図-7 2-word RT 親密度グループ別正答率

からも、実用に近い測定が行われていると考える。呈示音圧が低い場合の了解度の状況については別の検討に譲ることとする。

3.1.2 親密度の影響に関する考察

図-4 に前記 3 試験方法による了解度評価結果を示す。また、図-5 には Intell. Test, 図-6 には 4-word RT, 図-7 には 2-word RT の親密度グループ別了解度を示す。ここで、Intell. Test では正答率は単純な正答数と全回答数の比で表している。2-word RT では前記式 (1) により、また 4-word RT では 2-word RT と同様に以下の式により偶発性を調整した正答率を用いている。

$$S = \frac{100(R - \frac{W}{3})}{T} \quad (2)$$

ここに、 S : 調整後正答率 [%], R : 正答数, W : 不正解数, T : 全試行回数である。

これらの図から以下のことが考察できる。

- 1) 3種類の試験法の全ての SNR において、了解度は似た傾向は示すものの、2-word RT, 4-word RT, Intell. Test の順になっている。試験方法 (Intell. Test, 4-word RT, 2-word RT の 3 水準) と SNR (5 水準) の 2 要因分散分析の結果、

$F(2, 159) = 47.80, p < .001$ で有意水準 5% で試験方法により有意差が現れている。交互作用は有意 ($F(8, 135) = 5.59, p < .001$) だが、全ての SNR 水準で試験方法の単純主効果は有意であった。

- 2) 全ての試験方法において、高親密度グループの了解度は低親密度グループより高い。親密度グループ (2 水準) と SNR (5 水準) の 2 要因分析では、全ての試験方法において交互作用が有意であった (2 word RT: $F(4, 90) = 5.40, p < .001$, 4-word RT: $F(4, 90) = 3.84, p = .006$, Intell. Test: $F(4, 90) = 8.21, p < .001$)。親密度の単純主効果の検定を SNR 水準ごとに行なったところ、2word RT では SNR が 0 dB, -10 dB, -15 dB の水準で親密度の単純主効果が有意であった。(0 dB で $F(1, 18) = 6.08, p = .024$, -10 dB で $F(1, 18) = 12.30, p = .003$, -15 dB で $F(1, 18) = 8.67, p = .009$) 4-word RT では SNR が +10 dB と -15 dB で主効果が有意であった (+10 dB で $F(1, 18) = 6.78, p = .018$, -15 dB で $F(1, 18) = 5.91, p = .026$)。Intell. Test では全ての SNR 水準で親密度の主効果が有意

であった (No noise で $F(1, 18) = 17.22, p < .001$, +10 dB で $F(1, 18) = 15.50, p < .001$, 0 dB で $F(1, 18) = 47.50, p < .001$, -10 dB で $F(1, 18) = 41.65, p < .001$, -15 dB で $F(1, 18) = 31.66, p < .001$)。

以上から、選択肢の数が少ない場合でも親密度の影響はあり、理解度試験に用いる単語の親密度統制の必要があることが確認できた。

4. 代表的な外乱に対する日本語 DRT の評価

表-A.2 の単語リストを用いて日本語 DRT の評価を行った。まずは 3 種類のノイズを各 SNR で混入し、特徴別理解度の傾向を見た [13]。

4.1 実験方法

日本語を母国語とする非訓練話者を用いて、男性 4 名 (全員 20 代) 女声 4 名 (20 代 3 名, 50 代 1 名) の音声を表-A.2 の全単語について収録した。この時も単語文字と読みは呈示し標準語で読み上げるように支持した。アクセントは指示せず、検聴後問題があった場合のみ指示して再収録した。収録にはヘッドマウント型のエレクトレットマイクを用いて標準化周波数 16kHz, 16 ビットで行なった。これに白色雑音, 男女複数名の音声を混合したマルチトーカーノイズ (babble)[2], 音声帯域に似せた周波数特性をもつ擬似音声雑音 [2] を SNR が -15, -10, 0, 10 [dB] となるように混合した。白色雑音はコンピュータで発生し, 擬似音声雑音は補聴器適合評価用音源 CD, TY-89 [2] より収録したものを, またマルチトーカー・ノイズは Rice 大ホームページ [20] より入手したものを 16kHz にダウンサンプルして用いた。以上の音声をランダムな順序で再生し, 被験者に単語対内の両単語を呈示し, 正答を選ばせた。音声刺激は全てヘッドホンを用いてダイオティック刺激とした。各試験内の試験音声呈示順は被験者毎にランダムに並べ替えた。また, 選択肢の提示順も呈示毎にランダムに並べ替えた。各試験とも一つの試験音に対する応答が完了した時点で次の試験音を呈示するようにした。呈示音圧は被験者毎に試験音が十分可聴なレベルに調整させた。聴取試験者はいずれの雑音についても日本語を母国語とする 20 代学生 10 名とした。各試験者は 8 話者 × 5 ノイズ混入条件 × 6 音素特徴 × 20 単語, 計 4800 単語を評定した。すなわち 1 音素特徴あたり 800 単語の評定を行なっている。

4.2 結果と考察

図-8 にノイズ種別の全音素特徴を含めた平均正答率を示す。また図-9, 図-10, 図-11 にノイズ種ごとの音素特徴別正答率を示す。SNR(5 水準) と音素特徴分類(6 水準) で 2 要因分散分析を行なったところ, 全ての

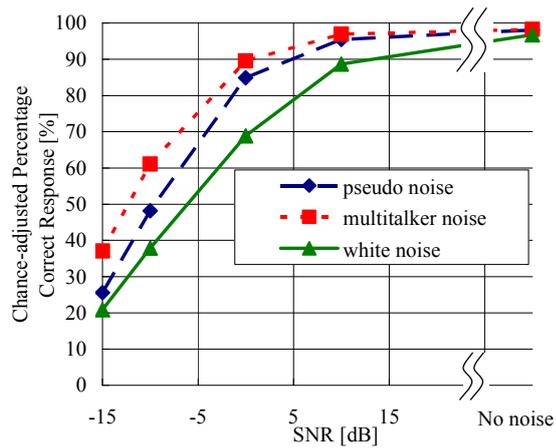


図-8 3 種類の雑音を付加した音声の DRT 評価結果

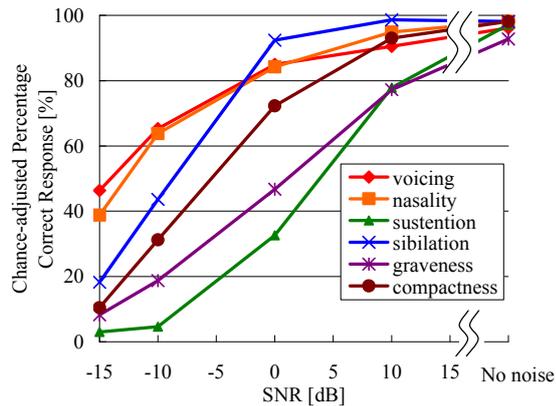


図-9 白色雑音を付加した音声の DRT 評価結果

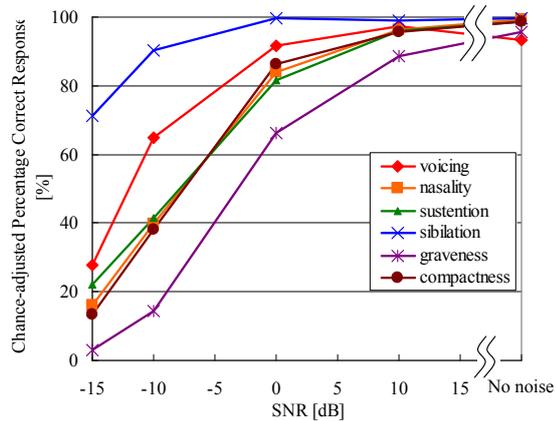


図-10 擬似音声雑音を付加した音声の DRT 評価結果

ノイズ種で SNR, 音素特徴分類の両要因とも有意であった。以下に統計量を示す。

- 白色雑音 : SNR ($F(4, 270) = 864.04, p < 0.001$), 音素特徴分類 ($F(5, 270) = 137.24, p < 0.001$)
- 擬似音声雑音 : SNR ($F(4, 270) = 1824.19, p < 0.001$)

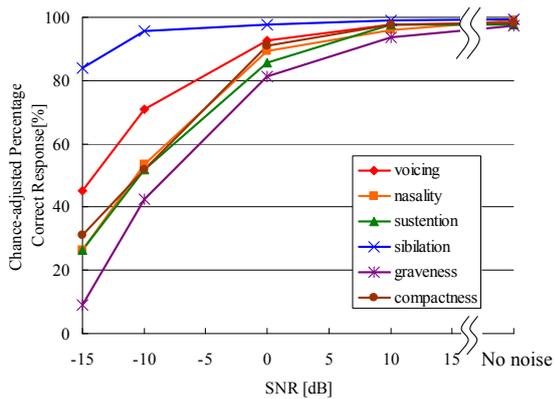


図-11 マルチトーカー・ノイズを付加した音声の DRT 評価結果

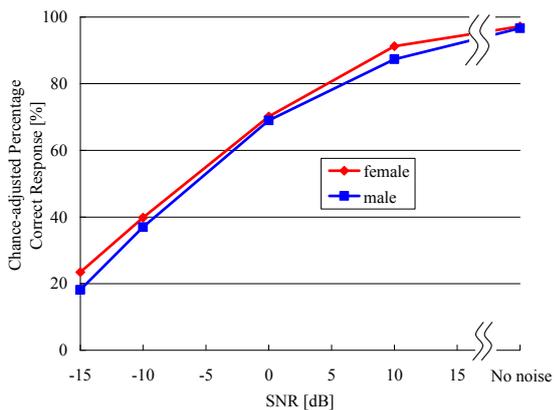


図-12 白色雑音を付加した男女話者音声間の DRT 結果比較

0.001), 音素特徴分類 ($F(5, 270) = 238.23, p < 0.001$)

- マルチトーカー・ノイズ:SNR ($F(4, 270) = 424.627, p < 0.001$), 音素特徴分類 ($F(5, 270) = 53.90, p < 0.001$)

また各水準間の有意差を Tukey 法を用いて有意水準 5% で検定した。SNR に関しては白色雑音では各水準間で有意差があったが、擬似雑音とマルチトーカー・ノイズにおいては”No noise” と ”+10 dB” 以外で有意差が見られた。音素特徴分類に関しては、白色雑音で voicing と nasality 以外で有意差が見られた。一方、擬似雑音とマルチトーカー・ノイズでは nasality と sustention, nasality と compactness, compactness と sustention の組み合わせで有意差が見られなかったが、これ以外の組み合わせでは有意差が見られた。

これらの結果から以下の考察ができる。

- ノイズ種によらず SNR に対し正答率は似た傾向を示す。SNR 5 水準とノイズ種 3 水準による 2

元配置分散分析では、交互作用が有意であった ($F(8, 885) = 5.03, p < 0.001$)。SNR 水準ごとにノイズ種の単純主効果を検定したところ、全ての SNR においてノイズ種の単純主効果は有意であった (SNR -15 dB で $F(2, 177) = 7.48, p < 0.001$, -10 dB で $F(2, 177) = 14.06, p < 0.001$, 0 dB で $F(2, 177) = 27.20, p < 0.001$, +10 dB で $F(2, 177) = 23.99, p < 0.001$, No noise で $F(2, 177) = 3.54, p = 0.031$)。ノイズ種による影響は白色、擬似音声、マルチトーカーの順に大きい傾向にある。

- Sibilant は特に白色雑音を混入した場合、高 SNR では比較的高い正答率を示すが、SNR の低下とともに急激に劣化する。これは Voiers も指摘しているように [6], この特徴が白色雑音に近い特徴のため、加算白色雑音が増加すると特徴の有無が区別がつかなくなるためであろう。白色でない混入雑音の場合はそれほど影響を受けないのも同じ理由からと考えられる。
- マルチトーカーノイズと擬似雑音では正答率の特徴間の差が全体的に見られなくなっている。前記のように、特に nasality, sustention, compactness 間に有意差が見られなくなっている。これはやはり Voiers が英語で観測した傾向と同様である。白色雑音では音素特徴による差が見られる。この傾向の差は加算ノイズ帯域幅、特に高域ノイズ量が影響しているのではないかと考えられる。

また、図-12 に白色雑音を付加した男女話者別音声の DRT 評価結果を示す。この図からもわかるように、DRT 評価に大きな差は見られず、評定音声の話者性別依存性が極めて少ないことが確認された。SNR (5 水準) と話者性別の 2 要因分散分析の結果、交互作用は有意ではなく、かつ話者性別による正答率の有意差は見られなかった ($F(1, 90) = 3.49, p = 0.065$)。

5. む す び

本論文では比較的熟練していない被験者でも安定した了解度を得ることができるとされている Diagnostic Rhyme Test (DRT) について検討した。DRT は語頭 1 音素のみ異なる単語対の候補の内一方を被験者に聴取させ単語対の内一方を選ばせる了解度試験方法である。英語においては検討が進んでおり、一定の成功を収めて米国内標準として確立されている。ここでは日本語でも同等の試験手法を検討した。まず日本語の音素を英語 DRT で用いられている特徴分類にあわせて分類を試みた。次に、語頭の音素が各特徴の内 1 つのみ異なり、かつ他の音素は同じである 2 モーラ単語対

候補リストを提案した。単語リストは親密度が比較的高い単語より選んだ。

次に DRT のような正答を選ぶ選択肢を与える closed-set テストに対する親密度の影響を、従来の自由に正答を記入させる open-set 単語了解度試験と比較した。その結果、選択肢が二者、あるいは二者択一でも親密度の影響は、少なくなる傾向はあるものの、有意に見られることが分かった。よって、了解度試験に用いる単語は、親密度を統制する必要があることを確認した。

提案した日本語 DRT を白色、マルチトーカー、擬似音声の各雑音を混入して正答率を調べた。白色、擬似音声、マルチトーカー雑音の順に雑音の影響が大きかった。これは雑音の帯域幅に関係していることが予想される。また白色雑音では音素特徴により雑音の影響に差が見られるが、他の雑音ではその差が小さいことも分かった。今後は提案した単語リストを用いて更に多くの評定者、ノイズ種、SNR についてデータを積み上げていきたいと考えている。また、ノイズの帯域と各特徴への影響を更に詳細に調べていくことも考えている。また、主要な音声符号化 CODEC の特徴別影響も評価していきたいと考えている。

謝 辞

本編は山形大学工学部電気電子工学科における長年の卒業研究および修士研究の成果をまとめたものである。度々の議論、ご支援頂いた電気電子工学科の中川研究室の学生、職員一同に感謝する。

文 献

- [1] 飯田 茂隆: 明瞭度試験方法について. 音響学会誌, 43, 7, 532 - 536 (1987).
- [2] 田中 美郷: 補聴器適合評価機器の試作に関する研究, 昭和 63 年度科学研究費補助金成果報告書 (1989).
- [3] 坂本 修一, 小澤 賢司, 他: 単語の親密度が了解度に及ぼす影響, 音講論, 399 - 400 (1997).
- [4] A. S. House, C. E. Williams, M. H. L. Hecker and K. D. Kryter: Articulation Testing Methods: Consonantal Differentiation with a Closed-Response Set, J. Acoust. Soc. Am., 37, 1, 158-166 (1965).
- [5] W. D. Voiers: Diagnostic Evaluation of Speech Intelligibility, in Speech Intelligibility and Speaker Recognition, M. E. Hawley, ed., (Dowden, Hutchinson & Ross, PA., 1977) 374-387.
- [6] W. D. Voiers: Evaluating Processed Speech using the Diagnostic Rhyme Test, Speech Technology, 1, 30-39 (1983).
- [7] ANSI Standard S3.2-1989, Method for Measuring the Intelligibility of Speech over Communication Systems, (1989, reaffirmed 1995).
- [8] G. Fairbanks: Test of Phonemic Differentiation: The Rhyme Test, JASA, 30, 596-600 (1958).
- [9] 鈴木 陽一, 近藤 公久, 他: 親密度を統制した単語了解度試験における反応傾向, 信学技報, H-98-47 (1998).
- [10] G. A. Miller, G. A. Heise, and W. Lichten: The Intelligibility of Speech as a Function of the Context

of the Test Materials, Journal of Experimental Psychology, 41, 329-335 (1951).

- [11] R. Jakobson, C. G. M. Fant, M. Halle: Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates, Tech. Rep. No. 13, Acoustics Laboratory, MIT (1952).
- [12] 近藤 和弘, 中川 清司: 新しい日本語明瞭度試験方法の基礎検討, 信学技報, SP2000-37 (2000).
- [13] M. Fujimori, K. Kondo, K. Takano, K. Nakagawa: On a Revised Word-pair List for the Japanese Intelligibility Test, Proc. International Symposium on Frontiers in Speech and Hearing Research (2006).
- [14] 坂本 修一, 鈴木 陽一, 他: 親密度と音韻バランスを考慮した単語了解度試験用リストの構築, 音響学会誌, 54, 12, 842-849 (1998).
- [15] 天野 成昭, 近藤 公久: 日本語の語彙特性, (三省堂, 東京, 1999) .
- [16] NHK 放送文化研究所編: 日本語発音アクセント辞典, (NHK 出版, 東京, 1998) .
- [17] A. S. House, C. Williams, M. H. L. Hecker, and K. D. Kryter: Psychoacoustic Speech Tests: A Modified Rhyme Test, U. S. Air Force Syst. Command, Hanscom Field, Electron. Syst. Div. Tech. Doc. Rept., ESD-TDR-63-403(1963). Defense Technical Information Center (www.dtic.mil) より入手可能.
- [18] 近藤 和弘, 泉 良, 中川 清司: 新しい日本語了解度試験方法の評価, 信学技報, SP2001-163 (2001).
- [19] 渡辺 俊朗: 単語了解度による規則合成音の評価法に関する検討, 信学論, J72-A, 10, 1503-1509 (1989).
- [20] Rice University: Signal Processing Information Base (SPIB), http://spib.rice.edu/spib/select_noise.html.

付 録

A.1 日本語 DRT 単語対リスト

A.1.1 フルリスト

表-A.1 に英語 DRT リストの単語数にあわせたフルリストを示す。単語対の前者は、先頭子音が評価対象音素特徴を含むもの、後者は含まないものである。

A.1.2 縮小リスト

表-A.2 に一部対象にふさわしくない単語を見直し、かつ日本語母音数に応じた必要最低限の単語数とした縮小リストを示す。

A.2 親密度グループ別単語リスト

表-A.3, 表-A.4 に 3.1 で述べた実験に用いた親密度グループ別の単語対, および 4 単語グループリストを示す。

近藤 和弘

昭 57 早大・理工・電子通信卒. 昭 59 同大学院理工学研究科修士課程了. 同年, (株) 日立製作所入社. 同中央研究所で通信音声信号処理システム, およびデジタルビデオ画像符号化システムの研究開発に従事. 平 4 (株) テキサス・インスツルメンツ筑波研究開発センター入社. 平 8 米国テキサス・インスツルメンツ入社, Media Tech. Lab., Member of Technical Staff. 平 11 山形大・工・助教授. 主として通信音声・音響信号処理システムの研究教育に従

表-A.1 日本語 DRT 用単語対 (フルリスト)

Voicing	Nasality	Sustention	Sibilation	Graveness	Compactness
財 - 才 (ザイ - サイ)	万 - 番 (マン - バン)	箸 - 菓子 (ハシ - カシ)	ジャム - ガム (ジャム - ガム)	杵 - 楽 (ワク - ラク)	焼く - 沸く (ヤク - ワク)
抱く - 炊く (ダク - タク)	泣く - 抱く (ナク - ダク)	旗 - 型 (ハタ - カタ)	着 - 角 (チャク - カク)	パイ - タイ (パイ - タイ)	貝 - パイ (カイ - パイ)
害 - 貝 (ガイ - カイ)	梨 - 出汁 (ナシ - ダシ)	掃く - 書く (ハク - カク)	社旗 - 破棄 (シャキ - ハキ)	番 - 段 (バン - ダン)	山羊 - 和議 (ヤギ - ワギ)
議事 - 記事 (ギジ - キジ)	無い - 台 (ナイ - タイ)	四季 - 知己 (シキ - チキ)	酌 - 箔 (シャク - ハク)	バン - 短 (バン - タン)	官 - バン (カン - バン)
銀 - 金 (ギン - キン)	ミス - ビス (ミス - ビス)	私利 - 地理 (シリ - チリ)	陣 - 銀 (ジン - ギン)	バラ - 鱈 (バラ - タラ)	銀 - 瓶 (ギン - ビン)
髓 - 粹 (ズイ - スイ)	ミリ - ビリ (ミリ - ビリ)	紐 - 肝 (ヒモ - キモ)	式 - 引き (シキ - ヒキ)	店 - 偽 (ミセ - ニセ)	金 - ビン (キン - ビン)
図示 - 寿司 (ズシ - スシ)	無理 - 鱈 (ムリ - ブリ)	昼 - 着る (ヒル - キル)	事務 - 義務 (ジム - ギム)	未知 - ニ値 (ミチ - ニチ)	気障 - ピザ (キザ - ピザ)
寓 - 食う (グー - クー)	面 - 弁 (メン - ベン)	粹 - 対 (スイ - ツイ)	知事 - 記事 (チジ - キジ)	剥く - 抜く (ムク - ヌク)	郡 - 文 (グン - ブン)
税 - 生 (ゼイ - セイ)	メタ - ベタ (メタ - ベタ)	好き - 月 (スキ - ツキ)	銃 - 寓 (ジュウ - グー)	無視 - 主 (ムシ - ヌシ)	黒 - プロ (クロ - プロ)
善 - 千 (ゼン - セン)	練る - 出る (ネル - デル)	砂 - 綱 (スナ - ツナ)	中 - 空 (チュウ - クウ)	面 - 年 (メン - ネン)	百合 - 瑠璃 (ユリ - ルリ)
下駄 - 桁 (ゲタ - ケタ)	猫 - 凸 (ネコ - デコ)	変 - 剣 (ヘン - ケン)	純 - 郡 (ジュン - グン)	ベン - 天 (ベン - テン)	弦 - 弁 (ゲン - ベン)
出刃 - 手羽 (デバ - テバ)	年 - 伝 (ネン - デン)	下手 - 桁 (ヘダ - ケタ)	シエア - ヘア (シエア - ヘア)	ベロ - 出る (ベロ - デロ)	剣 - ベン (ケン - ベン)
像 - 僧 (ゾウ - ソウ)	持つ - ボツ (モツ - ボツ)	へり - 蹴り (へり - ケリ)	条 - 号 (ジョウ - ゴウ)	毛 - 脳 (モウ - ノウ)	語気 - 簿記 (ゴキ - ボキ)
続 - 即 (ゾク - ソク)	門 - ボン (モン - ボン)	星 - 腰 (ホシ - コシ)	腸 - 香 (チョウ - コウ)	ボン - 豚 (ボン - トン)	紺 - ボン (コン - ボン)
誤字 - 孤児 (ゴジ - コジ)	退く - 退く (ノク - ドク)	掘る - 凝る (ホル - コル)	書庫 - 矛 (ショコ - ホコ)	凡 - 鈍 (ボン - ドン)	洋 - 劣 (ヨウ - ロー)
ドル - 取る (ドル - トル)	野良 - 銅鑼 (ノラ - ドラ)	補佐 - 濃さ (ホサ - コサ)	所持 - 保持 (ショジ - ホジ)	ポロ - トロ (ポロ - トロ)	余暇 - ろ過 (ヨカ - ロカ)

事. 工博. 信学会会員, IEEE シニアメンバー.

泉 良

平 13 山形大・工・電子情報卒. 同年 YCC 情報システム入社. 在学中日本語音声了解度試験方法の研究に従事.

藤森 雅也

平 16 山形大・工・電気電子卒. 平 18 山形大・理工学研究科修士課程了. 同年 (株) トヨタコミュニケーションシステム入社. 在学中日本語音声了解度試験方法の研究に従事.

加賀 類

平 16 山形大・工・電気電子卒. 平 17 より山形大・理工学研究科修士課程在学. 日本語音声了解度試験方法の研究に従事.

中川 清司

昭 43 阪大・通信卒. 昭 45 同大大学院修士課程了. 同年電信電話公社 (現 NTT) 電気通信研究所入所. ミリ波導波管伝送方式および光ファイバディジタル中継伝送方式の研究実用化に従事. 平 8 より山形大・工・教授. 情報通信システムの研究・教育に従事. 工博 (阪大). 昭 52 年度本学会学術奨励賞, 昭 53, 平 2 英国 IEE 論文賞, 昭 59 および平 3 本学会業績賞. 著書「光増幅器とその応用」(オーム社, 共著) など. IEEE シニアメンバー, 信学会フェロー.

表-A.2 日本語 DRT 用単語対 (縮小リスト)

Voicing	Nasality	Sustention	Sibilation	Graveness	Compactness
財 - 才 (ザイ - サイ)	万 - 番 (マン - パン)	箸 - 菓子 (ハシ - カシ)	ジャム - ガム (ジャム - ガム)	杵 - 楽 (ワク - ラク)	焼く - 沸く (ヤク - ワク)
抱く - 炊く (ダク - タク)	無い - 台 (ナイ - ダイ)	旗 - 型 (ハタ - カタ)	着 - 角 (チャク - カク)	パイ - 鯛 (パイ - タイ)	貝 - パイ (カイ - パイ)
議事 - 記事 (ギジ - キジ)	ミス - ビス (ミス - ビス)	私利 - 地理 (シリ - チリ)	式 - 引き (シキ - ヒキ)	見栄 - 煮え (ミエ - ニエ)	銀 - 瓶 (ギン - ビン)
銀 - 金 (ギン - キン)	見る - ビル (ミル - ビル)	昼 - 着る (ヒル - キル)	知事 - 記事 (チジ - キジ)	ミス - ニス (ミス - ニス)	気障 - ピザ (キザ - ピザ)
髓 - 粹 (ズイ - スイ)	無理 - 鱈 (ムリ - ブリ)	好き - 月 (スキ - ツキ)	中 - 空 (チュウ - クウ)	剥く - 抜く (ムク - ヌク)	黒 - プロ (クロ - プロ)
ぐう - 食う (グウ - クウ)	無視 - 武士 (ムシ - ブシ)	砂 - 綱 (スナ - ツナ)	純 - 群 (ジュン - グン)	無視 - 主 (ムシ - ヌシ)	夕 - ルー (ユウ - ルー)
税 - 生 (ゼイ - セイ)	面 - 弁 (メン - ベン)	変 - 剣 (ヘン - ケン)	シェア - ヘア (シェア - ヘア)	面 - 年 (メン - ネン)	弦 - 弁 (ゲン - ベン)
出歯 - 手羽 (デバ - テバ)	練る - 出る (ネル - デル)	縁 - けり (ヘリ - ケリ)	シエル - 経る (シエル - ヘル)	ペン - 天 (ペン - テン)	剣 - ペン (ケン - ペン)
像 - 僧 (ゾウ - ソウ)	門 - 盆 (モン - ボン)	星 - 腰 (ホシ - コシ)	条 - 号 (ジョウ - ゴウ)	毛 - 脳 (モウ - ノウ)	語気 - 簿記 (ゴキ - ボキ)
誤字 - 孤児 (ゴジ - コジ)	野良 - 銅鑼 (ノラ - ドラ)	掘る - 凝る (ホル - コル)	所持 - 保持 (ショジ - ホジ)	ポロ - とろ (ポロ - トロ)	余暇 - ろ過 (ヨカ - ロカ)

表-A.3 親密度別単語対リスト (抜粋)

低親密度グループ	高親密度グループ
世路 (セロ)	テロ (テロ)
理路 (リロ)	白 (シロ)
ゲル (ゲル)	ベル (ベル)
けら (ケラ)	寺 (テラ)
ぬら (ヌラ)	村 (ムラ)
野良 (ノラ)	虎 (トラ)
摩羅 (マラ)	皿 (サラ)
埤 (トヤ)	小屋 (コヤ)
苦味 (クミ)	罪 (ツミ)
地味 (チミ)	意味 (イミ)
死魔 (シマ)	今 (イマ)
ラマ (ラマ)	生 (ナマ)
寡婦 (カフ)	タフ (タフ)
師父 (シフ)	皮膚 (ヒフ)
流布 (ルフ)	主婦 (シュフ)
レフ (レフ)	シェフ (シェフ)
塗布 (トフ)	オフ (オフ)
実 (サネ)	種 (タネ)
うね (ウネ)	船 (フネ)
悦 (エツ)	熱 (ネツ)
八つ (ヤツ)	夏 (ナツ)
六つ (ムツ)	靴 (クツ)
珍 (チン)	金 (キン)
寸 (スン)	運 (ウン)
伝 (デン)	ペン (ペン)
ノン (ノン)	門 (モン)
吾子 (アコ)	過去 (カコ)
擬古 (ギコ)	事故 (ジコ)

表-A.4 親密度別 4 単語グループリスト

低親密度グループ	高親密度グループ
雨露 (ウロ)	黒 (クロ)
グロ (グロ)	プロ (プロ)
がる (ガル)	猿 (サル)
だる (ダル)	春 (ハル)
不磨 (フマ)	熊 (クマ)
夢魔 (ムマ)	妻 (ツマ)
斬 (ザン)	ガン (ガン)
シャン (シャン)	パン (パン)
座主 (ザス)	ガス (ガス)
ラス (ラス)	パス (パス)
樟 (クス)	ブス (ブス)
繻子 (シュス)	留守 (ルス)
些々 (ササ)	朝 (アサ)
輪さ (ワサ)	傘 (カサ)
下院 (カイン)	サイン (サイン)
ダイン (ダイン)	パイン (パイン)
弛緩 (シカン)	時間 (ジカン)
離間 (リカン)	痴漢 (チカン)