

Title	発話能力と知覚能力の関連についての考察
Author(s)	木場, 祐樹
Citation	
Issue Date	2008-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/4302
Rights	
Description	Supervisor: 党建武, 情報科学研究科, 博士

修 士 論 文

発話能力と知覚能力の関連についての考察

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

木場 祐樹

2008年3月

修 士 論 文

発話能力と知覚能力の関連についての考察

指導教官 党建武 教授

審査委員主査 党建武 教授

審査委員 徳田功 准教授

審査委員 赤木正人 教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

木場 祐樹

2008年2月

概要

音声生成と音声知覚には密接な関係があると考えられており、両者を音声のエンコード/デコードの過程と捉えた様々な実験、提案がされてきた。その代表的なものに発話機能が音声知覚に本質的な役割を果たすとする Liberman らの知覚運動理論が挙げられる。これについても未だ明確な証拠は見つかってはいないものの、近年これを支持するような様々な研究結果が得られて来ている。そこでもしこの仮説が正しいとするなら、非常に自然な考え方として、発話技能が高い人物は同時に高い知覚能力も持っているのではないかと推測できる。そこで本研究では早口言葉を用いて被験者の発話技能を評価し、一方音韻修復現象を用いることで知覚能力を評価し、発話技能と知覚能力の関連性について調査した。その結果発話技術の高い被験者は同時に優れた知覚能力を持っており、また発話の難しい文章は同時に知覚することも難しいという結果が得られた。さらに音声生成時の発話計画能力と知覚時の予測との関連についても明らかにした。

目次

第1章	序論	1
1.1	研究の目的と背景	1
1.2	論文構成	2
第2章	音韻修復現象	3
2.1	本研究のアプローチ	5
2.2	音韻修復現象の特徴と性質	5
第3章	発話能力と知覚能力の相関	5
3.1	音韻修復実験	5
3.1.1	実験プロセス	5
3.1.2	実験条件	6
3.2	発話評価実験	6
3.2.1	実験プロセス	6
3.2.2	実験条件	7
3.3	実験結果	7
3.3.1	文章の発話しやすさと修復しやすさ	9
3.3.2	個人の持つ発話技能と修復能力	10
3.3.3	発話の成否と修復の成否のリンク	11
3.4	まとめと考察	11
3.4.1	修復に至った要因	12
3.4.2	知覚に影響を与える発話のプロセス	12
第4章	発話計画と認知の予測	14
4.1	音韻修復実験	14
4.1.1	実験プロセス	14
4.1.2	予測によって修復される音素	15
4.1.3	実験条件	15
4.2	発話計画能力評価実験	15
4.2.1	実験プロセス	15

4.2.2	発話改善率	16
4.2.3	発話音声	16
4.2.4	実験条件	18
4.3	実験結果	18
4.3.1	発話未計画時	18
4.3.2	発話計画可能時	18
4.3.3	発話改善率と修復	18
4.4	まとめと考察	22
4.4.1	発話における発話計画の役割	22
4.4.2	発話計画可能時の発話技能と修復率	22
4.4.3	発話計画能力と音韻修復時の予測	22
第5章	結論	24
5.1	本研究の成果	24
4.1	今後の課題	24
	謝辞	25

第1章 序論

1.1 研究の目的と背景

音声によるコミュニケーションは、話者の脳内にある言語的メッセージが調音運動によって音響信号に変換され、聴取者がそれを聴覚系で受容して脳で解読して言語的メッセージを理解するという一連の処理の連鎖「スピーチ・チェーン」によって実現されるという考え方が 1960 年代に Danes らによって提唱された。[1]半世紀近く立った今日でもこの大枠事態は概ね不変であり、音声研究の格好の出発点となっているが、近年の認知神経科学の発展に伴いその詳細が急速に明らかになってきている。特にオリジナルのスピーチ・チェーンでは話者間の音声によるつながりに着目して、話者側では脳から口、聴取側では耳から脳までそれぞれ独立して一方向に処理が進んでいるとされているが、このエンコードとデコードの過程には相互作用があり、双方向の処理であると考えられるようになった。[2]そこでこれまでに明らかにされた発話と音声知覚の関連の一旦について紹介する。

古くには騒音のある場所では発話音声の性質が変化する Lombard 効果よく知られている。[3]また、自分の発話音声を数百 ms 程度の遅延を与えてフィードバックさせると多くの被験者が発話障害を起こすことや、第二言語の学習時にある音素の知覚訓練の効果がそのまま生成能力へ転移することが分かっている。[4][5] 音声を聴取する時に舌筋や脳の発話に関与する運動野も活性化しているという報告もある。[6][7]さらに生成過程が音声の知覚においてより本質的な役割を果たすとする有名な仮説に「音声知覚の運動理論」がある。[8]知覚運動理論では人間には音声のみを扱う特別な処理機構、音声モードが生理的に備わっており、音声の知覚は音響信号からそれを生成した発話者の調音ジェスチャーを推定する事に基づく、としている。これも未だ明確な証拠こそ見つかっていないものの、数多くの研究結果から一定の支持を集めてきている。[9]

そこでもしこの知覚運動理論の主張が正しいのであれば、また全面的に正しくないまでも数多くの研究結果が示すように生成過程が音声知覚へ少なからず影響を与えているのであれば、非常に自然な考え方として「優れた発話技能を持つ人物は同時に優れた音声知覚能力も持っているのではないだろうか？」という推測が成り立つ。しかしこれまでの研究ではその手法の難しさもあり、個人の発話や音声認知の巧拙を直接的に評価して論ずるような研究はほとんど行われてこなかった。そこで本稿では早口言葉を用いて被験者の発話技能を評価し、一方音韻修復現象を用いて知覚能力を評価することでこれを論じ、被験者の持つ発話技能と知覚能力の相関について調査する。またその結果から具体的に発話過程のどのようなプロセスが知覚に影響を与えているか考察し、発話と知覚の関連性を明らか

にする。

1.2 論文構成

本稿ではまず第 2 章で音韻修復の特徴と性質について述べる。以降第 3 章で発話能力と知覚能力の相関，第 4 章に発話計画と音韻修復の予測について述べ、そして第 5 章に結論をまとめる。

第2章 音韻修復現象

2.1 本研究のアプローチ

本研究では「優れた発話能力を持つ人物は同時に優れた音声知覚能力も持っているのではないだろうか?」という仮説に基づき、被験者の発話能力を早口言葉の巧拙で評価、一方で知覚能力を人間の持つ音韻修復と呼ばれる機能を用いることで評価し、発話能力と知覚能力の関連を明らかにしていく。そこで本章では音韻修復現象についてその特徴と性質を述べる。

2.2 音韻修復現象

音声信号の一部が削除され明らかに了解度が低下する場合でも、削除された部分に別の音が挿入されるとあたかも音声が無傷のように知覚されることがある。これが音韻修復現象と呼ばれ、聴覚における錯覚の代表的存在である。この錯覚は画像を認知するときにも同様の現象が起こることが知られている。画像の認知であれば図示しイメージが掴みやすいため、欠損した画像の修復例として図 2.1 に示す。図 2.1 の左図はある文字の一部分を欠損させたものである。この状態だと欠損部分が大きく元の文字が何であったか判断することが難しい。しかしその欠損部分に右図のように別の形を補完してやると明らかに認識が容易くなる (A,B,C,D と書いてある)。この文字の部分を音声、保管した円の部分を雑音に置き換えたものが音韻修復である (図 2.2 参照) [9]。

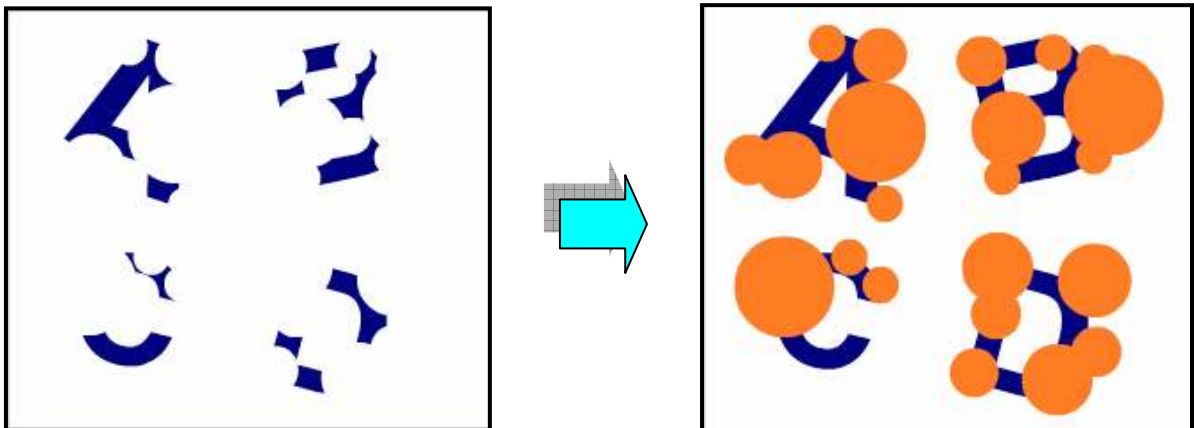


図 2.1 : 音韻修復の画像によるイメージ

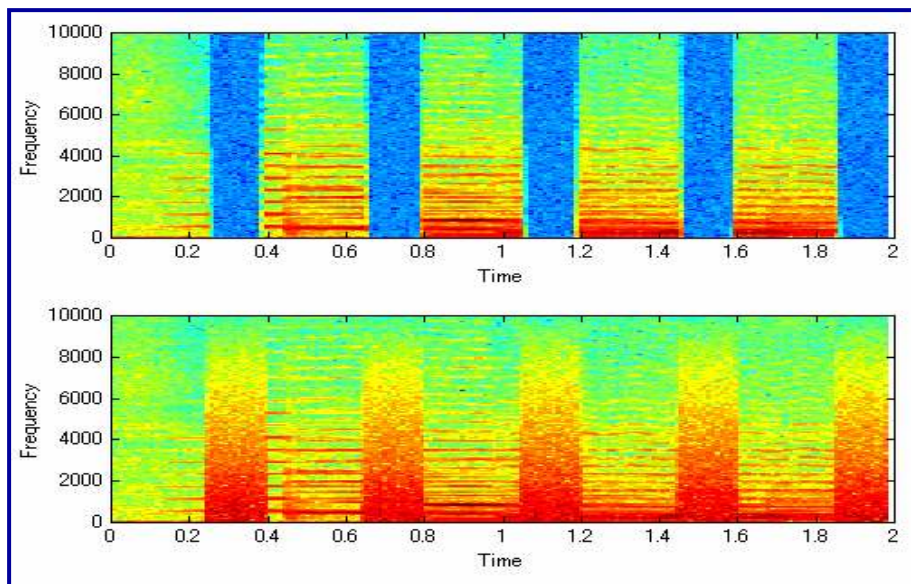


図 2.2 : 音韻修復の音声によるイメージ

欠落部分に別の音が挿入されることによって、例えある音素の情報が完全に欠落していても両隣の音素の調音結合や前後の文脈情報を手掛かりとして修復が起こる。この音韻修復のお陰で聞きたい音声の一部が大きな音によってマスキングされても、頑健に音声を知覚することができる。そのため音韻修復は様々な音で溢れている日常生活において極めて重要な役割を果たしている[10]。また、個人によって修復の度合いに差が出ることも知られており、雑音の多い実環境下ではその巧拙がそのまま知覚能力の差として表れる。本研究ではこのような性質を持つ音韻修復現象を用いてその修復率を調べることにより、被験者の知覚能力を評価する。

第3章 発話能力と知覚能力の相関

本章では被験者の発話能力を早口言葉の巧拙で評価し、一方知覚能力を音韻修復の修復率で評価することにより被験者の持つ発話能力と知覚能力に相関が見られるかどうか調査した。

3.1 音韻修復実験

3.1.1 実験プロセス

本実験での音声資料は被験者の持つ言語知識による影響を極力小さくするため、一般的に知られていない 8 つの早口言葉を元の音声として選んだ。また被験者の発話技能によって巧拙により大きな差が出ることを狙い、比較的発話が難しいものにした。以下がその早口言葉である。[11]

1. 陰間から傘陰間下駄(かげまからかさかげまげた)
2. 野葵家葵(のあおいえあおい)
3. 刃釜に生米(はがまになまごめ)
4. わやや八百屋におあやまりになった
5. 鴨米かむ小鴨が小米かむ(かもこめかむこがもがこごめかむ)
6. 規格価格か懸引価格か(きかくかかくかかけひきかかくか)
7. 第五交響曲に観客驚愕
8. 南から猿来よった皆見さる

これら早口言葉に対して十分に訓練を積み、スムーズに早口で発話できるようになった著者自身の発話音声を録音した。その録音した音声の一部分を **Matlab** を用いて周期的に **225ms** 間で **75ms** 欠落させ、削除した部分に雑音を挿入して本実験で用いる音声データベースとした。尚雑音には音声資料の平均パワーを **15~20dB** 程度上回る白色雑音を用いている。それらを **12** 人の本学の日本人男子学生に聞かせ、それぞれの文章を **10** 回聴取させ聞き取れた音についてひらがな、又は音素によって自由に書き出してもらった。ただし、正解音素ラベルとの対応を正しく集計するため、正しく認知できない場合、又は聞こえなかった場合はその箇所は空欄にしてもらった。その後元の早口言葉と比較して修復率を調査し、それを知覚能力の評価として用いた。尚本稿では音韻修復実験での正解率を修復率として表記している。

3.1.2 実験条件

音韻修復の実験条件を表 3.1 にまとめる。

表 3.1 音韻修復実験条件

音声データ	著者の発話早口言葉
音声欠損	白色雑音
欠損区間	225ms 周期, 75ms 間
被験者	日本人男子大学院生 12 名
文章数	8 個
聴取回数	10 回
聴取条件	ヘッドホン(Sony:MDR-Z600)

3.2 発話評価実験

3.2.1 実験プロセス

音韻修復実験後に音韻修復実験で用いた早口言葉について 12 人の同じ被験者にそれぞれ 5 回連続発話させてそれを録音した。その際日本語として不自然な文章もあるので発話ではなく認知の時点で失敗してしまうことも考えられた。そのためそれを避けるために、発話の前に十分時間を取り被験者に文章を理解してもらった上で発話を開始してもらった。その後そこで録音した音声文を用い、新たに評価者として参加してもらった学生 5 人に 5 段階評価で各被験者の各文章に対する発話技能を評価してもらった。5 回連続の発話の中で発話に頻繁に詰まる、発音を頻繁に間違えるなど発話技能が最も低い状態が 1、終始詰まることも無く比較的早口で発話できている発話技能が最も高い状態が 5 である。そしてその発話技能の平均値を各被験者の発話技能の評価値として用いた。

3.2.1 実験条件

発話評価実験の実験条件を表 3.2 にまとめる。

表 3.2 発話評価実験条件

音声データ	各被験者の発話早口言葉
被験者	日本人男子大学生 12 名
評価者	日本人男子大学生 5 名
文章数	8 個
発話回数	5 回
MOS	1~5 の 5 段階評価

3.3 実験結果

まず文章 1 と文章 2 の各被験者の音韻修復率と 5 人の評価者による平均発話評価値、及びその評価値の分散をそれぞれ表 3.3, 3.4 に示す。同様の結果が文章 8 までであるが、ここでは煩雑さを避けるために文章 1 と文章 2 だけを示した。表 3.3 を見ると各被験者の発話技能に対する評価値の分散は非常に低い値となっているのが分かる。これはそれぞれの評価者がそれぞれの発話に対して評価者間で概ね差のない評価値を与えていることを意味している。この傾向は文章 1, 2 だけではなく全ての文章において見られた。そのためこれ以降 5 人の評価者による評価値の分散は無視して議論を進めていく。また個人差はあるものの、ある 1 人の被験者にとって発話しにくい、又は修復しにくい文章は他の 11 人の被験者にとっても同様に発話しにくい、修復しにくいという傾向が見られた。

図 3.1 に 12 人の被験者のそれぞれの文章に対する発話評価値と音韻修復率を示す。この図では発話評価値と修復率に相関係数にして 0.637 の右に肩の相関が見られた。このことは被験者の発話能力と知覚能力に関連があることを示唆している。

表 3.3 文章 1 の音韻修復率と平均発話評価値、及び評価値の分散

被験者	A	B	C	D	E	F	G	H	I	J	K	L
修復率	45.8	33.3	29.1	29.1	37.5	33.3	50.0	37.5	50.0	37.5	41.6	41.6
平均評価値	3.6	2.4	2.0	1.6	3.0	2.6	3.4	2.4	3.2	2.2	2.2	2.8
評価値の分散	0.24	0.24	0.00	0.24	0.00	0.24	0.24	0.24	0.16	0.16	0.16	0.16

表 3.4 文章 2 の音韻修復率と平均発話評価値, 及び評価値の分散

被験者	A	B	C	D	E	F	G	H	I	J	K	L
修復率	54.5	72.7	63.6	45.4	64.3	49.1	66.6	60.6	63.6	45.4	63.6	54.5
平均評価値	4.8	4.6	4.6	3.6	4.4	4.0	4.6	4.2	4.2	4.6	4.0	3.8
評価値の分散	0.16	0.24	0.24	0.24	0.24	0.00	0.24	0.16	0.16	0.24	0.00	0.16

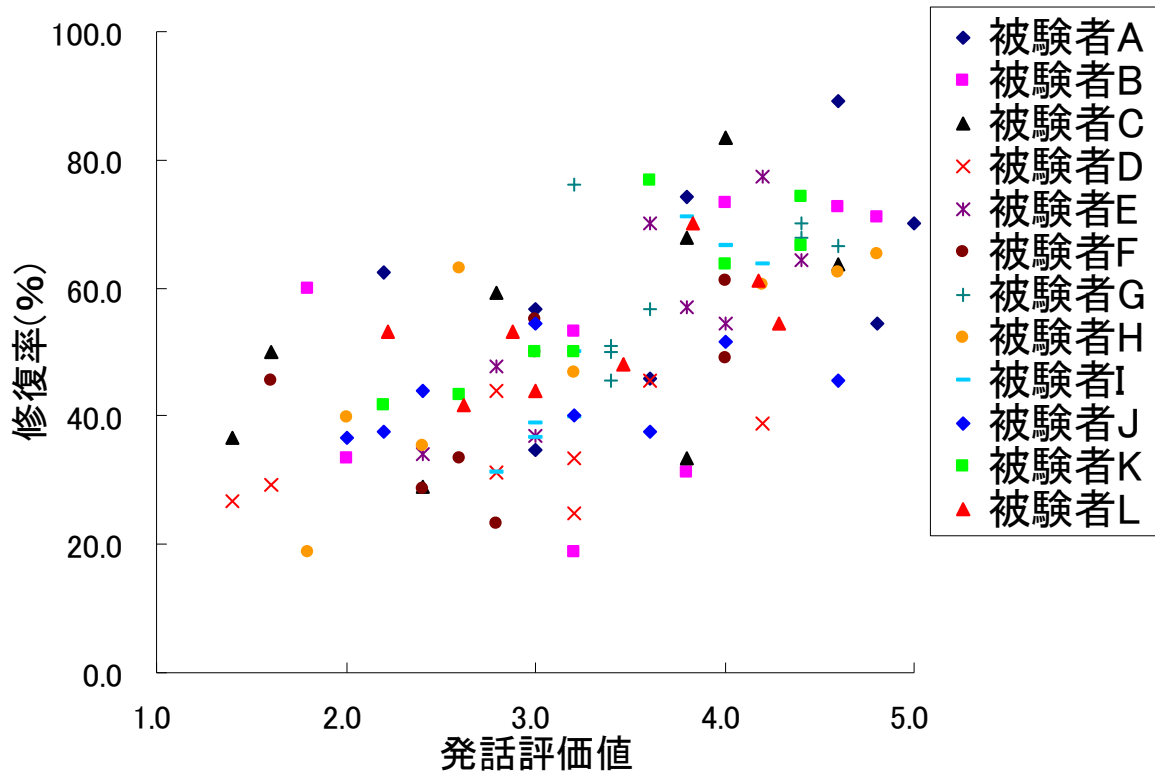


図 3.1 各被験者のそれぞれの文章における発話評価値と音韻修復率

3.3.1 文章の発話しやすさと修復しやすさ

文章の発話しやすさと修復しやすさの相関を調査するために、図 3.2 に 8 つの文章の平均発話評価値と平均修復率を示す。尚点線は標準偏差から求めたエラーバーを示している。8 つの点は 8 つの文章を表しており、1 つの文章において 12 人の被験者で平均している。ある文章において平均発話評価値が高いということは、多くの被験者にとってその文章は発話しやすかったということである。それに対し修復率が高いということは多くの被験者にとってその文章は修復しやすいということである。図 3.2 を見ると発話評価値と修復率に相関係数にして 0.872 の正の相関があり、発話評価値が高ければ高いほど、修復率も高くなっている。このことは即ち、被験者にとって発話しやすい文章は同時に修復もしやすかったことを示している。

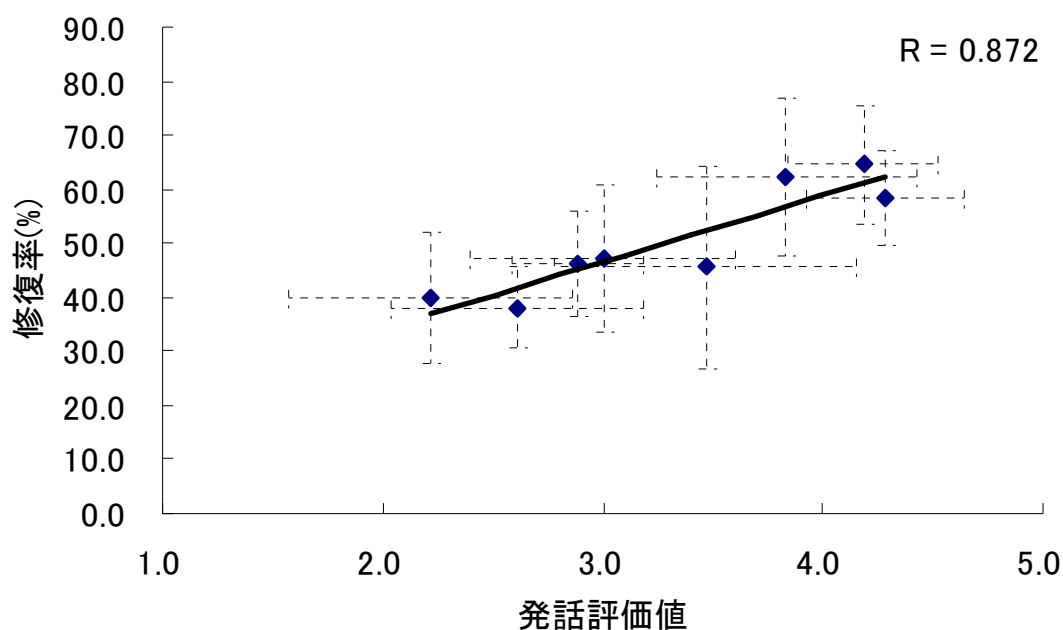


図 3.2 各文章の平均発話評価値と平均修復率

3.3.2 個人の持つ発話技能と修復能力

個人の持つ発話能力と修復能力の相関を調べるために、図 3.2 と同様にして被験者 1 人において 8 つの文章で平均化し、被験者の平均発話評価値と平均修復率を示したい。ところがここで 1 つ問題がある。表 3.3, 3.4 から分かるように今回の実験ではそれぞれの文章において発話しやすさ及び修復しやすさが大きく異なっている。そのためそのまま異なる実験試料である文章自体で平均化するとエラーバーが非常に大きくなり、統計的信頼性を著しく損なってしまう。そこで各文章においての各被験者の正解率と評価値をあらかじめ被験者全体における偏差を用いて表し、後に 8 つの文章の評価偏差値と正解偏差値を平均化し、各被験者の平均発話偏差値と平均修復偏差値として表した。図 3.3 に 12 人の被験者の平均発話偏差値と平均修復偏差値を示す。12 の点はそれぞれ 12 人の被験者を表している。発話偏差値が高い被験者は発話技能が高いということである。それに対し修復偏差値が高い被験者は修復が上手いということである。図 3.2 を見ると発話評価値と修復率に相関係数にして 0.875 の正の相関があり、発話評価値が高ければ高いほど、修復率も高くなっている。このことは即ち、発話が上手い被験者は同時に修復も上手いということを示唆している。

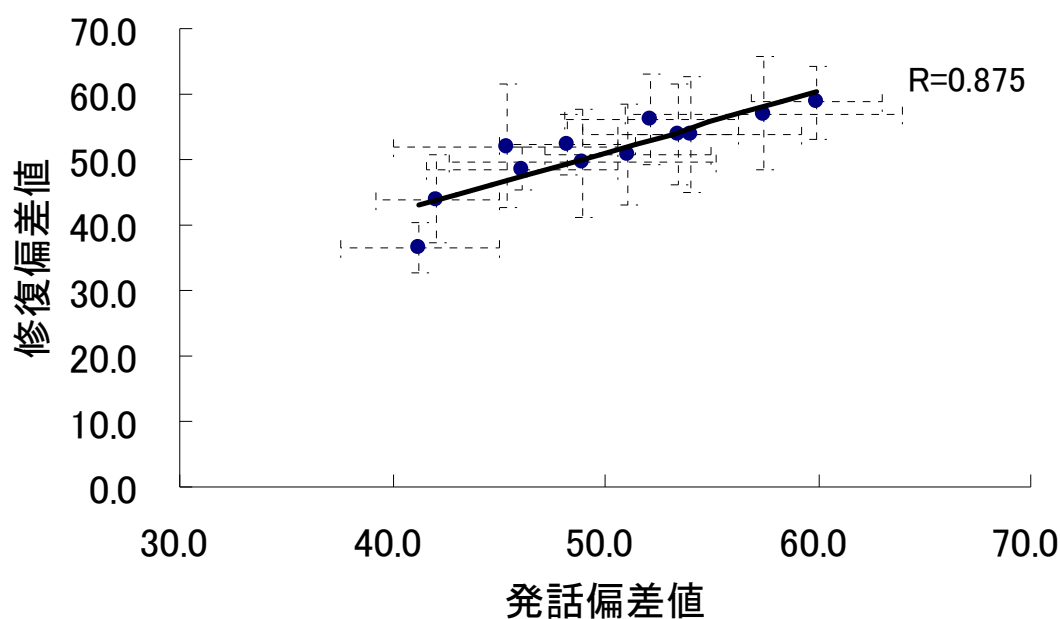


図 3.3 各被験者の平均発話偏差値と平均修復偏差値

3.3.3 発話の成否と修復の成否のリンク

表 3.5 に各被験者の音素全体の修復率と発話を成功した箇所だけの修復率, 及び発話を失敗した音素だけの修復率の比較を示す。ここで発話を失敗した音素の定義であるが、5回の連続発話の内一度でも直前で詰まる, 又は発音を間違えた音素として著者自身が評価した。この結果を見ると D と H を除く全ての被験者で発話に失敗した音素の修復率が発話に成功した音素の修復率より低くなっているのが分かる。発話に成功した音素の修復率の平均値は 52.5%であったのに対し、発話に失敗した音素の修復率は 42.5%であった。このことは被験者にとって発話しにくい箇所は同時に修復もしにくかったことを示している。

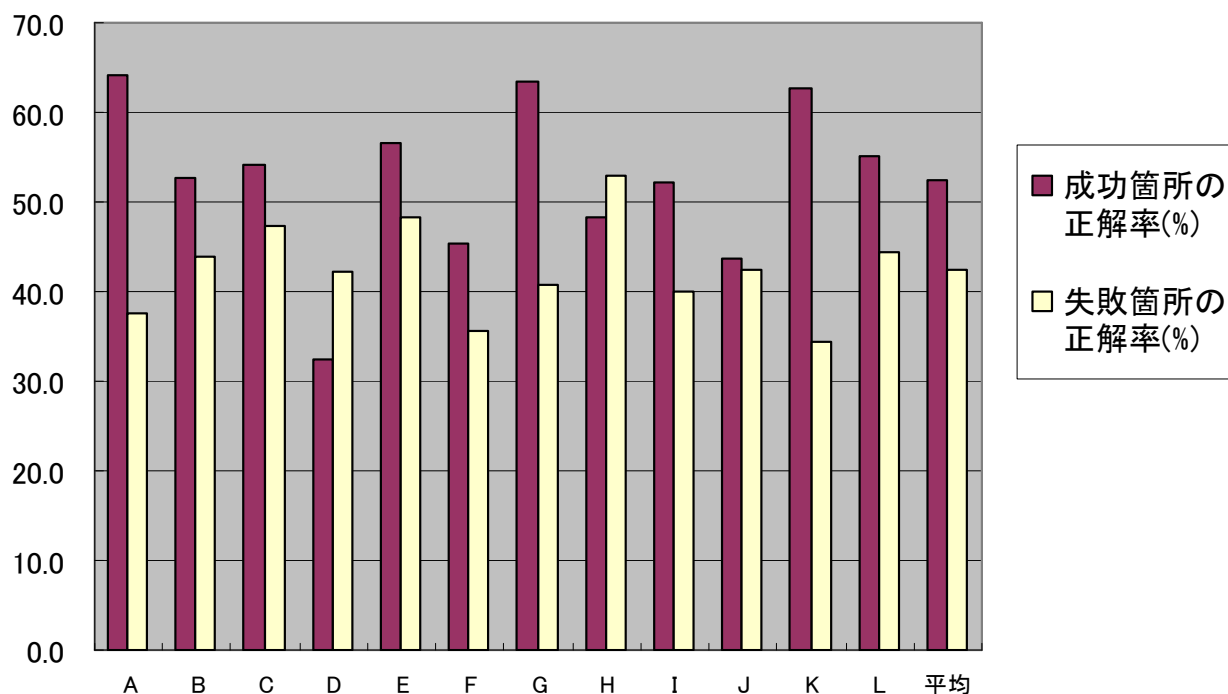


図 3.5 発話を成功した音素の修復率と、失敗した音素の修復率の比較

3.4 まとめと考察

本実験によって高い発話技能を持っている人物は同時に高い知覚能力も保持していること、また発話しやすい文章, さらに言えば本人にとって発話しやすい箇所では同時に知覚もしやすいということが示された。これらの結果は生成と知覚には相互作用があることを示唆しており、発話過程が音声知覚に影響を与えていることを明らかにしている。それで

は発話の具体的にどのようなプロセスが認知に影響を及ぼしているのだろうか？それを知るためには、被験者が文章を修復するにあたり何を手掛かりにしたか、その要因を調べる事が重要である。

3.4.1 修復に至った要因

音韻修復実験で被験者が修復に至った要因として、大別すると自身の音素の残存している情報から、両隣の音素の調音結合から、そして言語情報からの予測による3つが考えられる。この内言語情報からの予測の影響について単語親密度を用いて検討する。

単語親密度は、ある単語のなじみがあると感じる程度を複数の被験者が7段階で評定したときの平均値として求められる。[12]この単語親密度が単語認知の容易さに深くかかわっていることが、これまでの研究によって明らかにされている。図3.5に各文章に使われている平均単語親密度と平均修復率の関係を示す。

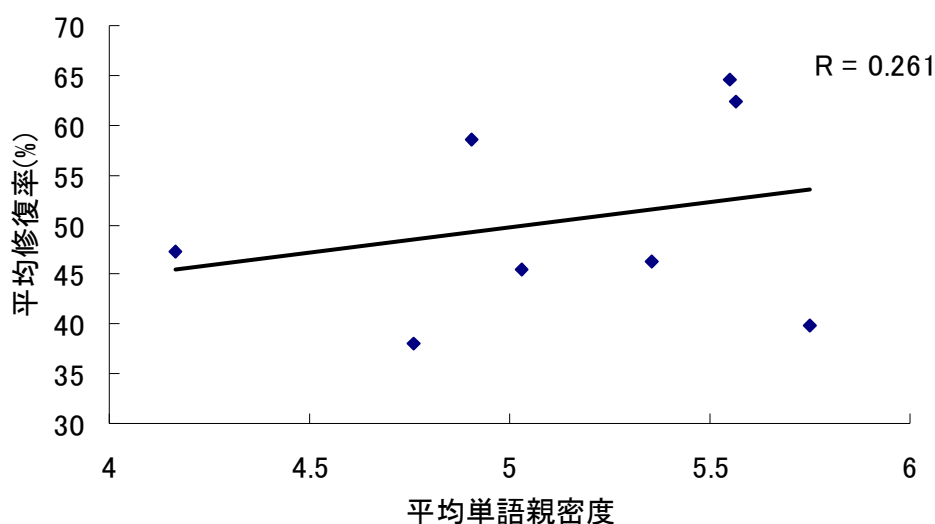


図 3.5 各文章に使われている平均単語親密度と修復率

図3.5から単語親密度と修復率について相関係数にして0.261の若干の相関が見て取れるが、文章内で使われている単語のなじみの程度が修復に影響を及ぼしているとは言い難い。そのため修復の手掛かりとなった要因、調音結合や文脈情報の影響について更なる検討が必要である。しかしながら本実験では音韻修復で用いた実験試料に音素のラベリングが無く、また雑音を周期的に挿入しているため、どの音素にどれだけの情報が残っているか不明でどのような要因から修復に至ったのか判断がつきにくい。そのためより明確な議論をするためには音素のラベリングがついている音声データを使用し、又雑音を周期的ではなく計算して挿入し、欠落させた部分を明確にした音声データベースを用いた検討が必要である。

3.4.2 知覚に影響を与える発話のプロセス

ここで一旦発話過程のどのようなプロセスが修復に影響を及ぼしたのか考察するために、発話と音韻修復のプロセスについて再考する。オリジナルのスピーチ・チェーンによると話者は脳内にある言語的メッセージを調音運動によって音響信号に変換しているとしている。言語的メッセージが脳内にあるということは言い換えれば、話者は発話する前に既に自分が喋る文章を先読みしているということに他ならない。話し手は文を発する時、最初に立てた発話計画を基にして話し始めから話し終わりまで同一の意識をもってその計画を遂行している。一方音声知覚の過程においても言語情報による先読みが大きく認識を助けている。特に音韻修復では文脈情報による予測の影響が非常に大きいことが分かっている[10]。つまり文章を先読みする、という点は一見お互いに全く異なる処理を施していると思える発話と知覚の過程において唯一共通しているプロセスなのである。そのためこの発話時の「発話計画」と音韻修復時の「予測」について相関が見られる可能性が考えられる。そこで次の章ではこの「発話計画」と「予測」に着目して、その関連を調査した。

第4章 発話計画と認知の予測

本章では被験者の発話計画能力を早口言葉を用いた手法で評価、一方一般的な日本語の文章を用いた音韻修復実験により認知時の予測能力を評価し、両者の相関について調査した。

4.1 音韻修復実験

4.1.2 実験プロセス

修復に至った要素をはっきりさせるため、今回加工して用いる音声資料は周期的ではなく計算して欠落させる必要がある。そのため音韻修復実験でも早口言葉を用いた第3章とは異なり元の音声データベースには全音素のラベリングがつけてある日本語の文章が収録されている ATR-B セットを用いた。またこの ATR-B セットではプロの人物の発話が収録されているために、音声資料としての信頼性が高い。以下が音韻修復実験で用いた文章である。

1. お金を入れてボタンを押すと切符が出てくる
2. スピーチが下手なのだから原稿を用意した方がいい
3. ブラブラと球場まで10分足らずの道を歩いていく
4. 救急車で病院に運ばれる途中、息を引き取ったのだという
5. 小さい雪は速く、大きい雪はふわふわと落ちてくる
6. 怒れば怒るほど火に油を注ぐようなものだった
7. 運転免許を取って三年目になります
8. 立春が過ぎても厳しい寒さの日々が続く

今回の実験では予測による修復について調査することが目的であるため、中でも日本語として非常に自然で且つ日常生活において耳にし、被験者の予測が良く働くと予想される文章を選んだ。その際被験者が保持している知識量による影響を避けるため、難しいまたは専門的な用語は排除するよう心掛けた。

これら選び出した音声文の一部分を Matlab を用いて欠落させ、削除した部分に雑音を挿入して本実験で用いる音声データベースとした。尚雑音には前実験同様音声資料の平均パワーを 15~20dB 程度上回る白色雑音を用いている。それらを 8 人の本学の日本人男子学生に聞かせ、それぞれの文章を複数回聴取させた。聴取回数は被験者にもうこれ以上聞いて

も変化が無いと思う 10 回以内の任意の回数まで続けさせているので、各被験者それぞれの文章によって異なっている。その後前章と同様に聞き取れた音について書き出してもらい、全体の音素の修復率と文脈情報からの予測によってのみ修復可能な音素の修復率を調査した。

4.1.1 予測によって修復される音素

文脈情報からの予測のみによって修復可能な音素とは自身の情報が 100% 欠落しており、且つ直前の音素の後半部分と直後の音素の前半部分が完全に欠落している音素と定義している。このような音素では自身の情報からは当然認識することはできない。また自身に隣接している音素の結合部分が大きく欠落しているため調音結合による修復もまず不可能である。尚、過去の無意味後を用いた研究結果ではその調音結合による情報が前後 200ms 程度の範囲内に分散しているとされているが、これらの音素ではそれを大きく上回る範囲で欠落させられている[13]。そのためこのような条件下の音素では物理的に文脈情報からの予測によってのみしか修復が起こりえない。

4.1.3 実験条件

音韻修復の実験条件を表 4.1 にまとめる。

表 4.1 音韻修復実験条件

音声データ	ATR-B セット
音声欠損	白色雑音
欠損割合	全音素の 40~50%程度
被験者	日本人男子大学生 8 名
文章数	8 個
聴取回数	1~10 回
聴取条件	ヘッドホン(ATH-AD500)

4.2 発話計画能力評価実験

4.2.1 実験プロセス

音韻修復実験に参加した 8 人の同じ被験者に対して、それぞれ 2 つの異なる状況下で 8 つの早口言葉を発話してもらい録音した。1 つ目の状況は被験者がそれぞれの文章について

任意のタイミングで一部分ずつディスプレイに表示させ、表示後即座に発話してもらったものだ。被験者は実験を行う前段階で、表示されている部分を発話し終わる前に次の部分を表示させ、連続的に発話できるように練習してもらっている。この状況下では被験者は発話計画を立てる時間をほとんど得られず、文章の先読みをすることが非常に困難な状況であると考えられる。本稿では便宜上この状況下を発話未計画時と呼ぶことにする。この発話未計画時の音声を録音した4日ないし5日後に、同じ早口言葉について今度は初めから全文見えている状況下での発話してもらった。(予備実験の段階で1度発話した早口言葉を3日後に再度発話してもらったところ、学習効果が表れていないことを確認している。)被験者には文章の表示後10~12秒の間に任意のタイミングで発話を開始してもらっている。また被験者には表示後発話までの間に発話する早口言葉をきちんと理解し、発話することに集中するよう指示が出されている。これによって被験者は発話計画を立てる時間を十二分に与えられ、文章を先読みできる。本稿ではこの状況下を便宜上、発話計画可能時と呼ぶことにする。これら録音した音声文について3章と同様に新たに評価者として参加した5人の被験者に5段階評価で主観評価実験を行い、両状況下での発話評価値を調査した。また発話計画を立てられることによる発話改善率を用いて、発話計画能力と音韻修復の予測の関係について調査した。

4.2.2 発話改善率

一部分ずつの表示によって発話計画を立てにくい場合の評価値をX、発計画時間を十分に取れる場合の発話評価値をYとすると発話改善率Z(%)は次のように表される。

$$Z = \frac{Y - X}{5 - X} \times 100 \dots \dots \dots (1)$$

例えば発話計画を立てづらい時の発話評価値が3、発話計画を立てやすい時の評価値が4であったなら、発話改善率は50%となる。両状況下での実験条件は発話計画の有無以外は等しいことから、この発話改善率の値はより発話計画能力と直結していると考えられる。本稿ではこの発話改善率を発話計画能力と同義と捉えて論じていく。

4.2.3 発話音声

発話未計画時の発話音声はあくまで表示後即座に発話することが重要である。そこで表示後認知に手間取るなどして何かアクシデントが起り、時間が経過してしまった音声はその被験者の発話音声から省かなくてはならない。そこで本来必要な数よりも多い14の早口言葉をあらかじめ用意し、各被験者が即座に発話することが出来た8つの早口言葉について評価した。漢字が読めなかった、日本語として理解できなかった等認知の段階で失敗した、またその早口言葉について既知であった場合等は除外している。以下がその早口言

葉である。

1. あのアイヌの女の縫う布の名は何
2. 宿直室に宿直して四苦八苦
3. 横町の七曲り長い七曲り
4. 茶会社の社長茶瓶で茶会社の社長
5. 隣の戸棚においてある柿は客の食う柿か客の食わぬ柿か
6. どじょうによろによろ三によろによろ
7. 新人歌手新春シャンソンショー
8. 全車種に標準装備
9. 向こうの細どぶに細どじょうがによろり
10. むこうの柿の木へ竹立てかけた
11. 魔術師魔術修業中
12. 古栗の木の古切れ木
13. 夜飛ぶ鳥取りたい
14. 服作る夫婦、靴作る夫婦

被験者は1~14まで順番に発話していき問題なく発話に成功した文章が8つに至った時点で発話をやめてもらっている。つまり 1~8 までの文章は全被験者で発話に失敗したものを除いて共通であるが、9~14 は被験者によって発話していたり、していなかったりとまちまちである。しかしこれらの早口言葉は各被験者の発話評価値を与える 5 人の評価者によって事前に難易度が調査されており、9~14 までの文章は発話の難易度にほとんど差が無いものが並べられている。また 9~14 までの文章はそれまでの全被験者の発話回数によって順番を調整され、各文章に発話被験者が分散するよう配慮されており、各文章において最低 3 回は発話被験者が存在している。これによって各被験者は発話した早口言葉によって、評価に差が出るような事態は避けられている。

4.2.4 実験条件

表 4.2 に発話計画能力評価の実験条件をまとめる。

表 3.2 発話計画能力評価実験条件

発話音声	早口言葉
被験者	日本人男子大学生 8 名
評価者	日本人男子大学生 5 名
文章数	14 のうち 8 個を評価
MOS	1~5 の 5 段階評価

4.3 実験結果

4.3.1 発話未計画時

発話未計画時の発話技能と修復の相関を調べるために、発話未計画時の発話評価値と予測による修復率の相関を図 4.1 に、発話評価値と音素全体の修復率の相関を図 4.2 に示す。発話未計画時には被験者の発話評価値と予測による修復に相関は見られなかった。またそれにつられるようにして 3 章で見られていた発話評価値と音素全体の修復の相関も失われている。

4.3.2 発話計画可能時

発話計画可能時の発話技能と修復の相関を調べるために、発話評価値と予測による修復率の相関を図 4.3 に、発話評価値と音素全体の修復率の相関を図 4.4 に示す。発話未計画時には見られなかった発話評価値と予測による修復に相関係数にして 0.598 の相関が見られた。また発話評価値と音素全体の修復率にも 3 章同様に相関が見られた。

4.3.3 発話改善率と修復

発話計画能力と予測による修復の相関を調べるために、発話改善率と予測による修復の相関を図 4.5 に、発話改善率と音素全体の修復率の相関を図 4.6 に示す。尚発話改善率は式 (1) で表される。

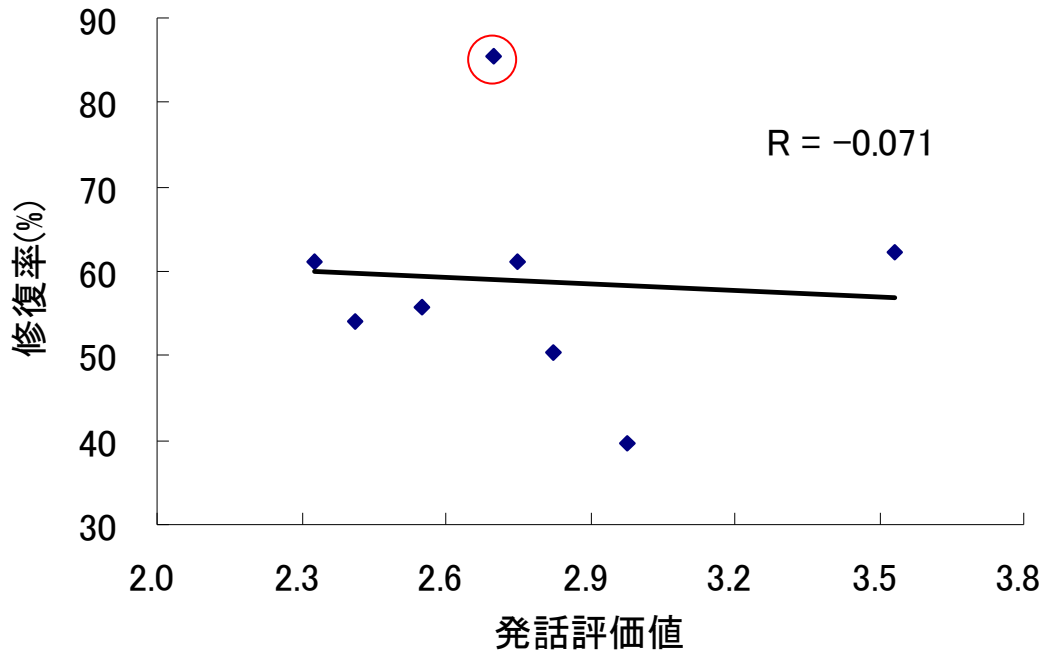


図 4.1 発話未計画時の発話評価値と予測による修復率

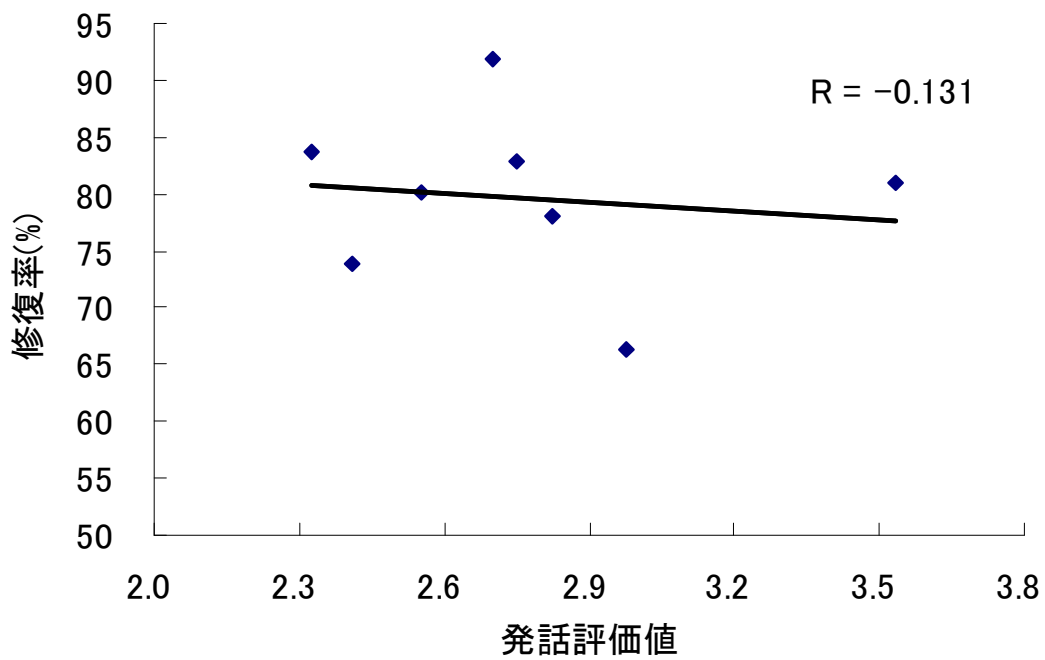


図 4.2 発話未計画時の発話評価値と音素全体の修復率

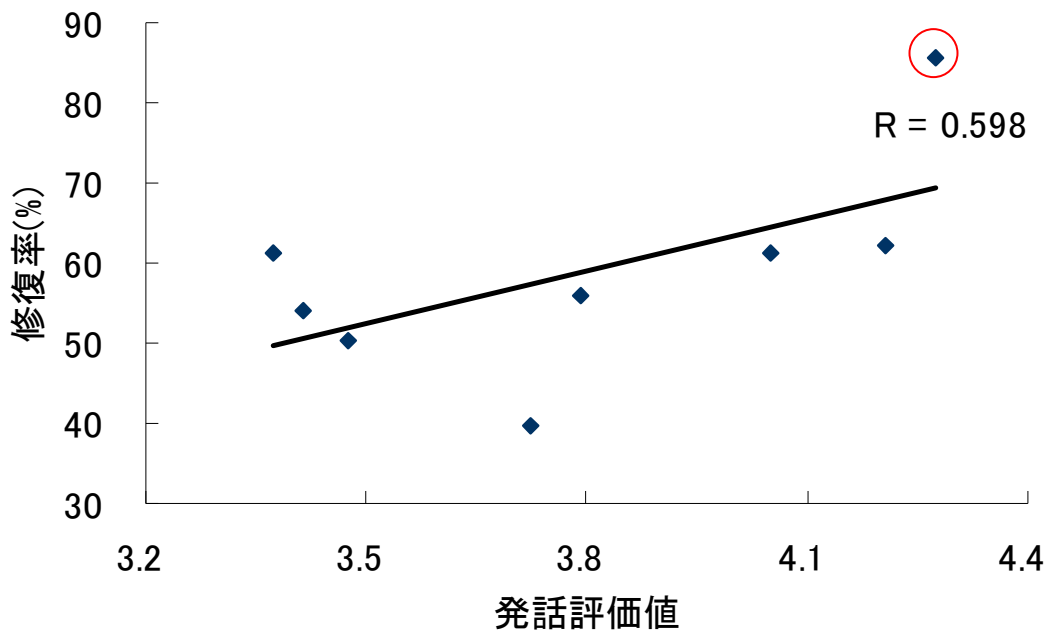


図 4.3 発話計画可能時の発話評価値と予測による修復率

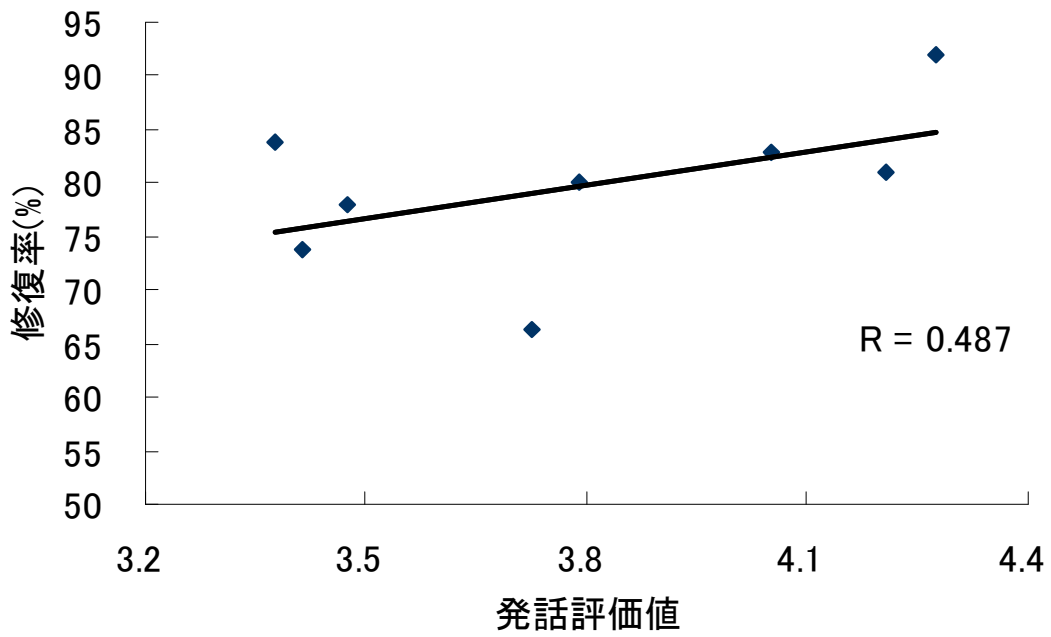


図 4.4 発話計画可能時の発話評価値と音素全体の修復率

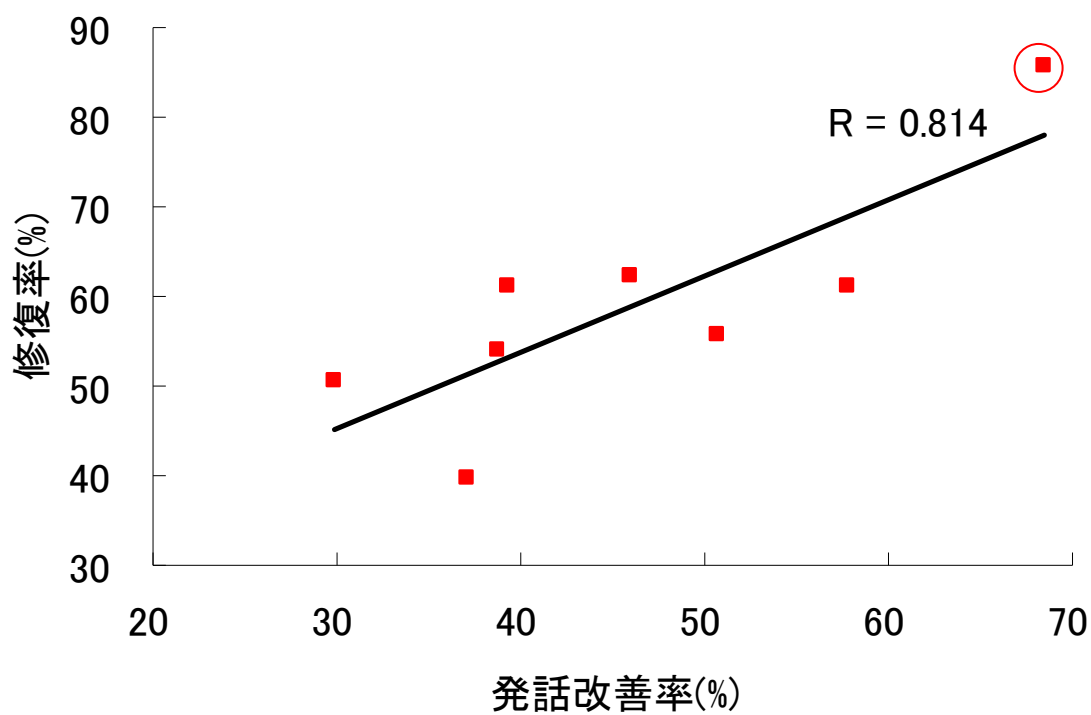


図 4.5 発話改善率と予測による修復率

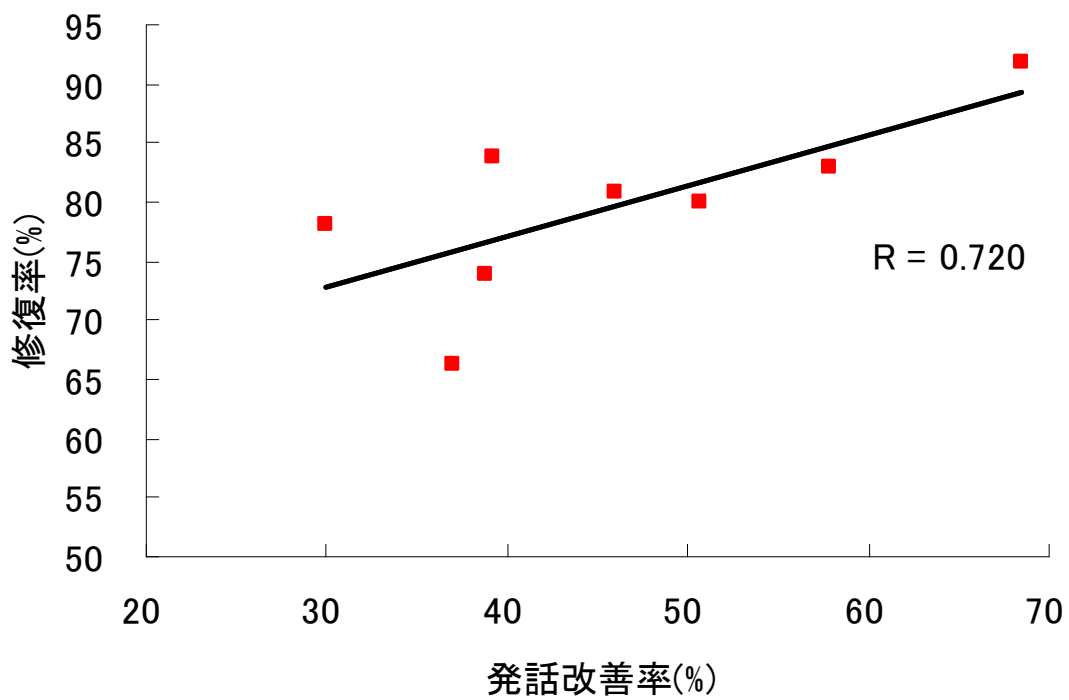


図 4.6 発話改善率と音素全体の修復率

4.4 まとめと考察

4.4.1 発話における発話計画の役割

早口言葉を一部分ずつ表示させた発話未計画時に比べ、始めから全文表示させた発話計画可能時には、図 4.5 からわかるように発話改善率にして 30~70%程度、全ての被験者で発話の上達が見られた。この事は発話、特に早口言葉では発話計画が重要な役割を果たしていることを示唆している。

4.4.2 発話計画可能時の発話技能と修復率

発話計画可能時の被験者の発話評価値と音素全体の修復率に相関係数 0.487 の相関が見られた。言い換えれば高い発話技能を持つ被験者ほど同時に高い知覚能力を持っているということである。この事は前章の結論と矛盾しない結果ではあるが、図 3.3 に比べて弱い相関となっている。その理由は 3 章では音韻修復実験と発話評価実験で同じ早口言葉を用いたのに対し、本章では一般的な日本語の長文と早口言葉というように互いに異なる音声資料を用いたためであると考えられる。同じ音声資料を用いた前章では節 3.3.3 で述べられているように発話失敗箇所と修復失敗箇所とのリンクが取れているため、発話技能が直接的に修復の可否に繋がり、高い相関を示している。それに対し異なる音声資料を用いた 4 章では発話失敗箇所と修復箇所のリンクが取れない。また音韻修復実験で一般的な日本語の文章を用いているため、発話しづらい箇所というのがほとんど存在しない。そのため発話評価値と修復率が直接的には結びつかず、図 3.3 に比べ低い相関となって表れていると考えられる。

4.4.3 発話計画能力と音韻修復の予測

早口言葉が一部分ずつ表示され、発話計画をほとんど立てられない発話未計画時には相関係数-0.071 と全く見られなかった発話技能と音韻修復時の予測の関係に、始めから全文表示されており、発話計画を立てることが十分可能な発話計画可能時では相関係数 0.598 の相関が見られるようになった。さらに発話改善率と予測の間には相関係数 0.814 の強い相関が見られた。このことは文章の予測は音声の生成、知覚の両過程において重要な役割を果たしており、両過程でプロセスを共有していることを示唆している。さらに発話過程の発話計画が音韻修復時の予測に影響を及ぼしており、高い発話計画能力を持つ被験者ほど高い予測能力、ひいては高い認知能力を持っていることが明らかになった。

ここで象徴的な例として図 4.1, 4.3, 4.5 のそれぞれ○で囲まれている被験者のケースについて言及する。この被験者は全て同一人物であるが(仮に被験者 A とする)、A の特筆すべき点はその修復率の高さだけでなく、音韻修復実験での聴取回数の少なさにある。被験

者 A の 1 つの文における平均聴取回数は 1.6 回であった。他の被験者の平均聴取回数が 4.6 回であること、この極端に少ない聴取回数で完全に欠落している音素を予測のみによって 85.9%修復していることを考えると、この値は実に驚異的である。図 4.1 を見ると分かるように発話未計画時には発話評価値が 8 人中 5 番目に過ぎなかった極端に高い修復能力を持つ被験者 A が、発話計画可能時になると劇的に発話が改善され、最も高い発話評価値且つ、発話改善率を示したのはこの結論に一定の説得力を与えている。

第5章 結論

5.1 本研究の成果

本稿では「優れた発話技能を持つ人物は同時に優れた音声知覚能力も持っているのではないだろうか？」という仮説から端を発し、早口言葉と音韻修復現象を用いることによって生成と知覚のメカニズムの解明を試みた。その結果、第3章では実際に高い発話技能を持っている人物は同時に高い知覚能力も保持していること、また発話しやすい文章、本人にとって発話しやすい箇所では同時に知覚もしやすいということが明らかになった。さらに第4章では発話過程の発話計画が音韻修復時の予測に影響を及ぼしており、高い発話計画能力を持つ被験者ほど高い予測能力、ひいては高い認知能力を持っていることが明らかになった。これらの結果はスピーチ・チェーンの生成と知覚において、両プロセスに相互作用があるという多くの先行研究で示唆されている知見と一致している。また本成果は知覚運動理論を直接的に証明する証拠とはなっていないものの、発話と知覚で予測という共通のプロセスを共有していることを示したことによって一定の支持を与えている。このことは今後音声知覚の運動理論及び認知メカニズムの解明に繋がることが期待される。

5.2 今後の課題

本研究では発話と知覚の関連について一部明らかにしたが、これで生成と知覚のメカニズムの全貌が明らかになったとは言い難い。発話と知覚で予測というプロセスを共有しているのであれば、両過程でそれぞれの程度依存しているのか、そこで使われている共通のコードはなんであるか等、明らかにしなくてはならない課題は依然多い。また本研究では早口言葉という特殊な文章を用いて発話について評価したが、他の尺度でも発話と知覚に同様の相関が見られるか調査し、議論をより一般化させる必要がある。早口言葉を用いないで被験者の発話技能を評価することは工夫が必要だが、例えば被験者に発話に熟達しているアナウンサーと一般人を用いて音韻修復実験を行い、その修復率を比較するといった検証が考えられる。

謝辞

本研究を遂行するに当たりご多忙の中、終始懇切な御指導，惜しみない御援助を賜りました党建武教授に深く感謝すると共に厚く御礼申し上げます。また、徳田功准教授には本研究の問題点への適切な助言や御指導を頂き、厚く御礼申し上げます。さらに、研究活動を行う上で多くの暖かい御厚意を頂きました党研究室の皆様にご心よりの感謝を致します。

参考文献

- [1] P.B. Denes, Elliot N. Pinson, *The speech chain: The physics and biology of spoken language*, Worth Publishers (1993).
- [2] 柏野牧夫, “「スピーチ・チェイン」と脳”, 電子情報通信学会技術研究報告, Vol.106, No.332(20061102) pp. 23-26 (2006).
- [3] H.Lane, B.Tranel, “The Lombard sign and the role of hearing speech”, *Journal of Speech and Hearing Research*, 12, 677-709 (1971).
- [4] 内山田太一, “変形聴覚フィードバックに対する発話系の反応の計測”, 修士論文 (2006).
- [5] 山田玲子, “第二言語の音声学習: 知覚と生成および処理階層間の相互作用” 電子情報通信学会技術研究報, Vol.104, No.503, pp. 41-46 (2004).
- [6] L. Fadiga, L. Fadiga L, Craighero L, Buccino G, Rizzolatti G., “Speech listening specifically modulates the excitability of tongue muscles: a TMS study”, *Eur. J. Neurosci.*, 15, 399-402 (2002).
- [7] Wilson, S.M, Saygin, A.P, Sereno, M.I, Iacoboni, M., “Listening to speech activates motor areas involved in speech production,” *Nat. Neurosci.*, 7, 701-702 (2004).
- [8] A.M. Liberman, I. G. Mattingly, “The motor theory of speech perception revised,” *Cognition*, 21, 1-36 (1985).
- [9] 柏野牧夫, “音声知覚の運動理論をめぐって”, 日本音響学会秋季論文集, PP243-246 (2004).
- [10] 菅野隆, “音韻修復に影響を与える言語的要因の調査”, 修士論文 (2006)
- [11] “早口言葉 720 発 七度返シ,舌ヅリ”
<http://www.nobi.or.jp/i/kotoba/hayakuchi/index.html>

[12] 天野成昭, 近藤公, “NTT データベースシリーズ 日本語の語彙特性 第 1 期 CD-ROM 版単語親密度”, 三省堂, (2003)

[13] 柏野牧夫, “音韻修復”, 日本音響学会誌, Vol.61, No.5, pp263-268 (2005)