

Title	音韻修復に影響を及ぼす言語的要因の調査
Author(s)	菅野, 隆
Citation	
Issue Date	2006-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1984
Rights	
Description	Supervisor:党 建武, 情報科学研究科, 修士

修 士 論 文

音韻修復に影響を与える言語的要因の調査

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

菅野 隆

2006 年 2 月

修 士 論 文

音韻修復に影響を与える言語的要因の調査

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 助教授
審査委員 鵜木祐史 助教授

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

410065 菅野 隆

提出年月: 2006 年 2 月

概 要

本研究は音韻修復と呼ばれる現象の調査を通して、認知メカニズムの解明を試みた。本稿ではいくつかの日本語連続音声文を用い、周期的な無音区間または白色雑音を挿入して音声試料を作成した。これを聴かせ、日本語の母国語話者と非母国語話者で行われる音韻修復の違いについて四つの実験を通し、言語情報の利用の違いを考察した。その結果、言語情報が音韻修復に与える影響は大きく、文章の予測が主な認知の手がかりになっていたことを示唆し、音韻修復現象の解明自体が人の認知過程の解明に多くの知見を提供出来る可能性を示した。

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 音韻修復現象における言語情報の役割	3
2.1 音韻修復における了解度の向上要因	3
2.2 音韻修復と言語情報との関連性	4
第3章 日本語の特色と音韻修復	5
3.1 他言語との比較	5
3.2 母国語話者と他言語話者の比較実験	5
3.2.1 実験プロセス1	5
3.2.2 実験条件1	6
3.2.3 実験結果1	6
3.3 考察1	10
3.3.1 長音と促音の音韻修復	10
3.3.2 助詞の音韻修復	11
3.3.3 雑音挿入による逆効果	11
第4章 音情報と言語情報	12
4.1 聴取実験の再調査	12
4.1.1 実験プロセス2	12
4.1.2 実験条件2	12
4.1.3 実験結果2	13
4.2 考察2	16
第5章 文章の完成度と理解	17
5.1 言語情報量の操作	17
5.2 文章理解と認知	18
5.3 言語情報量操作を行った聴取実験	18
5.3.1 実験プロセス3	18
5.3.2 実験条件3	18

5.3.3	実験結果 3-1：文章の完成度と修復度	19
5.3.4	実験結果 3-2：認知に至った主要因	19
5.4	考察 3	19
第 6 章	言語修得度と文章予測	23
6.1	文章理解度と予測	23
6.2	言語情報と言語修得度	23
6.3	言語修得度の違いによる文章予測を対象とした聴取実験	23
6.3.1	実験プロセス 4	23
6.3.2	実験条件 4	24
6.3.3	実験結果 4	25
6.4	考察 4	25
第 7 章	まとめ	26
7.1	結論	26
7.2	今後の課題	26
謝辞		28

第1章 序論

1.1 研究の目的と背景

雑音環境において、人間の優れている能力の一つは音韻修復機能である。この音韻修復とは、突発的な雑音などで音韻情報が欠落したにも関わらず、その音韻がある条件下で知覚されるという現象である。この能力のおかげで、人間は日常生活において数多くの不安定な音韻を、安定して知覚することが出来る。このような人間の持つ雑音に対する頑健性を音声認識システムで活かすことが出来れば、実環境での使用に耐えうる認識性能が期待でき、高度なマン・マシン・インターフェースの実現に向けての大きな前進となる。そのためには、まず、人間の持つ音韻修復の認知メカニズムを解明する必要がある。

音韻修復機能に関する研究は1970年代から行われてきた [1]。従来研究の大半は、音韻修復現象の起きる条件、特に音声の欠落区間に雑音を挿入することによる効果に着目していた [2] [3]。主な研究では、音韻知覚の手がかりはスペクトルの時間的推移に基づく音声の特徴に着目しているが、この特徴量が我々の知覚系の中の表現と良く一致している保証がないため、このような音韻知覚過程の頑健性を説明することは困難である [4]。音韻修復と言語との関連性に関する研究も英語をはじめ、応用に対しても数多く行われてきた [5][6]。日本語は英語と音韻体型も発声の仕方も異なるため、音響的にも知覚的にも英語における知見がそのまま当てはまるとは限らないが、意味的文脈等の知識が了解度を高めるプロセスに貢献していると示唆されている [7]。これらの研究は主に音韻修復を音声の知覚（聴覚）過程として着目していた。ところが、音韻修復は音声の知覚過程だけではなく、音声の認知過程でもある。後者の場合、聴者は欠落音韻を修復する際、音以外の手がかり、例えば、言語情報や知識などをどうやって利用しているのか、すなわち、音韻修復の現象自身でなく、そのメカニズムはどうなっているのかを考えてゆく必要がある。それゆえ、本稿では言語情報の音声修復への影響に着目して考察する。

本研究では、いくつかの文を用い、周期的な無音区間または白色雑音を挿入して音声試料を作成した。これを聴かせ、日本語の母国語話者と非母国語話者で行われる音韻修復の違いについて考察した。この様に、言語に対する理解の度合いと音韻修復との関係に基づいて音韻修復のメカニズムの解明を試みる。ただし、我々の目標は音韻修復に関わっているメカニズムの究明であるので、無音と雑音を挿入した音声試料による修復率の差異は本研究の主な議論対象にしない。

1.2 論文構成

本論文は、第2章で音韻修復現象の現在までに確認されている知見から言語情報への関連について述べ、第3章に母国語の違いによる音韻修復、第4章に音韻修復における音情報と言語情報、第5章に音韻修復における言語情報量と文章理解、第6章に言語修得度と予測による音韻修復の順に述べ、そして最後に結論をまとめる。

第2章 音韻修復現象における言語情報の役割

2.1 音韻修復における了解度の向上要因

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であるか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名詞的存在である。図 2.2 に画像によるイメージをしめす。

図 2.1 の左側の状態では、欠損部分が多く何が書いてあるのか読み取ることは難しいが、右側のように欠損部分に別の何かを補うことで、何が書いてあるのか理解し易いはずである。(左上部→A：右上部→B：左下部→C：右下部→D と書いてある) 欠損した文字を欠損した音声に、補った別のものを雑音へと音声に置き換えたものが音韻修復現象である(図 2.2 参照)。

人は欠損区間に雑音を挿入することで、失ったはずの音韻を聞き取ることが出来る。この現象の起源については様々な説があるが、最も初期の報告として Miller and Licklider による「垣根効果」であるされている。これは音声信号の中断部分に白色雑音を挿入することで、無音区間を設定した状態に比べ、音声が続くかのように聞こえ、あたかも垣根を通して景色を見たかのように聞こえたためにこのように呼ばれた。しかしこの報告では了解度は変化しないとされている。これは意味のある文ではなく、単音節単語のリストを読み上げた物を刺激として用いたことが要因になっていた可能性がある。その後、雑音挿

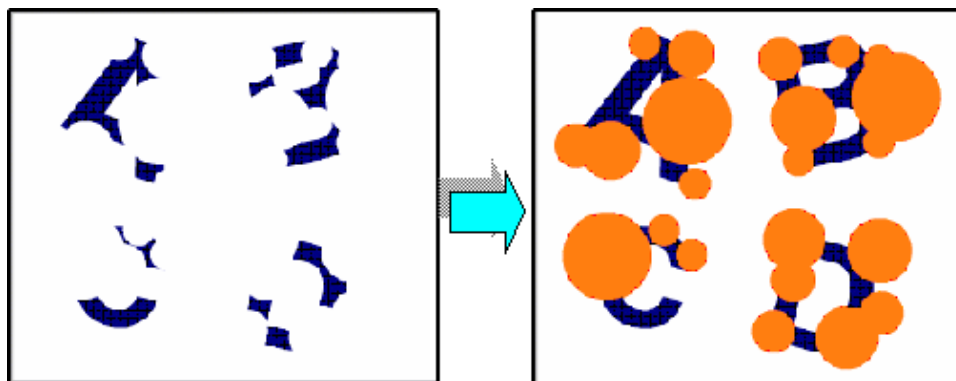


図 2.1: 音韻修復の画像によるイメージ

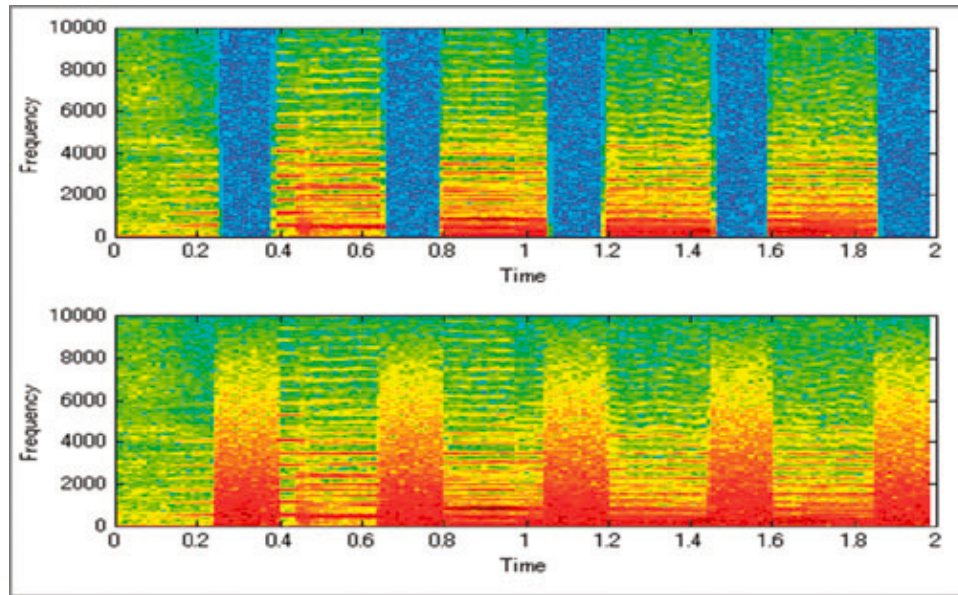


図 2.2: 音韻修復の音声によるイメージ

入によって了解度も向上するとされたが、連続性・自然性の向上と了解度の向上とは別のプロセスが知覚に働いていると考えられてきた。連続性を知覚するプロセスには音一般に働く聴覚特性（マスキングやラウドネス低下等）が関与しているとされている。これに対し、了解度が向上するプロセスには、文脈情報等を含む言語情報や調音結合といった知識が、知覚の手がかりになっていると考えられる [7]。しかしながら、言語の特性は世界共通ではなく手がかりの傾向は言語によって比重が変わる。このため、了解度の向上についても違いがあると言える。

2.2 音韻修復と言語情報との関連性

了解度が向上するプロセスには様々な情報が密接に関わっている。英語であれば、リズム・音程等が重要な要素となっており、言語情報は単に修復効果を引き上げるための独立した要因と考えられている。しかし、日本語の場合一般に音節構造が単純なうえ、ヨーロッパ語にある強弱アクセントがないという言語特性であるために、認識するにあたり文の意味を理解することの比重が他言語に比べ大きいとされ、日本語の音韻修復における了解度が向上するプロセスにも言語情報が重要な手がかりになっていると考えられる。そのため、音韻修復の解明に対し、日本語による音韻修復と言語情報の結びつきを調査することは、認知の手がかりを探る意味でも有用な手段になると考えられる。

第3章 日本語の特色と音韻修復

3.1 他言語との比較

本節では言語に対する理解の度合と音韻修復との関係について議論する．言語情報が認知の手がかりになるのであれば，まず異なる言語情報を持つ人との比較を行い，言語情報が使われているのかを調査する必要がある．本稿では，日本語の言語情報をもつ母国語話者と，他言語話者として中国語話者を対象として比較実験を行った．言語的な違いとして，中国語には助詞の概念がなく，促音や長音が短い．母国語話者と他言語話者について，これらの音がスムーズに修復されるかどうか，言語情報を使っているかどうかの指標となる．

3.2 母国語話者と他言語話者の比較実験

日本語の母国語話者と他言語話者における音韻修復の違いについて，日本語連続音声に周期的な無音・雑音区間を設定した音声資料を用いた聴取実験を行う．

3.2.1 実験プロセス 1

実験1では，被験者は日本語が母国語である話者5名と中国語話者5名を選択した．5名の中国語話者の間に日本語の習得状況に個人差はあるが，いずれも日本語音声をも音素表記で書くことができる．以降では中国語話者を他言語グループと称する．音声試料は，ATR 日本語連続音声データベース setB から，長さが約4秒間である文章から3文選んで，それぞれに200msの周期で100msの音声を無音で置換した．その結果，全音声区間の1/2が周期的に欠損する状況となっている．また，上記の無音区間を，音声の平均パワーより15～20dB上回る白色雑音に入れ替えて音声試料を作成した．その例を図3.1に示す．

実験では，まず，周期的に無音と入れ替えた音声試料を聴かせ，聴き取れた音をひらがな，漢字，または音素表記で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確に集計するため，正しく認知出来ない場合，または聞こえなかった場合，その箇所は空欄にしてもらった．次に上記の無音区間を白色雑音に入れ替えた音声試料を用いて，同様の聴き取り実験を行った．聴取実験では，音声試料を繰り返し聴き直すことが可能である．なお，文章が長く，一度に全文を書き出すことが困難であるため，文章を適当

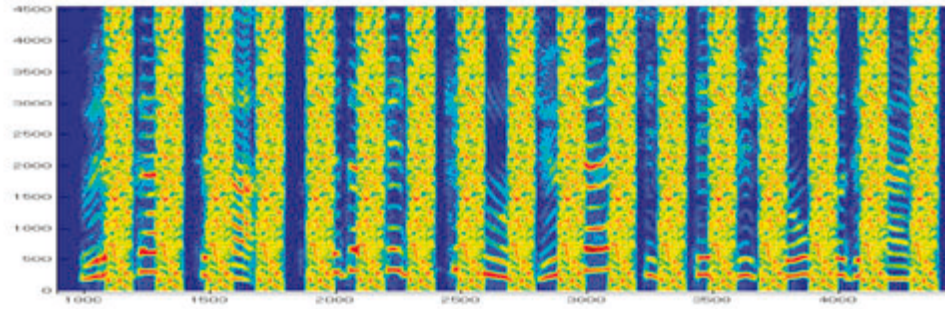


図 3.1: 音声資料のサウンドスペクトルグラム

表 3.1: 実験条件 1

音声データ	A T R 音声データベース setB
音声欠損	200ms 周期, 100ms 間
欠損区間	白色雑音または無音の挿入
被験者	日本語話者 5 名, 中国語話者 5 名
聴取条件	ヘッドホン (Sony:MDR-Z600)

な 3 つの文節に区切り, 先行の文節の聴取後に後続の文節をつなげて, 追加した部分のみを評価する方式とした。

3.2.2 実験条件 1

実験条件を表 3.1 にまとめる。

音声資料には比較的文章長の近い「文章 1: 裏返せばそれだけ事件を重大視している証拠である (49 音素)」「文章 2: 天然記念物級の規模という学者の発言もあった (49 音素)」「文章 3: 勝てる試合に何故勝てそうなピッチャーを先発させないのだろうか (55 音素)」を用いた。また, 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おり, Matlab を使用し音声資料を作成した。また, 実験風景を図に示す。

3.2.3 実験結果 1

各文章の音素ごとの正解率を図 3.3~3.5 に, 無音区間時と雑音挿入時の正解率の差分を図 3.6~3.8 に示す横軸は各文章の音素を順に並べており, 縦軸はその正解率である。また, それぞれの音素の左部は無音区間設定時の正解率で斜線で表し, 右部は雑音挿入時の正解率で黒塗りで表している。そして, 各文章のグラフ上段は母国語グループの結果, 下



図 3.2: 実験風景

段は他言語グループの結果である．また，差分は無音区間設定時と雑音挿入時の正解数の差分であり，両グループの修復の効果を表している．表 3.2 は各文章全体の無音区間時と雑音挿入時の正解率（無音区間時の正解率 [%] / 雑音挿入時正解率 [%]）と，それによる誤り削減率（括弧内 [%]）を表している．ここで正解率及び誤り削減率は次のように定義している．

- ・ 正解数 = $\frac{\text{正解音素数}}{\text{総音素数}}$
- ・ 誤り削減率 = $(1 - \frac{\text{雑音挿入後の正解数}}{\text{雑音挿入前の正解数}})$

表 3.2: 正解率統計

被験者	文章 1	文章 2	文章 3
母国語話者	84.4 / 2.2 (50.0)	67.6 / 78.0 (32.1)	84.0 / 94.5 (59.4)
他言語話者	38.8 / 59.2 (35.0)	47.2 / 57.2 (18.9)	55.6 / 65.2 (21.6)

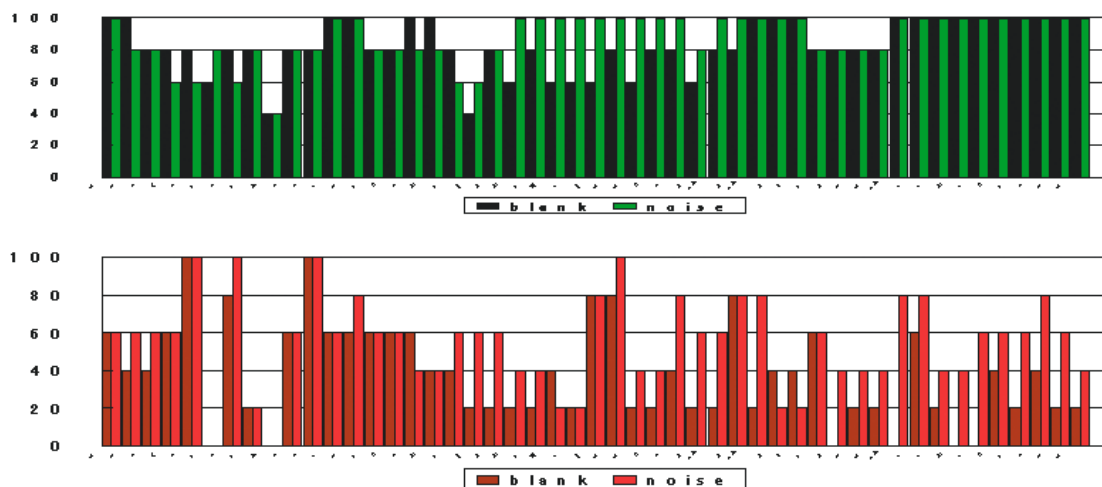


図 3.3: 文章 1 の各音素の正解率

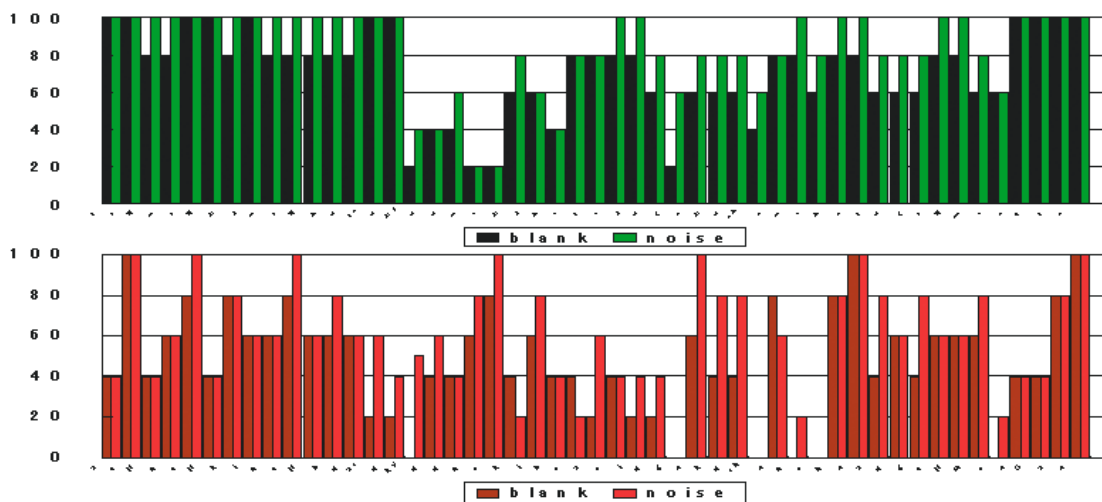


図 3.4: 文章 2 の各音素の正解率

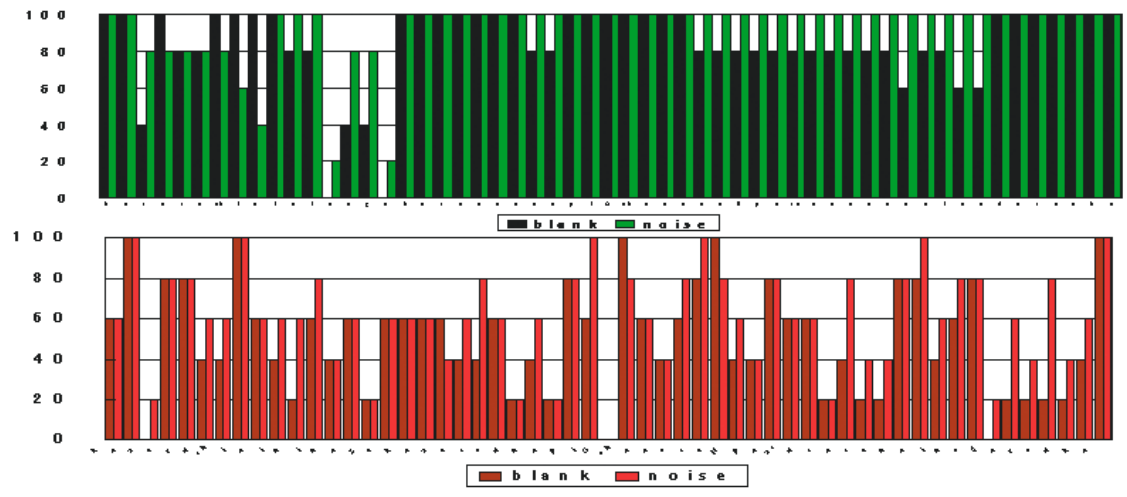


図 3.5: 文章 3 の各音素の正解率

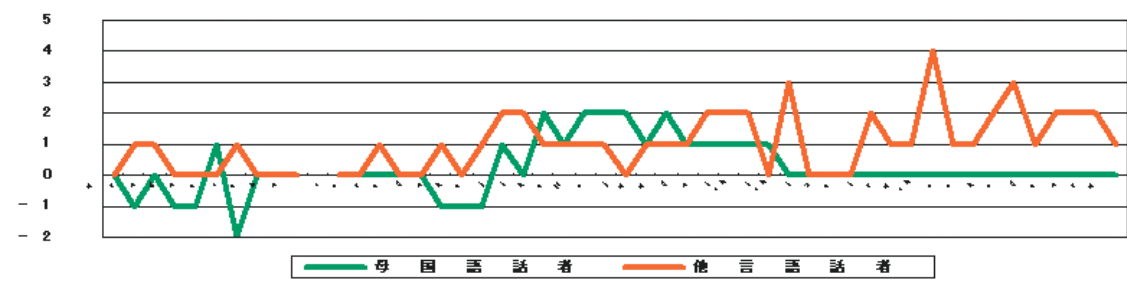


図 3.6: 文章 1 の正解率差分

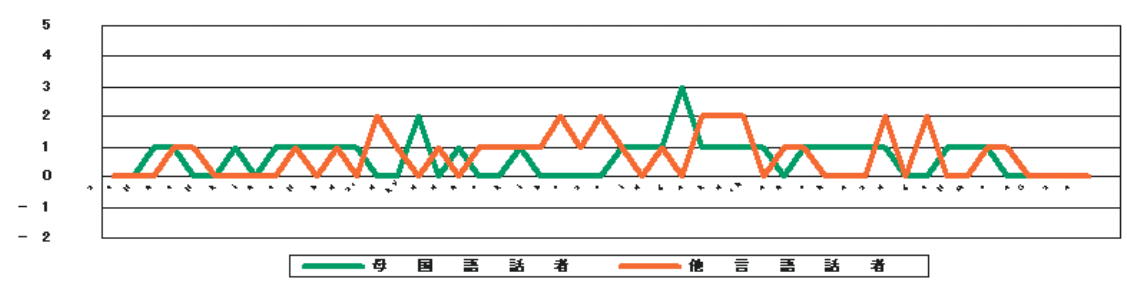


図 3.7: 文章 2 の正解率差分

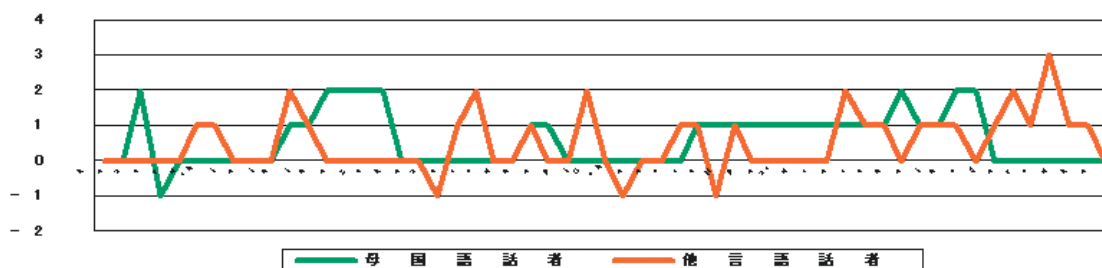


図 3.8: 文章 3 の正解率差分

表 3.3: 長音と促音の正解率統計

被験者	長音	促音
母国語話者	72.0 /85.0 (46.4)	100.0 /100.0 (0.0)
他言語話者	40.0 /44.0 (6.7)	50.0 /50.0 (0.0)

全体的に、雑音挿入の有無に関係なく母国語グループの正解率は他言語グループよりはるかに高くなっている。この結果より、言語情報が音声修復によく使われていることが分かる。

3.3 考察 1

3.3.1 長音と促音の音韻修復

日本語の言語的特徴として長音と促音が存在する。もし言語情報が重要であるとしたら、それらの音韻の修復は他言語話者にとって困難であると予想される。表 3.3 では、長音 12 音素、促音 2 音素に対する無音区間時と雑音挿入時の正解率 (無音区間時の正解率 [%] / 雑音挿入時正解率 [%]) と、それによる誤り削減率 (括弧内 [%]) を表している。

長音と促音のいずれにおいても、日本語話者は他言語話者よりかなり高い正解率を得た。これは日本語話者が日本語に関する言語的特徴を自然に理解しているからである。日本語話者の場合、雑音の挿入により音以外の情報を利用して長音の音韻修復誤りをさらに 46.4%削減している。それに対して他言語話者は無音時の正解率と誤り削減率共に低かった。これは言語情報が音韻修復に重要であることを示している。しかしながら、促音に関しては実験で用いた音素数が少ないため、今後これに注目した実験を行う必要がある。

表 3.4: 助詞の正解率統計

被験者	助詞
母国語話者	71.3 / 86.3 (52.2)
他言語話者	38.8 / 52.2 (24.1)

3.3.2 助詞の音韻修復

本実験での他言語グループは、全員中国語話者となっている。中国語には助詞という概念がないので、音声試料にある助詞の修復度合いは、言語情報を利用しているかどうかの指標となる。表 3.4 では、助詞 17 音素に対する無音区間時と雑音挿入時の正解率 (無音区間時の正解率 [%] / 雑音挿入時正解率 [%]) と、それによる誤り削減率 (括弧内 [%]) を表している。

雑音を挿入することで、被験者は音以外の情報を用いるようになり、正解率が向上している。結果では非母国語話者の場合、音以外の情報による誤り削減率は 24.1%であったのに対し、母国語話者の場合では 52.2%となっている。その差は、欠損した助詞に対して母国語話者は、言語情報を利用して復元することが可能であるが、中国語話者はそのような修復がやや困難であることを示唆している。

3.3.3 雑音挿入による逆効果

本実験では、雑音を挿入することで、本来修復された音韻を逆に認識出来なくなってしまうケースがあった。その詳細を見てみると、非母国語話者の場合、雑音を挿入することで、もともと聞こえた音が完全に聞こえなくなってしまう。母国語話者の場合、雑音の挿入により、文章の先頭部分の正解率が悪化した。それらの現象について以下の原因が考えられる。通常の意味の“音韻修復”は、音声がある程度欠落すると断続的な音素片のように聞こえて言語的な意味の理解に支障をきたしたとしても、音声欠落区間に雑音を入れることによって音声が連続的に聞こえ、言語情報を理解しやすくなる現象である。この場合、雑音の役目は欠落した音の修復ではなく、“元の音”が隠されたような錯覚を起こさせることであり、聞き手が言語情報や知識などを積極的に利用してその欠落した文を補充する。もし音声の欠落区間が短くて聴覚的に断続的な音素片になっていない場合、あるいは修復に利用できる言語情報等が無かった場合、雑音の挿入は音韻の修復度合を向上させる役目よりむしろ悪化させる可能性がある。

第4章 音情報と言語情報

4.1 聴取実験の再調査

最初の実験では、無音区間設定時ですでに正解率が高く、雑音を挿入したことによる効果が判別しづらいという結果であった。そのため、文章の聴取難易度を上げ、音韻修復の効果を明確化させる必要がある。これは、母国語話者と他言語話者の音韻修復の傾向を明確化することで、音情報と言語情報の手がかりの比重を調査することに繋がる。

4.1.1 実験プロセス2

音韻を修復する際には音情報と言語情報のどちらを使っているのかを判断するため、欠損区間を 100ms から 150ms と増やし、周期を 300ms にした。また、被験者の 10 人も同じであるため、別の文章を用意した。

4.1.2 実験条件2

実験条件を表 4.1 にまとめる。

音声資料には新たに用意した「文章4：彼も初球を本塁打し劇的なサヨナラ勝ちを収めた(52音素)」「文章5：だまされる方の知識水準に問題があるのでございます(53音素)」「文章6：己を超えなければ勝利の女神は微笑んでくれないらしい(52音素)」を用いた。

表 4.1: 実験条件2

音声データ	A T R 音声データベース setB
音声欠損	300ms 周期, 150ms 間
欠損区間	白色雑音または無音の挿入
被験者	日本語話者 5 名, 中国語話者 5 名
聴取条件	ヘッドホン (Sony:MDR-Z600)

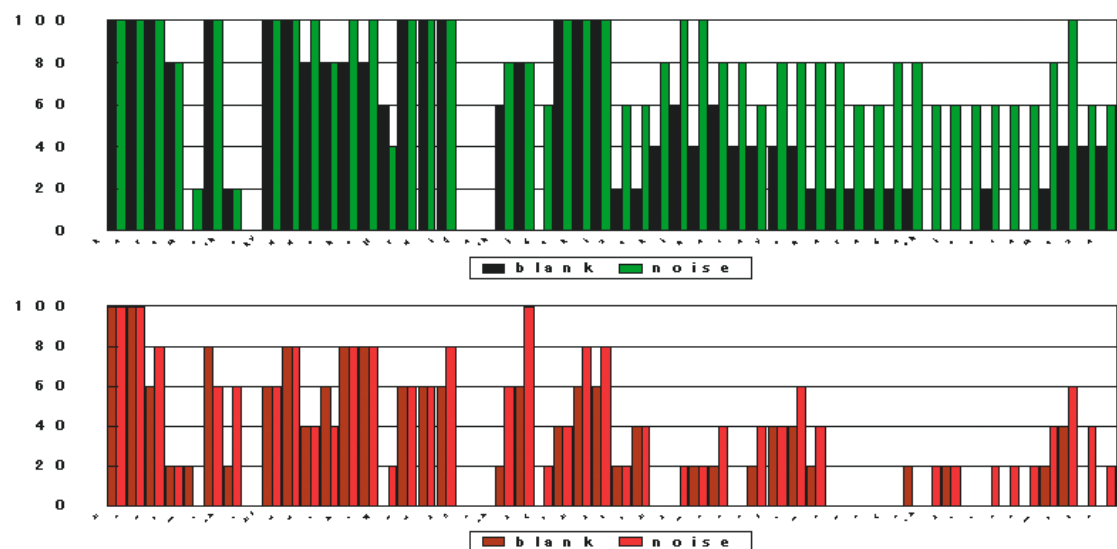


図 4.1: 文章 4 の各音素の正解率

表 4.2: 正解率統計

被験者	文章 1	文章 2	文章 3
母国語話者	82.8 / 90.3 (43.6)	69.2 / 79.2 (32.5)	88.0 / 92.0 (33.3)
他言語話者	37.9 / 56.6 (30.1)	53.3 / 60.0 (14.0)	60.7 / 64.7 (10.2)

4.1.3 実験結果 2

各文章の音素ごとの正解率を図 4.1～4.3 に、無音区間時と雑音挿入時の正解率の差分を図 4.4～4.6 に示す。横軸は各文章の音素を順に並べており、縦軸はその正解率である。また、それぞれの音素の左部は無音区間設定時の正解率で斜線で表し、右部は雑音挿入時の正解率で黒塗りで表している。そして、各文章のグラフ上段は母国語グループの結果、下段は他言語グループの結果である。また、差分は無音区間設定時と雑音挿入時の正解数の差分であり、両グループの修復の効果を表している。表 4.2 は各文章全体の無音区間時と雑音挿入時の正解率 (無音区間時の正解率 [%] / 雑音挿入時正解率 [%]) と、それによる誤り削減率 (括弧内 [%]) を表している。

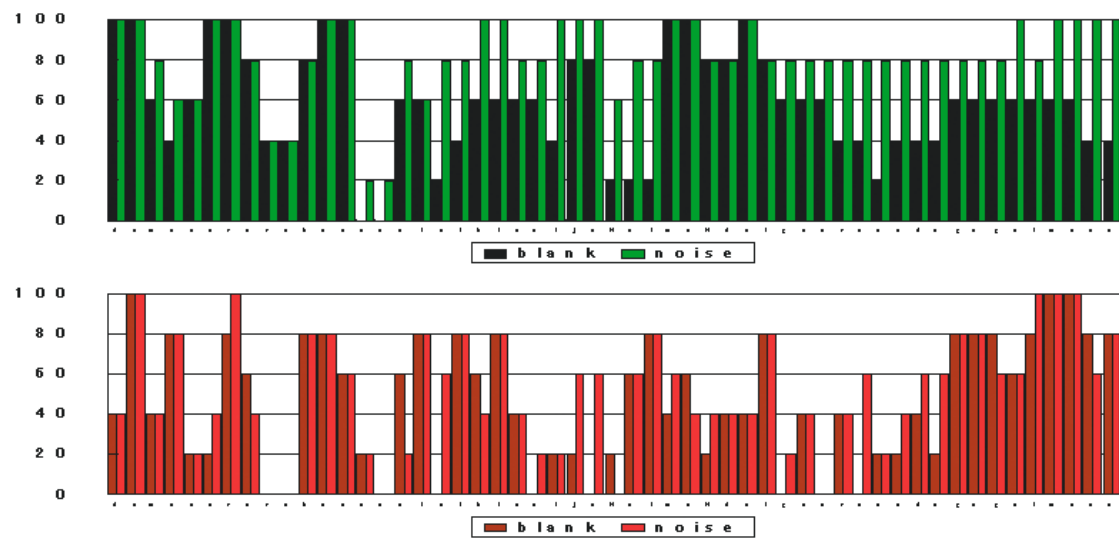


図 4.2: 文章 5 の各音素の正解率

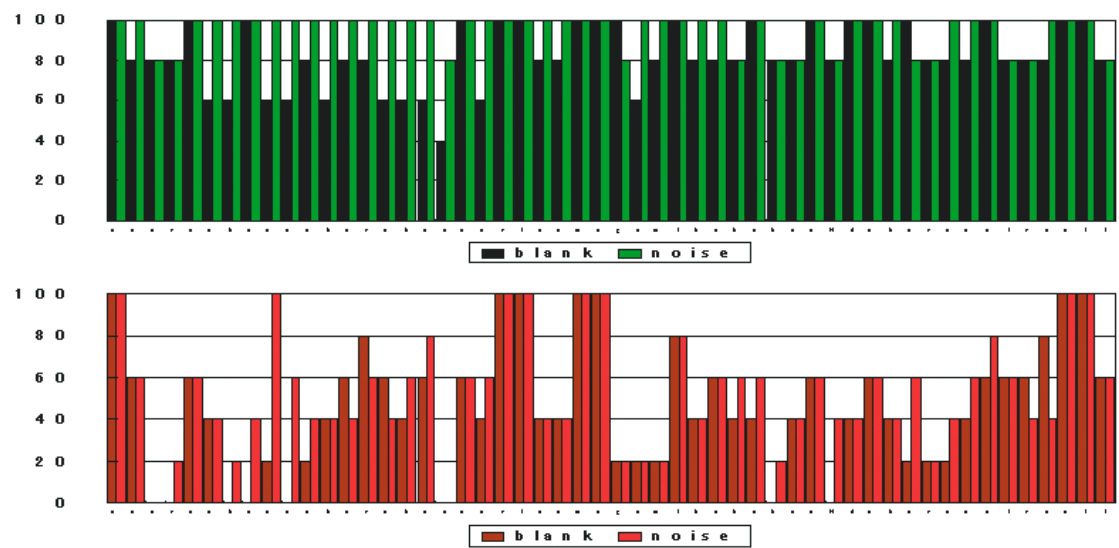


図 4.3: 文章 6 の各音素の正解率

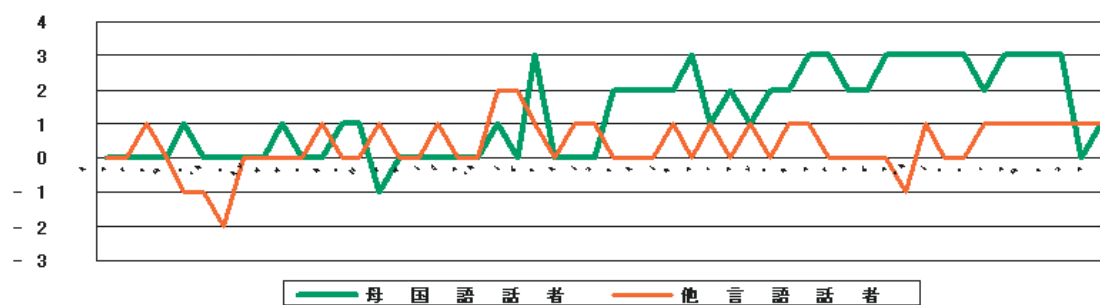


図 4.4: 文章 4 の正解率差分

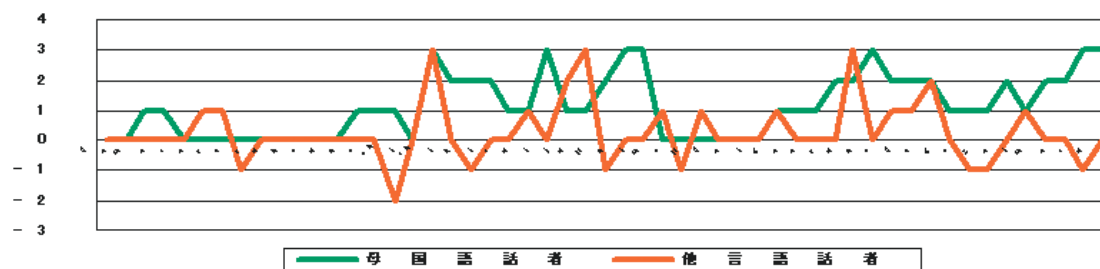


図 4.5: 文章 5 の正解率差分

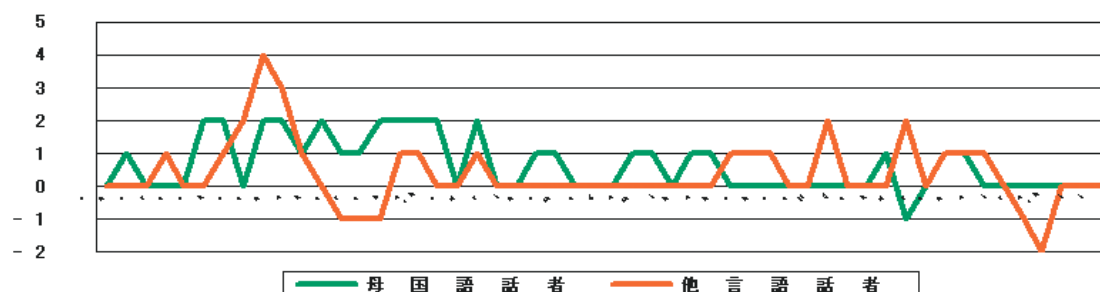


図 4.6: 文章 6 の正解率差分

4.2 考察2

欠損区間を 150ms に増やすことで、情報量をひどく損なっている。無音設定時では、両グループとも文の後半部の正解率がかなり悪くなったが、雑音の挿入により母国語話者の正解率が大きく向上した。それに対して、他言語グループの正解率はほとんど改善できなかった。この差は言語情報を持っているかどうかによるものと思われる。この結果により、日本語における音声資料への雑音挿入による通常の意味での音韻修復は、主に言語情報の寄与が大きいと言える。

しかし、文章6については文章の内容的に聴取難易度が低くなっている。これは「勝利の女神が…」という音声を少しでも聞いてしまうと「…微笑む」という予測がし易くなっているために、前実験と同様な無音区間時の正解率が高い結果を得ている。このため、今後文章の内容にも深く吟味する必要がある。

第5章 文章の完成度と理解

5.1 言語情報量の操作

これまでの実験結果の中で、文章量を増加したことで誤りが訂正されることが度々あった。実験1で用いた文章1の先頭部の例を図5.1示す。

図5.1では、母国語話者の修復効果が、後続文節を追加するたびに増加していることが分かる。「裏返せば」という単語であるが、実際には「うらめしさ」、「うらしま」、「うらがしま」等と聞こえていた音が、後続音節を接続することで「裏返した」や正解である「裏返せば」と聞き取り、正解率が上昇していた。また、無音設定時ではさほど正解率は変動していない。これは聴取実験の手法上、後続文節を次々追加してゆき長文の聞き取りを行い、先行文節について正解率が上昇していたために、文章の始めの文節は終わりの文節よりも聴取回数が必然的に多いという要因がある。しかし、被験者にはその時点で納得するまで音声資料を何度も繰り返し聞いているため、文節が増えるたびに認識が変化していたと考えるのが自然である。言い換えると、同じ文節を聞いているにもかかわらず、文章の完成度が増すことで、文のまとまりを理解し、ふさわしい言葉を埋めようとすることで、聞き取る音に変動する可能性がある。それには、文章の完成度を上昇させることで、雑音置換の効果が表れると考えられる。

そのため本節では、同じ音声聞いた状態でも変化して聞こえる点に着目し、言語情報量を操作することで聞こえ方及びそのときの文章理解について調査を行う。

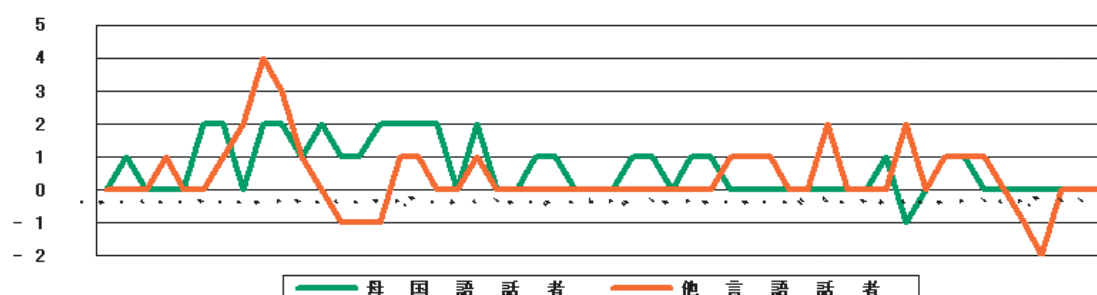


図 5.1: 先行文節の訂正

5.2 文章理解と認知

以前から音韻修復は言語的要因の影響の方が強いとされている [1]. これは、同一の音声聞いても、意味のある文になっているほど、その音の雑音置換による検知限が長くなることである。これまで接続する文節が増えることで誤りが改善する点、文章後半部の修復度が高い点から、文章が長くなることで正解率が上昇したのは、文章の完成度が高くなることで情報量が増え、理解が深まることに起因すると考えられる。

5.3 言語情報量操作を行った聴取実験

言語情報量と音韻修復の関係を調査する。音声試料の作成では、前もって文を幾つかの文節に分けて、文節の組み合わせにより試料の言語情報量を変化させた。また、その組み合わせによる文章の理解を特定の要因に分け調査を行う。

5.3.1 実験プロセス 3

音声試料は、新たに用意した3つの文章をそれぞれ4つの文節に分割し、時間順に A, B, C, D とする。被験者には D, C, B, A の順でそれぞれ聴かせ、聴き取れた音を書き出してもらった。逆順にしたのは先行文節の理解による言語情報を抑えるためである。次に文節を CD, BC, AB, BCD, ABC, ABCD のように接続し、順に提示して聴取実験を行った。これらのパターンについて、欠損区間を無音に設定した場合の聴取実験を終了した後、雑音を加えて同様な実験を行った。被験者には聴取結果に変化が無くなるまで何回も聴いてもらった。すなわち、被験者は確認のため、最低一回は繰り返して聴いている。また、正解した音素に対して、それぞれ認知するに至ったと思われる主要因を以下の3項目より選択してもらった。・「音」：音そのものが、その通りに聞こえた・「単語」：特定の単語を連想し、その単語の一部であると認識した・「文法」：文法的に理解した場合。特定の名詞の後の助詞や、構文の一部であると理解し聞こえた

5.3.2 実験条件 3

実験条件を表 5.1 にまとめる。

また、音声資料には「文章7：音楽が好きな俺のためにわざわざロックミュージックを用意してきてやった (64 音素)」「文章8：両手の指は変形し間接のあたりが瘤状に盛り上がっていた (59 音素)」「文章9：俺のプライベート時間を俺がどう過ごそうと全て俺の責任なんだ (61 音素)」を用意した。

表 5.1: 実験条件 3

音声データ	A T R 音声データベース setB
音声欠損	300ms 周期, 150ms 間
欠損区間	白色雑音または無音の挿入
被験者	日本語話者 5 名
聴取条件	ヘッドホン (Sony:MDR-Z600)

5.3.3 実験結果 3-1：文章の完成度と修復度

文章 7～9 の各文節における正解率（上段：無音時 / 下段：雑音時）をそれぞれ図 5.2～5.4 に示す。組み合わせは上から単音節，2 音節，3 音節，全音節時の順となっている。

5.3.4 実験結果 3-2：認知に至った主要因

音韻修復に至った要因「音/単語/文法」の占める内訳と聴取回数を図 5.5 に示す。横軸はそれぞれの組み合わせを示している。

5.4 考察 3

音声試料の作成では，単文節に対して前後の文節の追加により言語の情報量は増える。その効果は修復結果に表れた。特に先行文節を追加した場合，修復の効果が明らかに向上した。それは先行文節の追加により意味として分かりやすくなる。言い換えると，後続文節に比べ，先行文節の追加により言語の情報量がおおいに増加し，正解率を向上させた。ただし，今回の実験パラタイムでは，2～3 文節を接続して，後から提示した音声試料は先に聴取した音声試料の区間と重複があるので，同じ文節に対しての聴取回数は累積的に増えることになる。文節の組み合わせが長くなると，特に 3 つまたは 4 つに文節を組み合わせた場合，先行文節の聴取回数が自然に増える。すなわち，慣れによる影響も考えられ，正解率の増加はすべて言語情報によるものであるとは言い切れない。

また，それぞれのパターンに含む文節が多いほど，単語により修復した音韻の割合はかなり増加した。接続する区間数を増やすにつれ，「音」としてしか認識できなかった部分が「単語」・「文法」として認識するケースが増えている。つまり，言語情報量が増えることにより文章の理解度が増していると言える。特に文節 A の場合，図 5.5 に示した正解率が後続文節の追加によりそれほど増加していないが，単語により認知した割合は 2 倍程度に，文法による割合は 3 倍以上に増加した。それは言語情報が音韻修復に重要な役割を果たしていることを示している。

図 5.5 より，接続する文節を増やすにつれて，聴取回数は減少する傾向にある。同様に，

	A	B	C	D
・文節毎	75.7% 78.6%	32.6% 59.0%	92.9% 92.9%	58.8% 67.1%
・2文節接続			100.0% 100.0%	64.7% 72.9%
		54.3% 65.3%	87.4% 92.9%	
	75.7% 84.3%	54.7% 71.6%		
・3文節接続		52.6% 64.2%	100.0% 100.0%	72.9% 72.9%
	80.0% 84.3%	54.7% 72.6%	91.4% 92.9%	
・4文節接続				
	75.7% 85.7%	70.5% 72.6%	100.0% 100.0%	71.8% 75.3%

図 5.2: 文章 7 の正解率統計

	A	B	C	D
・文節毎	58.5% 72.3%	72.9% 82.4%	43.8% 56.3%	55.4% 60.0%
・2文節接続			53.8% 65.3%	69.2% 70.8%
		68.2% 81.2%	56.3% 65.3%	
	72.3% 72.3%	75.3% 84.7%		
・3文節接続		74.1% 84.7%	58.8% 66.3%	66.2% 69.2%
	72.3% 72.3%	76.5% 84.7%	55.0% 63.8%	
・4文節接続				
	72.3% 72.3%	74.1% 84.7%	60.0% 67.5%	69.2% 72.3%

図 5.3: 文章 8 の正解率統計

	A	B	C	D
・文節毎	66.7% 88.0%	68.6% 78.6%	76.0% 84.0%	62.4% 87.1%
・2文節接続			76.0% 88.0%	90.0% 91.8%
		72.9% 100.0%	82.7% 94.7%	
	88.0% 88.0%	92.9% 100.0%		
・3文節接続		91.4% 100.0%	81.3% 100.0%	82.4% 95.3%
	86.7% 88.0%	90.0% 100.0%	86.7% 100.0%	
・4文節接続				
	88.0% 89.3%	92.9% 100.0%	86.7% 100.0%	88.2% 97.7%

図 5.4: 文章 9 の正解率統計

文章を間違っ理解した場合でも、文章量が多い場合では、少ない場合よりも認識した結果の変動が少なく、誤認識した状態で落ち着く。このことから、文章量が少ない場合では予測の可能範囲が広く、認識結果が様々に変動しているのに対し、文章量が多い場合では言語情報を最大限に利用することで、予測可能範囲を狭めることが可能となり、聴取回数も減少し、正解率がある高い水準へ収束すると考えられる。

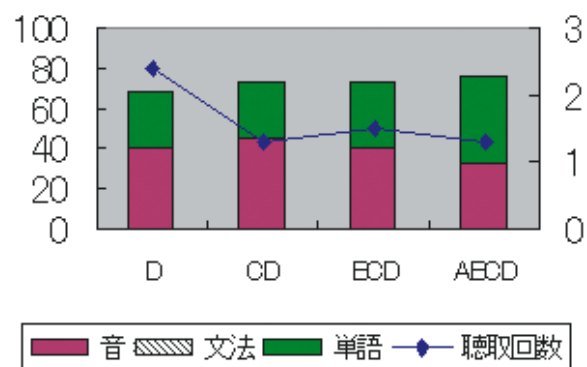
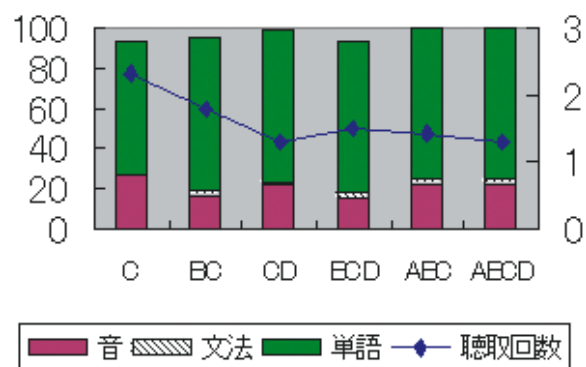
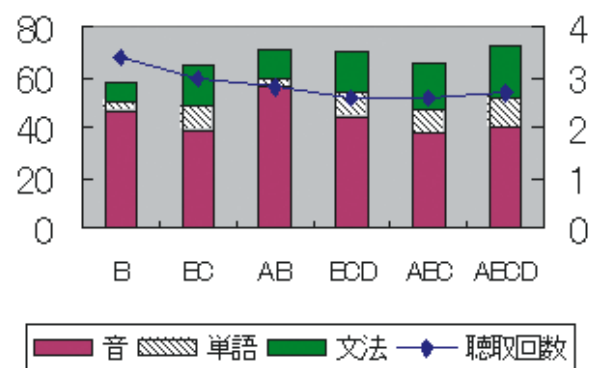
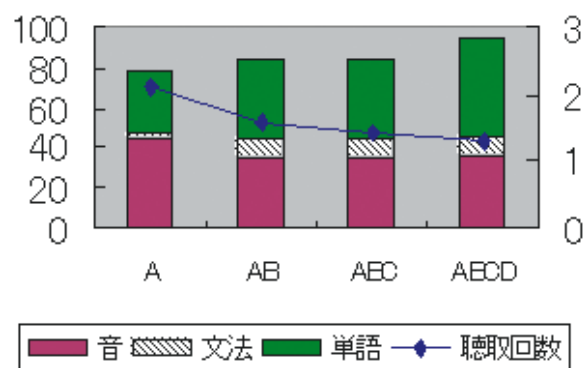


図 5.5: 主要因の内訳と聴取回数

第6章 言語修得度と文章予測

6.1 文章理解度と予測

音韻修復は基本的にリアルタイムで行われる。文章理解が雑音置換による検知限の上昇を行わせるためには、被験者は聴取中にそれまでの文章をある程度理解したうえで、次の文章を予測し音を聞き入れる準備を行っていると考えられる。そのため、言語情報は音の予測と直結し、逆に予測がはずれるような文章では、たとえ文章の完成度が高くても、かなり聴取が困難になると予想される。

6.2 言語情報と言語修得度

本研究の他言語話者の被験者についても日本語の言語情報を少なからず持っている。言語情報が音の予測と直結するのであれば、言語情報を多く持つ人ほど文章の予測は容易になり、正解率は上昇すると考えられる。そして、言語情報に頼るために予測が困難・裏切られるような文章では聞き取りづらくなり正解率は低下する。これに対し、言語情報を持たない人ではどの文章に対しても予測を行わないため、ランダム的な修復を行うと予想される。

6.3 言語修得度の違いによる文章予測を対象とした聴取実験

他言語話者であっても言語情報をある程度持つ被験者の場合、音韻修復の際に文脈の理解による単語の予測が少なからず存在すると考えられる。本実験では予測難易度の異なる文章2組用い、日本語の習得度合いと音韻修復率との関係を調査する。

6.3.1 実験プロセス4

予測し易い組は以下の3文からなる。「文章10：スピーチが下手なのだから原稿を用意した方がよい（45音素）」「文章11：お金を入れボタンを押すと切符が出てくる（34音素）」「文章12：運転免許をとって3年目になります（33音素）」そして、予測し難い組は以下の3文からなる。「文章13：ずっしりと手応えのある体験をつむ青年が増えて欲しい（50音素）」「文章14：山の裾に休息している哀れな漁村の屋根をみた（49音素）」「文章15：

表 6.1: 実験条件 4

音声データ	A T R 音声データベース setB
音声欠損	300ms 周期, 150ms 間
欠損区間	白色雑音または無音の挿入
被験者	日本語話者 2 名, 中国語話者 3 名
聴取条件	ヘッドホン (Sony:MDR-Z600)

表 6.2: 検査に用いた 50 問の漢字一覧

番号	漢字	番号	漢字	番号	漢字	番号	漢字	番号	漢字
1	文	11	数える	21	遊ぶ	31	破ける	41	争い
2	水	12	聞く	22	冷たい	32	妨げる	42	誘い
3	入る	13	売る	23	低い	33	育てる	43	貧しい
4	早い	14	北	24	断る	34	磨く	44	接する
5	鳥	15	表す	25	必ず	35	描く	45	潮
6	三	16	感じる	26	望み	36	携える	46	腰
7	男	17	決める	27	通る	37	腹	47	絡む
8	教える	18	登る	28	最も	38	株	48	旨
9	楽しむ	19	都	29	熱	39	秀でる	49	握る
10	馬	20	命	30	欲しい	40	置く	50	添う

缶詰はやや細めの 2 本のまきを使って引き上げた (50 音素)」上記の 6 文章を用いて音声資料とした。

日本語の習得にばらつきがあると思われる 3 名の中国語話者に、習得度を定のため事前に漢字の日本語読みのテストを受けてもらい修得度とした。これは中国人は漢字からその日本語の意味をほぼ理解できるため正しく読めるかどうかを基準になると考えたためである。また、日本と中国で使われる漢字に意味と書き方の一致がないものも存在するため、一致性のあるものを用意した。使用した漢字については、小学 1 年から中学 3 年までに習う常用漢字を、難易度が均等になるよう 50 問 (100 点) 選択した。参照のため 100 点を獲得した日本人被験者を 2 名加えている。

6.3.2 実験条件 4

実験条件を 6.1 にまとめる。

また、日本語の言語修得度検査を行うために用いた漢字を表 6.2 以下にまとめる。

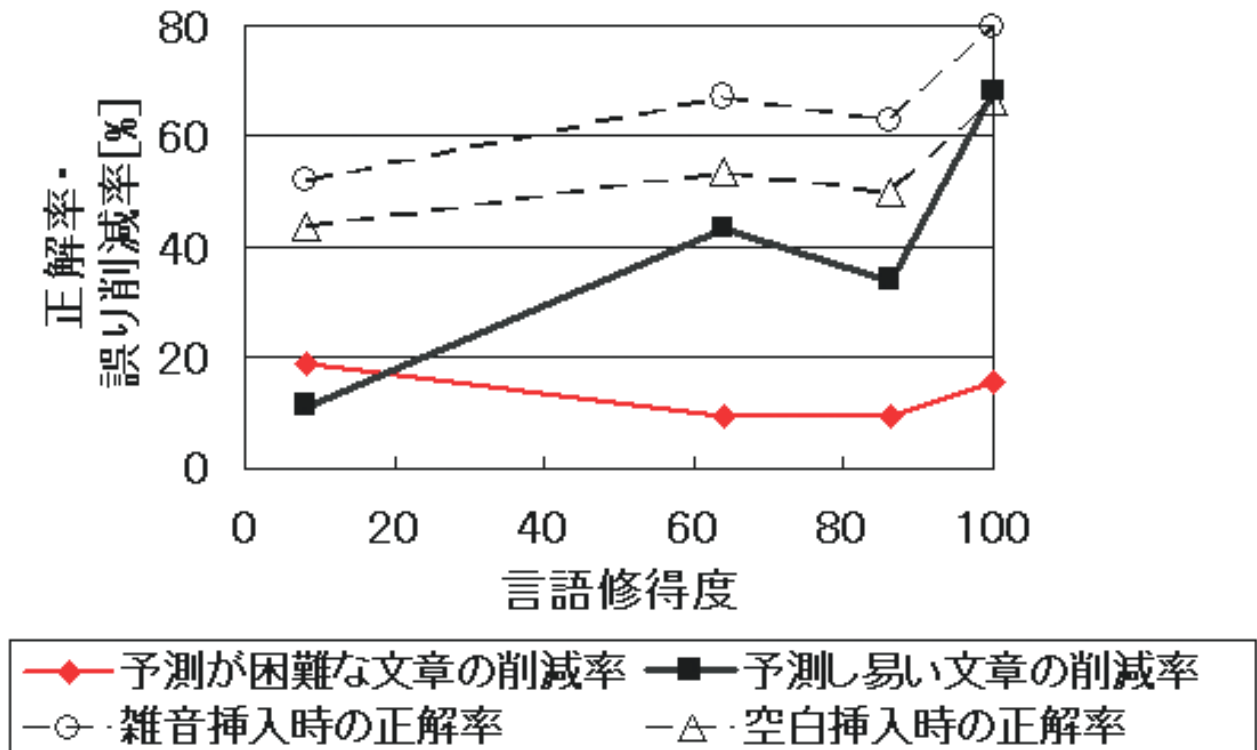


図 6.1: 言語修得度と正解率

6.3.3 実験結果 4

実験結果を図 6.1 にまとめる．縦軸は雑音/無音区間挿入時の正解率，及び音声の欠損区間に雑音を重畳したことによる誤りの削減率（修復の度合）である．日本人被験者 2 名の漢字テストは共に 100 点であり，正解率も同様な傾向にあったため 2 名の平均値としている．

6.4 考察 4

概ね言語習得度が高くなるにつれて，全体の正解率，及び予測し易い文章の修復の度合いが高くなり，予測し難い文章では若干ではあるが低下している．おそらく言語習得度の高い被験者は単語の出現を予測しながら文章を聞いており，予測し易い文章において修復の度合に顕著な差が現れたと考えられる．一方，予測し難い文章では出現率の低い単語に対して誤った予測が行われた可能性があり，言語情報が逆効果に働いたとも考えられる．

第7章 まとめ

7.1 結論

本稿は、言語情報が音韻修復に与える影響を確認するため、ある程度長い文章について母国語・非母国語話者の修復がどう違うか、文章量・位置により修復出来る音がどう変化するかを調べた。第3章と第4章の実験結果より、母国語話者の場合、多大な言語情報を持っているので、無音・雑音を加えたいずれの場合も正解率が高くなっており、他言語話者より高くなっている。その中でも、母国語話者の雑音挿入時の長音や助詞の正解率の高さは、日本語と中国語の言語的背景の違いによる言語情報の利用の差であり、音韻修復において言語情報の影響の高さを物語っている。第5章の実験では、文章の情報量を操作して聴取を行った結果、特定の文節は、特に先行文節に大きく影響されることを確認し、同時に文節の接続数を増やすことで被験者は文章の単語や文法の認識が上昇し、聴取回数も減少していることから、音韻修復は文章の理解の上に大きく向上することが分かった。第6章で言語修得度の違いによる言語情報の利用について調べた聴取実験結果から、言語情報を利用する被験者ほど、予測し易い文章と予測が困難な文章の正解率に大きな開きがあり、言語情報を利用しない被験者ほど、特に言語情報を利用している形跡は見られなかった。これは言語情報の利用が主に文章の予測に対して働いており、音韻修復もこの予測と関連性があることを確認した。これらの結果から、言語情報が音韻修復に与える影響は大きく、文章の予測が主な認知の手がかりになっていたことが分かった。

7.2 今後の課題

音韻修復の了解度が向上するプロセスの解明は、主に次の3つが影響していると考えられる。・言語情報・調音結合・音声知覚の運動理論

本研究では言語情報の影響の存在及びその手がかりを確認したが、まだ解明に至ったとは言い難い。なぜなら、文章の予測が音韻修復の手がかりとして、どのような文章・どのような文法・どのような構文によって引き起こされる要因であるのか、またその影響の大きさがまだ分かっていない。このため、これらに対する優劣や音声認識への応用のために、定量的表現の必要性がある。それには、単語親密度と音韻修復率の対応関係を調査すること等が挙げられる。さらに、言語情報が欠損部分の検知限を広げるとはいえ、雑音置換直後の検知限延長は説明できるが、元の音声資料の時間軸上手前の検知限延長効果についての説明は、人間の音声知覚に対する決着がついていないことから説明が困難であ

る．これについて，特定位置における雑音置換についてさらなる調査を行うことで，認知メカニズムの解明に対しさらなる知見を提供することが可能であると考えられる．

また，調音結合と言語情報の関わり方及びその影響の違いや，舌の運動を考察し，それによる音声と音韻修復の可否についての相関を調査することで，音韻修復現象の解明から認知過程へ知見を増やすことも不可能ではなく，これらを調査することが今後の課題となる．

謝辞

本研究を遂行するにあたり，御多忙な中，全般的な御指導と惜しめない御援助を頂いた党健武教授に心から感謝致します。また，徳田功助教授には，積極的な御配慮，御意見を頂き深く感謝致します。中井満助手には，本研究の問題点への適切な御助言や高度な研究指導を承りました。ここに感謝の意を表します。そして，党・赤木研究室，さらに下平研究室の皆様には，日ごろから御討論や実験へ御協力，御卒業された後も常に激励を頂き，意義深い研究生生活を送ることが出来ました。ここに感謝の気持ちを表し結びと致します。

参考文献

- [1] R.M.Warren, Perceptual restoration of missing speech sounds, Science Vol. 168, pp392–293, 1970.
- [2] Minoru Tsuzaki and Hiroaki Kato, Apparent duration of speech segments interrupted by a noise burst, Acoust.Soc.Jpn, Vol. 15, No. 3, 1994.
- [3] Bhiksha Raj Ramakrishnan, Reconstruction of Incomplete Spectrograms for Robust Speech Department of Elctrical and Computer Engineering Carnegie Mellon University, 2000.
- [4] 柏野牧夫, 閉鎖区間の前後に分散する手がかりに基づく日本語語中閉鎖子音の知覚, 日本音響学会誌, Vol. 48, No. 2, pp76–86, 1992.
- [5] J.A.Bashford.Jr.,and R.M.Warren, Multiple,phonemic restorations follow the rules for auditory induc -tion, Perception, Perception. Psychophysics, Vol. 42, pp114–121, 1987.
- [6] Daniel P.W. Ellis, Using Knowledge to organize sound: The prediction-driven approach to computational auditory scene analysis and its application to speech/nonspeech mixtures, Speech Communication 27, pp281–298, 1999.
- [7] 柏野牧夫, 音韻修復, 日本音響学会誌, Vol. 61, No. 5, pp263–268, 2005.
- [8] 柏野牧夫, 音声知覚の運動理論をめぐって, 日本音響学会秋季講演論文集, pp243–246, 2004.
- [9] Minoru Tsuzaki,and Hiroaki Kato, Durational shrinkage by noise replacement in quasi-isochronous and hyper-isochronous contexts, Acoust.Sci and Tech, Vol. 26, No. 1, 2005.