

発話の聞き取り過程における相互的性質について

高橋 俊章

On Interactive Nature of Speech Perception

Toshiaki TAKAHASHI

(Received September 28, 2001)

要 旨

本論文では、発話の聞き取り (speech perception) に焦点を絞り、その処理過程を理論的に考察した。その結果、聞き取りには、ボトムアップとトップダウンの両側面が必要であることが明らかになった。また、ボトムアップ処理が効率的であればあるほど、トップダウン情報に依存する必要が減り、逆に、トップダウン情報が多ければ多いほど、ボトムアップ処理に依存する必要が減ることがわかった。すなわち、両者の関係がトレードオフの関係にあることが示された。ただし、一定レベル以上、ボトムアップ処理からの情報が不足すると、トップダウン処理の活用が困難となることから、効率的なボトムアップ処理の獲得が発話の聞き取りに不可欠であると考えられる。

1 ボトムアップ処理とトップダウン処理

聞き取りについては、ボトムアップ処理によって行われるという意見とトップダウン処理によって行われるという意見がある。ボトムアップとは、認知処理の方向が低次レベル (ボトム) から高次レベルに進むことであり、トップダウンは、高次レベル (トップ) から低次レベルに処理が進むことを意味している。

例えば、電光掲示板に流れていく文字を読む場合のように、/t/、/e/、/n/と順番に音声信号を聞き取って、/ten /という単語を知覚するのであれば、聞き取りはボトムアップ処理で行われていると考える。逆に、Seven plus three equals **n の最後の単語が文脈から ten だと分かるように、高次の情報を用いて単語の知覚が行われている場合、聞き取りはトップダウン処理で行われていると考える。

以下では、聞き取りがボトムアップとトップダウンのどちらの処理過程によって行われているのかを詳しく検討する。

2 ボトムアップ的に音声信号を処理する上での3つの困難点

これまでの研究によれば、聞き取りの処理過程がボトムアップであると主張する人たちにとって都合の悪い音声的事実が3つある。

1 番目は、音声信号は音素と一対一の対応関係にはないことである (lack of invariance condition/acoustic-phonetic invariance: Clark and Clark, 1977, p.176; Sawusch, 1986, p. 52)。印刷されたアルファベットの場合「a」はいつも「a」であり、文字と書記素を一文字ずつ対応させて読んでいくことが可能である。しかし、聞き取りでは、物理的には同じ音でも違う音素として知覚される場合や、同一の音素でも異なる物理的な音で表されることがある。例えば、1400 Hz を中心とするバースト音 (破裂音) が /i/ の前に置かれた場合には /pi/ と知覚されるのに、全く同じバースト音でも /a/ の前に置かれた時には /ka/、そして /u/ の前に置かれた場合には /pu/ と知覚される (Liberman, Delattre, & Cooper, 1952)。

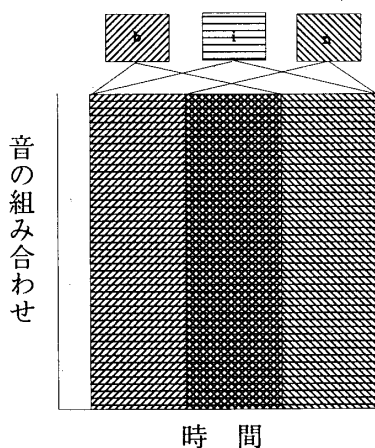


図1 同時調音 (coarticulation)

(Matlin 1983, p. 134から引用)

次は、1 番目に述べたこととも関係するが、音素は単独に1つ1つ発音されるわけではなく (coarticulation)、隣接する他の音素によって変化する (context conditioned variation) ため、物理的な音声信号の情報だけに基づいて音素に分けること (segmentation) が非常に困難であるということだ (Pisoni and Luce 1986, p. 4; Matlin 1983, pp.133-4が挙げられる)。例えば、bin という単語の場合、/b/ という発音が完了した後に、/i/ や /n/ が発音されていくのではない。実際には、b の発音は i と n の発音に重なって発音され、i や n も同様に他の2つの音素と重なって発音される (図1参照)。このように、1つ以上の音素が同時に伝達される (parallel transmission: Liberman, 1970) ため、1つの音素を処理した後に、次の音素を処理していくことは出来ない。このため、隣接する音素の影響を除外して、個々の音素を認知することはできないのである。

3 番目の問題は、たとえ話し手が音声信号を音素の単位で1つ1つ順番に処理していくことが可能であるとしても、このような処理を聞き手が行っているとは考え難いことである。なぜなら、聞き手が音声信号を処理している速度を考えると、音素単位ごとに音声信号の処理を進めているとは考えられないからである。つまり、音素の単位で処理を進め、最終的に単語の認知にいたるまでに必要とされる時間よりはるかに短い時間で、聞き手が単語を聞き取れることが明らかになっているからである (Liberman et al., 1967; Pisoni and Luce, 1986, p.19)。したがって、たとえ1 番目と2 番目の問題がなかったとしても、聞き手は、音素を左から順番にひとつひとつ処理

しているとは考えられない(lack of linearity conditions)。

3 ボトムアップの役割

前述した3つの問題点が事実であれば、聞き手が1つの音素を処理するために、その音素に隣接する他の音素や単語の綴字規則等のより高次の情報を利用しなければならないことになる。また、音素を単独で取り出すことが不可能であることや、時間的に効率が悪いことを考えると、聞き手が音素を順番にひとつひとつ聞き取っていくことによって、単語の知覚にいたるとは考えられないことになる。しかし、このことからすぐに聞き取りにおけるボトムアップの役割を否定することはできないであろう。

2節の1に従えば、「音素」は固有の音声特徴がないので、あとに続く音声信号の影響によって、どのような音として知覚されるかが異なることになる。しかし、英語の子音の音素にはすべて、それに固有の特徴がいくつか存在しており、例えば、/s/は前後にどんな母音が来ても、スー音(歯擦音)で区別出来るし、/z/と/s/の場合も/z/は/s/より音の高低(ピッチ)が低いことで区別が可能である。

また、2節の1や2は音素が音声知覚の処理単位ではない根拠になるかもしれないが、単語の知識等のより高次の情報が必要であることを示す直接の根拠にはならない。実際、Cole, Rudnick, Zue & Reddy (1980)が行った実験で、Victor Zueは事前にどんな語が含まれているか知らないでスペクトログラムだけを見て解読したが、その結果は音声学者が音声を聞いて書き取った大まかな音声表記(broad phonetic transcription)と80%~90%一致した。この実験結果は、音響的信号(acoustic signal)には十分な情報が含まれており、どのような単語が使用されているかといったトップダウンの情報がなくても音声信号の解読(ボトムアップ)だけでかなりの程度の聞き取りが可能であることを示している。

4 トップダウン： 文脈の役割

3で考察したように、単語の知識などがなくても、音声信号(ボトムアップ)の解読だけでかなり聞き取りは可能であると思われるが、ボトムアップの役割に限界があるのも事実である。

例えば、Black(1970, pp.74-96)が行った英語子音間の距離(interconsonantal distance)知覚に関する実験結果に基づき、竹蓋(1984, p.38)が指摘している。英語のネイティブスピーカーでも音が単独で取り出されてテストされた場合には、/m/と/n/や/ʒ(-ge)/と/dʒ(-dge)/、そして/f/や/θ(-th)/などは私たち日本人が/l/と/r/を聞いた場合と同じくらいに混同して聞いている。

また、統語や意味等のより高次の情報が聞き取りに関与していることを示す研究が他にもたくさんある。例えば、Garnes & Bond(1975, pp. 214-225)は自然な場面('informal settings')での会話から集められた聞き取りにおける誤り(約200)を分析して報告している。次にあげる

のはそのうち Clark & Clark (1977, p.214) によって紹介されている 4 例である。

| オリジナル | 聞き違い |
|-----------------------------|----------------------------|
| wrapping service | wrecking service |
| meet Mr. Anderson | meet Mr. Edison |
| I'm covered with chalk dust | I'm covered with chocolate |
| get some sealing tape | get some ceiling paint |

上記の例で「オリジナル」と「聞き違い」には音声的な類似点がある。Garnes & Bond (1975) (Garnes & Bond, 1980 も参照されたい。) が分析した誤りのほとんどすべての例において、「聞き違い」は「オリジナル」に意味的にも文法的にも適合している。また、chalk dust が chocolate だと誤って聞き取られたように、音声的な特徴の類似性や音声的な変化では説明できないほど、「オリジナル」と「聞き違い」が異なっている場合であっても文法的、意味的には適切である場合が多い。このことから、Clark & Clark (1977, p.214) も指摘しているように、聞き取りにおける誤りが音声的な特徴だけでなく、意味的な要因によっても左右されていることがわかる。もちろん、上で紹介した例はすべて「聞き違い」の例であるため、この種の誤りが通常の聞き取りプロセスを反映している証拠にはならない。しかし、実際の会話を聞き取る時に、私たちが文脈などの助けが必要であることを示すような事例はこれ以外にもたくさんある。

例えば、以下の Warren & Warren (1970, pp. 30-36) の実験は、音声信号の明瞭度が低く、個々の音素の識別が困難な場合には、統語的制約や意味的制約などを利用して聞き手が単語を聞き取っていることを示している。

彼らの実験で使われた 4 文は、最後の単語だけが異なり、文中の印 (*) は咳の音が入ってしまったために聞き取れなくなった部分を表している。

It was found that *eel was on the axle.

It was found that *eel was on the shoe.

It was found that *eel was on the orange.

It was found that *eel was on the table.

実験の結果、*eel は、上から順番に wheel, heel, peel, meal として知覚された。このことは、聞き手の受容した音声的情報が不完全な場合、文脈などからその不完全な情報を補完しようとすることを示していると言えるだろう。

このように、音声的情報が欠けても、聞き手は統語的、意味的制約などを利用して、不足した情報を補うことが出来ると考えられる。また、統語的制約や意味論的制約に限らず、音韻的制約

にも従うと考えられる。Ladefoged and Broadbent (1957) の実験では、被験者はヘッドホンをつけ、両耳からそれぞれ別の無意味語を聞いた。例えば、片方のヘッドホンからは "tass" を聞かされ、もう一方のヘッドホンからは "tak" を聞かされた。2つの語を両耳から同時に聞かされた被験者は、"task" か "tacks" を聞いたと報告した。この時、"tass" が "tak" よりほんのわずかだけ早くヘッドホンから流されれば、被験者は "task" を聞いたと報告し、逆に、"tak" が "tass" より早く流された場合には "tacks" を聞いたと報告した。しかしながら、"lanket" と "banket" が両耳から同時に流れた場合には、常に "blanket" と聞いたと被験者は報告した。しかも、"banket" が "lanket" よりも少しだけ早く提示された場合でも、"lanket" が "banket" よりもわずかだけ早く提示された場合でも、被験者は "blanket" と聞いたと報告し、"lbanket" と聞いたと報告した者はいなかった。これは、Hatch (1893, p.17) が指摘しているように、英語の音韻規則で許される音の組み合わせから成る単語しか聞き手の耳に入らないからだと考えられる。

5 余剰性：ボトムアップとトップダウンとの間にあるトレードオフの関係

これまで、音声信号の一部が聞こえなかったり、あるいは音声信号そのものの明瞭度が低い場合、聞き手は、意味論的制約や統語論的制約や音韻規則などを用いて、欠けている音声信号を補っていることを述べてきた。このことは、言語の持つ余剰性 (redundancy) とも関係している。余剰性とは、伝えようとする信号やメッセージの中に、伝達する時に理論的に必要とされる以上の情報が含まれていることである。本論文では詳述しないが、余剰性と内容理解も密接な関係があり、余剰性を増せば、内容理解も増すという関係にある。例えば、Chaudron (1983) や Pica, Young, and Doughty (1987) は内容語の繰り返しにより聞き手の内容理解が増したと報告している。ただし余剰性を与えるものであっても下位の学生にはかえって負担となる場合があることに注意しなければならない。例えば、Chiang and Dunkel (1992) は英語を外国語として学ぶ中国の学生を対象にした実験において、高・中級の聞き手は構成要素の繰り返し (unmodified: The food of the Pennsylvania Dutch Country is very hearty and delicious / modified: The food of the Pennsylvania Dutch Country is very hearty and delicious. Hearty and delicious food is nourishing and tasty) の使用によって利益を得たが、低・中級の聞き手はそうではなかったと報告している。また、Derwing (1989) も、内容語の繰り返しは効果があったが、中心となるメッセージの内容を理解をするのに、不必要な詳しい背景知識を与えることは却って聞き手に混乱を与えたり、余分な負荷を与えることになるのではないかと論じている。

例えば、Tom pl*y** t*nn*s yest**day. と書いた文を Tom played tennis yesterday. に復元することは容易である。つまり、単語の長さや単語全体としての形に関する情報、現在か過去かといった時制に関する情報、t*nn*s が名詞であろうという情報などが利用できるために、「*」の部分が欠けてもメッセージが復元できるのである (cf. Spolsky, 1973,

p. 167)。言い換えれば、音声的信号の明瞭度が低くても言語の理解が可能であるのは、意味論的制約や統語論的制約などによって余剰性が高くなっているためだと言えるであろう。ノイズのある電車の中や雑音の多い電話で会話が可能であるのも、文脈や文法的な制約によって語彙や話題の内容の幅が限られていることや、話者が意識的に言葉をより明瞭に発音するからであろう。

実際、Miller, Heise & Lichten (1951) は、ノイズがある場合、単独で提示された単語よりセンテンス文脈の中で提示された語の方がはるかに聞き取り易かったと報告している。

逆に、通常の会話では、発話は文脈の中でなされるため、単語は明瞭に発音されない。そのため、単語を文脈から切り離すと、聞き取りにくくなるのだと考えられる。Fowler & Housum (1987) がラジオ番組のモノログを分析したところ、2度目に登場した単語 ("old" words) は1度目に登場した単語 ("new" words) に比べ短く発音され、明瞭さが乏しくなった。また、そのため、単独で提示された場合には、聞き取りの誤りが多くなることを報告している。

Pollack & Pickett (1964, pp. 79-84) が、テープに録音された会話から単語だけを抜きだし、英語の母国語話者に聞かせて、同定することを求めたところ、約半数しか正しく認識することが出来なかったと報告しているが、これも同じ理由で説明出来るであろう。

また、Lieberman (1963) は、"The word that you will hear is ____ のように文脈情報が乏しい文から取り出された下線部の単語 (nine) は、"A stitch in time saves ____" のように文脈情報が豊富な文から取り出された同じ単語 (nine) に比べた場合、文脈が無い場合でも、ノイズがある場合でも、より明瞭で聞き取りやすかったと報告している (Hunnicut, 1985 も参照のこと)。この結果は、文脈のある文が、文脈のない文に較べて音声信号が明瞭でないので、文脈がない場合にはわかり難くなることを示している。

このように、通常のスピーチでは文脈から聞き手に解りそうなことは明瞭に話されない場合が多く、文脈なしで音声情報だけに頼らなければならなくなったときは逆に聞き取りが困難になるわけである。

6 音声情報の不足は、トップダウン情報で補えるか

英語を外国語として学ぶ日本人の場合、音声情報を正確に聞き取ることが不得意なので音声情報に依存できないため、文脈や統語的な制約などを利用して単語を聞き取ることが必要である。このようなことは可能であろうか。

実際に、これに関し、Marslen-Wilson & Welsh (1978) が行った実験結果が、興味深いデータを提供している。被験者はヘッドホンから誤った発音 (e.g., *cikarette*, *mizfortune*) を所々に含むスピーチを聞き、出来るだけ時間を置かず、聞いた通りに声に出して反復するように要求された (これを *shadowing task* という)。実験の結果、文脈の流れから使われる単語がかなり予測出来る場合 (e.g., *Still, he wanted to smoke a cikarette.*) の方が、その様な予測が出来ない場合 (e.g., *It was his mizfortune that they were stationary.*) に比べ「再生 (誤っ

た発音をそのまま繰り返すこと)] (e.g., *cikarette, mizfortune*) より「修復 (誤った発音を修正して繰り返すこと)] (e.g., *cigarette, misfortune*) がより多く行われた。つまり、予測可能性が高い場合には、被験者は、誤った発音を修正して繰り返したわけである。

この実験結果について、D. W. Carroll(1986, p. 131)は、聞き取りにおけるトップダウンの役割だけではなく、ボトムアップの変数も重要な働きをしていることを示すものであると述べている。つまり、意図された単語(e.g., *tragedy*) と実際に提示された単語の音素とが、1つの識別的特徴で食い違うとき (e.g., *trachedy*) には、「修復」がかなりの率 (約74%) で発生するのに対し、3つの識別的特徴で食い違う場合 (e.g., *travedy*) には、はるかに少ない率 (約24%) でしか「修復」は起こらなかった。言い換えれば、実際に聞いた音素と聞き手が予測していた音素が似ている時には誤った発音を訂正して繰り返すのに対し、もし予想した音素が実際に聞いた音素と音声的に十分似通っていると思えなかったときには、たとえ文脈には合っている、「修復」が起こりにくくなるのである [(注) 音素は文脈によって推測されているのであり、物理的に聞き取っているわけではない。逆にありもしない誤った音を聞き取ったと勝手に想像してしまう可能性さえある。Samuel (1986) が音素の「修復」と文脈に関する類似の報告をしている]。従って、ボトムアップが聞き取りにおいて果たしている役割はけっして否定することは出来ない。また、ノイズなど物理的・外的な要因や、音声識別能力の不足などの内的要因により、音声的情報などのボトムアップがある程度以上不足した場合、意味的情報などのトップダウンが利用出来なくなることも示していると考えられる (cf. Clarke, 1978)。

逆に、Hayashi (1991) の実験では、アメリカの英語母国語話者の場合、ノイズが減少しボトムアップからの情報が増加すればするほど、絵によって得られる文脈 (トップダウン) に依存する度合いが減少した。このことから、ボトムアップだけで十分な情報が得られる場合には、トップダウンへの依存が減少すると考えられる。これは、トップダウンより、ボトムアップ処理の方が効率的であるため、ボトムアップだけで十分な情報が得られる場合にはトップダウンに依存しないで処理が進められるためだと考えられる。同様に、Wolff (1987) は German ESL learners (ages 12-15) を対象にした実験において、テキストが困難になればなるほど、イラストレーション (トップダウン) を用いるが、テキストが簡単な場合には、ボトムアップが阻害されていないので、イラストレーションを無視したと報告している (p. 316)。また、Chiang and Dunkel (1992) は、英語を外国語として学ぶ中国の学生を対象にした実験で、(passage-dependent な質問項目においてよりも) passage-independent な質問において、背景的知識の効果 (トップダウンストラテジーの使用) が観察され、しかも、その効果は unfamiliar な話題より familiar な話題において顕著という結果を得た。

さらに、Henrichsen(1984)の研究において、音変化 (=Sandhi) がある時 (eg. Give them a few days and they will be back. [gIvəm:sandhi あり]) は

英語母語話者 >有意差あり> 上位 ESL 学習者 >有意差なし> 下位 ESL 学習者

であるのに対し、音変化がない時 (eg. Give them a few days and they will be back. [givð ε m:sandhi なし]) は

英語母語話者 >有意差なし> 上位 ESL 学習者 >有意差あり> 下位 ESL 学習者

という結果が出た。このように、音変化 (sandhi) がなく、音声信号の解読に支障がない場合には、英語母語話者と差がないような上位 ESL 学習者さえ、音変化があることによってボトムアップ情報が不十分になると、英語母語話者と聞き取りの能力で差が出てしまうことがわかる。

以上のことから、ボトムアップからの情報がある程度以上不足する場合には、トップダウンの情報が利用出来ないこと、また、ボトムアップとトップダウンからの情報に対する依存度がトレードオフの関係にあるため、どちらか一方からの情報が不足すると、他の一方に対する依存度が増し、逆にどちらか一方からの情報だけで十分な場合には、他の一方への依存度が減る関係にあることがわかる。

7 左から右への処理は可能か：オンライン処理におけるトップダウンの役割

2節では、3番目の問題点として、「音素」という単位で左から右に1つ1つ順番に処理して単語を認知することは困難であると述べた。しかし、それにも関わらず、実際には英語の母国語話者はスピーチを聞いた順に直線的にかなりの早さで処理していくことができるようである。

例えば、Shadowing Task (Marslen-Wilson & Welsh, 1978 参照) を用いた Marslen-Wilson の研究 (1973, 1975, 1976) では、連続したスピーチを聞き、それを約 250 ミリ秒ずつ遅れるだけでつぎつぎに追唱することが出来る被験者 ('fast shadowers') のことが報告されている。この被験者達はスピーチを単に早く繰り返しただけでない。彼らはスピーチを追唱するとき、その中に誤った発音の単語が含まれていた場合には、(その単語の発音が誤っていたことにさえ気付かず) 無意識に訂正した。また、彼らはスピーチを繰り返す時、誤った発音の単語をまったく別の単語に置き換えてしまったが、それらの単語は文法的にも正しく、文脈上から考えてもおかしくない場合がほとんどであった。したがって、音素を処理単位としてスピーチを処理していくことは不可能かもしれないが、スピーチを左から右に直線的に処理していくことは可能だということになる。しかも、このように非常に高速にスピーチを処理しながら、意味や文脈などの情報を十分活用することが出来ることをも示している。

このように、スピーチをオンラインで (即時的に) 処理していくことが可能なのは、文脈 (背景知識の利用を含む) などのトップダウンの役割が大きいからだと考えられる。例えば、Marslen-Wilson and Tyler (1980) の実験はオンラインでスピーチを処理していく上でのトッ

プダウンの役割を示している。この実験では、被験者は以下の(1)~(3)に示したような3タイプの文を聞き、ある単語(lead)が聞こえたらすぐボタンを押すように指示された。

- (1) 通常の散文(normal prose) (273ミリ秒)

eg., *Some thieves stole most of the lead off the roof*

- (2) 文法的ではあるが意味論的に不自然な散文(syntactic prose) (331ミリ秒)

eg., *No buns puzzle some in the lead off the text*

- (3) ばらばらの散文(random word-order) (358ミリ秒)

eg., *some the no puzzle buns in lead the text off*

実験の結果、ばらばらの散文の条件より文法的ではあるが意味的に不自然な散文の条件の方が反応時間が短かった。さらに、この2つの条件より通常の散文の条件の方がはるかに反応時間は短かった。反応ボタンを押すこと自体に50~75ミリ秒かかることを考慮に入れると、実質の反応時間はもっと短かったことになる(200ミリ秒弱)。しかし、語頭の音を聞いてから200ミリ秒間の音声情報だけでは、単語の候補を1つに絞るだけの情報を含んでおらず、まだ平均29の単語の候補が残されているのである。言い換えれば、意味的、統語的制約が利用できる場合には、目標とする単語を最後まで聞き終わらないうちに、それらの制約を用いて単語の候補を1つに絞ることが出来ることを示している [(注) 語彙認知に関しては Johnston & McClelland (1980)、語認知の処理単位については Rumelhart & McClelland (1981) (phoneme) や Klatt (1979) (speech wave) 参照]。また、上記の実験で、(1)と(2)の場合にはターゲットワードの位置が後になればなるほど反応時間は短くなったのに、(3)の場合にはならなかった。このことは意味的、統語的表象が一語一語累積的に形成されていくので、後ろの位置の単語の認知が促進されることを示していると考えられる。

8. リスニングのモデル

伝統的なリスニングのモデルでは、リスニングのプロセスは、直線的で逐次的なボトムアップな処理で行われていると説明されてきた。つまり、音素が順番に処理されていくことだけで単語が認知されるとされ、単語認知は文脈などからの影響を受けないと主張されてきた(伝統的なモジュラーモデルの批判に関しては Elman & McClelland, 1988 の実験を参照)。しかし、すでに考察してきたように、単語認知は音声の情報だけでなく文脈の影響を受けるのである。従って、リスニングの単語認知に関しても、リーディングの場合と同様、インタラクティブモデルが現時点では最も魅力的なモデルとなっている。リスニングにおける語彙認知に関するインタラクティブモデルには、ロゴジェンモデル(Morton 1982)の他に、コーホート理論(Marslen-Wilson 1975, 1980, 1987; Marslen-Wilson & Welsh, 1978; Marslen-Wilson & Tyler, 1975,

1980)、トレース インラクティブ アクティベーション モデル (McClelland & Elman, 1986) が提唱されている。しかし、どのモデルも基本的な概念は共通しており、文脈の影響が単語認知に及ぼす影響を説明出来ることや、音声信号の特徴を音素のような単位ごとに検出する検出器があり、それらの検出器の一つ一つから送られてくる信号の強さによって候補となる単語だけが次第に活性化され、単語の認知にいたるなどの特徴を持っている。

ただ、コーホート理論だけは他のモデルと異なり、音声信号を語頭の音から順番に処理していくことを想定しているため、モデルとしての妥当性を欠いているかも知れない。コーホート理論によると、最初の時点では同じ音素で始まるすべての語（仮に、sin, sit, sincere, simple の4つだけとする）が活性化されるが、音声信号が順番に1つ1つ処理され、/sins.../の時点まで処理が進んだ段階では他に候補がなくなるので、sincere に特定されると説明されている。しかし、Pisoni & Luce (1986, 1987) が批判しているように、コーホート理論では、語頭の音を聞き間違えてもその単語を認知することが出来ることが説明出来ない。

さらに、Salasoo & Pisoni (1985) の実験により、単語の初めの音響-音素的情報の方が単語の終わりに含まれる情報に比べ単語の聞き取りには重要ではあるが、語頭の音声的情報がなくても聞き取りが可能であることを示している。この結果は語頭的情報を必須とするコーホート理論に反するものである。

要約すれば、リスニングにおける単語認知のモデルとしては、コーホート モデルを除くインタラクティブ モデルが現時点では適切だということになりそうである。

9. 同時調音 (Coarticulation) の問題

リーディングの場合にはスペースによって語と語の境界は明確に示されている。リスニングの場合にも単語と単語の間にはポーズがあるように知覚されるかも知れないが、多くの場合、単語と単語とは音響的に連続しており、境界は明確ではない。Fraser (1992, pp. 29-30) は、単語の境界の位置は異なるが発音は同じであるような英語の刺激 ('white shoes' vs. 'why choose' や 'grey day' vs 'grade A' など。cf. Lehiste, 1960) を用いた諸研究において、被験者は65%~85%しか単語の境界を同定することが出来なかったと報告している。つまり、ボトムアップだけの情報だけでは完全に単語と単語の境界を同定することは出来ず、単語の使用頻度や統語など高次の情報が不可欠となる。

10. まとめ

文脈等に過度に依存して聞き取る場合、極端になると直感や推測に近くなる危険性が生じる。従って、正確に聞き取っていくためにはある程度のボトムアップの情報が不可欠である。また、音声的情報などボトムアップからの情報が一定限度以上に不足すると、文脈などトップダウンの情報が利用出来なくなる。しかし、ある一定のレベル以上に単語認知など低次レベルの処理能力

が高い場合には、背景知識や文脈（意味論的、統語的制約）などが豊富であるほど聞き取れる確率は高くなると考えられる。しかし、ボトムアップからの情報が非常に豊富である場合には、文脈などのトップダウンからの情報に依存する必要は減少する。

参考文献

- Albert, M. L., & Bear, D.(1974). Time to understand: A case study of word deafness with reference to the role of time in auditory comprehension. *Brain*, 97, 373-384.
- Allan, M. (1985). *Teaching English with video*. London: Longman.
- Bever, T. G.(1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.
- Black, John W. (1970). Interconsonantal differences. In Arthur J. Bronstein, Claude L. Shaver & Cj Stevens(Eds.), *Essays in honor of Claude M. Wise* (pp. 74-96). Columbia, Missouri: Artcraft Press.
- Caplan, D. (1972). Clause boundaries and recognition latencies for words in sentences. *Perception and Psychophysics*, 12, 73-76.
- Carroll, D. W.(1986). *Psychology of Language*. Brooks/Cole Publishing Company.
- Carroll, J. M., Tanenhaus, M. K., & Bever, T. G.(1978). The perception of relations : The interaction of structural, functional, and contextual factors in the segmentation of sentences. In W. J. M. Levelt & G. B. Floresd'Arcais (Eds.), *Studies in the perception of language*. New York: Wiley.
- Chaudron, Craig (1983). Simplification of input: Topic reinstatements and their effects on L2 learners' recognition and recall. *TESOL Quarterly*, 17, 437-458.
- Chiang, C. S., & Dunkel, P. (1992) The effect of speech modification, prior knowledge, and listening proficiency on EFL lecture learning. *TESOL Quarterly*, 26, 345-374.
- Clark, H. H., & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Clarke, M. A.(1978). Reading in Spanish and English: Evidence from adult ESL students. *Language Learning*, 29, 1, 121-150.
- Cole, R. A., Rudnicky, A. I., Zue, V. W., & Reddy, D. R.(1980). Speech as patterns on paper. In R. A. Cole(Ed.), *Perception and production of fluent speech*(pp. 3-50). Hillsdale, NJ: Erlbaum.
- Cotton, J. (1935). Normal "visual hearing". *Science*, 82, 592-593.
- Cutler, A. (1982a). Prosody and sentence perception in English. In J. Mehler, E. C. T. Walker, & M. Garrett(Eds.), *Perspectives on mental representation*. Hillsdale.

- N.J.: Lawrence Erlbaum.
- Derwing, T. M. (1989). Information type and its relation to nonnative speaker comprehension. *Language Learning*, 39, 157-172.
- Dodd, B. (1977). The role of vision in the perception of speech. *Perception*, 6, 31-40.
- Eastman, J. K.(1991). Learning to listen and comprehend: The beginning stages. *System*, 19, 3, 179-187.
- Elman, J. L., & McClelland, J. L.(1985). An architecture for parallel processing in speech recognition: The TRACE Model. *Bibliotheca Phonetica*, 12, 6-35.
- Elman, J. L., & McClelland, J. L.(1988). Cognitive penetration of the mechanisms of perception: Compensation for Coarticulation of lexically restored phonemes. *Journal of Memory and Language*, 27, 143-165.
- Ellis, A. W., & Beattie, G. (1986). *The psychology of language and communication*. The Guilford Press.
- Fodor, J. A., Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 4, 414-420.
- Fowler, C. A., & Housum, J.(1987) Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26, 489-504.
- Fraser, H.(1992). *The subject of speech perception*. London: the Macmillan Press, Ltd.
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291-325.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178-210.
- Garnes, S., & Bond, Z. S.(1975). Slip of the ear: Errors in perception of casual speech. *Papers from eleventh regional meeting*(pp. 214-225). Chicago Linguistic Society.
- Garnes, S., & Bond, Z. S.(1980). Slip of the ear: A snip of the ear? A slip of the year? In V. A. Fromkin (Ed.) *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*(pp. 231-239), London: Academic Press.
- Garrod, S.(1986). Language comprehension in context: A psychological Perspective. *Applied Linguistics*, 7, 3, 226-238.
- Garrett, M. F., Bever, T. G., & Fodor, J. A.(1966). The active use of grammar in speech perception. *Perception and Psychophysics*, 1, 30-32.

- Griffiths, R.(1990). Facilitating listening comprehension through rate-control. *RELC Journal*, 21, 55-65.
- Harris, M., & Coltheart, M. (1986). *Language processing in children and adults*. Routledge & Kegan Paul.
- Hatch, E. M.(1983). *Psycholinguistics: A second language perspective*. Newbury House Publishers, Inc.
- Hayashi, T.(1991). Interactive processing of words in connected speech in L1 and L2. *IRAL*, 29, 2.
- Henrichsen, L. E. (1984). Sandhi-variation: A filter of input for learners of ESL. *Language Learning*, 34, 103-126.
- Holmes, V. M., & Forster, K. I. (1970). Detection of extraneous signals during sentence recognition. *Perception and Psychophysics*, 7, 297-301.
- Hunnicut, S.(1985). Intelligibility versus redundancy -- Conditions of dependency. *Language and Speech*, 28, 45-56.
- Jarvella, R. J. (1970). Effects of syntax on running memory span for connected discourse. *Psychonomic Science*, 19, 235-236.
- Jarvella, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 10, 409-416.
- Johnston, J. C., & McClelland, J. L.(1980). Experimental tests of a hierarchical model of word identification. *Journal of Verbal Learning and Verbal Behavior*, 19, 503-524.
- Kellerman, S.(1990). Lip service: The contribution of the visual modality to speech perception and its relevance to the teaching and testing of foreign language listening comprehension. *Applied Linguistics*, 11, 3, 272-280.
- Klatt, D. H.(1979). Speech perception: a model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7, 279-312.
- Kohno, M.(1990). The role of rhythm in the first and second language acquisition. *Proceedings of International Conference on Spoken Language Processing*, 1, 1-4.
- Ladefoged, P., & Broadbent, D. E.(1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, 29, 98-104.
- Lehiste, L.(1960). An acoustic-phonetic study of internal open juncture. *Phonetica*, Supplement to vol. V(54 pages).
- Lieberman, P.(1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172-187.

- Lieberman, A. M.(1970). The grammars of language and speech. *Cognitive Psychology*, 1, 301-323.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M.(1967). Perception of the speech code. *Psychological Review*, 74, 431-461.
- Lieberman, A. M., Delattre, P. C., & Cooper, F. S.(1952). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs*, 68, 1-13.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.
- Marslen-Wilson, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, 244, 522-523.
- Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science*, 189, 226-268.
- Marslen-Wilson, W. D. (1976). Linguistic descriptions and psychological assumptions in the study of sentence perception. In R. J. Wales and E. C. T. Walker (Eds.), *New approaches to language mechanisms*. Dordrecht: D. Reidel.
- Marslen-Wilson, W. D. (1980). Speech understanding as a psychological process. In J. C. Simon(Ed.), *Spoken language generation and understanding*(pp. 39-67). Dordrecht, Holland: Reidel.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition*, 25, 71-102.
- Marslen-Wilson, W. D., & Tyler, L. K.(1975). Processing structure of sentence perception. *Nature*, 257, 784-785.
- Marslen-Wilson, W. D., & Tyler, L. K.(1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.
- Marslen-Wilson, W. D., & Welsh, A.(1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29-63.
- Matlin, M. W.(1983). *Cognition*. New York: Holt, Rinehart and Winston.
- McClelland, J. L., & Elman, J. L.(1986). The TRACE model of speech perception. *Cognitive Perception*, 18, 1-86.
- McClelland, J. L., John, M. S., & Taraban, R.(1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes*, 4, 3/4 ,SI 287-335.
- McNeill, D.(1987). *Psycholinguistics: A new approach*. New York: Harper and Row.

- Miller, G. A.(1956). The magical number seven plus or minus two: Some limitations on our capacity for processing information. *Psychological Review*, 63, 2, 81-97.
- Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41, 329-335.
- Morton, J.(1982). Disintegrating the lexicon: An information processing approach. In J. Mehler, E. Walker, & M. Garrett (Eds.), *Perspectives on mental representation: Experimental and theoretical studies of cognitive processes and capacities* (pp. 89-109). Hillsdale, NJ: Erlbaum.
- Mueller, G. A.(1980). Visual contextual cues and listening comprehension: An experiment. *The Modern Language Journal*, 64, 3, Autumn, 335-340.
- Omaggio, A. C.(1979). Pictures and second language comprehension. *FLA*, 12, 2.
- O'Malley, J. M., Chamot, A. U., & Küpper, L.(1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics*, 10, 4, 418-437.
- Pollack, I., & Pickett, J. M.(1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior*, 3, 79-84.
- Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly*, 21, 737-758.
- Pimsleur, P., Hancock, C., Furey, P.(1977). Speech rate and listening comprehension. In M. K. Burt, H. C. Dulay, & M. Finocchiaro(Eds.), *Viewpoints on English as a second language* (pp. 27-34). New York: Regents.
- Pisoni, D. V., & Luce, P. A.(1986). Speech perception: Research, theory, and the principal issues. In E. C. Schwab & H. C. Nusbaum(Eds.), *Pattern recognition by humans and machines: Volume 1, speech perception*(pp. 1-50). New York: Academic Press.
- Pisoni, D. V., & Luce, P. A.(1987). Acoustic-phonetic representations in word recognition. *Cognition*, 25, 21-52.
- Reich, S. S. (1980). Significance of pauses for speech perception. *Journal of Psycholinguistic Research*, 9, 379-389.
- Rivers, W. M.(1986). Comprehension and production in interactive language teaching. *The Modern Language Journal*, 70, i, 1-7.
- Rivers, W. M., & Temperley, M. S.(1978). *A practical guide to the teaching of English as a second or foreign language*. Oxford University Press.

- Rubin, Joan (1994). A review of second language listening comprehension research, *MLJ*, 199-221
- Rumelhart, D. E., & McClelland, J. L. (1981). Interactive processing through spreading activation. In A. M. Lesgold & C. A. Perfetti(Eds.), *Interactive processes in reading*. Hillsdale, N. J.; Lawrence Erlbaum Associates.
- Sachs, J. D.(1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, 2, 437-442.
- Salasoo, A., & Pisoni, D. B.(1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory and Language*,24, 210-231.
- Samuel, A. G.(1986). The role of the lexicon in speech perception. In E. C. Schwab & H. C. Nusbaum(Eds.), *Pattern recognition by humans and machines:Volume 1, speech perception*(pp. 89-111). New York: Academic Press.
- Sawusch, J. R.(1986). Auditory and phonetic coding of speech. In E. C. Schwab & H. C. Nusbaum(Eds.), *Pattern recognition by humans and machines: Volume 1, speech perception*(pp.51-88). New York: Academic Press.
- Sheerin, S.(1987). Listening comprehension: teaching or testing? *ELT Journal*, 41/2, 126-131.
- Shiffrin, R. M., & Schneider, W.(1977). Controlled and automatic human information processing: II Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84, 127-190.
- Sonnenschein, S.(1982). The effect of redundant communications on listeners: When more is less. *Child Development*, 53, 717-729.
- Sonnenschein, S., & Whitehurst, G. J. (1982). The effects of redundant communications on the behavior of listeners: Does a picture need a thousand words? *Journal of Psycholinguistic Research*, 11, 2, 115-125.
- Spolsky, B.(1973). What does it mean to know a language? Or, how do you get someone to perform his competence? In J. W. Oller & J. C. Richards(Eds.), *Focus on the learner: Pragmatic perspectives for the language teacher*(pp. 164-176). Newbury House Publishers, Inc.
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26, 212-215.
- Tallal, P., & Piercy, M.(1975). Developmental aphasia: the perception of brief vowels and extended stop consonants. *Neuropsychologia*, 13, 69-74.
- Taraban, R., & McClelland, J. L.(1990). Parsing and comprehension: A multiple-

- constraint view. In D. A. Balota, G. B. Flores d'Arcais, & Rayner, K.(Eds.), *Comprehension processes in Reading*(pp.231-263). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Tyler, L. K.(1992). *Spoken language comprehension: An experimental approach to disordered and normal processing*. London: The MIT Press.
- Tyler, L. K., & Warren, P.(1987). Local and global structure in spoken language comprehension. *Journal of Memory and Language*, 26, 638-657.
- Waltz, D. L., & Pollack, J. B. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, 9, 51-74.
- Wanner, E., & Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. Miller(Eds.), *Linguistic theory and psychological reality*. Cambridge, MA: M.I.T. Press.
- Warren, R. M., & Warren, R. P.(1970). Auditory illusions and confusions. *Scientific American*, 223, 30-36.
- Wolff, Dieter. (1987). Some assumptions about second language text comprehension. *SSLA*, 9, 307-326.
- 小池生夫. (1993). 『英語のヒアリングとその指導』大修館.
- 河野守夫・沢村文雄. (1985). 『Listening & Speaking: 新しい考え方』山口書店.
- 河野守夫(1990). 「リズム知覚のメカニズムと Listening Comprehension」『文部科学研究費平成元年度研究成果報告書』.
- 鈴木 寿一. (1990). 「聴解力養成におけるポーズの有効性の再検討」第16回全国英語教育学会静岡研究大会口頭発表資料.
- 竹蓋 幸生. (1981). 「日本人大学生の米語音聴取にみる”Acquired Similarity”と”Acquired Distinctiveness” 子音間距離知覚実験による観察」『Language Laboratory』, 18号, pp. 11-28.
- 竹蓋 幸生. (1984). 『ヒアリングの行動科学 実践的指導と評価への道標』. 研究社出版.
- 竹蓋 幸生. (1989). 『ヒアリングの指導システム』. 研究社出版.