

マルチモーダルな情報の処理過程 —音韻知覚におよぼす視覚情報の影響—

近藤 公久

e_mail: kondo@av-hp.brl.ntt.jp

NTT基礎研究所

243-01 神奈川県厚木市森の里若宮3-1

本研究は、視覚情報が音声の知覚に及ぼす影響を心理実験によってとらえることにより、人間の知覚過程におけるマルチモーダルな情報の処理過程の基本を明らかにすること目的としている。音声と同時に提示される文字情報や唇の画像情報がその音声の知覚に影響を及ぼすことが知られている。文字情報はそれ自体音韻情報を持っているが、唇の画像情報は単独では音韻を確定するのに十分な情報を有しない。しかし、唇情報は音韻の弁別素性となりうる調音位置の情報を持っている。さらに、唇の動きは音声の時間的情報も持っている。本稿では、これらの視覚情報がどのように音韻の知覚に影響を及ぼしているのかを、視覚と聴覚の両処理の時間的關係を考慮に入れて検討する。

Multi-modal information processing: Effects of visual information on phoneme perception.

Tadahisa Kondo

NTT Basic Research Laboratories

3-1 Wakamiya, Morinosato, Atsugi, Kanagawa, 243-01 Japan

In this study, the nature of intermodal interaction in speech perception is investigated by examining the effects of simultaneously presented visual information on speech perception. It is known that speech perception is influenced by simultaneously presented visual information-- letters or visual image of moving lip. A letter contains phonological information, while the image of the moving lip is not enough to determine the phoneme. On the other hand, the image of the lip could provide the information as to the place of articulation and the timing of articulation. In this paper, how such information from visual modality-- letters, and the visual image of lip movement-- affect phoneme perception is investigated. In particular, the time course of the multi-modal information processing is discussed.

1. はじめに

McGurkとMacDonald¹⁾がいわゆるマガーク効果(McGurk Effect)を報告して以来、音韻知覚に対し唇の画像情報が影響を及ぼしていることが広く認識されるようになってきている。また、文字情報が音声知覚に影響を及ぼしていることも実験的に確かめられている^{2),3),4)}。本稿では、筆者自身の行った実験結果を含め、これまでに知られている視覚情報が音声知覚に影響を及ぼす現象を示しながら、視覚と聴覚の情報が統合されて知覚が生じる過程について検討する。特に、視聴覚情報の統合が起こる際の時間的關係に注目する。最後に、唇の情報や文字情報が音声知覚に影響を及ぼす過程を総合的に説明することを試みる。

マルチモーダルな知覚過程を明らかにするという事は、それぞれの感覚で入力される情報がどのように処理されて、どのような形式で脳内で表現され、どのような脳内の知識と照合されるのか、そして、互いの処理過程のどの段階でどのように情報が統合されるのかを明らかにすることである。ここでは、音韻を知覚する過程を考えてみる(図1)。発声者が発した音声は音波として聴者の耳に到達する。この音波には音声の知覚に必要なさまざまな特徴情報が含まれている。聴者が音声の中の音韻を知覚する過程を模式的に示すと、まず入力された音声の特徴情報を抽出し、次にこれらの特徴情報から音韻の決定に関係するものを選択・統合し、最終的に音韻を決定しなければならない。一方、

視覚情報が音韻の知覚に影響を及ぼすということは、音声情報処理側の音韻の決定に影響を及ぼすということであるから、視覚から得られる情報は音韻の決定以前に音韻の特徴情報と統合可能な形式で表現されていなければならない。たとえば、マガーク効果であれば、唇の画像情報から得られる音韻に関する情報が抽出され、音声情報と統合され得る表現形式に変換されなければならない。また、文字情報であれば、文字の持つ音韻に関する情報が抽出されて音声情報と統合され得る表現形式に変換されなければならない。それでは、この音声情報と統合され得る表現形式はどんなもので、音声情報のどのような特徴と統合されるのであろうか。また、それは音声情報処理過程のどの段階で起こるのであろうか。

2. マガーク効果

2.1 口の開閉情報と音韻知覚

マガーク効果は単音節あるいは2連音節を聴取する際に、音声と同期して音節を発声している顔の画像を提示することによって音声提示された音節の知覚が変化する効果である。例えば、/ba/または/pa/という音声に対し、/ga/または/ka/を発声している画像を同時に提示すると、/da/または/ta/に聞こえる率が高くなる。この効果はどのようなことを意味するであろうか。音素/b/と/p/は両唇、/d/と/t/は歯茎と舌の先、/g/と/k/は軟口蓋と舌の奥を一旦つけてから再度開くことによって発声される。そして、この声道を狭める場所は調音位置と呼ばれ、これらの音素はこの調音位置の違いに起因した異なる音素特徴を持っている。しかし、/b/, /d/, /g/は有声破裂子音、/p/, /t/, /k/は無声破裂子音という同じ調音様式に分類され、この意味ではそれぞれの3音素のグループ内では似通った音素特徴を持っている。一方、顔画像が持っている音声に関連する情報は基本的には口の形状と動きであり、これから得られる音韻的な情報は調音位置の情報である。すなわち、唇の開閉情報はそこから発声される音韻の唇音/非唇音の弁別を、

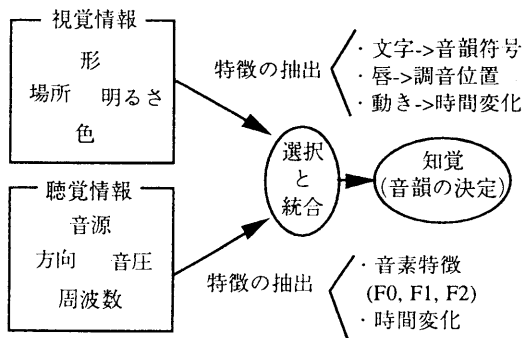


図1 視覚情報と聴覚情報の統合と知覚

舌やあごの動きの情報は、歯茎音と軟口蓋音との弁別を可能とする情報を含んでいる。マガーク効果は、この調音位置情報と音声の特徴情報が融合されて起こると説明できる。ここでは、まず、唇画像から最も明確にとらえられると考えられる口の開閉情報のみを与えた場合でもマガーク効果が起こることを示す。

聴覚刺激： 女性一名が自然発声した/pa/, /ta/, /ka/の単音節に白色雑音を重畳し、全体のSN比で0 dBと5dBの2種を用意した(16 bit, 16 kHz サンプリング音声)。雑音重畳後の刺激のヘッドフォンによる提示音圧は約70db SPLであった。

視覚刺激： 音節/ba/を発声している顔画像ビデオから、両唇を閉じている画像(C)、半開きの画像(H)、完全に開いている画像(O)の3つの静止画像を取り出し、音声の開始時点でCからOに変化するCO条件とHからOに変化するHO条件の2条件を用意した(図2)。CO条件は両唇音情報を、HO条件は非唇音情報を有すると考えられる。しかし、これらの2枚の静止画像の提示では唇の開閉以外の情報はほとんど有しないと考えてよい。

装置： 音声および画像の提示には岩通アイセル社製視聴覚刺激提示装置(IS-801Az)を用いた。本装置は、8 msの精度で画像と音声を任意のタイミングで提示することが可能である。

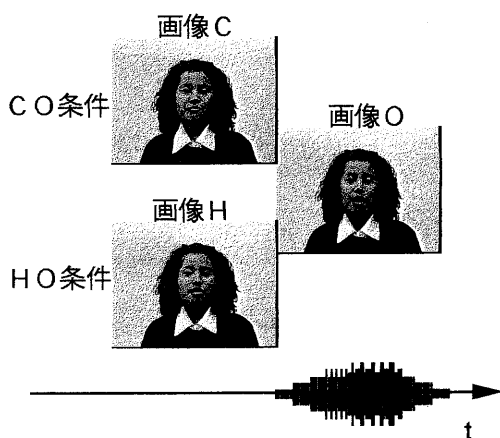
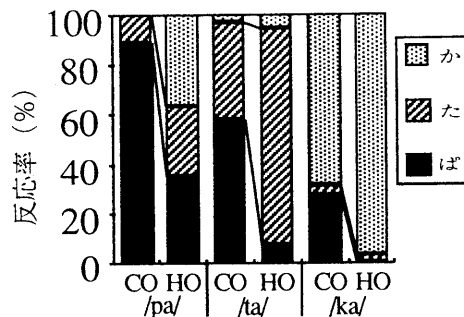


図2 2.1の実験に用いた視覚刺激と音声とのタイミング



音声と画像の種類のみ組み合わせ

図3 音声知覚に対する顔画像の影響

被験者： 20代前半の女性4名であった。

タスク： 被験者は提示された音声か、「ば」、「た」、「か」、「その他」のいずれであったかを反応ボタンによって回答した。

手続き： 音声3種*画像2種*雑音2種の12通りの組み合わせを5回繰り返す60試行を1つのセッションとし、各被験者は全部で10セッション、計600試行を行った。提示順はランダムとした。

結果： 図3に雑音条件S/N=5dBの場合の音声/pa/, /ta/, /ka/に対するCO条件とHO条件の結果を示す。図3をみてわかるとおり、音声/pa/に対しCO条件では90%近くが「ば」と答えるが、HO条件では35%程度しか「ば」という反応がなくなり、そのかわり、「た」と「か」の反応が合わせて60%以上を占める。逆に音声/ta/および/ka/に対しHO条件ではそれぞれほとんど「た」および「か」と反応するが、CO条件では「ば」と反応する率が音声/ta/および/ka/に対し60%および30%程度まで増加する。

この結果から、2枚の静止画で与えられた唇の開閉情報によってもマガーク効果が起こることが確認されたといつてよい。また、中園⁵⁾は、音声/pa/に対し音節/ka/の/k/を発声する瞬間の顔の静止画を提示してもマガーク効果が生じることを報告している。一方、GreenとKuhl⁶⁾は、唇情報から得られる調音位置情報と音声情報から得られる調音位置とが統合され、さらに有声・無声子音の弁別情報(VOT: Voice Onset

Time)が統合されることを示した。例えば、VOTを変化させて/ibi/-/ipi/の範囲で聞こえる音声と同時に/igi/を発生している顔画像を提示すると、/idi/-/iti/の範囲で音声聞こえるというものである。

以上のように唇の開閉が音声知覚における両唇音と非唇音の弁別に影響を与えているのは明らかである。もちろん、あごの動きや舌の動きなども調音位置情報としてとらえられるならば同様な影響を与えることは容易に想像できる。では、唇の動きの情報から調音位置情報を抽出し、音声情報と統合するのはどの程度自動的な処理なのかという問題がでてくる。SekiyamaとTohkura⁷⁾は、日本人は米国人よりマガー効果が起こりにくいという結果を報告し、この結果を、対面対話状況において相手の顔を直視する欧米人との文化的違いによって説明している。これは、唇の動きとそこから発せられる音声を経験に基づいて関連付けされることによって、画像情報と音声情報の統合され易さが強化されるということであり、マガー効果が経験に基づいて形成される過程であるように思われる。しかし、日本人でもマガー効果が全く起きないわけではなく、音声に雑音を重畳した場合には顕著に現れる。以上から、画像情報と音声情報を統合する過程はある程度自動的な過程であるが、唇の動きから調音位置を抽出する過程は経験に依存していると考えることができる。

2. 2 口の開閉情報と音声のタイミング

次に、唇の動きと音声のタイミングについて考えてみる。音声も唇の動きも時間的変化をともなった情報として入力される。そのため、どの時点での情報が統合されるのかが重要なポイントとなる。坂口ら⁸⁾は音声と画像の同期のずれが200ms以上になるとマガー効果の起こる率が有意に低下することを報告した。また、橋本ら⁹⁾は、文の了解度が顔画像を提示することによって上昇するが、音声を画像に対して遅延した場合、120ms以上遅延すると画像提示による了解度の上昇効果が低下することを示した。

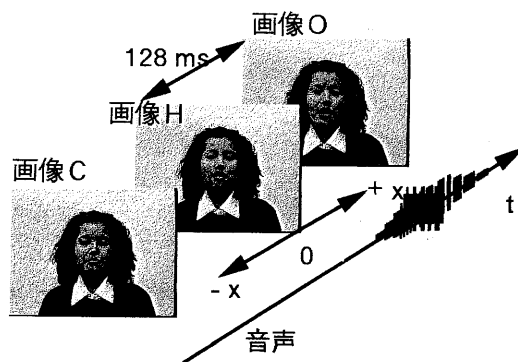


図4 2. 2の実験に用いた視覚刺激と音声とのタイミング

このずれの限界である120から200msは日本語音声の基本的単位であるといわれるモーラの長さに関係していると坂口らも橋本らも結論している。すなわち、モーラ長を越えて同期がずれる場合にマガー効果が現れなくなると考えられる。確かにモーラを単位として唇情報と音声を同期させて統合を行っていると考えるのは妥当であるように思える。しかし、ここではさらに細かく検討してみる。たとえば、どんな音声を発生するにしろ、発声前は唇を閉じており発声が終わればまた閉じるのが普通である。両唇音である/ba/を発生する場合であっても、唇を完全に閉じている時間は非常に短く、その発声の途中では半開きの状態もあれば、完全に開口している状態もあるのである。そこで次のような実験を行った。

視覚刺激: 2. 1の静止画像C, H, Oを連続して提示する。Hの提示時間は128msで固定し、音声の開始時刻に画像がCからHに変わるタイミングを0、音声開始される以前にCからHに変わる条件を負、音声開始された後にCからHに変わる条件を正で表して、-320, -256, -160, -128, -96, -64, 0, 64, 128, 256msの10種類のタイミングで提示した(図4)。この実験設定では、画像自体が持っている情報はすべての実験条件で全く等しい。しかし、その提示タイミングを変化させることによって、音声開始時の唇の形を変化させることとなる。例え

ば、音声開始時点の画像は、タイミング条件が正では画像Cの両唇が閉じた画像、-96、-64、0 msでは画像Hの半開きの画像、-128 ms以前では唇が大きく開いた画像Oが提示されていることになる。また、音声刺激長が230、240、260 ms (/pa/, /ta/, /ka/の順)であったことから、タイミング条件256 msの場合は音声の終了時に画像CからHに変化するタイミングである。

手続き：音声3種*画像1種*タイミング10種の30通りの組み合わせを3回繰り返す90試行を1つのセッションとし、各被験者が全部で10セッション、計900試行を雑音種毎に別々に行った。最初に雑音重畳5dB条件、別の日に0dB条件を行った。提示順はランダムとした。

聴覚刺激、被験者、装置、タスクはすべて2. 1と同様であった。

結果：雑音重畳条件S/N=5dBでの音声/pa/に対する各視覚条件の結果を図5に示す。音声の開始から64 ms以上後に画像CからHに変化するタイミング64ms以上の条件で「ば」と反応する率が一番高く約90%を占める。また、音声開始時点で既に画像CがHに変化している-128から0 msの条件では「ば」の反応が減少し、その代わりに「た」、「か」が増加する。すなわち、マガーク効果が起こる率が増加する。そして、タイミング-160 ms以下では、50%近くマガーク効果が起きていることがわかる。

本実験結果は、マガーク効果において音声開

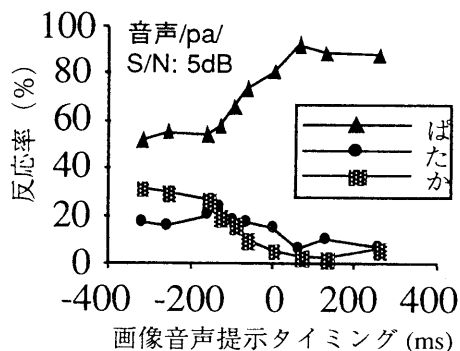


図5 顔画像の提示タイミングの違いによる音声知覚への影響

始時点の唇の形状が音声知覚に影響を及ぼしていることを示している。すなわち、視覚情報と聴覚情報の統合は音声の開始時に同期して行われていることが示唆される。

MacDonaldとMcGurk¹⁰⁾は、マガーク効果をモーター理論¹¹⁾によって説明可能であると述べている。モーター理論は、知覚過程を、情報を受けとり解析する受け身の処理としてとらえるだけではなく、音声を発声する際の調音過程と関連づけた能動的な過程としてとらえるものである。すなわち、音声の発声に伴う唇のうごきや舌の動きの指令は、音声知覚においても聴者の脳内で模擬的に起動しているとする。そして、この模擬的に起動する聴者の脳内の調音活動は、音声と同期して行われるのみならず、知覚的に音声の予測を可能なものにする。この意味で、ここで示した音声と画像のタイミングのずれに関する結果の説明には、モーター理論による説明が一応有効であるように思われる。

3. 文字情報の音韻知覚への影響

Dijkstraら¹²⁾は、文字提示によって音節の二者択一選択反応時間が短くなることから、様相間プライミング効果(cross-modal priming effect)の存在を示した。そして、文字は即座に音韻的な符号に変換されて音声処理に影響を及ぼしているとしている。近藤と笈¹²⁾は、雑音によって明瞭度の低下した音節の知覚が同時に提示される文字によって変化することを報告した。しかし、文字の影響によって、マガーク効果にみられる融合現象は起きず、文字情報の影響は音韻範疇的であることを主張している。たとえば、音声/ba/に対して文字「が」を提示すると/ga/と聞こえる率は有意に上昇するが、/da/と聞こえるようにはならない。また、彼らは音声自体の明瞭度が極端に低く音韻が全く聞き取れない場合は文字提示によって音韻が聞こえるようにはならないことを示し、文字の影響は、雑音などの影響で音声に曖昧性があり音韻知覚の決定段階においても候補が複数存在する場合において候補の選択に文字情報が影響を及ぼしてい

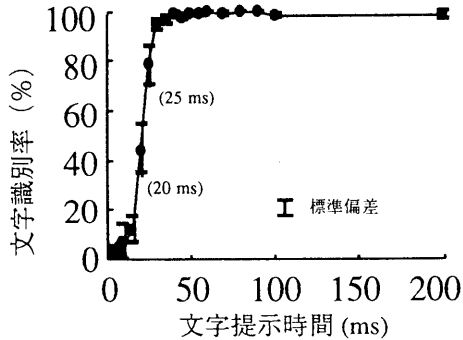


図6 文字の提示時間に対する文字識別率

らとしている。本稿では、文字の影響をさらに詳細に検討するため、文字提示時間による文字「ば」、「た」、「か」の識別閾、雑音重畳量による音節/pa/, /ta/, /ka/の識別閾と、文字提示が音韻知覚に及ぼす影響との関係を検討する。

3. 1 文字と音節の識別閾

まず、文字、音声それぞれ単独に提示した場合の、ひらがな文字「ば」、「た」、「か」の文字提示時間に対する識別閾、および音節/pa/, /ta/, /ka/の雑音重畳量に対する識別閾を測定した。被験者は4名、用いた音声は、実験2. 1および実験2. 2で用いた音声の雑音重畳量を変化させたものである。また、その他の実験装置や実験手続きも先の実験と同様であった。結果を図6と図7にそれぞれ示す。被験者は識別不能な場合に「X」と反応することが許されたため、最低値は0%となっている。図6から文字提示時間が30 msより短くなると文字「ば」、「た」、「か」の識別率が低下し、25 msで約80%、20 msでは50%以下となることがわかる。また、図7から音節/pa/, /ta/, /ka/の識別率は、音声のSN比が5 dBで約80%まで低下し、0 dBでは50%以下になることがわかる。

3. 2 文字提示時間と文字が音韻知覚に及ぼす影響

文字と音声を同時に提示した際に、文字の提示時間を変化させた(10, 20, 30, 40, 50 ms)場合

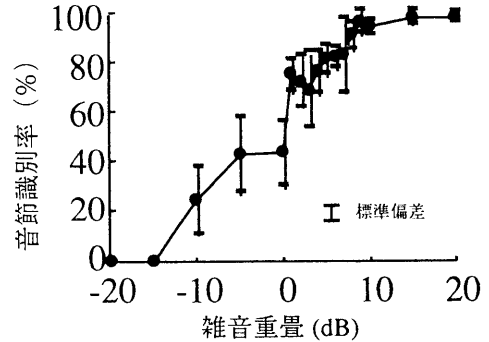


図7 雑音重畳量に対する音節識別率

の音節の識別実験を行った。文字刺激条件は、音声と一致しているひらがな(文字・音声一致条件)、音声と一致していないひらがな(文字・音声不一致条件)、および文字でない「*」を提示した条件(コントロール条件)の3条件であった。音声刺激は/pa/, /ta/, /ka/の3種類、雑音重畳量は0 dBであった。雑音重畳方法その他の条件および手続きはすべて実験2. 1と同様であった。図8に結果を示す。文字の識別閾と思われる25 msで音声と一致した文字提示の影響がなくなっている。すなわち、文字情報も音声情報も曖昧な場合に、曖昧な情報同士が統合されて音韻知覚が促進されることはないことを示している。

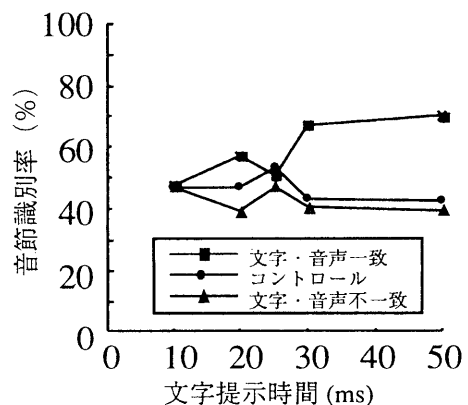


図8 音節知覚に対する文字提示の効果(文字の提示時間による違い)

3. 3 音節の雑音重畳量と文字が音韻知覚に及ぼす影響

文字と音声と同時に提示した際に、音声の雑音重畳量を変化させた(-10, -5, 0, 2, 5 dB) 場合の音声 /pa/, /ta/, /ka/ の識別実験を行った。文字提示時間は50 msであった。その他の条件はすべて、実験3. 2と同様であった。図9に結果を示す。雑音重畳量が0 dB以下の場合でも音声・文字一致条件では文字情報の影響によって識別率が低下しないのがわかる。

3. 4 文字の提示タイミングと文字が音韻知覚に及ぼす影響

近藤と筑¹³⁾は、文字を音声開始に対して1秒前に提示した場合においても、文字を提示しない場合に比較して、音声と一致した文字を提示した場合には音節の認識率が有意に上昇し、音声と異なる文字を提示した場合には異聴率が有意に上昇することを報告した。しかし、文字を同時に提示した場合と比較するとその影響は小さいとしている。

3. 5 文字が音韻知覚に及ぼす影響とマガーク効果の比較

本節では、文字が音韻知覚に及ぼす影響について検討した。音声を発する唇の動きはそこか

ら発声される音声と一致している必然性がある。しかし、文字情報と音声情報はたとえ同時に提示されたとしても、それらが全く同じ情報を有する必然性はない。マガーク効果に比較して文字情報は時間的に幅広く音韻知覚に影響を与える原因の一つとして、この視聴覚情報間の共起必然性の違いを挙げることができる。また、文字情報が音韻の情報を持っているのに対し、唇画像情報はそれ自体では音韻を確定するだけの情報は持っておらず、調音位置の情報という音韻の決定に関しては部分的弁別素性情報のみを持っている。この特性から、文字提示ではマガーク効果の様な融合現象が起きずその影響が音韻範疇的であったと言える。逆に、唇音と非唇音の区別に比べ、歯茎音と軟口蓋音の区別は顔画像からだけでは明確でないことから、マガーク効果における融合現象は、融合というよりは、単に調音位置情報と音韻特徴情報の統合だと考えることも可能である。以上の唇画像と文字が音韻知覚に影響を及ぼす現象は、視覚から得られる特徴情報と音声から得られる特徴情報が統合されるという観点では共通していると言えるものの、視聴覚情報の入力タイミングに対する依存性や統合される情報の質や統合のレベルにおいては異なっていると考えられる。

4. 視覚情報と音声情報の統合モデル

知覚における情報の統合過程はMcllerandとRumelhart¹⁴⁾の相互活性化モデル(Interactive Activation model)やMassaro¹⁵⁾のFLMP (Fuzzy Logical Model of Perception)によって説明が試みられている。しかしながら、これらのモデルでは視聴覚処理のタイミングについては明示的には考慮されていない。例えば、実験2. 2の同じ画像情報をタイミングだけを変えて提示した場合の知覚の変化は、これらのモデルを直接適用しても説明できない。なぜならば、視覚情報の特徴から音韻知覚において統合されるべき調音位置情報を得るためには音声とのタイミング情報が必要であると考えられるからである。

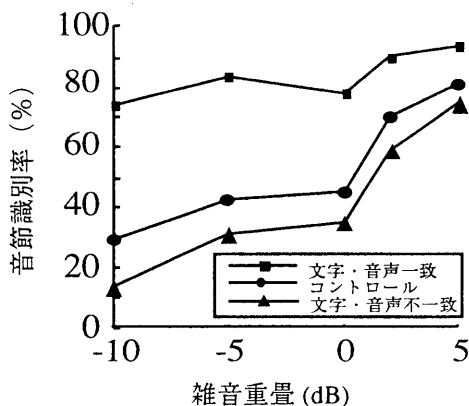


図9 音節知覚に対する文字提示の効果 (雑音重畳量による違い)

また、これらのモデルでは、どんな特徴情報が統合され、どんな情報は統合されないのかは、適用される知覚現象毎に仮定される特徴情報に依存してモデルを適用する者が定めることができ、モデル自体には統合されるべき情報をそのときどきに応じて選択する機能は表現されていない。

一方、Bregman¹⁶⁾の聴覚の情景分析(Auditory Scene Analysis)では、特徴情報の分凝(segregation)と関連しあう特徴の群化(grouping)、さらに、時間を含む特徴の流れ(stream)を形成することによって、同時に複数存在する特徴情報や時間的に分散する特徴情報から知覚が生じる過程を説明している。この考え方は、聴覚の比較的低次のレベルにおける現象、例えば、混在する音の中からそれぞれの音源に関する音を分離する過程等を説明する理論として現在注目を集めている。本稿で取り上げたマルチモーダルな知覚過程においても、どんな特徴情報同士が統合されて最終的に一つの事象として知覚されるべきであるか、特に互いの情報のタイミングに依存した統合過程を説明する上では、このような聴覚の情景分析の考え方が適用可能であると筆者は考えている。

最後に、本稿で示した視覚情報が音韻知覚に影響を及ぼす現象におけるマルチモーダルな知覚過程をまとめてみると以下の様になる。

視覚および聴覚から得られる特徴情報は、互いに関係する情報のみではない。まず、視覚情報から唇の動きや文字を知覚するが、これ自体は視覚内で閉じた処理である。そして、唇の形や動きから調音位置の情報を抽出するには、唇の形や動きと調音位置を関連づける知識が、同様に、文字から音韻的符号情報を抽出するには、文字と読みの対応の知識が脳内に形成されていなければならない。そして、情報の入力タイミング等の時間情報やモータ理論的な口の動きと音声の関係の知識等が制約条件となって、得られた調音位置情報や音韻符号情報と聴覚処理によって音声から抽出された音韻の特徴情報が、一つの事象としてとらえられるかどうか

判断される。そして、最終的に一つの事象としてとらえられる範囲の特徴情報が統合された結果が知覚として生じると考えることができる。

謝辞

常日頃からご指導頂くNTT基礎研究所の平原達也主幹研究員、本論文の執筆にあたり有意義な御助言を頂いた同柏野牧夫研究主任並びに名古屋大学寛一彦教授に感謝いたします。

文献

- 1) McGurk, H. and MacDonald, J., "Hearing lips and seeing voices," *Nature* 264, 746-748 (1976).
- 2) Dijkstra, T., Schreuder, R. and Frauenfelder, U.H., "Grapheme Context Effects on Phonemic Processing," *Language and Speech* 32 (2), 89-108 (1989).
- 3) Frost, R., Repp, B. H. and Katz, L., "Can speech perception be influenced by simultaneous presentation of print?" *J. of Memory and Language* 27, 741-755 (1988).
- 4) 近藤公久, 箕一彦, "音声情報と同時に提示される文字情報の音声知覚に与える影響", *音響誌* 51, 7, 425-436 (1995).
- 5) 中園薫, "マルチメディア通信における動画像フレームレートの音声知覚への影響", *信学技報 HC93-68*, 17-24 (1994).
- 6) Green, K.P. and Kuhl, P.K., "The role of visual information in the processing of place and manner features in speech perception" *Perception and Psychophysics* 45 (1), 34-42 (1989).
- 7) Sekiyama, K., Tohkura, Y., "McGurk effect in non-English listeners: Few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility" *J. Acoust. Soc. Am.* 90 (4), Pt.1, 1797-1805 (1991).
- 8) 坂口裕昭, 世木秀明, 出口利定, "音声知覚における視覚の影響", *音講論集* 1-7-4, 277-278 (1993-3).
- 9) 橋本正浩, 林安紀子, 世木秀明, "聴覚と視覚を併用した音声知覚とモダリティ間のタイミングの関係", *信学技報 SP92-39*, 1-6 (1992).
- 10) MacDonald, J. and McGurk, H., "Visual influences on speech perception processes," *Perception & Psychophysics* 24, 3, 253-257 (1978).
- 11) Liberman, A.M., Cooper, F.S., Shankweiler, D.P. and Studdert-Kennedy, M., "Perceptin of the speech code," *Psychological Review* 74, 431-461 (1967).
- 12) 近藤公久, 箕一彦, "単音節中の子音知覚に及ぼす文字提示の影響", *信学技報 SP93-43*, 65-72 (1993).
- 13) 近藤公久, 箕一彦, "音節知覚に及ぼす文字提示の影響—文字提示のタイミングによる影響—", *音講論集* 3-7-9, 477-478 (1994).
- 14) McIlerand, J. L., Rumelhart, D. E., "An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings", *Psychological Review* 88 (5), 375-407 (1981).
- 15) Massaro, D. W., "Integration of multiple sources of information in language processing," *Attention and Performance XVI*, Inui, T. and McClelland, J. L., ed., MIT press, MA (to appear).
- 16) Bregman, A.S., "Auditory Scene Analysis- The perceptual organization of sound", MIT press, MA (1990).