

1. はじめに

第二言語習得における聴解の重要性は、多くの研究で言及されてきた (Lynch, 1998 ; 2011 ; Vandergrift, 2004 ; 2007 ; 福田, 2004 ; 横山, 2004b)。聴解は、生理的・認知的プロセスに関与しており、文脈や非言語的・文化的な手がかりにも注意を払わなければならないという特質があるが、近年聴解に関する研究の多くは、その認知的プロセスについて解明しようとしてきた (Vandergrift, 2004 ; 2007)。その中で、映像メディアの発達と普及により、聴解への視覚情報 (Visuals) の関与が徐々に注目されるようになった。なぜなら、殆どの目標言語使用場面において、学習者は聴覚情報と視覚情報を統合 (integrate) しているため、視覚情報は聴解において重要な側面であると考えられているからである (Baltova, 1994 ; Buck, 2001 ; Kellerman, 1990 ; 1992 ; Lynch, 1998 ; Rost, 2011 ; Suvorov, 2015 ; Vandergrift, 2007 ; 竹原・野村, 2004)。

一方、我々の生活には通常視覚情報が存在するが、教室指導や聴解試験などの制約により、音声情報とともに与えられず、聴解授業で扱う場面の現実性の欠如が指摘されている (Buck, 2001 ; Kellerman, 1990 ; 松崎, 2008)。また、英語の大型語学試験の聴解問題は、パソコンでの受験の実現により、視覚情報の一つである静止画を取り入れたが、動画の使用に関しては慎重な姿勢を見せている。第二言語習得の聴解への視覚情報の影響を調査した研究は多いものの、視覚情報の捉え方を初め、学習者の学習背景や習熟度、聞くテキストの種類など、様々な要因が存在するため、いまだに視覚情報のどの部分がどのように働いているかが解明されておらず、多くの課題が残っている。それゆえ、Buck (2001) と Lynch (1998) は、聴解力の構成概念としての妥当性が保証されない限り、安易に視覚情報を取り入れるより、従来の言語的知識に集中するように主張した。

しかし、第二言語の聴解指導の最終目標は、学習者が日常生活においてあらゆる情報を駆使して目標言語を理解し、コミュニケーションの目的を達成できることにあると思われる。特に実生活で多く遭遇すると想定される対面聴解においては、視覚情報が必然的に存在し、常に何らかの形で学習者の理解に影響していると考えられる。従って、従来音声の言語的情報しか含まれていない聴解力の構成概念に、視覚情報を取り入れる前段階として、その視覚情報が学習者の聴解に影響する諸様相を丁寧に確かめていく必要がある。

そこで、本稿はまず第二言語習得の観点から聴解における視覚情報の定義と性質を整理し、次に視覚情報に関する第二言語習得の聴解研究の変遷と成果を概観し、最後に先行研究に残された課題及び今後の展望について述べる。

2. 聴解における視覚情報

2.1. 視覚情報とは何か

聴解における視覚情報は、大きく四つに分類できる。「言語活動が行われた環境」、「話し手の顔による情報（表情や口の動き）」、「話し手の身体による情報（ジェスチャーや身体の動き）」と「発話内容に関する写真・文字情報」である。第二言語習得の聴解研究では、各研究者の視覚情報に対する捉え方により、どの用語で表現し、どこまでの範囲を指しているかが異なる。表1では、Buck (2001)、Kellerman (1990)、Lynch (1998)、Rost (2011) と Vandergrift (2007) が使用した用語及びその詳細の整理を試みた。

表1 第二言語習得の聴解研究における視覚情報に当たる用語及びその詳細

	Buck (2001)	Kellerman (1990)	Lynch (1998)	Rost (2011)	Vandergrift (2007)
視覚情報に当たる用語	Non-verbal signals	Visual information	Setting・ Behavioral environment	Non-verbal cues/Visual signals	Visuals
用語の詳細					
言語活動が行われた環境 (コンテキスト)	—	✓	✓	✓	✓
話し手の表情・口の動き	✓	✓	✓	✓	✓
話し手のジェスチャー・ 身体の動き	✓	✓	✓	✓	✓
発話内容に関する写真・ 文字情報	—	—	—	✓	✓
その他	—	話し手の 外見	—	話し手の 視線	—

表1で示されたように、各研究者が捉えている視覚情報の範疇が相違している。まず、今回取り上げた5つの文献ともに「話し手の顔と身体からの情報」に言及しており、次に多く言及されたのは「言語活動が行われた環境」である。そして「発話内容に関する写真・文字情報」が含まれているのは、2007年以降の文献のみである。

本稿では、聴解における視覚情報の研究を包括的に分析できるよう、Vandergrift (2007) に従い、聴解における視覚情報を以下のように定義する。聴解における視覚情報 (Visuals) とは、目標言語を聞いて理解する際に、その言語活動が行われた環境と、話し手の顔による情報、話し手の身体による情報、そして、そこで提示された発話内容に関する写真・文字情報のことである。

2.2. 発話理解からみた視覚情報の特徴

奥野 (2016) は、学習者の自然会話のデータから、彼らが興味や目的に合わせ、意味交渉をすると同時に、視覚情報などあらゆる手段で発話を理解し反応しようとしているのが観察できたと述べている。殆どのコミュニケーション場面に存在する視覚情

報は、相手の発話を理解する際に、音声情報を補足し、発話を解釈する文脈の範囲を制限してくれる(Buck, 2001 ; Kellerman, 1990 ; Lynch, 1998 ; Vandergrift, 2007) 。特に対面コミュニケーションでは、熟知している相手なら、視覚情報はその発話を解釈する証拠をより多く提供すると Rost (2011) が述べている。次に、この音声情報の補足と解釈範囲の制限についてみていく。

2.2.1. 音声情報の補足

音声情報の補足に関して、最もよく挙げられる視覚情報は相手の顔情報である。積山 (2014) によると、「一般に、聴覚情報だけでの正確な聞き取りが困難な状況では、顔が見えていることによって音声聴取の正確さがおおいに向上する。(中略) 顔の中でも、口や顎を含む下半分が、言語的な視覚情報を多くもたらすといわれる。(p.32)」。McGurk & MacDonald (1976) では、英語母語話者に音声と口形が異なる映像を使用した英語の単音リピート課題で実験した結果、視覚と聴覚の情報が系統的に融合された誤答が見られ、子供より成人のほうが視覚の影響を受けたという。この McGurk らにより初めて報告された視覚と聴覚の情報が統合した現象を「マガーク効果」ともいわれ、口形による単音の調音情報使用の可能性が示された。

積山 (2014) は、視覚情報による補足ができるには、声と口の動きの対応関係が分かることが前提にあり、「音声-口形マッチング」ができるかどうかは、視覚と聴覚の情報が統合された知覚が生じるための第一歩であると述べている。では、この視覚情報とのマッチングはどのように習得されるのか。Lewkowicz & Hansen-Tift (2012) は、4-12 か月の幼児を対象に、母語の英語と母語でないスペイン語の独白を聞かせ、選好注視法で調査した。その結果、英語で話しかけられた場合には、4-8 か月の幼児は目より口元のほうを注視するが、12 か月を過ぎた途端に口元を注視する割合が減っていく。これに対しスペイン語独白の場合には、12 か月になっても口元を注視していたことが分かった。この結果について、Lewkowicz らは以下のように推論している。4-12 か月の幼児は音声と口形のマッチングを習得するため、発話者の口元から視覚の音韻的手がかりにアクセスしていたが、12 か月後にはそのマッチングの習得が終了し、幼児は他の知識を習得するため再び目を注視し、社会的手がかりを求めようになる。しかし、母語でない言語を聞く場合、専門化した母語の知識と知覚狭小化¹⁾により、その言語の処理が困難なため、12 か月以後の幼児でも、引き続き視覚の音韻的手がかりを求め口元を見ていたという。

成人の第二言語学習者の「音声-口形マッチング」習得に関する研究は僅かであるが、日本に滞在する中国語母語話者を対象とした Sekiyama (1997) の研究が挙げられる。Sekiyama (1997) は、日本に留学で滞在している中国語母語話者を対象に、マガ

¹⁾ 母語にない音韻対立を聞き分ける能力は、生後6ヶ月以降に次第に失われ、音声処理は生後12か月までに母語に特化したものとなるが、これは母語の効率的な処理に集中できる体制を整えるためだと考えられる(積山2014)。

ーク効果を測定した。その結果、彼らの滞在期間とマガーク効果の相関は非常に高い ($r=.723$) ことが判明した。そこから、Sekiyama (1997) は日本での生活経験を通して口形による調音情報が積極的に利用されるようになり、非母語においての視覚情報の付加の影響がより顕著に見られたと考察している。

McGurk & MacDonald (1976) を始め、Sekiyama (1997) 等の研究は、非言語的の情報として考えられていた視覚情報にも、発話と直接に関わる言語的な部分が存在していることを示した。しかし、単音レベルに現れる視覚情報は、大量な情報が絶えずに流れてくる実際のコミュニケーションの場合、学習者の理解にどのような影響するのかという疑問が残る。

2.2.2 解釈範囲の制限

Lynch (1998) と Rost (2011)、Vandergrift (2007) は、視覚が提供した文脈の手がかりと非言語的情報は非常に重要であり、相手の発話を理解する際に考慮すべき潜在的な解釈制限をしてけると述べている。これにより、発話のトップダウン処理が優位になり、全体主旨の把握が（特に初級レベルの学習者に対し）促進されるという。

島田 (2010) は、「聞く」能力について、学習者の自己評価 *Can-do statements* と客観テスト（大学のプレースメントテスト）の相関を比較した結果、「郵便局や銀行の窓口の説明・ゼミや討論の議論・学校職員の事務連絡」といったよりあらたまった内容の場面との相関が高く、「引き受けてくれたか断っているか・買い物時の値段・親しい人同士のくだけた日本語」といったよりインフォーマルで社会文化的知識を必要とする場面との相関が低いということを明らかにした。これらの結果は、学習者の習熟度・学習背景や、タスクの定型性や双方向性の度合いなどといった場面の質的な違いが影響している可能性が考えられる。また、フォーマルとインフォーマル場面にはいずれも視覚情報が含まれているが、よりインフォーマルの場面において、視覚が提供した文脈の手がかりと非言語的情報はそこで必要となる社会文化的知識を補い、トップダウン処理がより促進されるため、テストで測定された聞く能力の高低に関わらず、学習者自身はより達成できると自己評価した一因となった可能性も指摘できる。

Lynch (2011) や Rost (2011)、Suvorov (2015) は、日常場面に限らず、アカデミックの場合も、講演や講義中の講師の顔・ジェスチャーや板書、教室の前方に提示された図表やスライド等は、内容を理解する上で重要な視覚情報だと考えられている。例えば、学会会議でのプレゼンテーションでは、短時間で難易度の高い命題内容を聞く側に伝達しなくてはならないが、内容のアウトラインなどを視覚で提示すれば、非母語話者の聞き手にとって記憶の負担が軽減され、自分の言語能力でカバーできない部分も補われる。このような工夫により、非母語話者の聞き手がプレゼンテーションの大事なポイントを理解できる可能性が高くなる (Rowley-Jolivet, 2002)。

Baltova (1994) は動画を提示し多肢選択問題で学習者の理解度を測定した結果、発話の解釈範囲を制限する視覚情報は全体主旨の把握を促進するが、細部までの理解はそれほど促進しないと指摘した。一方、呉 (2016) は視覚情報の効果を説明文の母語

翻訳課題で測定した際に、音声群より映像群のほうが説明文にあるカタカナ語を多く再生したと確認した。発話理解における視覚情報の大きな役割：音声情報の補助と解釈範囲の制限には、各々の場面の質的特徴や、学習者の習熟度・学習背景、課されるタスクの種類とそれに対する熟知度などといった様々な要因が関与し、一概には言えないはずである。これら複雑に絡み合った要因と視覚情報の働きを解明するのに、第二言語習得における実証的な聴解研究が必要となるだろう。

3. 聴解と視覚情報に関する第二言語習得の実証研究

1990年代頃から、ビデオの普及により、文脈の手がかりを提供してくれる視覚情報を取り入れ、聴解教材の真正性を高めようと試みるいくつかの研究がなされてきた。その後、パソコンを媒体とした試験の出現で、使用される視覚情報の種類や、テキストの種類の違いを視野に入れ、視覚情報を取り入れる聴解テストの構成概念 (construct) の妥当性を検討した研究もいくつか見られる。この節では、これらの実証的研究を「文脈としての視覚情報の効果」と「使用される視覚情報の種類」、「テキストの種類による使用差」という三つの側面から概観していく。

3.1. 文脈としての視覚情報の効果

Baltova (1994) は、カナダ在住の8年生のFSL46人を対象に、動画群(映像と音声)と音声群(音声のみ)に分け、フランス語の物語の動画(15分)を提示した後、多肢選択問題で理解度を測った。その結果、動画群のほうが成績が高かったが、事後アンケートでは、話された内容が難しく、話す速度も速かったとの意見が多かったと述べている。また、動画群の4人が、ジェスチャー等を含むシーンと比べ、音声のみに集中した方が理解しやすく、長い対話シーンのほうが理解が難しいとの報告もあった。このことから、Baltova (1994) は映像を加えた方が物語の理解を促進するものの、

表2 「文脈としての視覚情報の効果」に関するSLAの実証研究

著者	素材	提示手法			参加者	測定方法	主な結果
		音声のみ	静止画	映像			
Baltova (1994)	子供向けの15分の動画(仏語)	✓	—	✓	8年生のFSL/46名	多肢選択問題及びアンケート	動画群は：①音声群より得点が高かったが、②細かい内容の理解までは促進されなかった。
Progosh (1996)	日常会話と料理番組(英語)	—	—	✓	日本人大学生のESL/62名	7段階尺度評価	ビデオは内容的に面白いとの答えが多かったが、「ビデオテストはより簡単だ」という問いに答えが分かれた。
Coniam (2001)	30分のトーク番組(英語)	✓	—	✓	NNS 英語教師/104名	記述問題及びアンケート	音声群と映像群の記述問題の点数において有意差が出なかった。

必ずしも話された内容を理解するとは限らないと結論づけた。

Coniam (2001) は、香港のNNS 英語教師 104 人を対象に、音声群とビデオ群に分け、30 分のトーク番組を提示し、記述問題及びインタビューを行った。その結果、記述問題での両群の点数に差が見られず、ビデオ群のインタビューでは「映像が邪魔だ」という答えもあったと述べている。

Progosh (1996) は学習者の視点から調査を行った。62 名の日本人大学生に、英語の会話と料理番組のビデオを視聴させた後、多肢選択問題・筆記問題及び7段階の尺度でビデオに対する印象を評価してもらった。その結果、圧倒的にビデオを使用したテストのほうが面白くてよいとの答えが多かったものの、「ビデオテストのほうがテープテストより簡単だ」という質問には答えが分かれたと報告している。これは Baltova (1994) の結果と一致していると考えられる。

Baltova (1994) と Progosh (1996) 、Coniam (2001) のいずれも、視覚情報を分類せずに大きな「文脈に関する情報」として捉えた。Baltova (1994) の扱った素材には物語性が強く、言語に頼る比重が小さくなるため、細かい言語情報が聞き取れなくても、動画群のほうがより理解できたと推測できる。それに対し、Coniam (2001) の対象者は一般の学習者より習熟度が高いと思われる教師であること、理解度を再認ではなく再生で測定したこと、さらに再生された内容の採点方法が明確に記述されていないことから、文脈としての視覚情報（話者・背景等の情報）が理解を促進させた原因になったかどうかは確認できない。また、大きな文脈情報として捉えたため、視覚情報のどの部分がどう働いたのかも結果から見えない。聞くテキストの特性や学習者の習熟度との関係を明らかにするためにも、視覚情報をさらに分類し、それぞれの種類の有用度をみていく必要がある。

3.2. 使用された視覚情報の種類

Ockey (2007) は、アメリカ在住の ESL 学習者 6 人（初・中・上級 2 名ずつ）を対象に、政治学の講義を「静止画＋音声」と「ビデオ」の二つの形式で提示し、画面注視時間・視聴後の口頭内省・アンケートを通し、視覚情報の使用状況を質的に分析した。その結果、静止画より動画の注視時間が長く、口頭内省から使用された視覚情報に口・手・体の動き及び表情などが挙げられた。しかし、使用された視覚情報の数と種類、視覚情報が補助として役に立ったかという点に関しての個人差が大きく、初級レベルにも上級レベルにもまったく画面を注視しない参加者がいたと述べている。

Wagner (2008) は、アメリカ在住の ESL 8 人を対象に、日常会話と講義のビデオを視聴した後、口頭内省と多肢選択問題をさせ、質的分析を行った。口頭内省では、ハンドジェスチャー・ボディランゲージ（表情を含む）・登場人物と背景による文脈情報の使用が報告されたが、Ockey (2007) の結果と同様に、使用能力に個人差が観察されたと述べている。特に登場人物や背景などの文脈情報に頼り、最初の段階で話題の仮説を立て、さらに音声のみでは判定しにくい、二人以上の登場人物を区別しやすくなる点から、視覚情報は学習者の理解を促進するものだと主張した。

表3 「使用された視覚情報の種類」に関する SLA の実証研究

著者	素材	提示手法			参加者	測定方法	主な結果
		音声のみ	静止画	映像			
Sueyoshi & Hardison (2005)	陶器入門の講義 (英語)	✓	—	✓	ESL/42名	多肢選択問題及びアンケート	下位群は顔+手の条件で、上位群は顔のみの条件が最も点数が高かった。アンケートでは、口の動きを見ていたとの意識が上下位群ともに低かった。
Ockey (2007)	政治学の講義 (英語)	—	✓	✓	初中上級 ESL/2名ずつ	画面注視時間、事後口頭内省、インタビュー	静止画群と動画群の画面注視時間に有意差が認められたが、視覚情報の使用種類と数に個人差が見られた。
Wagner (2008)	会話・講義 (英語)	—	—	✓	ESL/8名	事後口頭内省及び多肢選択問題	ハンドジェスチャーや表情、文脈情報などの使用が報告されたが、使用状況に個人差が見られた。

上記の質的研究と異なる観点から視覚情報の種類を探ろうとする Sueyoshi & Hardison (2005) は、視覚情報を「顔から」と「顔とジェスチャーから」の二種類に分け、学習者の聴解への影響を調べた。彼らは42名のESLを上位群と下位群に分け、それぞれ音声条件、映像/顔条件と映像/顔ジェスチャー条件のグループに分け、陶器入門の講義を受けさせ、多肢選択問題で理解度を測定した。理解度に関して、上下位群ともに音声条件より映像条件のほうが優位だった。さらに映像条件の中で、上位群は映像/顔条件が最も成績が高かったことから、口形による調音情報を手がかりとして正確に使用できたと推測された。一方、下位群は映像/顔ジェスチャー条件のほうが成績が高かったことから、調音情報が効率的に利用できないため、ジェスチャーなどに頼れない映像/顔条件のほうが成績が低かったと考察された。しかしながら、視聴後のアンケートでは、映像視聴中にジェスチャーと口の動きを見ていたかとの質問に対し、上下位群ともに口の動きよりジェスチャーを見ていたとの答えが多かった。上位群のほうが口の動きからの調音情報を効率的に使用したため、映像/顔条件の得点が最も高くなると考察された一方、アンケートの回答からうかがわれた上位群の使用意識の低さが興味深いものであった。

以上三つの研究で抽出した視覚情報の種類「話し手の顔による情報」と「話し手の身体による情報」、「言語活動が行われた環境」は、Kellerman (1990) と Lynch (1998) の分類と一致した。Sueyoshi & Hardison (2005) の結果から、学習者の習熟度により、異なる種類の視覚情報が彼らの内容理解へ及ぼす影響の度合いが変化する可能性があること示された。また、Baltova (1994) が指摘した「長い対話シーンの理解は困難である」との意見と、Sueyoshi & Hardison (2005) の結果を合わせれば、文脈の手がかり

が豊富でジェスチャー等を伴う日常会話とそうでない学術講義に対し、学習者への視覚情報の影響も異なってくるのが考えられる。言い換えれば、異なるテキストに対して、それぞれの視覚情報が持つ内容理解への有用度も変わる可能性が示唆された。

3.3. テキストの種類から見た視覚情報の使用

Ginther (2002) は TOEFL の聴解の出題項目に使用される静止画の有無が学習者の聴解へどう影響するかを調査した。調査対象は TOEFL550 点で上下位群に分けられた ESL160 名で、全員米国での滞在期間が 2 か月以下であった。調査素材は生活会話・学術討論・講義の 3 種類で、静止画有りの条件では、登場人物と環境を提示する「講義/文脈条件」と、内容に関する図表を提示する「講義/内容条件」との 2 種類の講義素材を設けた。視聴後の多肢選択問題で分析した結果、講義/内容条件のみ、若干の促進が認められたが、その他の条件では有意差が認められなかった。また、講義/文脈条件は、音声のみの講義条件より得点が僅かに下回っていたという。

Wagner (2010) は 202 名の ESL 学習者を音声群と映像群に分け、生活会話と学術講義における視覚情報の影響を調べた。映像群に関して、生活会話には文脈情報（登場人物など）、学術講義には文脈と内容情報（図表など）を提示した。その結果、素材の種類に関係なく映像群のほうが、理解度問題の得点が有意に高かったと述べている。

Suvorov (2015) は、33 人の ESL 大学院生と学部生を対象に、実際に行われた大学の講義ビデオを視聴させ、その過程のアイムーブメントを記録した。提示する素材には、文脈情報と内容情報を含む内容ビデオと、文脈情報のみを含む文脈ビデオの 2 種類であった。その結果、情報量がより多いと思われる内容ビデオは、文脈ビデオより、「参加者が秒ごとにビデオを見る回数」と「参加者がビデオを凝視する総時間」に有

表 4 「テキストの種類から見た視覚情報の使用」に関する SLA の実証研究

著者	素材	提示手法			参加者	測定方法	主な結果
		音声のみ	静止画	映像			
Ginther (2002)	生活会話・学術討論・講義 (英語)	✓	✓	—	ESL/160 名	多肢選択問題	講義/図表条件のみ低い程度の促進が認められ、講義/文脈情報条件は、講義/音声のみ条件より得点が僅かに下回った。
Wagner (2010)	生活会話・講義 (英語)	✓	—	✓	ESL/202 名	多肢選択問題及び内容記述問題	生活会話、講義ともに視覚を与えた群のほうが得点が高かった。
Suvorov (2015)	実際の大学講義 (英語)	—	—	✓	大学院生及び学部生の ESL/33 名	アイムーブメントと多肢選択問題	秒ごとの凝視回数及び総凝視時間において文脈ビデオと内容ビデオの間に有意な差が見られたが、理解度問題の得点との相関が認められなかった。

意な差が見られた。しかし、両群におけるこれらのアイムーブメントの数値と視聴後の理解度問題の得点との間に相関が見られなかったと報告している。Suvorov (2015) は図表の補助の有無により、参加者の内容理解の得点に差がなかったことから、それまでの先行研究で扱われてきた「文脈か内容か」という視覚情報の分類を再検討し、新たな分類を探る必要性を指摘した。

日常会話と学術講義という異なるテキストの理解への視覚情報の影響に関して、Ginther (2002) と Wagner (2010) の結果は一見相反するようだが、Sueyoshi & Hardison (2005) の考察と合わせれば、学習者の滞在期間の長短が視覚情報の使用にも影響していたと推測できる。滞在期間が平均2か月以下の Ginther (2002) の参加者にとって、視覚情報による文脈の手がかりと音声情報の補助は、学術講義より日常会話の意味解釈を促進したと言えよう。滞在期間が短いと、音声情報の補助のための口形と音声のマッチングもまだ完全に習得されていない可能性があるため、テキストの困難度と情報量が上がれば、内容促進どころか、逆に視覚情報の存在が意味を理解するのを阻害するものとなりうる。

この三つの研究では、異なるテキストの種類における視覚情報の影響の違いについて分析を試みたが、それぞれのテキストが提供する視覚情報の種類と量の変動的であれば、学習者の理解に作用した視覚情報の影響についての説明が困難となる。また、3.1.でも触れたように、再認か再生かという参加者の理解を測定する方法によっても得られる結果が変わる可能性があるため、習熟度と学習背景が異なる学習者における、異なるテキストに対する視覚情報の有用度を分析するには、さらに要因を明確に絞り、慎重に研究を進めていくべきである。

3.4. 先行研究の問題点

3.4.1. 視覚情報の分類の妥当性

Suvorov (2015) が指摘したように、先行研究で扱われている視覚情報の種類が様々であり、分類の妥当性が検討されたことはなかった。全体を一つの大きな文脈として捉えるものもあれば、文脈（登場人物・環境）と内容（図表等）との二種類に分けるものもあった（表5参照）。特に大規模語学試験の聴解問題に視覚情報を取り入れる妥当性について検討した Wagner (2010) と Suvorov (2015) では文脈と内容の二分類を採用した。しかし、学習者の実際の使用状況を質的に分析してみると、文脈（登場人物・環境）のほかに、ハンドジェスチャー、表情を含むボディーランゲージという

表5 ESLを対象とした先行研究で扱われた視覚情報の分類及びテキストの種類

視覚情報の分類 テキストの種類	静止画 vs. 動画	<動画> 顔 vs. ジェスチャー	<動画> 文脈 vs. 内容
日常会話 (傍聴者として)	—	—	Wagner 2008 ; 2010
学術講義 (当事者として)	Ockey 2007	Sueyoshi & Hardison 2005	Wagner 2010 Suvorov 2015

視覚情報が使用されていることが分かった (Ockey, 2007 ; Wagner, 2008) 。そのような情報をさらに細かく分類し、顔 (主に口の動き) とジェスチャーに注目して調査を行った Sueyoshi & Hardison (2005) は、習熟度によってそれらの視覚情報の使用が異なることを証明した。

大型語学試験に動画を取り入れるかどうかという論争に関して、確かに視覚情報のどの部分がどのように働いているかが分からない限り、安易に試験に入れるのは危険である。しかし、語学試験の聴解問題作成を念頭に考えられた視覚情報の分類と、実際に学習者に使用されている視覚情報の分類が一致していないのであれば、目標言語使用域での言語運用を予測すべき言語テストの妥当性が問題となってくる。構成概念として試験に取り入れる前に、実際の生活の中での学習者の使用を見据えたうえで、聴解に関与する視覚情報の諸様相を確かめる必要がある。

3.4.2 学習者要因という視点

いままでの先行研究は視覚情報の種類やテキストの種類という視点から調査と考察を行ったが、その中で視覚情報の使用に大きな個人差が観察された研究は少なくない。参加者が示した視覚情報使用の回数と種類にばらつきが存在し、まったく画面を見ず聴覚情報だけに集中した参加者まで観察された。これらの現象を包括的に考察するには、視覚情報の種類だけでは足りない。例えば、先行研究でも言及された学習者の習熟度や、文化的影響、目標言語の国での滞在時間なども視覚情報の使用に影響する要因の一つだと考えられる。Sekiyama (1997) は、学習者の日本での滞在期間とマガーク効果 (McGurk & MacDonald, 1976) の相関が非常に高い ($r=.723$) ことから、外国での接触経験が口形による調音情報の使用に関与しており、滞在時間が長くなると口形による調音情報の使用も効率的になると考察している。さらに、いままで先行研究で扱われた参加者は殆どが目標言語の国で学ぶ学習者であり、教室外の豊富な言語的刺激と対面コミュニケーションの必要性があるために、母国で学ぶ学習者より視覚情報の使用が効率的であることも考えられる。このように、外からの情報という外的要因のみならず、「学習者」という視点に立ち返り、彼らの中に存在する内的要因に着目した研究が俟たれる。

3.4.3 データ収集と分析の方法

表 2~4 で示した先行研究の殆どは、聴解時の学習者の理解を聴解テストの点数に基づいて量的に比較したものである。それに対し、Ockey (2007) と Wagner (2008) は学習者がモニターに注目する時間や内省などで質的に分析した。Suvorov (2015) はこれらの研究結果の解釈が困難であることから、アイムーブメントという手法を取り入れ、より客観性の高いデータを通し、聴解能力の構成概念として視覚情報の妥当性を検証した。しかし、Suvorov (2015) のデータからは、学習者がどのように見ているかという行動における相違点は見られたものの、その相違点がどのように学習者の内容理解に影響したかという分析はなされていない。また、Sueyoshi & Hardison (2005)

で見られた視覚情報使用に関する学習者の意識と実際の運用状況のズレは、無意識に使用された視覚情報を引き出すことに対するプロトコルやアンケートの限界を示している。今後、学習者の聴解における視覚情報の使用意識と使用実態を研究するにあたって、反応時間の測定や母語翻訳課題など新たな方法を検討し、より客観的にデータ収集と分析を行うべきであろう。

3.4.4 対面聴解に関する第二言語習得研究

聴解における視覚情報に関する第二言語習得研究が扱うコミュニケーションの場面は、基本的に対面聴解を前提としている。対面聴解はさらに、相手とのインターアクションが可能かどうかにより、一方向性と双方向性に分かれる（表6参照）。一般に聴解研究は対面であれ非対面であれ、一方向性聴解に集中していた。また、雑談など対面であつた双方向性聴解を対象とした研究でも、参加者が当事者ではなく傍聴者として聞くものが多く、双方向性の特徴である「相手とのインタラクション」を考慮した研究は非常に限られている。

Rost & Ross (1991) や Vandergrift (1997)、横山 (2004a) などは、学習者が対面聴解において、意味交渉に使用した明確化の質問と、意味交渉に使用したその他のストラテジー、理解を構築していく過程と習熟度との関係などについて実証的研究を行った。これらの研究は双方向性聴解におけるインタラクションに着目し、それが意味交渉していく上での役割を考察していたが、対面聴解に伴う視覚情報の存在が学習者の持つ認知資源の再分配への影響には目を向けなかった。

Vandergrift (2004 ; 2007) が指摘したように、あらゆる情報に注意を向けつつ、聞いた内容の即時の処理とそれに基づいた適切な応答が強いられている学習者の立場から考えると、インタラクションができる双方向性聴解のほうがより認知的負荷が大きい。会話に参加する当事者として意味交渉をしたり応答したりする前に、話し手の発話内容をできるだけ正確に聞き取り、要点を把握することも重要である。双方向性の対面聴解となれば、常に音声情報に伴う視覚情報が存在し、両方の即時的な処理に追われた学習者の認知的負荷はさらに大きなものとなると推測できる。このように、当

表6 視覚情報を伴う「聞く」行動例の分類

一方向・双方向 対面・非対面	一方向	双方向
対面	<ul style="list-style-type: none"> ・講義 ・プレゼンテーション ・スピーチ ・落語等 	<ul style="list-style-type: none"> ・雑談 ・店員とのやりとり ・スカイプ、テレビ電話での会話
非対面	<ul style="list-style-type: none"> ・テレビ（インタビュー、ドラマの会話、ドキュメンタリー、ニュース等） ・映画 	<ul style="list-style-type: none"> ・電話（宅急便の不在票、ピザの広告等を見ながらの問い合わせや注文）

*奥野 (2016) に基づき、筆者が作成したものである。

事者としての双方向性の対面聴解と、教室で習う「傍聴者として聞くだけでよい」聴解とは質的に大きく異なっている。果して実際の双方向性の対面聴解で、大きな認知的負荷を強いられている学習者はどのように発話を理解しているのか。今後、双方向性の対面聴解という観点から学習者の「聞く」行動についての検討も必要となろう。

4. まとめ及び今後の展望

いままで聴解は、言語的知識（音韻、語彙、統語など）と非言語的知識（文脈、一般知識など）を使用し、知覚した音声情報を受動的に分析・解釈していくことだと認識されているが、Rost (2011) は聴解を「receptive (話された内容を確実に受ける)」、「constructive (意味の構築と叙述)」、「collaborative (意味交渉し、反応する)」と「transformative (掛り合いと共感を通し意味を創造する)」の四段階と定義している。相手が見えない、相手の反応を確かめられない聴解から、相手が存在し、お互い相互作用のできる、より真正性の高い聴解へと移行するには、いままで欠けていた視覚情報を取り入れる必要がある。しかし、聴解に視覚情報という新たな要素が加わると、会話か講義か、当事者として聞くか傍聴者として聞くか、学習者の内容理解に関与している要因が非常に多くて複雑である。このような視覚情報を教室指導や聴解評価に取り入れる前に、まず目標言語使用域での言語運用実態についての解明を急がねばならない。それと同時に、聴解指導に関して、あらゆる場면을非対面聴解として練習する代わりに、学習者に聞く目的と状況を与え、普段なら目に映る視覚情報が見られるようにし、彼らをより自然な文脈に置くことが必要である (Lynch, 1998; 松崎, 2008; 奥野, 2016)。「幅広い場面において自然なスピードの、まとまりのある会話やニュース、講義を聞いて、話の流れや内容、登場人物の関係や内容の論理構成などを詳細に理解したり、要旨を把握したりすることができる。」という日本語能力試験 N1 で求められる聴解力を育てるのに、どんな教材や指導が適切だろうか。特に、小柳 (2016) が指摘しているように、学習環境が母国の教室のみの学習者に対し、目標言語の国での対面聴解に必要な能力を事前に練習し、自然習得環境に効率よく橋渡しできるような教室指導が重要である。近年、学習者を当事者として話しかける映像やインタラクティブな練習を取り入れた聴解教材²も開発されているが、教室指导向けの聴解教材というより、教室指導後のサポートサイト的な位置づけとして強く認識されている。

今後、1) 学習者の実生活での使用を見据えた視覚情報の分類の再検討、2) 滞在期間や習熟度など新たな分析視点の導入、3) 反応時間や母語翻訳課題などより客観的にかつ包括的なデータ収集、4) 対面聴解で会話の当事者として聞く学習者の行動の特徴、という先行研究の問題点を踏まえ、対面聴解における視覚情報のどの部分がどう学習者の理解に働きかけるかについて究明する第二言語習得研究が期待される。

² 国際交流基金が作成した『まるごと 日本のことばと文化』教科書のサポートサイト「まるごと+」のことである。

引用文献

- 呉佳穎 (2016) 説明文の母語翻訳課題から見た対面聴解と非対面聴解の違い 『日本語教育方法研究会誌』 22 (3) ,14-15
- 奥野由紀子 (2016) 「聞く」プロフィシエンシーを高める練習方法の転換 —学習者の会話事例からの考察— 『日語教学研究』,252-274
- 小柳かおる (2016) 教室指導は習得にどんな影響を与えるか —混合環境 (自然習得環境+教室習得環境) の第二言語習得研究— 『現場を支える日本語教育研究—学ぶ・教える・評価する—NINJAL 国際シンポジウム予稿集』,68-71
- 島田めぐみ (2010) 自己評価 Can-do statements に関する一考察:客観テストとの比較を通して 『東京学芸大学紀要. 総合教育科学系』 61 (2) ,267-277
- 積山薫 (2014) 「第2章 音声情報の視聴覚統合処理」 『視聴覚融合の科学』 コロナ社,31-61
- 竹原卓真 野村理朗編著 (2004) 「第3章 顔と知覚:運動情報」 『「顔」研究の最前線』 北大路書房,40-55
- 福田倫子 (2004) 日本語教育における聴解研究の展望 『広島大学日本語教育研究』 (14) ,73-79
- ボイクマン聡子 (2009) 「聞いて話す」プロフィシエンシーとその養成のための教室活動 『プロフィシエンシーと日本語教育』,189-219
- 松崎寛 (2008) 聞く教育 (話し言葉の日本語) - (話し言葉と日本語教育) 『日本語学』 27 (5) ,229-237
- 横山紀子 (2004a) 対面聴解における理解過程の分析:理解構築の「流れ」に注目して 第28回日本言語文化学会 発表要旨 『言語文化と日本語教育』 (28) ,119-122
- 横山紀子 (2004b) 第2言語における聴解ストラテジー研究:概観と今後の展望. 『言語文化と日本語教育.増刊特集号,第二言語習得・教育の研究最前線』,184-201
- Baltova (1994) The Impact of Video on the Comprehension Skills of Core French Students. *Canadian Modern Language Review*, 50 (3) ,507-31.
- Buck (2001) *Assessing listening*. Cambridge University Press.
- Coniam (2001) The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: A case study. *System*, 29 (1) ,1-14.
- Ginther (2002) Context and content visuals and performance on listening comprehension stimuli. *Language Testing*, 19(2), 133-167.
- Kellerman (1990) Lip Service: The Contribution of the Visual Modality to Speech Perception and its Relevance to the Teaching and Testing of Foreign Language Listening Comprehension1. *Applied Linguistics*, 11 (3) ,272-280.
- Kellerman (1992) 'I see what you mean': The role of kinesic behaviour in listening and implications for foreign and second language learning. *Applied linguistics*, 13 (3) ,239-258.

- Lewkowicz & Hansen-Tift (2012) Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*, 109(5), 1431-1436.
- Lynch (1998) Theoretical perspectives on listening. *Annual Review of Applied Linguistics*, 18, 3-19.
- Lynch (2011) Academic listening in the 21st century: Reviewing a decade of research. *Journal of English for Academic Purposes*, 10 (2) , 79-88.
- McGurk & MacDonald (1976) Hearing lips and seeing voices. *Nature*, 264, 746-748.
- Ockey (2007) Construct implications of including still image or video in computer-based listening tests. *Language Testing*, 24 (4) , 517-537.
- Progosh (1996) Using video for listening assessment: Opinions of test-takers. *TESL Canada Journal*, 14 (1) , 34-44.
- Rost & Ross (1991) Learner Use of Strategies in Interaction: Typology and Teachability. *Language learning*, 41 (2) , 235-268.
- Rost (2011) Teaching and researching: Listening. Routledge.
- Rowley-Jolivet (2002) Visual discourse in scientific conference papers A genre-based study. *English for specific purposes*, 21 (1) , 19-40.
- Sekiyama (1997) Cultural and linguistic factors in audiovisual speech processing: The McGurk effect in Chinese subjects. *Perception & Psychophysics*, 59 (1) , 73-80.
- Sueyoshi & Hardison (2005) The role of gestures and facial cues in second language listening comprehension. *Language Learning*, 55 (4) , 661-699.
- Suvorov (2015) The use of eye tracking in research on video-based second language (L2) listening assessment: A comparison of context videos and content videos. *Language Testing*, 32 (4) , 463-483.
- Vandergrift (1997) The Cinderella of communication strategies: Reception strategies in interactive listening. *The Modern Language Journal*, 81 (4) , 494-505.
- Vandergrift (2004) Listening to Learn or Learning to Listen? *Annual Review of Applied Linguistics*, 24, 3-25.
- Vandergrift (2007) Recent developments in second and foreign language listening comprehension research. *Language teaching*, 40(03), 191-210.
- Wagner (2008) Video listening tests: what are they measuring? *Language Assessment Quarterly*, 5 (3) , 218-243.
- Wagner (2010) The effect of the use of video texts on ESL listening test-taker performance. *Language Testing*, 27 (4) , 493-513.

(うー ちゃーいん・首都大学東京大学院博士後期課程)