

視聴覚情報の非同期と視聴覚統合：洋画の吹き替えにあまり違和感を覚えないのはなぜか？

望月 要 大西 仁
帝京大学 メディア教育開発センター

映像と音声の呈示タイミングがずれると視聴者に違和感を与える。この際特徴的なのは、映像の遅れに関しては小さなずれでも違和感が生じるのに対して、音声の遅れに関しては比較的大きなずれでもあまり違和感が生じないことである。

これに似た現象として、フラッシュする視覚刺激とパルス音を連続して呈示し、被験者にそれらが同時に発生したか否かを判断させると、光のほうやや先に呈示される時に最も高い確率で同時と判断されるということが知られている。ヒトの顔と言葉のような複雑な刺激でも、フラッシュする視覚刺激とパルス音のような単純な刺激でも似た現象が起こるが、視聴覚情報の同期外れに関して、視覚情報の遅れに敏感で聴覚情報の遅れに鈍感なのは、一般的に成り立つのだろうか。

本研究は、ヒトの顔のような複雑な刺激呈示時の視聴覚統合に関し、非同時性の検出と違和感の発生要因を明らかにすることを目的としている。今回は、話者の言語が与える影響について予備的検討を行った。

実験 1

方法

参加者： 成人男女 8 名が実験に参加した。

材料： 素材は、日本語の落語、英語のニュース、ロシア語のニュース、韓国語のニュース、韓国語のドラマのビデオクリップを用いた。各ビデオクリップは 10 s の長さで、同期したもの、映像が 200 ms 遅延したもの、音声 が 200 ms 遅延したものが用いられた。映像の大きさは 640×480 画素であった。

日本語の落語は東アジア系黄色人が演じた。英語、ロシア語、韓国語のニュースキャスターは、各々黒人、白人、東アジア系黄色人であった。

手続き： 刺激は、プロジェクタおよびスピーカー、ないしは 12 インチ液晶モニターおよびヘッドホンで呈示された。参加者は、各刺激が呈示される度に映像と音声のずれおよびずれに伴う違和感に関する質問紙に回答した。質問項目は、映像と音声のずれの検知については、(1) ずれていなかった / わからなかった、(2) 映像が遅れていた気がする、(3) 映像が遅れていた、(4) 音声が遅れていた気がする、(5) 音声が遅れていた、の 5 件法で回答した。違和感については、(1) な

かった、(2) あった、(3) (他に比べて) 特に大きくあった、の 3 件法で回答した。各刺激の呈示は 1 回のみであった。

結果

ずれの検知については (2)~(5) をまとめて「ずれ検知」とした。違和感については (2) と (3) をまとめて「違和感あり」とした。表 1 に各刺激に対する反応確率を示す。落語刺激に対しては、違和感是一般に知られている傾向と一致した。また、ずれの検知に関しても違和感と同様のパターンを示した。英語およびロシア語ニュース刺激に対しては、映像遅延に対する違和感が小さく、音声遅延に対する違和感が大きいという、落語とは逆のパターンを示した。また、ずれの検知に関しても音声遅延に対する検知率が高く、ロシア語ニュースに関しては映像遅延に対する検知率が低かった。韓国語の刺激に対しては、ニュース刺激の音声遅延の検出率が高いことを除けば、落語刺激と同様のパターンを示した。参加者は、日本語と韓国語刺激ではほぼ同様のパターンを示したのに対して、英語とロシア語刺激では明らかに異なる反応パターンを示した。

実験 2

方法

参加者および材料： 成人男女 6 名が実験に参加した。実験に用いた素材は、(1) 日本語の落語中継、(2) 韓国語のドラマ、(3) ロシア語ニュース番組を録画したもので、タイムラグアジャスター (エレテックス EDS 3310) を用いて映像と音声の間に 1 フレーム単位で相対時間差を設けた。映像と音声の間の感覚間時間差 (SOA) は、日本語では -5 (音進み) から +10 フレーム (音遅れ) の 16 種類、外国語では ±11 の 23 種類を用いた。映像の大きさは 720×480 画素で、実験に用いた 17 型画面上で 91mm×140mm の大きさだった。日本語の落語と韓国語ドラマは東アジア系黄色人が演じ、ロシア語のニュースキャスターは白人であった。手続き：

映像刺激は 17 型 CRT モニタ (DELL P991) 上に呈示した。1000 ms の試行間間隔の後ピープ音と共に画面中央に凝視点を 500 ms 呈示し、その消失と同時に動画映像を 4800 ms 呈示した。映像と同時にヘッ

表 1: 各刺激に対する反応確率

日本語落語					
「ずれ」反応確率			「違和感」反応確率		
同期	映像遅延	音声遅延	同期	映像遅延	音声遅延
0.250	1.000	0.250	0.250	1.000	0.125
英語ニュース					
「ずれ」反応確率			「違和感」反応確率		
同期	映像遅延	音声遅延	同期	映像遅延	音声遅延
0.125	0.875	1.000	0.000	0.500	0.875
ロシア語ニュース					
「ずれ」反応確率			「違和感」反応確率		
同期	映像遅延	音声遅延	同期	映像遅延	音声遅延
0.500	0.125	0.875	0.000	0.000	0.750
韓国語ニュース					
「ずれ」反応確率			「違和感」反応確率		
同期	映像遅延	音声遅延	同期	映像遅延	音声遅延
0.125	1.000	0.750	0.000	1.000	0.714
韓国語ドラマ					
「ずれ」反応確率			「違和感」反応確率		
同期	映像遅延	音声遅延	同期	映像遅延	音声遅延
0.250	1.000	0.000	0.125	0.875	0.000

ドフォンアンプ (Audio-Technica AT-HA2) を通してヘッドフォンから音声刺激を呈示した。刺激呈示終了と同時に、映像と音声の関係に「違和感を感じたか」否かを2件法でキーボードを押して回答させた。以上を1試行として、各SOAを15回ずつ無作為な順に呈示して判断を求めた。試行数と実験所要時間は、日本語では240試行30分程、外国語では345試行40分程であった。反応時間には特に制限を設けなかった。

結果

4種類の刺激について、参加者が違和感を感じないと判断した確率を求め、これをSOAの関数として最小2乗法によりガウス関数を当てはめた。ガウス関数の平均(曲線の頂点に対応するSOA値)は、最も違和感を感じない映像と音声の時間差と考えることができる。この値は日本語の落語で4.27フレームであった(図1)。これは映像が音声と同じ量の時間差があっても、映像が先行する方が、違和感を感じる数が少ないことを示しており、この結果は、評定尺度法を用いた先行研究の知見と一致している。

韓国語とロシア語の分布平均値は、それぞれ-0.49と4.33フレームであった(図2)。外国語については、参加者数と素材の内容の点から今回の結果だけから明確な結論を述べることはできないが、ガウス分布の標

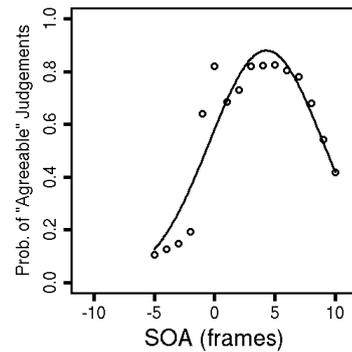


図 1: 日本語落語についての判断の結果。6名の参加者の平均データを心理物理学関数による回帰。

準偏差を見ると、日本語の6.65に対して、ロシア語が9.73、韓国語が9.14と大きな値を示している。これは、外国語では違和感を感じる閾値が高くなることを示唆するものである。またロシア語では、音声先行の方がより鋭敏に検出されるという非対称性が認められず、韓国語では非対称性は維持されるものの、音声10フレーム遅れても、全く違和感を感じていない、といった点も注目に値する。

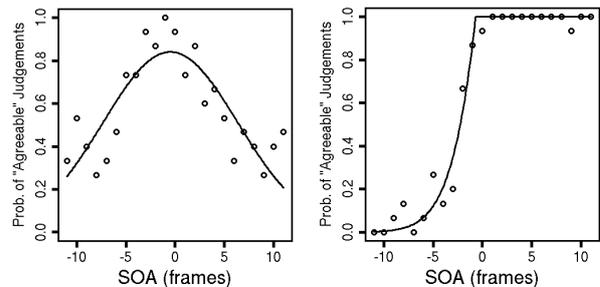


図 2: ロシア語(左)と韓国語(右)、各1名の参加者の判断結果。

考察

今回の予備的検討で、映像と音声がずれたときに生じる違和感は、日本語と外国語でパターンが異なった。外国語の中でも韓国語の場合は比較的日本語と似たパターンを示し、英語とロシア語は日本語とまったく異なるパターンを示した。洋画の吹替えにはそれほど違和感を覚えないのに対し、韓国ドラマの吹替えには違和感を覚えると言われているが、今回の結果はこれらと符合する。ここで問題となるのは、韓国語と日本語は音が似ており、また韓国人と日本人は顔が似ていることで、言語の音響と顔のどちらの類似性が効いているかということである。この問題については、言語と顔を統制した課題により検討したい。