

視覚情報が音声の持つ感性情報に与える影響についての検討

(指導教員 世木 秀明 助教授)

世木研究室 9710083 豊田 大輔

1.はじめに

音声には様々な感情を含ませることができ、他者との日常的なコミュニケーションにおいて感情表現は重要な役割を果たしている。近年、コンピュータのヒューマンインターフェイスの高度化のために、音声の感性情報に関する研究が盛んに進められている。

本研究では、このような研究の一環として視覚情報が音声の持つ感性情報に与える影響について様々な感情を表す視覚刺激を用いた聴取実験により検討を行うことを目的とした。

2.聴取実験

・発話資料

発音の正常な関東方言の成人男性話者 1 名が『平静』、『怒り』、『悲しみ』、『驚き』および、『喜び』の感情を込めて 5 回ずつ発話した単音「あ」25 音声と単語「どうして」25 音声を用意した。各音声 2 回ずつランダムに、聴力が正常な成人男女 17 名に呈示し、5 つの感情のうち、いずれかを強制選択させ、発話者が意図した発話と最も良く一致した音声と最も一致度が悪かった音声を聴取実験に使用する感情音声とした。これらの音声の音声パワー、基本周波数の変化および、持続時間を測定した結果、先行研究で報告されている音響パラメータとほぼ同様の傾向が見られた。

・視覚刺激

単純な線画による顔の絵を元に、眉毛の傾きと位置、口唇形状を様々に変化させて作成した 16 種類の顔画像を用意した。各画像 3 回ずつランダムに、視力が正常な成人男女 20 名に呈示し、『平静』、『怒り』、『悲しみ』、『驚き』および、『喜び』の 5 つの感情のうち、どれに感じるかを強制選択させ、最もそれぞれの感情として感じる顔画像を聴取実験に使用する視覚刺激画像とした。図 1 に聴取実験に使用した 5 種類の感情表現の顔画像を示す。

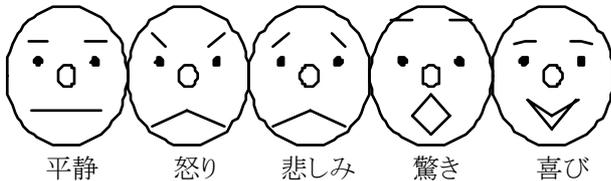


図 1. 視覚刺激として用いた顔画像

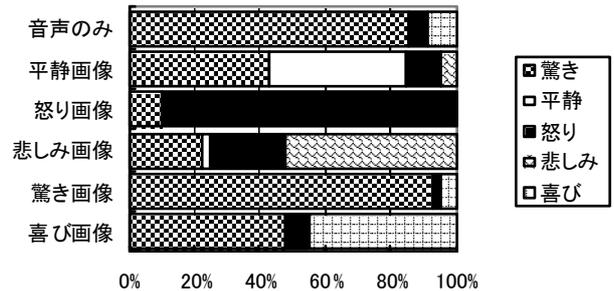
・実験方法

聴取実験は、5 種類の顔画像に対して全ての感情音声を組み合わせた刺激を用いて行った。刺激数は、単音「あ」で 50 種類、単語「どうして」で 50 種類である。被験者は聴力、視力共に正常な 20 代成人男女 20 名である。刺激呈示条件は、上述の単音、単語 50 種類の刺激を 2 回ずつランダムに 4 秒間隔で呈示し『平静』、『怒り』、『悲しみ』、『驚き』および、『喜び』の 5 つの感情のうち、どれに感じるかを強制選択させた。音声はヘッドホ

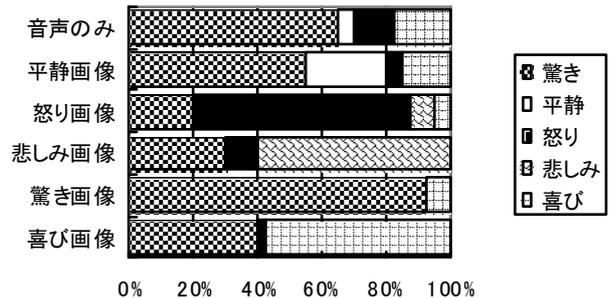
ンで画像は被験者前方 30cm 前に配置した 17 インチディスプレイにより呈示した。また、被験者には、顔画像を注目して音声を聞くように教示した。

3. 実験結果と考察

図 2 に聴取実験の代表的な結果を棒グラフで示す。図 2 a)は、呈示音声として発話した意図と一致率が高かった音声を使用した場合であり、図 2 b)は、一致率が低かった音声を使用した場合である。



a) 音声のみを呈示した場合に発話した意図と一致率が高かった音声を使用した聴取実験結果



b) 音声のみを呈示した場合に発話した意図と一致率が低かった音声を使用した聴取実験結果

図 2. 単語「どうして」の驚き音声に種類の感情顔画像を組み合わせたときの被験者の回答の分布

図 2 より、感情音声に一致した感情の画像を呈示すると、発話者が意図した感情に受け取られる割合は高くなる。しかし、その割合は音声のみを呈示した場合の一致率の高低にほとんど影響を受けていない。一方、音声の持つ感情と異なった感情をもつ顔画像を呈示した場合、顔画像の持つ感情に受け取られる割合が大きくなる傾向が観察された。この結果を詳細に検討すると、特に『怒り』および、『悲しみ』音声は他の感情音声よりも顔画像の影響を受けにくい傾向があることが観察された。これは、音声の持つ感情表現に関する音響的特徴が影響しているのではないかと考えられた。また、単音「あ」でも同様であった。

これらのことから、視覚刺激が音声の持つ感性情報の聴取に大きく影響を与えること、音響的に顕著な特徴が見られる音声の場合は、視覚刺激の影響が比較的小さいことが観察された。また、意図した感情を聞き手に正しく伝えるためには、感情に対応した視覚刺激を呈示することが非常に有用であると考えられた。