

顔と声の相互作用によるマガーク様効果の検討

092G012 柿本幸里・092G044 藤田信子・092G062 渡邊健太

問題と目的

人間がコミュニケーションを円滑に行う為には人物の識別・同定をおこなう必要があるが、我々はこれらを無意識的に、あるいは意識的に行っている。また、他人と接する際に、相手の顔表情と声は識別・同定の手がかりだけではなく、相手の感情を推測する手がかりとしても用いられ、さらには性別、年齢、職業など様々な個人属性の推測や判断にも用いられていることなどが指摘されている。

McGurk & MacDonald (1976) によって最初に報告され、後に Harry McGurk の名前をとってマガーク効果と呼ばれるようになった現象は、視覚 (画像) 情報と聴覚 (音声) 情報とを矛盾させた場合に、情報の統合が生じて元の情報とは異なった音声として知覚されるというものであり、聴覚に対する視覚優位性を示す現象として有名である。

彼らは、「ga, ga・・・」と発話している人物の映像に重ねて、「ba, ba・・・」という音声をスピーカーから流して、それがどのように聞こえるかという参加者の反応を調べた。

実験の結果、参加者は映像上の唇の動きにだまされて「ga」でも「ba」でもなく、「da, da・・・」という音を聞いたように錯覚を起こすという現象が生じた。この現象が示唆しているのは、音声情報と映像情報が同時に提示された場合、音声の印象が映像情報にひきずられてしまうということである。

本研究では、人物の映像と音声の間でマガーク効果のような視聴覚情報の統合効果が生じるかどうかを検討したい。また、視覚的な情報が聴覚的な情報に対する知覚をどれだけ左右するかを心理物理学的に実証するために、河原 (1998) によって開発された STRAIGHT を応用する。STRAIGHT は、音声知覚の研究用に開発されたもので、音声のピッチや周波数特性、発話タイミングなどを操作するためのツールである。

方法

参加者

実験材料の声の人物と交流のある大学生 20 名 (男女 10 名ずつ) が本研究の実験に参加した。

刺激

参加者にとって既知の 2 人の人物 (K, F) の発話 (“もみじまんじゅう”) をビデオ採録した。刺激としては、K あるいは F の映像に重ねて、K+F の混合率が 0%~100% の音声流れるビデオクリップを作成した。

手続き

参加者はタッチパネル式のノートパソコンの正面に着席し、実験の概要の説明を受けた。次に、ヘッドホンを装着し、映像と音声が K100%, F100% の刺激を納得いくまで何度も聞いてもらった。参加者が刺激の確認を終えると、プログラムを起動し、操作方法を説明し、実験を行った。実験中は、画面中央に K, F いずれかの映像が表示された。実験では、K, F の中間音声ヘッドホンを通して再生されるが、参加者はこれらの刺激を見聞きして、K あるいは F のどちらの声に聞こえたかを判断して、パソコン画面右側に表示されている K, F の顔写真をタッチすることで回答するよう教示された。参加者が刺激を聞き逃したり、もう一度聞きたいと思った場合は、刺激が表示される画面中央をタッチすることで何度でも聞き直しをすることが出来るようにした。

結果と考察

実験の結果、参加者によって K の声と判断された割合を従属変数とし、映像の人物 (2 水準: K, F) × K の音声割合 (11 水準: 0%~100% の 11 段階) の 2 要因分散分析を行った。

その結果、まず映像の人物の主効果が有意で ($F(1,19) = 13.44, p < .005$), F の映像が流れていたときの方が、K の映像が流れていたときよりも全体的に音声 K のものと判断される割合が高かったことがわかった。また、K の音声割合の主効果も有意で ($F(10,190) = 354.48, p < .0001$), K の音声割合が増えるにつれて有意に K の声と判断される割合が増加していた。さらに、映像の人物 × K の音声割合の交互作用も有意であった ($F(10,190) = 5.71, p < .0001$)。そこで単純主効果の下位検定を行ったところ、映像の人物の効果は、K の音声割合が 30% のとき ($F(1,209) = 4.30, p < .05$), 50% のとき ($F(1,209) = 57.75, p < .0001$), 60% のとき

($F(1,209) = 11.93, p < .001$) に有意であることがわかった。

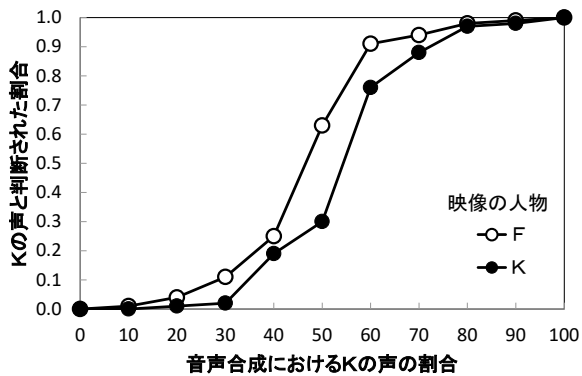


図1 全体の実験結果

本実験の結果, 二人の中間音声の流れたときには, 参加者はその声を映像の人物の音声と判断するのではなく, むしろ, もう一方の人物の音声と認識する傾向が強いことがわかった。この結果は, 人物の映像と音声の間にもマガーク様効果が存在するのではないかという本研究の仮説とは異なる方向の結果である。本研究の参加者はすべて, 刺激となった二人の人物の友人であったため, 日常のコミュニケーションで形成された人物と音声の結合したイメージが, 合成された中間音声に対して, 何らかの違和感をもたらし, それが結果に影響したのかもしれない。

そこで, 刺激人物との日常の交友の程度を3段階(交友頻度1:あまり合わない, 7人;交友頻度2:時々会う, 8人;交友頻度3:よく会う, 5人)に分け, それを要因に加えて, 交友の程度×映像の人物×Kの音声割合の3要因分散分析を行った。その結果, 交友の程度×映像の人物×Kの音声割合の2次の交互作用が有意となった ($F(20,170) = 7.06, p < .05$)。単純交互作用の低位検定を行ったところ, 先の2要因分散分析で認められた映像の人物×Kの音声割合の交互作用は, 交友頻度3 (よく会う友人のグループ) に対してのみ有意であり ($F(10,170) = 7.95, p < .0001$), 交友頻度2 (時々会う友人) ($F(10,170) = 1.76, ns$), 交友頻度1 (あまり合わない友人) ($F(10,170) = 0.79, ns$) のグループでは認められないことがわかった。

以上の結果から, 人物とその音声の間には, 「逆マガーク」様の効果が存在することがわかった。また, その効果には, 対象人物との日常のかかわりによって形成されるであろう人物と音声の結合したイメージが影響しているらしいことがわかった。この

ことから, 人は他者との交流が深くなるにつれて, その人物を声だけ, 顔だけではなく, 声と顔を一つのセットとして, 人物を判断するようになるのではないかと考えられる。

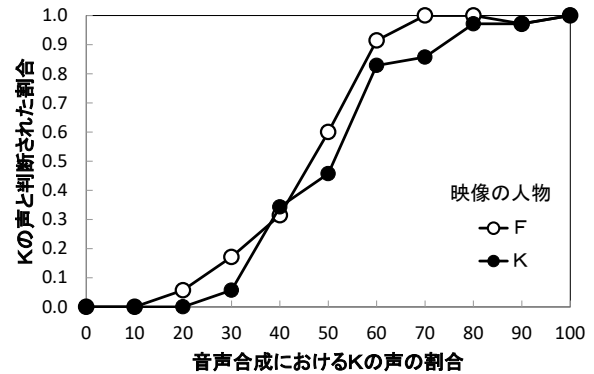


図2 あまり合わない友人による判断結果

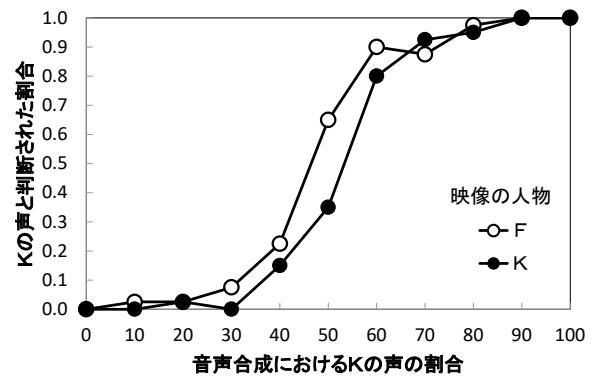


図3 時々会う友人による判断結果

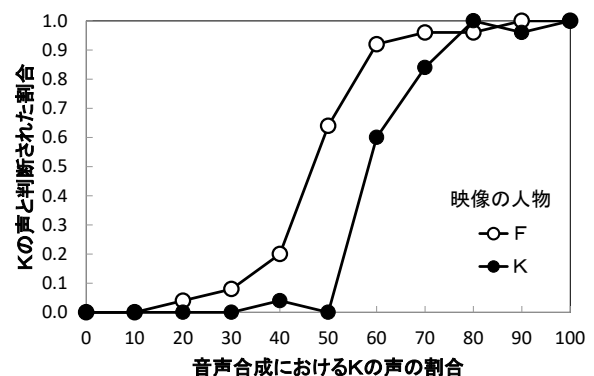


図4 よく会う友人による判断結果

引用文献

- McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, **264**, 746-748.
- 河原英紀 (1998). 聴覚の情景分析が生み出した高品質 VOCODER : STRAIGHT 日本音響学会誌, **54**, 521-526