

科学研究費助成事業 研究成果報告書

平成 30 年 6 月 19 日現在

機関番号：21301

研究種目：基盤研究(C) (一般)

研究期間：2014～2017

課題番号：26330311

研究課題名(和文) オフスクリーン音に注目したシーン認識における視聴覚統合過程の検討

研究課題名(英文) Study of audio-visual integration process revealed by eye-tracking experiment with the "off-screen" sound presentation.

研究代表者

茅原 拓朗 (Kayahara, Takuro)

宮城大学・事業構想学群(部)・教授

研究者番号：00345026

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：より包括的に視聴覚統合過程を検討するために、視野外に音源が存在する(オフスクリーン条件)場合を含む映像と音の様々な組み合わせにおいて被験者の視線の動きとシーンの意味解釈を実験的に検討したところ、映像につけられた音はシーンに向けられる視線の集中度合いを規定し、視線が向けられた視覚情報と音源の意味とが結びつくことでシーンの意味解釈が成立する過程が明らかとなった。また、その際の音の適度なオフスクリーン性がシーンの意味解釈の複雑さや豊かさに寄与している可能性が示唆された。

研究成果の概要(英文)：In this study, whether the sound induce subjects' gaze and affect meaning of the scene was examined when the subjects was presented with audio-visual stimulus, including the condition in which the sound source is outside of scene ("off-screen" sound condition). The data from eye-track recording revealed that the sound induce subjects' gaze and affect meaning of the scene. Interestingly, in off-screen sound condition, the sound didn't induce subjects' gaze to specific visual object in the scene, but, made meaning of the scene read by subjects more rich and complex.

研究分野：知覚・認知心理学

キーワード：シーン認識 視聴覚統合 視線計測 サウンドトラック オフスクリーン音

1. 研究開始当初の背景

人間は目の当たりをしているシーン内で何が存在し何が起きているかを即座に把握し、適切な行動をとることができるが、それがどのように達成されているかは未だによく分かっていない。特に、具体的な視覚オブジェクトで構成された実シーンでその認識がどのように行われているかの検討は視線計測などを指標によく途についたところ[1]である。また、実際のシーンは視覚のみならず使えるモダリティを総動員してマルチモーダルに認識されているはずであるが、実シーン認識におけるマルチモーダル過程は、単純な視聴覚情報の相互作用・統合過程すらまだほとんど研究されていないのが現状である。

一方、視聴覚のマルチモーダル研究では、視覚刺激は常に聴覚刺激の「音源」であるという前提で（音源となる視覚対象が常に視野内に「見えている」条件下で）主としてその時空間的特性が検討されてきた。しかし実際には、視覚には視野がありシーンは基本的に前方に限定されるのに対し、聴覚は全方位的に音を捉えることができるために、実際の環境では（その音が明確に意識にのぼるかどうかは別としても）視野内に視覚対象としては存在しない音源の音が聞こえている状況の方が頻度としては多いと考えられる。従って、これまでのようにシーン内に存在する視覚対象と音との関係を検討するだけでなく、シーン内に音源となる対象が存在しない場合の視聴覚統合についても併せて検討しなければ、実際のシーン認識における視聴覚統合プロセスの包括的な解明に至ることはできない。

2. 研究の目的

そこで本研究課題では、特に映画において“シーンの外側”を意味するオフスクリーンという用語を援用し、シーン外から聞こえてくる音をオフスクリーン音と名付けて、シーン内に音源がある場合のオンスクリーン音と区別・対比させながら視線計測を主な指標とした心理物理実験によりシーン認識における視聴覚統合過程を検討する。

3. 研究の方法

本研究課題では、実シーン認識における視覚と聴覚の統合過程に着目し、視線計測を主な指標として実験的な検討を行う。特に本課題では、音源となる視覚対象がシーン内に存在する（オンスクリーンと呼ぶ）場合だけでなく、それがシーン内に存在しない場合（オフスクリーンと呼ぶ）にも光をあて、視聴覚統合過程の包括的な理解を目指す。その際のアプローチとして主に次に掲げる2つの方法によって検討を進めた。

視覚には限定された視野があるのに対し、聴覚が全方位的であるということから考えれば、オフスクリーン音を聞く状況が存在すること自体は明らかである。しかし、それが実際の我々の認知や行動にどのように役立

てられているかについてはまだよく分かっていない。その一方で、映画などの映像制作においては、オフスクリーン音が次のシーンの展開を予期させたり現在のシーンの情動的意味づけを行うなど重要な役割を担ってきた[2]。そこで、1つめの方法として映画やアニメーションなど、実際の映像作品を用いた検討を行った。視聴覚作品におけるサウンドトラックをシーンの意味づけとつきあわせながら分析して仮説構築・実験パラメータの抽出を行うとともに、初期のトーキー（音声付き）・アニメーション作品を視聴している際に視線計測を行うことで検討を行う。

さらに2つめの方法として、オフスクリーン性を操作しながら音を付与した映像を視聴している際の視線計測を行い、かつそのシーンの意味解釈を問うことで、オフスクリーン音も含む音響と映像の統合過程を実験的に検討した。オフスクリーン性は1つめの実際の映像作品の分析の結果から、ある音の音源となる視覚オブジェクトが映像中に存在するかどうかによって定義することとした。

4. 研究成果

1つめの方法として実際の映画作品の分析を行った。文献等からある程度音響デザインの制作意図が明らかになっている映像作品を取り挙げ、映像に対する音の付与のされ方を詳細に検討した。その結果、オフスクリーン音は、音響制作において「アンビエンス」とも呼ばれている背景音・環境音と、なんらかの音源が特定できる音の大きく2つに大別できることが明らかになった。この結果を受けて、本研究課題ではとりあえず音源が特定できる音に着目して、音源となる視覚オブジェクトがシーン内に存在する場合を「オンスクリーン」状態、存在しない場合を「オフスクリーン」状態と定義してさらに検討を進める。

実際の映像作品を使った次の段階の分析として、初期のアニメーション作品を取り挙げ、ナイーブな（その作品を未見の）被験者に対して無音の場合と通常通り音が付与されている場合の2つの場合に視線と映像の解釈がどのようになるかを検討した。

初期アニメーションの音響デザインには、背景音・環境音がほとんど使われず、登場キャラクターのほとんど全ての動きに音が付与され音源が特定できる音だけでサウンドトラックが構成される、所謂「ミッキーマウシング」と呼ばれる特徴がある。そのような作品で音がある場合とない場合の効果を調べることで、本研究課題でオフスクリーン性の定義に用いる音源が特定できる音が視線とシーンの意味解釈に与える影響についての今後の検討の指針を得ることができる。

図1は、最初期のトーキー・アニメーション作品である『蒸気船ウィリー』の1シーンを音を鳴らさずに視聴した時のゲイズ・プロット（視線が向けられた場所を線で結びながら順番にプロットしたもの）である。画像中

央に位置するメインキャラクターがスプーンを使ってバケツやフライパンなどを打楽器に見立てて演奏するシーンであるが、被験者の視線は、メインキャラクターとシーンの意味解釈にはほとんど寄与していない手前の鳥のキャラクターに集中している。またシーンの意味解釈としても「バケツをたたいて騒いでいる」「鳥のキャラクターに何かを教えている」というように必ずしも演奏には結びつかないものとなった。

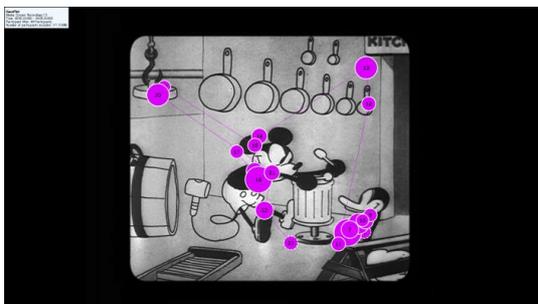


図1：『蒸気船ウィリー（音無し条件）』視聴時のゲイズ・プロット（視線が向けられた場所を線で結びながら順番にプロットしたもの。各円の大きさはその場所での視線の滞留時間の長さを示す）。

一方、図2は、同じ『蒸気船ウィリー』の1シーンにもととの音をつけて視聴した時のゲイズ・プロットである。ここではストーリーに寄与していない鳥のサブキャラクターには視線が向けられず、メインキャラクターの中でも特に演奏動作に視線が向けられ、さらにはメインキャラクターが演奏するバケツや壁に掛けられたフライパン等、ミッキーマウスが叩いて音を鳴らすものへ順番に視線が移動しているのが分かる。

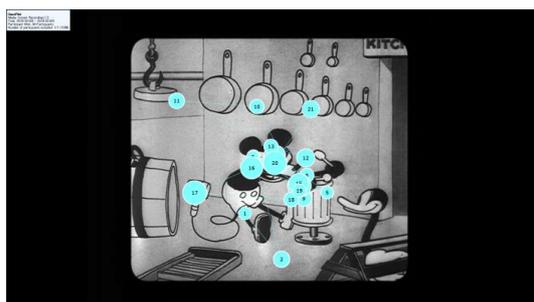


図2：『蒸気船ウィリー（音有り条件）』視聴時のゲイズ・プロット。

以上の結果から、シーン内に音源となる視覚オブジェクトが存在する場合の音（オンスクリーン音）の存在は、視線を音源となる視覚オブジェクトにむけて強く誘導すると同時に、シーンの意味解釈を可能にし、また方向付けることが分かった。

それらもうけて、2つめの方法として、映像作品ではない任意の画像に対してオンスクリーン音とオフスクリーンを付与したと

きの視線を計測し、また被験者に画像の意味解釈を尋ねる実験を行った。



図3：公園のシーンに犬の鳴き声をオンスクリーン条件で付与した際の視線のヒートマッププロット（画像中の場所毎に視線が向けられた頻度を等高線的に表現したもの。この画像では視線が集中している場所の背後に音源となる視覚オブジェクトとしての犬がいる）。

図3は、公園のシーンに犬の鳴き声をオンスクリーン条件で付与した際のヒートマッププロットである。視線が集中している場所の背後に犬がおり、提示時間を通じて視線はほぼ犬のみに集中していることが見てとれる。また映像の意味解釈としては、「公園で犬が鳴いている」のように視線が集中した視覚オブジェクト（この場合は犬）にのみ言及されるような単純なものとなった。



図4：住宅のリビングルームのシーンに食器を洗う音をオフスクリーン条件（シーンとの意味的連関有り）で付与した際の視線のヒートマッププロット。

図4は、住宅のリビングルームのシーンに食器を洗う音をオフスクリーン条件で付与した際のヒートマッププロットである。食器を洗うことに対応する音源はシーン中には存在しないことからオフスクリーン音ではあるものの、同じ住宅内で起こりうるアクティビティということでシーンとの意味的連関は有る条件として設定した。図4のヒートマップが示す視線の傾向は、図3のオンスクリーン条件ほど一つの視覚オブジェクトには集中しないものの、画面右手の食器が収まっている右側のキャビネットや、画面左手の部屋の奥まった部分（キッチンがあることが想定できる）など、音から想起される場所に

向けられていることが分かる。さらにシーンの意味解釈を被験者に尋ねると興味深いことに図3のオンスクリーン条件よりもストーリー性にとむより複雑な解釈が回答される傾向があることが見出された。



図5：街の交差点のシーンに野球場の打球音をオフスクリーン条件（シーンとの意味的連関無し）で付与した際の視線のヒートマッププロット。

図5は、街の交差点のシーンに野球場の打球音をオフスクリーン条件で付与した際のヒートマッププロットである。打球音に対応する音源はシーン中には存在しないことからオフスクリーン音であり、同時にシーンのおかれた文脈とも連関が無い条件として設定した。図5のヒートマップが示す視線の傾向は、図4のオフスクリーン条件（シーンとの意味的連関有り）よりもさらに分散的なものになり、ヒートマップから解釈可能な傾向は見られなかった。また、シーンの意味解釈も「よくわからない」など明瞭な回答が得られないことがほとんどであった。

以上のことから、映像につけられた音がシーンに向けられる視線を誘導し、視線が向けられた視覚情報と音源の意味とが結びつくことでシーンの意味解釈が成立する過程が明らかとなり、またその際の音の適度なオフスクリーン性がシーンの意味解釈の複雑さを生み出している可能性が示唆された。

【文献】

- [1] Hwang, A.D. et al. (2011) Semantic guidance of eye movement in real-world scene. *Vision research*, 51, 1192-1205
- [2] Bordwell, D. & Thompson, K. (2012) *Film art: an introduction*, McGraw Hill (New York)

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計0件）

〔学会発表〕（計3件）

①茅原拓朗、心理学におけるワークショップの古くて新しい可能性～19世紀知覚研究とエンタテインメントメディアの温故知新～、日

本基礎心理学会第36回大会（招待講演）、2017年12月2日、立命館大学（大阪府吹田市）

②Takuro Kayahara, "Sound track" brings new approaches of the psychological study of audio-visual interaction and a new track of outreach for psychology., 31st International Congress of Psychology (ICP2016, Invited talk), 27Jul.2016, Yokohama, Japan.

③茅原拓朗、映像を伝える音、日本基礎心理学会公開シンポジウム（招待講演）、2015年10月25日、慶應義塾大学（東京都港区）

6. 研究組織

(1) 研究代表者

茅原 拓朗 (KAYAHARA, Takuro)

宮城大学・事業構想学部・教授

研究者番号：00345026