

映像の分節的な知覚における経験的理解の影響

大山 潤爾

東京大学大学院工学系研究科

日本学術振興会特別研究員

jo@fennel.rcast.u-tokyo.ac.jp

東京大学先端科学技術研究センター

渡邊 克巳

独立行政法人産業技術総合研究所

科学技術振興機構

In this study, we examine whether contextual segmentation affects visual perception. Fodor and Bever (1965) discovered the temporal synchronization effect on speech perception of contextual segments in spoken language. We converted Fodor's auditory paradigm into a visual task by converting the auditory streams into visual streams, and the brief sounds into brief visual flashes. The detection of a similar temporal synchronization effect would strongly support the perception of a visual stream as the context of the visual event. We conducted a new experiment using movies with natural scenes and edited scene segments to examine whether a semantic movie segment in the action context had a temporal synchronization effect on flash timing. The results showed that the visual stream could be perceived within contextual segmentation.

Keywords: event perception, attention, visual cognition, temporal synchronization.

問題・目的

Fodor & Bever (1967)は、スピーチに挿入したクリック音が、スピーチの文脈的な文節に同期して知覚されることを報告している。彼らはこの研究で、音声言語が文脈的な文節単位に区切られて知覚されていると提唱している。我々は、この音声言語の実験手法を視覚課題に応用することで、映像が、音声言語と同様に、文脈的に分節化されて知覚されている可能性を調べた。映像呈示中のランダムなフレームで、白いバー刺激（フラッシュ）を呈示し、フラッシュの映像内容の文脈的な分節に対する同期効果を調べることで、映像文脈内容の知覚特性を調べることができると考えられる。

cd/m2)を呈示した。呈示時間が短い(1 frame = 40ms)ため、バー刺激はフラッシュのように知覚された。

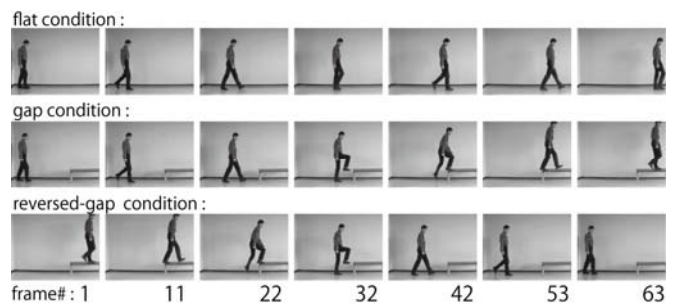


Figure 1. Sequence of three conditions of movies stimuli.

方法

実験協力者：実験には、視力または矯正視力が1.0以上の健常成人7名が参加した。実験は3条件あり、被験者は、3つの実験全てに参加する。3つの実験を行う順序は、被験者間で入れ替える事で、順序効果を相殺した。

実験環境：暗室内に、実験PC、22インチCRT モニタ(周波数100 Hz)、操作キーボード、椅子を設置し、実験を行った。椅子の位置と高さは、モニタの中心と眼の高さが平行で、モニタから眼までの距離が57 cmになるように、被験者ごとに調整した。

刺激：実験を通して、モニタ画面の中央には、常に注視点を提示した。実験は3つあり、それぞれの実験では、歩行、台上り、そして、台上り映像の逆再生映像 (Figure 1参照) を、モニタ中央に呈示した。映像のサイズは視野角 $4.0^\circ \times 3.0^\circ$ で、フレームレートは25 Hz (40 ms / frame)、フレーム数は63、よって、映像の長さは2520 ms (40 * 63 ms)であった。また、知覚の変化を検出するプローブとして、映像再生中の1つのフレームにおいて、映像の上下に白いバー刺激(80

手続き：実験は3つあり、それぞれの実験では、歩行、台上り、そして、台上り映像の逆再生映像を用いた。各試行は、被験者がスタートボタンを押すことで始まり、映像が再生される。被験者は注視点を固視しながら、映像を観察し、その間のランダムなフレームに呈示されるフラッシュ（バー刺激）のタイミングを覚える (Figure 2参照)。フラッシュの挿入されるフレームは、映像中のランダムな偶数番号のフレームであり、映像の開始から、400 ms (10 frame) から1680 ms (42 frame) の間の17箇所から選ばれる。

映像の呈示後400 msのブランクがあり、映像中のフレームのうち偶数番号のフレームの中からランダムな一つのフレームの画像が呈示される。被験者は、左右の方向キーを使って、呈示されたランダムな偶数番号のフレーム画像から、その前後の偶数番号のフレーム画像に切り替えていくことができる。被験者は、方向キーを操作し、映像全体から、フラッシュが呈示されたフレームを選び、決定ボタンで回答した。

3つの実験は、それぞれ6つのブロックからなり、各ブロックでは、フラッシュが、映像中のランダムな17箇所それぞれ1回ずつ呈示される(17試行)ため、

実験全体は306試行であった[3実験*6ブロック*17箇所=306試行].

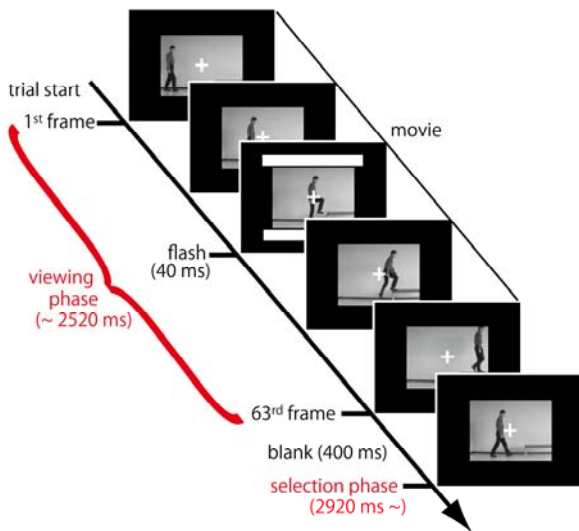


Figure 2. Experimental procedure of an gap condition task.

結果

Figure 3(A)は、各フレームでフラッシュが知覚された頻度を示す。実験の結果、台上りの順再生映像の、38番フレームと同時にフラッシュが知覚されたと回答する頻度は、実際の呈示頻度より高く、その前後のタイミングに呈示されたフラッシュの知覚頻度は、実際の呈示頻度より低いことがわかった。

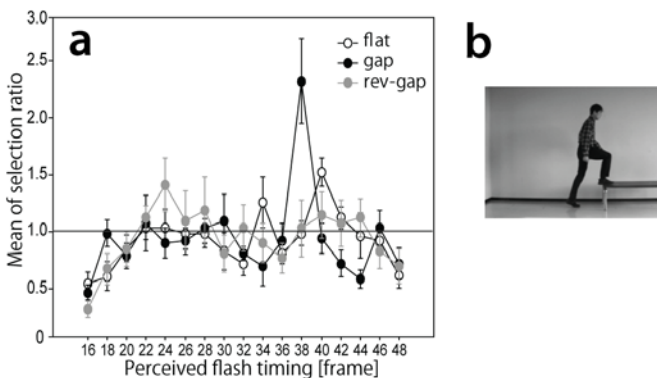


Figure 3. (A) Result of the distribution of a flash perception. (B) Scene figure of 38th frame of the gap condition movie.

二次元配置分散分析(An ANOVA with repeated measures)による解析の結果、映像の主効果は見られず($F[2, 12] = .44, p = .63$)、フレーム系列の主効果と交互作用に有意差が見られた(timing ($F[16, 96] = 5.34, p < .01$), interaction ($F[32, 192] = 2.43, p < .01$)). さらに、

多重比較による検定の結果、台上りの順再生映像の、38番フレームにおけるフラッシュの知覚頻度が有意に高いことが示された(Post-hoc Tukey's HSD tests; $p < .05$). フラッシュの知覚頻度が特異的に高い38番フレームは、台に足をのせた瞬間の画像であった(Figure 3(B)参照). 一方、歩行映像や、台上りの逆再生映像では、こうした知覚の変容は見られなかった。

考察

実験の結果、映像中の人の行動が変化する特定の場面にフラッシュが同期して知覚されることがわかった。映像を逆再生で呈示した場合は、同期効果は見られなかった。台上りの逆再生では、呈示されるフレームごとの画像は、順再生とまったく同じであり、台上りの順再生映像における38番目のフレーム画像は、台上りの逆再生映像中の26番目のフレーム画像と同じである。しかし、逆再生映像では、フラッシュが特定のフレーム画像に同期する現象は見られなかった。よって、知覚の変容は、映像の時系列的な前後関係、つまり文脈的分節によって引き起こされると示唆される。

結論

映像が、音声言語と同様に、文脈的に分節化されて知覚されている可能性について、新しい実験手法を導入して検討した。実験の結果、映像中の人の行動が変化する特定の場面にフラッシュが同期して知覚されることがわかった。映像を逆再生で呈示した場合は、同期効果は見られなかった。本研究は、映像が、学習や経験による理解に基づき、文脈的に分節化して知覚されている可能性を示唆する。

今後は、年齢や理解力の異なる被験者を対象に、フラッシュの分節同期効果を調べ、映像の文脈的理解が知覚に及ぼす影響とその学習発達段階を解明していく。本研究内容はOhyama & Watanabe (2010)に掲載された論文に基づいている。

Acknowledgement

本研究は、日本学術振興会及び文部科学省「はぐりく健康創造クラスター」の支援を受けて行われた。

引用文献

- Fodor, J.A., & Bever, T.G. (1965) The psychological reality of linguistic segment. *Journal of Verbal Learning and Verbal Behavior*, 4, 414-420.
- Ohyama, J., & Watanabe, K. (2010) Contextual segmentation in the visual stream produces a temporal synchronization effect on visual perception. *Psychologia*, 53(2), 125-137.