

# 意味共活性化による視聴覚プライミング Audio-visual priming effect by semantic co-activation

成川 額史<sup>†</sup>, 松香 敏彦<sup>‡</sup>  
Takashi Narikawa, Toshihiko Matsuka

<sup>†</sup>九州大学, <sup>‡</sup>千葉大学  
Kyushu University, Chiba University  
narikawa.takashi.846@s.kyushu-u.ac.jp, matsuka@chiba-u.jp

## 概要

プライミング効果は、先行して呈示された刺激（先行刺激）が、後に続いて呈示される刺激（後続刺激）の処理を促進する心理効果である。本研究において実施したプライミング課題では、先行する聴覚刺激と後続の視覚刺激との間に直接的な意味の関係がある場合、直接的な意味の関係がないが間接的に意味的な結びつきがある場合、意味的な関係がほとんどまたは全くない場合の3条件について比較を行った。その結果、先行刺激と後続刺激との間に直接的な関係がある場合のみならず、間接的な結びつきのみがある場合においても反応が一定促進されることが示唆された。

キーワード：意味的プライミング、視聴覚統合、注意

## 1. はじめに

プライミング効果は、先行刺激が後続刺激の処理を促進する心理効果である。特に、両刺激の間に意味的な関係がある場合にプライミング効果が生じることが示されており、これは意味的プライミング (semantic priming) とよばれる。意味的プライミングは非言語刺激やモダリティの異なる刺激どうしにおいても生じることが近年示されており (e.g., Chen & Spence, 2018; Manesse et al., 2020), とりわけ先行して呈示される聴覚刺激がそれに続いて呈示される視覚刺激の処理に与える影響について検討がなされてきた。

また、刺激同士の意味的な結びつきが処理に影響を与える現象の一つに、意味的競合 (semantic competition) がある。これは、ターゲットとなる視覚刺激に対応する聴覚刺激が呈示される時、ターゲット刺激と同時に呈示されていてかつ意味的に関係のある視覚刺激に注意が向きやすくなる心理効果である。Huettig & Altmann (2005) は視覚世界パラダイム (world visual paradigm) を用いて、「ピアノ」と言う言語音声呈示されたときには、ピアノの画像だけでなく、同じく楽器であるトランペットの画像に向けられる視線継続時間も増加することを示した。さらに Toon & Kukona (2019) は非言語音が言語音声よりも強固な意味的競合の効果をもたらすことを示唆した。以上のように、一見すると聴覚刺激との対応関係が無い視覚刺激についても、聴覚刺激に

よって直接的に想起されるものとの意味的な結びつきを有している場合には、意味共活性化 (semantic co-activation) により、それらに対する認知的処理をトップダウン的に促進することが示されている。しかし、これについて非言語音を用いて検討した先行研究は未だ少ない。意識的な反応について検討する課題を用いた実験はほとんど行われていない。本研究では、先行する聴覚刺激の呈示が、直接的な対応関係をもたないが意味的に結びつきのある視覚刺激に対する処理および反応においてもプライミング効果をもたらすことを、2つの実験により示した。

## 2. 実験 1

### 2.1. 参加者

18-25歳の学生15名 ( $M = 21.33$ ,  $SD = 1.99$ ) が参加した。いずれの参加者も聴力・視力または矯正視力は正常であった。ただし、参加者15名のうち3名については、先に募集された参加者3名のデータの分析対象からの除外（後述）を理由に、参加者間における条件順序のカウンターバランスを維持するために追加で募集した。

### 2.2. 装置

パーソナルコンピュータ (Apple, MacBook Air M1), およびパーソナルコンピュータに接続されたヘッドフォン (Shure, SRH840) を使用した。

### 2.3. 刺激

ターゲット画像として POPORO (Kovalenko et al., 2012) に含まれる画像を用いた。POPORO は、ある物体（以下、Object と表記）を示す「Object 画像」、Object との意味的な結びつきの強い物体を示す「Match 画像」、Object との意味的な結びつきの弱い物体を示す「MisMatch 画像」からなるセットが400組含まれる画

像データセットである。本研究ではそれらの画像をグレースケール変換し、そのうち21組を、左右反転したものと併せて用いた。またマスク画像として、ランダムな模様をもつグレースケール画像を1種類使用した。ターゲット画像およびマスク画像はいずれも400×400ピクセルであった。実験1で用いた音刺激については、抽出された各 Object 画像が示す物体に対応する音をそれぞれ作成して用いた（例えば Object 画像がイヌの場合、イヌの鳴き声）。音刺激の継続長は最大350msであった。なお、それぞれの音刺激と Object 画像との対応関係が4名の実験協力者によって事前に確認された。

## 2.4. 手続き

Chen & Spence (2018) に倣った実験課題を実施した。試行の流れは以下の通りであった：最初にブランク画面が1483ms 呈示された後、50% の確率でターゲット画像が17ms 呈示され、続けてマスク画像が1500ms 呈示された。このとき参加者は、ターゲット画像が呈示されたときのみ、出来るだけ素早くキーを押すことが求められた。ターゲット画像は Object 画像が表示される congruent 条件, Match 画像が表示される semi-congruent 条件, MisMatch 画像が表示される incongruent 条件を設定した。これら条件の試行をそれぞれ同一のブロックに割り当て、それら3ブロックからなる連続した試行を1シーケンスとした。1ブロックに含まれる試行は、21の各音刺激に対して画像あり(正画像)が1試行、画像あり(反転画像)が1試行、画像なしが2試行の計84試行がランダムな順序で行われた。画面の観察距離は0.3mであった。音刺激はターゲット画像に先行してヘッドフォンから呈示され、視覚刺激との SOA (刺激開始時間間隔, stimulus onset asynchrony) は350ms または500ms であった。SOA は各シーケンス内の試行で同一であり、ブロックおよびシーケンスの順序は参加者間でカウンターバランスがとられた。各シーケンスの前にはそれぞれ21試行からなる練習が設けられ、本試行で用いられる音刺激が1回ずつ呈示された。練習試行において、ターゲット画像は各試行で独立に50% の確率で呈示され、その後続くシーケンスと等しい SOA で実施された。練習試行におけるターゲット画像呈示時間は51ms であった。練習試行では、ターゲット画像が呈示された試行では「おす」、ターゲット画像が呈示されなかった試行では「おさない」と言う文字列が、マスク画像と同時に画面下部

に呈示された。1回目の練習試行のみ実験者が立ち合い、1回目の本試行以降は参加者が一人で取り組んだ。ただし、1回目の練習試行終了時に参加者が課題の内容を理解していないことが実験者によって判断された場合には、再度練習試行が実施された。

## 2.5. 結果

課題の正答率が50% 以下であった参加者1名、および課題の内容を誤解していたことを実験後に申告した参加者2名の計3名のデータを分析の対象から除外した。

Hit 試行(ターゲット画像が表示された試行のうち、参加者が正確にキーを押した試行)の反応時間を分析の対象とした。ここでは反応時間データが ex-Gaussian 分布に従うことを仮定し、以下の一般化線形混合モデルを用いて分析を行った。

$$\begin{aligned} RT &\sim \text{ex-Gaussian}(\mu, \sigma, \tau) \\ \mu &= \beta_0 + \beta_{1i} \cdot SOA_i + \beta_{2j} \cdot Relation_j \\ &\quad + \beta_{3ij} \cdot (SOA_i \cdot Relation_j) + r_{participant} \end{aligned}$$

ここで、 $\beta_0$ ,  $\beta_{1i}$ ,  $\beta_{2j}$ ,  $\beta_{3ij}$  は0または1をとるダミー変数であり、 $r_{participant}$  は平均0, 分散 $\sigma_r^2$ の正規分布に従う被験者ランダム効果である。推定には R (ver. 4.4.1) および rstan パッケージ (ver. 2.32.6), brms パッケージ (ver. 2,22,0: Bürkner, 2017) を用い、NUTS (No-U-Turn Sampler: Hoffman & Gelman, 2014) によるマルコフ連鎖モンテカルロ法でサンプリングを実施した。事前分布には弱情報事前分布が用いられた。サンプリングにおいて、開始から1000サンプルをバーンイン期間として除外し、バーンイン期間後の1000サンプルを使用した。このサンプリングを4回繰り返す、合計4000サンプルを得た。各 SOA 条件について得られた点推定値を比較した結果、SOA 350ms では、semi-congruent 条件は incongruent 条件に比べて12ms 短く、congruent 条件は incongruent 条件に比べて27ms 短かった。また SOA 500ms では、semi-congruent 条件は incongruent 条件に比べて3ms 短く、congruent 条件では incongruent 条件に比べて16ms 短かった(図1・表1)。

なお、いずれの参加者データについても反応正答率が100% の条件が複数あったため、信号検出理論による分析は行わなかった。

図1 各条件の反応時間 (実験1)

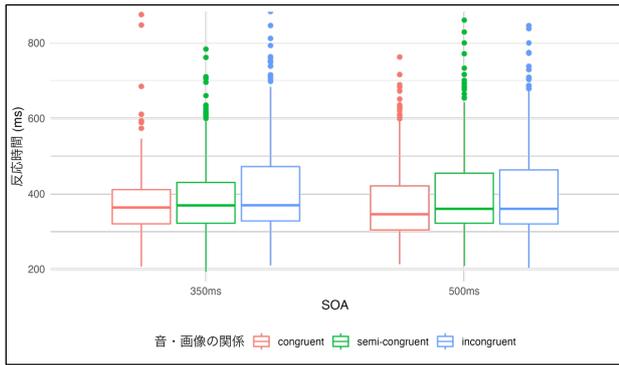


表1 各条件間における反応時間の点推定値の差 (実験1)

	点推定値の差	95% 信用区間
SOA 350ms		
incongruent – semi-congruent	0.012	[0.006, 0.020]
incongruent – congruent	0.027	[0.020, 0.035]
semi-congruent – congruent	0.015	[0.008, 0.022]
SOA 500ms		
incongruent – semi-congruent	0.003	[-0.004, 0.010]
incongruent – congruent	0.019	[0.012, 0.026]
semi-congruent – congruent	0.016	[0.009, 0.023]

### 3. 実験2

#### 3.1. 参加者

実験1で分析対象となった18-24歳の参加者12名 ( $M=21.00, SD=1.91$ ) が引き続き参加した。いずれの参加者も日本語話者であり、実験1への参加から30日以上が経過していた。

#### 3.2. 刺激

ターゲット画像について、実験1で用いた21組の画像の中から17組を抽出して用いた(残り4組の画像については、Objectを4モーラ以下の単語で表現できず、

対応する音声刺激の継続長が長くなりすぎると判断したため、除外した)。実験2では先行する聴覚刺激として言語音声を用いた。音声刺激については、各Object画像が示す物体にそれぞれ対応する単語(例えばObject画像がイヌの場合、「イヌ」)を、同一の話者が読み上げた音声を加工したものを使用した。継続長は2モーラ語が250ms、3モーラ語が350ms、4モーラ語が450msであった。

#### 3.3. 手続き

実験1と同様の手続きで課題を実施した。ただしSOAについては実験1と異なり、音声刺激の継続長を考慮して500msまたは1000msとした。

#### 3.4. 結果

Hit試行における反応時間について、実験1と同様のモデルを用いて分析を行った。各SOA条件において得られた点推定値を比較した結果、SOA500msでは、semi-congruent条件はincongruent条件に比べて3ms長かった一方、congruent条件はincongruent条件に比べて5ms短かった。またSOA1000msでは、semi-congruent条件はincongruent条件に比べて8ms短く、congruent条件ではincongruent条件に比べて15ms短かった(図2・表2)。

図2 各条件の反応時間 (実験2)

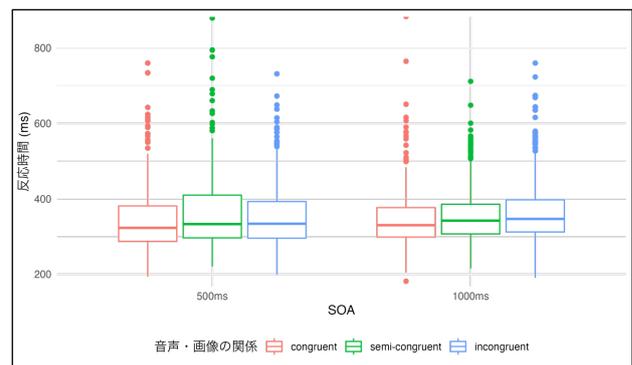


表2 各条件間における反応時間の点推定値の差  
(実験2)

	点推定値の 差	95% 信用区間
SOA 500ms		
incongruent – semi-congruent	-0.003	[-0.009, 0.004]
incongruent – congruent	0.005	[-0.001, 0.012]
semi-congruent – congruent	0.008	[0.002, 0.015]
SOA 1000ms		
incongruent – semi-congruent	0.008	[0.002, 0.015]
incongruent – congruent	0.015	[0.009, 0.022]
semi-congruent – congruent	0.007	[0.000, 0.013]

#### 4. 総合考察

Chen & Spence (2018) の結果と同様、実験1ではSOAが350msのとき、実験2ではSOAが1000msのときにcongruent条件とincongruent条件との反応時間の間に比較的大きな差が見られた。これらの条件では、言語刺激・非言語刺激の両方において、congruent条件では音と画像との一致効果によって処理が促進されたことが推察される。Chen & Spence (2018) によると、言語刺激では処理の際に語彙体系 (lexical system) によって解釈されるための時間を要するのに対し、非言語刺激では直接的に記憶表象を活性化するため、処理に要する時間が短縮される。本研究の結果は、この仮説を支持するものである。また本研究において新たに設定されたsemi-congruent条件においても、プライミング効果の一定生じたことが推察される。実験1ではSOAが350msの場合において、実験2ではSOAが1000msの場合においてsemi-congruent条件はcongruent条件とincongruent条件との中間的な効果を示した。このことから、実験2においてもsemi-congruent条件では意味共活性化によって弱いプライミング効果の生じたことが予想される。以上より、意味共活性化が無意識的な注意を促進するのみにとどまらず、意識的な反応行動をも促進しうることが示された。そして実験参加者の意識

的な反応を求めるプライミング課題を用いることによっても、視線継続時間を指標とした Huettig & Altmann (2005) などと同様に、意味共活性化の特徴について検討できることを示唆した。

本研究の結果は、先行して呈示される聴覚刺激が、後に続く視覚刺激と直接的な対応関係をもたない場合にも、意味共活性化を通じて一定のプライミング効果をもたらすことを示唆した。しかし、そのような先行刺激が後続刺激の処理にもたらす効果の大きさはどれほどか、即ち直接的な対応をもつ刺激におけるそれと同等のプライミング効果をもたらすのかについては議論の余地が大きい。加えて、本研究において実施した実験では使用した刺激の数に限りがあったことから、意味共活性化を生じやすい要因やその普遍性については検討するに至らなかった。今後の研究では、多様な刺激を用いて実験を行うことが期待される。

#### 5. 文献

- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Chen, Y. C., & Spence, C. (2018). Dissociating the time courses of the cross-modal semantic priming effects elicited by naturalistic sounds and spoken words. *Psychonomic Bulletin & Review*, 25(3), 1138–1146. <https://doi.org/10.3758/s13423-017-1324-6>
- Kovalenko, L. Y., Chaumon, M., & Busch, N. A. (2012). A pool of pairs of related objects (POPORO) for investigating visual semantic integration: behavioral and electrophysiological validation. *Brain Topography*, 25(3), 272–284. <https://doi.org/10.1007/s10548-011-0216-8>
- Manesse, C., Fournel, A., Bensafi, M., & Ferdenzi, C. (2020). Visual priming influences olfactomotor response and perceptual experience of smells. *Chemical Senses*, 45(3), 211–218. <https://doi.org/10.1093/chemse/bjaa008>
- Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn Sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(47), 1593–1623.
- Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, 96(1), B23–B32. <https://doi.org/10.1016/j.cognition.2004.10.003>
- Toon, J., & Kukona, A. (2019). Activating semantic knowledge during spoken words and environmental sounds: Evidence from the visual world paradigm. *Cognitive Science*, 44(1), e12810. <https://doi.org/10.1111/cogs.12810>