

# 視聴覚特徴連合の知覚学習

林 明日美  
横澤 一彦

東京大学大学院人文社会系研究科  
東京大学大学院人文社会系研究科

統計的知覚学習の研究ではモダリティの異なる刺激の連合を学習できることが示唆されているが、1対1の連合のみならず、その背後にある刺激間の関連を決定づける規則までも学習できるかどうかは明らかになっていない。そこで本研究では、視聴覚の関連性を決める規則を知覚学習できるかどうかを調べた。実験では、画面上の2次元表示位置が変わる円とピッチの高さが変わる純音を用い、そのペアを連続的に同時提示した。円の位置と音の高さは等確率で4つのうちいずれかの規則(高い音ほど上の方の位置とその逆、高い音ほど右の方の位置とその逆)に基づいて決まっているか、無関連だった。その後、純音のみ提示し、参加者は円の位置を予想した。その結果、参加者が予想した円の左右位置または上下位置は規則に沿う形で音の高さによって線形的にマッピングされていたことから、視聴覚刺激 10 ペア程度の知覚経験から規則的な特徴連合を知覚学習できることが示唆された。

Keywords: Multi-sensory processing, Cross-modal correspondence, Audio-visual association, Perceptual learning.

## 問題・目的

我々は日常生活の中で複数の感覚から膨大な情報を受け取っているが、それらの情報をまとめ、1つの意味のある世界を構成することができる。では、その情報のまとめ方はどのように学習されるのだろうか。

Parise et al. (2014) は、音の高さから予想される音源の位置が自然界及び生活環境の聴覚的情境を反映していることを実証し、日常生活の中で音の高さと空間的位置の対応関係を無意識に学習している可能性を示した。また、いくつかの研究で、統計的に規則性のある刺激に接触するだけで刺激間の関係性を学習できることが示されている (Fiser & Aslin, 2001, 2002; Seitz et al., 2007)。このように、刺激に関する情報や学習を促す課題が提供されなくても受動的に刺激に接触しているだけで成り立つ学習は「知覚学習」と呼ばれている。統計的知覚学習の研究で典型的に用いられる手法は、一定間隔で連続的に提示される刺激のシーケンスの中で時間的または空間的に接触して現れる2つまたは3つの刺激の結合確率を操作し、刺激のシーケンスに接触した後で結合確率の高い刺激のペアとそうでないものを区別できるかを調べるといったものである。

これらの研究では2つまたは3つの刺激の連合を知覚学習できることが示唆されているが、刺激の特徴間に規則的な連合がある場合にその規則を学習できるかどうかは明らかになっていない。

そこで本研究では、視覚刺激の空間位置と聴覚刺激の周波数が連続的に対応している場合に、その対応規則を知覚学習できるかどうかを調べることを目的とした。

## 方法

**参加者** 30名(平均21.63歳、 $SD = 1.40$ )が参加した。

**刺激** 視覚刺激は半径30pxの黒塗りの円で、画面上中央の760×760pxの枠内に提示した。提示される場所は水平・垂直の両方向とも1px単位で変化した。聴覚

刺激は周波数が200~900Hzの範囲を1Hz単位で変化する純音を用い、ヘッドホンを通して提示した。視覚刺激と聴覚刺激は同時に300ms提示され、100msのインターバルを空けて次の視聴覚ペアが提示された。

**手続き** 視覚刺激と聴覚刺激のペアがいくつか連続的に同時提示された後、聴覚刺激のみが提示された。参加者は、その時の視覚刺激の位置を予想し、マウスで円を動かして予測した位置でクリックすることによって回答した。

円の表示位置はランダムであったが、同時提示される音の周波数は、4つのうちいずれかの規則に基づき円の位置によって決まっていた。その規則とは、(1)円が上の方にあるほど音は高い、(2)円が下の方にあるほど音は高い、(3)円が左の方にあるほど音は高い、(4)円が右の方にあるほど音は高い、の4つであった。従って、(1)と(2)の規則に基づいている場合は円のy座標のみが音の高さと関連しており、(3)と(4)の規則の場合はx座標のみが関連していたことになる。さらに、課題を難しくするため円の位置と音の高さが無関連の条件も含めた。この場合は、円の位置と音の高さがそれぞれランダムに決まっていた。連続提示される視聴覚刺激の数は5/10/15ペアの3条件があり、聴覚刺激のみが提示される時の周波数は200/375/550/725/900 Hzの5条件であった。視聴覚刺激の連合関係を決める規則は、無関連の場合も含め全て等確率であった。

また、参加者の意欲を持続させるため、毎試行フィードバックを行った。フィードバックとして提示された赤い円の位置は、x座標/y座標のどちらか一軸は規則に従って決められた座標であり、もう一方の軸は参加者がクリックした座標を採用した。円の位置と音の高さが無関連の場合はランダムな位置にフィードバックを提示した。

視聴覚の対応規則5条件×連続提示される視聴覚刺激の数3条件×聴覚刺激の周波数5条件=75条件を3回ずつランダムな順序で繰り返し、全部で225試行を行った。

## 結果

画面中央の座標を(0, 0)、右上を(350, 350)、左下を(-350, -350)とし、4つの対応規則ごとに5種類の周波数の音に対して予想された円の中心座標の平均値を求めた。そして、横軸に聴覚刺激の周波数を取り、(1)上高音条件と(2)下高音条件では予想された位置のy座標を、(3)左高音条件と(4)右高音条件では予想された位置のx座標をプロットし、回帰直線を求めた。4つの対応規則においては1pxの変化と1Hzの変化が対応していたため、もし規則に沿った位置の予測が完璧に出来ていれば、直線の傾きは1または-1になるはずである。一方、円の位置を無作為に回答してれば、全ての音に対して平均座標は0付近になるため直線の傾きは0に近づくと予想される。従って、回帰直線の傾きが1または-1に近いほど規則に沿った予想をより正確にできたことが考えられる。

分析の結果、全ての条件で回帰直線が対応規則に沿ったものとなり( $t(148) \geq 9.70, p < .0001$ )、参加者は音の高さと円の上下位置または左右位置の対応関係を学習できたことが示唆された。表1に回帰直線の傾きを示す。

視聴覚刺激の提示数による違いを見ると、提示数が5ペアの時より10ペアの時のほうが傾きの絶対値が大きい(上高音条件:  $t(296) = 2.72, p < .0069$ ; 下高音条件:  $t(296) = 3.87, p < .0001$ ; 左高音条件:  $t(296) = 4.35, p < .0001$ )、10ペアの時と15ペアの時では差がなかった( $t(296) \leq 0.58, n.s.$ )。このことから、提示刺激数が5ペアから10ペアに増えると視覚刺激の位置の予測精度は上がるが、10ペアから15ペアに増えても予測精度はあまり変わらないことが示された。

また、5ペアの視聴覚刺激が提示された条件について、下高音条件より上高音条件の方が傾きの絶対値が大きい傾向にあり( $t(296) = 1.87, p < .0632$ )、左高音条件より右高音条件の方が傾きの絶対値が大きかった( $t(296) = 3.74, p < .0003$ )。このことから、特徴連合の規則によって学習のしやすさに違いがあり、高い音は空間的に上の位置や右の位置と関連づけられやすいことが考えられる。

表1. 条件ごとの回帰直線の傾き

	上の方が高音	下の方が高音	左の方が高音	右の方が高音
刺激の提示数				
5	0.54	-0.43	-0.42	0.61
10	0.68	-0.66	-0.65	0.69
15	0.67	-0.62	-0.67	0.72

## 考察

実験の結果、参加者は音の高さと視覚刺激の上下位置または左右位置との対応関係を学習し、それに沿った予測ができたことが示された。学習段階で接触した音と全く同じ音がテスト刺激として提示された可能性は低いにも関わらず、音の高さに対応する視覚刺激を予想できたことは、参加者が視聴覚刺激の1対1の連合を学習するのみならず、特徴が連続的に対応しているという構造を読み取れたことを示唆している。

また、提示される視聴覚刺激の数による予測精度の違いから、視聴覚刺激が10ペア程度あれば十分学習ができることが窺える。これは、10分弱~25分程度の学習を必要としていた従来の知覚学習の研究と比較するとかなり少ない。参加者の内観報告によると、頭在的に対応規則に気付けた試行数は2~3割程度であったが、分析の結果から、少しの知覚経験から視聴覚刺激の連合を無意識に学習できていたことが予想できる。

また、上高音条件と右高音条件でより精度の高い予測ができていたことから、特徴連合の規則によって学習のしやすさに違いがあることが示唆された。前者については、日常生活の中で「高い音は上の方から発せられ、低い音は下の方から発せられる」という経験を数多くしてきたため、「上の方ほど高音」という対応関係の枠組みを参加者が初めから持っていたためではないかと考えられる。この結果は、「高い音ほど空間的に上の方に音源定位しやすく、それは生活環境における実際の音源位置と周波数のマッピングを反映している」という Parise et al. (2014)の知見と一致する。

以上のように、少しの刺激経験から視覚刺激の2次元的空间位置と聴覚刺激の周波数というモダリティを超えた特徴連合を知覚学習できること、さらに、特徴連合を決める対応規則によって学習のしやすさが異なることが明らかになった。この結果は、人間は異なるモダリティから刺激を受け取ったときに何らかの枠組みや構造の中に位置付けようとする傾向があることや、そうした知覚経験を通して刺激の関連性を示す枠組みを学習していることを示唆している。

## 引用文献

- Fiser J., Aslin, R., (2001). *Psychological Science*, 12 (6), 499-504.
- Fiser J., Aslin, R., (2002). *Journal of Experimental Psychology: Learning Memory and Cognition*, 28 (3), 458-467.
- Parise, C., Knorre, K., Ernst, M., (2014). *Proceedings of the National Academy of Sciences*, 111(16), 6104-6108.
- Seitz, A., Kim, R., Van, Wassenhove, V., Shams, L., (2007). *Perception*, 36 (10), 1445-1453.