

博士学位論文

感覚刺激の処理速度が
視聴覚統合にもたらす影響の心理物理学的研究

東北大学大学院文学研究科人間科学専攻

竹 島 康 博

目 次

| | |
|--|-----|
| 論文要約 | iii |
| | |
| 第 1 章 視聴覚情報の処理過程の理論モデルおよび諸現象 | 1 |
| 1-1 視聴覚統合に利用される感覚情報 | |
| 1-2 視聴覚情報の統合過程 | |
| 1-2-1 視聴覚統合の多段階生起 | |
| 1-2-2 視聴覚統合と注意の関連 | |
| 1-3 視聴覚統合によって生じる錯覚現象 | |
| 1-4 本論文の目的 | |
| | |
| 第 2 章 視聴覚統合の生起を制約する時空間情報の一致と感覚情報の信頼性 | 14 |
| 2-1 感覚刺激間の時空間情報の一致 | |
| 2-2 感覚刺激間の情報信頼性の強度による影響 | |
| 2-3 研究 1：情報信頼性の時間変化が視聴覚相互作用に与える影響 (Takeshima, Y., & Gyoba, J. (2013b). Changing pitch of sounds alters perceived visual motion trajectory. <i>Multisensory Research</i> , 26, 317-312. doi:10.1163/22134808-00002422) | |
| 2-3-1 目的 | |
| 2-3-2 実験 1：情報信頼性の時間変化に依存した運動軌道錯覚 | |
| 2-3-3 実験 2：視野による情報信頼性の違い | |
| 2-3-4 実験 3：運動の滑らかさの影響 | |
| 2-3-5 考察 | |
| | |
| 第 3 章 低次過程における時間特性の操作が視聴覚統合に与える影響の検討 | 34 |
| 3-1 感覚刺激の処理速度と視聴覚統合の関連 | |
| 3-2 研究 2：視覚刺激の空間周波数操作による視聴覚間の同期判断への影響 | |
| 3-2-1 目的 | |
| 3-2-2 方法 | |
| 3-2-3 結果 | |
| 3-2-4 考察 | |
| 3-3 研究 3：視覚刺激の空間周波数処理特性が分裂錯覚に与える影響 (Takeshima, Y., & Gyoba, J. (in press). Spatial frequency modulates the degree of illusory second flash perception. <i>Multisensory Research</i> . doi: 10.1163/22134808-00002468) | |
| 3-3-1 目的 | |
| 3-3-2 方法 | |
| 3-3-3 結果 | |

3-3-4 考察

3-4 総合考察

第 4 章 高次過程における視覚刺激特性の操作が視聴覚統合に与える影響の検討・48

4-1 研究 4：視覚刺激の複雑さが分裂錯覚の生起頻度に与える影響

(Takeshima, Y., & Gyoba, J. (2013c). Complexity of visual stimuli affects visual illusion induced by sound. *Vision Research*, 91, 1-7. doi.org/10.1016/j.visres.2013.07.013)

4-1-1 目的

4-1-2 実験 1：視覚刺激の複雑さの影響

4-1-3 実験 2：Connectivity の統制

4-1-4 実験 3：2 音の SOA による影響

4-1-5 考察

4-2 研究 5：視覚刺激を構成する物体数が視聴覚統合による促進効果に与える影響

(Takeshima, Y., & Gyoba, J. (2014). Pattern dot quantity affects auditory facilitation effects on visual object representations. *Perception*, 43, 107-116. doi:10.1068/p7540)

4-2-1 目的

4-2-2 実験 1：聴覚刺激による視覚表象形成促進への影響

4-2-3 実験 2：処理段階の推定

4-2-4 考察

4-3 総合考察

第 5 章 視聴覚情報の統合過程における制約の機能的必要性 73

5-1 感覚刺激の処理速度の視聴覚統合への影響のメカニズム

5-1-1 視覚処理の処理速度が視聴覚統合に与える影響

5-1-2 感覚刺激の処理速度の影響の機能的必要性

5-1-3 視聴覚統合における処理速度が与える影響の神経基盤

5-2 視聴覚統合を制約する要因の役割

5-2-1 重みづけと知覚形成の安定化および効率化

5-2-2 処理速度および時空間情報の一致と情報源の同一性との関連

5-3 今後の展望と本知見の応用可能性

5-3-1 視聴覚統合に関連した究極要因の検討

5-3-2 処理速度の影響を考慮した視聴覚相互作用の応用

5-3-3 臨床分野への応用

5-4 結論

謝辞 91

引用文献 92

感覚刺激の処理速度が視聴覚統合にもたらす影響の心理物理学的研究

論文要約

我々が外界を知覚する際には、単一の感覚情報だけを利用するのではなく、複数の感覚情報を統合して知覚を形成している。これまでの研究から、この多感覚情報は常に統合されるわけではなく、統合の生起を規定する制約とよべる要因が存在することが明らかとなっている。本論文の目的は、視覚情報と聴覚情報の統合（視聴覚統合）に焦点をあて、統合を規定する新たな制約として感覚刺激の処理速度を取り上げ、視聴覚統合の生起に関連した諸要因が果たしている役割を解明することである。さらに、得られた知見とこれまでの研究の知見をまとめることで、視聴覚情報の統合過程でそれぞれの感覚情報をどのように利用、統合して知覚を形成しているのかを明らかにすることを試みた。

第1章「視聴覚情報の処理過程の理論モデルおよび諸現象」では、まず、視聴覚統合において利用される視覚情報と聴覚情報について概説した。特に、視聴覚統合における各感覚情報の利用方略について説明を試みた考え方として「モダリティ適切性仮説」を取り上げた。この仮説では、多感覚統合において判断すべき事象に対して優位なモダリティの感覚情報が優先的に利用されると主張されている。例えば、視覚は空間解像度が優れており、一方、聴覚は時間解像度が優れている。つまり、物体が存在する位置といった空間情報を判断する場合には視覚情報が、ある時間内の物体の点滅回数といった時間情報を判断する場合には聴覚情報が優先されると推測される。実際に、「腹話術効果」、「Flutter-driving」、「時間腹話術効果」のように、モダリティ適切性仮説を支持する現象も報告されていることから、視聴覚統合における感覚情報の利用過程を考える上で有益な知見を提供している仮説である。

次に、視聴覚情報の統合過程の特徴の1つとして、視聴覚統合が感覚処理の複数の段階において生起することについて議論した。まず、この視聴覚統合の複数段階生起を示す現象として、「Crossmodal correspondence」の影響について取り上げた。Crossmodal correspondence とは、音のピッチの高低が物体の位置の高低と対応するように、感覚情報の特徴どうしが結びつく現象である。Crossmodal correspondence により、運動方向が変容する錯覚現象が報告されており、知覚レベルで影響していることを示す。一方、この現象によって刺激に対する反応が速くなることも示されており、認知レベルでも影響が生じることを示している。また、運動知覚では信号検出理論を利用して、運動方向の検出感度と判断基準が共に視聴覚統合によって変化することが報告されており、このことから統合が複数段階で生起することが示されている。

続いて、視聴覚情報の統合過程を考える上で重要な注意機能との関連についても議論した。これまでの研究から、視聴覚統合によって注意の瞬きや反復の見落としといった見落とし現象の効果が低減することが知られている。この背景には、視覚情報と聴覚情報が注意によって結び付けられて単一感覚時よりも頑健な表象を形成することが関連していると

考えられる。また、視覚刺激に対する注意が同期した聴覚入力にも拡散して感覚処理全体を促進する「Crossmodal attentional spread」と呼ばれる現象も、視聴覚統合によって課題成績が向上する 1 つの要因であると考えられる。

最後に、視聴覚情報の統合によって生じる視聴覚相互作用について取り上げた。聴覚刺激によって視知覚が変容する現象は、通過／反発現象や **Sound-induced visual motion (SIVM)** をはじめとして数多く報告されており、運動知覚や大きさ知覚、速度知覚など変容が生じる事象も様々である。また、神経基盤の解明が積極的に行われている視聴覚相互作用も存在し、本論文でも実験的検討を行った「分裂錯覚」もそのような視聴覚相互作用の 1 つである。この錯覚では、視覚刺激を短く 1 回提示するのに合わせて短い聴覚刺激を 2 回提示することにより、視覚刺激も 2 回提示されたように知覚される。機能的磁気共鳴画像 (fMRI) や事象関連電位 (ERP)、脳磁図 (MEG) など様々な脳機能計測手法による研究から、聴覚刺激によって実際には提示されていない視覚刺激の知覚が脳活動として生じていることが裏付けられている。この錯覚現象は、複数の視聴覚情報を統合して知覚の曖昧さを低減する過程の中で生じていると考えられる。

このように、視覚情報と聴覚情報を統合することによって単一感覚時よりも安定した知覚を形成することが可能である。しかし、視聴覚情報の統合は常に生じるわけではなく、統合の生起を制約する要因が存在することが報告されている。視聴覚情報を統合することによって安定した知覚を形成することが可能であるにもかかわらず、その生起を抑制するようにはたらく要因が存在することは、機能的に重要な役割があると考えられる。したがって、この制約について検討を進めることは、視聴覚情報の統合過程の解明やその応用を行っていくうえで重要である。そこで、本論文では視聴覚統合の制約として検討が行われてきた「時空間情報の一致」と「感覚情報の信頼性」について第 2 章で議論をした上で、新たな制約として「感覚刺激の処理速度」を取り上げて、視聴覚統合に対して果たす役割の解明を目的に第 3 章および第 4 章で心理物理実験による検討を行った。本研究で実施した心理物理実験から得られた知見とそれらの機能的・応用的意義について、第 5 章で総合考察を行った。

第 2 章「視聴覚統合の生起に関連する時空間情報の一致と感覚情報の信頼性」では、これまでの研究で明らかとなっている視聴覚統合の生起を制約する 2 つの要因について議論を行った。時空間情報の一致については、ネコの単一細胞記録の研究から始まり、ヒトの行動実験からも、視覚刺激と聴覚刺激の提示される位置が一致していることや、提示タイミングが同一であること（同期していること）必要であることが明らかとなっている。ただし、各感覚刺激の提示される位置の一致は常に必要なわけではなく、刺激が周辺視野で提示される条件では必要となる場合が多いが、中心視野では必ずしも必要ではないことも報告されている。この時空間情報の一致は、各感覚刺激が同一の事象から発生していることを保証するために必要となると考えられる。したがって、時空間情報の一致は感覚情報どうしを正しく結びつけるための手がかりの役割を果たしていると推測される。ただし、視覚と聴覚はそれぞれ時間や空間の解像度の問題から、時空間情報に対する精度が異なっている。また、それぞれの適応刺激の伝達速度や、神経レベルの反応潜時の違いがあるこ

とから、時空間情報の一致の判断には刺激の提示される位置や提示タイミングの「ずれ」に対してある程度、寛容であることが必要となる。これまでの研究から、空間情報と時間情報はともにずれに対して一致判断の許容範囲が存在することが明らかとなっている。また、多くの視聴覚相互作用は、視覚刺激と聴覚刺激の提示タイミングが物理的に同期していなくても、時間情報の許容範囲内であれば生起することが報告されている。したがって、視聴覚相互作用の生起においても、視覚刺激と聴覚刺激が主観的に同期していると知覚されることが重要であると推測される。

また、感覚情報の信頼性も視聴覚情報の統合に深く関連していることが知られている。感覚情報の信頼性は感覚モダリティの優位性ともつながりがあり、事象に対して優位性を持つモダリティの情報信頼性は必然的に高くなる。ただし、統合時に判断される感覚情報の信頼性がモダリティの優位性と必ずしも一致するとは限らない。多感覚統合では、それぞれの感覚入力に対して最尤推定に基づいて重みづけが行われ、その重みづけに従った感覚情報の優位性に基づいた知覚が形成され则认为られている。そのため、提示頻度（点滅回数）のような時間情報はモダリティ適切性仮説に従えば聴覚情報が優先されるはずであるが、それに反して視覚情報が優先される場合があることが報告されている。反対に、モダリティ適切性仮説に従えば視覚情報が優位になるような場合でも、聴覚刺激の影響が生じることも報告されている。特に、視覚は周辺視野ほど正確性が損なわれるため、周辺視野では視覚情報の信頼性が低くなる。そのため、中心視野では聴覚刺激の影響が生じないが、周辺視野では視聴覚相互作用が生じるようになる事例が報告されている。したがって、視聴覚統合の生起には視覚と聴覚の相対的な情報信頼性の高さが関連しており、感覚情報の信頼性が高くて単一感覚の情報だけで十分な場合には統合は生じないが、信頼性が低くて知覚が曖昧になる場合には他の感覚情報との統合が生じると考えられる。

しかし、これまでの研究では、感覚情報の信頼性が継時的に変化するような場合の多感覚情報の統合過程については検討が行われていない。中心視野と周辺視野では視覚情報の信頼性が異なることから、視野の中心から周辺に向かうような運動刺激では、情報の信頼性が運動と共に低くなる。そこで、運動刺激を用いて、視覚情報の信頼性が継時的に変化する時の聴覚情報との統合過程について検討を行った（研究1）。実験では、仮現運動刺激を上下視野に提示し、聴覚刺激のピッチの上昇／下降による運動軌道が上昇／下降する錯覚を利用した。下視野では、運動方向が下降する場合には信頼性が継時的に低くなり、上昇では継時的に高くなる。一方、上視野では運動方向と信頼性の継時変化は下視野と逆の対応関係となる。実験の結果、上下視野ともに聴覚刺激のピッチ変化による運動軌道の錯覚は、周辺視野へと向かう運動軌道の場合のみ生じることが確認された。また、信号検出理論を用いて運動軌道の弁別感度を調べたところ、周辺視野の感度は中心視野と比較して低くなることが示された。したがって、情報信頼性が運動とともに継時的に変化する場合には、一連の変化の中で信頼性が相対的に低くなる状態になったときに、視覚と聴覚の感覚情報の統合が行われることが示された。さらに、運動が滑らかに知覚されない場合には、運動につれて情報信頼性が低くなる場合であっても運動軌道の錯覚が生じないことが示された。このことから、一連の運動が1つの事象として脳内で表象形成され、その中で情報の信頼性の変化が検知された場合に、視聴覚統合処理が生起することが示唆された。

第3章「低次過程における時間特性の操作が視聴覚統合に与える影響の検討」では、視聴覚統合の生起に「感覚刺激の処理速度」も関連すると仮定して、その影響について検討を行った。視覚と聴覚では、適応刺激の伝達速度や神経レベルの反応潜時が異なるため、物理的同期よりも主観的同期に依存して視聴覚相互作用が生起する。そのため、物理的に同期していても主観的に同期と知覚されなければ、視聴覚情報の統合が生じないと推測される。したがって、視聴覚情報が統合するまでの過程において、一方の感覚情報処理が大きく遅延すれば、提示タイミングが物理的に同期していたとしても、主観的な同期は損なわれると推測される。そこで、視覚処理の低次過程で処理速度を規定している空間周波数を取り上げ、視覚と聴覚の同期判断に与える影響について検討を行った（研究2）。空間周波数の低いガボールパッチと高いガボールパッチを視覚刺激として、様々なタイミングで聴覚刺激とともに提示して同期判断課題を行った。得られたデータに二乗誤差が最小になるようにガウス関数の当てはめを行って主観的同期点（PSS）を算出したところ、主観的同期と物理的同期とのずれは空間周波数の高い刺激の方が大きくなっており、同期判断のためには視覚刺激を聴覚刺激よりも先行して提示する必要がある。空間周波数の高い刺激に対して反応する持続型チャンネルは信号の伝導速度が遅いため、高空間周波数の視覚刺激に対しては低空間周波数の刺激と比べて視覚処理が遅くなる。この処理速度の違いは、単純反応時間の比較からも裏付けられた。視覚処理が遅くなる場合、聴覚刺激との同期に対して処理速度が遅延する分だけ視覚刺激を先行して提示する必要があると推測される。つまり、視覚刺激の空間周波数による PSS の違いは、視覚処理の処理速度の違いを補正した結果として生じていると考えられる。

さらに、この補正の結果は視聴覚相互作用の生起にも影響を与えると推測される。そこで、視覚刺激の空間周波数の異なる2つのガボールパッチを用いて分裂錯覚の生起頻度の比較を行った（研究3）。その結果、空間周波数が高い視覚刺激では、低い刺激と比較して分裂錯覚が生じにくくなっていた。信号検出理論を用いて実験参加者の判断基準についても測定したが、空間周波数間で判断基準の違いは見られなかった。つまり、空間周波数による分裂錯覚の生起頻度の違いは、実験参加者の反応バイアスによるものではなかった。処理速度の遅い高空間周波数の刺激では、物理的に同期するタイミングで提示しても聴覚刺激との同期知覚が損なわれやすいと考えられる。したがって、視覚刺激の処理速度が同期知覚に影響を与えることが、視聴覚相互作用の生起頻度を変容させていると考えられる。研究2および研究3より、空間周波数による視覚処理の処理速度の違いによって視覚刺激と聴覚刺激の同期知覚が変容し、それによって視聴覚相互作用の生起にも影響が波及することが明らかとなった。

第4章「高次過程における視覚刺激特性の操作が視聴覚統合に与える影響の検討」では、高次過程における視覚処理の処理速度を規定する要因として、視覚刺激の複雑さと物体数を取り上げた。視聴覚統合は感覚処理の複数の段階で生起することが確認されているため、感覚刺激の処理速度の影響も低次過程の段階だけに見られるものではないと推測される。そこで、まず、視覚処理のより高次過程で処理速度を規定している視覚刺激の複雑さが分

裂錯覚の生起に与える影響について検討を行った（研究 4）。視覚刺激の複雑さは、本来質的な変数であるが、同等集合サイズという概念を用いることによって量的に操作することが可能である。実験では、同等集合サイズを利用した複雑さの異なるパターンを視覚刺激として、分裂錯覚の生起頻度の比較を行った。実験の結果、複雑な視覚刺激では、単純な刺激と比べて分裂錯覚が生じにくくなった。また、視覚パターンの **connectivity** を統制した上で再度検討を行ったが、それでも複雑な視覚刺激では分裂錯覚が生じにくくなっていた。したがって、視覚刺激の複雑さによって分裂錯覚の生起頻度が変容することが明らかとなった。一方、2 回の聴覚刺激のオンセット非同期時間（SOA）を操作したところ、SOA が長くなると視覚刺激の複雑さによる分裂錯覚の生起頻度の違いが観察されなくなった。分裂錯覚には、1 回目の視覚刺激と聴覚刺激の統合が必要であることから、視覚処理が遅延することによって聴覚刺激との統合に必要な時間が伸長すると考えられる。したがって、視覚刺激の複雑さによる分裂錯覚の生起頻度の違いは、視覚処理の処理速度による聴覚刺激との統合処理時間の変容によって生じていると考えられる。

また、高次過程における視覚処理の処理速度を規定する要因は、視覚刺激の複雑さ以外にも存在する。そこで、複雑さ以外に視覚刺激の処理速度を規定する要因として物体数を取り上げた。さらに、処理速度の影響の多様性を検討するために、視聴覚相互作用の指標として聴覚刺激による視覚表象形成の促進効果を利用して物体数の影響を検討した（研究 5）。実験では、異同判断課題における聴覚刺激による逆行マスキング効果の低減を指標とした。さらに、物体数による処理速度の影響を調べるため、異同判断を行うターゲットとマスクとの間の ISI を操作した。実験の結果、物体数が増えるほどより長い ISI で聴覚刺激による逆行マスキング効果の低減が見られた。加えて、物体数による処理速度の影響が生じる段階を推定するために、異同判断の反応時間の計測を行った。その結果、物体数が増えるほど反応時間が長くなっていた。異同判断に対する反応時間は視覚刺激の符号化効率の指標となることから、物体数による処理速度の影響は、刺激の符号化段階において生じていると推測される。以上の結果より、視覚刺激を構成する物体数が増えるほど、聴覚刺激による視覚表象形成の促進効果が生じる最適なターゲットとマスク間の時間窓が長い方にシフトしていくと考えられる。

研究 4 および研究 5 より、感覚刺激の処理速度が視聴覚統合に与える影響は、より高次の過程における処理速度の操作によっても生じることが確認された。つまり、視聴覚統合が感覚処理の複数段階で生起するように、感覚刺激の処理速度の影響も複数の段階で生じていることが明らかとなった。また、その影響は分裂錯覚のような錯覚現象から、逆行マスキングに対して頑健になるような視覚表象形成の聴覚刺激による促進効果まで、多様に生起することが示された。最後に、分裂錯覚における 2 音の SOA の操作や、逆行マスキング課題のターゲットとマスクの ISI を操作した結果により、視覚刺激の処理速度が遅延することによって視聴覚情報の統合処理に必要な時間が伸長し、視聴覚相互作用が変容することが明らかとなった。

第 5 章「視聴覚情報の統合過程における制約の機能的必要性」では、本研究で得られた感覚刺激の処理速度が視聴覚統合に及ぼす影響についてまとめた上で、これまでの研究が

ら明らかになっている統合過程に制約をあたえる諸要因が果たす役割について考察した。

まず、研究 1 では、視聴覚統合処理における制約の 1 つである情報信頼性の継時的変化の影響について、運動刺激を用いた検討を行った。その結果、一連の運動を 1 つの事象として脳内で表象の形成を行い、その過程で情報信頼性の変化が検知された場合に視聴覚統合処理が生起することが示された。研究 1 より、それぞれの感覚情報の信頼性が視聴覚統合処理過程でどのように利用されるかについて、補足的なデータを得ることができた。

続いて、研究 2 では、視覚処理の低次過程における処理速度の規定因である空間周波数の違いによって、聴覚刺激との同期判断の変容について検討を行った。その結果、視覚刺激の空間周波数が高い場合には聴覚刺激との同期における物理的同期と主観的同期のずれが大きくなり、安定した同期知覚のために聴覚刺激に対して視覚刺激をより先行して提示する必要があることが明らかとなった。

また、研究 3 より、高空間周波数の視覚刺激では視聴覚相互作用の 1 つである分裂錯覚が生じにくいことが示された。研究 2 および研究 3 の結果より、空間周波数の高い視覚刺激では物理的に同期したタイミングで提示しても聴覚刺激との同期知覚が損なわれやすく、そのために視聴覚情報の統合が生じにくくなると考えられる。さらに、そのために、分裂錯覚のような視聴覚相互作用もまた生じにくくなることが示された。

研究 4 では、視覚刺激の複雑さが視聴覚情報の統合過程に与える影響について、分裂錯覚を用いて検討を行った。その結果、処理速度の遅い複雑な視覚刺激では、単純な刺激と比べて分裂錯覚が生じにくいことが明らかとなった。一方、2 音の SOA を長くすることによって、複雑さによる分裂錯覚の生起頻度の違いは観察されなくなった。

研究 5 では、視覚刺激を構成する物体数が聴覚刺激による逆行マス킹の低減効果に与える影響について検討を行った。その結果、物体数が増えるほど、低減効果が生じるための最適なターゲットとマスク間の ISI が長くなることが示された。視覚刺激を構成する物体数が増えることで処理時間も伸長するため、ターゲットとマスク間の ISI が短いと統合前にマス킹の影響が生じてしまうと推測される。したがって、物体数が増えるほど、視聴覚統合のために十分な長さのターゲットとマスク間の ISI を設定する必要があるといえる。研究 4 の結果も合わせると、視覚処理に遅延が生じる場合には、聴覚刺激との統合に必要な処理時間も伸長することが示唆された。したがって、処理速度の遅い視覚刺激では分裂錯覚が生じにくくなるのは、視聴覚情報の統合に必要な処理時間が伸長するため、2 音の時間間隔が短いと錯覚を生起させるほど十分に統合が行われていないためであると考えられる。

以上のような感覚刺激の処理速度の影響は、視聴覚情報の統合に必要な感覚刺激どうしの同期知覚に密接に関連していると考えられる。研究 1 より、視聴覚統合では、感覚入力形成する事象をベースに感覚情報の信頼性の変化を検知することが示された。視覚刺激と聴覚刺激の同期は、それぞれの感覚入力同一の事象から生じていることの判断基準となると推測される。しかし、視覚と聴覚では適応刺激の伝達速度や感覚皮質の反応潜時に違いがあるため、同期判断時には情報源との距離に依存した時間窓の変容による補正を働かせていることが示唆されている。研究 2 より、視覚刺激の処理速度によって PSS や時間窓が変容することが明らかとなっている。したがって、同期の時間窓の補正には感覚

刺激の処理速度も関連しており、生体の同期判断に対する柔軟な対応機能の1つであるといえる。一方、感覚刺激の処理速度が遅くなると、聴覚刺激との統合に必要な処理時間も伸長し、視聴覚相互作用が生じにくくなるような場合があることが、研究3、研究4、研究5より示されている。これは、視聴覚統合が安定した知覚を形成する役割を持つ点から考えると不利益となるような現象ともいえる。しかし、視覚刺激と聴覚刺激の提示頻度が一定以上に密になると、刺激間の同期知覚が生じなくなるため、統合に必要な処理時間が多少伸長しても、現実場面ではあまり大きな問題は生じないと推測される。

本研究では実際に神経基盤を検討する実験は行っていないものの、これまでの知見と照らし合わせることで、感覚刺激の処理速度の影響の神経基盤を推定することがある程度、可能である。前述したように、感覚刺激の処理速度の影響は視覚と聴覚の同期判断と関連している。視覚と聴覚の同期には、背外側前頭前皮質 (dorsolateral prefrontal cortex: DLPFC) と側頭・頭頂接合部 (temporal parietal junction: TPJ) の活動の協応が関連している。背外側前頭前皮質は、視聴覚統合に関連した領域である前頭前皮質 (prefrontal cortex: PFC) の一部であることから、統合時には視聴覚刺激間の同期判断が行われていることを裏付けている。一方、研究2および研究3で操作した空間周波数の処理は、外側膝状体 (lateral geniculate nucleus: LGN) や一次視覚野 (V1) で行われている。研究4で操作した複雑さの処理にはV4や下側頭皮質 (inferotemporal cortex: IT) が、研究5で操作した物体数の影響は視覚情報の符号化段階で生じていることから下頭頂溝 (inferior parietal sulcus) や頭頂間溝 (intraparietal sulcus) の活動が関連している。視聴覚統合では、視覚処理と聴覚処理が並列に行われる中で相互に作用すると考えられる。視覚と聴覚の入力は各感覚皮質における処理が行われた後に、背外側前頭前皮質を含む前頭葉の領域へと送られる。感覚皮質における処理の過程で、視覚刺激については空間周波数の処理が外側膝状体や一次視覚野で、複雑さなどの形態処理がV4や下側頭皮質で行われている。背外側前頭前皮質と同時に側頭・頭頂接合部や下頭頂小葉にも信号が送られて、視覚刺激と聴覚刺激の同期判断が行われるが、その時に視覚刺激の処理速度による影響が生じると考えられる。また、下頭頂小葉は視聴覚統合に関連している頭頂間溝に近接していることから、これらの領域の活動は視聴覚統合自体にも影響を与え、同期判断の変容がそのまま視聴覚統合に要する処理時間を調整していると推測される。物体数については、視覚的ワーキングメモリ負荷が生じることで頭頂間溝の活動が抑制されて聴覚刺激との同期判断を変容させ、視聴覚統合に要する処理時間を調整すると考えられる。加えて、上側頭溝はV4や下側頭皮質、一次聴覚野と近接しており、上側頭溝で行われる視聴覚統合の処理もまた、視覚刺激の処理速度の影響を受けると考えられる。

以上のような感覚刺激の処理速度の影響の機能的必要性や神経基盤を踏まえた上で、視聴覚情報の統合過程を制約する諸要因の役割について議論を進める。研究1より、視聴覚統合における各感覚情報の重みづけは、知覚される事象をもとに行われていることが示唆されている。仮現運動の場合は、感覚入力としては視覚刺激の出現と消失の繰り返しであるが、視覚表象としてはある位置からある位置までの運動として形成される。聴覚刺激でも音情報から運動軌道を脳内で復元して聴覚表象を形成していると考えられるため、運動刺激の情報信頼性は一連の運動の時空間情報から算出されて重みづけが行われていると考

えられる。感覚入力ではなく知覚される運動のような事象の脳内表象をベースに重みづけを行うことは、統合処理における負荷を軽減する役割があると考えられる。また、事象の知覚は生体の感覚処理の最終目標であるため、あらかじめ目標を設定した上でそれに最適な感覚情報の統合プロセスを経ることができるという点で有益であると結論付けられる。視聴覚刺激の物体表象は、それぞれの感覚入力から物体情報と時空間情報を並列に処理を行い、その後に各情報を統合することで形成される。したがって、視聴覚統合処理では、高次過程において形成された視聴覚表象からそれぞれの感覚情報の信頼性に基づいた入力の重みづけを行い、フィードバック経路を利用して知覚を再構成していると推測される。

時空間情報の一致に関しては、視聴覚統合における視覚と聴覚の情報源の同一性を保証する役割を担っていると考えられる。適応刺激の物理的特性や感覚刺激の反応潜時の影響から同期知覚に許容範囲をもっており、感覚刺激の処理速度によって統合に必要な処理時間に変容するのも、視覚情報と聴覚情報を正しく統合しようとする同じはたらきによるものと推測される。感覚処理の遅延によって統合を含めた全体的な処理も遅延するとすれば、様々な感覚入力が絶えず行われる場合には異なる情報源の入力どうしが統合される可能性も出てくる。しかし、実際には同一の情報源の入力どうしが統合されていることから、情報源の同一性を感覚処理のかなり低次のレベルでも行っていると推測される。したがって、感覚刺激の処理速度による統合の処理時間の伸長により、感覚処理の低次過程にも誤った統合を抑える機能が存在していることが示唆される。

最後に、視聴覚情報の統合過程を解明していく上での今後の展望、および本研究で得られた知見の応用の可能性について議論を展開した。本研究で取り上げた感覚刺激の処理速度の影響の検討も含めて、視聴覚統合の分野で行われている研究の多くは、感覚刺激の物理特性や時空間情報の影響、神経科学的基盤の解明を目的としたものである。これらは、視聴覚統合が行われる直接的な要因、つまりメカニズムといった「至近要因」の解明を目指した研究である。対して、視聴覚統合の機能を生態学的妥当性のような「究極要因」の観点から研究を行うアプローチも必要であるが、そのような研究の数はまだまだ少ない。そこで、今後は生態学的妥当性の観点から視聴覚統合の機能の解明を目指した研究を行っていく必要がある。生態学的妥当性を踏まえた検討を行っていく上で重要となると考えられる要因として、情動刺激や社会的刺激の影響である。このような刺激は、情動的・感性的特性として感覚情報処理に基づくアウェアネスとは独立して素早く自動的に処理されるが、その背景には配偶者選択や防衛反応などの環境に対する適応行動の遂行が存在している。そこで、今後、視覚刺激や聴覚刺激に含まれる情動や感性の情報が視聴覚統合処理に与える影響の検討を行うことで、環境に対する適応的機能として研究を大きく発展させることができると考えられる。

視聴覚統合によって生じる視聴覚相互作用は、さまざまな現象を生じさせるため、映像コンテンツ作成に対して応用することができると考えられる。例えば、近年、映像コンテンツに関する新しい感性概念として、刺激情報を還元した本物らしさを創出する「迫真性」が提案されている。迫真性は、前景情報に対して適切に注意が向けられ、同時に別の感覚モダリティ情報と調和することが必要となるが、視聴覚統合を利用することで喚起のしやすさを操作できると考えられる。また、本研究で取り上げた分裂錯覚は、聴覚刺激によっ

て提示されていない視覚刺激を知覚させられることから、少ないリソースで実際以上の映像を見せることができるという点で有益な現象といえる。しかし、研究 3 及び研究 4 より、処理速度が遅い視覚刺激の場合には 2 音の SOA を適切に調整する必要がある。本研究で検討した空間周波数と複雑さはどちらも定量的な変数であるため、今後さらなる検討を行うことで、処理速度による分裂錯覚が生起するために最適な SOA の変化をアルゴリズム化することも可能である。分裂錯覚の生起に最適なパラメーターの算出が可能になれば、映像コンテンツへの応用が進むと考えられる。

視聴覚統合の研究では、臨床分野への応用が盛んに検討されている。例えば、自閉症スペクトラム障害児は定型発達児と比べて分裂錯覚の生起の時間窓が広いことが報告されている。この時間窓の拡張は、継時的な入力による多感覚処理の不正確さを引き起こし、自閉症スペクトラム障害の特徴の 1 つである社会的コミュニケーションの不全につながっている可能性が示唆されている。研究 4 で示された視覚刺激の処理速度による分裂錯覚の生起頻度の変容も時間的機能に関連していることから、自閉症スペクトラム障害における視聴覚相互作用の時間窓の拡張が果たす役割の解明に大きく貢献できる。また、統合失調症患者では、健常者と比べて視聴覚統合による音声知覚の促進効果が生じにくく、低次過程における感覚の統合処理が不得意であることも示されている。一方で、統合失調症患者には感覚野の不適切な連合があり、幻覚を引き起こす一因となっている。つまり、統合失調症患者は健常者と比べて多感覚情報を適切に処理できないといえる。本研究で検討した感覚刺激の処理速度の影響もまた、統合失調症患者の多感覚統合処理の時間的機能を検討する一助となると考えられ、統合失調症患者の処理速度の異なる感覚刺激の統合処理の特性を調べていくことで、スクリーニングテストの開発などの応用を行うことが可能であると考えられる。

以上のように、本論文では、視聴覚統合の生起に関連した制約について、これまで検討がなされてきた時空間情報の一致、情報信頼性の強度の観点に加え、感覚刺激の処理速度という新たな視点からも検討を行った。今後は、メカニズムの解明のような至近要因からの検討だけでなく、感覚刺激の情動的・感性的特性の影響の検討といった生態学的妥当性の観点からの視聴覚統合についてアプローチを行っていくことで、多感覚処理に対する理解が深まることが期待される。また、本研究から得られた知見は、感覚情報の統合過程で必要となる処理時間をアルゴリズム化することで、視聴覚相互作用を効果的に利用した映像コンテンツの作成への応用が期待される。さらに、自閉症スペクトラム障害などについて多感覚統合処理からアプローチすることや、統合失調症などのスクリーニングテストの開発など、臨床場面への応用も可能である。

第 1 章 視聴覚情報の処理過程の理論モデルおよび諸現象

本論文の目的は、視覚情報と聴覚情報の統合（視聴覚統合）に焦点をあて、統合を規定する新たな制約として感覚刺激の処理速度を取り上げ、視聴覚統合の生起に関連した諸要因が果たしている役割を解明することである。

第 1 章では、これまで提案されてきた視聴覚統合の処理モデルから、視覚情報と聴覚情報がどのように用いられ、統合が行われていくのかについて議論を行っていく。加えて、視聴覚情報の統合の結果により、単一感覚だけの提示とはことなる知覚が生じる現象についても概説していく。

1-1 視聴覚統合に利用される感覚情報

私たちは、日常生活において感覚器官を通して外界の状況を把握している（図 1-1）。これらの感覚器官は一般的に五感と呼ばれているが、心理学では 9 つの器官に分類されることが多い（行場, 2007）。私たちがこのような感覚を通じて外界の環境を知覚する場合には、単一の感覚情報だけを利用するのではなく、複数の感覚からの情報を統合していると考えられる。このような複数の感覚情報の統合は「多感覚統合 (Multisensory integration)」と呼ばれ、1980 年代から研究が行われてきた。近年、多感覚統合の分野の研究の数は爆発的に増えており、論文検索エンジン「PubMed」において「Multisensory」というキーワードの検索結果は、2000 年には 50 件に満たなかったのが、2010 年には 250 件近くまでに達している (Stein, 2012)。

多感覚統合に含まれるものとして、視覚情報と聴覚情報の統合である「視聴覚統合 (Audio-visual integration)」があげられる。視聴覚統合について検討している研究も数多くあり (Spence, 2007 for a review)、本論文においても視聴覚統合について議論を行っていく。まず、それぞれの感覚について概略する。視覚は受容器である眼に適応刺激の光波が入ることによって生じる。視覚処理では、最初に局所的な方位、動き、奥行き、色彩などの情報が検出され、位置や動きの情報が背側経路 (dorsal pathway) で、物体や顔の識別のための色や細かい形の情報が腹側経路 (ventral pathway) で並列に処理され、これらの情報の統合によって最終的に視覚認識が行われる (花沢, 2007)。一方、聴覚は受容器である耳に適応刺激の音波が入ることによって生じる。聴覚処理は、内耳において音波の神経発火への変換や周波数

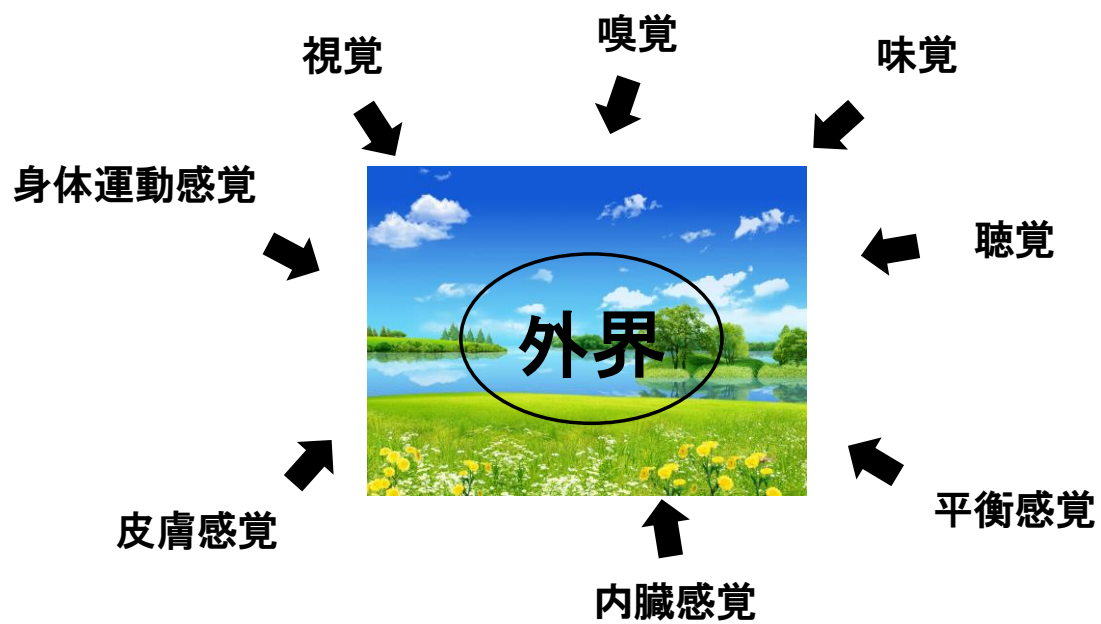


図 1-1. 多感覚情報を利用した外界の知覚のイメージ（風景画像は、<http://free-images.gatag.net/2013/02/08/200000.html> より転載）。

の解析を行われ、その後の聴神経の発火頻度から音の大きさの解析を行っていることは明らかになっているが、聴覚野における細かな情報処理に関しては未解明な部分が多い（入野・津崎，2007）。人間の情報処理では、一般的に視覚の情報が優先的に利用される傾向があり、「視覚優位」と呼ばれている（行場，2007）。

視覚情報と聴覚情報の統合は、単一の感覚情報の弱さ、曖昧さを低減するように行われる。例えば、背景の騒音（ノイズ）が多い場合には、聴覚情報だけでは知覚が不鮮明で音声聞き取りにくい、口唇の動きという視覚情報が加わることによってこの不鮮明さが低減する（Sumbly & Pollack, 1954）。また、視覚と聴覚にはそれぞれ時空間解像度において相補的な関係がある。視覚は聴覚と比べた場合に、空間解像度は優れるが時間解像度に劣るという特性があり、反対に聴覚では、視覚と比べて時間解像度は高いが空間解像度は低いという特性を持つ。そのため、視覚刺激と聴覚刺激が同時に提示された場合には、感覚情報どうしが統合される過程において、空間情報については視覚情報が、時間情報については聴覚情報が優先されると考えられてきた。この考えは、知覚を形成する際により優れているモダリティ（感覚様相）の情報が優先されるという、「モダリティ適切性仮説（Modality appropriateness hypothesis）」という名前で提案されている（Welch & Warren, 1980）。視聴覚刺激の同時提示における時空間的な知覚を検討した研究では、音源（聴覚刺激の提示位置）と、人の顔のような音源と推測される物体の位置（視覚刺激の提示位置）が近接しているが異なる場合には、音源位置の知覚は視覚情報に引きずられる、「腹話術効果（Ventriloquism effect）」と呼ばれる現象が報告されている（Jack & Thurlow, 1973）。また、視覚刺激と聴覚刺激をそれぞれ異なる頻度で連続提示した場合には、視覚刺激の提示頻度は聴覚刺激の提示頻度に引きずられて知覚される、「Flutter-driving」と呼ばれている現象が生じる（Shipley, 1964）。さらに、「腹話術効果」とは反対に、視覚刺激と聴覚刺激を時間的に近接しているが異なるタイミングで提示すると、視覚刺激の提示タイミングは実際よりも聴覚刺激の提示タイミングに引きずられる「時間腹話術効果（Temporal ventriloquism effect）」が存在する（Morein-Zamir, Soto-Faraco, & Kingstone, 2003; Vroomen & de Gelder, 2004）。このように、空間的な知覚には視覚情報が、時間的な知覚には聴覚情報が優先されるモダリティ適切性仮説を支持するような現象は数多く報告されている。したがって、優位な感覚情報を優先的に利用するモダリティ適切性仮説は、視聴覚統合における感覚情報の利用過程を考える上で有益な知見を提供している。

1-2 視聴覚情報の統合過程

1-2-1 視聴覚統合の多段階生起

視聴覚情報の統合過程については、これまで数多くの研究が行われている。その中で、視聴覚統合は、感覚処理のある特定の段階で生起しているのではなく、視覚処理と聴覚処理が並列で行われる中で、複数の段階において生起することが示唆されている。

視聴覚統合が感覚処理の複数段階において生起していることは、知覚レベルで生じていると推測される現象と、認知（決定）レベルで生じていると推測される現象の両方が報告されていることから明らかである。視聴覚統合の複数段階生起を示す現象として、

「Crossmodal correspondence」の影響があげられる。Crossmodal correspondence とは、感覚刺激の特性どうしの内的な結びつきのことで、例えば、音の高さ（ピッチ）の高低は視覚刺激の提示位置の高低や、大きさ、明るさと結びついていることが報告されている (Spence, 2011 for a review)。音のピッチと視覚刺激の各特性との結びつきは、知覚レベルと認知レベルの両方の処理に影響を与えることが報告されている。知覚レベルの研究としては、運動方向の知覚への影響が報告されており、運動方向が上方向と下方向の運動縞を重ね合わせた視覚刺激に、ピッチが変化する聴覚刺激を同時に提示すると知覚される運動方向が変容する (Maeda, Kanai, & Shimojo, 2004)。Maeda *et al.* (2004) では、視覚刺激単体では運動方向が曖昧であるが、ピッチが下降する聴覚刺激では下方向の運動が、上昇する聴覚刺激では上方向の運動が知覚され、これらは知覚レベルで生じている現象であることが報告されている。また、聴覚刺激のピッチ変化による運動方向の変容について、機能的磁気共鳴画像装置 (functional Magnetic Resonance Imaging: fMRI) を用いて神経活動を調べた結果、ピッチの変化による運動情報でも hMT+ や V5+ といった運動知覚に関連した領域の活動が変化することが示されている (Sadaghiani, Maier, & Noppeney, 2009)。一方、認知レベルでの研究としては、先行提示した聴覚刺激のピッチの相対的な高さと、後続するプローブの提示位置の視野の上下が一致していれば、プローブに対する反応潜時が短くなることが報告されている (Chiou & Rich, 2012)。また、明るさとの対応（ピッチが低い音は暗い刺激と、高い音は明るい刺激と対応する）を利用した視覚探索課題の実験では、ターゲット刺激の明るさの変化に合わせて聴覚刺激を提示した場合、明るさとピッチの高低が一致している条件の方が不一致の条件よりも探索時間が短くなることが示されている (Klapetek, Ngo, & Spence, 2012)。この効果は、一致の試行と不一致の試行をブロックで分けたブロックデザ

インを使った実験の中でのみ見られることから、Klapetek *et al.* (2012) では、視覚探索課題において聴覚刺激のピッチは認知レベルで探索時間に影響を与えていると結論付けられている。

また、運動知覚における視聴覚統合について検討した研究においても、知覚レベルでの影響を報告している研究 (e.g., Kitagawa & Ichihara, 2002; Soto-Faraco, Spence, & Kingstone, 2005; Vroomen & de Gelder, 2003) と、認知レベルでの影響を報告している研究 (e.g., Alais & Burr, 2004a; Meyer & Wuerger, 2001; Wuerger, Hofbauer, & Meyer, 2003) の両方がある。そのため、Sanabria, Spence, and Soto-Faraco (2007) は信号検出理論を用いて、1つの実験課題の中で、運動知覚における視聴覚統合の知覚レベルと認知レベルの両方の影響の検討を行った。分析に信号検出理論を用いることにより、知覚レベルの影響の指標である検出感度 (d') と認知レベルの指標である判断基準 (c) の2つの指標を算出することができる (Macmillan & Creelman, 1991)。実験の結果、聴覚刺激の運動方向検出において、同時に提示される視覚刺激の運動方向の一致性によって、検出感度と判断基準の両方の指標が変動しており、知覚レベルと認知レベルの両方において視聴覚統合の影響が生じていた (Sanabria *et al.*, 2007)。ここでは、感覚処理を知覚レベルと認知レベルと大きく2つに分けて議論したが、実際には知覚レベルや認知レベルの中でも複数の段階において視聴覚情報の統合が生じていると考えられる。

1-2-2 視聴覚統合と注意の関連

視聴覚情報の統合過程においては、注意 (attention) も非常に重要な役割を果たしている (Koelewijn, Bronkhorst, & Theeuwes, 2010; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010 for a review)。Chen and Spence (2011) では、逆行マスキングを用いたターゲット刺激の同定課題に対する聴覚刺激の影響の実験結果から、注意を利用した視聴覚表象形成過程のモデルを提案している (図 1-2)。1-1 で取り上げたように、視覚処理では刺激を物体情報と時空間情報に分けて並列に処理を行い、それぞれを統合して視覚表象を形成する (Bowman & Wyble, 2007)。Chen and Spence (2011) のモデルでもこの処理過程を踏襲し、入力された視覚情報と聴覚情報はともに物体情報と時空間情報に分かれて処理が行われる (視覚における物体情報は色や形など、聴覚における物体情報は音の高さや音圧など)。物体情報では視聴覚間で意味的に一致している場合、時空間情報では視聴覚間で時空間的に一致している場合には、注意が感覚処理を促進するようにはたらし、物体表象と時空間表象はそれぞれ

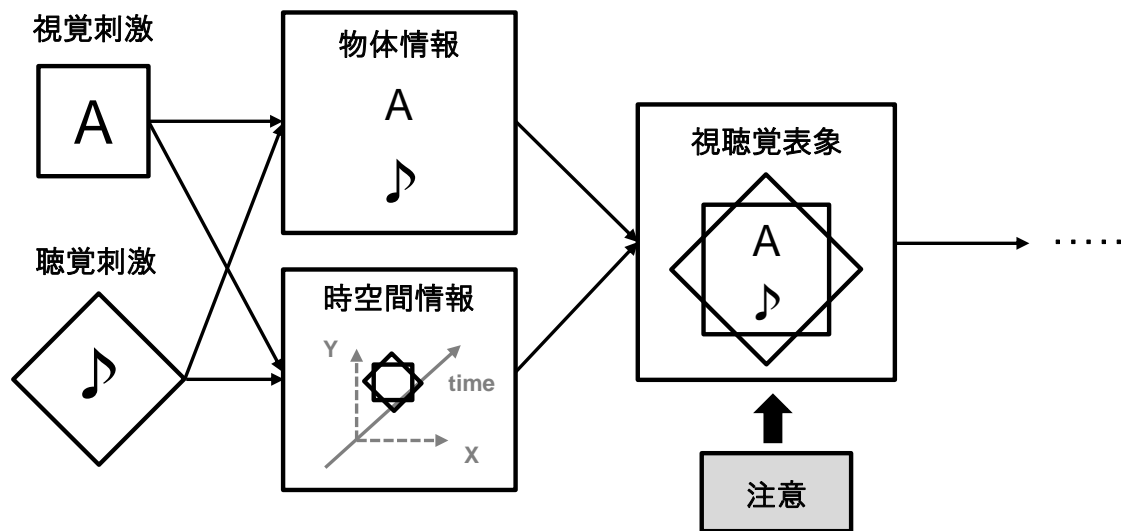


図 1-2. 聴覚刺激による視覚表象形成の促進が生じるメカニズムの説明（Chen and Spence (2011) より一部改変）。視覚刺激と聴覚刺激はそれぞれ物体情報と時空間情報に分けられて処理，統合が行われた後に，2つの情報を統合して1つの視聴覚表象が形成される。

単一感覚時よりも頑健なものとして形成される。最後に、物体表象と時空間表象が注意を「糊」として統合され、1つの視聴覚表象(オブジェクトファイル:Treisman & Gelade, 1980)として形成される。視聴覚刺激の物体表象と時空間表象のどちらか、もしくはどちらもが単一感覚時と比べて頑健となっていた場合、統合された後の視聴覚表象も単一感覚表象と比べてより頑健となる。

感覚表象が頑健になることにより、刺激の見落とし現象が抑制されることも知られている。見落とし現象には、視覚刺激系列を高速提示する課題 (rapid serial visual presentation task: RSVP task) において、2つのターゲットを設定した場合に1番目のターゲット (T1) の正答率に比べて2番目のターゲット (T2) の正答率が著しく低下する「注意の瞬き (Attentional blink)」(Raymond, Shapiro, & Arnell, 1992) や、同一の刺激を2回提示すると繰り返しが見落とされる「反復の見落とし (Repetition blindness)」(Kanwisher, 1987) が有名である。視聴覚統合の研究では、T2の提示と同期して聴覚刺激を提示することで注意の瞬きが低減すること (Olivers & Van der Burg, 2008) や、同一の刺激にそれぞれ同期して聴覚刺激を提示することで反復の見落としが低減される (Chen & Yeh, 2008; 2009) ことが報告されている。

視聴覚統合によって単一感覚の時よりも課題成績が向上することを説明する現象としては、他にも「Crossmodal attentional spread」(Busse, Roberts, Crist, Weissman, & Woldorff, 2005) があげられる。Busse *et al.* (2005) では、視覚刺激を左右どちらかの視野にランダムな順番で提示し、一方の視野だけに注意を向けてターゲットを検出するよう教示をし、同期した聴覚刺激がターゲット検出成績に与える影響を、事象関連電位 (Event-related potential: ERP) と fMRI を用いて検討を行っている。ターゲットの検出成績は、聴覚刺激を提示した時の方が高くなっており、加えて注意を向けている視野にターゲットと聴覚刺激が提示された時の聴覚野の活動が、注意が向いていない視野にターゲットと聴覚刺激が提示された時と比べて高くなっていた (図 1-3)。この結果から、Busse *et al.* (2005) では視覚刺激に向けられた注意が同期した聴覚刺激に対しても拡散し、感覚処理全体を促進すると解釈している。

また、視聴覚統合と注意との相互作用も感覚処理の複数段階で生起すると考えられており、Talsma *et al.* (2010) によって処理モデルも提案されている (図 1-4)。このモデルでは、視聴覚情報は各感覚の初期処理、時空間情報の再構成、整合性の調整、意味の分析の順に処理が行われていくとされており、各段階で視聴覚間の相互作用が生じるとされている。

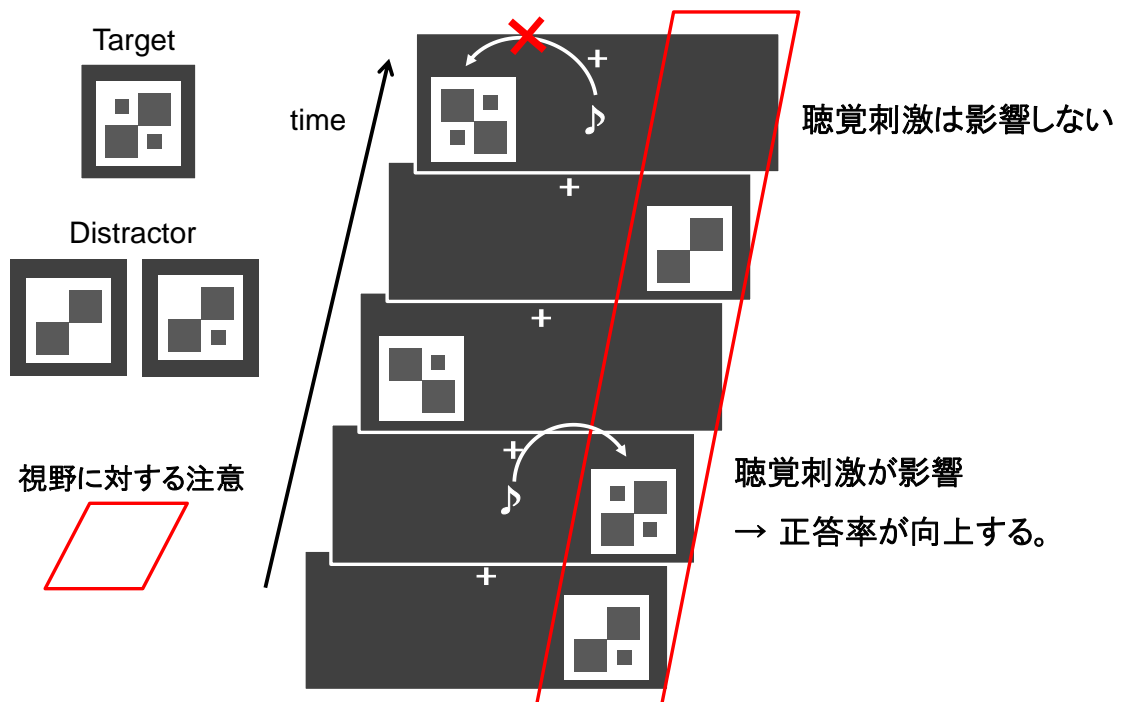


図 1-3. Cross-modal attentional spread の説明 (Busse et al. (2005) より一部改変)。注意が向けられた視野にターゲットが提示される場合には検出の正答率が同期した音によって向上するが、注意が向いていない視野に対しては音によるターゲットの検出への影響は生じない。

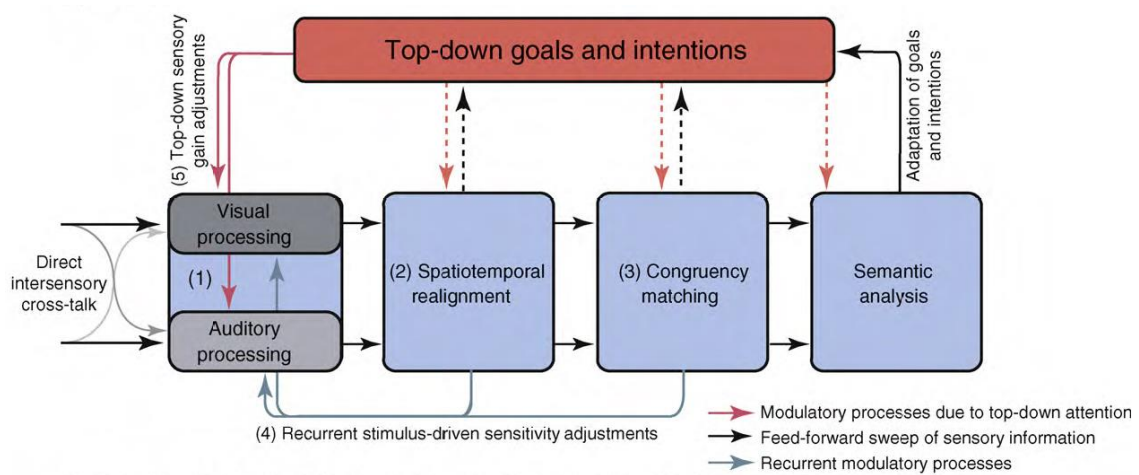


図 1-4. 視聴覚統合と注機能との多段階相互作用のモデル (Talsma et al. (2010) より抜粋)。
 視覚と聴覚の感覚入力は、「感覚の初期処理」、「時空間情報の再構成」、「整合性の調整」、
 「意味の分析」の順に処理が行われ、その過程でボトムアップやトップダウンの注意と相
 互に影響を与える。また、後に行われる高次の処理によって低次の初期感覚入力にも影響
 が生じるため、多感覚入力による順応や学習が生起すると考えられる。

同時に、ボトムアップやトップダウンの注意が、各段階の処理に対して各感覚情報の不鮮明さを低減するように働くとされている。加えて、このモデルでは後期段階の処理の結果が初期段階の処理に影響を与え、視聴覚情報の順応や学習の過程も説明している。

1-3 視聴覚統合によって生じる錯覚現象

視聴覚情報が統合されることによって、視覚や聴覚だけの単一感覚時とは知覚が変容する。1-2で議論した聴覚刺激の提示によって視知覚がより頑健で鮮明になるような現象は、この変容現象の例といえる。本論文では、視聴覚統合によって単一感覚時とは知覚が変容する現象を、総称して「視聴覚相互作用 (Audio-visual interaction)」と呼ぶこととする。視聴覚相互作用は、一方の感覚刺激の提示による他方の感覚刺激の知覚が頑健で鮮明になるだけでなく、錯覚現象として生じることもある。

錯覚現象の視聴覚相互作用として、通過／反発現象があげられる (図 1-5a)。多くの通過／反発現象の実験では、画面の左右から2つの同じ物体を互いに接近するように提示し、画面の中央で重なり、その後互いに離れるように運動させる。このような視覚刺激画面を提示すると、2つの物体が画面の中央で交差するような運動が知覚される (通過知覚)。しかし、2つの物体が重なるタイミングで短く聴覚刺激を提示すると、2つの物体が衝突し反発したような運動が知覚されるようになる (反発知覚 : Sekuler, Sekuler, & Lau, 1997)。この現象は、通過と反発の多義的な解釈が可能な刺激画面の知覚が、聴覚刺激によって変容することを示す現象である。通過と反発の知覚が切り替わるのは、提示される聴覚刺激によって注意の切り替えが生じているため (e.g., Watanabe & Shimojo, 1998) という説明もあるが、注意の切り替えだけではこの現象について十分に説明できないことも報告されている (Grassi & Casco, 2009)。

通過／反発現象は、聴覚情報の付加によって多義的に解釈される運動刺激画面の知覚が切り替わる錯覚であったが、視聴覚統合によって生じる錯覚は多義性をもたない刺激でも生じる。例えば、静止した視覚刺激に運動情報をもつ聴覚刺激を同時に提示することによって視覚刺激の運動が知覚される、Sound-Induced Visual Motion (SIVM) と呼ばれる現象 (図 1-5b) がある (Hidaka, Manaka, Teramoto, Sugita, Miyauchi, Gyoba, Suzuki, & Iwaya, 2009)。この現象では、視覚刺激は同じ位置で点滅を繰り返しているだけであるが、水平方向の運動情報を含む聴覚刺激を提示することで、静止しているはずの物体の水平方向の仮現運動が知覚される。加えて、聴覚刺激の運動情報が左方向であれば左方向の、右方向の運動情報

であれば右方向の仮現運動が知覚されることが報告されている。また、運動知覚だけでなく大きさ知覚においても聴覚情報によって一義的な視覚刺激の変容が生じる（図 1-5c）。聴覚刺激にも物体の大きさに関連した情報が含まれていることが示唆されており（e.g., Carello, Anderson, & Kunkler-Peck, 1998）、特に音の音圧の大きさが物体の大きさと結びついていると考えられる（Lipscomb & Kim, 2004）。そのため、音圧の大きな聴覚刺激を同期して提示することにより、視覚刺激の大きさが実際よりも大きく知覚されることが示されている（Takeshima & Gyoba, 2013a）。他にも速度知覚における視聴覚相互作用もあり（e.g., Manabe & Riquiaroux, 2000; Takeshima & Gyoba, 2011）、多様な視聴覚統合によって生じる錯覚が報告されている。これらはいずれも、複数の視聴覚情報を統合して知覚の曖昧さを低減する過程の中で生じていると考えられる。

錯覚の視聴覚相互作用には、生起メカニズムの神経基盤の検討が行われているものあり、「分裂錯覚 (fission illusion)」もその 1 つである（図 1-5d）。分裂錯覚とは、視覚刺激を短く 1 回提示するのに合わせて短い聴覚刺激を 2 回提示することにより、視覚刺激も 2 回提示されたように知覚される現象である（Shams, Kamitani, & Shimojo, 2000; 2002; Andersen, Tiippana, & Sams, 2004）。聴覚刺激によって提示されていない視覚刺激が知覚される時、一次視覚野 (primary visual area: V1) の活動は物理的に視覚刺激を 2 回提示した場合と同様であることが、fMRI を用いた研究から報告されている（Watkins, Shams, Tanaka, Haynes, & Rees, 2006）。また、分裂錯覚生起時の V1 の活動に関しては、注意など他の要因によるものではないことも裏付けられている（Watkins, Shams, Josephs, & Rees, 2007）。分裂錯覚の生起時に視覚野が活動することは、錯覚である視覚刺激に対して視覚誘発電位 (Visual evoked potentials: VEPs) が生じることからも確認されている（Shams, Kamitani, Thompson, & Shimojo, 2001）。聴覚野（特に A1）から V1 に対してフィードバックする神経経路が存在する（Clavagnier, Falchier, & Kennedy, 2004; Falchier, Clavagnier, Barone, & Kennedy, 2002; Rockland & Ojima, 2003）ことから、2 回目の聴覚刺激の提示が V1 の活動を引き起こし、提示されていない 2 回目の視覚刺激を知覚させていると考えられる。また、分裂錯覚の生起過程においても、刺激提示の直後や一連の刺激の提示後の複数段階において視聴覚情報の統合が行われていることが、ERP や脳磁図 (Magnetoencephalography: MEG) を用いた研究から示唆されている（Mishra, Martinez, Sejnowski, & Hillyard, 2007; Shams, Iwaki, Chawla, & Bhattacharya, 2005）。

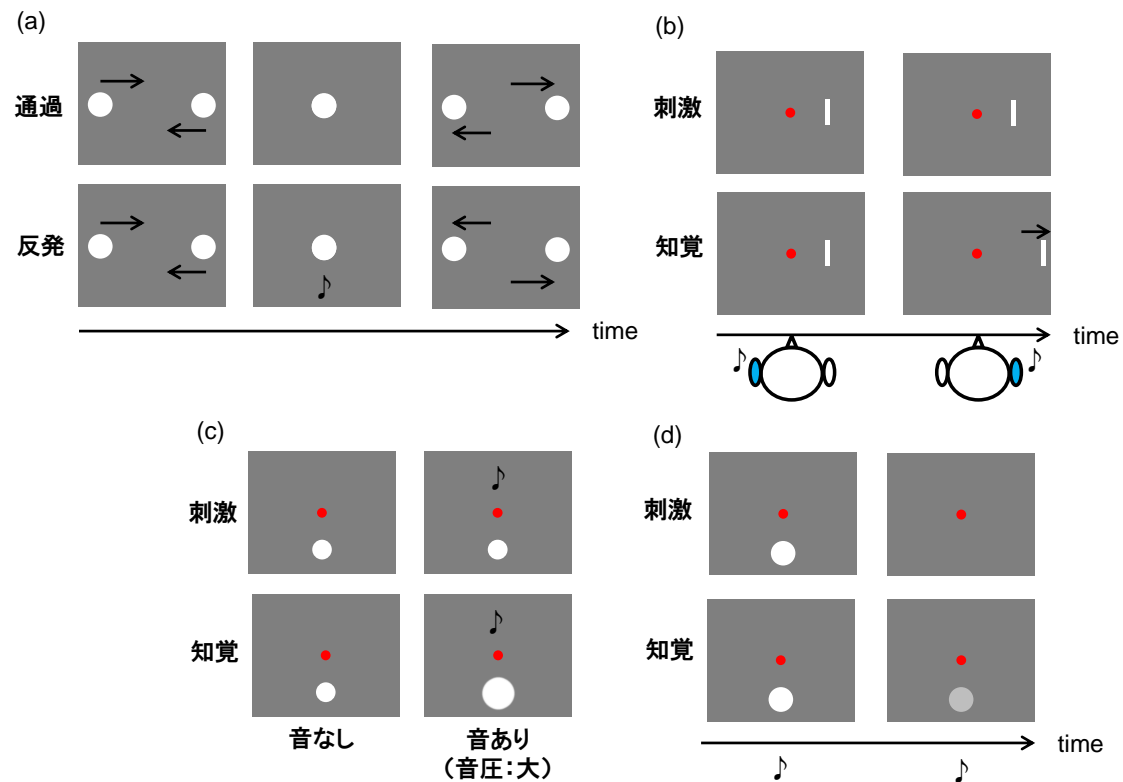


図 1-5. 錯覚現象としての視聴覚相互作用の例。(a) 通過／反発現象。音が提示されない場合は 2 つの物体は互いに交差して通り抜けたように知覚される（通過知覚）が、重なるタイミングで音を提示することによって 2 つの物体が衝突して跳ね返ったように知覚される（反発知覚：Sekuler et al. (1997) を基に作成）。(b) SIVM。左右の耳に交互に音を提示することにより、実際には静止している（同じ位置で点滅している）物体が仮現運動しているように知覚される（Hidaka et al. (2009) を基に作成）。(c) 大きさ知覚の変容。音圧の大きな音を同期して提示することにより、物体の大きさが実際よりも大きく知覚される（Takeshima and Gyoba (2013a) を基に作成）。(d) 分裂錯覚。視覚刺激を短く 1 回提示するのに同期して 2 回の短い音を提示することにより、視覚刺激も 2 回提示されたように知覚される（Shams et al. (2000) を基に作成）。

1-4 本論文の目的

このように、視覚と聴覚の統合によって、課題成績が向上する、知覚が変容するといった現象が数多く報告されている。しかし、このような視聴覚情報の統合はどのような場合にでも生じるわけではなく、制約が存在することが知られている。視聴覚情報を統合することによって安定した知覚を形成することが可能であるにもかかわらず、その生起を抑制するようにはたらく要因が存在することは、機能的に重要な役割があると考えられる。したがって、この制約を検討していくことは、視聴覚情報の統合過程の解明やその応用を行っていくうえで重要であると考えられる。本論文では、まず、これまで視聴覚統合の制約として検討されてきた時空間情報の一致と感覚情報の信頼性について、第2章で議論していく。さらに、感覚情報の信頼性について、時間的な変化という要因を取り入れた研究について報告する（研究1）。続いて、これまで検討されてこなかった要因として、感覚刺激の処理速度を取り上げ、視聴覚情報の統合過程に与える影響について検討していく。第3章では、視覚処理の低次過程における処理速度を規定する要因として刺激の空間周波数を操作し、同期判断および分裂錯覚の生起に与える影響について検討を行った研究について報告する（研究2および研究3）。第4章では、視覚処理の高次過程における処理速度の規定因として、形の複雑さの操作（研究4）と物体数の操作（研究5）を行った研究について報告する。最後に、第5章において本研究から得られた知見を踏まえた、視聴覚統合の生起を制約する諸要因についての機能的および応用的意義の議論を行っていく。

第 2 章 視聴覚統合の生起を制約する時空間情報の一致と感覚情報の信頼性

第 1 章では、視聴覚情報の統合過程に関連したモデルや、統合によって生じる相互作用について議論を行った。視聴覚統合は、外界の情報から安定した知覚を形成するために行われると考えられる。一方、視覚情報と聴覚情報の統合は常に生じるわけではなく、統合の生起の制約となる要因が存在している。第 2 章では、視聴覚統合の生起を制約する要因について議論を行っていく。

2-1 感覚刺激間の時空間情報の一致

視聴覚統合の生起に関連する制約として、まず感覚刺激間の時空間情報の一致があげられる。これは、視聴覚情報の統合が生起するためには、視覚刺激と聴覚刺激の提示位置が空間的に一致していることと、視覚刺激と聴覚刺激が同期していること、これらのいずれか一方もしくは両方が必要となることが報告されている。この制約は、ネコを対象とした単一細胞記録の研究で明らかとなっている。視聴覚刺激に限らず、複数の感覚刺激を同時に提示すると、ネコの脳の上丘 (superior colliculus: SC) が反応することが報告されている (Meredith & Stein, 1983)。このことから、上丘には複数の感覚からの入力に反応する多感覚ニューロンが存在すると考えられている。この多感覚ニューロンの反応には、感覚間の時空間情報の一致に関連している。ネコの上丘の活動の記録では、視覚刺激と聴覚刺激を同一の位置から提示した場合の方が、異なる場所から提示したときよりも細胞の発火量が増えることが報告されている (Meredith & Stein, 1986)。時間情報の一致に関しても検討が行われており、視覚刺激と聴覚刺激を同期して提示した場合に、非同期のときと比べて上丘の細胞の発火量が増えることが示されている (Meredith, Nemitz, & Stein, 1987)。この時空間情報の一致に関しては、ヒトを対象とした行動実験においても検討が行われている。視覚刺激の提示に LED を、聴覚刺激の提示にスピーカーを用いて各感覚刺激の提示位置の一致性を操作した上で逆行マスクングに対する視覚刺激の検出感度を検討した研究では、視聴覚刺激が同じ位置から提示した場合の方が不一致のときよりも検出感度が高くなることが示されている (Frassinetti, Bolognini, & Ladavas, 2002)。また、似たような実験装置を用いて、視聴覚刺激の時間情報の一致が逆行マスクングに対する視覚刺激の検出感度に与える影響を検討した研究では、視聴覚刺激の提示が同期している場合に、非同期の場合と比べて検出感度が高くなることが報告されている (Bolognini, Frassinetti, Serino, & Ladavas, 2005)。ヒ

トにおいても視聴覚情報の統合には上丘の活動が関連していることが報告されている (Fairhall & Macaluso, 2008) ことから、視聴覚統合には時空間情報の一致が重要であると考えられる (図 2-1)。ただし、中心視野における聴覚刺激の提示による視覚刺激の検出成績の向上は、聴覚刺激の提示位置が視覚刺激と異なっても生じるということも報告されている (Noesselt, Bergmann, Hake, Heinze, & Fendrich, 2008)。したがって、視聴覚統合における空間情報の一致は、周辺視野では必要となる場合が多いが、中心視野では必ずしも必要とはいえない (Koelewijn et al., 2010)。これは、視聴覚統合には多くの脳領域が関連しており、その中には空間情報を必要としない機能に関連した領域も含まれている (Stein, London, Wilkinson, & Price, 1996) ことから明らかである。

視聴覚統合における時空間情報の一致の必要性の背景には、感覚刺激間の情報源の同一性を保証する働きが関連していると考えられる。第 1 章で議論したように、多感覚統合はノイズを含んだ感覚情報から安定した知覚を形成するために必要となる。しかし、そのためには複数の感覚間の情報を正確に結びつける必要がある。つまり、同一の事象の感覚情報どうしを統合しなければならない。対象となる感覚情報が同一の事象によるものであれば、それらの感覚情報は同一の情報源から生じているはずである。情報源が同一であるならば、各感覚情報は同一の位置から、同じタイミングで生じているはずである。したがって、感覚情報どうしを正確に統合する上で、各感覚刺激の時空間情報の一致は情報源の同一性を保証する重要な手がかりとなる。

時空間情報の一致を感覚刺激間の情報源の同一性の手がかりとする場合、それぞれの感覚において情報源の位置や刺激の提示タイミングを正確に把握しなければならない。しかし、視覚と聴覚を例にとっても、視覚は時間解像度が劣っており、聴覚は空間解像度が劣っていること (e.g., Welch & Warren, 1980) から、全ての感覚が同じ精度で刺激の提示位置やタイミングを把握できるとはいえない。また、視覚の適応刺激である光波は秒速 300,000,000 メートル、聴覚の適応刺激である音波は秒速 330 メートルと、それぞれ伝達速度が大きく異なっている。さらに、反応潜時も視覚ニューロンと聴覚ニューロンで異なることも報告されている (King, 2010)。したがって、物理的には刺激の提示位置やタイミングがずれていたとしても、主観的には多感覚刺激の提示が時空間的に一致していると知覚される許容範囲が存在する。空間位置に関しては、視覚刺激と聴覚刺激の提示位置が視角 14 度程度離れていても同一の事象から視覚と聴覚の刺激が生じていると判断され、また、提示位置の一致を判断する課題では視角 11 度程度離れていても視聴覚刺激の提示位置が

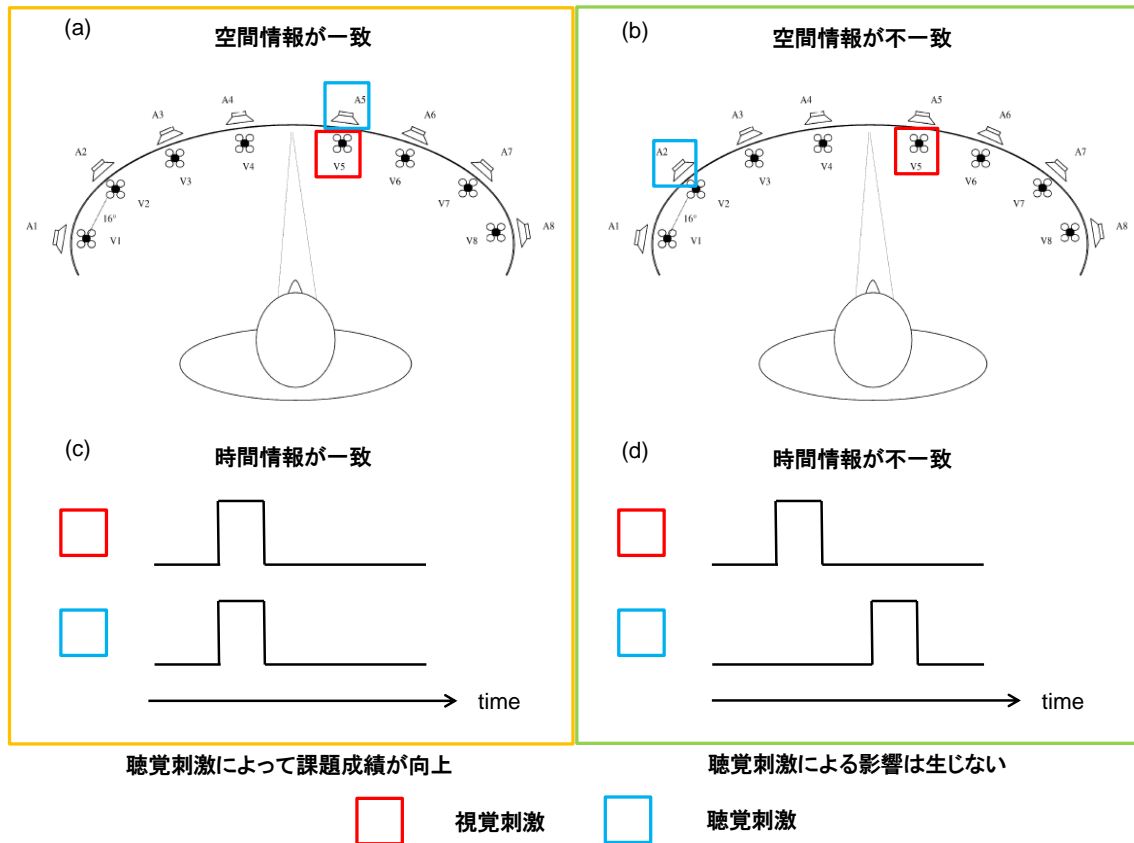


図 2-1. 時空間情報の一致／不一致による視聴覚統合への影響 (Frassinetti et al. (2002) より一部改変)。(a) 空間情報が一致の条件, (b) 空間情報が不一致の条件。赤枠と青枠が視覚刺激と聴覚刺激の提示位置を示している。(c) 時間情報が一致の条件, (d) 時間情報が不一致の条件。凸が視覚刺激と聴覚刺激の提示タイミングを示している。

同じであると判断されることが報告されている (Lewald & Guski, 2003)。一方、同期に関しては視覚刺激に対して聴覚刺激がおおよそ 130 ミリ秒先行から 250 ミリ秒後続までの範囲（この範囲について、本論文では同期の時間窓と呼ぶ）であれば、視聴覚刺激の提示が同期していると判断されることが報告されている (Guski & Troje, 2003; Lewald & Guski, 2003)。また、光と音の伝達速度に違いがあることから、視覚刺激と聴覚刺激の同期の時間窓は、情報源との距離によっても変容する (Sugita & Suzuki, 2003)。加えて、同じ時間ずれの視聴覚刺激を一定時間何度も提示されることによっても同期の時間窓が変化（再校正）される (Fujisaki, Shimojo, Kashino, & Nishida, 2004)。ただし、同期知覚にも限界があり、1 秒間に 4 回 (4 Hz) 以上の時間密度で視聴覚刺激が提示されると同期知覚ができなくなる (Benjamins, Van der Smagt, & Verstraten, 2008; Fujisaki & Nishida, 2005)。

1-3 で議論した視聴覚相互作用の生起にも、同期の時間窓が大きく関連している。例えば、通過／反発現象に関しては、聴覚刺激の提示が 2 つの運動刺激の接触時から 66.5 ミリ秒先行から 66.5 ミリ秒後続までの範囲で生起する (Remijn, Ito, & Nakajima, 2004)。SIVM の生起に関連しては、視覚刺激と聴覚刺激の提示タイミングのずれの許容範囲が 95%信頼区間推定によって検討されており、聴覚刺激が 32.55 ミリ秒先行から 82.27 ミリ秒後続までの範囲となっている (Hidaka et al., 2009)。また、聴覚刺激による視覚的な大きさ知覚の変容現象では、聴覚刺激が 100 ミリ秒先行から 100 ミリ秒後続までの範囲で、音圧の効果が生じることが示されている (Takeshima & Gyoba, 2013a)。さらに、分裂錯覚においては、錯覚としての視覚刺激を知覚させる聴覚刺激の提示が 115 ミリ秒先行から 115 ミリ秒後続までの範囲であれば錯覚が十分に生起することが報告されている (Shams et al., 2000; 2002)。これらの視聴覚相互作用の生起の時間窓は、いずれも同期の時間窓 (Guski & Troje, 2003 を参照) の範囲内に収まっている。加えて、通過／反発現象では、視聴覚刺激の提示タイミングがずれた反発画面に順応させることによって、反発知覚の時間窓が順応に依存して変容することも示されている (Fujisaki et al., 2004)。したがって、視聴覚相互作用の生起においても、視覚刺激と聴覚刺激が主観的に同期していると知覚されることが重要であるとまとめることができる。

2-2 感覚刺激間の情報信頼性の強度による影響

視聴覚統合の生起に関連する制約は、視覚刺激と聴覚刺激の時空間情報の一致だけではない。各感覚情報の信頼性もまた、視聴覚統合に深く関連している。感覚情報の信頼性は、

各感覚モダリティの優位性ともつながりがある。1-1において、空間情報では視覚、時間情報では聴覚と、それぞれ優位なモダリティの情報が優先的に利用されることで感覚情報が最適に利用されるという、モダリティ適切性仮説について述べた。しかし、統合過程における感覚情報の優位性は、モダリティ本来の優位性と必ずしも一致するとは限らない。視覚と聴覚による研究では、大きさ知覚における統合過程において、最尤推定を用いて各感覚入力の重みづけを行い、その重みづけに従った感覚情報の優位性に基づいた知覚が形成されることが報告されている (Ernst & Banks, 2002)。一方の感覚情報 (モダリティ 1) の感覚入力を s_1 、重みづけを w_1 、他方の感覚情報 (モダリティ 2) の感覚入力を s_2 、重みづけを w_2 とすると、推定される統合結果 \hat{s}_B は、 $\hat{s}_B = w_1\hat{s}_1 + w_2\hat{s}_2$ で表される。この重みづけに感覚情報の信頼性が関連しており、モダリティ 1 の信頼性を r_1 、モダリティ 2 の信頼性を r_2 とすると、モダリティ 1 に対する重みづけは、 $w_1 = r_1 / (r_1 + r_2)$ で表すことができる。この時、各信頼性の値は感覚情報の分散の逆関数で表され、 $r_1 = 1 / \sigma_1$ 、 $r_2 = 1 / \sigma_2$ と定義される (図 2-2)。

この最尤推定に基づく重みづけにしたがった感覚情報の統合は、視覚と聴覚においても行われていることが示されており、1-3 で取り上げた分裂錯覚も視覚情報と聴覚情報を最尤推定に従って重みづけした結果として生じていることが主張されている (Shams, Ma, & Beierholm, 2005)。視聴覚統合における感覚情報の信頼性は、多くの場合はモダリティ本来の優位性と一致する場合が多いが、最尤推定に従った重みづけに基づいているため、この2つが乖離する場合も生じる。例えば、刺激の提示頻度は時間情報であるため、一般的には聴覚情報が優位となって 1-1 で議論したように「Flutter-driving」という錯覚が生じる。しかし、聴覚刺激の提示頻度の変化が不鮮明な場合には、視覚刺激の提示頻度の情報の重みづけが大きくなり、聴覚刺激の提示頻度の知覚が視覚情報に引きずられるようになる (Wada, Kitagawa, & Noguchi, 2003)。

モダリティの優位性に基づけば視覚情報が優位になるような場合でも、聴覚情報の影響を受けることがある。例えば、1-3 で取り上げた SIVM や大きさの変容現象は、運動知覚や大きさ知覚は視覚が優位となるため、聴覚刺激の影響を受けないはずである。しかし、これらの現象では、視覚刺激の知覚を曖昧にすることで、視覚情報の重みづけを下げ、相対的に聴覚情報の重みづけが高くなるようにしている。視覚においては、対象の大きさが同一であっても、その正確性は視野によって異なり、周辺に行くほど正確性は低くなる (Anstis, 1974)。そのため、中心視野と比べて周辺視野では知覚が曖昧になり、視覚情報の

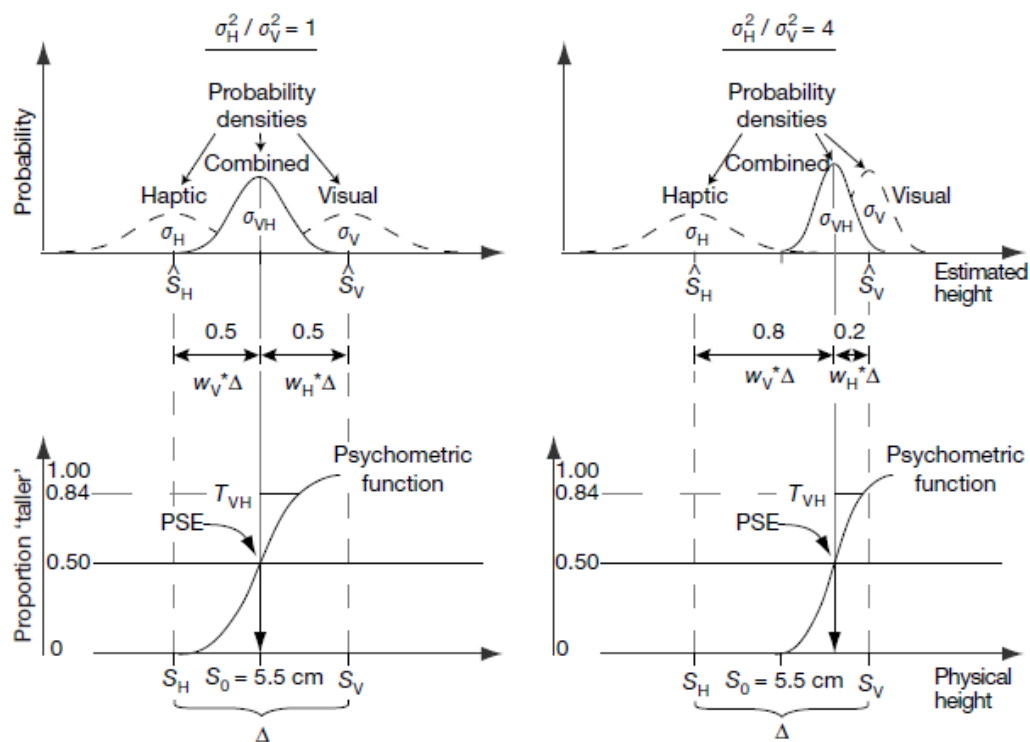


図 2-2. 最尤推定に基づいた感覚入力の重みづけによる知覚の変容の説明モデル (Ernst and Banks (2002) より抜粋)。(左) 視覚入力と触覚入力の情報信頼性が等しく，視覚と触覚に同じ重みづけが行われる条件。(右) 触覚入力と比べて視覚入力の情報信頼性が高いため，視覚に対してより重みづけが行われる条件。

信頼性が低くなる。SIVM や大きさの変容現象では、視覚刺激を周辺視野に提示する（偏心度を大きくする）ことによって視覚単体での知覚を曖昧にし、聴覚刺激に含まれる運動情報や大きさ情報が知覚に影響するようにしている。そのため、視覚刺激を中心視野に提示して視覚情報の信頼性を高くすると、聴覚刺激との統合が行われず、錯覚が生じないことが示されている (Hidaka et al., 2009; Takeshima & Gyoba, 2013a)。また、1-1 で取り上げた視覚優位の現象である腹話術効果も、最尤推定に従った感覚情報の重みづけによっては情報源の定位が聴覚情報に引きずられることが報告されている (Alais & Burr, 2004b)。視聴覚統合の生起には視覚と聴覚の相対的な情報信頼性の高さが関連しており、感覚情報の信頼性が高く単一感覚の情報だけで十分な場合は統合が生じないが、信頼性が低く知覚が曖昧になる場合には他の感覚情報との統合が生じる。

2-3 研究 1：情報信頼性の時間変化が視聴覚相互作用に与える影響

2-3-1 目的

2-2 で議論したように、生体の多感覚情報処理においては各感覚情報の情報信頼性に基づき、信頼性の低い感覚情報による正確性の欠いた曖昧な知覚表象を他の感覚情報によって補うことでより安定した知覚を形成している。その過程において、知覚が物理入力と異なる場合に錯覚が生じていると推測される。これまでの研究では、感覚情報の信頼性を最尤推定に基づいて固定されたパラメーターとして操作し、その統合過程について検討が行われてきた (e.g., Alais, Newell, & Mamassian, 2010)。一方、感覚情報の信頼性が継時的に変化する場合も想定できる。2-2 で議論したように、視覚においては対象の大きさが同一であっても、その正確性は視野によって異なるため、中心視野と比べて周辺視野では視覚情報の信頼性が低くなる (Anstis, 1974)。したがって、視野の中心から周辺に向かうような運動刺激では、情報の信頼性が運動と共に低くなる。このような感覚情報の信頼性が継時的に変化する場合の多感覚情報の統合過程については検討が行われていない。そこで、研究 1 では運動刺激を用いて、視覚情報の信頼性が継時的に変化する時の聴覚情報との統合過程について調べた。なお、ここで報告するデータは, Takeshima and Gyoba (2013b) に基づいている。

実験 1 では、仮現運動刺激を上下視野に提示し、聴覚刺激のピッチ変化による運動の錯覚 (Maeda et al., 2004) を利用して検討を行った。下視野では、運動方向が下降の運動刺激

では信頼性が継時的に低くなり、上昇の運動刺激では継時的に高くなる。一方、上視野では運動方向と信頼性の経時変化は下視野と逆の対応関係となる。また、Maeda *et al.* (2004) では上下の運動方向が曖昧な運動刺激に対して、ピッチが下降する聴覚刺激を同時提示することで下方向の運動が、上昇する聴覚刺激を同時提示することで上方向の運動が優位に知覚されるようになることが示されている。したがって、上下視野でピッチ変化する聴覚刺激によってどのような錯覚が生じるかを調べることで、視覚情報の信頼性が経時変化する時の聴覚情報との統合過程を検討することが可能である。実験2と実験3では、実験1で観察された現象の妥当性について検討を行った。実験2では、視覚における運動情報の信頼性が中心視野と周辺視野で異なることについて、信号検出理論を用いて検討を行った。実験3では、滑らかな運動が知覚されない場合には視聴覚情報の統合が行われにくい (e.g., Sadaghiani *et al.*, 2009; Soto-Faraco, Spence, & Kingstone, 2004) という知見を利用して、刺激提示間時間間隔 (ISI) の操作による知覚される仮現運動の滑らかさの影響の検討を行った。

2-3-2 実験1：情報信頼性の時間変化に依存した運動軌道錯覚

実験参加者：東北大学の学生18名（男性5名、女性13名）が実験1に参加した。そのうち9名（男性3名、女性6名）は視覚刺激が下視野に提示されるブロック（LVFブロック）に、残りの9名は上視野に提示されるブロック（UVFブロック）に割り当てられた。実験参加者は全員、正常な視力（矯正視力を含む）と聴力を有していた。また、実験の目的を知っている実験参加者はいなかった。

装置：実験制御用のPC (XPS720, Dell; OS: Windows Vista, Microsoft; 解像度：1024×768 pixels; リフレッシュレート：60 Hz) を使用し、視覚刺激はCRTディスプレイ (Trinitron GDM-F520, Sony) 上に提示し、聴覚刺激は聴覚デバイス (Edirol FA-66, Roland) とヘッドホン (HDA200, Sennheiser) を通して提示した。刺激の制御には、MATLAB (MathWorks, Inc.) と実験用ライブラリ (Cogent Graphics and 2000 toolbox) を利用した。視覚刺激と聴覚刺激は、デジタルオシロスコープ (TS-80600, Iwatsu) を使用して開始が同期していることを確認した。実験は案室内で行い、暗騒音は43.6 dB SPL (A) であり、視距離はアゴ台を使用して60 cmに固定した。実験参加者の応答は、キーボードのテンキーを使用して行った。

刺激：灰色 (17.9 cd/m²) の背景に、赤色 (15.3 cd/m²) の注視点と仮現運動刺激を提示した

(図 2-2a)。仮現運動刺激は、白色 (43.5 cd/m²) のドット (0.8 × 0.8 deg) で構成されていた。注視点の大きさは直径 1.1 deg で、運動刺激はその上下 10.0 deg 離れた位置に提示された。ドット間の距離は、3.0 deg で、11 ヶ所にドットが提示された (運動距離は 30 deg)。ドットの提示時間は 33 ms で、ISI は 100 ms であった。運動軌道は、上昇、直進、下降で、上昇と下降では最後 2 ヶ所のドットが 0.5 deg ずつ上か下にずれた位置に提示された (図 2-2b-d)。聴覚刺激は、純音 (立ち上り/立ち下り時間 : 10 ms, サンプル周波数 : 22050 Hz) で、周波数は 500 Hz (L-tone), 1000 Hz (M-tone), 2000 Hz (H-tone), 音圧は全て 75 dB (A) で、視覚刺激と同期して 100 ms 提示された。ピッチ変化のパターンは上昇、一定、下降の 3 水準であった。ピッチが変化する場合、仮現運動刺激の最後 2 ヶ所のドットと同期する音に変化し、上昇では L→M→H, 下降では H→M→L とピッチが変化した。一定の場合は、すべてのドットに L か H の音が同期して提示された。

手続き : 各試行は、「0」キーを押すことによって開始された。注視点が 1000 ms 提示された後、仮現運動刺激が提示された (図 2-2e)。実験参加者は、Trajectory (3; upward, straight, or downward) × Pitch (3; ascending, constant, or descending) × Direction (2; leftward or rightward) × 繰り返し (10) の、計 180 試行を実施した。実験参加者の課題は、提示された仮現運動の運動軌道を、上昇か下降の強制二肢選択法で答えることであった。

結果 : 視野ブロック別に、運動軌道が上昇と判断された割合を算出した (図 2-3)。LVF ブロックで得られたデータに対して Trajectory (3) × Pitch (3) の 2 要因分散分析を行ったところ、Trajectory ($F(2, 16) = 255.53, p < .001, \eta_p^2 = .97$) と Pitch ($F(2, 16) = 8.44, p < .005, \eta_p^2 = .51$) の主効果、Trajectory × Pitch の交互作用 ($F(4, 32) = 3.21, p < .05, \eta_p^2 = .29$) が有意であった。2 要因の交互作用が有意であったので単純主効果検定を行ったところ、各 Pitch 条件において Trajectory の単純主効果が有意であった (ascending: $F(2, 48) = 129.24, p < .001, \eta_p^2 = .84$; constant: $F(2, 48) = 123.50, p < .001, \eta_p^2 = .84$; descending: $F(2, 48) = 150.61, p < .001, \eta_p^2 = .86$)。多重比較 (ライアン法) の結果、いずれの Pitch 条件においても upward, straight, downward の順で割合が有意に高くなっていた ($ps < .001$)。したがって、実験参加者は上昇と下降の運動軌道を正確に弁別できていた。また、straight 条件において Pitch の単純主効果が有意であり ($F(2, 48) = 13.69, p < .001, \eta_p^2 = .36$)、多重比較の結果、ピッチが descending 条件の割合が constant, ascending と比べて有意に低くなっていた ($ps < .001$)。一方、constant

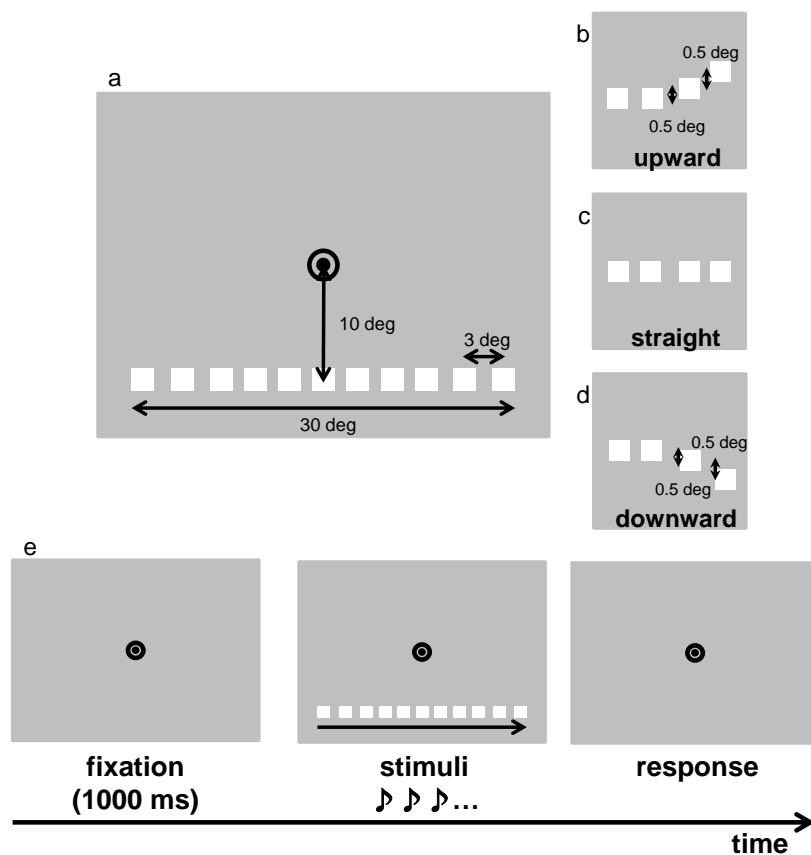


図 2-2. (a) 実験 1 の刺激提示画面図。LVF ブロックでは仮現運動刺激が注視点の下に，UVF ブロックでは注視点の上に提示された。運動方向は右方向と左方向であった。(b-d) 仮現運動刺激の軌道変化の図。(b) は上昇，(c) は直進，(d) は下降の軌道を示す。(e) 実験 1 の 1 試行の流れ。(Takeshima and Gyoba (2013b) より転載。)

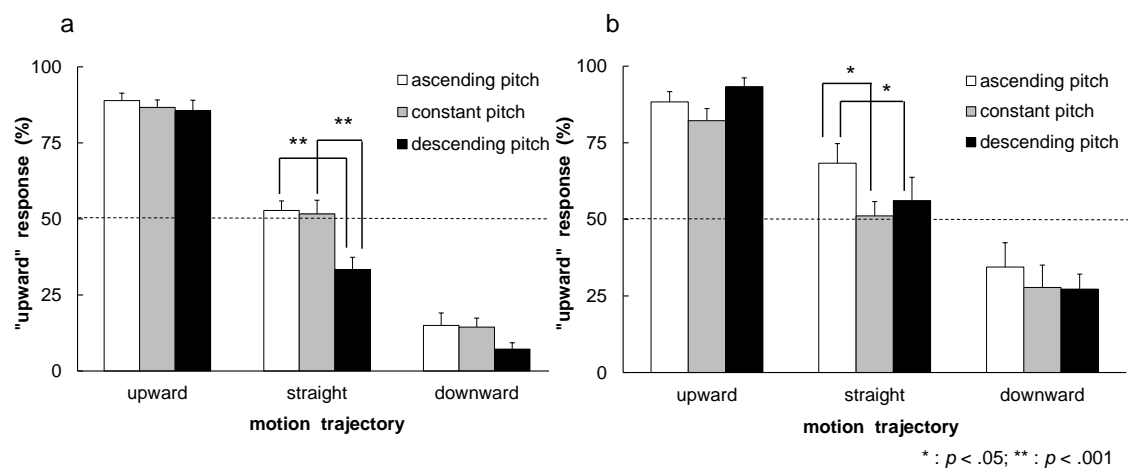


図 2-3. 実験 1 の結果。(a) LVF ブロックで運動軌道が”上昇”と判断された割合。(b) UVF ブロックで運動軌道が”上昇”と判断された割合。エラーバーは標準誤差を示す。(Takeshima and Gyoba (2013b) より転載。)

と ascending との割合の差は有意ではなかった ($p = .79$)。以上の結果より、ピッチが下降する音を付随して提示することにより、運動軌道が下降するように知覚される錯覚が生じることが示された。

一方、UVF ブロックで得られたデータに対しても同様に Trajectory (3) \times Pitch (3) の 2 要因分散分析を行ったところ、Trajectory ($F(2, 16) = 63.56, p < .001, \eta_p^2 = .89$) と Pitch ($F(2, 16) = 4.03, p < .05, \eta_p^2 = .34$) の主効果、Trajectory \times Pitch の交互作用 ($F(4, 32) = 4.02, p < .01, \eta_p^2 = .33$) が有意であった。2 要因の交互作用が有意であったので単純主効果検定を行ったところ、各 Pitch 条件において Trajectory の単純主効果が有意であった (ascending: $F(2, 48) = 42.42, p < .001, \eta_p^2 = .64$; constant: $F(2, 48) = 42.65, p < .001, \eta_p^2 = .84$; descending: $F(2, 48) = 62.79, p < .001, \eta_p^2 = .72$)。多重比較の結果、いずれの Pitch 条件においても upward, straight, downward の順で割合が有意に高くなっていた ($ps < .005$)。したがって、UVF ブロックにおいても、実験参加者は上昇と下降の運動軌道を正確に弁別できていた。また、straight 条件において Pitch の単純主効果が有意であり ($F(2, 48) = 7.55, p < .001, \eta_p^2 = .24$)、多重比較の結果、ピッチが ascending 条件の割合が constant, descending と比べて有意に低くなっていた ($ps < .005$)。一方、constant と descending との割合の差は有意ではなかった ($p = .28$)。以上の結果より、ピッチが下降する音を付随して提示することにより、運動軌道が下降するように知覚される錯覚が生じることが示された。

2-3-3 実験 2：視野による情報信頼性の違い

実験参加者：東北大学の学生 8 名（男性 5 名、女性 3 名）が実験 2 に参加した。そのうち 7 名は実験 1 に参加していなかった。実験参加者は全員、正常な視力（矯正視力を含む）と聴力を有していた。また、実験の目的を知っている実験参加者はいなかった。

刺激：実験 1 と同様の視覚刺激と聴覚刺激を使用した。ただし、運動距離は実験 1 の半分 (15 deg) で、ドットは 7 ヶ所に提示された。運動刺激は画面の左右の端どちらかから提示され、画面の中央に向かって運動した（図 2-4a 参照）。運動刺激の提示位置は、注視点から上下 2 deg（中心視野）もしくは 10 deg（周辺視野）であった。運動軌道は上昇と下降で、最後 2 ヶ所のドットが 0.2 deg ずつ上か下にずれた位置に提示された（図 2-4b-c）。ピッチ変化のパターンは、実験 1 と同様に上昇、一定、下降の 3 水準であった。

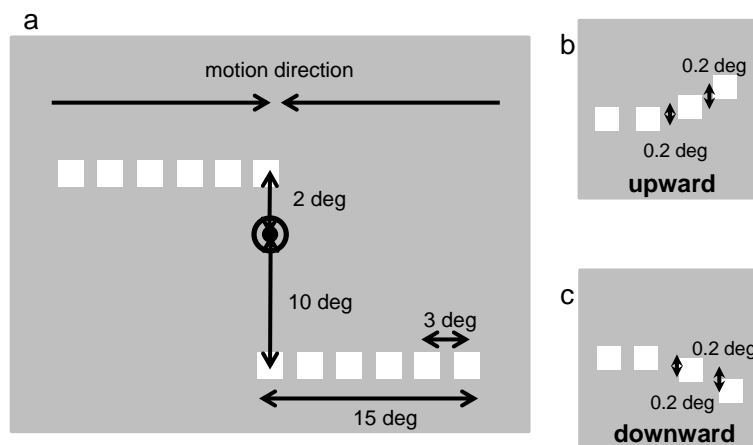


図 2-4. (a) 実験 2 の刺激提示画面図。LVF ブロックでは仮現運動刺激が注視点の下に，UVF ブロックでは注視点の上に提示された。刺激は画面の端から中心に向かって運動した。(b) は上昇，(c) は下降の仮現運動刺激の軌道変化を示す。(Takeshima and Gyoba (2013b) より転載。)

手続き：1 試行の流れは実験 1 と同様であった。実験参加者は、運動刺激が下視野に提示されるブロック（LVF ブロック）と上視野に提示されるブロック（UVF ブロック）の両方を実施された。各ブロックの実施順番は、実験参加者間でカウンターバランスが取られた。実験参加者は、Visual field (2; center or periphery) × Trajectory (3; upward, straight, or downward) × Pitch (3; ascending, constant, or descending) × Direction (2; leftward or rightward) × Block (2; LVF or UVF) × 繰り返し (10) の、計 480 試行を実施した。実験参加者の課題は実験 1 と同様で、提示された仮現運動の運動軌道を上昇か下降の強制二肢選択法で答えることであった。

結果：視野ブロック別に、信号検出理論を用いて運動軌道の弁別感度 (d') と判断基準 (c) を算出した（図 2-5, 表 2-1）。実験参加者が運動軌道を正しく弁別できていた場合には、 d' 値は高くなる。加えて、実験参加者の応答にバイアスが生じていない場合には、 c 値は 0 になる。一方、 c 値が正の値になった場合には上昇と応答しやすいというバイアスが、負の値になった場合には下降と応答しやすいというバイアスが生じていたことを示す。

LVF ブロックの弁別感度のデータに対して Visual field (2) × Pitch (3) の 2 要因分散分析を行ったところ、Visual field の主効果が有意で ($F(1, 7) = 24.59, p < .005, \eta_p^2 = .78$), center の d' 値が periphery と比べて高くなっていた。一方、Pitch の主効果 ($F(2, 14) = 2.06, p = .37, \eta_p^2 = .13$) と Visual field × Pitch の交互作用 ($F(2, 14) = 0.07, p = .93, \eta_p^2 = .01$) は有意ではなかった。また、LVF ブロックの判断基準のデータに対して 95 % 信頼区間推定を行ったところ、中心視野ではいずれのピッチ条件においても応答にバイアスは生じていなかった。一方、周辺視野では constant と descending 条件において運動軌道を下降と応答するバイアスが生じていた。

UVF ブロックの弁別感度のデータに対しても同様に Visual field (2) × Pitch (3) の 2 要因分散分析を行ったところ、Visual field の主効果が有意で ($F(1, 7) = 50.29, p < .001, \eta_p^2 = .88$), center の d' 値が periphery と比べて高くなっていた。一方、Pitch の主効果は有意ではなく ($F(2, 14) = 1.39, p = .28, \eta_p^2 = .17$), Visual field × Pitch の交互作用は有意傾向であった ($F(2, 14) = 1.39, p = .09, \eta_p^2 = .30$)。また、UVF ブロックの判断基準のデータに対して 95 % 信頼区間推定を行ったところ、中心視野ではいずれのピッチ条件においても応答にバイアスは生じていなかった。一方、周辺視野では constant と ascending 条件において運動軌道を上昇と応答するバイアスが生じていた。

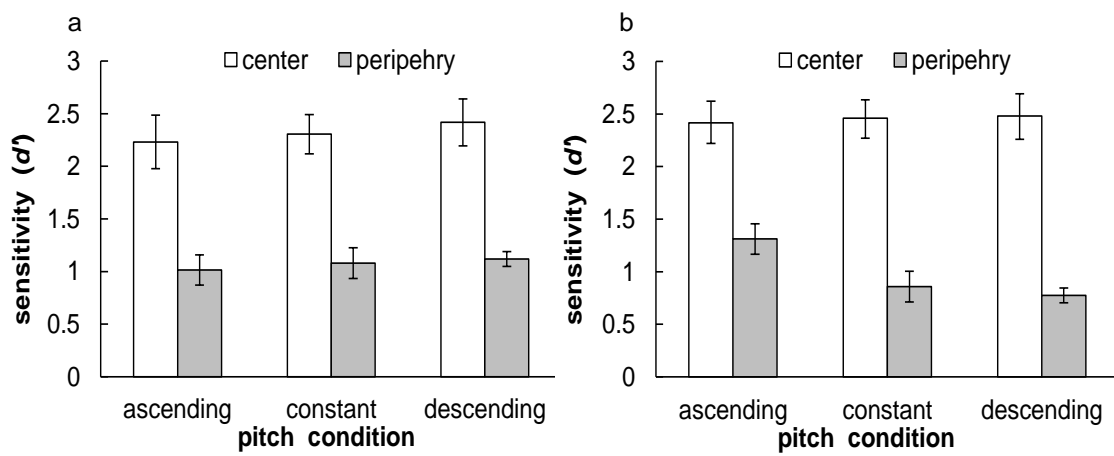


図 2-5. 実験 2 の結果。(a) LVF ブロックの運動軌道の検出感度。(b) UVF ブロックの運動軌道の検出感度。エラーバーは標準誤差を示す。(Takeshima and Gyoba (2013b) より転載。)

表 2-1. 実験 2 の運動軌道の判断基準の結果。上段は中心視野の，下段は周辺視野の判断基準を示す。斜体の値は，正の値であれば 95%信頼区間推定によって参加者の判断が上昇に，負の値であれば下降にバイアスがかかっていたことを示す。(Takeshima and Gyoba (2013b) より転載。)

| Visual field | | center | | |
|------------------------|--------------------------|--------------------|----------------------|----------------------|
| Pitch | | ascending | constant | descending |
| the lower visual field | criterion | -0.04 | -0.12 | -0.18 |
| | 95 % confidence interval | -0.23 – 0.15 | -0.38 – 0.15 | -0.38 – 0.01 |
| the upper visual field | criterion | 0.12 | 0.04 | -0.06 |
| | 95 % confidence interval | -0.11 – 0.37 | -0.18 – 0.26 | -0.24 – 0.11 |
| Visual field | | periphery | | |
| Pitch | | ascending | constant | descending |
| the lower visual field | criterion | 0.22 | -0.16 | -0.20 |
| | 95 % confidence interval | -0.09 – 0.53 | -0.29 – -0.03 | -0.36 – -0.04 |
| the upper visual field | criterion | 0.68 | 0.33 | 0.28 |
| | 95 % confidence interval | 0.44 – 0.92 | 0.03 – 0.64 | -0.05 – 0.61 |

2-3-4 実験 3 : 運動の滑らかさの影響

実験参加者 : 実験者 1 名を含む東北大学の学生 9 名 (男性 3 名, 女性 6 名) が実験 3 に参加した。そのうち 7 名は実験 1 および 2 に参加していなかった。実験参加者は全員, 正常な視力 (矯正視力を含む) と聴力を有していた。また, 実験者を除き, 実験の目的を知っている実験参加者はいなかった。

刺激 : 実験 1 と同様の視覚刺激と聴覚刺激を利用した。運動刺激は, 注視点の下 10 deg の位置に提示された (LVF ブロックのみ実施)。ただし, 仮現運動の ISI を操作し, 100 ms と 500 ms の 2 水準を設定した。

手続き : 実験 1 と同様の 1 試行の流れで, 実験を実施した。実験参加者は, ISI (2; 100 or 500 ms) \times Trajectory (3; upward, straight, or downward) \times Pitch (3; ascending, constant, or descending) \times 運動方向 (2; leftward or rightward) \times 繰り返し (10) の, 計 360 試行を行った。実験参加者の課題は, 実験 1 および 2 と同様で, 提示される運動刺激の軌道が上昇か下降かを強制二肢選択法で応答することであった。

結果 : 運動軌道が上昇と判断された割合を算出した (図 2-6)。ISI 条件別に, Trajectory (3) \times Pitch (3) の 2 要因分散分析を行ったところ, ISI が 100 ms の条件では, Trajectory ($F(2, 16) = 77.28, p < .001, \eta_p^2 = .90$) と Pitch ($F(2, 16) = 5.24, p < .005, \eta_p^2 = .40$) の主効果, Trajectory \times Pitch の交互作用 ($F(4, 32) = 3.41, p < .005, \eta_p^2 = .30$) が有意であった。2 要因の交互作用が有意であったので単純主効果検定を行ったところ, 各 Pitch 条件において Trajectory の単純主効果が有意であった (ascending: $F(2, 48) = 54.17, p < .001, \eta_p^2 = .69$; constant: $F(2, 48) = 48.50, p < .001, \eta_p^2 = .67$; descending: $F(2, 48) = 61.81, p < .001, \eta_p^2 = .72$)。多重比較の結果, いずれの Pitch 条件においても upward, straight, downward の順で割合が有意に高くなっていた ($ps < .001$)。したがって, 実験参加者は上昇と下降の運動軌道を正確に判別できていた。また, straight 条件において Pitch の単純主効果が有意であり ($F(2, 48) = 10.74, p < .001, \eta_p^2 = .31$), 多重比較の結果, ピッチが descending 条件の割合が constant, ascending と比べて有意に低くなっていた ($ps < .001$)。以上の結果より, ピッチが下降する音を付随して提示することにより, 運動軌道が下降するように知覚される錯覚が生じることが示された。一方, ISI が 500 ms の条件では, Trajectory の主効果が有意であった ($F(2,$

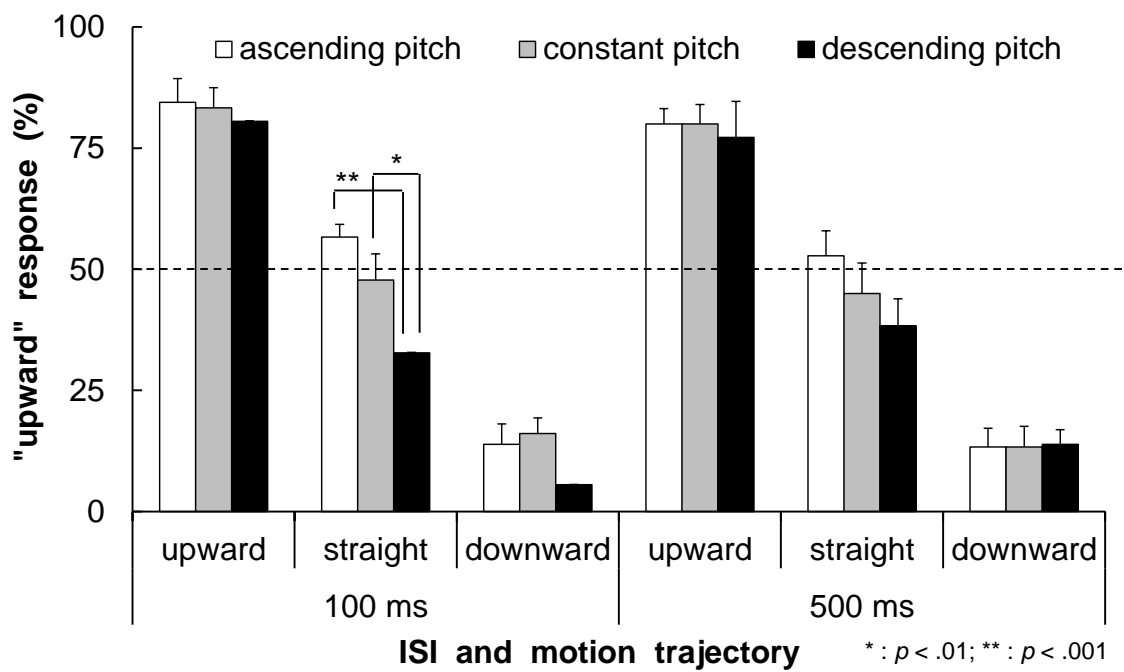


図 2-6. 実験 3 の結果。運動軌道が”上昇”と判断された割合。エラーバーは標準誤差を示す。
(Takeshima and Gyoba (2013b) より転載。)

16) = 106.83, $p < .001$, $\eta_p^2 = .93$)。多重比較の結果, upward, straight, downward の順で割合が有意に高くなっていた ($ps < .001$)。したがって, 実験参加者は上昇と下降の運動軌道を正確に判別できていた。しかし, Pitch の主効果 ($F(2, 16) = 0.85, p = .44, \eta_p^2 = .10$) と Trajectory \times Pitch の交互作用 ($F(4, 32) = 1.61, p = .19, \eta_p^2 = .17$) は見られなかった。

2-3-5 考察

研究 1 では, これまであまり検討されてこなかった, 感覚情報の信頼性が時間変化する場合の影響について, 仮現運動を用いて検討を行った。

実験 1 では, 運動刺激を下視野と上視野に提示して, 付随する聴覚刺激のピッチ変化による運動軌道の錯覚への影響について検討を行った。実験の結果, 上下どちらの視野に提示した場合でも物理的に運動軌道が変化している場合には, その違いを正確に弁別できていた。一方, 上下両視野ともに, 視野の外側に軌道が変化する錯覚のみが付随する聴覚刺激のピッチ変化によって生じていた。周辺視野は, 中心視野と比べて視覚情報の正確性が低くなることが知られている (Anstis, 1974)。そのため, 中心視野と比べて周辺視野では視覚運動軌道の知覚が曖昧になると考えられる。したがって, 聴覚刺激のピッチ変化による運動軌道の錯覚は, 軌道の変化によって視覚情報の信頼性が低くなる場合に生じると推測される。

実験 2 では, 運動軌道の弁別における中心視野と周辺視野の正確性の違いについて検討を行った。実験の結果, 下視野, 上視野ともに周辺視野では中心視野と比べて運動軌道の弁別感度が低くなっていた。これは, 周辺視野では視覚情報の信頼性が低くなることによって生じると考えられる。

実験 3 では, 知覚される運動の滑らかさが聴覚刺激のピッチ変化による運動軌道の錯覚に与える影響について検討を行った。実験の結果, 仮現運動が滑らかに見える場合には聴覚刺激のピッチ変化によって運動軌道の錯覚が生じるが, 滑らかに見えない場合には影響を受けにくくなることが示された。先行研究においても, 聴覚刺激が視覚の運動知覚を変容させるためには, 運動情報が必要であることが示されている (e.g., Sadaghiani et al., 2009; Soto-Faraco et al., 2004)。したがって, 運動知覚において聴覚情報を利用されるためには, 視覚入力によって滑らかな運動が知覚されることが必要であることが示唆される。

以上の結果より, 中心視野と比べて周辺視野では運動軌道判断の正確性が低減するため, 視覚情報の信頼性が低くなっていた。また, 聴覚刺激のピッチ変化による運動軌道の変容

は、視野の中心から周辺に向かう、情報信頼性が運動とともに低くなるような場合にのみ生じていた。視聴覚統合では、聴覚情報は視覚情報の曖昧さを低減するために用いられ、その結果として錯覚が生じると考えられる(Takeshima & Gyoba, 2013a)。したがって、聴覚刺激が引き起こす錯覚によって生じる運動軌道の変化が情報信頼性の変化を伴う場合には、その変化を考慮して信頼性が低減する場合にのみ聴覚情報を用いていると推測される。一方、聴覚刺激が視覚運動知覚に与えるためには、運動が滑らかに知覚されることが必要となる (Sadaghiani et al., 2009; Soto-Faraco et al., 2004)。本研究の運動軌道の変化によって情報信頼性が低減する場合でも、運動が視覚的に滑らかに知覚されなければ聴覚情報を用いられないことが確認されている。これは、一連の運動刺激を1つの事象として感覚表象を形成し、その中で感覚情報の信頼性の変化が検知された場合に、視聴覚情報の統合を行っていることを示唆している。

第 3 章 低次過程における時間特性の操作が視聴覚統合に与える影響の検討

第 2 章では、これまでの研究において検討が行われてきた視聴覚統合の生起に関連する制約について議論を行ってきた。特に、時空間情報の一致は視聴覚刺激の情報源の同一性を保証し、1 つの知覚表象を形成する上で重要であると考えられる。一方、視聴覚刺激の同期に注目すると、「感覚刺激の処理速度」も視聴覚統合の生起に関連する重要な要因となることが推定できる。第 3 章では、視覚刺激の処理速度に焦点を当て、特に初期段階で処理速度を規定する要因について議論を行う。

3-1 感覚刺激の処理速度と視聴覚統合の関連

2-1 で議論したように、視聴覚統合の生起に対しては視覚刺激と聴覚刺激の同期が必要となる (e.g., Bolognini et al., 2005; Meredith et al., 1987)。また、視聴覚刺激間の同期判断には幅をもった時間窓が存在し (Guski & Troje, 2003; Lewald & Guski, 2003), 多くの視聴覚相互作用もこの時間窓の中で生じている (e.g., Hidaka et al., 2009; Remijn et al., 2004; Shams et al., 2000; 2002; Takeshima & Gyoba, 2013a)。そのため、視聴覚統合の生起に必要な同期に関しては、物理的な同期よりも主観的同期の方が重要であると考えられる。

視覚は光を、聴覚は音を適応刺激としているが、光は秒速 300,000,000 メートル、音は秒速 330 メートルと、それぞれ伝達速度が大きく異なっている。さらに、視覚ニューロンと聴覚ニューロンの反応潜時にも違いが存在している (King, 2010)。一方、視覚刺激と聴覚刺激を正確に統合するためには、空間的一致や同期を手がかりとして情報源が同一であることを保証する必要があると考えられる。したがって、視覚刺激と聴覚刺激の情報源の同一性をより柔軟に保証するために、視聴覚間の同期判断には幅をもった時間窓が存在すると考えられる。この同期の時間窓は、視聴覚刺激の情報源との距離によっても変容するため (Sugita & Suzuki, 2003), 柔軟な補正が働く仕組みであると推測される。

視聴覚刺激の同期に関連した研究では、視聴覚間の同期判断がどのような場面で変容するかについて検討が行われている。例えば、視覚刺激と聴覚刺激の提示位置が同じであれば、提示位置が異なる場合と比べて同期と判断が行われやすい (Zampini, Guest, Shore, & Spence, 2005)。また、視覚刺激と聴覚刺激の提示時間が異なると、それに合わせて主観的同期点 (Point of subjective simultaneity: PSS) も変容することも示されている (Kuling, Van Eijk, Juola, & Kohlrausch, 2012)。このように、刺激の時空間特性によって視聴覚刺激間の

PSS が変容する。したがって、視聴覚刺激の提示タイミングが多少ずれていても同期と知覚されるように、物理的に同期していても非同期と知覚されてしまう場合も考えられる。本研究では、視聴覚情報の統合が行われるまでの内的処理に注目して、感覚処理の処理速度に焦点を当てた。視聴覚情報が統合するまでの過程において、一方の感覚情報処理が大きく遅延すれば、提示タイミングが物理的に同期していたとしても、主観的な同期は損なわれると推測される。

3-2 研究 2：視覚刺激の空間周波数操作による視聴覚間の同期判断への影響

3-2-1 目的

研究 2 では、視覚処理の低次過程で処理速度を規定している空間周波数に焦点をあてる。視覚処理において、空間周波数は外側膝状体 (lateral geniculate nucleus: LGN) や V1 において処理が行われる。外側膝状体では、低空間周波数に対しては一過型チャンネル、高空間周波数に対しては持続型チャンネルと、それぞれ異なる細胞が反応する (Breitmeyer & Julesz, 1975; Kulinkowski & Tolhurst, 1973; Tolhurst, 1973)。持続型チャンネルの伝導速度は、一過型チャンネルと比較して遅い (Hoffmann, 1973) ことから、高空間周波数の処理速度は低空間周波数よりも遅くなると考えられる。単純反応時間の比較でも、高空間周波数の刺激は低空間周波数の刺激と比べて遅くなることも示されている (Breitmeyer, 1975)。本実験では、視覚刺激の空間周波数による処理速度の違いが、視聴覚刺激間の同期判断に与える影響について検討を行った。実験では、同期判断 (simultaneous judgment: SJ) 課題と、単純反応時間計測課題を行った。なお、ここで報告するデータは、Takeshima and Gyoba (submitted) に基づいている。

3-2-2 方法

実験参加者：実験者 1 名を含む東北大学の学生 9 名（男性 5 名，女性 4 名）が本実験に参加した。実験参加者は全員、正常な視力（矯正視力を含む）と聴力を有していた。また、実験者を除いて実験の目的を知っている実験参加者はいなかった。

装置：実験制御用の PC (XPS720, Dell; OS: Windows Vista, Microsoft; 解像度：1024 × 768 pixels; リフレッシュレート：60 Hz) を使用し、視覚刺激は CRT ディスプレイ (Trinitron

GDM-F520, Sony) 上に提示し、聴覚刺激は聴覚デバイス (Edirol FA-66, Roland) とヘッドホン (HDA200, Sennheiser) を通して提示した。刺激の制御には、MATLAB (MathWorks, Inc.) と実験用ライブラリ (Cogent Graphics and 2000 toolbox) を利用した。視覚刺激と聴覚刺激は、デジタルオシロスコープ (TS-80600, Iwatsu) を使用して開始が同期していることを確認した。実験は案室内で行い、暗騒音は 43.6 dB SPL (A) であり、視距離はアゴ台を使用して 60 cm に固定した。実験参加者の応答は、キーボードのテンキーを使用して行った。

刺激：各試行では、赤色 (15.3 cd/m^2) の注視点 (直径約 1.1 deg) と視覚刺激が、灰色 (17.9 cd/m^2) の背景に提示された。視覚刺激は空間周波数の異なる 2 種類のガボールパッチを使用し、空間周波数はそれぞれ 1 c/deg と 5 c/deg であった (図 3-1a)。ガボールパッチの大きさは直径 2.0 deg で、提示時間は 17 ms であった。聴覚刺激は純音 (立ち上り/立ち下り時間 : 1.7 ms, サンプリング周波数 : 22050 Hz) で、周波数は 2000 Hz, 音圧は 75 dB (A), 提示時間は 17 ms であった。視覚刺激と聴覚刺激のオンセット非同期時間 (SOA) は 9 水準で、-350, -233, -150, -67, 0, +67, +150, +233, +350 ms を設定した (マイナスの記号は聴覚刺激が先行して、プラスの記号は視覚刺激が先行して提示されることを示す)。

手続き：各試行は、「0」キーを押すことによって開始された。まず、注視点に 500 ms 提示された後、500 ~ 1000 ms のランダムな長さのブランク画面を挟み、視覚刺激であるガボールパッチが提示された (図 3-1b)。また、9 水準の SOA で聴覚刺激も提示された。実験参加者は、視覚刺激と聴覚刺激が同期している試行 (SOA = 0 ms) を 320 試行、非同期の試行を 320 試行 (0 ms 以外の各 SOA を 40 回ずつ繰り返し) の計 640 試行を実施した (同様の手続きの研究として、e.g., Santangelo & Spence, 2008; Zampini et al., 2005; Van der Burg, Olovers, Bronkhorst, & Theeuwes, 2008)。実験参加者の課題は、視覚刺激と聴覚刺激が同期しているかどうかを強制二肢選択法で応答することであった (同期判断課題)。

同期判断課題を実施した後に、実験参加者は各周波数のガボールパッチに対する反応時間を計測する課題を行った (反応時間計測課題)。同期判断課題と同様に、注視点に 500 ms 提示された後に、500 ~ 1000 ms のランダムな長さのブランク画面を挟み、ガボールパッチが提示された (図 3-1c)。実験参加者は、ガボールパッチが提示されたら素早く「5」キーを押すように教示され、各空間周波数に対して 30 回ずつの計 60 試行をランダムな順番で実施した。

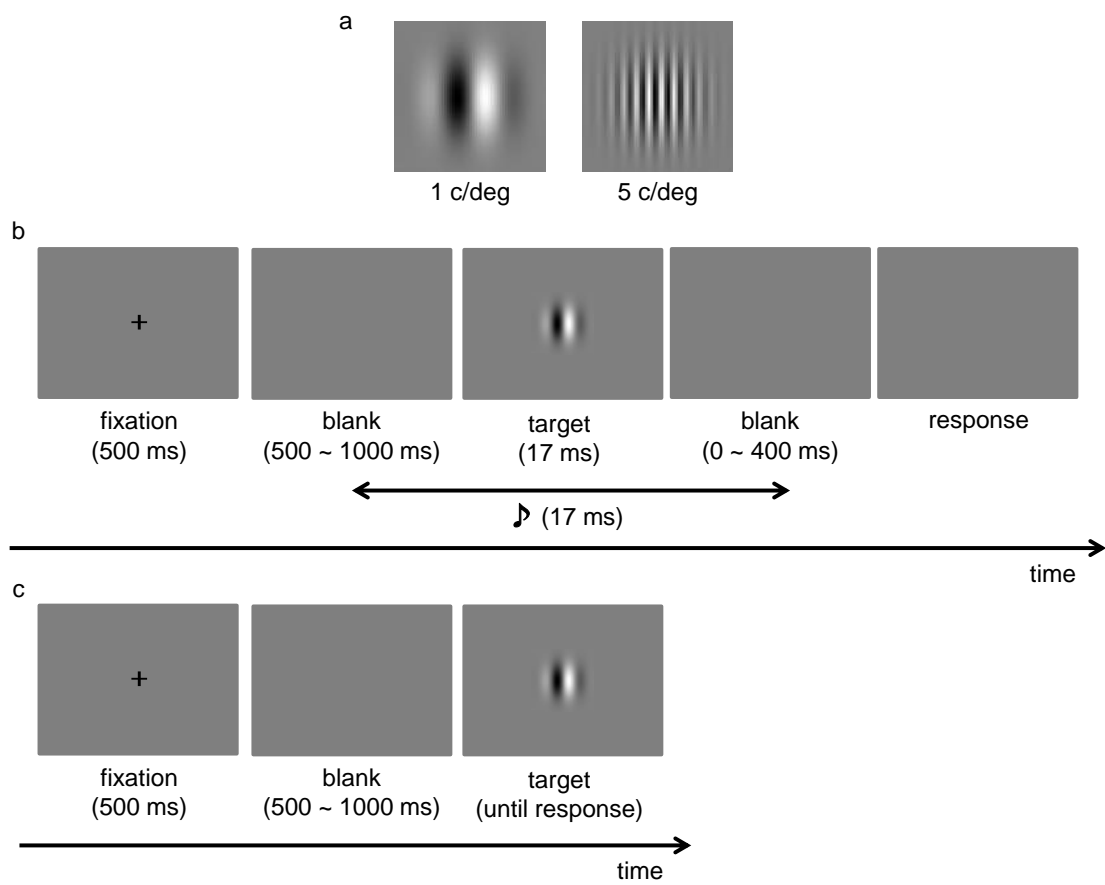


図 3-1. (a) 研究 2 で使用したガボールパッチ。(b) 視覚刺激と聴覚刺激の同期判断課題の 1 試行の流れ。(c) ガボールパッチに対する反応時間計測課題の 1 試行の流れ (Takeshima and Gyoba, submitted より転載)。

3-2-3 結果

視覚刺激と聴覚刺激が同期と判断された割合を算出した（図 3-2a）。PSS を算出するために、Excel ソルバー（Microsoft）を使用して、二乗誤差（root-mean-square-error: RMSE）が最小になるように各参加者のデータに下記の数式に基づくガウス関数の当てはめを行った（Santangelo & Spence, 2008; Van der Burg et al., 2008 を参照）。

$$P(\text{response}|\text{SOA}) = \text{blink rate} + a \cdot e^{\left[-.5\left(\frac{\text{SOA}-\text{PSS}}{b}\right)^2\right]}$$

「blink rate」は、データの中に混在するノイズ（例えば、瞬き）を示すパラメーターで、他のパラメーターを安定して推定するために用いられる（Swanson & Birch, 1992）。SOA は実験条件と同様の値（-350 ~ +350）を使用し、blink rate は 0 ~ 2.5 % の範囲内、その他のパラメーターは 0 以上になるように推定値を制限した。

関数の当てはめを行ったところ、blink rate は 1 c/deg では 1.5 %、5 d/deg では 2.1 %、RMSE は 1 c/deg では 0.10 ($SD = 0.06$)、5 d/deg でも 0.10 ($SD = 0.06$) であった。算出された同期判断率の値に対して、Spatial frequency (2) × SOA (9) の 2 要因分散分析を行ったところ、SOA の主効果が有意で ($F(8, 64) = 122.87, p < .001, \eta_p^2 = .93$)、SOA の値が 0 に近いほど同期判断率が高くなっていた。また、Spatial frequency × SOA の交互作用も有意であった ($F(8, 64) = 3.84, p < .001, \eta_p^2 = .33$)。2 要因の交互作用が有意であったため単純主効果検定を行ったところ、SOA が +233 ms ($F(1, 72) = 7.39, p < .05, \eta_p^2 = .09$) と +350 ms ($F(1, 72) = 5.94, p < .05, \eta_p^2 = .08$) の条件において Spatial frequency の単純主効果が有意であり、5 c/deg の方が 1 c/deg と比べて同期判断率が高くなっていた。一方、Spatial frequency の主効果は有意ではなかった ($F(1, 8) = 0.35, p = .57, \eta_p^2 = .04$)。

さらに、推定した PSS の平均値を算出した（図 3-2b）。PSS を空間周波数間で比較するために、両側 t 検定を行ったところ、5 c/deg の方が 1 c/deg と比べて PSS の値が有意に大きくなっていた ($t(8) = 2.82, p < .05, d = .23$)。95 % 信頼区間を算出したところ、1 c/deg 条件では 74.53 ms ~ 112.12 ms で、5 c/deg 条件では 80.90 ms ~ 119.35 ms であった。

最後に、反応時間計測課題における反応時間の平均値を算出した（図 3-2c）。空間周波数間で比較するために、両側 t 検定を行ったところ、5 c/deg の方が 1 c/deg と比べて反応時間が有意に遅くなっていた ($t(8) = 2.82, p < .05, d = .30$)。

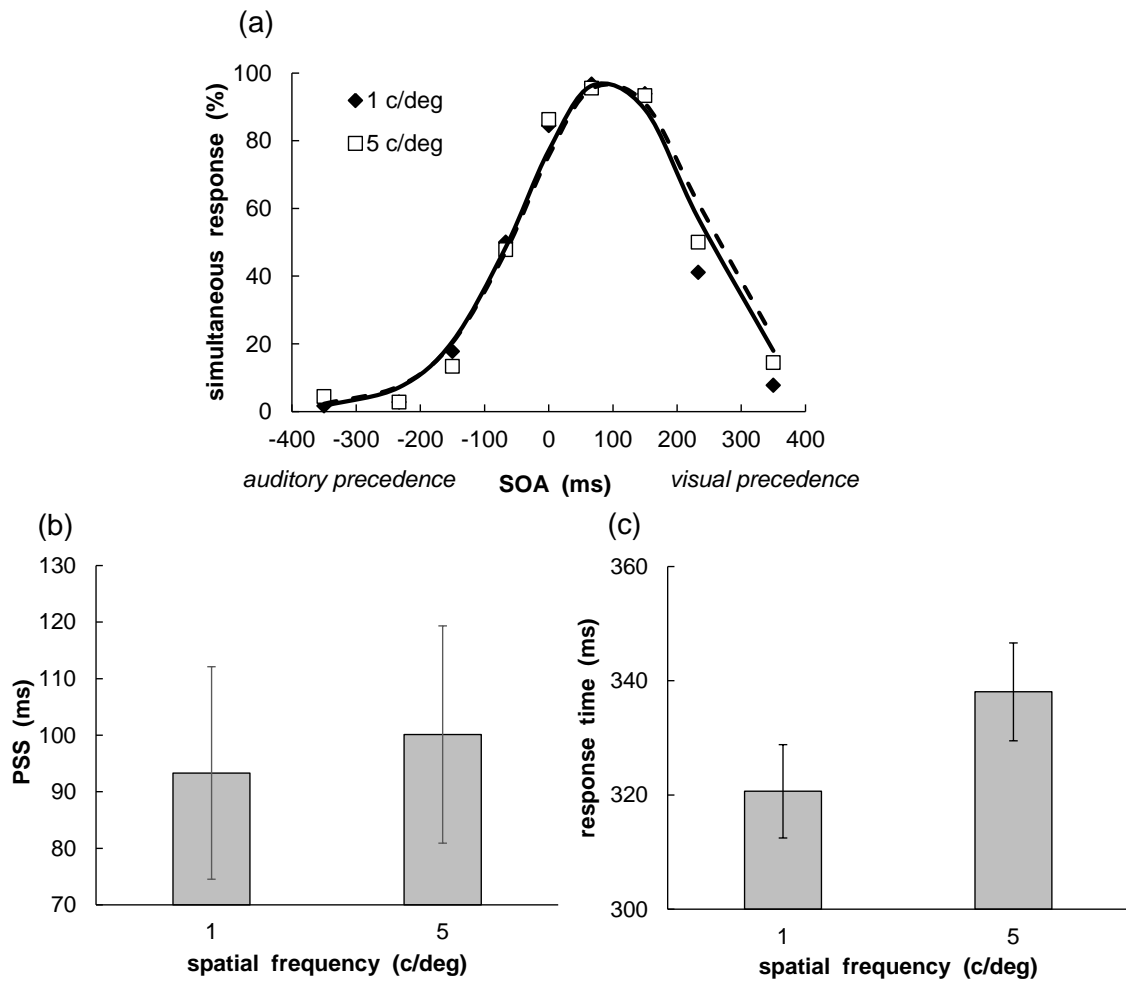


図 3-2. 研究 2 の結果。(a) 視覚刺激と聴覚刺激が同期していると判断された割合。(b) 同期判断率から算出された主観的同期点の結果。エラーバーは 95%信頼区間を示す。(c) ガボールパッチに対する反応時間計測の結果。エラーバーは標準誤差を示す (Takeshima and Gyoba, submitted より転載)。

3-2-4 考察

研究 2 では、視覚刺激の空間周波数を操作することにより、視聴覚間の同期判断に与える影響について検討を行った。実験の結果、空間周波数が高い視覚刺激では、低い視覚刺激と比べて PSS の値が大きくなっていた。また、95 % 信頼区間推定から同期の時間窓を推定したところ、空間周波数が高い視覚刺激の方では低い視覚刺激よりも時間窓がプラス方向にずれていた。PSS が 0 になる場合、視聴覚間の物理的な同期と主観的な同期判断が一致すると考えられる。対して、PSS がマイナスになれば聴覚刺激が先行する場合に同期と判断され、プラスになれば視覚刺激が先行する場合に同期と判断されると考えられる。本研究では、どちらの空間周波数のガボールパッチでも PSS の値はプラスになっており、視覚刺激が先行する場合に聴覚刺激と同期していると判断されていた。日常生活においては、光と音の伝達速度の違いから視覚入力聴覚入力よりも早く入力が行われる。また、視覚ニューロンと聴覚ニューロンでは反応潜時が異なる (King, 2010)。このような視覚と聴覚の違いを補正する機能がはたらくため (Vroomen & Keetels, 2010)、視覚刺激と聴覚刺激の PSS を測定すると、視覚刺激先行に値が偏る場合が多い (e.g., Kayser, Petkov, & Logothetis, 2008; Lewald & Guski, 2003; Zampini et al., 2005)。本研究の結果も、これらの先行研究を再現していた。

一方、空間周波数が高いガボールパッチの方が PSS の値が有意に大きくなっていたことから、空間周波数による違いも生じていた。高空間周波数に対して反応する持続型チャンネルは、低空間周波数に対して反応する一過型チャンネルと比べて信号の伝導速度が遅い (Hoffman, 1973) ため、高空間周波数の視覚刺激に対しては視覚処理が遅くなる。これは、Breitmeyer (1975) や本研究における反応時間の計測からも裏付けられている。したがって、視覚処理の遅延によって、視聴覚刺激の時間的同期のためには視覚刺激と聴覚刺激の提示タイミングの物理的なずれを大きくする必要があると考えられる。つまり、視覚刺激の処理速度を考慮した同期の時間窓の補正が働いているといえる。

以上のように、視覚刺激の空間周波数による処理速度の違いは同期知覚を変容させることが明らかになり、さらには視聴覚情報の統合過程にも影響が波及することが示唆された。

3-3 研究 3：視覚の空間周波数処理特性が分裂錯覚に与える影響

3-3-1 目的

研究 2 では、処理速度が遅い高空間周波数の視覚刺激は、低空間周波数の刺激と比べて PSS の値が有意に大きくなっていた。つまり、視聴覚刺激が同期していると判断されるためには、聴覚刺激に対して視覚刺激をより先行して提示する必要があることを示している。したがって、視覚刺激の空間周波数が高くなると、視覚刺激と聴覚刺激が物理的に同期していても主観的な同期は損なわれやすくなる。研究 3 では、視覚刺激の空間周波数による主観的な同期知覚への影響が、視聴覚相互作用にまで及ぶかについて検討を行った。実験では、視聴覚間の同期という時間的な側面について検討するため、短い時間間隔の中で視聴覚統合を必要とする視聴覚相互作用である分裂錯覚を利用した。なお、ここで報告するデータは、Takeshima and Gyoba (in press) に基づいている。

3-3-2 方法

実験参加者：東北大学の学生 13 名（男性 8 名、女性 5 名）が本実験に参加した。実験参加者は全員、正常な視力（矯正視力を含む）と聴力を有していた。また、参加者全員が実験の目的について知らされていなかった。

装置：実験制御用の PC（XPS720, Dell; OS：Windows Vista, Microsoft; 解像度：1024 × 768 pixels; リフレッシュレート：60 Hz）を使用し、視覚刺激は CRT ディスプレイ（Trinitron GDM-F520, Sony）上に提示し、聴覚刺激は聴覚デバイス（Edirol FA-66, Roland）とヘッドホン（HDA200, Sennheiser）を通して提示した。刺激の制御には、MATLAB (MathWorks, Inc.) と実験用ライブラリ (Cogent Graphics and 2000 toolbox) を利用した。視覚刺激と聴覚刺激は、デジタルオシロスコープ (TS-80600, Iwatsu) を使用して開始が同期していることを確認した。実験は案室内で行い、暗騒音は 43.6 dB SPL (A) であり、視距離はアゴ台を使用して 60 cm に固定した。実験参加者の応答は、キーボードのテンキーを使用して行った。

刺激：赤色 (15.3 cd/m^2) の注視点（直径約 1.1 deg）と視覚刺激を、灰色 (17.9 cd/m^2) の背景に提示した。視覚刺激は、空間周波数が 1 c/deg と 5 c/deg のガボールパッチであった。ガボールパッチの大きさは直径 2.0 deg であった。聴覚刺激は純音（立ち上り/立ち下り時

間：1.5 ms, サンプリング周波数：22050 Hz) で、周波数は 3500 Hz, 音圧は 80 dB (A) であった。視覚刺激は、各試行で 1 回もしくは 2 回提示され、各提示時間は 17 ms であった。一方、聴覚刺激は、提示されない (no-beep), 1 回目の視覚刺激と同じタイミングで 1 回提示 (1-beep), 2 回の視覚刺激と同じタイミングで 2 回提示 (2-beeps) の 3 水準であった。聴覚刺激の 1 回の提示時間は 15 ms であった。2 回の刺激の SOA は、視覚刺激、聴覚刺激共に 50 ms であった。

手続き：各試行は、「0」キーを押すことによって開始された。まず、注視点が 500 ms 提示された後、視覚刺激が 1 回もしくは 2 回提示された (図 3-3)。実験条件は、Flash (2; 1-flash or 2-flashes) \times Spatial frequency (1 c/deg or 5 c/deg) \times Beep (3; no-beep, 1-beep, or 2-beeps) で、実験参加者は各条件を 20 回繰り返して、計 240 試行を行った。実験参加者の課題は、視覚刺激の提示回数が 1 回か 2 回かを、強制二肢選択法で応答することであった。

3-3-3 結果

信号検出理論 (Macmillan & Creelman, 2005) に基づいて、視覚刺激の提示回数の弁別感度 (d') と判断基準 (c) を算出した (図 3-4a-b)。検出感度は $d' = z(H) - z(F)$ の式により、判断基準は $c = -0.5 * (z(F) + z(H))$ の式によって算出される。ここで、 $z(p)$ は実験参加者の反応の割合 p の逆正規分布変換を行った値、 H と F はそれぞれ「hit」(2 回の視覚刺激の正しい検出) と「False-alarm」(2 回の視覚刺激の誤った検出) の値を示している。実験参加者が視覚刺激の提示回数を正しく弁別できていた場合には、 d' 値は高くなる。加えて、実験参加者の応答にバイアスが生じていない場合には、 c 値は 0 になる。一方、 c 値が正の値になった場合には 2 回と応答しやすいというバイアスが、負の値になった場合には 1 回と応答しやすいというバイアスが生じていたことを示す。また、分裂錯覚が生じる場合には、弁別感度が低くなる (Rosenthal, Shimojo, & Shams, 2009)。ゆえに、2-beeps 条件の d' 値が高かった (3.0 以上) 3 名の参加者のデータについては、分裂錯覚が生じていなかったとして以降の分析から除外した。

算出された弁別感度を従属変数として、Spatial frequency (2) \times Beep (3) の 2 要因分散分析を行ったところ、Beep の主効果が有意であった ($F(1, 9) = 32.55, p < .01, \eta_p^2 = .78$)。また、Spatial frequency \times Beep の交互作用も有意であった ($F(2, 18) = 5.80, p < .05, \eta_p^2 = .39$)。2 要因の交互作用が有意であったので単純主効果検定を行ったところ、各 Spatial frequency 条

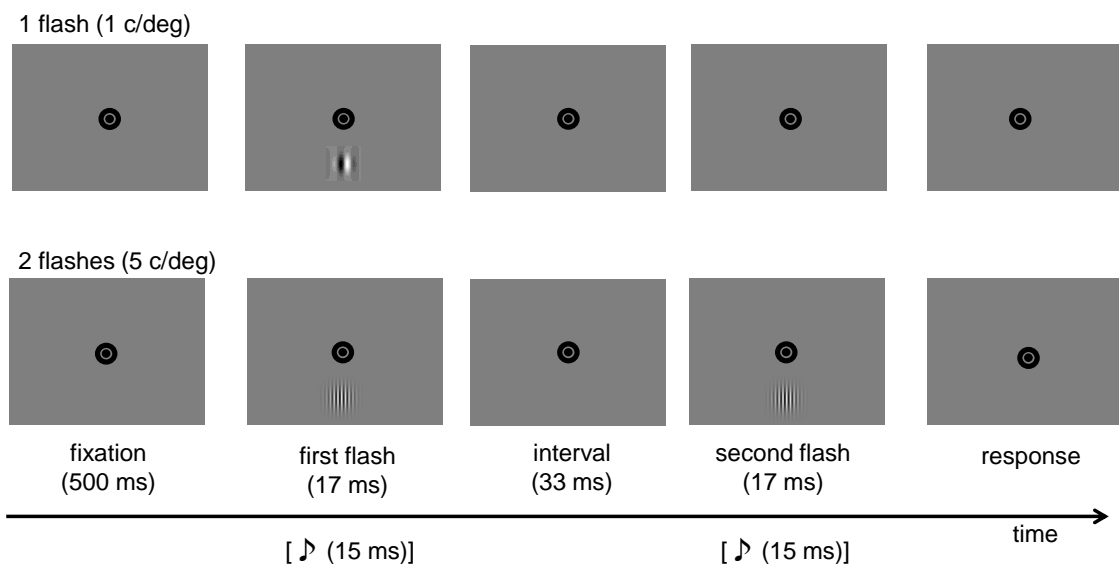


図 3-3. 研究 3 の 1 試行の流れ。上段は視覚刺激の提示回数が 1 回の場合，下段は提示回数が 2 回の場合を示す。(Takeshima and Gyoba (in press) より転載。)

件における Beep の単純主効果が有意であった (1 c/deg: $F(2, 36) = 34.23, p < .01, \eta_p^2 = .66$; 5 c/deg: $F(2, 36) = 21.89, p < .01, \eta_p^2 = .55$)。多重比較 (シェファー法) を行ったら、いずれの空間周波数でも 2-beep 条件の d' 値が no-beep, 1-beep 条件と比べて低くなっていた ($p < .01$)。一方, no-beep と 1-beep との差は, 1 c/deg では有意ではなかったが ($p = .25$), 5 c/deg では有意であった ($p < .01$)。また, 2-beeps 条件における Spatial frequency の単純主効果も有意であり ($F(1, 27) = 20.76, p < .01, \eta_p^2 = .43$), 5 c/deg 条件の d' 値が 1 c/deg 条件と比べて有意に高くなっていた。Spatial frequency の主効果は有意ではなかった ($F(1, 9) = 0.28, p = .61, \eta_p^2 = .03$)。

次に, 算出された判断基準を従属変数として, Spatial frequency (2) \times Beep (3) の 2 要因分散分析を行ったら、Spatial frequency の主効果が有意であり ($F(1, 9) = 9.31, p < .05, \eta_p^2 = .51$), 1 c/deg の c 値が 5 c/deg と比べて高くなっていた。また, Beep の主効果も有意であった ($F(2, 18) = 23.97, p < .01, \eta_p^2 = .73$)。多重比較を行ったら、1-beep 条件の c 値が no-beep 条件よりも低くなっており ($p < .05$), また, 2-beeps 条件の c 値が no-beep 条件よりも高くなっていた ($p < .01$)。しかし, Spatial frequency \times Beep の交互作用は有意ではなかった ($F(2, 18) = 1.40, p = .27, \eta_p^2 = .13$)。

最後に, 信号検出理論による分析の確認のため, 視覚刺激の提示回数が 2 回と判断された割合を算出した (図 3-4c-d)。Flash (2) \times Spatial frequency (2) \times Beep (3) の 3 要因分散分析を行ったら、Flash \times Spatial frequency \times Beep の交互作用が有意であった ($F(2, 18) = 5.05, p < .05, \eta_p^2 = .36$)。3 要因の交互作用が有意であったので, 視覚刺激の空間周波数が分裂錯覚の生起頻度に与える影響を検討するため, 視覚刺激の提示回数別に Spatial frequency (2) \times Beep (3) の 2 要因分散分析を行った。視覚刺激の提示回数が 1 回の条件では, Spatial frequency ($F(1, 9) = 7.81, p < .05, \eta_p^2 = .46$) と Beep ($F(2, 18) = 29.39, p < .001, \eta_p^2 = .77$) の主効果が有意であった。さらに, Spatial frequency \times Beep の交互作用が有意であったため ($F(2, 18) = 6.36, p < .01, \eta_p^2 = .41$), 単純主効果の検定を行ったら、それぞれの空間周波数で Beep の単純主効果が有意であった (1 c/deg: $F(2, 36) = 34.59, p < .01, \eta_p^2 = .66$; 5 c/deg: $F(2, 36) = 22.73, p < .01, \eta_p^2 = .56$)。多重比較を行ったら、いずれの空間周波数においても, 2-beeps 条件の割合が no-beep 条件や 1-beep 条件と比較して有意に高くなっていた ($p < .001$)。一方, no-beep 条件と 1-beep 条件の割合には有意な差は見られなかった ($p > .19$)。さらに, 2-beeps 条件において spatial frequency の単純主効果が有意であり ($F(1, 27) = 15.00, p < .01, \eta_p^2 = .36$), 5 c/deg の割合が 1 c/deg と比較して有意に低くなっていた。視覚刺激の

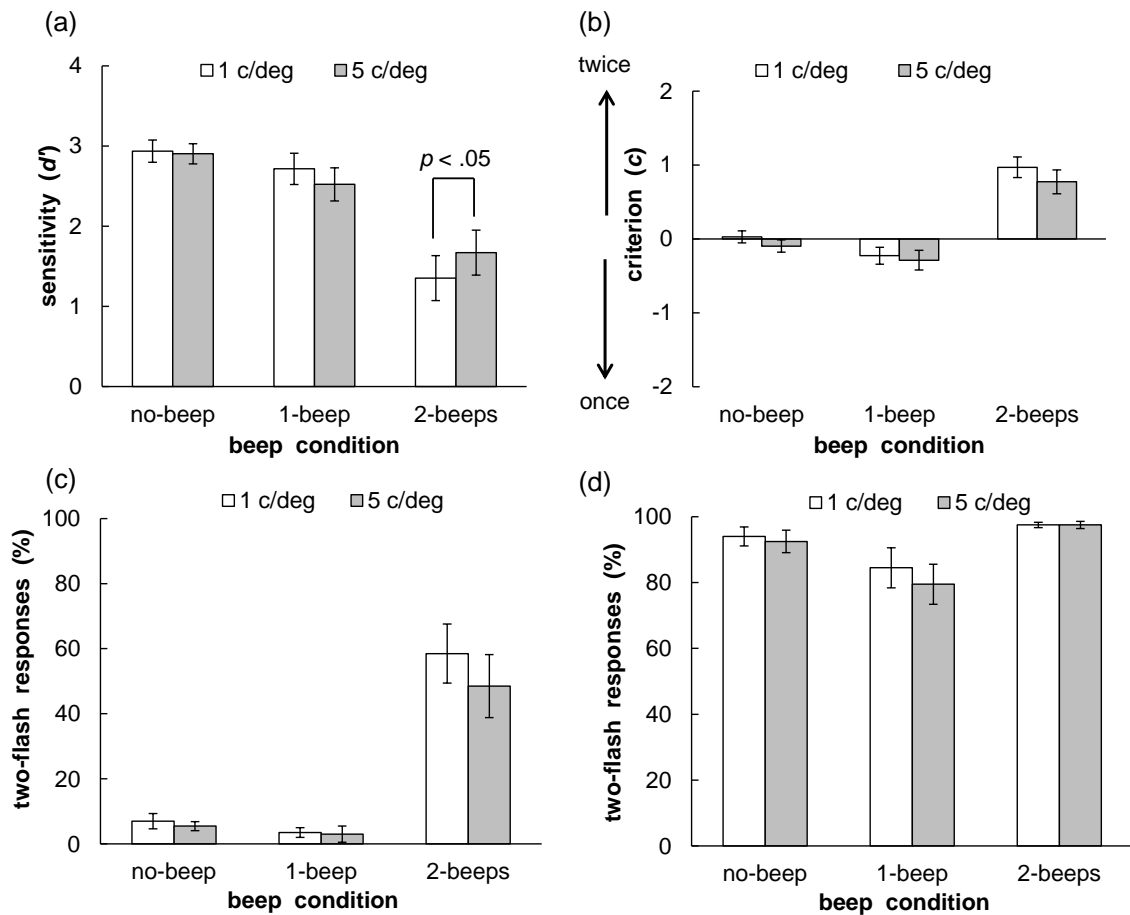


図 3-4. 研究 3 の結果。(a) 視覚刺激の提示回数の弁別感度。(b) 視覚刺激の提示回数の判断基準。正の値は参加者の判断が 2 回、負の値は 1 回にバイアスがかかっていることを示す。(c) 視覚刺激の提示回数が 1 回の時に 2 回と判断された割合。(d) 視覚刺激の提示回数が 2 回の時に 2 回と判断された割合。エラーバーは標準誤差を示す。(Takeshima and Gyoba (in press) より転載。)

提示回数が 2 回の条件では, Beep の主効果が有意であった ($F(2, 18) = 7.01, p < .01, \eta_p^2 = .43$)。多重比較を行ったところ, 1-beep 条件の割合が no-beep 条件や 2-beeps 条件と比較して有意に低くなっていた ($ps < .05$)。一方, no-beep 条件と 2-beeps 条件との差は有意ではなかった ($p = .19$)。Spatial frequency の主効果 ($F(1, 9) = 2.62, p = .14, \eta_p^2 = .23$) と Spatial frequency \times Beep の交互作用 ($F(2, 18) = 1.45, p = .26, \eta_p^2 = .14$) は有意ではなかった。

3-3-4 考察

研究 3 では, 視覚の空間周波数処理特性が, 視聴覚統合によって生じる分裂錯覚の生起頻度に与える影響の検討を行った。分析では, 信号検出理論を用いて反応バイアスを考慮した分析と, 視覚刺激の提示回数別に分けて錯覚の生起頻度を算出した分析の両面から検討を行った。視覚刺激の提示回数の弁別感度を算出したところ, どちらの空間周波数でも 2 回の聴覚刺激提示によって他の条件よりも弁別感度が低くなっていた。分裂錯覚が生じる場合には弁別感度が低くなる (Rosenthal et al., 2009) ことから, ガボールパッチを使用した本実験でも分裂錯覚が生じていた。視覚刺激の提示回数が 1 回の条件において, 2 回の聴覚刺激提示によって提示回数が 2 回と判断される割合が高くなっていたことから, 本実験において分裂錯覚が生起していたと考えられる。一方, 空間周波数の高い視覚刺激では低い刺激と比べて弁別感度が高く, 提示回数が 2 回と判断された割合が低くなっていた。この結果より, 高空間周波数の視覚刺激では分裂錯覚が生じにくいことが示された。ただし, 判断基準では聴覚刺激を 2 回提示した場合でも空間周波数間で有意な差が生じていなかったことから, 視覚刺激の空間周波数による分裂錯覚の生起頻度の違いは, 実験参加者の反応バイアスによるものではない。したがって, 分裂錯覚の生起頻度の違いは, 視覚の空間周波数処理の特性によるものと考えられる。高空間周波数に対して反応する持続型チャンネルは, 信号の伝導速度が遅い (Hoffmann, 1973) ため, 高空間周波数の視覚刺激に対しては視覚処理が遅くなる。研究 2 でも示したように, 視覚処理が遅くなると時間的同期知覚のためには, 視聴覚刺激の提示タイミングを大きくずらす必要がある。したがって, 高空間周波数の視覚刺激では同期知覚が生じにくくなることから, 分裂錯覚の生起頻度が低くなると推測される。また, 持続型チャンネルは, 低空間周波数に対して反応する一過型チャンネルと比べて反応の立ち上がり (潜時) が遅い (Breitmeyer, 1975)。分裂錯覚は非常に短い時間間隔の中で生じる錯覚であるため, 潜時が遅いと 2 回目の聴覚刺激の提示時にもまだ 1 回目の視覚刺激への反応が持続し, 錯覚となる 2 回目の視覚刺激が知覚されに

くくなると考えられる。したがって、処理速度の違いを含めた視覚の空間周波数処理特性が視聴覚情報の統合過程に影響を及ぼし、分裂錯覚の生起頻度に違いを生じさせることが示された。

3-4 総合考察

本章では、視覚刺激の空間周波数が高くなることによる視覚処理速度の遅延が視聴覚情報の統合過程に与える影響について、研究2と研究3の2つの研究結果の報告を行った。

研究2では、視覚刺激の空間周波数が高くなることによって、視聴覚刺激の同期判断における PSS がプラス側（視覚刺激先行）にシフトすることが示された。これは、視覚刺激の空間周波数が高くなることによって処理速度が遅延し、同期の知覚のためには聴覚刺激に対して視覚刺激を先行して提示する必要があることを示唆している。すなわち、視覚刺激の処理速度の影響を受けて、聴覚刺激との同期の時間窓を柔軟に補正する仕組みが備わっていると考えられる。同時に、視聴覚刺激を物理的に同期して提示した場合には、空間周波数の高い視覚刺激ではより同期の知覚が損なわれやすくなる。

研究3では、空間周波数が高い視覚刺激を用いた場合、低い刺激と比べて分裂錯覚が生じにくいことが示された。分裂錯覚は、非常に短い時間間隔の中での視聴覚情報の統合を必要とする錯覚である。研究2より、視覚刺激の空間周波数が高くなると物理的に同期していても視聴覚刺激間の同期の知覚が損なわれやすい。したがって、2回目の聴覚刺激の提示までに1回目の視聴覚情報の統合が生じにくいために、分裂錯覚の生起頻度が低くなる。また、空間周波数の高い視覚刺激に対して反応する持続型チャンネルは反応の潜時が長い。そのため、2回目の聴覚刺激提示時に1回目の視覚入力に対する反応が持続してしまっていることも関連している。

視覚処理では、空間周波数が高い刺激に対する処理速度が低い刺激と比べて遅くなる。この処理速度の違いによって、空間周波数が高い視覚刺激では聴覚刺激との同期の知覚が損なわれやすくなり、それにより視聴覚相互作用も生じにくくなることが明らかとなった。

第 4 章 高次過程における視覚刺激特性の操作が視聴覚統合に与える影響の検討

第 3 章で議論したように、視覚刺激の空間周波数による処理速度の違いは、視聴覚刺激間の同期知覚や視聴覚相互作用の生起頻度に影響を与える。一方、空間周波数は視覚処理の初期段階で処理されるため、視覚処理の低次過程における処理速度の規定因といえる。1-2 で議論したように、視聴覚情報の統合は感覚処理の特定の段階で生じるのではなく、視覚処理と聴覚処理が並列で行われる中の様々な段階で生じることが示唆されている。したがって、感覚刺激の処理速度の影響も空間周波数処理のような視覚処理の低次過程の処理段階だけでなく、より高次過程の処理段階においても生じると推測される。第 4 章では、視覚処理のより高次過程における処理速度の規定因による視聴覚統合への影響について議論を行う。

4-1 研究 4：視覚刺激の複雑さが分裂錯覚の生起頻度に与える影響

4-1-1 目的

視覚処理における空間周波数よりも高次過程で処理速度を規定する要因の 1 つとして、物体（視覚刺激）の複雑さがあげられる (Markovic & Gvozdenovic, 2001)。物体の複雑さを処理する領域としては、前下側頭皮質 (anterior inferotemporal cortex)、後下側頭皮質 (posterior inferotemporal cortex)、V4 が知られている (Kobatake & Tanaka, 1994)。これらの領域は、視覚処理においては空間周波数を処理する外側膝状体や V1 よりも高次過程の領域である。

物体の複雑さは、一般的には質的な変数である。しかし、エントロピーや冗長度といった情報理論によって量的に定義することが可能である (Attneave, 1954)。Garner and Clement (1963) は、冗長度を用いて同等集合サイズ (Equivalent set size: ESS) という概念を提案し、複雑さを量的に操作することが可能な視覚パターンを作成している。彼らは、 3×3 の仮想マトリクス上に 5 個のドットを配置して視覚パターンを作成している。ESS は、これらのパターンに回転変換と鏡映変換を行うことで心的に推測されるパターンの総数によって複雑さを定義している。総数が少なければ複雑さが低く、総数が多くなるほど複雑さが高いパターンとなる。回転変換や鏡映変換を行うため、ESS によって定義される複雑さには、パターンの対称性なども関連している (Garner & Clement, 1963)。

この ESS の概念を利用して、複雑さによる視覚処理の処理速度について検討が行われている。例えば、視覚探索課題では複雑なパターンの方が単純なパターンと比べて探索時間が長くなる (e.g., Makovski & Jiang, 2008; Markovic & Gvozdenovic, 2001; Rauschenberger & Yantis, 2006)。また、複雑なパターンでは単純なパターンよりも異同判断の反応時間も長くなる (e.g., Garner & Sutliff, 1974; Howe, 1980; Markovic & Gvozdenovic, 2001)。加えて、複雑なパターンの方が、符号化効率が低い (Clement & Vanadoe, 1967; Garner & Sutliff, 1974; Rauschenberger & Yantis, 2006) ことや、形の再認課題でも複雑なパターンの方がより時間がかかる (Kayaert & Wagemans, 2009) ことが報告されている。したがって、複雑な物体の方が単純な物体よりも処理速度が遅いと考えられる。そこで、研究 4 では視覚刺激の複雑さを ESS の概念を用いて統制し、視聴覚相互作用の生起頻度に与える影響について検討を行った。なお、ここで報告するデータは、Takeshima and Gyoba (2013c) に基づいている。

実験 1 では、複雑さの異なる視覚パターンを視覚刺激として使用し、分裂錯覚の生起頻度の比較を行った。視覚刺激の複雑さによる処理速度の違いが空間周波数と同様の影響を与えるとすれば、複雑な視覚刺激では単純な刺激と比べて分裂錯覚が生じにくいと予測される。

実験 2 では、実験 1 で統制できなかったパターンの connectivity の統制を行い、再度、視覚刺激の複雑さが分裂錯覚の生起頻度に与える影響について検討を行った。実験 1 の結果がパターンの connectivity ではなく複雑さによって生じたのであれば、connectivity を統制しても実験 1 の結果を再現できると予測される。

実験 3 では、視覚刺激の複雑さによる処理速度の違いが分裂錯覚の生起頻度に影響を与えることを確認するため、視覚刺激に同期する 2 音の SOA を操作した。分裂錯覚の生起頻度が視覚処理の処理速度の影響を受けるとすれば、2 音の SOA によって複雑さによる影響が異なると予測される。

4-1-2 実験 1：視覚刺激の複雑さの影響

実験参加者：実験者 1 名を含む東北大学の学生 7 名（男性 2 名、女性 5 名）が実験 1 に参加した。実験参加者は全員、正常な視力（矯正視力を含む）と聴力を有していた。また、実験者 1 名を除き、実験の目的について知らされている者はいなかった。

装置：実験制御用の PC (XPS720, Dell; OS: Windows Vista, Microsoft; 解像度：1024 × 768

pixels; リフレッシュレート:60 Hz) を使用し, 視覚刺激は CRT ディスプレイ (Diamondtron M² RDF223G, Mitsubishi) 上に提示し, 聴覚刺激は聴覚デバイス (Edirol FA-66, Roland) とヘッドホン (HDA200, Sennheiser) を通して提示した。刺激の制御には, MATLAB (MathWorks, Inc.) と実験用ライブラリ (Cogent Graphics and 2000 toolbox) を利用した。視覚刺激と聴覚刺激は, デジタルオシロスコープ (TS-80600, Iwatsu) を使用して開始が同期していることを確認した。実験は案室内で行い, 暗騒音は 39.1 dB SPL (A) であり, 視距離はアゴ台を使用して 57 cm に固定した。実験参加者の応答は, キーボードのテンキーを使用して行った。

刺激: 灰色 (24.4 cd/m²) の背景に, 黒色 (0.6 cd/m²) の注視点と白色 (105.0 cd/m²) の視覚パターンを提示した。視覚パターンは, Garner and Clement (1963) で用いられているドットパターンのドットを四角 (0.7 × 0.7 deg) に置き換え, ESS4 と ESS8 のパターンを 6 種類ずつ作成した (図 4-1)。同等集合サイズ概念より, ESS4 のパターンが単純な視覚刺激, ESS8 のパターンを複雑な視覚刺激とした。視覚パターンは注視点の下に提示し, 注視点から視覚パターンの中心までの距離は 7.0 deg であった。聴覚刺激は純音 (立ち上り/立ち下り時間: 1.0 ms, サンプリング周波数: 22050 Hz) で, 周波数は 3500 Hz, 音圧は 95 dB (A) であった。視覚刺激は, 各試行で 1 回もしくは 2 回提示され, 各提示時間は 17 ms であった。一方, 聴覚刺激は, 提示されない (no-beep), 1 回目の視覚刺激と同じタイミングで 1 回提示 (1-beep), 2 回の視覚刺激と同じタイミングで 2 回提示 (2-beeps) の 3 水準であった。聴覚刺激の 1 回の提示時間は 10 ms であった。2 回の刺激の SOA は, 視覚刺激, 聴覚刺激共に 50 ms であった。

手続き: 各試行は, 「0」キーを押すことによって開始された。注視点に 1000 ms 提示された後, 視覚刺激が 1 回もしくは 2 回提示された (図 4-2)。実験条件は, Flash (2; 1-flash or 2-flashes) × ESS (2; ESS4 or ESS8) × Beep (3; no-beep, 1-beep, or 2-beeps) で, 実験参加者は各条件を 24 回 (各視覚パターンを 4 回ずつ) 繰り返して, 計 288 試行を行った。実験参加者の課題は, 視覚刺激の提示回数が 1 回か 2 回かを強制二肢選択法で応答することであった。

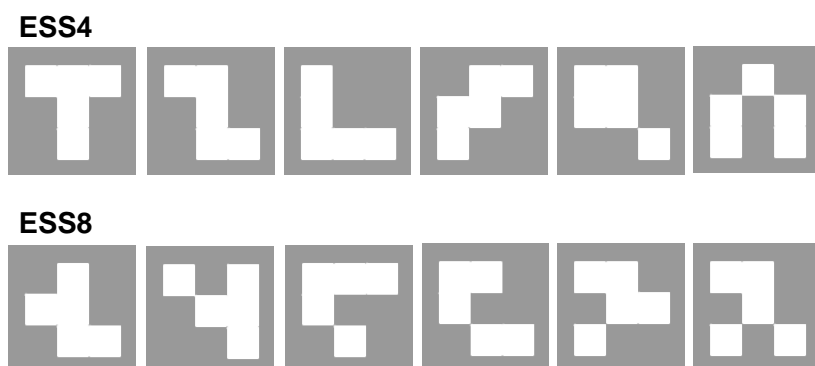


図 4-1. 研究 4 で使用した視覚パターン。上段は ESS4 の単純なパターンを，下段は ESS8 の複雑なパターンを示す。ESS4 のパターンは，回転変換と鏡映変換によって 1 つのパターンにつき 4 種類のパターンが，ESS8 のパターンは 8 種類のパターンが内的に推測される。
(Takeshima and Gyoba (2013c) より一部改変。)

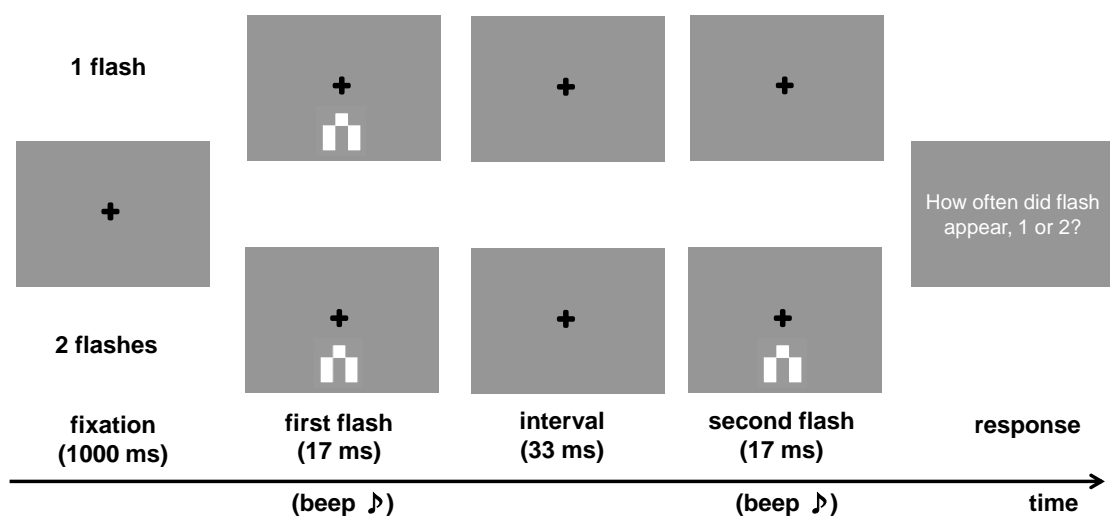


図 4-2. 実験 1 の 1 試行の流れ。上段は視覚刺激の提示回数が 1 回の場合を，下段は 2 回の場合を示す。(Takeshima and Gyoba (2013c) より一部改変。)

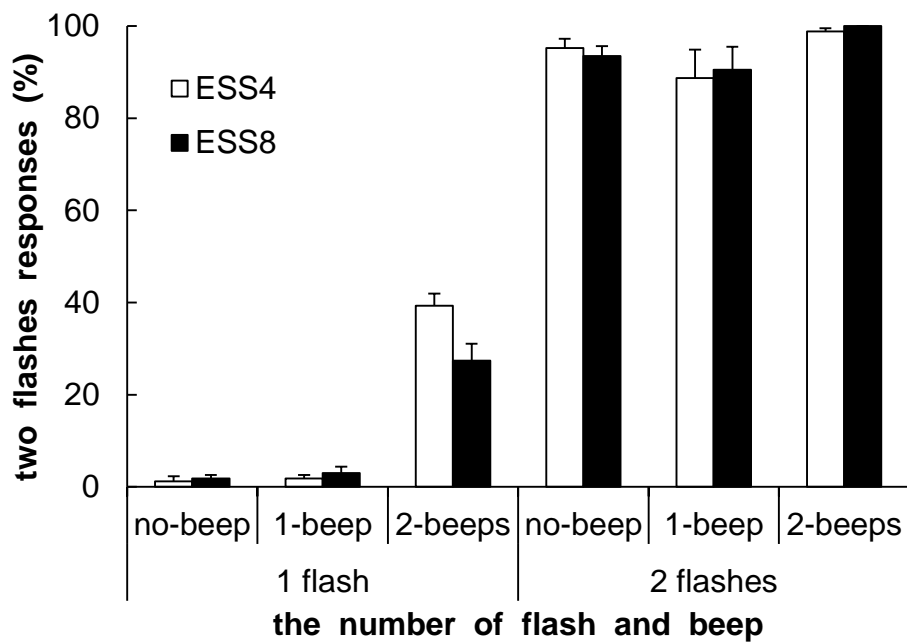


図 4-3. 実験 1 で視覚刺激の提示回数が 2 回と判断された割合。エラーバーは標準誤差を示す。(Takeshima and Gyoba (2013c) より転載。)

結果：視覚刺激の提示回数が2回と判断された割合を算出した（図4-3）。Flash (2) × ESS (2) × Beep (3) の3要因分散分析を行ったところ、Flashの主効果 ($F(1, 6) = 1531.18, p < .001, \eta_p^2 = .99$) と Beepの主効果 ($F(2, 12) = 33.22, p < .001, \eta_p^2 = .87$) が有意であった。一方、ESSの主効果は有意ではなかった ($F(1, 6) = 2.89, p = .14, \eta_p^2 = .32$)。また、Flash × ESSの交互作用 ($F(1, 6) = 7.85, p < .05, \eta_p^2 = .56$)、Flash × Beepの交互作用 ($F(2, 12) = 12.12, p < .005, \eta_p^2 = .67$)、ESS × Beepの交互作用 ($F(2, 12) = 8.42, p < .01, \eta_p^2 = .58$) も有意であった。さらに、Flash × ESS × Beepの交互作用も有意であった ($F(2, 12) = 8.10, p < .01, \eta_p^2 = .57$)。3要因の交互作用が有意であったので、視覚刺激の提示回数別にESS (2) × Beep (3) の2要因分散分析を行った。視覚刺激が1回提示では、ESSの主効果 ($F(1, 6) = 8.99, p < .01, \eta_p^2 = .60$) と Beepの主効果 ($F(2, 12) = 99.44, p < .001, \eta_p^2 = .94$) が有意であった。また、ESS × Beepの交互作用も有意であった ($F(2, 12) = 14.78, p < .01, \eta_p^2 = .71$)。2要因の交互作用が有意であったため、単純主効果検定を行ったところ、各ESS条件におけるBeepの単純主効果が有意であった (ESS4: $F(2, 24) = 112.55, p < .001, \eta_p^2 = .90$; ESS8: $F(2, 24) = 49.31, p < .001, \eta_p^2 = .80$)。多重比較（ライアン法）を行ったところ、いずれのESS条件でも2-beep条件の割合がno-beep、1-beep条件と比べて高くなっていた ($ps < .001$)。一方、いずれのESSにおいても、no-beepと1-beepとの差は有意ではなかった ($ps > .21$)。また、2-beeps条件におけるESSの単純主効果も有意であり ($F(1, 18) = 37.90, p < .001, \eta_p^2 = .68$)、ESS8条件の割合がESS4条件と比べて有意に低くなっていた。視覚刺激が2回提示ではESSの主効果 ($F(1, 6) = 0.14, p = .72, \eta_p^2 = .23$)、Beepの主効果 ($F(2, 12) = 2.12, p = .16, \eta_p^2 = .26$)、ESS × Beepの交互作用 ($F(2, 12) = 1.08, p = .37, \eta_p^2 = .15$) のいずれも有意ではなかった。

4-1-3 実験2：Connectivityの統制

実験参加者：実験者1名を含む東北大学の学生7名（男性2名、女性5名）が実験2に参加した。7名のうち3名は、実験1とは異なる参加者であった。実験参加者は全員、正常な視力（矯正視力を含む）と聴力を有していた。また、実験者1名を除き、実験の目的について知らされている者はいなかった。

刺激：実験1と同様の視覚刺激と聴覚刺激を使用した。

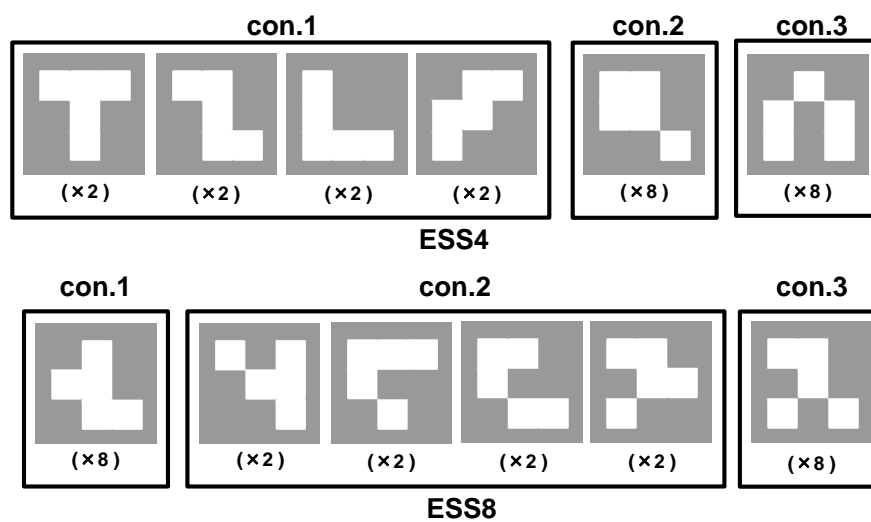


図 4-4. 実験 2 における視覚パターンの connectivity による分類。上段は ESS4 のパターンの分類を，下段は ESS8 のパターンの分類を示す。カッコ内の数字は，全試行における各パターンの繰り返し使用回数を表す。（Takeshima and Gyoba (2013c) より転載。）

手続き：実験 1 と同様の手続きで実験を行った。実験参加者は、Flash (2; 1-flash or 2-flashes) × ESS (2; ESS4 or ESS8) × Beep (3; no-beep, 1-beep, or 2-beeps) × 繰り返し (24) の計 288 試行を実施した。ただし、実験 1 と異なり視覚パターンを各 ESS で connectivity ごとに分類し、ESS 間で connectivity の平均が揃うように各パターンの使用回数を操作した (図 4-4)。

結果：視覚刺激の提示回数が 2 回と判断された割合を算出した (図 4-5)。Flash (2) × ESS (2) × Beep (3) の 3 要因分散分析を行ったところ、Flash の主効果 ($F(1, 6) = 261.59, p < .001, \eta_p^2 = .98$) と Beep の主効果 ($F(2, 12) = 12.04, p < .005, \eta_p^2 = .67$) が有意であった。一方、ESS の主効果は有意ではなかった ($F(1, 6) = 0.17, p = .70, \eta_p^2 = .03$)。また、Flash × ESS の交互作用 ($F(1, 6) = 26.03, p < .005, \eta_p^2 = .61$)、Flash × Beep の交互作用 ($F(2, 12) = 14.55, p < .001, \eta_p^2 = .70$)、ESS × Beep の交互作用 ($F(2, 12) = 4.12, p < .05, \eta_p^2 = .41$) も有意であった。さらに、Flash × ESS × Beep の交互作用も有意であった ($F(2, 12) = 5.03, p < .05, \eta_p^2 = .47$)。3 要因の交互作用が有意であったので、視覚刺激の提示回数別に ESS (2) × Beep (3) の 2 要因分散分析を行った。視覚刺激が 1 回提示では、ESS の主効果 ($F(1, 6) = 12.79, p < .05, \eta_p^2 = .68$) と Beep の主効果 ($F(2, 12) = 15.95, p < .001, \eta_p^2 = .73$) が有意であった。また、ESS × Beep の交互作用も有意であった ($F(2, 12) = 10.68, p < .005, \eta_p^2 = .64$)。2 要因の交互作用が有意であったため、単純主効果検定を行ったところ、各 ESS 条件における Beep の単純主効果が有意であった (ESS4: $F(2, 24) = 19.69, p < .001, \eta_p^2 = .62$; ESS8: $F(2, 24) = 11.93, p < .001, \eta_p^2 = .50$)。多重比較を行ったところ、いずれの ESS 条件でも 2-beeps 条件の割合が no-beep, 1-beep 条件と比べて高くなっていた ($ps < .001$)。一方、いずれの ESS においても、no-beep と 1-beep との差は有意ではなかった ($ps > .79$)。また、2-beeps 条件における ESS の単純主効果も有意であり ($F(1, 18) = 30.70, p < .001, \eta_p^2 = .63$)、ESS8 条件の割合が ESS4 条件と比べて有意に低くなっていた。視覚刺激が 2 回提示では ESS の主効果は有意傾向であったが ($F(1, 6) = 4.67, p = .07, \eta_p^2 = .43$)、Beep の主効果 ($F(2, 12) = 1.91, p = .19, \eta_p^2 = .24$) と ESS × Beep の交互作用 ($F(2, 12) = 0.08, p = .92, \eta_p^2 = .01$) は有意ではなかった。

4-1-4 実験 3：2 音の SOA による影響

実験参加者：実験者 1 名を含む東北大学の学生 8 名 (男性 3 名、女性 5 名) が実験 3 に参加した。8 名のうち 3 名は、実験 1・2 に参加していない参加者であった。実験参加者は全員、正常な視力 (矯正視力を含む) と聴力を有していた。また、実験者 1 名を除き、実験

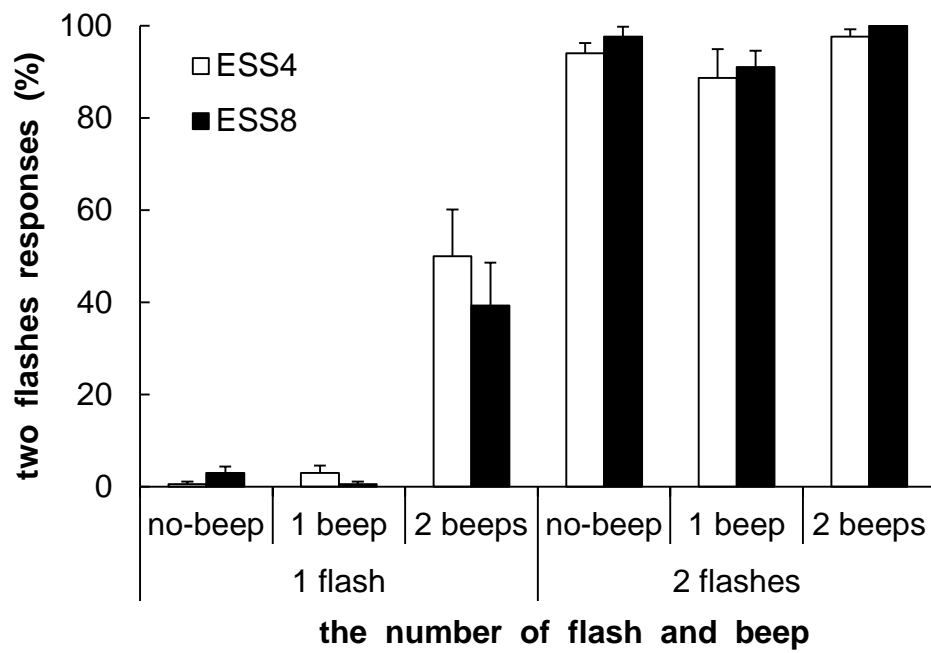


図 4-5. 実験 2 における視覚刺激の提示回数が 2 回と判断された割合。エラーバーは標準誤差を示す。(Takeshima and Gyoba (2013c) より転載。)

の目的について知らされている者はいなかった。

刺激：実験 1 と同様の視覚刺激と聴覚刺激を使用した。ただし、視覚刺激および聴覚刺激の 2 回の刺激間の SOA は、実験 1・2 とは異なり、50 ms, 85 ms, 117 ms の 3 水準を設定した。

手続き：実験 1 と同様の手続きで実験を行った。実験参加者は、Flash (2; 1-flash or 2-flashes) × SOA (3; 50, 85, or 117 ms) × ESS (2; ESS4 or ESS8) × Beep (3; no-beep, 1-beep, or 2-beeps) × 繰り返し (24) の計 864 試行を実施した。

結果：視覚刺激の提示回数が 2 回と判断された割合を算出した (図 4-6)。視覚刺激の提示回数別に、SOA (3) × ESS (2) × Beep (3) の 3 要因分散分析を行った。視覚刺激の提示回数が 1 回の場合には、SOA の主効果 ($F(2, 14) = 12.92, p < .001, \eta_p^2 = .65$), ESS の主効果 ($F(1, 7) = 7.32, p < .05, \eta_p^2 = .51$), Beep の主効果 ($F(2, 14) = 15.48, p < .001, \eta_p^2 = .69$) が有意であった。さらに、SOA × Beep の交互作用も有意であった ($F(4, 28) = 2.95, p < .05, \eta_p^2 = .30$)。2 要因の交互作用が有意であったため単純主効果検定を行ったところ、各 SOA 条件において Beep の単純主効果が有意であった (50 ms: $F(2, 42) = 19.61, p < .001, \eta_p^2 = .48$; 85 ms: $F(2, 42) = 10.23, p < .001, \eta_p^2 = .33$; 117 ms: $F(2, 42) = 11.47, p < .001, \eta_p^2 = .35$)。多重比較の結果、いずれの SOA 条件においても 2-beeps 条件の割合が no-beep, 1-beep 条件と比べて高くなっていた ($ps < .001$)。一方、no-beep と 1-beep 条件との割合の差は、いずれの SOA 条件においても有意ではなかった ($ps > .42$)。また、2-beeps 条件において SOA の単純主効果が有意であり ($F(2, 42) = 11.16, p < .001, \eta_p^2 = .35$)、多重比較を行ったところ、50 ms 条件における割合が 85 ms, 117 ms 条件と比較して有意に高くなっていた ($ps < .001$)。一方、85 ms と 117 ms 条件の間の割合の差は有意ではなかった ($p = .69$)。

視覚刺激の提示回数が 2 回の場合には、SOA × Beep の交互作用が有意であった ($F(4, 28) = 5.31, p < .005, \eta_p^2 = .43$)。2 要因の交互作用が有意であったので単純主効果検定を行ったところ、Beep が no-beep と 1-beep 条件の時に SOA の単純主効果が有意であった (no-beep: $F(2, 42) = 4.70, p < .05, \eta_p^2 = .18$; 1-beep: $F(2, 42) = 5.74, p < .01, \eta_p^2 = .21$)。多重比較の結果、SOA が 117 ms の時の割合が 50 ms と比べて高くなっていた ($ps < .05$)。一方、85 ms と 117 ms の SOA では、割合の差は有意ではなかった ($ps > .19$)。また、SOA が 50 ms の時に Beep の単純主効果が有意で ($F(2, 42) = 7.50, p < .001, \eta_p^2 = .26$)、多重比較を行ったところ、

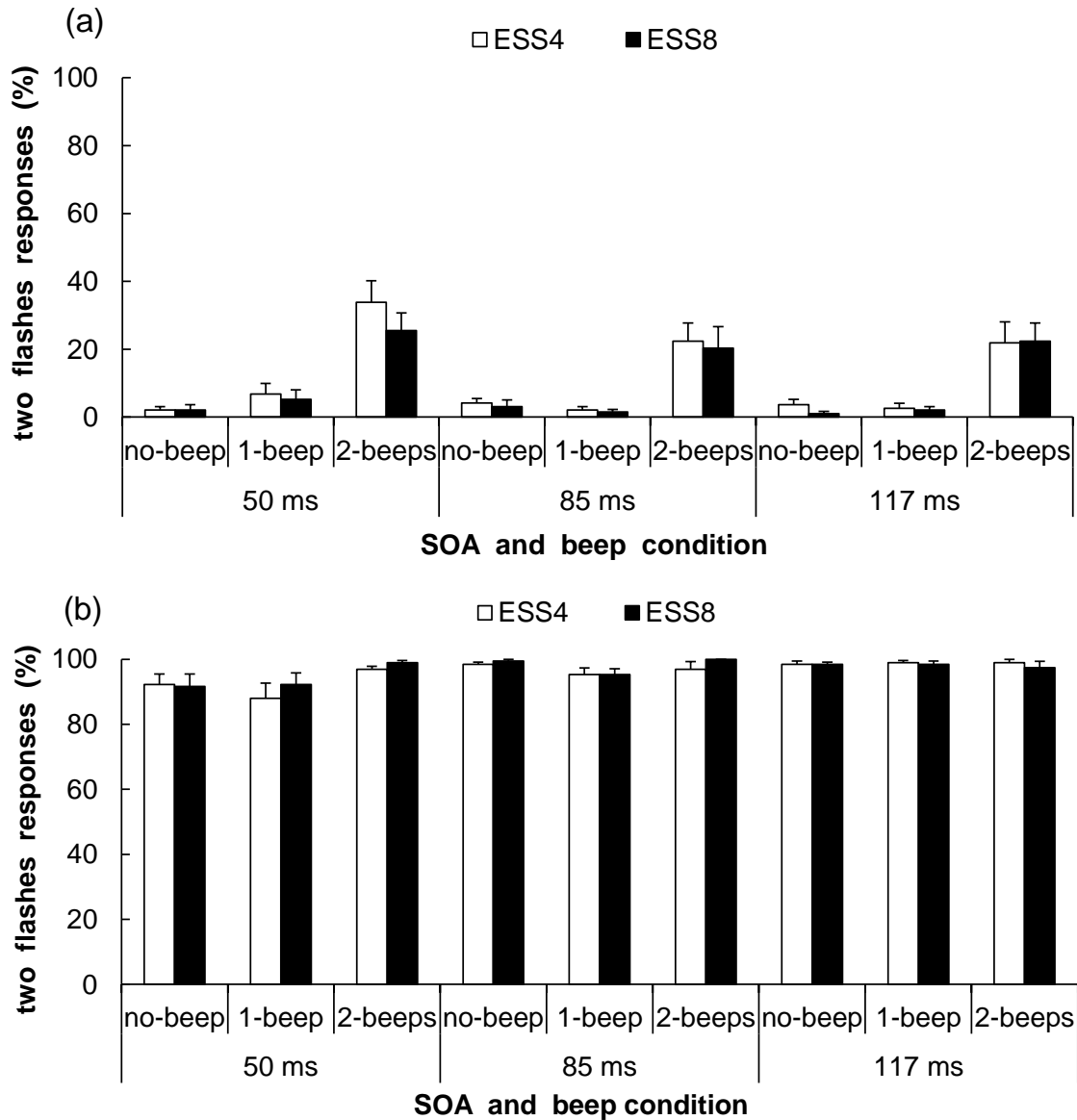


図 4-6. 実験 3 の結果。(a) 視覚刺激の提示回数が 1 回の場合に 2 回と判断された割合。(b) 視覚刺激の提示回数が 2 回の場合に 2 回と判断された割合。エラーバーは標準誤差を示す。
(Takeshima and Gyoba (2013c) より転載。)

no-beep 条件の割合が 1-beep, 2-beeps 条件よりも低くなっていた ($p < .05$)。一方, 1-beep と 2-beeps 条件との割合の差は有意ではなかった ($p = .84$)。

視覚パターンの ESS による分裂錯覚の生起頻度の違いに SOA が与える影響を検討するために, 視覚刺激の提示回数が 1 回の場合について, 各 SOA 別に ESS (2) \times Beep (3) の 2 要因分散分析を行った。SOA が 50 ms の場合には, ESS の主効果 ($F(1, 7) = 15.83, p < .01, \eta_p^2 = .69$), Beep の主効果 ($F(2, 14) = 17.66, p < .001, \eta_p^2 = .71$), ESS \times Beep の交互作用 ($F(2, 14) = 3.81, p < .05, \eta_p^2 = .35$) が有意であった。2 要因の交互作用が有意であったので単純主効果検定を行ったところ, どちらの ESS でも Beep の単純主効果が有意であった (ESS4: $F(2, 28) = 21.12, p < .001, \eta_p^2 = .60$; ESS8: $F(2, 28) = 11.63, p < .001, \eta_p^2 = .45$)。多重比較の結果, 2-beeps 条件の割合が no-beep, 1-beep 条件と比べて有意に高くなっていた ($ps < .001$)。No-beep と 1-beep 条件との割合の差は有意ではなかった ($ps > .38$)。また, 2-beeps 条件において ESS の単純主効果が有意であり ($F(1, 21) = 16.85, p < .001, \eta_p^2 = .45$), ESS8 の割合が ESS4 と比較して有意に低くなっていた。一方, SOA が 85 ms と 117 ms の場合には, Beep の主効果が有意であった (85 ms: $F(2, 14) = 10.98, p < .001, \eta_p^2 = .61$; 117 ms: $F(2, 14) = 11.98, p < .001, \eta_p^2 = .63$)。多重比較を行ったところ, 2-beeps 条件の割合が no-beep, 1-beep 条件と比べて有意に高くなっており ($ps < .005$), no-beep と 1-beep 条件との割合の差は有意ではなかった ($ps > .70$)。しかし, ESS \times Beep の交互作用はどちらの SOA でも有意ではなかった (85 ms: $F(2, 14) = 0.08, p = .92, \eta_p^2 = .01$; 117 ms: $F(2, 14) = 0.75, p = .48, \eta_p^2 = .10$)。

4-1-5 考察

研究 4 では, 視覚刺激の複雑さが視聴覚統合によって生じる分裂錯覚の生起頻度に与える影響について検討を行った。

実験 1 では, Garner and Clement (1963) によって提案された同等集合サイズ概念によって複雑さが定量的に操作された視覚パターンを用いて, 複雑さの異なる ESS4 と ESS8 のパターン間で分裂錯覚の生起頻度の比較を行った。実験の結果, ESS4 と ESS8 のいずれのパターンにおいても, 聴覚刺激が 2 回提示することによって視覚刺激の提示回数が 1 回であっても 2 回と判断される割合が, 聴覚刺激を提示しない場合と比較して有意に高くなっていた。したがって, 本研究で用いた視覚パターンにおいても分裂錯覚が生起することが確認された。しかし, ESS 間で同条件における視覚刺激が 2 回と判断される割合を比較したところ, ESS4 と比べて ESS8 のパターンでは割合が有意に低くなっていた。ゆえに,

単純な ESS4 のパターンよりも複雑な ESS8 のパターンでは、分裂錯覚が生じにくいことが示された。

実験 2 では、各パターンの使用回数を操作することで ESS4 と ESS8 間の connectivity の統制を行った上で、再度、分裂錯覚の生起頻度の比較を行った。実験の結果、視覚刺激の提示回数は 1 回でも聴覚刺激を 2 回提示することによって、聴覚刺激を提示しない場合よりも視覚刺激が 2 回と判断される割合が有意に高くなっていた。したがって、connectivity を操作しても分裂錯覚は生起することが確認された。また、同条件における視覚刺激が 2 回と判断される割合を ESS 間で比較したところ、実験 1 と同様に ESS8 では ESS4 と比べて割合が有意に低くなっていた。以上より、実験 1 において見られた ESS4 と ESS8 の間の分裂錯覚の生起頻度の違いは、connectivity の違いによるものであることが示された。

実験 3 では、複雑さによる分裂錯覚の生起頻度が、視覚刺激の処理時間の遅延によって生じていることを確認するために、2 回の刺激間の SOA の操作を行った。実験の結果、2 回の聴覚刺激によって 1 回の視覚刺激が 2 回と判断される割合は、SOA が長くなるほど有意に低くなっていた。2 回の刺激間の SOA が長くなるほど、分裂錯覚は生じにくくなることが報告されている (Shams et al., 2000; 2002)。したがって、この結果は先行研究の知見を再現していた。さらに、各 SOA において ESS4 と ESS8 の間で視覚刺激が 2 回と判断される割合を同条件において比較したところ、SOA が 50 ms では ESS8 の方が ESS4 よりも割合が有意に低くなっていた。しかし、SOA が 85 ms と 117 ms においては、2 回と判断される割合の差は ESS 間で有意ではなかった。ゆえに、複雑さによる分裂錯覚の生起頻度の違いは、2 回の刺激間の SOA が長くなると生じなくなることが示された。

以上の結果より、視覚刺激の複雑さが分裂錯覚の生起頻度に影響を与えることが示された。また、2 回の刺激間の SOA の操作によって複雑さによる分裂錯覚の生起頻度の違いが見られなくなることから、視覚刺激の処理速度が関連していると考えられる。分裂錯覚は、非常に短い時間間隔における視聴覚統合によって生起する錯覚である。錯覚としての 2 回目の視覚刺激が知覚される際には、物理的に 2 回の視覚刺激を提示した場合と同程度の V1 の活動が生じる (Watkins et al., 2006)。分裂錯覚の生起メカニズムとしては、1 回目の視聴覚刺激が統合されることによって 1 つの刺激表象を形成し、2 回目の聴覚入力 V1 を活動させ、形成された表象にしたがった錯覚としての視覚刺激を知覚させるというメカニズムが考えられる。このメカニズムの推定は、聴覚野（特に A1）から V1 に対してフィードバックする神経経路が存在する (Clavagnier et al., 2004; Falchier et al., 2002; Rockland & Ojima,

2003) ことから妥当であるといえる。視覚刺激が複雑な場合には視覚処理に時間を要する (Markovic & Gvozdenovic, 2001) ことから、複雑な視覚刺激では聴覚刺激との統合にも単純な視覚刺激と比べて時間を要すると推測される。この視覚刺激の複雑さによって視聴覚間の統合が遅延するため、1 回目の視聴覚刺激の統合が不十分なまま 2 回目の聴覚刺激が提示され、錯覚としての 2 回目の視覚刺激が知覚されにくくなると考えられる。

4-2 研究 5：視覚刺激を構成する物体数が視聴覚統合による促進効果に与える影響

4-2-1 目的

研究 4 では、視覚刺激の複雑さによる視覚処理速度の操作によって、分裂錯覚の生起頻度に変容することが示された。研究 5 では、感覚刺激の処理速度が視聴覚統合に与える影響について、以下の点から更なる検討を行った。1 点目は、感覚刺激の処理速度の影響の多様性の検討である。研究 3 および研究 4 では、いずれも分裂錯覚の生起頻度を視聴覚相互作用の指標として検討を行っている。研究 5 では、視聴覚統合によって生じる錯覚以外の現象を指標とした検討を試みる。1-3 で議論したように、視聴覚統合においては、課題に無関連な聴覚刺激を同期することによって、視覚課題の遂行成績が向上する現象が報告されている。本研究では、同期した聴覚刺激による視覚表象形成の促進効果 (Chen & Spence, 2011) を指標として用いる。Chen and Spence (2011) ではこの促進効果を、逆行マスキングを利用したターゲットの同定課題において示している。本研究でも、逆行マスキングを用いた課題を使って検討を行う。

2 点目は、高次過程の処理速度の規定因の影響の多様性の検討である。感覚刺激の処理速度の影響が複数の処理段階において生起するのであれば、物体の複雑さ以外の規定因を操作しても生じると推測される。そこで研究 5 では、視覚刺激を構成する物体数の操作を行った。視覚的短期記憶や一目で判断できる物体数 (即時把握) の容量には限界がある (e.g., Alvarez & Cavanagh, 2004; Oyama, Kikuchi, & Ichihara, 1981) ことから、物体数が増えるほど負荷が大きくなり処理速度が遅くなることが予測される。以上の点を踏まえ、研究 5 ではターゲットを構成する物体数を操作し、逆行マスキングに対する同期した聴覚刺激の影響について検討を行った。なお、ここで報告するデータは、Takeshima and Gyoba (2014) に基づいている。

実験 1 では、2 つのターゲットの異同判断を逆行マスクングによって困難にし、ターゲットの提示と同期した聴覚刺激が各物体数のターゲットの正答率に与える影響について検討を行った。その際、ターゲットとマスクの ISI も操作することによって、物体数毎の異同判断の正答率の向上効果と ISI との関連についても検討を行った。

実験 2 では、視覚刺激を構成する物体数によって処理速度が異なることを確認するため、異同判断に要する反応時間を物体数間で比較した。同時に、物体数による処理速度の違いが、視覚処理のどの段階で生起するかの推定も行った。

4-2-2 実験 1：聴覚刺激による視覚表象形成促進への影響

実験参加者：東北大学の学生 13 名（男性 6 名、女性 7 名）が実験 1 に参加した。実験参加者は全員、正常な視力（矯正視力を含む）と聴力を有していた。また、実験の目的を知っている実験参加者はいなかった。

装置：実験制御用の PC（XPS720, Dell; OS: Windows Vista, Microsoft; 解像度：1024 × 768 pixels; リフレッシュレート：60 Hz）を使用し、視覚刺激は CRT ディスプレイ（Trinitron GDM-F520, Sony）上に提示し、聴覚刺激は聴覚デバイス（Edirol FA-66, Roland）とヘッドホン（HDA200, Sennheiser）を通して提示した。刺激の制御には、MATLAB (MathWorks, Inc.) と実験用ライブラリ（Cogent Graphics and 2000 toolbox）を利用した。視覚刺激と聴覚刺激は、デジタルオシロスコープ（TS-80600, Iwatsu）を使用して開始が同期していることを確認した。実験は案室内で行い、暗騒音は 43.6 dB SPL (A) であり、視距離はアゴ台を使用して 60 cm に固定した。実験参加者の応答は、キーボードのテンキーを使用して行った。

刺激：各試行では、白色 (43.5 cd/m^2) の注視点 ($0.4 \times 0.4 \text{ deg}$)、ターゲットパターン、マスクパターンが黒色 (1.7 cd/m^2) の背景に提示された。実験で使用したパターンは、 5×5 ($2.0 \times 2.0 \text{ deg}$) の仮想マトリクス上にドット（直径 0.26 deg ）を配置することで作成した（図 4-7）。ターゲットパターンのドットの数量は 3 水準で、5 個、7 個、9 個であった。また、ターゲットパターンは「基準パターン」と「変形パターン」の 2 種類を用意した。「変形パターン」は、「基準パターン」の中心のドットの位置を、隣接したセルに 1 セル分移動することで作成した。ターゲットパターンは、注視点の左右に提示し、注視点からパターンの中心までの距離は、それぞれ 4.0 deg であった。マスクパターンは、仮想マトリクス上の

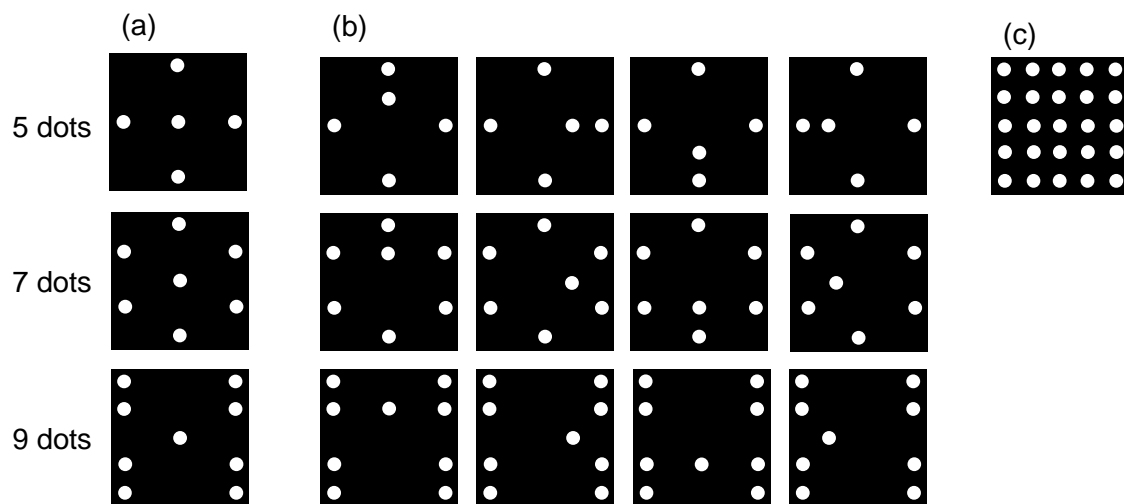


図 4-7. 研究 5 で使用した視覚パターン。各パターンは、 5×5 の仮想マトリクス上にドットを配置することで作成した。(a) は基準パターンを、(b) は変形パターンを、(c) はマスクパターンを示す。(Takeshima and Gyoba (2014) より転載。)

全てのセルにドットを配置することで作成した。ターゲットパターンとマスクパターンの ISI は 4 水準で、33 ms, 67 ms, 100 ms, 150 ms であった。聴覚刺激は純音（立ち上り/立ち下り時間：5.0 ms; サンプリング周波数：22050 Hz）で、周波数は 1000 Hz, 音圧は 80 dB (A) であった。

手続き：各試行は、「0」キーを押すことで開始された。画面の中央に注視点が 1000 ms 提示された後、2 つのターゲットパターンが 50 ms 提示された。ターゲットパターンが消えた後、ブランク画面を挟んでマスクパターンが 50 ms 提示された（図 4-8）。実験参加者は、2 つのターゲットパターンが同じ試行と異なる試行を半数ずつ行った。ターゲットパターンが同じ試行では、同じドット数の「基準パターン」が注視点の左右に提示された。一方、異なる試行では同じドット数の「基準パターン」と 4 種類の中からランダムに選ばれた「変形パターン」が提示された。「基準パターン」と「変形パターン」の配置は、半数ずつランダムに設定された。聴覚刺激は、半分の試行においてターゲットパターンの開始と同期して 50 ms 提示された。実験参加者は、Tone (2; without tone or with tone) × Quantity (3; 5, 7, or 9 dots) × ISI (4; 33, 67, 100, or 150 ms) × Pattern type (2; same or different) × 繰り返し (18) の計 864 試行を実施した。また、実験参加者の課題は左右のパターンの異同を強制二肢選択法で応答することであった。

結果：異同判断の正答率を算出し（図 4-9）、Tone (2) × Quantity (3) × ISI (4) の 3 要因分散分析を行ったところ、Tone の主効果 ($F(1, 12) = 5.01, p < .005, \eta_p^2 = .29$), ISI の主効果 ($F(3, 36) = 24.56, p < .001, \eta_p^2 = .67$) が有意であった。また、Quantity の主効果が有意であったため ($F(2, 24) = 33.63, p < .001, \eta_p^2 = .74$), 多重比較（ライアン法）を行ったところ、ドット数が増えるほど正答率が低くなっていた ($ps < .001$)。さらに、Tone × ISI の交互作用 ($F(3, 36) = 4.21, p < .05, \eta_p^2 = .26$), Tone × Quantity の交互作用 ($F(6, 72) = 2.85, p < .05, \eta_p^2 = .19$), Tone × Quantity × ISI の交互作用 ($F(6, 72) = 2.27, p < .05, \eta_p^2 = .16$) も有意であった。

3 要因の交互作用が有意であったので、ドット数別に Tone (2) × ISI (4) の 2 要因分散分析を行った。ドット数が 5 個の場合では、ISI の主効果が有意であった ($F(3, 36) = 20.52, p < .001, \eta_p^2 = .63$)。多重比較を行ったところ、ISI が 150 ms の時の正答率が他の ISI 条件よりも高くなっていた ($ps < .001$)。一方、Tone の主効果 ($F(1, 12) = 0.57, p = .47, \eta_p^2 = .05$) と Tone × ISI の交互作用 ($F(3, 36) = 0.36, p = .78, \eta_p^2 = .03$) は有意ではなかった。ドット数が

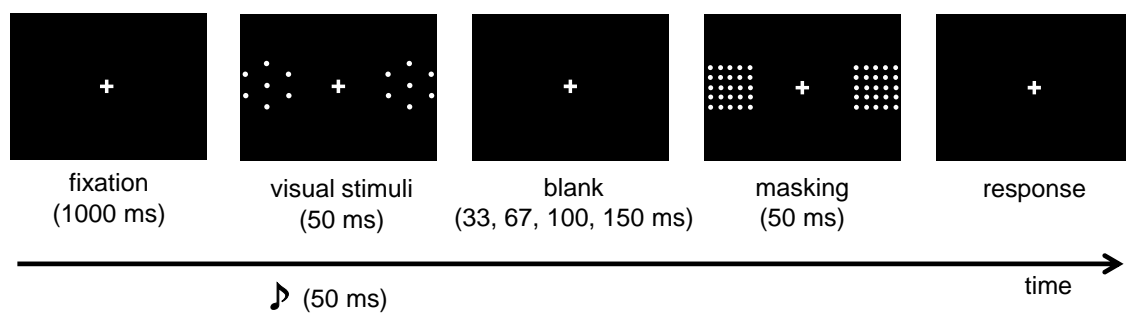


図 4-8. 実験 1 の 1 試行の流れ。(Takeshima and Gyoba (2014) より転載。)

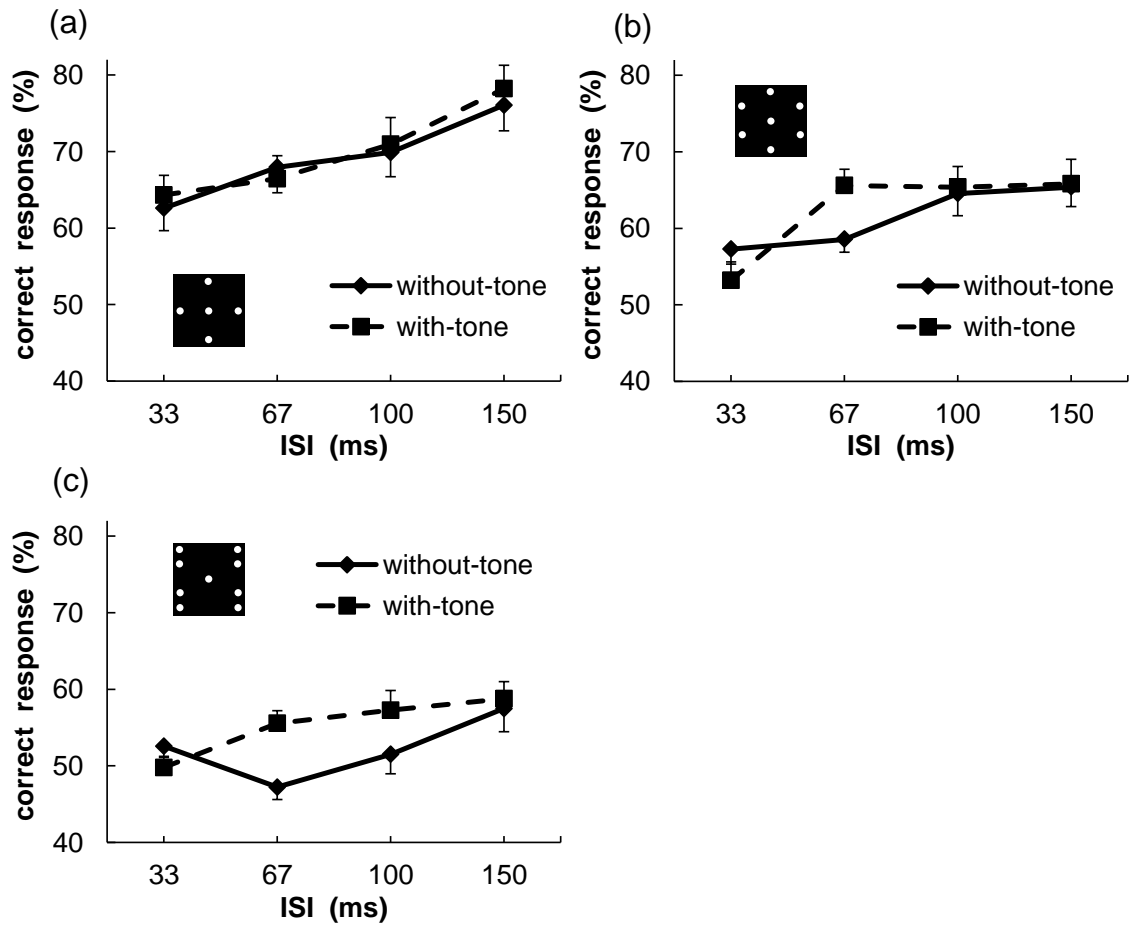


図 4-9. 実験 1 の結果。(a) ドット数が 5 個のパターンの正答率。(b) ドット数が 7 個のパターンの正答率。(c) ドット数が 9 個のパターンの正答率。エラーバーは標準誤差を示す。

(Takeshima and Gyoba (2014) より一部改変。)

7 個の場合では ISI の主効果が有意であり ($F(3, 36) = 14.77, p < .001, \eta_p^2 = .55$), 多重比較を行ったところ, 33 ms 条件の正答率が他の ISI 条件よりも低くなっていた ($ps < .001$)。また, Tone \times ISI の交互作用も有意であった ($F(3, 36) = 3.68, p < .005, \eta_p^2 = .23$)。2 要因の交互作用が有意であったため単純主効果検定を行ったところ, ISI が 67 ms で Tone の単純主効果が有意であり ($F(1, 48) = 7.94, p < .001, \eta_p^2 = .14$), 聴覚刺激によって正答率が向上していた。一方, Tone の主効果は有意ではなかった ($F(1, 12) = 0.38, p = .55, \eta_p^2 = .03$)。ドット数が 9 個の場合では, Tone の主効果 ($F(1, 12) = 7.22, p < .005, \eta_p^2 = .38$) と ISI の主効果 ($F(3, 36) = 4.86, p < .001, \eta_p^2 = .29$) が有意であった。多重比較を行ったところ, 150 ms の正答率が 33 ms と 67 ms 条件の正答率よりも高くなっていた ($ps < .005$)。また, Tone \times ISI の交互作用も有意であった ($F(3, 36) = 6.18, p < .005, \eta_p^2 = .25$)。2 要因の交互作用が有意であったので単純主効果検定を行ったところ, 67 ms と 100 ms の ISI において Tone の単純主効果が有意であり (67 ms: $F(1, 48) = 16.13, p < .001, \eta_p^2 = .25$; 100 ms: $F(1, 48) = 7.73, p < .001, \eta_p^2 = .14$), 聴覚刺激によって正答率が向上していた。

さらに, 67 ms ISI 条件における聴覚刺激による向上効果の強さについて 7 dots と 9 dots で比較するために, with-tone 条件内で Quantity (3) \times ISI (4) の 2 要因分散分析を実施したところ, Quantity \times ISI の交互作用が有意であった ($F(6, 72) = 3.03, p < .05, \eta_p^2 = .20$)。2 要因の交互作用が有意であったため単純主効果検定を行ったところ, 67 ms ISI 条件において Quantity の単純主効果が有意であった ($F(2, 96) = 8.74, p < .001, \eta_p^2 = .15$)。多重比較を行ったところ, 9 dots 条件の正答率が他のドット数よりも低くなっていた ($ps < .001$)。

4-2-2 実験 2 : 処理段階の推定

実験参加者 : 実験 1 に参加していた東北大学の学生 13 名のうち, 7 名 (男性 3 名, 女性 4 名) が実験 2 に参加した。実験参加者は, 全員実験の目的を伝えられていなかった。

刺激 : 実験 1 のターゲットパターンを, 実験 2 においても使用した。ターゲットパターンは注視点の左右に提示し, 注視点からパターンの中心までの距離は, それぞれ 4.0 deg であった。

手続き : 各試行は, 「0」キーを押すことで開始された。画面の中央に注視点 が 1000 ms 提示された後, ブランク画面を挟んで 2 つのターゲットパターンが提示された。ブランク画

面は、100 ~ 200 ms の間でランダムな長さで提示された。ターゲットパターンは、実験参加者の応答があるまで提示され続けた。実験参加者は、2つのターゲットパターンが同じ試行と異なる試行を半数ずつ行った。ターゲットパターンが同じ試行では、同じドット数の「基準パターン」が注視点の左右に提示された。一方、異なる試行では同じドット数の「基準パターン」と4種類の中からランダムに選ばれた「変形パターン」が提示された。「基準パターン」と「変形パターン」の配置は、半数ずつランダムに設定された。実験参加者は、Quantity (3; 5, 7, or 9 dots) × Pattern type (2; same or different) × 繰り返し (20) の計 120 試行を実施した。実験参加者の課題は、実験1と同様に左右のパターンの異同を強制二肢選択法で応答することであった。実験では、異同判断の正答率と反応時間を計測した。

結果：反応時間の平均と標準偏差を算出し、平均から標準偏差の2倍逸脱したデータに関しては、以降の分析から除外した。その結果、全体の3.8%のデータが除外された。ドット数ごとに正答率と平均反応時間を算出し（図4-10）、Quantity (3) の2要因分散分析を実施した。その結果、正答率ではQuantityの主効果が有意ではなかった ($F(2, 12) = 1.35, p = .30, \eta_p^2 = .18$)。一方、反応時間ではQuantityの主効果が有意であり ($F(2, 12) = 11.81, p < .001, \eta_p^2 = .66$)、多重比較の結果、9 dots 条件の反応時間が他のドット数よりも長くなっていた ($ps < .05$)。

4-2-4 考察

研究5では、視覚刺激を構成する物体数を操作することによる視覚刺激の処理速度の違いが、聴覚刺激による視覚表象形成の促進効果 (e.g, Chen & Spence, 2011) に与える影響について検討を行った。

実験1では、逆行マスクング課題における2つのパターンの異同判断の正答率を指標として、視覚刺激の物体数、同期した聴覚刺激、ターゲットとマスクのISIの関連について調べた。その結果、物体数が5個の時には聴覚刺激による正答率の向上効果は生じなかったが、7個の時にはISIが67 ms の条件において聴覚刺激による正答率の向上が生じていた。一方、物体数が9個の時にはISIが67 ms と100 ms の条件において、同期した聴覚刺激によって正答率が向上していた。ISIが67 ms における聴覚刺激による向上効果の大きさを物体数が7個と9個の条件で比較するために、with-tone 条件におけるISIが67 ms 時の正答率を7個と9個で比べたところ、7個の正答率が9個よりも有意に高かった。したがって、

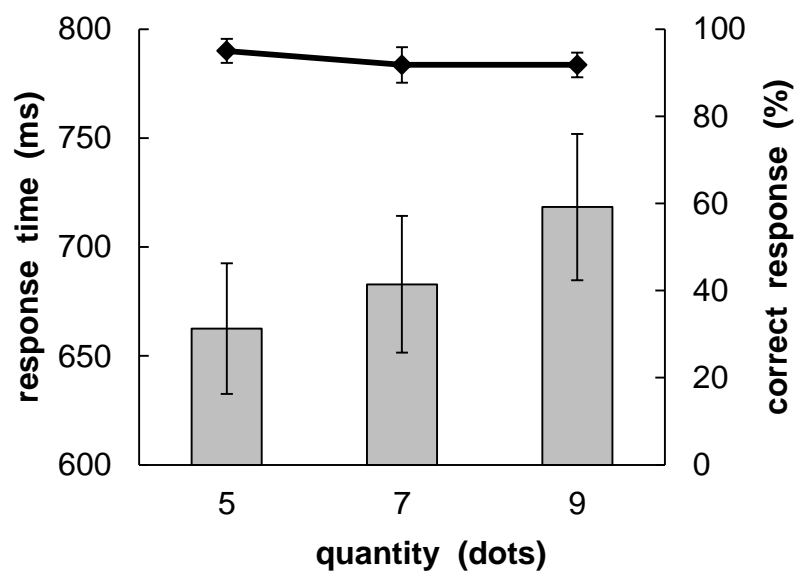


図 4-10. 実験 2 の結果。棒グラフは反応時間を，折れ線グラフは正答率の結果を表している。エラーバーは標準誤差を示す。(Takeshima and Gyoba (2014) より転載。)

ISI が 67 ms の条件における聴覚刺激の向上効果は、9 個と比較して 7 個の方が大きいことが示された。以上の結果より、聴覚刺激による視覚表象形成の向上効果が生じるための最適な ISI は、視覚刺激の物体数が増えるほどに長くなることが示された。

実験 2 では、異同判断に必要な処理時間について、物体数間で比較を行った。その結果、物体数が 9 個の時には、5 個や 7 個と比べて異動判断に要する時間が有意に長かった。5 個と 7 個の間の差は有意ではなかったが、異動判断に要する時間の平均値は 7 個の方が長くなっていた。異動判断に要する時間が長い場合、視覚刺激の符号化効率が低いことが報告されている (e.g., Garner & Sutliff, 1974)。したがって、視覚刺激を構成する物体数による処理速度の違いは、視覚処理の効率化の段階で特に生じていると推測される。

以上の結果より、視覚刺激を構成する物体数による符号化段階の処理速度の違いにより、聴覚刺激による視覚表象形成の促進効果が影響を受けることが示された。物体数が増えるほど視覚刺激の符号化処理に要する時間が長くなり、合わせて聴覚刺激との統合にも遅延が生じると考えられる。本研究では、逆行マスキング課題においてターゲットとマスクの ISI が短い場合には、物体数が多いと視覚刺激と聴覚刺激の統合がマスク提示前までに完了しないため、聴覚刺激による促進効果が生じないと考えられる。ISI が長くなるほど、多い物体数でも聴覚刺激による促進効果が生じるようになると考えられる。ただし、ISI が長くなるほどマスキングの効果が弱くなる (e.g., Enns & DiLollo, 2000) ため、視覚入力だけで異動判断が可能になり、聴覚刺激の効果が生じなくなる。したがって、視覚刺激を構成する物体数が増えるほど、聴覚刺激による視覚表象形成の促進効果が生じる最適なターゲットとマスク間の時間窓が長い方にシフトしていくと考えられる。

4-3 総合考察

本章では、第 3 章で議論した視覚刺激の空間周波数処理よりも高次の視覚処理過程における処理速度の遅延が視聴覚情報の統合過程に与える影響について検討した 2 つの研究の報告を行った。

研究 4 では、視覚刺激の複雑さが高くなることで分裂錯覚が生じにくくなることが示された。また、提示する 2 回の聴覚刺激の SOA を長くすることによって、複雑さによる錯覚の生起頻度の違いが生じなくなった。したがって、複雑さによって視覚処理の処理速度の遅延することによって、1 回目の視聴覚刺激の統合が遅延、もしくは損なわれることが錯覚の生起頻度の違いを生じさせていることがわかった。ただし、分裂錯覚は 1 回目と 2

回目の刺激間の SOA が長くなるほど生起頻度が低くなり、SOA が 117 ミリ秒程度までが限界と考えられている (Shams et al., 2000; 2002)。そのため、視覚刺激の複雑さによる影響が見られなくなる背景には、2 回の聴覚刺激間の SOA の伸張によって分裂錯覚そのものが生起しにくくなる効果も含まれてしまっていると考えられる。

研究 5 では、視覚刺激を構成する物体数によって、聴覚刺激による視覚表象形成の促進効果のタイムコースが変容することが示された。物体数が増えることによって聴覚刺激との統合に必要な時間が長くなり、逆行マスクングに対する促進効果の生起のためには、マスクの提示までの時間間隔を長くする必要がある。これは、物体数が増えることによって視覚刺激の符号化までの視覚処理が遅延し、その分だけ聴覚刺激との統合が遅くなることによって生じると考えられる。

以上のように、視聴覚情報の統合過程に視覚刺激の処理速度が与える影響は、視覚刺激の空間周波数よりも高次の規定因の操作によって生じることが確認された。研究 2 では、視覚処理の遅延によって同期の知覚が損なわれやすくなることが示されていたが、研究 4 および研究 5 より視聴覚情報の統合が遅延することも示唆される。また、感覚刺激の処理速度が与える影響は、分裂錯覚のような視聴覚統合によって生じる錯覚現象だけでなく、逆行性マスクングに対して頑健になるような聴覚刺激による視覚表象形成の促進効果においても生じることが示された。したがって、感覚刺激の処理速度は視聴覚統合に対して多様な影響を与えることが示された。

第 5 章 視聴覚情報の統合過程における制約の機能的必要性

本論文の目的は、視聴覚統合の生起を制約する諸要因について調べることであった。第 2 章では、視聴覚統合の生起を制約する時空間情報の一致と感覚情報の信頼性について議論を行い、第 3 章および第 4 章では視覚処理の処理速度と視聴覚統合との関連について検討を行った。第 5 章では、視覚処理の処理速度の影響についてまとめた上で、改めて視聴覚統合の生起に関連する諸要因について議論を行っていく。

5-1 感覚刺激の処理速度の視聴覚統合への影響のメカニズム

5-1-1 視覚処理の処理速度が視聴覚統合に与える影響

第 2 章では、視聴覚統合の生起を制約する要因である時空間情報の一致と感覚上の信頼性について包括的に議論を行った上で、視聴覚統合処理に対する感覚情報の信頼性の継時的変化の影響について、研究 1 で検討を行った。その結果、運動軌道が周辺視野に向かうために運動とともに視覚情報の信頼性が低下する事態でのみ、聴覚情報による影響が生じることが明らかとなった。この影響は、一連の運動を 1 つの事象として脳内で表象の形成を行い、その過程で情報信頼性の変化が検知された場合に視聴覚統合処理が生起することを示唆する。

第 3 章では、視覚処理の低次過程における処理速度の規定因である視覚刺激の空間周波数を操作することにより、視覚処理速度が視聴覚統合に与える影響の検討を行った。視覚刺激の空間周波数が高くなると、聴覚刺激との同期判断において物理的同期と主観的同期のずれがより大きくなり、安定した同期の知覚には視覚刺激をより先行して提示する必要があることが研究 2 より明らかとなった。さらに、高空間周波数の視覚刺激では、分裂錯覚が生じにくくなることも研究 3 より示された。分裂錯覚の生起には、1 回目の視覚刺激と聴覚刺激の統合が重要であると考えられる。したがって、空間周波数が高い視覚刺激では物理的な同期タイミングで聴覚刺激を提示しても同期の知覚が損なわれてしまい、視聴覚情報の統合が生じにくくなると考えられる。さらに、視聴覚統合が生起しにくくなることによって、統合を必要とする視聴覚相互作用もまた生起しにくくなると考えられる。

第 4 章では、空間周波数よりも高次過程の処理速度の規定因である視覚刺激の複雑さ、および視覚刺激を構成する物体数を操作することによる視聴覚統合への影響について検討

を行った。視覚刺激の複雑さの影響を検討した研究 4 からは、複雑な視覚刺激でも分裂錯覚が生じにくくなることが示された。同時に、1 回目と 2 回目の聴覚刺激の SOA を長くすることによって、複雑さによる分裂錯覚の生起頻度の違いが生じなくなることも明らかとなった。視覚刺激を構成する物体数の影響を検討した研究 5 では、物体数が増えるほど聴覚刺激による視覚表象形成の促進 (Chen & Spence, 2011) が生じるターゲットとマスクの ISI が長くなることが示された。視覚刺激を構成する物体数が増えるほど処理時間も長くなるため、ターゲットとマスクの ISI が短いと視聴覚刺激の統合が生じる前にマスキングの影響が生じてしまうと考えられる。したがって、マスキングの効果が生じる前に統合が行われるためにはターゲットとマスクの ISI を十分に長くする必要があると考えられる。

研究 3 および研究 4 より、処理速度の遅い視覚刺激では分裂錯覚が生じにくいことが示された。視覚処理の処理速度が視聴覚統合に影響を与える背景には、研究 2 より同期知覚への影響が関連していると考えられる。視聴覚統合では、事象をベースに感覚情報を処理していることが研究 1 から示唆されており、それぞれの感覚入力が同一の事象から生じていることの判断基準として同期の情報が使われると推測される。しかし、視覚刺激の処理速度に遅延が生じた場合には同期が損なわれ、視聴覚統合が生じにくくなると考えられる。そのため、視聴覚情報の統合の結果として生じる分裂錯覚などの視聴覚相互作用が生じにくくなると推測される。視覚処理の遅延によって視聴覚情報が統合されにくくなることが示されたが、同時に統合に必要な処理時間も伸長することが、研究 4 および研究 5 の結果より示されている。分裂錯覚における 2 回の聴覚刺激間の時間間隔や逆行マスキングにおけるターゲットとマスク間の時間間隔の伸長によって、視覚刺激の処理速度の影響が生じなくなる。したがって、処理速度の遅い視覚刺激を聴覚刺激と物理的同期タイミングで提示した場合には、十分な処理時間があれば視聴覚相互作用が生起する程度まで統合が行われるようになることから、視覚処理が遅い場合には視聴覚統合に必要な時間が長くなり、短い処理時間では統合が不十分であると考えられる。つまり、処理速度が遅い視覚刺激では分裂錯覚が生じにくいのは、統合が不十分な状態で 2 回目の聴覚刺激が提示されてしまうためであるといえる (図 5-1)。

5-1-2 感覚刺激の処理速度の影響の機能的必要性

視聴覚統合の役割は単一感覚時よりも安定した外界の知覚を行うことであり、そのため研究 1 の結果が示唆するように、感覚情報は事象をベースとして統合処理が行われる。感

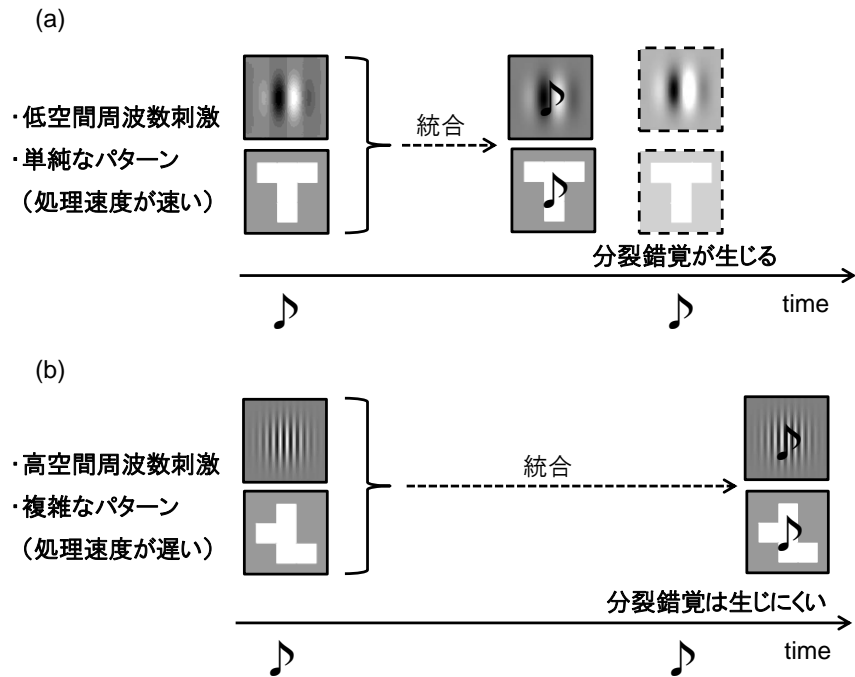


図 5-1. 処理速度による分裂錯覚の生起頻度の違いの説明図。(a) 処理速度が速い視覚刺激の場合。2 回目の聴覚刺激が提示される前に 1 回目の視覚刺激と聴覚刺激の統合処理が完了するため、分裂錯覚が生起する。(b) 処理速度が遅い視覚刺激の場合。2 回目の聴覚刺激が提示される段階ではまだ統合処理が不十分なため、分裂錯覚が生起しにくい。

覚入力の情報源が同一の事象であることの保証として、2-1 でも議論したように視覚と聴覚の感覚刺激どうしの同期が必要となる。ただし、光波と音波の伝達速度に大きな違いがある、視覚と聴覚では皮質における反応潜時に違いがある (King, 2010) ことから、同期知覚を行う場合には柔軟な補正が必要となる。適応刺激の伝達速度の違いという物理的な問題に対しては、情報源との距離によって同期の時間窓を変容させる (Sugita & Suzuki, 2003) ことによって対応していると考えられる。一方、皮質の反応潜時の違いという神経レベルの問題に対しては、同期の時間窓そのものに幅 (e.g., Guski & Troje, 2003; Lewald & Guski, 2003 を参照) を持たせることで対応していると考えられる。同期判断のピーク (PSS) を測定すると、視覚刺激先行側に偏ることが多い (e.g., Kayser et al., 2008; Lewald & Guski, 2003; Zampini et al., 2005)。神経レベルでの刺激に対する反応は聴覚の方が視覚と比べて速いことが、同期判断のピークが視覚刺激先行側に偏る原因の 1 つと考えられている (Vroomen & Keetels, 2010)。加えて、研究 2 では視覚刺激の処理速度によって PSS および時間窓がシフトすることが示されている。つまり、感覚刺激の処理速度を考慮して同期の時間窓を補正しており、生体には視聴覚情報を統合するために柔軟に対応する仕組みが備わっていると推測される (図 5-2)。

一方、視覚処理が遅い分、聴覚刺激との統合にも処理が速い刺激と比べてより長い処理時間が必要であることが示唆されている。この統合に必要な処理時間が伸長することによって、研究 3 や研究 4 のような分裂錯覚の生起頻度の低減、研究 5 のような視覚表象形成の促進が生じる ISI のシフトなどの影響が生じると考えられる。分裂錯覚の生起頻度の低減は、一見すると視聴覚情報を統合する上で不利益となるように感じられる。しかし、視覚刺激と聴覚刺激の提示頻度が一定以上に密になると、刺激間の同期知覚が生じなくなる (Benjamins et al., 2008; Fujisaki & Nishida, 2005)。このように、視覚と聴覚の統合における時間解像度があまり高くないために、統合に必要な処理時間が多少伸長しても、むしろ現実場面ではあまり大きな問題は生じないと考えられる。

神経レベルにおける処理速度が視覚と聴覚で異なるという性質があるため、視聴覚刺激間の同期知覚には、かなり柔軟な補正が必要となるのは明らかである。本論文では、従来考えられてきた視覚と聴覚の処理速度の違いを時間窓の幅によって対応する働きに加えて、視覚と聴覚の処理速度の違いがより大きくなる場合には、PSS をシフトさせるという補正が働くことを新たに見出した。視覚と聴覚という 2 つの感覚情報を利用して外界を知覚することは、単一感覚よりも情報量が多いためにより安定した知覚の形成が可能である。そ

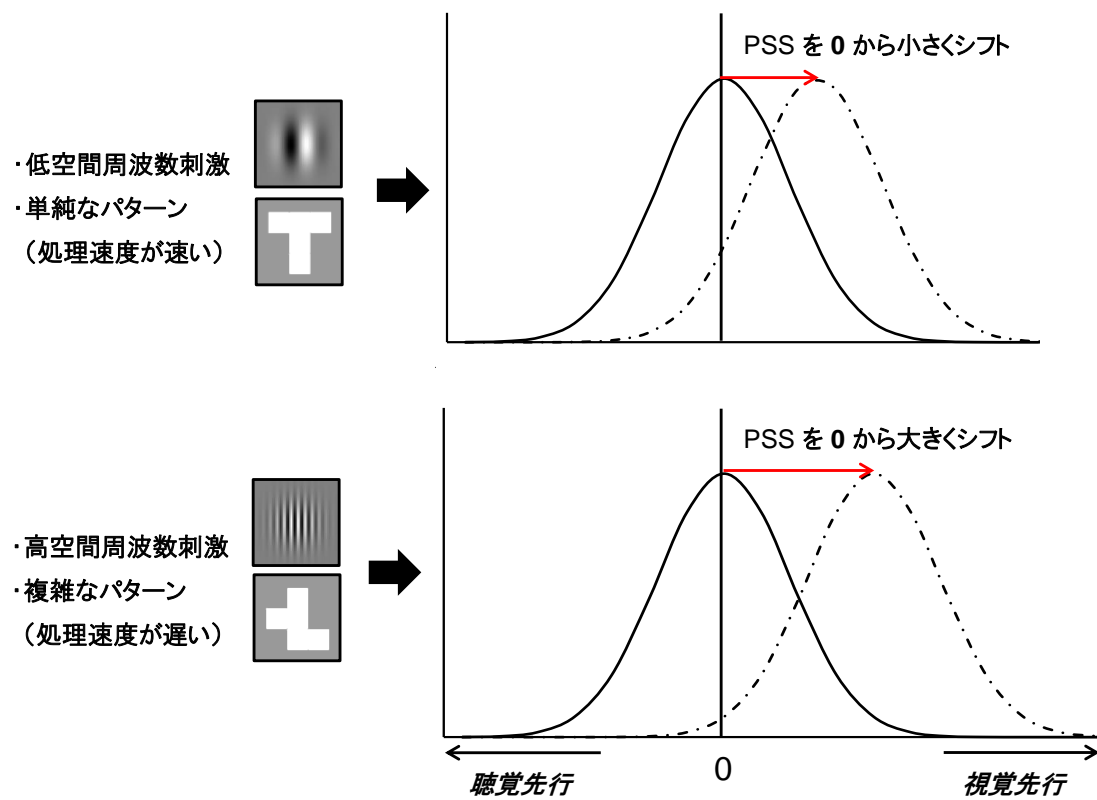


図 5-2. 感覚刺激の処理速度による同期判断の変容のモデル。処理速度が遅い視覚刺激の場合には、視覚刺激をより先行して提示しないと同期知覚が生じにくい。

のため、感覚刺激の処理速度に対する補正は、より柔軟な同期知覚を行うための重要な役割の一端を担っているといえる。

5-1-3 視聴覚統合における処理速度が与える影響の神経基盤

1-2-1 でも議論したように、視聴覚情報の統合は感覚処理の複数段階で生起している。そのため、視聴覚統合に関連している脳の領域も多岐にわたる。視聴覚統合に関連する脳領域として、上丘、上側頭溝 (superior temporal sulcus: STS)、頭頂間溝 (intraparietal sulcus: IPS)、前頭前皮質 (prefrontal cortex: PFC) を含む前頭葉の部位といった、高次過程の領域があげられる (e.g., Barraclough, Xiao, Baker, Oram, & Perret, 2005; Saito, Yoshimura, Kochiyama, Okada, Honda, & Sadato, 2005; Sugihara, Diltz, Averbeck, & Romanski, 2006)。また、近年の研究ではこれらの高次過程の領域から V1 や A1 といったより低次過程の感覚皮質へと、統合過程において信号のフィードバックが生じていることが主張されている (e.g., Foxe & Schroeder, 2005; Kayser & Logothetis, 2007)。したがって、感覚刺激の処理速度が視聴覚統合に与える影響に関連した脳活動も、感覚処理の低次過程から高次過程までの様々な段階において生じていると考えられる。

感覚刺激の処理速度が視聴覚情報の統合過程に影響を与える背景には、視聴覚刺激の同期知覚に関連している。視覚刺激と聴覚刺激の同期の判断においては、背外側前頭前皮質 (dorsolateral prefrontal cortex: DLPFC) と側頭頭頂接合部 (temporal parietal junction: TPJ) の活動の協応、および下頭頂小葉 (inferior parietal lobule: IPL) の活動に関連している (Adhikari, Goshorn, Lamichhane, & Dhamala, 2013)。背外側前頭前皮質は、前述した視聴覚統合に関連した領域である前頭前皮質の一部であることから、統合時には視聴覚刺激間の同期判断が行われていることを裏付けている。

本論文において、視覚刺激の処理速度の規定因として操作したのは、空間周波数、複雑さ、および物体数である。空間周波数の処理は、外側膝状体や V1 で行われている。複雑さには視覚刺激の形態処理に関連しており、V4 や下側頭皮質 (inferotemporal cortex: IT) で処理が行われている (Kobatake & Tanaka, 1994)。また、物体数の影響は視覚刺激の符号化の段階で生じていることが、研究 5 より示唆されている。視覚情報の符号化に関連して、視覚的ワーキングメモリ負荷が生じることで下頭頂溝 (inferior parietal sulcus) の活動の上昇 (Todd & Marois, 2004)、および頭頂間溝の活動の抑制 (Todd & Marois, 2005) が報告されている。感覚刺激の処理速度の影響は、これらの領域と前述した同期および統合に関連し

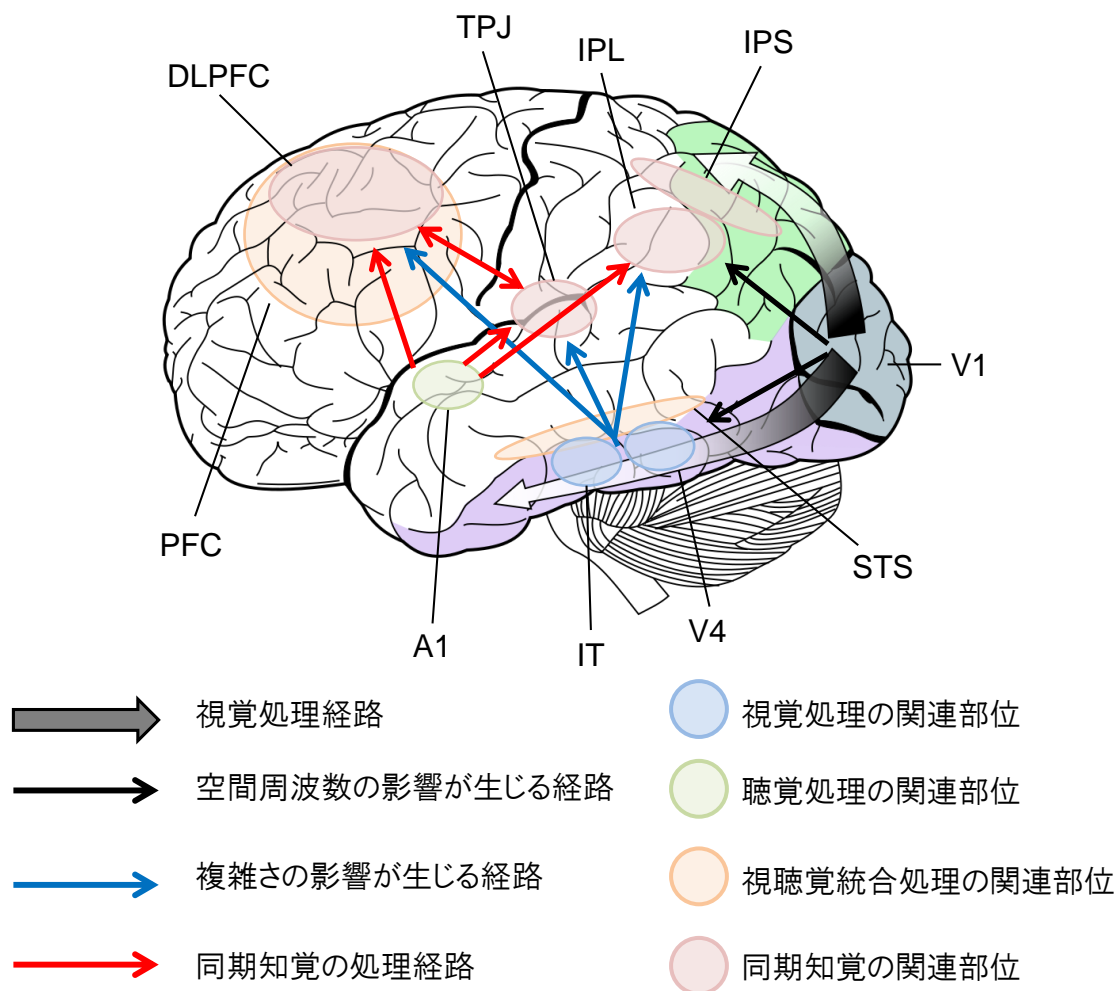


図 5-3. 感覚刺激の処理速度が視聴覚統合処理に与える影響の神経基盤のモデル図（脳の模式図は、http://ja.wikipedia.org/wiki/%E8%A6%96%E8%A6%9A%E9%87%8E#mediaviewer/File:Ventral-dorsal_streams.svg より転載）。空間周波数や視覚刺激の複雑さは、同期知覚処理を行う部位や近接した視聴覚統合処理を司る部位に影響を与えると推測される。また、物体数が影響を与える符号化処理は同期知覚処理と部位が重複しているため、物体数の影響が同期知覚に直接影響すると推測される。

た脳領域の活動が関連していると推測される（図 5-3）。

視聴覚統合においても、感覚刺激の入力段階ではそれぞれ単一の感覚処理が行われる。その中で、視覚刺激については外側膝状体や一次視覚野で空間周波数の処理が、続く腹側経路の V4 や下側頭皮質で複雑さを含む形態処理が行われる。1-1 でも述べたように、視覚処理では腹側経路における物体処理と背側経路における時空間処理を統合して視覚認識が行われるが、この視覚情報をさらに高次の認知機能において活用するために、視覚処理の信号は背外側前頭前皮質を含む前頭葉の領域へと送られる。また、聴覚刺激との同期判断を行う場合には、側頭頭頂接合部や下頭頂小葉にも信号が送られると考えられる。聴覚刺激もまた、聴覚野における処理を経てこれらの領域へと信号が送られると考えられる。背外側前頭前皮質と側頭頭頂接合部の協応や下頭頂小葉の処理によって視聴覚刺激間の同期が判断されるが、その時に視覚刺激の処理速度によって同期判断が変容すると考えられる。同時に、背外側前頭前皮質は前頭葉の領域であり、下頭頂小葉は頭頂間溝に近接していることから、これらの領域の活動は視聴覚統合自体にも影響を与えると考えられる。したがって、同期判断の変容がそのまま視聴覚統合に要する処理時間を調整していると推測される。また、視覚的ワーキングメモリ負荷が生じることで頭頂間溝の活動が抑制されることも、聴覚刺激との同期判断を変容させ、視聴覚統合に要する処理時間を調整すると考えられる。つまり、物体数が増えることで処理速度が遅くなると同時に、処理負荷も高くなることで視聴覚統合に影響を与えていると推測される。加えて、上側頭溝は V4 や下側頭皮質、一次聴覚野と近接している。したがって、上側頭溝で行われる視聴覚統合の処理もまた、視覚刺激の処理速度の影響を受けて、統合に必要な処理時間が変容すると考えられる。

視聴覚統合では、高次の脳領域から低次の感覚野への信号のフィードバックが仮定されている (e.g., Foxe & Schroeder, 2005; Kayser & Logothetis, 2007)。つまり、視聴覚統合の処理が感覚刺激の処理速度による影響を受けた場合には、一次視覚野や一次聴覚野へのフィードバック信号も遅延するといった影響を受けると考えられる。視聴覚相互作用の生起にはこのフィードバック信号が重要であると考えれば、研究 3 および研究 4 の分裂錯覚の生起頻度の低減や、研究 5 の物体数による視覚表象形成の促進が生じる ISI の変容が起こる過程を説明できる。

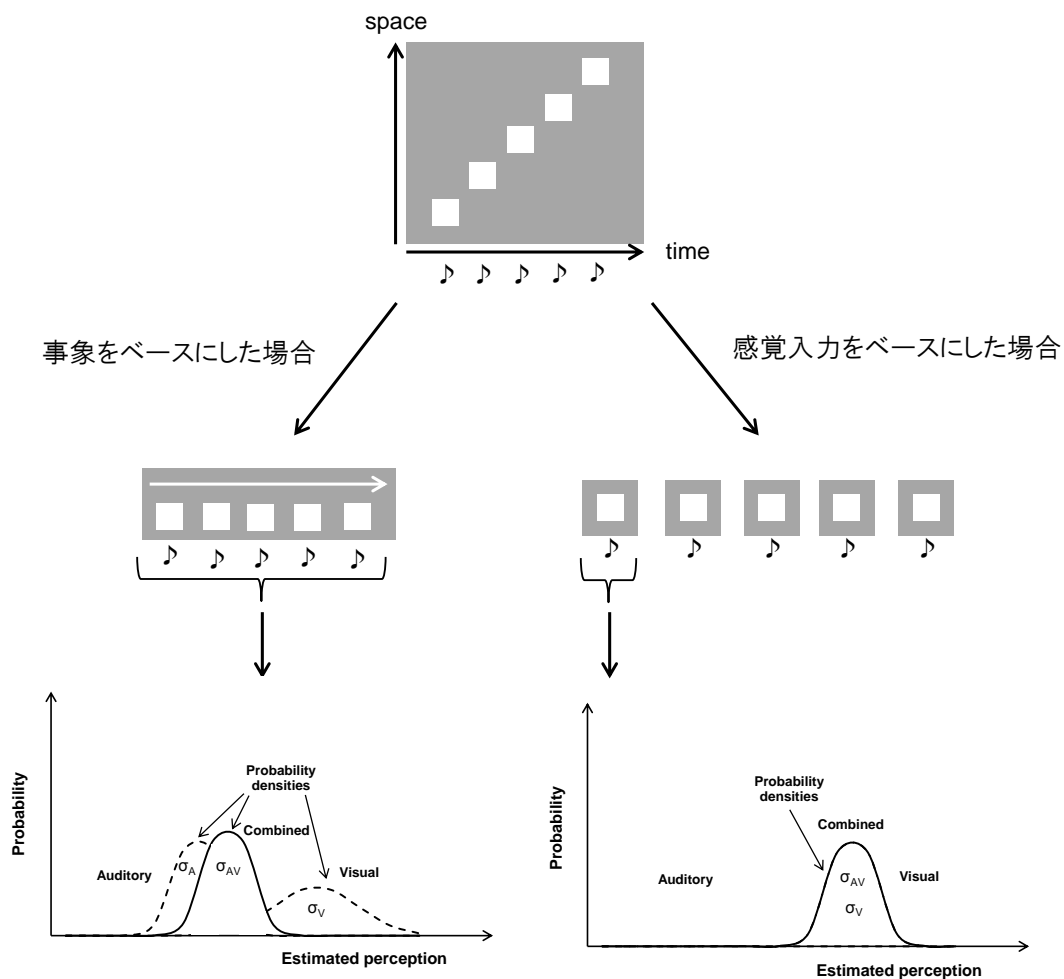
5-2 視聴覚統合を制約する要因の役割

視覚と聴覚の統合は、周囲の環境の知覚をより安定した形で形成するために行われている。そのため、同一の事象から生じた刺激どうしを正しく結びつける必要がある。視聴覚情報の統合過程に、視覚と聴覚の情報信頼性の強度や時空間情報の一致、感覚刺激の処理速度が関連する背景をこの観点から議論する。

5-2-1 重みづけと知覚形成の安定化および効率化

第2章でも述べたように、多感覚統合では統合される各感覚情報を最尤推定に従って信頼性の重みづけを行い、その重みづけを考慮して統合された結果が知覚として生じる。ゆえに、判断の対象となる感覚情報の重みづけが小さい場合には統合が生じて他の感覚情報の影響を強く受け、大きい場合には統合は生じずに他の感覚情報の影響を受けないと考えられる。研究1では、この重みづけに提示刺激の時空間的側面を加えた検討を行い、感覚情報の信頼性が継時的に変化する場合であっても、重みづけにしたがった統合が行われることを示している。また、研究1では仮現運動刺激を用いたが、ISIを操作して滑らかな運動が知覚されない場合には統合が生じなくなることから、一連の運動を一つの事象として重みづけを行っている可能性が示唆される（図5-4）。

一連の運動刺激の入力を一つの感覚情報とすることで、多感覚統合による効率のよい感覚情報処理を行っていると考えられる。仮現運動は、感覚入力としては視覚刺激の提示と消失の繰り返しであるが、視覚表象としてはある位置からある位置までの運動として形成される。視覚刺激と聴覚刺激の運動情報が一貫している時には hMT/V5+ といった脳の運動知覚に関連した領域の活動が促進する (Alink, Singer, & Muckli, 2008; Sadaghiani et al., 2009) ことから、聴覚刺激においても音情報から運動軌道を脳内で復元して聴覚表象を形成していると考えられる。そのため、運動刺激の場合には、情報信頼性は一連の運動の時空間情報から算出されて重みづけが行われ、統合の結果が知覚されることが考えられる。つまり、一連の運動刺激の入力が完了後、情報信頼性から統合が必要とされた場合には感覚情報が統合され、それにより時空間情報の再構成 (Talsma et al., 2010)、およびワーキングメモリの書き込み (Sheth & Shimojo, 2000) が行われて錯覚としての運動軌道の知覚が生じることが考えられる。もし運動刺激において感覚入力が行われるたびに統合処理を行ったとしたら、多感覚統合によって逆に負荷が大きくなって効率が悪くなってしまうと考えられる。



視覚情報の信頼性が低く、聴覚情報が重みづけられる。 視覚情報のみで知覚が形成される。

図 5-4. 研究 1 を例とした、知覚される事象をベースとした感覚情報の重みづけと感覚入力をベースとした感覚情報の重みづけの概念図。一連の運動という事象をベースにした場合は、周辺視野に向かうにつれて視覚情報の信頼性が低くなり、聴覚情報に重みづけが行われる（左）。一方、ドットと音という感覚入力をベースにした場合には、視覚情報だけで知覚が成立するため、聴覚情報が利用されない（右）。

感覚表象をベースに重みづけを行うことは、事象のより安定した知覚を行う上でも有益であると考えられる。生体の感覚処理は、外界の事象を知覚することを目的として行われる。感覚表象をベースに重みづけを行うことは、生体の感覚処理の最終目標である事象の知覚をベースにして統合処理を行っていることを意味している。したがって、感覚処理の結果として知覚される事象をベースとして統合処理を行うことで、あらかじめ目標を設定した上でそれに最適な感覚情報の統合プロセスを導いているといえる。1-2-2でも議論したように、視聴覚統合処理では、入力された視覚情報と聴覚情報をともに物体情報と時空間情報に分けて処理を行っており、それぞれで視聴覚間の情報が一致している場合には、物体表象と時空間表象はそれぞれ単一感覚時よりも頑健なものとして形成される。その後、物体表象と時空間表象を統合して1つの視聴覚表象を形成するため、この表象形成は感覚処理では比較的高次の処理過程で行われている。したがって、高次過程において形成された視聴覚表象から、視覚と聴覚の感覚情報の信頼性に基づいた入力の重みづけを行い、知覚を再構成していると推測される。高次過程の情報によって知覚が再構成されるというメカニズムは、Talsma *et al.* (2010) が提案しているモデルとも合致している。

以上のように、多感覚統合の処理過程における各感覚情報の重みづけは、高次過程における感覚表象をベースに行うことで、知覚形成の効率化および外界の事象知覚の安定化の役割を果たしていると考えられる。

5-2-2 処理速度および時空間情報の一致と情報源の同一性との関連

本論文では、視聴覚統合の制約に感覚刺激の処理速度を新たな要因として取り上げて検討を行ってきた。感覚刺激の処理速度の影響の基盤には、感覚刺激間の同期が関連している。研究2では、処理速度の遅い視覚刺激では、視覚刺激と聴覚刺激を物理的に同期したタイミングで提示しても同期知覚が損なわれやすく、むしろ処理速度の遅さを考慮して視覚刺激を先行して提示する必要があることが示された。処理速度の影響は、一方の感覚処理が遅延する分、感覚間の統合に必要な処理時間にも遅延が生じることによって引き起こされていることも、研究4および研究5から明らかとなっている。

第2章でも議論したように、各感覚刺激の時空間情報の一致は情報源の同一性を保証する重要な手がかりとなるといえる。しかし、刺激の物理的性質や、5-1-2で述べたような神経レベルの影響から、空間一致や同期の知覚にある程度の許容範囲を設けることで、物理的には時空間情報が一致していなくても、知覚的には一致しているように判断される。つ

まり、感覚処理によって感覚刺激の情報源が同一と判断されれば統合が行われると考えられる。処理速度の遅い視覚刺激では統合に必要な処理時間が伸長するのは、視覚情報と聴覚情報を正しく統合しようとするはたらきによるものといえる。視覚処理と聴覚処理は一次感覚野どうしの結びつきがあるものの、統合による知覚が生じるまでにはそれぞれの感覚処理も並列に行われ、その過程で複雑に関連し合っていると考えられる。そのため、一方の感覚処理が遅く、他方がそれに合わせるように処理が行われていればその分だけ統合を含めた全体の処理も遅延する。統合の時間分解能という制約 (Benjamins et al., 2008; Fujisaki & Nishida, 2005; 2010) はあるものの、様々な感覚入力が続え行われる日常場面を想定した場合、一方の感覚処理の遅れによって異なる情報源の入力どうしが統合してしまう可能性もある。しかし、そのような現象が生じないことから、多感覚統合においては情報源の同一性を感覚処理のかなり低次のレベルでも行っていると推測される。例えば、それぞれの刺激の一次感覚野に入力があつた段階で、反応潜時の違いを補正した上で同期していれば、各感覚刺激の情報源が同一と判断される可能性が考えられる。したがって、感覚刺激の処理速度による統合の処理時間の伸長は、低次過程にも誤った統合を抑える機能が存在することを示唆している (図 5-5)。

5-3 今後の展望と本知見の応用可能性

本論文では、感覚情報の処理速度が視聴覚情報の統合過程に影響を与えること、そこから視聴覚統合の生起に関連する諸要因の役割についての再考察を行った。ここでは、本研究の今後の展望、および本研究で得られた知見の応用可能性について議論を行っていく。

5-3-1 視聴覚統合に関連した究極要因の検討

1-1 でも取り上げたように、近年、視聴覚統合の分野の研究が非常に増えてきている。一方で、多くの研究は感覚刺激の物理的特性や時空間情報の影響、神経科学的基盤に焦点を当てたものである。本研究で行った感覚刺激の処理速度の影響の検討も、感覚刺激の物理的特性について調べた研究に含まれる。生体の行動や内的処理を説明する上では、大きく2つのアプローチがある。一方は、その行動や内的処理が生じる直接的な原因、つまりメカニズムを説明するもので、「至近要因」と呼ばれる。他方は、その機能の生存する上で必要性を説明するもので、「究極要因」と呼ばれる。これまでの視聴覚統合の研究は、至

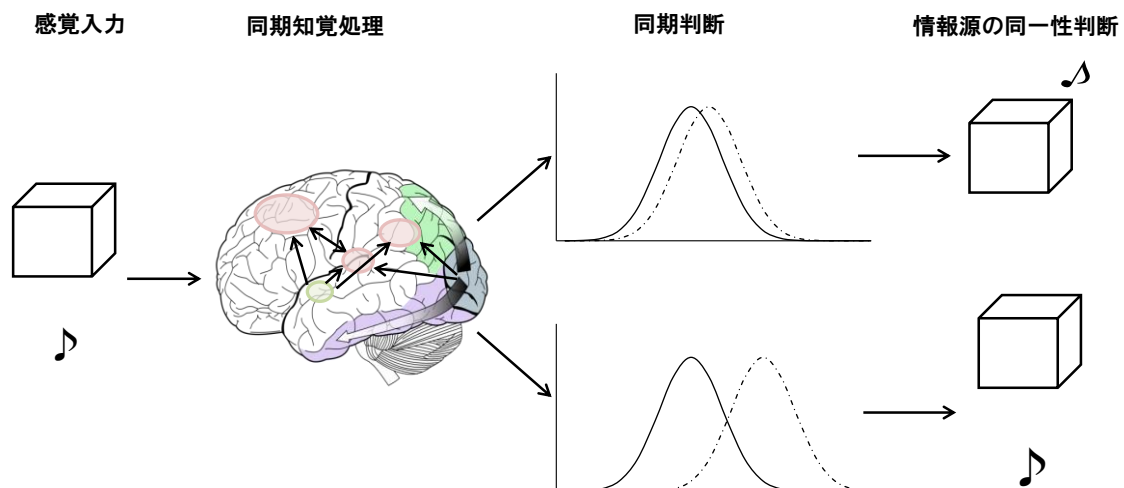


図 5-5. 視覚と聴覚の入力から判断されるそれぞれの感覚入力の情報源の同一性判断の概念図（脳の模式図は、http://ja.wikipedia.org/wiki/%E8%A6%96%E8%A6%9A%E9%87%8E#mediaviewer/File:Ventral-dorsal_streams.svg より転載）。視覚と聴覚の感覚入力が行われ、2つの感覚入力の同期が知覚された場合には、同一の情報源（事象）から入力が生じていると判断され、視聴覚情報の統合処理が行われる（上）。ただし、処理速度が遅い場合には同期知覚が生じないために、情報源が同一ではないと判断され、視聴覚統合は行われなくなる（下）。

近要因によるアプローチがほとんどであり、究極要因によるアプローチはほとんど行われてきていない。近年、刺激の親近性によって分裂錯覚の生起頻度が変容する (Setti & Chan, 2011) といった究極要因に迫るアプローチの研究が行われるようにはなっているものの、その例はまだまだ少ない。したがって、生態学的妥当性の観点から見た視聴覚統合についても検討を行っていく必要がある。

生態学的妥当性を踏まえたアプローチを行っていく上で重要と考えられるのが、情動刺激や社会的刺激が視聴覚統合に与える影響の検討である。情動刺激や社会的刺激は、情動的・感性的特性として、感覚情報処理に基づくアウェアネスとは独立して素早く自動的に処理される (e.g., Tamietto & de Gelder, 2010)。その背景には、配偶者選択や防衛反応などの環境に対する適応行動の遂行が存在すると考えられる。配偶者選択に関わる情動的・感性的特性の1つとして、「魅力 (Attractiveness)」があげられ、顔刺激の魅力認知は周辺視野に提示された刺激や瞬間提示された刺激であっても可能である (Guo, Hong, & Roebuck, 2011; Olson & Marshuetz, 2005) ことからその重要性がうかがえる。また、魅力的な顔は注意を自動的に引きつけて無視できないことが報告されている (Sui & Liu, 2009)。また、防衛反応を行う上では、自分に対して脅威となる刺激を素早く検出することが重要となる。そのため、視覚情報処理においては網膜から上丘・視床枕を経て扁桃体へと至る処理系の存在が指摘されている (e.g., Pessoa & Adolphs, 2010)。この処理経路では、前知覚的に行われる情動処理が後続する知覚処理に対して促進的にはたらき (佐藤, 2004)、脅威刺激の迅速な検出および対応を可能にしていると考えられる。したがって、魅力刺激も情動刺激もともに注意を引き付け、優先的な処理が行われる刺激であるといえる。

本研究においても実験的検討を行った分裂錯覚は、見慣れている視覚刺激のように、視覚情報が優先的に処理される事態では生じにくいことが報告されている (Setti & Chan, 2011)。そのため、魅力顔や情動顔を視覚刺激とした時には、顔刺激に対して注意が引き付けられて視覚処理が優先されるため、分裂錯覚が生じにくくなると推測される。一方、視覚刺激だけでなく聴覚刺激も注意を引き付けるような事態における検討は行われていない。情動顔と情動を喚起する不快音、魅力顔と魅力に関連した女性らしい高い音声に組み合わせの場合には、視聴覚間の情報の一致によって統合処理が促進され、分裂錯覚がより生じやすくなることが予測される。また、視覚刺激に注意が向く場合には、Crossmodal attentional spread によって刺激の検出や同定の成績が向上する。この時、同様の組み合わせの視覚刺激と聴覚刺激を使用することで、視聴覚間の情報の一致によって統合処理が促進

されることでより検出や同定が促進されることも推測される。このように、視覚刺激や聴覚刺激に含まれる魅力や情動といった情報が、視聴覚統合処理を促進する可能性が考えられる。また、その背景には環境に適応的に行動するため、優先して処理すべき対象に対しては複数の感覚情報の統合を促進することで優先度を高める働きがあることも推測される。今後、視覚刺激や聴覚刺激に含まれる情動や感性の情報が視聴覚統合処理に与える影響の検討を行うことで、環境に対する適応的機能としてこの分野の研究を大きく発展させることができると考えられる。

5-3-2 処理速度の影響を考慮した視聴覚相互作用の利用

1-3 でも述べたように、視覚と聴覚の統合によって生じる視聴覚相互作用では、様々な現象が観察される。視聴覚相互作用を上手く利用することにより、効果的な映像コンテンツを作成することができる。近年、映像コンテンツに関する新しい感性概念として、「迫真性 (Verisimilitude)」が提案されている (行場・寺本, 2012)。迫真性は、刺激情報を可能な限り還元して本物らしさを創出することで生み出される。迫真性の喚起には、前景情報に対して適切に注意が向けられ、同時に別の感覚モダリティ情報と調和することが必要とされている (寺本・吉田・日高・浅井・行場・坂本・岩谷・鈴木, 2010)。注意に関する研究では、大きな物体は注意を引き付けやすいことが明らかとなっている (Peschel & Orquin, 2013 for a review)。また、1-3 で紹介したように、音圧の大きな音を付加することで物体の大きさが実際よりも大きく知覚される (Takeshima & Gyoba, 2013a)。つまり、視聴覚相互作用によって物体に対する注意を適切に操作することで、迫真性を喚起させる映像コンテンツを作成できる可能性がある。

本論文で取り上げた分裂錯覚は、映像コンテンツの作成へと応用する上で有益な視聴覚相互作用といえる。この錯覚では、聴覚刺激の提示によって本来提示されていない視覚刺激を知覚させることができる。したがって、少ないリソースで実際以上の映像や画像を見せることが可能である。しかし、研究3および研究4で検討したように、分裂錯覚は、空間周波数が高い、複雑であるといったように処理速度が遅い視覚刺激を用いた場合には、1回目と2回目の聴覚刺激の SOA を適切に調整する必要がある。そのため、分裂錯覚を映像コンテンツで活用する場合には、コンテンツ内で使用する刺激の処理速度を考慮する必要がある。本研究で処理速度の規定因とした視覚刺激の空間周波数と複雑さは、いずれも定量的に操作することが可能である。つまり、空間周波数や複雑さといった視覚刺激の

処理速度の操作によって、聴覚刺激との統合に必要な処理時間の違いをアルゴリズム化できる可能性がある。研究 2 のように、処理速度が異なる視覚刺激では聴覚刺激との同期判断における PSS も異なるため、処理速度の違いの指標となり得る。また、ERP や眼球運動のサッケードを利用した視覚刺激の処理時間の測定方法が提案されており (VanRullen & Thorpe, 2001; Kirchner & Thorpe, 2006), このような手法を用いた検討から処理速度の違いについて補足的なデータを得ることができる。一方、分裂錯覚の生起に関連した ERP 成分を検討した研究も盛んに行われている (e.g., Mishra et al., 2007; Shams et al., 2001; 2005)。そのため、2 回の聴覚刺激の SOA の操作に加えて、ERP 成分の潜時の比較によっても視覚刺激の処理速度による分裂錯覚における視聴覚統合の処理時間の違いを検討できる可能性がある。これらの補足的な研究を数多く行うことで、分裂錯覚の生起に最適なパラメータの算出が可能になり、映像コンテンツへと応用することが期待できる。この応用分野への影響は、映像コンテンツにとどまらずプレゼンテーション技術へと波及することにより、より広く活用される技術になると予想される。

5-3-3 臨床分野への応用

近年、多感覚統合の知見の臨床分野への応用が盛んに検討されている (Foxe & Molholm, 2009; Wallace, 2012 for a review)。本研究から得られる知見もまた、臨床分野への応用が可能であると考えられる。

例えば、自閉症スペクトラム障害児は、定型発達児と比較して分裂錯覚の生起の時間窓が広いことが報告されている (Foss-Feig, Kwakye, Cascio, Burnette, Kadivar, Stone, & Wallace, 2010)。自閉症スペクトラム障害とは、社会的コミュニケーションにおける欠陥、限局された反復的行動といった症状によって特徴づけられる障害である (APA, 2013)。Foss-Feig *et al.* (2010) では、自閉症スペクトラム障害者が知覚される情報がどの感覚からもたらされたのかの同定に困難を示す (Cesaroni & Garber, 1991) ことが、時間窓の拡張の原因となっていると主張されている。加えて、時間窓の拡張は継時的な入力による多感覚情報処理の不正確さにつながり、会話などで感覚情報を正確に処理できないことから社会的コミュニケーションの不全を引き起こしている可能性を提案している。研究 4 で示された視覚刺激の処理速度による分裂錯覚の生起頻度の変容も、視聴覚情報処理の時間的機能に関連して生じている現象である。したがって、自閉症スペクトラム障害に対する視聴覚統合によるアプローチに感覚刺激の処理速度の観点を加えることで、視聴覚相互作用の時間窓の拡張が

果たす役割の解明に大きく貢献できる。健常者と自閉症スペクトラム障害者との視聴覚統合処理の違いを解明していくことは、障害の理解および治療へとつながっていくと考えられる (Foss-Feig et al., 2010)。

また、統合失調症患者の視聴覚統合処理の特徴についても検討が行われており、健常者と比べて視聴覚統合による音声知覚の促進効果が生じにくいことが明らかとなっている (Ross, Saint-Amor, Leavitt, Molholm, Javitt, & Foxe, 2007)。統合失調症患者は、バイオロジカルモーションの処理システムに不全がある (Kim, Doop, Blake, & Park, 2005)。バイオロジカルモーションの処理には上側頭皮質後部 (posterior superior temporal cortex) の活動が関連している (Allison & Carthy, 2000) が、この部位は統合失調症患者では異常が見られる (Shenton, Dickey, Frumin, & McCarley, 2001)。上側頭皮質後部は口唇の動きも処理していることから、背側系の処理の機能不全が音声知覚における統合を阻害していると考えられている (Ross et al., 2007)。また、統合失調症患者では視覚刺激と聴覚刺激を同期して提示した場合の ERP 成分の振幅が健常者と比較して小さく、低次過程における統合処理が不得意である (Magnee, Oranje, Rngeland, Kahn, & Kemner, 2009)。一方で、統合失調症患者には感覚野の不適切な連合があり、幻覚を引き起こす一因となっていることが示唆されている (Jardri, Pins, Bubrovsky, Lucas, Lethuc, Delmaire, Vantighem, Desprez, & Thomas, 2009)。したがって、統合失調症患者は健常者と比べて多感覚情報を適切に処理することができないと考えられる。本研究で検討を行った感覚刺激の処理速度の影響は、統合失調症患者の多感覚統合処理の時間的機能を検討する一助になると考えられる。今後、処理速度の異なる感覚刺激の統合処理を統合失調症患者がどのように行うかを検討することで、スクリーニングテストの開発などの応用を行うことが可能である。

5-4 結論

本研究では、視聴覚統合の生起に関連した要因について、これまで検討がなされてきた時空間情報の一致、情報信頼性の強度の観点に加え、感覚刺激の処理速度という新たな視点からも検討を行った。情報信頼性の強度に関しては、信頼性が継時的に変化する場合には、時間的な変化を踏まえた感覚情報の利用および統合が行われることが示された。また、感覚刺激の処理速度の検討からは、一方の感覚処理に遅延が生じる場合には、統合過程においても処理時間が伸長することが示された。さらに、統合の処理時間の遅延により、短い時間間隔で感覚入力が必要となる視聴覚相互作用が生じにくいことが示された。

今後は、メカニズムの解明のような至近要因からの検討だけでなく、感覚刺激の情動的・感性的特性の影響の検討のように、生態学的妥当性の観点からの視聴覚統合についてアプローチを行っていくことで、生体の多感覚処理についてさらなる理解が深まることが期待される。また、本研究から得られた知見は、統合過程で必要となる処理時間についてアルゴリズム化することで、視聴覚相互作用を効果的に利用した映像コンテンツの作成への応用が期待される。さらに、将来的には、自閉症スペクトラム障害などについて多感覚統合処理からアプローチすることや、統合失調症などのスクリーニングテストの開発など、臨床場面への応用も可能であると考えられる。

謝 辞

この論文を執筆するにあたって、多くの方からご指導とご支援をいただいた。ここに、感謝の意を綴る。

まず、指導教員である行場次朗先生には、研究者として大事なことを数えきれないほどご指導いただいた。行場先生の時に優しく、時に厳しいご指導があったからこそ、この論文を執筆するまでに至れたと感じている。また、5年前、進路に悩んでいた私に対して、研究者としての道の魅力と陰しさの両方を教えていただいたことが、研究者としての私の原点になっている。深く、感謝の意を述べたい。

心理学研究室の大渕憲一先生、阿部恒之先生、坂井信之先生、辻本昌弘先生には、研究についての有益なアドバイスをいただき、また、私の研究を温かく見守っていただいた。他にも、日常生活の中で何気なく声をかけていただくなど、お気遣いいただいたことに感謝を述べたい。大学院生や学部生には、研究についての議論や実験に参加してもらうなど、研究を進めていく上で不可欠なサポートいただいた。研究室での生活が楽しいと感じられたのも、一緒に研究室に通った皆のおかげだと感じている。

先輩である室蘭工業大学の寺本渉さん、東北福祉大学の河地庸介さん、立教大学の日高聡太さん、福島大学の高橋純一さんには、学会や研究会など多くの場面において、研究に対する叱咤激励をいただいた。所属が変わってからも私を気遣ってくださった偉大な先輩方に、感謝の意を述べたい。言語学研究室の小泉政利先生と東北文化学園大学の柴田寛さんとは、共同研究をさせていただく機会をいただいた。お二人との共同研究を通して、自分の研究の視野を大きく広げることができた。また、日本学術振興会には、研究生活に資金的な援助をいただいた。

最後に、私の大学院進学を認め、多くの面でサポートをしてくれた家族に対して、心より感謝の意を伝えたい。

2015年1月6日

竹島 康博

引 用 文 献

- Adhikari, B., Goshorn, E., Lamichhane, B., & Dhamala, M. (2013). Temporal-order judgment of audiovisual events involves network activity between parietal and prefrontal cortices. *Brain Connectivity*, 3, 536-545.
- Alais, D., & Burr, D. (2004a). No direction-specific bimodal facilitation for audiovisual motion detection. *Cognitive Brain Research*, 19, 185-194.
- Alais, D., & Burr, D. (2004b). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14, 257-262.
- Alais, D., Newell, F. N., & Mammassian, P. (2010). Multisensory processing in review: From physiology to behavior. *Seeing and Perceiving*, 23, 3-38.
- Alink, A., Singer, W., & Muckli, L. (2008). Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *The Journal of Neuroscience*, 28, 2690-2697.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, 4, 267-278.
- Alvarez, G., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, 15, 106-111.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. 5th ed. Washington, DC: APA.
- Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fusion and fusion illusions. *Cognitive Brain Research*, 21, 301-308.
- Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, 14, 589-592.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, 61, 183-193.
- Barraclough, N. E., Xiao, D., Baker, C. I., Oram, M. W., & Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *Journal of Cognitive Neuroscience*, 17, 377-391.

- Benjamins, S. J., Van der Smagt, M. J., & Verstraten, F. A. J. (2008). Matching auditory and visual signals: Is sensory modality just another feature? *Perception*, 37, 848-858.
- Bolognini, N., Frassinetti, F., Serino, A., & Ladavas, E. (2005). "Acoustical vision" of below threshold stimuli: Interaction among spatially converging audiovisual inputs. *Experimental Brain Research*, 160, 273-282.
- Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, 114, 38-70.
- Breitmeyer, B. G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research*, 15, 1411-1412.
- Breitmeyer, B., & Julesz, B. (1975). The role of on and off transients in determining the psychophysical spatial frequency response. *Vision Research*, 15, 411-415.
- Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldordff, M. G. (2005). The spread of attention across modalities and space in multisensory object. *Proceedings of the National Academy of Sciences, USA*, 102, 18751-18756.
- Carello, C., Anderson, K. L., & Kunkler-Peck, A. J. (1998). Perception of object length by sound. *Psychological Science*, 9, 211-214.
- Cesaroni, L., & Garber, M. (1991). Exploring the experience of autism through firsthand accounts. *Journal of Autism and Developmental Disorders*, 21, 303-313.
- Chen, Y. -C., & Spence, C. (2011). The crossmodal facilitation of visual object representations by sound: Evidence from the backward masking paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 1784-1802.
- Chen, Y. -C., & Yeh, S. -L. (2008). Visual events modulated by sound in repetition blindness. *Psychonomic Bulletin & Review*, 15, 404-408.
- Chen, Y. -C., & Yeh, S. -L. (2009). Catch the moment: Multisensory enhancement of rapid visual events by sound. *Attention, Perception, & Psychophysics*, 198, 209-219.
- Clavagnier, S., Falchier, A., & Kennedy, H. (2004). Long-distance feedback projections to area V1: Implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, Affective, & Behavioral Neuroscience*, 4, 117-126.
- Clement, D. E., & Vanadoe, K. W. (1967). Pattern uncertainty and the discrimination of visual patterns. *Perception & Psychophysics*, 2, 427-431.

- Chiou, R., & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception*, *41*, 339-353.
- Enns, J., & DiLollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*, 345-352.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429-434.
- Fairhall, S. L., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *European Journal of Neuroscience*, *29*, 1247-1257.
- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *The Journal of Neuroscience*, *22*, 5749-5759.
- Frassinetti, F., Bolognini, N., & Ladavas, E. (2002). Enhancement of visual perception by cross-modal visuo-auditory interaction. *Experimental Brain Research*, *147*, 332-343.
- Foss-Feig, J., Kwakye, L., Cascio, C., Burnette, C., Kadivar, H., Stone, W., & Wallace, M. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, *203*, 381-389.
- Foxe, J., & Molholm, S. (2009). Ten years at the Multisensory Forum: Musings on the evolution of a field. *Brain topography*, *21*, 149-154.
- Foxe, J. J., & Schroeder, C. (2005). The case for feedforward multisensory convergence during early cortical processing. *NeuroReport*, *16*, 419-423.
- Fujisaki, W., & Nishida, S. (2005). Temporal frequency characteristics of synchrony-asynchrony discrimination of audio-visual signals. *Experimental Brain Research*, *166*, 455-464.
- Fujisaki, W., & Nishida, S. (2010). A common perceptual temporal limit of binding synchronous inputs across different sensory attributes and modalities. *Proceedings of the Royal Society B*, *277*, 2281-2290.
- Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, *7*, 773-778.
- Garner, W. R., & Clement, D. (1963). Goodness of pattern and pattern uncertainty. *Journal of Verbal Learning and Verbal Behavior*, *2*, 446-452.
- Garner, W. R., & Sutliff, D. (1974). The effect of goodness on encoding time in visual pattern discrimination. *Perception & Psychophysics*, *16*, 426-430.

- Grassi, M., & Casco, C. (2009). Audiovisual bounce-inducing effect: Attention alone does not explain why the discs are bouncing. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 235-243.
- Guo, K., Liu, C., & Roebuck, H. (2011). I know you are beautiful even without looking at you: Discrimination of facial beauty in peripheral vision. *Perception*, 40, 191-195.
- Guski, R., & Troje, N. F. (2003). Audiovisual phenomenal causality. *Perception & Psychophysics*, 65, 789-800.
- 行場次朗 (2007). 感覚の心理. 阿部恒之 (編) 心理学の視点 20. 国際文献印刷社 pp. 37-44.
- 行場次朗・寺本渉 (2012). 臨場感と迫真性. 映像メディア学会誌, 66, pp. 561-563.
- 花沢俊明 (2007). 神経生理Ⅱ—高次の視覚領野—. 内川恵二・篠森敬三 (編). 視覚Ⅰ 視覚系の構造と初期機能. 朝倉書店 pp. 45-63.
- Hidaka, S., Manaka, Y., Teramoto, W., Sugita, Y., Miyauchi, R., Gyoba, J., Suzuki, Y., & Iwaya, Y. (2009). Alternation of sound location induces visual motion perception of a static object. *PLoS ONE*, 4, e8188.
- Hoffmann, K. -P. (1973). Conduction velocity in pathways from retinal to superior colliculus in the cat: A correlation with receptive-field properties. *Journal of Neurophysiology*, 36, 409-424.
- Howe, E. S. (1980). Effects of partial symmetry, exposure time, and backward masking on judged goodness and reproduction of visual patterns. *The Quarterly Journal of Experimental Psychology*, 32, 27-55.
- 入野俊夫・津崎実 (2007). 聴覚初期過程の機能モデル. 大山正・今井省吾・和氣典二・菊地正 (編). 新編 感覚・知覚心理学ハンドブック Part 2. 誠信書房 pp. 349-358.
- Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the “Ventriloquism” effect. *Perceptual and Motor Skills*, 37, 967-979.
- Jardri, R., Pins, D., Bubrovszky, M., Lucas, B., Lethuc, V., Delmaire, C., Vantighem, V., Desprez, P., & Thomas, P. (2009). Neural functional organization of hallucinations in schizophrenia: Multisensory dissolution of pathological emergence in consciousness. *Consciousness and Cognition*, 18, 449-457.
- Kanwisher, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition*, 27, 117-143.

- Kayaert, G., & Wagemans, J. (2009). Delayed shape matching benefits from simplicity and symmetry. *Vision Research*, 49, 708-717.
- Kayser, C., & Logothetis, N. (2007). Do early sensory cortices integrate cross-modal information? *Brain Structure and Function*, 212, 121-132.
- Kayser, C., Petkov, C., & Logothetis, N. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex*, 18, 1560-1574.
- Kim, J., Doop, M., Blake, R., & Park, S. (2005). Impaired visual recognition of biological motion in schizophrenia. *Schizophrenia Research*, 77, 299-307.
- King, A. J. (2010). Multisensory integration: Strategies for synchronization. *Current Biology*, 15, R339-R341.
- Kirchner, H., & Thorpe, S. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, 46, 1762-1776.
- Kitagawa, N., & Ichihara, S. (2002). Hearing visual motion in depth. *Nature*, 416, 172-174.
- Klapetek, A., Ngo, M. K., & Spence, C. (2012). Does crossmodal correspondence modulate the facilitatory effect of auditory cues on visual search? *Attention, Perception, & Psychophysics*, 74, 1154-1167.
- Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, 71, 856-867.
- Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica*, 134, 372-384.
- Kulikowski, J. J., & Tolhurst, D. J. (1973). Psychophysical evidence for sustained and transient detectors in human vision. *The Journal of Physiology*, 232, 149-162.
- Kuling, I. A., Van Eijk, R. L. J. & Juola, J. F., & Kohlrausch, A. (2012). Effects of stimulus duration on audio-visual synchrony perception. *Experimental Brain Research*, 221, 403-412.
- Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive Brain Research*, 16, 468-478.
- Lipscomb, S. D., & Kim, E. M. (2004). Perceived match between visual parameters and auditory correlates: An experimental multimedia investigation. In *Proceedings of the 8th International Conference on Music Perception & Cognition*, (pp. 72-75). Evanston, IL: Society for Music Perception and Cognition.

- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. (2nd ed.). London: Psychology Press.
- Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology*, 14, R990-R991.
- Magnée, M., Oranje, B., Engeland, H., Kahn, R., & Kemner, C. (2009). Cross-sensory gating in schizophrenia and autism spectrum disorder: EEG evidence for impaired brain connectivity? *Neuropsychologia*, 47, 1728-1732.
- Makovski, T., & Jiang, Y. (2008). Indirect assessment of visual working memory for simple and complex objects. *Memory & Cognition*, 36, 1132-1143.
- Manabe, K., & Riquimaroux, H. (2000). Sound controls velocity perception of visual apparent motion. *Journal of the Acoustical Society of Japan (E)*, 21, 171-174.
- Marković, S., & Gvozdenovi, V. (2001). Symmetry, complexity and perceptual economy: Effects of minimum and maximum simplicity conditions. *Visual Cognition*, 8, 305-327.
- Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal Factors. *The Journal of Neuroscience*, 7, 3215-3229.
- Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, 221, 389-391.
- Meredith, M. A., & Stein, B. E. (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research*, 365, 350-354.
- Meyer, G. F., & Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *NeuroReport*, 12, 2557-2560.
- Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *The Journal of Neuroscience*, 27, 4120-4131.
- Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research*, 17, 154-163.
- Noesselt, T., Bergmann, D., Hake, M., Heinze, H. -J., & Fendrich, R. (2008). Sound increases the saliency of visual events. *Brain Research*, 1220, 157-163.
- Olivers, C. N. L., & Van der Burg, E. (2008). Bleeping you out of the blink: Sound saves vision from oblivion. *Brain Research*, 1242, 191-199.

- Olson, I., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion*, 5, 498–502.
- Oyama, T., Kikuchi, T., & Ichikawa, S. (1981). Span of attention, backward masking, and reaction time. *Perception & Psychophysics*, 29, 106-112.
- Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a ‘low road’ to ‘Many roads’ of evaluating biological significance. *Nature Reviews Neuroscience*, 11, 773-782.
- Peschel, A., & Orquin, J. (2013). A review of the findings and theories on surface size effects on visual attention. *Frontiers in Psychology*, 4, Article 902.
- Rauschenberger, R., & Yantis, S. (2006). Perceptual encoding efficiency in visual search. *Journal of Experimental Psychology: General*, 135, 116-131.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849-860.
- Remijn, G. B., Ito, H., & Nakajima, Y. (2004). Audiovisual integration: An investigation of the ‘Streaming-bounding’ phenomenon. *Journal of Physiological Anthropology and Applied Human Science*, 23, 243-247.
- Rockland, K. S., & Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, 50, 19-26.
- Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topography*, 21, 185-192.
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Molholm, S., Javitt, D. C., & Foxe, J. J. (2007). Impaired multisensory processing in schizophrenia: Deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophrenia Research*, 97, 173-183.
- Sadaghiani, S., Maier, J. X., & Noppeney, U. (2009). Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *The Journal of Neuroscience*, 29, 6490-6499.
- Saito, D., Yoshimura, K., Kochiyama, T., Okada, T., Honda, M., & Sadato, N. (2005). Cross-modal binding and activated attentional networks during audio-visual speech integration: A functional MRI study. *Cerebral Cortex*, 15, 1750-1760.

- Sanabria, D., Spence, C., & Soto-Faraco, S. (2007). Perceptual and decisional contributions to audiovisual integrations in the perception of apparent motion: A signal detection study. *Cognition*, 102, 299-310.
- Santangelo, V., & Spence, C. (2008). Crossmodal attentional capture in an unspeeded simultaneity judgment task. *Visual Cognition*, 16, 155-165.
- 佐藤弥 (2004). 知覚前情動処理により知覚を促進する神経機構. 心理学評論, 47, pp. 51-66.
- Sekular, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, 385, 308.
- Setti, A., & Chan, J. (2011). Familiarity of objects affects susceptibility to the sound-induced flash illusion. *Neuroscience Letters*, 492, 19-22.
- Shams, L., Iwaki, S., Chawla, A., & Bhattacharya, J. (2005). Early modulation of visual cortex by sound: An MEG study. *Neuroscience Letters*, 378, 76-81.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, 408, 788.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, 14, 147-152.
- Shams, L., Kamitani, Y., Thomposon, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *NeuroReport*, 12, 3849-3852.
- Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *NeuroReport*, 16, 1923-1927.
- Shenton, M., Dickey, C., Frumin, M., & McCarley, R. (2001). A review of MRI findings in schizophrenia. *Schizophrenia Research*, 49, 1-52.
- Sheth, B., & Shimojo, S. (2000). In space, the past can be recast but not the present. *Perception*, 29, 1279-1290.
- Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science*, 145, 1328-1330.
- Soto-Faraco, S., Spence, C., & Kingstone, A. (2004). Cross-modal dynamic capture: Congruency effects in the perception of motion across sensory modalities. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 330-345.
- Soto-Faraco, S., Spence, C., & Kingstone, A. (2005). Assessing automaticity in the audiovisual integration of motion. *Acta Psychologica*, 118, 71-92.

- Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science & Technology*, 28, 61-70.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971-995.
- Stein, B. E. (2012). Introduction. In Stein, B. E. (Ed.), *The New Handbook of Multisensory Processing* (pp. xi-xvi). London: The MIT Press.
- Stein, B. E., London, N., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience*, 8, 497-506.
- Sugihara, T., Diltz, M., Averbeck, B., & Romanski, L. (2006). Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *The Journal of Neuroscience*, 26, 11138-11147.
- Sugita, Y., & Suzuki, Y. (2003). Implicit estimation of sound-arrival time. *Nature*, 421, 911-912.
- Sui, J., & Liu, C. (2008). Can beauty be ignored? Effects of facial attractiveness on covert attention. *Psychonomic Bulletin & Review*, 16, 276-281.
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26, 212-215.
- Swanson, W. H., & Birch, E. E. (1992). Extracting thresholds from noisy psychophysical data. *Perception & Psychophysics*, 51, 409-422.
- Takeshima, Y., & Gyoba, J. (2011). Modification of velocity perception by loud sounds. *The Japanese Journal of Psychonomic Science*, 30, 19-24.
- Takeshima, Y., & Gyoba, J. (2013a). High-intensity sound increases the size of visually perceived objects. *Attention, Perception, & Psychophysics*, 75, 501-507.
- Takeshima, Y., & Gyoba, J. (2013b). Changing pitch of sounds alters perceived visual motion trajectory. *Multisensory Research*, 26, 317-312.
- Takeshima, Y., & Gyoba, J. (2013c). Complexity of visual stimuli affects visual illusion induced by sound. *Vision Research*, 91, 1-7.
- Takeshima, Y., & Gyoba, J. (2014). Pattern dot quantity affects auditory facilitation effects on visual object representations. *Perception*, 43, 107-116.

- Takeshima, Y., & Gyoba, J. (in press). Spatial frequency modulates the degree of illusory second flash perception. *Multisensory Research*.
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14, 400-410.
- Tamietto, M., & de Gelder, B. (2011). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, 12, 302-302.
- 寺本 渉・吉田和博・日高聡太・浅井暢子・行場次朗・坂本修一・岩谷幸雄・鈴木陽一 (2010). 「迫真性」を規定する時空間情報. 日本ヴァーチャルリアリティ学会論文誌, 15, pp. 483-486.
- Todd, J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, 428, 751-754.
- Todd, J., & Marois, R. (2005). Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 144-155.
- Tolhurst, D. (1973). Separate channels for the analysis of the shape and the movement of moving visual stimulus. *The Journal of Physiology*, 231, 385-402.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Audiovisual events capture attention: Evidence from temporal order judgments. *Journal of Vision*, 8 (5), 1-10.
- VanRullen, R., & Thorpe, S. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, 13, 454-461.
- Vroomen, J., & de Gelder, B. (2003). Visual motion influences the contingent auditory motion aftereffect. *Psychological Science*, 14, 357-361.
- Vroomen, J., & de Gelder, B. (2004). Temporal ventriloquism: Sound modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 513-518.
- Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception & Psychophysics*, 72, 871-884.
- Wada, Y., Kitagawa, N., & Noguchi, K. (2003). Audio-visual integration in temporal perception. *International Journal of Psychophysiology*, 50, 117-124.

- Wallace, M. T. (2012). The impact of multisensory alterations in human developmental disabilities and disease: The tip of an iceberg? In Stein, B. E. (Ed.) *The New Handbook of Multisensory Processing* (pp. 645-656). London: The MIT Press.
- Watanabe, K., & Shimojo, S. (1998). Attentional modulation in perception of visual motion events. *Perception*, 27, 1041-1054.
- Watkins, S., Shams, L., Josephs, O., & Rees, G. (2007). Activity in human V1 follows multisensory perception. *NeuroImage*, 37, 572-578.
- Watkins, S., Shams, L., Tanaka, S., Haynes, J. -D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage*, 31, 1247-1256.
- Welch, R. B., & Wallen, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88, 638-667.
- Wuerger, S. M., Hofbauer, M., & Meyer, G. F. (2003). The integration of auditory and visual motion signals at threshold. *Perception & Psychophysics*, 65, 1188-1196.
- Zampini, M., Guest, S., Shore, D., & Spence, C. (2005) Audio-visual simultaneity judgments. *Perception & Psychophysics*, 67, 531-544.