



Title	自動的文字-音統合の処理についての検討 : 二つの分析方法と読み習得における役割
Author(s)	エン, 冬陽; Yan, Dongyang
Citation	北海道大学大学院教育学研究院紀要, 140, 1-24
Issue Date	2022-06-25
DOI	https://doi.org/10.14943/b.edu.140.1
Doc URL	https://hdl.handle.net/2115/86253
Type	departmental bulletin paper
File Information	04-1882-1669-140.pdf



自動的文字-音統合の処理についての検討

— 二つの分析方法と読み習得における役割 —

エン冬陽*

【要旨】 文字-音 (L-SS) 統合に関する神経生理学的手法を用いた研究についてシステマティック・レビューを行い、その分析方法と母語や第二言語 (L2) の読み習得におけるL-SS統合の役割を検討した。2021年6月までに発表された論文のうち、25本の論文が該当した。2つの分析アプローチのうち、additive modelを用いた研究ではモデルから想定される早期の反応が確認されておらず、このモデルを間接的な神経活動の測定に用いるには限界のあることが示唆された。一方、一致効果は多くの研究で安定した結果が確認され、一致効果の方向性には読み経験と言語の規則性の両方が関わると考えられた。また、母語が背景としてL2の読み習得におけるL-SS統合の役割に影響を与えることが示唆された。今後、L2における読み困難の理解のために、母語やL2の特性を踏まえて、L2の習熟度とL-SS統合の関係を検討していく必要がある。

【キーワード】 読み、文字-音の統合、多感覚統合、視聴覚統合、ディスレクシア、fMRI、EEG/MEG

1. 背景

1.1. 読みの習得における文字-音の統合の重要性

文字を流暢に読むスキルは、現代の生活や教育に不可欠である。表音文字の一種であるアルファベット書記体系では、単語は音素（子音もしくは母音）を表す個々の文字で構成されている。定型発達児が最初に読むことを学び始めたとき、文字と音に対応する視聴覚の結合表象が習得される。この結合表象を習得するためには、まず個々の文字を識別して音との対応を覚える。そして、親近感のない文字列に出会ったとき、この知識を適用する。親近感のない単語を文字ごとに読み上げる練習をするにつれて、この過程はますます自動化され、効率的になる。また、この文字-音の結合表象が自動的に活性化されること（文字-音の統合：letter-speech sound integration, 以下L-SS統合）により、文字から音への効率的な変換が可能となる (Ehri, 2005)。従って、L-SS統合は熟達した読み手になるための重要なステップである。L-SS統合を含む多感覚統合の発達には、連合野の成熟が必要であるだけでなく、多感覚回路が感覚入力を用いてかなりの学習をすることも必要である (Stein et al., 2014)。このため、この結合表象から特定の情報を自動的に抽出できるようになるには、かなり長い時間がかかる。読みの重要な部分である自動化されたL-SS統合と多感覚統合の発達を把握することは、母語における読み習得の過程の理解だけでなく、L2の習得やディスレクシアの障害の理解にとっても不可欠である。

* 北海道大学大学院教育学院

ディスレクシアは、音韻障害を特徴とすることが広く知られている (e.g., Snowling, 2000)。Blomert (2011) と Richlan (2012) は、必ずしも音韻障害の知見に異議を唱えるわけではないものの、ディスレクシアの原因として自動的な文字-音統合の異常という代替理論を提案した。自動的な文字-音統合は、文字に対応する音を知っていることとは異なる。MMN (Mismatch Negativity) を用いた Froyen ら (2011) の研究では、ディスレクシア児では、どんな音に対応するかを知っている文字に対しても、文字-音の結合表象の自動的な活性化が認められなかった。Blomert (2011) は L-SS 統合に関する行動と神経生理学的研究のレビューを実施し、発達性ディスレクシアのある子どもでは L-SS 結合表象を形成できないことが、自動的な活性化が起こりにくい背景であると述べた。ディスレクシアの家族性リスクのある就学前の子どもを対象とした研究も、L-SS の結合表象を形成するためにより長期間の読み訓練が必要であることを示している (e.g., Karipidis et al., 2018)。さらに、近年の神経生理学と神経画像研究のレビューも、ディスレクシアのある人の脳活動が定型発達者と異なることを示し、L-SS 統合に異常があることを示唆している (Richlan, 2019)。

L-SS 統合は L2 の読み習得にも関わると推測されるが、L2 の習得と L-SS 統合の関係はまだ十分に明らかにされていない。また、L-SS 統合が流暢な読みの習得における重要な要因となるのであれば、L2 習得に困難を抱える学習者の中にディスレクシアと共通する L-SS 統合の異常が見られる可能性がある。

1.2. L-SS 統合に関わる脳領域

神経画像研究により、AV (audiovisual) の多感覚統合に関連するいくつかの脳領域が特定されている。L-SS 統合においても、両側の上側頭溝 (STS) と上側頭回 (STG) 領域にある多感覚領域 (parabelt cortex) に加えて聴覚皮質 (Heschl's sulcus/planum temporale) などの領域も関わるのが van Atteveldt ら (2004, 2007a) により報告された。さらに、下頭頂小葉 (IPL) と後頭側頭皮質 (OT) の活動も自動的な L-SS 統合に関わる (e.g., Brem et al., 2010; Schurz et al., 2014; Richlan, 2012)。特に、左腹側 OT は文字の特異的な処理や視覚正書法表象の形成に関与するとされるが、その活動は新たな視聴覚結合の習得やその後の自動化の影響を受ける (e.g., Madec et al., 2016; Brem et al., 2018)。

脳領域間の機能的結合に関する Richlan (2019) の最近のレビューは、下前頭領域、聴覚野および STG/STS などの多感覚領域と、左腹側 OT を連絡する連合線維束である弓状束 (arcuate fasciculus) が L-SS 統合において重要であると述べている。なお、従来、弓状束は下前頭領域 (ブローカ野) と側頭葉 (ウェルニッケ野) を結ぶ神経線維束を指し、上縦束 (superior longitudinal fasciculus) と一部を共有する。最近の拡散強調画像法 (DTI) や解剖学的研究から得られた知見は、下前頭野、下頭頂葉、側頭葉といった領域が上縦束や弓状束などの背側連合線維を通じて連絡され、さらに頭頂葉と側頭葉が下縦束 (inferior longitudinal fasciculus) の腹側連合線維で連絡されることを示している (Dick & Tremblay, 2012)。Richlan (2019) のレビューにおける“弓状束”は、この前頭野から頭頂葉を経て側頭葉と結ぶ言語処理ネットワークに関わる白質線維束を指している。

1.3. 分析における 2 つのアプローチ

fMRI や MEG/EEG の神経生理学的手法を用いたこれまでの研究では、一致効果 (congruency

effect) と additive モデル (additive model) という 2 つの異なるアプローチを用いて、lipreading などの音声の AV 統合や L-SS 統合を含む AV 統合の研究が行われている。

AV 音声統合研究においては、音声と一致した口の動き (一致条件)、音声と不一致な口の動き (不一致条件) が呈示され、一致刺激に対する反応と不一致刺激に対する反応の違いが一致効果として評価される。一方、additive モデルを用いた解析では、単感覚条件として音声 (A 条件) と音声を発する口の動き (V 条件)、多感覚条件として音声と一致した口の動き (AV 条件) が呈示され、単感覚条件に対する反応と多感覚刺激に対する反応を比較する。

これらのパラダイムは L-SS 統合を調べるさまざまな実験デザインでも用いられている。L-SS 統合研究において一致効果を評価する場合には、文字と文字の音が一致して提示された場合と、文字と文字の音が不一致な場合の反応の違いを比較する。additive モデルを用いた解析では、音のみ (A 条件)、文字 (V 条件)、文字とその文字と対応する音声 (AV 条件) が呈示される。さまざまな実験デザインが用いられているが、ほとんどの研究は受動的デザインで行われている。

1) 一致効果 (congruency effects)

AV 音声統合研究においては、マガーク効果という現象が知られている (McGurk and MacDonald, 1976)。マガーク効果とは、聴覚音声と異なる口の動きを見た場合に、不一致な口の動きに引きずられて聴覚情報が異なる音声として認識される現象である。すなわち、不一致条件において、視覚刺激によって聴覚処理に影響を受けることを示している。このことから、一連の L-SS 統合の研究が示す一致効果も、視覚文字が音声の初期処理に影響を与えたことを意味すると考えられている。具体的には、変化しない情報である視覚文字の入力処理よりも音声の入力処理が困難であるため、視覚文字が大きな手がかりとなり、STS と STG にある多感覚領域から一次聴覚野へのフィードバックによって、聴覚皮質での音声処理に影響を与えると考えられる。この証拠はサルに関する研究からも得られ、視覚刺激が一次や二次聴覚野の聴覚処理に影響を与えることが示されている (Kayser et al., 2007)。すなわち、一致効果が生じるのは、多感覚表象が視覚刺激によって自動的に活性化されるためであり、L-SS 統合が確立していることの証拠であると解釈される。

しかしながら、一致効果の逆パターンを示す研究もある (不一致効果)。不一致条件では、異なる聴覚、視覚刺激やその両方に反応する多くの神経処理過程が行われる可能性がある (Hickok, 2012)。例えば、不一致条件では前頭葉から頭頂葉に広がる聴覚の背側経路の活動の増加が見られることがメタ分析から報告されている。この経路はエラー検出などに関わることから、視覚刺激と聴覚刺激が一致しない場合に、これらの矛盾した情報の解決に関わると考えられる (Erickson et al., 2014)。そのため、一致効果の方向性 (一致効果か不一致効果か) に加えて、生じる部位を評価することが重要である。

なお、AV 刺激の不一致性を用いて L-SS 統合を調べる実験デザインとして、MMN パラダイム (規則な刺激配列が連続呈示する際に突然新規な刺激を単発提示する) を用いた研究がある。MMN 研究における不一致効果は、神経適応 (Grill-Spector & Malach, 2001) または多感覚反復抑制 (Henson, 2003) に起因すると考えられる。繰り返し呈示され、過剰学習された AV 刺激 (一致条件) の刺激に対しては細胞の活動が徐々に減少していく「順応」が生じるが、不一致刺激に対しては活性化の大幅な増加が観察される。L-SS 研究の MMN パラダイムを用

いた研究では文字と一致する音が呈示され、不一致条件ではどちらか一方（通常は音）が変化する。MMN研究での不一致効果は上記の不一致効果とは異なり、一致条件のL-SS結合表象が既に習得され、その処理が自動化されていることを示すものである。

2) additiveモデル (additive models)

多感覚統合の基礎的な機能を理解するために神経生理学的研究で用いられるもう1つのアプローチはadditiveモデルを用いたものである。このモデルはlipreadingなどのAV音声統合に関する研究にも用いられている (e.g., Calvert, Campbell, and Brammer, 2000)。具体的には、多感覚AV刺激に対する反応と、単感覚の聴覚刺激に対する反応と単感覚の視覚刺激に対する反応の総和を比較するものである[AV versus (A+V)]。このモデルは多感覚ニューロンの活動パターンを前提とする。additiveモデルを用いた多感覚統合の早期の研究は、動物の上丘におけるニューロンのリアルタイム活動に注目し、単一ユニット記録または局所電場電位のニューロンの発火率を収集することによって実施された（単一ユニット記録：Meredith and Stein 1983, 1985；局所電場電位：Mitzdorf 1987；Buzsaki and Draguhn 2004）。

これらの基礎的研究に基づき、JamesとStevenson (2012) は、多感覚脳領域には3種類の異なるニューロンがあると推定した。1つは単感覚ニューロンであり、感覚入力の一つのモダリティのみで有意な神経活動を生成し、この反応は他の感覚モダリティからの同時入力によって変調されない。もう一つはbimodal (or trimodal) ニューロンである。このニューロンは2つ以上の単感覚入力のモダリティに対し、有意な神経活動を生じる (Stein and Stanford 2008)。ほとんどのbimodalニューロンは、同時入力に対して2つの単感覚刺激の最大の反応よりも大きな活動を引き起こす[$AV > \text{Max}(A,V)$ ：多感覚増強]が、一部のbimodalニューロンは、多感覚抑制という単感覚刺激への最大の反応よりも低い活動を生じる[$AV < \text{Max}(A,V)$]。多感覚の増強と抑制は、いずれも情報が統合されていることを示す指標になる。最後のニューロンの種類はsubthresholdニューロンであり、このタイプは2つのパターンの活動を示す。単感覚刺激のみの場合、単感覚ニューロンと同じく一種類の単感覚入力のみに賦活を示すが、多感覚刺激の場合、bimodalニューロンのように単感覚条件より大きな賦活を示す (Allman and Meredith 2007；Allman et al. 2008；Meredith and Allman 2009)。多感覚条件で誘発されたbimodal neuronやsubthreshold neuronsの多感覚の増強または抑制はいずれも情報統合のサインと考えて良いことから (James&Stevenson, 2012)、AV統合の研究では、supra-additiveモデル [$AV > (A + V)$]とsub-additiveモデル[$AV < (A + V)$]が統合の指標として用いられる (Poeppl et al., 2014)。

しかし、Calvert, Campbell, Brammer (2000) が、AV音声統合 (lipreading) において左側のSTSにおけるBOLD信号がsupra-additiveモデルとsub-additiveモデルとも満たすことを報告したものの、多くの研究では同様の結果は示されていない (e.g., Beauchamp et al., 2004)。BOLD信号は異種ニューロンの大きなクラスターから生じ、bimodalニューロンまたはsubthresholdニューロンの単一ユニットだけではない (Laurienti et al., 2005)。さらに、単感覚ニューロンと比較して多感覚ニューロンの割合は少なく、これらのsupra-additiveを示す多感覚ニューロンは、sub-additive性 [$AV < (A + V)$] または抑制性 [$AV < \text{Max}(A,V)$] を示す他のニューロンと平均化される可能性がある (Stevenson et al., 2014)。このため、Hocking&Price (2008) は、より緩い基準 (e.g., [$AV > \text{Max}(A,V)$]) をAV統合の基準と

して受け入れるべきであると提案した。

Additiveモデルを用いた分析アプローチは読みにおけるAV統合（L-SS統合）の研究にも用いられている。しかしながら、視聴覚ともに変化するlipreadingなどのAV音声統合の基準を、L-SS統合に適応することができるのかについては検討する必要がある。

1.4. 目的

このレビューの目的は、一致効果とadditiveモデルを用いた読みにおける自動的なL-SS統合研究における脳活動パターンを比較し、この2つの分析アプローチの利点や欠点を評価することである。また、L2における読み困難の背景を理解するために、定型発達者の母語におけるL-SS統合の発達の軌跡とL2の習得におけるL-SS統合の役割を明らかにすることである。

2. 方法

このレビューでは、MEG, EEG, fMRIを用いて行われた、定型発達者の母国語の読みとL2の読みにおけるL-SS統合の役割を検討した研究について評価する。なお、定型発達者を対照群としてディスレクシアについて検討された研究も含んだ。ガイドラインとして、the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) (Page et al., 2021) に従った (Figure 1を参照)。

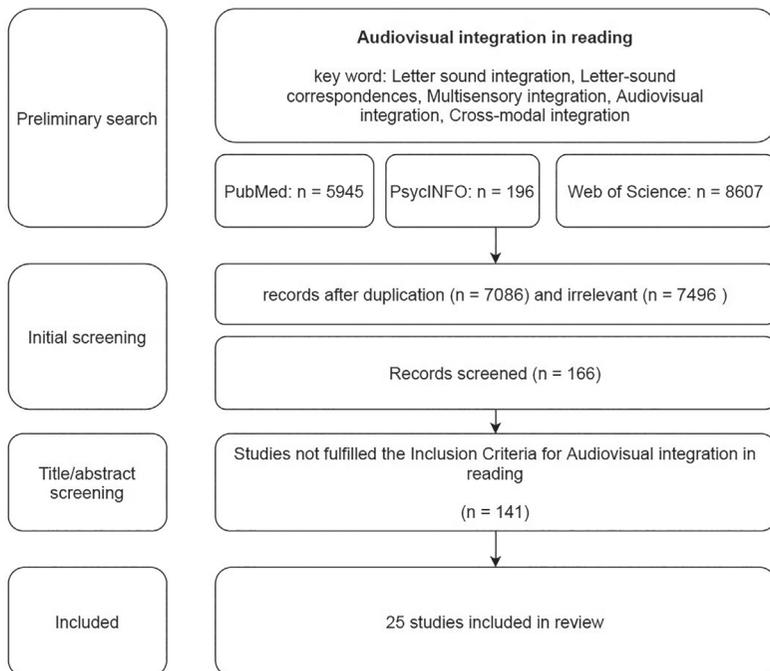


Figure 1. 検索とスクリーニングに関するフローチャート

Chart 1. 基準：

- 1) fMRI, MEG / EEGなどを研究手法として利用した研究。
 - 2) 課題が文字と単語のL-SS統合を測定する研究（例:文字, 音節, 母音-子音-母音 (CVC) の文字列, アルファベットの単語または偽単語)
 - 3) 読みに関連する課題を用いた研究（例, 語彙の決定, 意味の判断, 韻の判断など）
 - 4) ディスレクシアがある場合, DSM-IV (American Psychiatric Association, 2000) や ICD-10 (World Health Organization, 2007) を基準としてディスレクシアを評価する研究
 - 5) 実際にあるアルファベット言語フォントが刺激として使われているもの（偽文字が刺激としている研究は含まない）
 - 6) 一致効果とadditiveモデルいずれの分析方法を用いた研究
 - 7) 英語で書かれた研究のみ
-

まず, PubMed, PsycINFO, とWeb of Science のオンラインデータベースで2021年6月までに公開された論文を検索し, 本稿に用いる研究を検索した。Letter sound integration (n=200), Letter-sound correspondences (n=212), Multisensory integration (n=8110), Audiovisual integration (n=3961), Cross-modal integration (n=2265) をキーワードとしてデータベースに入力した。この最初の検索で14748件の研究が見つかった。次に, 7086件の重複研究と7496件の無関係な研究を除外した。除外された研究の多くは, 視聴覚以外の多感覚統合に関するものや, 細胞レベルでの多感覚統合に関するものであった。

残りの166の研究の要約とタイトルをChart 1 に示すような基準によってレビューした。基準の適用後, 分析のために25の論文が残った。次に, 原文と要約を具体的にレビューした。

3. 結果

本稿においてレビューした論文のリストをTable 1～3 に示した (fMRI研究についてはTable 1 を, EEGについてはTable 2 を, MEG研究についてはTable 3 を参照)。Table 1～3 では定型発達者 (TD) とディスレクシア (Dyslexia) を対象とした研究を分けて示している。各Tableには, 実験課題や参加者などの情報も示した。

fMRI研究は9本, EEG/MEG研究は15本あり, fMRIとEEGの同時計測研究は1本あった。一致効果を使った研究は25本あった。このうち, 通常の一一致効果を示す研究が14本 (fMRI研究が10本, EEG/MEG研究が4本, fMRIとEEGの同時計測研究が1本), MMNを用いた研究が8本, その他が3本であった。一方, additiveモデルを用いた研究は7本であり, そのうち, fMRI研究が3本, EEG/MEG研究が3本, fMRIとEEGの同時計測研究が1本であった。L2学習者を対象とした研究はMEGを用いた研究1本とMMNを用いた2本のみであり, その他は母語話者を対象としたものであった。

4. 考察

4.1. 二つの分析方法の比較

fMRI研究では一貫して、多感覚領域と聴覚皮質における一致効果が報告されていた (Table 1)。また、レビューしたほとんどのEEG/MEG研究では、頭頂側頭領域におけるかなり遅い時間帯 (約400ms) での一致効果が観察された (Table 2 研究 4, Table 3 研究 1, 3)。序論で述べたとおり、一致効果では、視覚情報により多感覚表象が活性化され、その後、多感覚領域から聴覚野へのフィードバックによって音声の処理が影響されるようになる。L-SS統合に関するfMRIを用いたeffective connectivity研究も、STSと聴覚野の結合性の程度が一致・不一致条件間での聴覚野の活動の違いを予測できるという結果を示しており、多感覚領域からのフィードバックが一致効果の背景であることを裏付ける (van Atteveldt et al., 2009)。このため、一致効果は比較的遅い時間帯で出現すると考えるのが妥当であろう。従って、この時間帯での一致効果はL-SS統合の結果により生じるものと解釈できる。

additiveモデルは、多感覚統合の一般的な進行中の処理を前提として使われてきた。序論で述べたように、supra-additiveモデルとsub-additiveモデルとも、AV統合の処理過程を解釈できる可能性がある。L-SS統合研究の数が少ないものの、fMRIを用いた研究においてsupra-additiveの基準を満たしたのはvan Atteveldtら (2004) の研究でのみである (Table 1 研究 1)。fMRIとEEGの同時計測によるKronshabelら (2014) のディスレクシアにおけるadditiveモデルを用いた唯一の研究 (Table 2 研究 9) では、ディスレクシアの子どもは、初期 (230~300ms) に定型発達者と同様のパターン (i.e., sub-additive効果) を示したが、後期成分での群間差を認め、300ms後に左後頭側頭領域にsupra-additive効果を示した。これが事実であれば、時間的分解能が低いfMRIでは、additiveモデルによって発達性ディスレクシアにおけるL-SS統合の欠陥を調べるのが難しいかもしれない。なぜなら、初期と後期の活動が加算され、この結果、定型発達群と比較した場合、群間差がないことを示している可能性がある。また、fMRI信号は単一のニューロンから生じるものではなく、異種ニューロンの大きなクラスターから生じる間接的なものである。このような間接的な神経活動の測定から生じる信号から多感覚脳領域のニューロンの正確な活動を推定することは難しいと思われる。

一方、ほとんどのEEG/MEGでは、一致AV刺激に対するsupra-additive効果ではなく、sub-additive効果が認められている (Table 2 研究 4, 9 と Table 3 研究 1, 2)。EEG/MEG研究では一致AV刺激に対して230ms以降にsub-additive効果がみられた。Stevensonら (2014) は、約200ms以降は感覚処理に厳密には関連していない (Hillyard et al., 1998) ことから、分析期間を200ms以内に制限することを提案した。また、Besleら (2008) による人間の頭蓋内EEGからは、40~200msの間でのsub-additive効果が確認されている。さらに、lipreadingにおけるAV音声統合を調べた研究では、50msから始まるsub-additive効果が観察された (Klucharev et al., 2003)。したがって、200ms以降に観察されたsub-additive効果は注意して扱う必要があると考えられる。

4.2. 一致効果の方向性

実験デザインや刺激の種類が異なるにもかかわらず、一致効果を用いた研究では多くの研究において比較的一貫した結果が得られていた。Hollawayら (2015) の研究を除き、定型発達

の成人を対象としたほとんどの神経生理学的研究では、一致した刺激に対する両側のPT/STG (STS) での反応が増加することを示している。一方、子どもを対象とした研究では不一致効果が観察された。我々は就学前の子どもと定型発達者からの研究の分析を通して、一致効果の方向性の違いは読み能力と言語の規則性に起因すると推定した。

一致効果の方向性と読み能力に関しては、子どもを対象とする研究から推察できる。一連の研究から、数か月の読み指導の後に不一致効果が現れることが示され、一致効果が出現するためには、長期間のトレーニングが必要であることが示唆される (e.g., Plewko et al., 2018; Wang et al., 2020; Caffarra et al., 2021; Figure. 2を参照; Table 1, 2, 3)。この不一致効果は子どもたちが文字と音の連合 (association) から自動L-SS統合への発達が完全に達成されていないためであると推測できる。子どもにおける不一致効果が異なる視聴覚情報の処理の抑制困難によるものではないことは、不一致刺激の抑制に関連する領域であるACCの活性化の増加が観察されないことから示唆される。さらに、不一致条件では子どもたちの下側頭回 (ITG; 腹側OTに相当) の活動の増加が見られた。序論でも述べたように、腹側OTは文字の特異的な処理に関連する。多感覚領域からフィードバック信号を送ることで、子どものITGにおけるL-SS統合の自動化の低さを補っている可能性があると考えられる (King et al., 2019)。

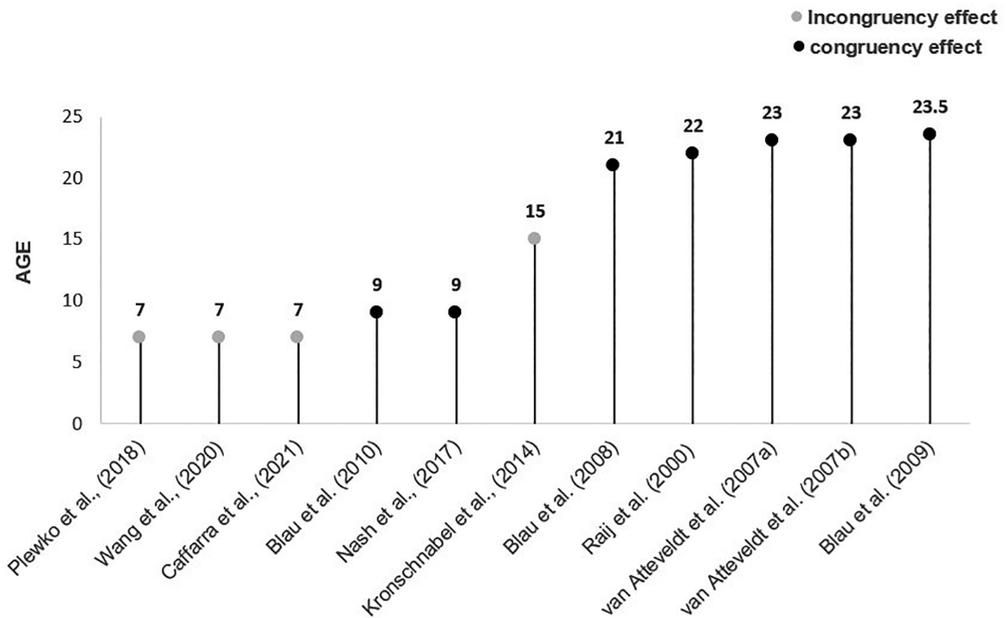


Figure 2. 年齢による一致効果の方向性

一方で、不一致効果から一致効果へいつ移行するのかについてはまだ不明である。重要なことに、不一致効果は、1年間の指導を受けた子どもで観察されている (Wang et al., 2020; Table 1 研究10)。Blauら (2010) の研究に参加した約2年の正式な識字教育を受けたと推察できる9歳の子どもは、不一致効果ではなく、一致効果を示した。同じ年齢であるが、より若い年齢で教育を始めた (つまり、より長い読み経験を持つ) イギリスの子どもも、同様の一致効果を示した (Nash et al., 2017; Table 2 研究11)。しかし、この研究では、不一致条件ではなく、一致条件をベースライン条件と比較したことに注意する必要がある (他のほとんどの研究は一致条件と不一致条件を比べて、一致効果を示している)。不一致効果から一致効果への移行を明らかにするには、学童期の子どもを対象としたさらなる縦断研究が必要であると考えられる。

なお、Kronschnabelら (2014) の研究では2年以上の読み指導を受けたと考えられる定型発達者において、単一の文字に対しては一致効果が観察できず、CVC文字列に対して不一致効果がみられている (Table 1 研究8)。この結果はドイツ語話者のCVC文字列の読み処理において、単一の文字とは異なる音韻処理などが行われたことが影響したと考えられる。

一致効果の方向性と言語の規則性に関しては、Hollowayら (2015) の研究から推測できる (Table 1 研究5)。文字と音の関係が1対1である言語は規則性が高く、文字と音の関係が複数対複数の言語は規則性が低いと考えられる。Hollowayら (2015) は文字と音の対応関係が規則性の低い言語 (英語) でのL-SS統合を調べており、STGにおけるL-SS刺激による不一致効果を示した。この研究で用いられた視聴覚刺激は英語における最も規則的なL-SS刺激であり、これは、英語の音素レベルでの不規則性が、最も規則的なL-SS刺激の処理にも影響していることを示唆している。聴覚刺激によるMMNを用いた研究においても、偽単語の統計的規則性 (その言語における音素配列の確率の高さ) により異なる反応が示されており、規則性の異なる言語は音声の神経処理の確立に異なる影響を与える可能性がある (Bonte et al., 2005)。

4.3. MMN研究による発達の軌跡

対象とした論文には8本のMMNを用いた研究が含まれた (Table 2 研究1~3, 5~8&10)。このうち3本の研究が、L-SS統合の発達の時間経過を調べることを目的にしていた (Table 2 研究2, 7, 8)。

定型発達成人 (Froyen et al., 2008; Andres et al., 2011; Jones et al., 2016; Table 2 研究1, 3, 10) においては100~250msに早期のMMNの増加を示した。この結果は早期の自動的なL-SS統合の証拠として解釈された。4年間の読み指導を受けた11歳のオランダの子どもたちにおいても、早期のMMNの増加が確認されている (Froyen et al., 2009, Table 2 研究2)。さらに、2.5年間の読み指導を受けたオランダの子ども (平均年齢9.0歳) でも、早期のMMNの増加が観察されている (Žarić et al., 2014, Table 2 研究8)。一方、1年間の指導を受けた8歳の子どもでは、MMNの増加は、300~700msのかなり遅い期間 (後期MMN) に現れ、その振幅が読み能力とともに向上した (Froyen et al., 2009)。この後期のMMNは自動化されていない文字と音の連合を示唆し、読み能力の向上とともに早期のMMNに移行すると考えられる。

これらの結果から、正式な読み指導を受けた後、2.5年間の読み指導を受けるまでに未熟な文字と音の連合から完全に自動化された統合への段階的な変化が起こる可能性があると考えられる。

きる。MMN研究でみられた発達に伴う変化は、先に述べた不一致効果から一致効果への移行と類似した結果を示していると言える。

4.4. L2 研究

L2読みに関する研究はEEGを用いた研究が2本、MEGを用いた研究が1本のみであった (Table 2 研究 5, 6 ; Table 3 研究 3)。MEGを用いたCaffarraら (2021) の研究から早期バイリンガルの子どもたちについて示された証拠は、L1とL2がともに文字と音の対応関係が規則的である書記体系 (Spanish/Basque) である場合には、両言語で同じようなL-SS統合処理が行われることを示唆している (Table 3 研究 3)。このことから、L-SS統合は流暢なL2の読み習得にも重要な役割を果たしていると考えられる。

しかし、MMN研究を用いた2つの研究からは、後期L2学習者におけるL-SS統合の発達は、言語の規則性及び母語によって影響を受ける可能性があるように思われる。Yangら (2016) のMMN研究は、韓国成人である英語学習者において早期MMNの増加を見出した (Table 2 研究 5)。この結果は、4.3で述べたFroyenら (2009) の11歳のオランダ語母語話者児童の研究結果と一致していた。一方、同様の英語能力を持つ中国人学習者は、早期MMNの増加を示していない。Froyenら (2009) の研究で用いられた後期MMNデータがないため、読み指導の年数が少ないオランダ語母語話者の子どもでみられた遅い時間帯でのL-SS統合が、中国人英語学習者においてみられるかどうかについては明らかでない。さらに、英語能力の高い中国人英語学習者を含んで行われたWangら (2019) らのMMN研究でも同様の結果が観察され、英語能力の程度に関わらず、中国人英語学習者では早期MMNが認められないことが示された (Table 2 研究 6)。

これらの結果から、母国語の背景が音素レベルのL2を読む際の脳の活動に影響することが考えられる。書記素と音素の対応が明確である韓国語とは異なり、中国語は書記素と音素の対応規則のない表語文字である。Caoら (2009) は中国の子どもが書記素に対して敏感であり、音素への依存度が低いと述べている。書記体系の違いのため、中国の読み手がアルファベットを読む際に、既に確立された母語読みの学習方略 (L-SS統合に依存しない) を利用する可能性がある。このために自動的なL-SS統合が中国の英語学習者では生じにくいと考えられる。

5. 結論と限界

本稿の目的はL-SS統合研究に一般的に使用されてきた二つの分析アプローチについて検討し、この2つのアプローチの利点や欠点を評価することであった。また、L2における読み困難の背景を理解するために、定型発達者の母語におけるL-SS統合の発達の軌跡とL2の習得におけるL-SS統合の役割を明らかにすることであった。

今回のレビューを通して、次のようなことが明らかになった。まず、多感覚領域の単一ニューロンの活動に関する研究から得られたモデルであるadditiveモデルは、多感覚統合の進行中の処理を示唆する早期の反応が確認されておらず、EEG/MEG, fMRIのような間接的な神経活動の測定から得られた信号の分析に用いるには限界があると考えられる。一方で、多感覚領域から聴覚領域へのフィードバックによる一致効果は、多くの研究で安定して確認されており、

EEG/MEGやfMRI研究において用いやすいアプローチと言える。定型発達者の母語におけるL-SS統合の発達軌跡に関しては、一致効果の方向性の違いに読み経験が関わると考えられた。また、母語の規則性も一致効果の方向性に寄与することが示された。L2の読み習得に関しては、L-SS統合は流暢なL2の読み習得にも関わるが、母語の背景がL2のL-SS統合の役割に影響を与えることが示唆された。ただし、本稿から示唆された言語の規則性や母語からL2への影響については研究が少なく、さらに検証を行う必要がある。

今回のレビューでは、ディスレクシアにおけるL-SS統合の異常については考察しなかった。しかし、L2習得に困難を抱える学習者の中にディスレクシアと共通するL-SS統合の異常がみられるかどうか検討するためには、定型発達者の母語とL2の読み習得におけるL-SS統合の役割を理解する必要がある。本研究はその基礎となるものである。今後、L2の規則性などの言語特性や母語の背景がL2の学習に与える影響を考慮しつつ、L2の習熟度とL-SS統合の関係を明確にする必要がある。

謝辞

本研究は、JST次世代研究者挑戦的研究プログラムJPMJSP2119の支援を受けたものである。

引用文献

- Allman, B. L., & Meredith, M. A. (2007). Multisensory processing in “unimodal” neurons: cross-modal subthreshold auditory effects in cat extrastriate visual cortex. *Journal of neurophysiology*, 98(1), 545-549.
- Allman, B. L., Keniston, L. P., & Meredith, M. A. (2008). Subthreshold auditory inputs to extrastriate visual neurons are responsive to parametric changes in stimulus quality: sensory-specific versus non-specific coding. *Brain research*, 1242, 95-101.
- Andres, A. J., Cardy, J. E. O., & Joanisse, M. F. (2011). Congruency of auditory sounds and visual letters modulates mismatch negativity and P300 event-related potentials. *International Journal of Psychophysiology*, 79(2), 137-146.
- Bakos, S., Landerl, K., Bartling, J., Schulte-Körne, G., & Moll, K. (2017). Deficits in letter-speech sound associations but intact visual conflict processing in dyslexia: results from a novel ERP-paradigm. *Frontiers in human neuroscience*, 11, 116.
- Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, 41(5), 809-823. [https://doi.org/10.1016/s0896-6273\(04\)00070-4](https://doi.org/10.1016/s0896-6273(04)00070-4)
- Bernstein, L. E., Auer Jr, E. T., Moore, J. K., Ponton, C. W., Don, M., & Singh, M. (2002). Visual speech perception without primary auditory cortex activation. *Neuroreport*, 13(3), 311-315.
- Bernstein, L. E., & Liebenthal, E. (2014). Neural pathways for visual speech perception. *Frontiers in neuroscience*, 8, 386.
- Besle, J., Fischer, C., Bidet-Caulet, A., Lecaigard, F., Bertrand, O., & Giard, M. H. (2008). Visual activation and audiovisual interactions in the auditory cortex during speech perception:

- intracranial recordings in humans. *Journal of Neuroscience*, 28(52), 14301-14310.
- Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: a functional magnetic resonance imaging study of dyslexic children. *Brain*, 133(3), 868-879.
- Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Current Biology*, 19(6), 503-508.
- Blau, V., van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2008). Task-irrelevant visual letters interact with the processing of speech sounds in heteromodal and unimodal cortex. *European journal of neuroscience*, 28(3), 500-509.
- Blomert, L. (2011). The neural signature of orthographic-phonological binding in successful and failing reading development. *Neuroimage*, 57(3), 695-703.
- Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., ... & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proceedings of the National Academy of Sciences*, 107(17), 7939-7944.
- Brem, S., Hunkeler, E., Mächler, M., Kronschnabel, J., Karipidis, I. I., Pleisch, G., & Brandeis, D. (2018). Increasing expertise to a novel script modulates the visual N1 ERP in healthy adults. *International Journal of Behavioral Development*, 42(3), 333-341.
- Bonte, M. L., Mitterer, H., Zellagui, N., Poelmans, H., & Blomert, L. (2005). Auditory cortical tuning to statistical regularities in phonology. *Clinical Neurophysiology*, 116(12), 2765-2774.
- Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *science*, 304(5679), 1926-1929.
- Caffarra, S., Lizarazu, M., Molinaro, N., & Carreiras, M. (2021). Reading-related brain changes in audiovisual processing: cross-sectional and longitudinal MEG evidence. *Journal of Neuroscience*.
- Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D., & David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*, 10(12), 2619-2623.
- Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current biology*, 10(11), 649-657.
- Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral cortex*, 11(12), 1110-1123.
- Cao, F., Peng, D., Liu, L., Jin, Z., Fan, N., Deng, Y., & Booth, J. R. (2009). Developmental differences of neurocognitive networks for phonological and semantic processing in Chinese word reading. *Human brain mapping*, 30(3), 797-809.
- Dick, A. S., & Tremblay, P. (2012). Beyond the arcuate fasciculus: consensus and controversy in the connective anatomy of language. *Brain*, 135(12), 3529-3550.
- Fitzgerald, K., & Todd, J. (2020). Making sense of mismatch negativity. *Frontiers in Psychiatry*, 11, 468.
- Froyen, D., Van Atteveldt, N., Bonte, M., & Blomert, L. (2008). Cross-modal enhancement of the

- MMN to speech-sounds indicates early and automatic integration of letters and speech-sounds. *Neuroscience letters*, 430(1), 23-28.
- Froyen, D. J., Bonte, M. L., van Atteveldt, N., & Blomert, L. (2009). The long road to automation: neurocognitive development of letter-speech sound processing. *Journal of Cognitive Neuroscience*, 21(3), 567-580.
- Froyen, D., Willems, G., & Blomert, L. (2011). Evidence for a specific cross-modal association deficit in dyslexia: an electrophysiological study of letter-speech sound processing. *Developmental science*, 14(4), 635-648.
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of reading*, 9(2), 167-188.
- Erickson, L. C., Heeg, E., Rauschecker, J. P., & Turkeltaub, P. E. (2014). An ALE meta-analysis on the audiovisual integration of speech signals. *Human brain mapping*, 35(11), 5587-5605.
- Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta psychologica*, 107(1-3), 293-321.
- Henson, R. N. (2003). Neuroimaging studies of priming. *Progress in neurobiology*, 70(1), 53-81.
- Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences*, 95(3), 781-787.
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature reviews neuroscience*, 13(2), 135-145.
- Hocking, J., & Price, C. J. (2008). The role of the posterior superior temporal sulcus in audiovisual processing. *Cerebral Cortex*, 18(10), 2439-2449.
- Holloway, I. D., van Atteveldt, N., Blomert, L., & Ansari, D. (2015). Orthographic dependency in the neural correlates of reading: evidence from audiovisual integration in English readers. *Cerebral Cortex*, 25(6), 1544-1553.
- James, T. W., & Stevenson, R. A. (2012). The use of fMRI to assess multisensory integration. In *The Neural Bases of Multisensory Processes*. CRC Press/Taylor & Francis.
- Jones, M. W., Kuipers, J. R., & Thierry, G. (2016). ERPs reveal the time-course of aberrant visual-phonological binding in developmental dyslexia. *Frontiers in human neuroscience*, 10, 71.
- Jost, L. B., Eberhard-Moscicka, A. K., Frisch, C., Dellwo, V., and Maurer, U. (2014). Integration of spoken and written words in beginning readers: a topographic ERP study. *Brain Topogr.* 27, 786-800. doi: 10.1007/s10548-013-0336-4
- Karipidis, I. I., Pleisch, G., Röthlisberger, M., Hofstetter, C., Dornbierer, D., Stämpfli, P., & Brem, S. (2017). Neural initialization of audiovisual integration in prereaders at varying risk for developmental dyslexia. *Human brain mapping*, 38(2), 1038-1055.
- Karipidis, I. I., Pleisch, G., Brandeis, D., Roth, A., Röthlisberger, M., Schneebeli, M., ... & Brem, S. (2018). Simulating reading acquisition: the link between reading outcome and multimodal brain signatures of letter-speech sound learning in prereaders. *Scientific reports*, 8(1), 1-13.
- Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2007). Functional imaging reveals visual modulation of specific fields in auditory cortex. *Journal of Neuroscience*, 27(8), 1824-1835.
- Kayser, C., & Shams, L. (2015). Multisensory causal inference in the brain. *PLoS biology*, 13(2),

- e1002075.
- Kemény, F., Gangl, M., Banfi, C., Bakos, S., Perchtold, C. M., Papousek, I., ... & Landerl, K. (2018). Deficient letter–speech sound integration is associated with deficits in reading but not spelling. *Frontiers in human neuroscience*, 12, 449.
- King, A. J., Hammond-Kenny, A., & Nodal, F. R. (2019). Multisensory Processing in the Auditory Cortex. In *Multisensory Processes* (pp. 105–133). Springer, Cham.
- Klucharev, V., Möttönen, R., & Sams, M. (2003). Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Cognitive Brain Research*, 18(1), 65–75.
- Kronshabel, J., Brem, S., Maurer, U., & Brandeis, D. (2014). The level of audiovisual print–speech integration deficits in dyslexia. *Neuropsychologia*, 62, 245–261.
- Lakatos, P., Chen, C. M., O’Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, 53(2), 279–292.
- Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., & Stein, B. E. (2005). On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Experimental Brain Research*, 166(3), 289–297.
- Maded, S., Le Goff, K., Anton, J. L., Longcamp, M., Velay, J. L., Nazarian, B., ... & Rey, A. (2016). Brain correlates of phonological recoding of visual symbols. *NeuroImage*, 132, 359–372.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746–748.
- Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, 221(4608), 389–391.
- Meredith, M. A., & Stein, B. E. (1985). Descending efferents from the superior colliculus relay integrated multisensory information. *Science*, 227(4687), 657–659.
- Meredith, M. A., & Allman, B. L. (2009). Subthreshold multisensory processing in cat auditory cortex. *Neuroreport*, 20(2), 126.
- Mitzdorf, U. (1987). Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex. *International Journal of Neuroscience*, 33(1–2), 33–59.
- Murray, M. M., & Wallace, M. T. (2011). *The neural bases of multisensory processes*. CRC Press.
- Nash, H. M., Gooch, D., Hulme, C., Mahajan, Y., McArthur, G., Steinmetzger, K., & Snowling, M. J. (2017). Are the literacy difficulties that characterize developmental dyslexia associated with a failure to integrate letters and speech sounds?. *Developmental Science*, 20(4), e12423.
- Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: visual articulatory information enables the perception of second language sounds. *Psychological research*, 71(1), 4–12.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Bmj*, 372.
- Plewko, J., Chyl, K., Bola, Ł., Łuniewska, M., Dębska, A., Banaszkiwicz, A., ... & Jednoróg, K. (2018). Letter and speech sound association in emerging readers with familial risk of dyslexia.

- Frontiers in human neuroscience, 12, 393.
- Poeppl, D., Overath, T., Popper, A. N., & Fay, R. R. (2014). The human auditory cortex. Springer.
- Raij, T., Uutela, K., & Hari, R. (2000). Audiovisual integration of letters in the human brain. *Neuron*, 28(2), 617-625.
- Richlan, F. (2012). Developmental dyslexia: dysfunction of a left hemisphere reading network. *Frontiers in human neuroscience*, 6, 120.
- Richlan, F. (2019). The functional neuroanatomy of letter-speech sound integration and its relation to brain abnormalities in developmental dyslexia. *Frontiers in Human Neuroscience*, 13, 21.
- Schurz, M., Kronbichler, M., Crone, J., Richlan, F., Klackl, J., & Wimmer, H. (2014). Top-down and bottom-up influences on the left ventral occipito-temporal cortex during visual word recognition: An analysis of effective connectivity. *Human brain mapping*, 35(4), 1668-1680.
- Snowling, M. J. (2000). *Dyslexia*. Blackwell publishing.
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. The MIT Press.
- Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature reviews neuroscience*, 9(4), 255-266.
- Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews Neuroscience*, 15(8), 520-535.
- Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., ... & Wallace, M. T. (2014). Identifying and quantifying multisensory integration: a tutorial review. *Brain topography*, 27(6), 707-730.
- van Atteveldt, N., Formisano, E., Goebel, R., Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron* 43, 271-282.
- van Atteveldt, N., Formisano, E., Blomert, L., Goebel, R. (2007a). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cereb. Cortex* 13, 962-974.
- van Atteveldt, N., Formisano, E., Goebel, R., Blomert, L. (2007b). Top-down task effects overrule automatic multisensory responses to letter-sound pairs in auditory association cortex. *Neuroimage* 36, 1345-1360.
- van Atteveldt, N., Roebroek, A., & Goebel, R. (2009). Interaction of speech and script in human auditory cortex: insights from neuroimaging and effective connectivity. *Hearing research*, 258(1-2), 152-164.
- Wang, C., Yang, Z., Cao, F., Liu, L., & Tao, S. (2019). Letter sound integration in native Chinese speakers learning English: Brain fails in automatic responses but succeeds with more attention. *Cognitive neuroscience*, 10(2), 100-116.
- Wang, F., Karipidis, I. I., Pleisch, G., González, G. F., & Brem, S. (2020). Development of print-speech integration in the brain of beginning readers with varying reading skills. *Frontiers in Human Neuroscience*, 14, 289.
- Xu, W., Kolozsvari, O. B., Monto, S. P., & Hämäläinen, J. A. (2018). Brain responses to letters and speech sounds and their correlations with cognitive skills related to reading in children. *Frontiers in human neuroscience*, 12, 304.
- Yang, Z., Wang, C., Feng, L., Li, S., & Tao, S. (2016). Unsuccessful letter-sound integration in

English reading by native Chinese speakers: Evidence from an event related potentials study. *Science Bulletin*, 61(24), 1855–1864.

Žarić, G., Gonzalez, G. F., Tijms, J., van der Molen, M. W., Blomert, L., & Bonte, M. (2014). Reduced neural integration of letters and speech sounds in dyslexic children scales with individual differences in reading fluency. *PLoS One*, 9(10), e110337.

Table 1 A summary of empirical findings of fMRI studies investigating automatic audiovisual integration

Details of the study		Evidence of multisensory integration								
Typical readers										
Authors	Participant status	Sample size	Language (L1/L2)	Task	AV stimulus	Significant "congruency effect" or "incongruent effect"?	Detail for congruency effect	AVc (bimodal) > A or V (unimodal) (AV > max (A, V))?	AVc > A + V (i.e., super-additivity)? Or AVc < (A + V) (i.e., subadditivity)?	Other detail
van Atteveldt et al.(2004)	TD adults	16	Dutch	Passive perception for simultaneously presenting AV	vowel/ consonant	Yes: bilateral HS/PT	Congruent AV > A > Incongruent AV	Yes: bilateral STG/STS, bilateral PT/HS	Yes: bilateral PT/HS (supra-additivity)	-
van Atteveldt et al. (2007a)	TD adults	8	Dutch	Passive perception for AV with varying SOA of ± 300 ms	vowel/ consonant	Yes: bilateral PT, bilateral aSTP (SOA=0)	Congruent AV (SOA=0) A > Incongruent AV	Yes: bilateral PT, bilateral aSTP, left STS (SOA=0)	<i>Not reported</i>	-
van Atteveldt et al. (2007b)	TD adults	25	Dutch	Passive perception for simultaneously presenting AV Active matching with varying SOA of ± 300 ms	vowel/ consonant	Yes: bilateral STG, bilateral PT, left aSTP	Congruent AV > A > Incongruent AV	<i>Not reported</i>	-	-
Biau et al.(2008)	TD adults	19	Dutch	Auditory identification task	vowel	Yes: left aSTG, left medial MTG, bilateral FG, bilateral PT/HS	Congruent AV > A > Incongruent AV	Yes: left aSTG, bilateral HS / PT, left MTG / STS, bilateral FG	<i>Not reported</i>	-
Holloway et al.(2015)	TD adults	18	English	Passive perception for simultaneously presenting AV	irregular letter phoneme (LP) fairly transparent letter name (LN) transparent digit number (NN)	Yes: bilateral pSTG, left aSTG, right STG	Incongruent AV > Congruent AV	Yes: bilateral pSTG, left aSTG, right STG	-	-
						No	Congruent AV = Incongruent AV	<i>Not reported</i>	-	-
						Yes: right STG/left aSTG/ left pSTG	Congruent AV > Incongruent AV	-	-	-

Dyslexia													
6	Blau et al. (2009)	TD adults	13	Dutch	Passive perception for simultaneously presenting AV	vowel							
		Dyslexic adults	13	Dutch				No	Yes: bilateral STG				Not reported
7	Blau et al. (2010)	TD children (aged 9)	16	Dutch	Passive perception for simultaneously presenting AV	vowel							
		Dyslexic children (aged 9)	18	Dutch				No	Yes: left HS/ PT, bilateral STS				Yes: left HS/ PT, bilateral STS
8	Kronshäbel et al. (2014)	TD adolescents (aged 15)	22	German	Target detection task (fMRI & EEG)	single letter/CVC							
		Dyslexic adolescents (aged 15)	13	German				As above	Yes: left ITG, left STG				Not reported
9	Plewko et al. (2018)	TD children (aged 7)	35	Polish	Passive perception for simultaneously presenting AV	vowel/ consonant							
		Prereading children at risk for dyslexia (aged 7)	50	Polish				As above	Yes: left PT/STG, right ITG				Not reported

10	Wang et al.(2020)	TD children (aged 7)	32	German	Target detection task pronounceable nonwords	Yes: left IFG, right MTG, left ITG	Incongruent AV > Congruent AV	Yes: left IFG, right MTG, left ITG	-
		Dyslexic adolescents (aged 7)	8	German		Yes: right MFG/ITG	Congruent AV > Incongruent AV	Yes: right MFG/ITG	-

Superior temporal gyrus (STG) / superior temporal sulcus (STS) / planum temporale (PT) / Heschl's sulcus (HS) / anterior superior temporal plane (aSTP) / middle temporal gyrus (MTG) / fusiform gyrus (FG) / inferior temporal gyrus (ITG) / inferior frontal gyrus (IFG) / medial frontal gyrus (MFG)

Audio (A)/ Visual (V)/ Audiovisual (AV)

Typical development (TD)

Table 2 A summary of empirical findings of EEG studies investigating automatic audiovisual integration

Details of the study		Evidence of multisensory integration					AVc > A + V (i.e., super-additivity)? Or AVc < (A + V) (i.e., subadditivity)?		
Authors	Participant status	Sample size	Language (L1/L2)	Task	AV stimulus	Significant "congruency effect" or "incongruent effect"?	Detail for congruency effect	Other detail	
EEG (Typical readers)									
Native									
1	Froven et al. (2008)	TD adults	Dutch	oddball	letter	Yes: earlier MMN at 100–250ms (max SOA=0)	Incongruent AV > Congruent AV	Not reported	-
2	Froven et al. (2009)	TD children (aged 8)	Dutch	oddball	letter	Yes: late MMN at 600ms (SOA=200)	Incongruent AV > Congruent AV	Not reported	-
3	Andres et al. (2011)	TD adults	English	oddball vowel decision task	letter	Yes: early MMN at 100–250ms (SOA=200)/late MMN at 600ms (SOA=0)	Incongruent AV > Congruent AV	Not reported	-
4	Jost et al. (2014)	TD children (aged 7)	German	oddball vowel decision task	word	Yes: early MMN / P300	Congruent AV > Incongruent AV (MMN) / Incongruent AV > Congruent AV (P300)	Not reported	-
Nonnative									
5	Yang et al. (2016)	Korean English adults	Korean/English	oddball	letter	Yes: early MMN at 100–250ms (SOA=200)	Incongruent AV > Congruent AV	Not reported	-
6	Wang et al. (2019)	Intermediate Chinese English adults	Chinese/English	oddball	letter	No	-	Yes: audiovisual suppression effects (AVc < A + V) at 300–324, 480–764ms in centroparietal electrode	-
		Advanced Chinese English adults	Chinese/English	oddball	letter	Yes: P300	Incongruent AV > Congruent AV		-
		Native English adults	English	oddball	letter	Yes: P300	Incongruent AV > Congruent AV	Not reported	-
				oddball	letter	Yes: early MMN at 100–250ms (SOA=0, 200)/P300	Incongruent AV > Congruent AV		-

Dyslexia											
7	Froyen et al. (2011)	Dyslexic children (aged 11)	16	Dutch	oddball	letter		Yes: late MMN at 600ms (SOA=200) Yes: early MMN at 100-250ms (SOA=0, 200)/ late MMN at 600ms (SOA=0, 200)	Incongruent AV > Congruent AV	Not reported	-
8	Žarić et al. (2014)	TD children (aged 9) Mildly dyslexic (aged 9) Severely dyslexic (aged 9)	20 18 18	Dutch Dutch Dutch	oddball oddball oddball	letter letter letter		Yes: early MMN at 100-250ms (SOA=0, 200) Yes: early MMN at 100-250ms (SOA=200)	Incongruent AV > Congruent AV Incongruent AV > Congruent AV	Not reported	-
9	Kronshabel et al. (2014)	TD adolescents (aged 15) Dyslexic adolescents (aged 15)	22 13	German German			Target detection task (fMRI & EEG) CVC	Yes: P1 at the left temporooccipital site No	Congruent AV > Incongruent AV -	Yes: audiovisual suppression effects (AVc < A + V) at left superior temporal electrodes (230-300 ms) Yes: audiovisual suppression effects (AVc < A + V) at left superior temporal electrodes (230-300 ms); supra-additivity (AVc < A + V) at left superior temporal electrodes (300-800ms)	-
10	Jones et al. (2016)	TD adults Dyslexic adults	18 18	English English	letter- matching task	letter		Yes: PNM/P3b/LRP Yes: P3b	Incongruent AV > Congruent AV Incongruent AV > Congruent AV	Not reported	-
11	Nash et al. (2017)	Age-matched TD children (aged 7-13) Reading-age-matched TD children (aged 7-13)	17 17	English English	Priming task	letter		Yes: P1 in left frontal electrodes Yes: P2 in the central posterior region	Congruent AV > Baseline AV Congruent AV > Baseline AV	No unimodal stimuli	-
12	Bakos et al. (2017)	Dyslexic children (aged 7-13) TD children (aged 9) Dyslexic children (aged 9)	13 37 36	English German German	Stroop-like interference paradigm	letter		Yes: P1 over fronto-central electrodes and P2 in the central posterior region Yes: N1 and conflict slow potential (cSP)	Congruent AV > Baseline AV Incongruent AV > Congruent AV	No unimodal stimuli	-

13 Kermény et al. (2018)	TD children (aged 9)	21	German		Yes: cSP	Incongruent AV > Congruent AV	-
	Children with combined reading and spelling deficit (aged 9)	10	German	Stroop-like interference paradigm letter	No	-	No unimodal stimuli
	Children with isolated spelling deficit (aged 9)	17	German		Yes: cSP	Incongruent AV > Congruent AV	-

Table 3. A summary of empirical findings of MEG studies investigating automatic audiovisual integration

Details of the study		Evidence of multisensory integration								
Authors	Participant status	Sample size	Language (L1/L2)	Task	AV stimulus	Significant "congruency effect" or "incongruent effect"?	Detail for congruency effect	AVc > A + V (i.e., super-additivity)? Or AVc < (A + V) (i.e., subadditivity)?	Other detail	
EEG (Typical readers)										
Native										
1	Raij et al. (2000)	TD adults	8	Finnish	Monitoring task	letter	Yes: left and right STS (415-515ms)	Congruent AV > Incongruent AV	Yes: audiovisual suppression effects (AVc < A + V) in left and right temporal and occipital areas (325-535 ms)	AV > A in left STS
2	Xu et al. (2018)	TD children (aged 6-11)	29	Finnish	Target detection task	letter	No	Congruent AV = Incongruent AV	Yes: audiovisual suppression effects (AVc < A + V) in bilateral inferior parietal, bilateral supramarginal, right precuneus, right postcentral, right superior and middle temporal cortices (after 250ms)	-
Nonnative										
3	Caffarra et al. (2021)	TD pre-school children (aged 4-9) TD school children (aged 4-9)	20 22	Spanish/ Basque	Target detection task	syllables	No	Incongruent AV > Congruent AV	Not reported	-

A systematic review of letter–speech sound integration: two analysis models and reading acquisition

Dongyang Yan

Key Words

Reading, Letter–speech sound integration, Multisensory integration, Audiovisual integration, Dyslexia, fMRI, EEG/MEG

Abstract

A systematic review of studies was undertaken to explore research investigating letter–speech sound (L-SS) integration via neurophysiological techniques (EEG, MEG, fMRI). Through this review, we are attempting to build a coherent comprehension of two commonly used evaluation analysis (i.e., congruency effects, additive models) and the role of L-SS integration in the reading acquisition of both native and second languages (L2). We collected 25 empirical studies published by June 2021 focusing on L-SS integration in typical readers and dyslexia. These results showed that the apparent inconsistency may arise by utilizing additive models, suggesting the limitations by using this model in EEG/MEG and fMRI studies. On the other hand, the findings of many neuroimaging studies showed a high degree of consistency in congruency effects. We noted that the direction of congruency effects (i.e., congruency/incongruency effects) tends to be attributed to reading experience and native language transparency. Meanwhile, the influence of native language background might be found in brain responses to L-SS integration during L2 reading. To understand the difficulties of L2 reading, the relationship between L2 proficiency and L-SS integration, in accordance with the linguistic features of native language and L2, need to be investigated in future research.