

日本語の単語認識における文字位置の重要性の違い

梅津 由紀子・宮下 雅史・森本和紗

大正大学心理社会学部人間科学科

指導教員：井関 龍太

要旨：Huebert & Cleary (2022) の実験では、単語のファミリーの判断において内側の文字と比べて外側の文字が重要であると示された。しかし、日本語は分かち書きを採用していないため、英語と違いファミリーの判断は文字の位置に影響されない可能性が考えられる。本研究では、四字熟語を断片の状態にし、穴埋めテストをした後に元の単語のファミリーを評価する実験を行った。実験の結果、日本語であっても外側の文字が内側の文字よりもファミリーに特権的に影響していることが示された。

問 題

日常生活において何かしらの文を読むとき、文字の位置が入れ替わっていたり、誤字や脱字が多少あったりしても、文の読解過程でそれらの語の意味を推測することで、正しく語を認識して読めることがある。特に文中にある単語の最初と最後の文字以外の順番が入れ替わっても正しく文が読めてしまう現象はタイポグリセミアと呼ばれている。中尾清月堂は、2018年にこのタイポグリセミアを利用した広告を出した。広告には「みなさまに だじいなおらせ。こたのび なかお せいげどつう が ぜったいに ばれない ように どやらきの リニューアル を おなこい ました。」と書かれていた（中尾清月堂, 2018）。この文章は単語内の文字の順番が入れ替えられているが、「みなさまに だじいなおらせ。このたび なかお せいげつどう が ぜったいに ばれない ように どらやきの リニューアル を おこない ました。」という正しい文章で読むことができる。

以下では、区別のために、単語の最初と最後の文字を外側の文字、それ以外を内側の文字と呼ぶことにする。読解の際の単語認識のメカニズムについて Johnson & Eisler (2012) は、文中にある単語の単語内での文字の位置を入れ替えることによって文の処理にどのような影響を与えるかを調べた。Johnson & Eisler の実験では参加者に、提示する文をできるだけ

速く正確に読ませ、読解時間を測定した。この時、提示される文の中には単語内の文字の位置が入れ替わっている転置のある文と転置のない通常の文があった。転置のある文は、それぞれ単語の最初の 2 文字が転置している文、最後の 2 文字が転置している文、内側の 2 文字が転置している文であった。結果として、単語の最初の 2 文字が転置された文と最後の 2 文字が転置された文は、内側の 2 文字が転置された文よりも読むのが遅くなった。このことは、単語の認識や文全体の読解において、外側の文字が内側の文字に比べ重要であることを示唆している。また、このような文字の位置情報について外側の文字の方が内側の文字よりも役割の重みがあることを以下では、外側の文字の特権性と呼ぶことにする。

単語の読解時以外における外側の文字の特権性について、Huebert & Cleary (2022) は、ファミリーリティにおいても外側の文字が重視されることを明らかにした。ファミリーリティとは、刺激への馴染みの程度のことである。例えば、単語の場合、見覚えがあるけれどもどこで見たか思い出せない場合には意識的な記憶はないがファミリーリティを感じているといえる。ファミリーリティを感じるには、その対象の細部を思い出す必要はない。断片的な記憶の情報と目の前の刺激が照合され、その適合具合がファミリーリティとして感じられる。Huebert & Cleary の実験では、参加者は、まず学習段階として単語断片を手がかりとして単語全体を完成させる穴埋め課題を行った。この際、提示された単語断片には外側の位置にある文字を残し内側の文字を欠落させた外側断片（例：「OVERALL」は「OV__LL」）と内側の文字を残し外側の文字を欠落させた内側断片（例：「VACCINE」は「_AC_IN_」）の 2 つの条件があった。その後、テスト段階として、穴の開いていない単語についてのファミリーリティを 11 段階で評価した。この単語には、学習段階で穴埋めをした断片の回答となる単語とまったく学習しなかった未学習単語があった。また、テスト段階で提示される単語のファミリーリティの度合いを調べるにあたり、単語を完璧には記憶していないという状況を作るため、穴埋めに正解した単語のファミリーリティ評価を分析から除外した。実験の結果、テスト段階でのファミリーリティの評価は、外側断片を学習済みの単語は内側断片を学習済みの単語よりも高い値であった。また、外側と内側断片の学習済みの単語はそれぞれ未学習の単語よりもファミリーリティの評価が高かった。このことから、参加者は単語をはっきりとおぼえていなくても、外側の文字についてよりファミリーリティを感じることを示唆された。

しかしながら、これらの実験は英語を母語とする者を対象とした英語での実験であり、日本語でも同様に単語認識における文字位置についてファミリーリティの違いがみられるのかは明らかでない。英語では分かち書きを行うので、文中ではふつう単語と単語の間にスペースが存在する。Johnson & Eisler (2012) は、外側の文字の特権性には、単語間のスペースが関係するとしている。スペースが存在することで、文字間の視覚的な混雑状態(crowding)の緩和や横方向の干渉の減少が期待できるためである。一方、日本語の場合は分かち書きをしないので、文中において単語と単語の間にスペースがない。このことから、日本語使用者はスペースによる外側の文字の特権性を示さない可能性がある。

そこで、本研究では2つの仮説が考えられる。1つ目の仮説は、スペースによる横方向の干渉が生じないことから日本語においては、外側の文字の特権性が英語に比べ弱いとするものである。これを仮説1とする。仮説1が正しければ、外側断片と内側断片の学習条件によるファミリーリティに差がみられないことが予想される。2つ目の仮説として、スペースの有無に関わらず日本語における外側の文字の特権性は英語と同様に存在するというものである。これを仮説2とする。仮説2が正しければ、ファミリーリティは外側断片で内側断片に比べ高くなることが予想される。いずれの仮説においても、外側断片と内側断片の刺激は学習段階で一度目にしているため、未学習条件よりもファミリーリティの評価値が高くなると考えられる。

本研究は、日本語の単語認識において外側の文字は内側の文字より重視されるのかについて検証する。実験では、Huebert & Cleary (2022) の実験2の刺激に用いられた7文字の英単語の代わりにひらがな7文字で表される四字熟語を用いる。学習段階ではひらがな四字熟語断片について穴埋めをし、テスト段階では学習していない単語も含めたひらがな四字熟語のファミリーリティを11段階で評価する。独立変数は、各学習条件（外側・内側・未学習）とし、従属変数を学習段階の各断片（外側・内側）の正答率とテスト段階での各学習条件のファミリーリティの評価とする。

方 法

実験参加者

日本語を母語とする大学生50人が実験に参加した。参加者には性別と年齢は尋ねなかった。

刺激と装置

lab.js (Henninger et al, 2020) で作成したプログラムをサーバー上にアップロードし、参加者にアクセス可能なURLを配布した。参加者は、各自のPCからwebブラウザでURLにアクセスしオンライン上で課題を行った。

Huebert & Cleary (2022) を参考に刺激を作成した。刺激に用いる単語は、高橋 (2005) の883個の四字熟語の中から捨て仮名を含まない、ひらがな7文字表記で1文字目にあたる漢字の読みがひらがな2文字で表されるものを熟知度の高い順から120個用いた。そのひらがな7文字の四字熟語を用いて4つの学習-テストブロックを作成した。学習ブロックでは、それぞれ20個の一部文字を欠落させたひらがな四字熟語断片を用い、テストブロックでは30個のひらがな7文字の四字熟語を用いた。学習ブロックの20個の断片は、外側条件の外側断片10個、内側条件の内側断片10個からなり、断片を区別するため必ず最初から2文字目と最後から2文字目を含む単語断片を作成した。そのため外側断片は、最初と最後からの2文字目を含むひらがな四字熟語の断片であった（例：「いんがおうほう」については、「いん__ __ほう」）。内側断片は、最初と最後からの2文字目を含み、隣接による影響を等しくするため、2文字ずつ隣り合う断片を作成した（例：「りんきおうへん」につ

いては、「__んき__うへ__」)。欠落している文字の部分成全角のアンダースコアに置き換え、連続するアンダースコアの間に半角スペースを挿入し、完成したときの文字数がわかるようにした。テストブロックに含まれる 30 個のうち 10 個は学習ブロックの外側断片に対応するひらがな四字熟語であり（例：ゆいがどくそん）、10 個は学習ブロックの内側断片に対応するひらがな四字熟語（例：りんきおうへん）であり、10 個は学習ブロックに対応する刺激がないひらがな四字熟語であった。各四字熟語が 3 つの条件（外側、内側、未学習）について同じ頻度で出現するような 3 種類のバージョンを作成し、参加者間でカウンターバランスをとった。各ブロック内の四字熟語は参加者ごとにランダムな順に提示した。

図 1 は「いんがおうほう」の学習ブロックでの画面の例である。学習ブロック時に画面中央付近に提示されるひらがな四字熟語の断片の文字サイズは、HTML のフォントサイズで 7 であった。フォントの種類は参加者の PC によった。参加者がキーボードを用いて回答を入力する際の文字サイズは 40px であり、この文字のフォントの種類は参加者の PC によった。



図 1 学習ブロックの画面の例

図 2 はテストブロックでの画面の例である。テストブロックでのファミリーアリティ評価は、中央付近に HTML フォントサイズ 7 で提示される四字熟語の下にスライダーを設け、マウスでカーソルを合わせクリックまたはドラッグすることによって行われた。また、テストブロックでの四字熟語の上には、「先ほど穴埋め問題をした覚えがありますか」という文を提示した。



図2 テストブロックの画面の例

手続き

参加者は、配布された URL を Web ブラウザから開くことで実験プログラムにアクセスした。参加者には、画面中央付近に四字熟語をひらがなにした単語の一部が穴あきになったものが提示された後、穴を埋めることでできあがる四字熟語をキーボードで入力する課題を実施することを教示した。その後テストブロックでひらがなの四字熟語全体が提示され、先ほどの穴埋め問題をした覚えがある四字熟語であるかを 10 段階で評価するように求めた。参加者は、学習ブロックではひらがなの四字熟語を見て、穴を埋めて完成した状態の四字熟語全体をキーボード入力した。キーボードは漢字に変換できないようにあらかじめ設定してあった。参加者は、答えを入力できたか、入力を諦めたときに **Enter** キーを押して次の試行に進んだ。最初の学習ブロック 20 試行を回答後、参加者は最初のテストブロックに取り掛かった。テストブロックでは、ひらがなで四字熟語全体が提示され、先ほど穴埋め課題をした覚えがある四字熟語であるかをマウスでスライダーをドラッグすることで 11 段階での評価をした。0 は「覚えがない」「見たことがない」10 は「覚えがある」「見たことがある」を意味する。ドラッグ後、**Enter** キーまたは右下にある「次へ」のボタンをマウスでクリックし、次の四字熟語の評価に進んだ。テストブロックは 30 試行から構成された。四字熟語はランダムな順に提示された。テストブロックの 30 試行を全て終了すると、次の学習ブロックが始まった。この手順でさらに 3 回、合計して 4 回の学習-テストブロックを繰り返した。

結 果

ファミリアリティ

学習段階での穴埋め課題に正答した単語のデータを除外した後、各条件のファミリアリティを集計し、平均したものを図3に示した。

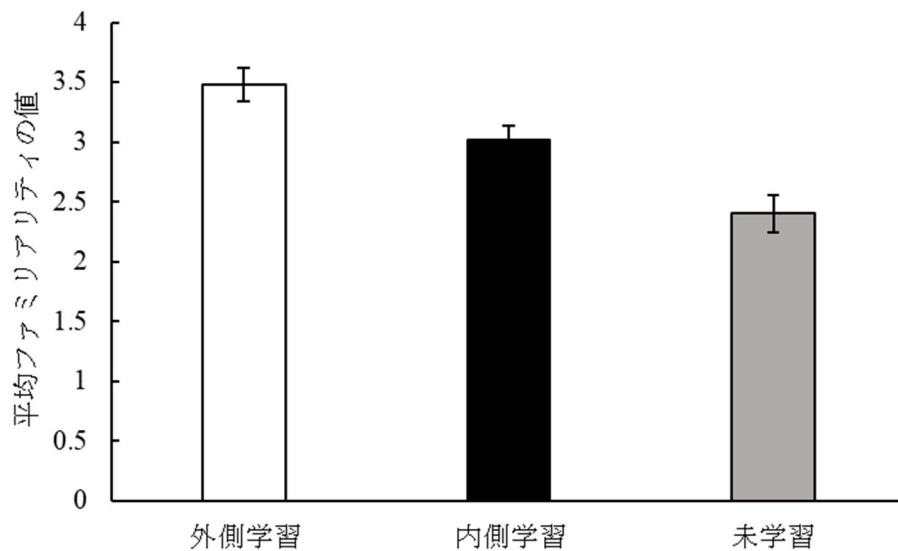


図3 各条件の平均ファミリーアリティ (エラーバーは95%信頼区間)

断片の種類を要因とした 1 要因 3 水準の参加者内分散分析を有意水準 5%で行った。その結果、主効果が有意であった ($F(2,98) = 30.35, p < 0.005$)。シェーファー法による多重比較を有意水準 5%で行ったところ、未学習断片よりも内側断片と外側断片のほうが平均ファミリーアリティが有意に高かった (内側 vs 未学習: $t(49) = 4.38$, 外側 vs 未学習: $t(49) = 6.99$)。そして、内側断片よりも外側断片のほうが平均ファミリーアリティが有意に高かった ($t(49) = 3.87$)。

断片完成率

各条件の学習時の平均正答率を図4に示した。対応のある t 検定を有意水準 5%で行った。その結果、外側断片のほうが内側断片よりも有意に平均正答率が高かった ($t(49) = 8.45$)。

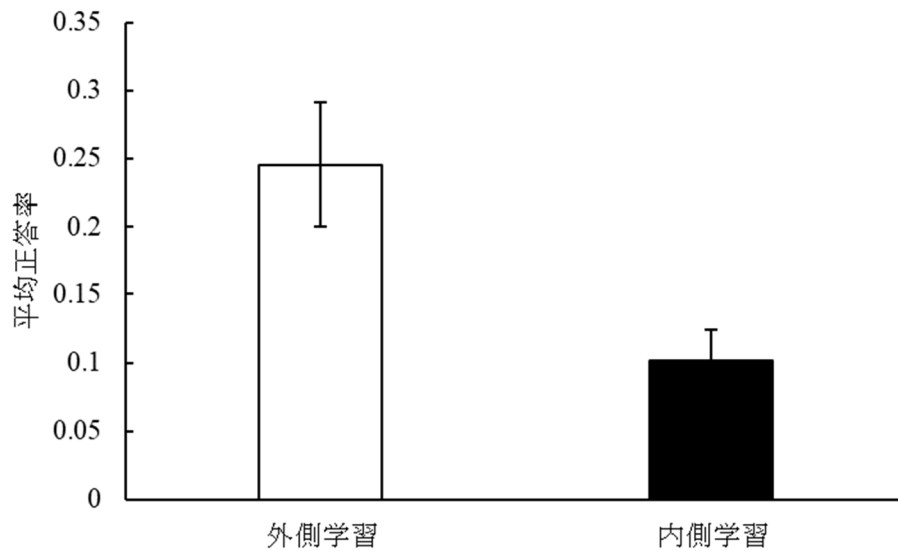


図4 各条件の平均正答率（エラーバーは 95%信頼区間）

考 察

本研究では、日本語における外側の文字のファミリーリティにおける特権性を検証することを目的とした。このため、単語を穴あきの状態にし、残った断片が外側にある外側断片条件と、内側にある内側断片条件、穴埋めテストをしなかった未学習断片条件の間でファミリーリティの評価を比較した。実験の結果、外側断片条件のほうが内側断片条件よりもファミリーリティが高いことが示された。また、外側・内側断片条件は未学習断片条件よりもファミリーリティが高いことも示された。この結果はHuebert & Cleary (2022) を支持するものである。

本研究では、2つの仮説を検証した。仮説1は、日本語はスペースによって単語間が区切られていないため、外側の文字の特権性が英語よりも薄いという仮説である。このことから、外側断片条件と内側断片条件のファミリーリティの評価に差がないことが予想された。本研究の結果はこの予測と一致しなかった。仮説2は、日本語であっても外側の文字の特権性が発揮されるという仮説だ。このことから、外側断片条件のほうが内側断片条件よりもファミリーリティの評定値が高くなることが予測された。本研究の結果はこの予測と一致した。また、どちらの仮説においても、外側・内側断片条件ともに未学習断片条件より評価が高くなることが予想された。この予想は支持された。これらのことから、日本語であっても単語から感じられるファミリーリティは文字の位置に影響されることが示唆された。Johnson & Eisler (2012) は、単語間にハッシュマーク (#) を挿入することでスペースを無くし、その状態で文字の順番が入れ替えられた文章を素早く読む実験を行った。その結果、最初の1・2番目の文字が入れ替わっている場合のほうが単語の内側の文字が入れ替わっている場合よりも読む時間が長くなることが示された。つまり、最初の文字が読みにくいと読解に手

間取るということだ。したがって、英語を母語とする者は、単語間のスペースがない状態でも最初の文字を優先的に処理すると考えられる。Johnson & Eislerは、最初の文字が語彙情報の処理方法やアクセス方法に関わっているから単語の処理に重要であるとしている。このことから、日本語を母語とする者も普段から分ち書きをされていない文章を用いていても外側の文字を優先的に処理すると考えられる。

しかし、本研究の断片完成の正答率には外側断片と内側断片の間に有意な差がみられた。したがって、集計したデータの数に条件間で差があり、このことが結果に影響した可能性がある。Huebert & Cleary (2022) の実験でも正答率に差がみられたため、この点を解消するには実験の方法を見直す必要があると考えられる。

実験の結果から、日本語の単語のファミリーティについても外側の文字に特権性がある可能性が示された。本研究の限界として、漢字をまたがる断片の効果が挙げられる。本研究では、最初の漢字の読みが2文字の四字熟語を選んだが、1文字のものも存在する。例えば、外側断片にした際、2文字のものは最初の漢字が残るが、1文字のものは2つ目の漢字まで断片が被る（例：「悪戦苦闘」であれば、「あく__ __とう」, 「他力本願」であれば「たり__ __がん」）。しかし、最初の読みが1文字の四字熟語の数が少なく、分析に適していないと判断した。もし1文字のものを使用した場合、3文字目が2文字目と関連付けられていることから、穴埋め課題の正答率やファミリーティの値が上昇すると考えられる。

引用文献

- Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2022). lab.js: A free, open, online study builder. *Behavior Research Methods*, 54(2), 556–573. <https://doi.org/10.3758/s13428-019-01283-5>
- Huebert, A. M., & Cleary, A. M. (2022). Do first and last letters carry more weight in the mechanism behind word familiarity? *Psychonomic Bulletin & Review*, 29(5), 1938–1945. <https://doi.org/10.3758/s13423-022-02093-1>
- Johnson, R. L., & Eisler, M. E. (2012). The importance of the first and last letter in words during sentence reading. *Acta Psychologica*, 141(3), 336–351. <https://doi.org/10.1016/j.actpsy.2012.09.013>
- 中尾清月堂 (2018) . いよいよ始まりました！どら焼きキャンペーン！ Retrieved February 13, 2023 from <https://nakaoseigetsudou.jp/nakao/blog/%e3%81%84%e3%82%88%e3%81%84%e3%82%88%e5%a7%8b%e3%81%be%e3%82%8a%e3%81%be%e3%81%97%e3%81%9f%ef%bc%81%e3%81%a9%e3%82%89%e7%84%bc%e3%81%8d%e3%82%ad%e3%83%a3%e3%83%b3%e3%83%9a%e3%83%bc%e3%83%b3%ef%bc%81/>
- 高橋 雅延 (2005). 日本語四字熟語 833語の熟知価 聖心女子大学論叢, 104, 280-244.