

THE IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS (JAPANESE EDITION)

IEICE | **電子情報通信学会**
D | **論文誌** 情報・システム

VOL. J99-D NO. 1

JANUARY 2016

本PDFの扱いは、電子情報通信学会著作権規定に従うこと。

なお、本PDFは研究教育目的（非営利）に限り、著者が第三者に直接配布することができる。著者以外からの配布は禁じられている。

情報・システムソサイエティ

一般社団法人 **電子情報通信学会**

THE INFORMATION AND SYSTEMS SOCIETY

THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS

研究速報

漢字二字熟語の語彙判断における概形の影響

矢内 浩文^{†a)} (正員) 林 健太^{†*}

Influence of Outline Shapes of Japanese Kanji Characters on Reactions to Lexical Decision Task of Two-Kanji Compound Words

Hiro-Fumi YANAI^{†a)}, Memberand Kenta HAYASHI^{†*}, Nonmember[†] 茨城大学工学部メディア通信工学科, 日立市

Department of Media and Telecommunications, Ibaraki University, Hitachi-shi, 316-8511 Japan

* 現在, (株) エア-

a) E-mail: hfy@ieee.org

DOI:10.14923/transinfj.2015HAL0003

あらまし 二字熟語を構成する漢字を転置してできる非熟語を見て, 瞬時に対応する熟語として認識できる, あるいは誤って認識してしまう現象に, 漢字の概形が影響することを示す。概形が同じ漢字でできた非熟語の方が, 異なる場合よりも熟語であると誤認識する率が高い。

キーワード 熟語, 単語, 転置, 認識, 空間周波数

1. ま え が き

文章に誤字や脱字があっても, 何とごもなく読めるものである。場合によっては, それに気づかないこともある。これは人間の語彙処理の程過において, 文字よりも大きな単位の情報が用いられている拠証であるといえる (このパラグラフには意図的に誤字が挿入されている。誤字はすべて隣接文字の転置で, 次の5箇所である: 「あてつも」「何とごもなく」「気づかない」「程過」「拠証」.)

Cambridge University の研究結果であるとして 2003 年頃にインターネット上で話題になった文章は, 単語内で文字を並べ替えた非単語から構成されている。その一部を抜粋すると “...aoccdnrig to a rscheearch at Cmabrigde Uinervtisy, it deons't mtttaer in waht oredr the ltteers in a wrod are...” であった。これは, 長年多くの研究者が取り組んできた単語認識メカニズムの研究の問題意識を例示する好例であったため, 言語研究の専門家も注目した [1]。その英文を読むに際して多少は負担が大きくなるものの, 比較的円滑に認識できるのは, 単語の最初と最後の文字が固定されていること [2], そして最初と最後以外の文字は並べ替えであって, そこにない別の文字との入れ替えではないこと [3] であることが指摘されている。

ここでは日本語の二字熟語における漢字の転置に限

定し, 語彙判断課題 (熟語か非熟語かをできるだけ速く判断する課題) を用いた研究の結果を報告する。冒頭の例でいえば, 「過程」↔「程過」, 「証拠」↔「拠証」である。ところで, 英語やスペイン語などでは, 単語 (日本語の熟語に相当するとみなせる) はより多くの文字で構成される (英語では平均単語長は 5 程度) から, 単語内の文字の並べ替え (転置) に限定しても, 組み合わせが多数あり, それゆえ, さまざまなレベルの非単語度が設定できる。それに対し, 日本語の漢字 2 文字の熟語では転置は 1 通りだけである。日本語の二字熟語を題材に用いることで考察の一般性を阻害するとの疑問が生じるかもしれない。しかしそもそも, 例えば英語の単語は表音文字で構成されているのに対し, 日本語の熟語は形態素文字から成るから, 音韻と形状 (図形的特徴) の影響の異なる刺激を用いることにより, 人間の言語情報処理の分析の観点を新たに設定できるという意味で, 制限ではなく一般化と捉えることもできるだろう。実際, 仮名の表音文字としての特性に注目した研究結果が外国の研究者によって報告されている [4]。

図形的特徴が語彙判断に及ぼす影響について述べておく。ミススペルの発見に影響する形状要因を調べた研究がある [5]。そこでは, 英語でのミススペル検出課題において, 上に飛び出している文字 (d や t) と, 下に飛び出している文字 (p や y), そして上にも下にも飛び出していない文字 (a や c) の 3 グループに分けることでミススペルの発見成績を評価している。それによれば, グループ内での別の文字との入れ替えよりも, グループ間で文字の入れ替えを行った方がミススペルの発見率が高いことが示されている。

ここで, 日本語の二字熟語を転置した非熟語を刺激に取り入れた語彙判断課題の研究の例を挙げておく。非熟語として 3 種類 (熟語を転置した非熟語, 同音の熟語がある非熟語, そのどちらでもない非熟語) を用意し, 熟語と交ぜて語彙判断させたところ, 転置の非熟語を含む場合に反応時間が長く, エラー率も高いことが示されている [6]。なお, そこでは概形の観点での分析はなされていない。

私たちは, 概形が三角形, 四角形, 五角形に分類できる漢字から成る二字熟語を刺激とした実験を行った。例えば, 「評価」は “四角形・四角形”, 「人権」は “三角形・四角形”, 「会談」は “五角形・四角形” という具合である (図 1 参照)。以下では, 熟語と, それを転置した非熟語の中からランダムに一つずつ刺激呈示する

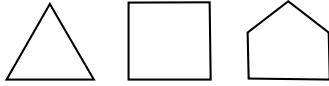


図 1 実験刺激に用いた漢字の概形。三角形、四角形、五角形

Fig. 1 Outline shapes of kanji characters used in the experiment. Triangle, rectangle, and pentagon.

ことにより、熟語か非熟語かをできるだけ速く判断させる課題を実施し、異なる概形の漢字から構成される熟語（または非熟語）と同じ概形の場合の反応時間、エラー率を比較した結果を報告する。エラー率が高いことは転置非熟語を熟語であると誤認識しやすいことを意味するのはもちろん、反応時間が長いということは脳内で誤認識の訂正が行われている可能性があるから、その場合もだまされやすい非熟語だということになる。

2. 実験

実験協力者は日本語を母国語とする大学生 23 名（男性 16 名、女性 7 名、平均年齢 21 歳）であり、実施前にインフォームド・コンセントをとった。実験参加者は 24 名であったが、全実験にわたるエラー率が他の実験参加者の平均エラー率の 4 倍を超えていた 1 名を分析から除外した。刺激はコンピュータ画面上に呈示した。画面の背景は黒色、熟語は横書き、白色で表示した。字体は MS ゴシックで、文字のサイズは、1 文字あたり視角で 1° となるようにし、文字間隔はゼロとした。実験協力者には、刺激が呈示されたらできるだけ速く正確に語彙判断を行い、刺激が熟語（正熟語）であるか非熟語であるかに応じてパソコン・キーボードの指定されたキーを押すよう指示した。

刺激は漢字 2 文字の熟語（80 個）及びそれら熟語の文字を転置した非熟語（80 個）を合わせて 160 個で構成した。80 個の熟語は、2 文字の漢字の概形が同じ熟語と異なる熟語が同数（40 個ずつ）であった。内訳は、概形が同じ熟語は“四角形・四角形”が 40 個（「問題」「評価」「計画」ほか）、概形が異なる熟語は、“三角形・四角形”（「土地」「人権」「小型」ほか），“五角形・四角形”（「企業」「全国」「会談」ほか），“四角形・三角形”（「以上」「導入」「個人」ほか），“四角形・五角形”（「国会」「試合」「記念」ほか）がそれぞれ 10 個であった。

刺激呈示順序はランダムで、実験協力者を 2 グループに分け、刺激呈示順序でカウンターバランスをとっ

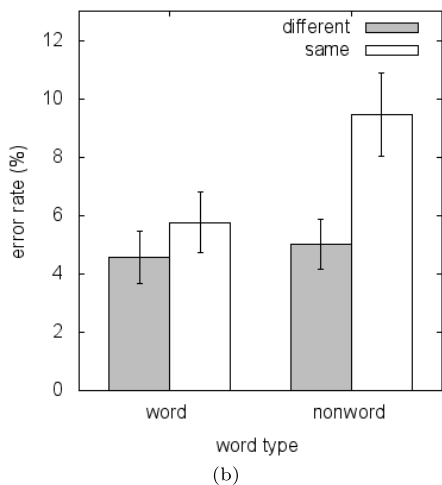
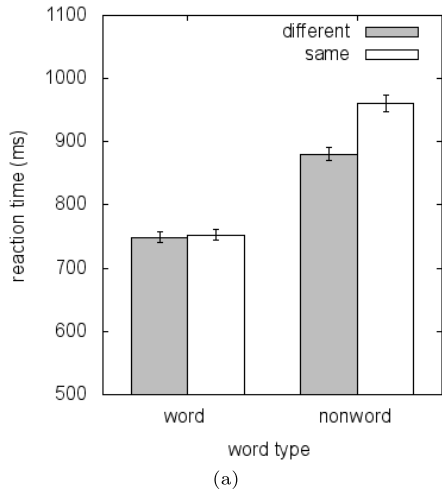


図 2 概形の異・同、熟語・非熟語に応じた反応時間（上段）及びエラー率（下段）。エラーバーは標準誤差
Fig. 2 Reaction time and error rate according as conditions for outline shapes and word/nonword.

た。実験は全 160 試行を実施し、10 試行ごとに 30 秒の休憩を取った。1 試行の流れは次のとおりであった。画面中央に注視点（十字）が表示され、3000ms 後に注視点が消える。注視点が消えてから 1000ms が経過すると刺激が呈示される。実験協力者が回答するか 5000ms が経過すると自動的に次の試行に移行する。

3. 結果

反応時間とエラー率を、概形の異・同、熟語性の正・非（正熟語・非熟語）に分けて整理したのが図 2 である。なお、反応時間については正答のみを用いた。平均で比較すると、反応時間については、概形が同じ

場合に非熟語の方が熟語よりも反応時間が長い。同じくエラーについては、熟語、非熟語の両方について、概形が同じ場合にエラー率が大きい。概形(異・同)×熟語性(正・非)の2要因分散分析(対応あり)の結果、反応時間については、概形の主効果は有意($F(1,22) = 25.012, p < .001$)、熟語性の主効果も有意である($F(1,22) = 78.430, p < .001$)。同じくエラー率については、概形の主効果は有意である($F(1,22) = 8.725, p < .01$)が、熟語性の主効果は有意ではない($F(1,22) = 3.174, p > .05$)。

4. む す び

熟語・非熟語の判断課題を用いて、二字熟語を構成する漢字の概形が同じ場合は、概形が異なる場合よりも判断に時間を要し、判断の誤りも多いことを示した。このことは、熟語の認識過程において、漢字を特定できる詳細な図形的特徴のみならず、漢字をおおまかに分類する全体形状も関与していることを示唆している。いわば、漢字を画像として見たときに、低空間周波数チャンネルの情報が熟語の認識過程にかかわっている可能性がある。

最後に今後の課題を挙げる。今回の実験では、概形の影響を検討する研究の入り口として、漢字の概形を人間の直感で分類した。フーリエ成分やラプラシアン・

ピラミッドなど客観的な手法による分類と直感的分類の関係を明らかにできれば、熟語の視覚的認識過程の理解が更に深まるであろう。

文 献

- [1] J. Grainger and C. Whitney, "Does the huamn mnid raed wrods as a wlohe," Trends in Cognitive Sciences, vol.8, no.2, pp.58-59, 2004.
- [2] K. Rayner, S.J. White, R.L. Johnson, and S.P. Liversedge, "Raeding wrods with jubmled letters," Psychological Science, vol.17, no.3, pp.192-193, 2006.
- [3] S. Schoonbaert and J. Grainger, "Letter position coding in printed word perception: Effects of repeated and transposed letters," Language and Cognitive Processes, vol.19, pp.333-367, 2004.
- [4] M. Perea and E. Perez, "Beyond alphabetic orthographies: The role of form and phonology in transposition effects in Katakana," Language and Cognitive Processes, vol.24, pp.67-88, 2009.
- [5] A.F. Monk and C. Hulme, "Errors in proofreading: Evidence for the use of word shape in word recognition," Memory and Cognition, vol.11, no.1, pp.16-23, 1983.
- [6] R. Mizuno and T. Matsui, "Orthographic or phonological?: Exploration of predominant information for native Japanese readers in the lexical access of kanji words," Psychologia, vol.56, pp.208-221, 2013.

(平成 27 年 3 月 27 日受付, 10 月 2 日早期公開)