

漢字語とカタカナ語における出現頻度効果と親密度効果の差異

楠瀬 悠*

日本語や英語を用いた研究において、視覚提示された語の読みの成績に、その単語の出現頻度や親密度が影響を及ぼすことが良く知られており、さらに、出現頻度と親密度の間には高い相関関係が認められている (e.g., 天野・近藤, 1999, 2000a; Gernsbacher, 1984; Gordon, 1985; Hino & Lupker, 1998). しかし, Hino & Lupker (submitted) は、漢字語とカタカナ語それぞれの出現頻度と親密度の相関を計算したところ、漢字語の方がカタカナ語よりも相関が高かったことを報告している。このことから、出現頻度と親密度が読みの成績に与える影響は、表記によって異なる可能性がある。

そこで本研究では、漢字語とカタカナ語の出現頻度と親密度を直交に操作した語彙判断課題を実施し、漢字語とカタカナ語における両要因の影響を再検討した。実験の結果、出現頻度と親密度との間に高い相関が認められた漢字語では、出現頻度効果と親密度効果がともに観察され、漢字語と比較して低い相関であったカタカナ語では親密度効果のみが観察された。この結果は、漢字語とカタカナ語の音韻一形態対応の一貫性の程度の違いによって出現頻度と親密度への影響が異なる可能性を示唆していた。これらの結果から、日本語の読みにおける出現頻度効果と親密度効果について再考した。

Key words : 出現頻度効果, 親密度効果, 音韻一形態間の一貫性

はじめに

私たちが日常的にことばを読むためには、与えられた視覚情報を基に、私たちが保持している語彙知識を検索し、それに対応することばの情報を正しく復元する必要がある。このような単語認知プロセスに関する古典的な研究では、私たちが保持している語彙知識のことを心的辞書 (mental lexicon) と呼び、その辞書内にはそれぞれの語に対応する語彙表象 (lexical representation) が仮定されている。そして、この語彙表象にアクセスすることで提示された単語の意味や発音、綴りなどの情報を復元できると考えられている。現在まで

に、心的辞書内における語彙表象へのアクセスがどのように行われているか、ということについて様々な単語認知モデルが提案されてきた (e.g., Becker, 1970; Forster, 1976; McClelland & Rumelhart, 1981; Morton, 1969; Paap, Newsome, McDonald & Schvaneveldt, 1982).

この単語認知モデルを提案する上で、単語が持つ様々な要因の影響があると考えられるが、最も重要視されている要因に出現頻度を挙げることができる。出現頻度とは、日常的にその単語をどの程度目にしたか、という指標のことである。一般的に、提示された文字列が単語かどうかの判断に要する時間を測定する語彙判断課題において、よ

* 人間学部心理学科

く目にする高頻度語の方が、あまり目にしない低頻度語よりも速く判断でき、かつ誤反応も少ないことがよく知られており、この効果は出現頻度効果と呼ばれている (e.g., Balota & Chumbley, 1984; Becker, 1979; Hino & Lupker, 1998, 2000; 広瀬, 1984, 2007; Monsell, Doyle & Haggard, 1989; Paap, McDonald, Schvaneveldt & Noel, 1987; Rubenstein, Garfield & Millikan, 1970).

こうした出現頻度の効果を見るためには、各単語の頻度を客観的な指標として計算した語彙コーパスが必要である。英語における研究では、古くから Thorndike & Lorge (1944) や Kucera & Francis (1967), Carroll, Davis & Richman (1971) などの語彙コーパスが存在しており、出現頻度など単語に関する要因の検討が行われてきた。また、日本でも国立国語研究所 (1970, 1993) や天野・近藤 (1999, 2000a, 2003a, 2003b) のデータベースなどがあり、日本語においても同様に研究されてきた。例えば、天野・近藤のデータベースは、朝日新聞の記事データ 14 年間分を形態素解析し、約 3 億語中におおよそ 34 万個の単語がどの程度含まれていたかの度数を算出しており、現在までに日本語のコーパスの中では最も多く使用されている。

実際、Rubenstein et al. (1970) は、Thorndike & Lorge (1944) の語彙コーパスを使用し、語彙判断課題における出現頻度効果の検討を行っている。実験の結果、語彙コーパスの頻度が 10 回異なることで 75ms の出現頻度効果を報告している。また日本語においても、Hino & Lupker (1998) は、国立国語研究所 (1970) を使用し、漢字語とカタカナ語における出現頻度効果の違いについて語彙判断課題、音読課題、Go/No-Go 語彙判断課題を実施して検討した。その結果、両表記ともに、いずれの課題においても大きな出現頻度効果を観察している。このように、英語においても、日本語においても、語彙コーパスに基づいた出現頻度効果は様々な課題で観察される非常に頑健な効果であり、したがって従来の単語認知モデルではこの要因を重視して作成されてきた。

それでは、この出現頻度効果は単語認知モデルにどのように反映されているのだろうか。古典的

な視覚的単語認知モデルは、大きく分けて直列検索モデルと活性化モデルの 2 つに分類される。直列検索モデルは、心的辞書における語彙表象選択の際、視覚的な情報と語彙表象の間に直列的な関係を仮定するモデルである (e.g., Becker, 1970; Forster, 1976; Paap et al., 1982)。このモデルでは視覚刺激が入力されると、その視覚刺激に類似する複数の語彙表象が活性化される。そして、その活性化された複数の語彙表象から候補セットが形成され、その候補セットと視覚刺激の間で照合がなされる。この候補セット内における照合の順番は、出現頻度が高い単語から低い単語へと順に照合されると仮定しており、この照合の順序によって出現頻度効果を説明している。他方、活性化モデルでは、語彙表象の選択は活性化のプロセスのみで行われると仮定している (e.g., McClelland & Rumelhart, 1981; Morton, 1969)。このモデルにおける語彙表象は、閾値を持った語彙ユニットとして表現されており、その閾値は出現頻度に応じて決定される。そして、直列検索モデルと同様に、視覚刺激の入力から刺激の類似性によって複数の語彙表象 (ユニット) が活性化され、その中で最初に閾値を越えて活性化した語彙ユニットが単語として認識される。つまり、活性化モデルでは、各語彙ユニットが持つ閾値を出現頻度に対応させ、高頻度語は閾値を低く、低頻度語ほど閾値を高く設定することで出現頻度効果を説明している。

このように、古典的な単語認知モデルではいずれにおいても、出現頻度効果を説明するシステムをモデルの核に据えており、私たちがことばを読むシステムを理解する上で、出現頻度は非常に重要な要因となっている。この背景には、出現頻度がその単語をどの程度学習したかという経験の程度を表していることが考えられる。実際、近年の認知モデルの主流となっているコネクショニズムは、その学習の程度によってユニットの結合強度を決定している。この理論は、単語認知においては並列分散処理 (Parallel Distributed Processing; PDP) モデルと呼ばれ、脳神経の構造を基にした形態・音韻・意味情報における多数のユニット間の結合関係で語彙情報を表現しているモデルである (e.g., Plaut, 1997; Plaut, McClelland, Seidenberg

& Patterson, 1996; Seidenberg & McClelland, 1989). このモデルの仮定として、ユニット間の結合強度は学習の程度、つまり少なくとも形態情報から他の情報への結合は視覚的な情報である出現頻度に依存する。つまり、多く学習したものがほど強く結合され、その結果、反応が速くなることで出現頻度効果を説明している。以上のことから、私たちがことばを読むというプロセスを理解する上で提案されてきた過去から現在までの様々な単語認知モデルにおいて、出現頻度は中心的な役割を果たしているのである。

出現頻度と文字単語親密度の関係

ことばの読みにおける重要な要因として出現頻度について概観してきたが、出現頻度に大きく関連する要因として親密度が挙げられる。親密度とは、提示された単語に対してどの程度親しみがあるかの程度のことであり、読者の主観的な指標とされている(天野・近藤, 1999, 2003a; Gernsbacher, 1984)。そして、この親密度も出現頻度と同様に、親しみやすい高親密度語の方が、あまり親しみのない低親密度語よりも速く判断することができ、誤反応も少ないことが知られており、この効果は親密度効果と呼ばれている(e.g., 天野・近藤, 1999; Connine, Mullennix, Chernoff & Yelen, 1990; Gernsbacher, 1984; Gordon, 1985)。実際、天野・近藤は、7段階で評定した親密度を4.0から0.5ずつのグループに分けた刺激を用いた語彙判断課題、音読課題を実施したところ、いずれの課題においても親密度効果を観察している。この結果から、親密度はその単語をどの程度目や耳にしているか、という日常的な学習の経験に基づくと考えられることから、出現頻度と同じく親密度も単語認知における重要な要因であると述べている。

このように、出現頻度と親密度はそれぞれ学習の程度を示す指標と考えられており、ある程度の共通性を持つ。このことから、出現頻度と親密度の間には相関関係が認められることが知られている(e.g., 天野・近藤, 2000a; Gordon, 1985)。Gordonは、英語の複数のコーパス(e.g., Carroll et al., 1971; Kucera & Francis, 1967)における出現

頻度と親密度の間の相関関係を調べたところ、 $r=.74$ から $.80$ という高い相関係数を報告している。また、天野・近藤(2000a)は、彼らのデータベースの中で両者の相関を確認したところ、 $r=.634$ であったことを報告している。この結果は英語と比較すると相関の程度は多少劣るが、日本語においても出現頻度と親密度の間には相関関係があることを示している。

このような関係性を持つ出現頻度と親密度は、ともに単語の読みの成績に大きく影響するため、言語を用いた心理学実験においては統制すべき重要な要因となっている(天野・近藤, 1999)が、それではどちらの影響がより重要なのだろうか。Gernsbacher(1984)は、様々な研究で報告されてきた出現頻度と他の要因(形態—音韻間の一貫性、意味的具体性、意味数)の間に観察されてきた非一貫的な出現頻度効果が、親密度を統制することによって解消することを報告している。このことから、語の読みでは出現頻度も重要な要素ではあるが、親密度の方がより重要な要因であると主張している。また、天野・近藤(2000a)は、出現頻度と親密度を系統的に変化させ、両要因における単語の認知率について検討したところ、親密度の方がより認知率の結果を反映すると報告している。この結果から、読みの成績には客観的な指標である出現頻度よりも主観的な評価である親密度の方が成績をより反映すると述べている。さらに、言語心理学実験のように刺激の要因を統制する必要がある場合には、出現頻度と親密度の間の相関が高いことから、両要因を統制する必要は無く、親密度のみを使用すれば良いと主張している。

漢字語とカタカナ語における出現頻度と親密度の相関の違い

今まで述べてきたように、英語においても日本語においても出現頻度と親密度の間には相関関係が認められているが、表記によって相関関係の程度が異なる可能性も示唆されている。英語などアルファベット言語とは異なり、日本語には漢字語と仮名語という複数の表記が存在している。天野・近藤(2000a)の結果は、表記に関わらず両要因間の相関を計算しており、この違いについて

ては検討していない。Hino & Lupker (submitted) はこの点に着目し、漢字語 (46,620 語) とカタカナ語 (11,103 語) のそれぞれの出現頻度と親密度の間の相関を計算したところ、漢字語の方がカタカナ語よりも相関が高かったことを報告している (漢字語: $r=.71$ v.s. カタカナ語: $r=.44$)。

Hino & Lupker (submitted) は、漢字語とカタカナ語の相関関係の違いについて、両表記における音韻—形態間の一貫性の違いから説明を試みている。カタカナ語はモーラと文字の対応関係が規則的であることから、音韻—形態間の一貫性の値は高い。一方、漢字語は多数の同音語を持つ (e.g., 高低に対する校庭・肯定・皇帝など) ことから、1つの音韻情報に対して複数の異なる形態情報を持つため、音韻—形態間の一貫性の値は低くなる。この一貫性の違いが親密度に与える影響が異なる可能性を指摘している。つまり、カタカナ語の親密度は視覚情報からの影響に加え、音韻情報からの影響が大きくなる。一方、漢字語の親密度は音韻情報の影響は小さく、視覚情報からの影響に頼ることになる。このことから、漢字語とカタカナ語における出現頻度と親密度の相関関係の程度の違いは、音韻情報の影響の違いによって生じている可能性を示唆している。

このことが正しければ、単語を読むプロセスにおいて、漢字語と仮名語の間では出現頻度と親密度の影響が異なる可能性がある。もしそうであるなら、天野・近藤 (2000) が提案する、刺激の統制において親密度のみを使用するという考えについても再検討する必要がある。また、それぞれの要因が読みのモデル構築に与える影響も異なる可能性が考えられ、どのような学習経験を基にしているのかという点についても考慮する必要があるだろう。以上のことから、本研究では、漢字語と仮名語における出現頻度効果と親密度効果の再検討を行う。

古典的な単語認知モデルでは、親密度を出現頻度と同様に視覚的な経験に基づくものとして仮定しており、音韻情報による影響の違いを説明することは難しいと考えられる。そのため、このモデルの場合、漢字語とカタカナ語の出現頻度と親密度の効果は同じ程度で観察されると予測される。

一方、PDP モデルはその影響を形態—音韻間の学習強度の違いとして表現が可能である。そのため、このモデルではその学習強度の違いから、漢字語とカタカナ語における出現頻度と親密度の影響が異なると思われる。

本研究の目的

現在までの出現頻度と親密度に関する多くの研究では、両要因間の相関の高さから同時に操作、統制されることが多く、それぞれの要因が与える影響を独立に検討した研究は未だない。しかし、天野・近藤 (2003a) を検索すると、一部ではあるが出現頻度が高く親密度が低い語や、親密度が高く出現頻度が低い語などが存在している。

実際、天野・近藤 (2000b) においても、「おにぎり」や「ごちそうさま」など親密度が 6.0 以上と非常に高く、出現頻度が 10 回未満という非常に低い語があることも報告されている。しかし、これらの語は全て仮名語であり、私たちが日常的な読みを使用する語は専ら漢字語であることが多い。漢字語とカタカナ語の間の出現頻度と親密度の相関の違いに加え、この点からも各表記における出現頻度と親密度の影響を検討する必要があるものと思われる。

そこで本研究では、漢字語とカタカナ語の出現頻度と親密度の高低について、可能な限り独立に操作した語彙判断課題を実施し、漢字語とカタカナ語における出現頻度効果と親密度効果の再検討を試みた。この検討を通して、視覚的単語認知の研究で示されてきた両要因の影響を再度明らかにすることが本研究の目的である。この目的を達成することは、今後の単語認知研究や失語症における単語の音読検査や記憶の成績などの研究にとっても非常に有用なものであると考えている。

漢字語とカタカナ語における出現頻度と親密度の影響が表記によって異なるなら、それぞれの効果の見え方も異なってくると考えられる。カタカナ語は、出現頻度と親密度との間の相関関係が低く、その親密度は音韻—形態間の一貫性の高さに基づいた音韻情報によるものと仮定するならば、視覚的な情報である出現頻度効果よりも親密度効果の方が観察されやすくなるはずである。一方、

漢字語は出現頻度と親密度との間の相関関係が高いことから、両要因が同時並列的に影響を与えることが考えられるため、出現頻度効果も親密度効果も同程度で観察されると予測される。

方法

実験参加者

大学生及び大学院生 30 名がこの実験に参加した (平均年齢 20.70 歳; $SD=1.72$)。実験参加者は全て日本語を母国語とし、眼鏡等による矯正も含めて正常の範囲の視力を有した。いずれの実験参加者も漢字語およびカタカナ語の両実験に参加した。

刺激

漢字語およびカタカナ語について、天野・近藤 (2003a; 2003b) のデータベースから選択した。出現頻度条件については、天野・近藤 (2003b) において 5000 以上の語を高頻度語、1000 未満の語を低頻度語とした。また、親密度条件については、天野・近藤 (2003a) における文字単語親密度の値が 7 件法で 6 以上の語を高親密度、6 未満の語を低親密度とした。モーラ数を統制するため、漢字語は漢字二字熟語を、カタカナ語は 3 ~ 5 文字

語を使用した。

そして、漢字語とカタカナ語それぞれについて、出現頻度と親密度の高・低を操作した各 4 条件につき 16 語ずつの刺激を選択した (漢字語の出現頻度: $F(1,60)=618.02, MSE=1,115,268, p<.001$; 漢字語の親密度: $F(1,60)=2243.78, MSE=0.002, p<.001$; カタカナ語の出現頻度: $F(1,60)=74.48, MSE=10,331,748, p<.001$; カタカナ語の親密度: $F(1,60)=125.47, MSE=0.04, p<.001$)。また、漢字語とカタカナ語の間において、出現頻度と親密度の値を比べたところ、表記と両要因との間の交互作用は観察されず、同程度であることが確認された (全ての $F<0.2$)。

また、Table1 に示したように、漢字語とカタカナ語それぞれの各条件の間で、モーラ数・語長 (e.g., 天野・近藤, 2003b), 国立国語研究所 (1993) を使って計算した形態隣接語数・音韻隣接語数を統制し、漢字語においては文字出現頻度 (e.g., 天野・近藤, 2003b) も統制した (全ての $F<1.8$)。

このようにして作成された漢字語およびカタカナ語の単語刺激セットに加えて、それぞれの表記条件に 64 個の非単語を作成した。それぞれの非単語刺激は、カタカナ語は一文字を置き換えることによって、漢字語は漢字一文字を組み合わせてすることで実在しない漢字二文字の非単語にすること

Table 1

Mean number of word familiarity rating (Fam), word frequency (Freq), word length (Len), word morae (Morae), orthographic neighborhood size (ON), phonological neighborhood size (PN), kanji character frequency (KCF) for the High (H) or Low (L) frequency condition and High (H) or Low (L) familiarity condition words used in Kanji and Katakana Experiment.

Word Type	Condition	Freq	Fam	Len	Morae	ON	PN	KCF
Kanji	H-H	7240	6.36	2.00	3.50	54.06	56.00	707307
	H-L	7140	5.76	2.00	3.50	54.88	60.56	668698
	L-H	588	6.34	2.00	3.50	52.56	44.44	681854
	L-L	665	5.76	2.00	3.50	53.44	48.38	681601
Katakana	H-H	7444	6.35	3.69	3.56	2.69	6.75	-
	H-L	7395	5.77	3.63	3.63	3.56	8.81	-
	L-H	474	6.35	3.63	3.63	2.88	7.38	-
	L-L	495	5.77	3.63	3.63	3.38	6.19	-

Note—Word frequencies, familiarity ratings, morae and kanji character frequencies were taken from Amano and Kondo (2003a, 2003b) and, thus, the frequencies were counts per 287,792,797 words and familiarity ratings were based on a 7 point scale ranging from 1 (Unfamiliar) to 7 (Familiar). Orthographic Neighborhood sizes and phonological neighborhood sizes were calculated using National Language Research Institute (1993), which involves 36,780 word entries.

で作成された。これらの非単語刺激は、モーラ数・語長を両表記語の単語条件と等しくした。漢字語とカタカナ語の各刺激セットは、64個の単語刺激、64個の非単語刺激の計128ペアでそれぞれ構成された。

手続き

実験参加者は個別に実験に参加した。実験参加者には、CRTモニター (Iiyama, HM204DA) 中央に提示される漢字もしくはカタカナ文字列が“語”であるか“非語”であるかを判断し、PC (IBM-AT 互換機) に接続された反応ボックス上の“語”キーあるいは“非語”キーをできるだけ迅速かつ正確に押すよう求める語彙判断課題を実施させた。実験参加者は、常に、利き手を使って“語”反応をするよう二つのキーを割り当てた。

各実験はそれぞれ128試行からなり、刺激提示順序は実験参加者毎にランダムであった。両実験に先立って、練習を16試行実施した。練習試行では、実験試行で使用していない刺激を提示した。

各試行は400Hzのピープ音を50ms間提示することで開始された。ピープ音に続いてCRTモニター中央に凝視点が1000ms提示され、即座にターゲットに置き換えられた。凝視刺激、ターゲットは全て黒色背景に白色で提示した。また、これらの刺激はCRTモニターの垂直同期信号に同期させて提示した。

実験参加者には、CRTモニターの前方向約50cmのところ座って実験を行うよう教示した。ターゲット提示からキー押し反応までの反応時間と反応の正誤が、PCにより自動的に記録された。試

行間間隔時間は2秒であった。

結果

漢字語およびカタカナ語の語彙判断に要した“語”試行の反応時間及び誤反応率を分析した。反応時間のデータのうち、各実験参加者の平均から2.5×標準偏差の範囲外にあるものを外れ値とみなしてデータ分析から除外した。その結果、漢字語の実験では45個(2.34%)のデータが、カタカナ語の実験では54個(2.81%)のデータが外れ値として除外された。さらに、誤反応は反応時間の分析から除外した。その結果、漢字語の実験では78個(4.06%)のデータが、カタカナ語の実験では147個(7.66%)の誤反応が反応時間の分析から除外された。

“語”試行における反応時間と誤反応率の実験参加者平均と項目平均のデータに対して、それぞれ実験参加者を無作為要因とする分析(F_1 :以後、実験参加者分析と略記)とターゲット刺激を無作為要因とする分析(F_2 :以後、項目分析と略記)を行った。実験参加者分析における反応時間と誤反応率の平均値をTable2に示す。そして、出現頻度(高・低)および親密度(高・低)をそれぞれ実験参加者内要因・項目間要因とした2要因の分散分析を実施した。

実験参加者分析における漢字語の両要因の高・低条件の平均反応時間をFigure1に、カタカナ語の両要因の高・低条件の平均反応時間をFigure2に示す。いずれも縦軸は反応時間(ms)、横軸は親密度と出現頻度の各条件を、エラーバーは標準

Table 2

Mean Lexical latencies (RT) in millisecond and error rates (ER) in percent for each condition in Kanji and Katakana Experiment.

Word type	H-H		H-L		L-H		L-L	
	RT	ER	RT	ER	RT	ER	RT	ER
Kanji	441.34 (12.10)	1.04 (0.43)	494.11 (13.27)	6.04 (0.48)	460.89 (12.25)	1.46 (1.20)	503.98 (12.98)	7.71 (1.49)
Katakana	509.44 (17.77)	3.13 (0.77)	552.00 (16.59)	12.71 (1.49)	518.04 (14.22)	2.92 (0.97)	545.02 (17.17)	11.88 (1.51)

Note—Standard errors of the means are in parenthesis(). H-H indicate high frequency-high familiarity condition, H-L indicate high frequency-low familiarity condition, L-H indicate low frequency-high familiarity condition and L-L indicate low frequency-low familiarity condition.

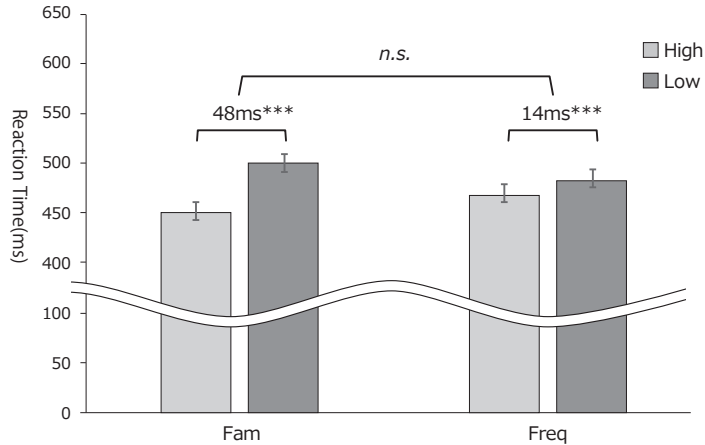


Figure 1. Mean lexical latencies in millisecond for familiarity(Fam) and frequency(Freq) for each(high and low) conditions in Kanji words experiment. Asterisks indicate significant differences(** $p < .001$, *n.s.* means non significant). Error bars reflect within-subjects *SEM*.

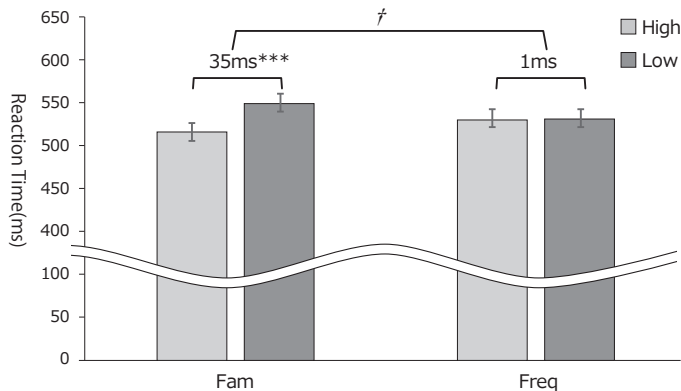


Figure 2. Mean lexical latencies in millisecond for familiarity(Fam) and frequency(Freq) for each(high and low) conditions in Kanji words experiment. Asterisks and dagger indicates significant differences (** $p < .001$, † $p < .10$). Error bars reflect within-subjects *SEM*.

誤差を示している。

漢字語の反応時間の分析では、まず出現頻度の主効果が実験参加者分析と項目分析の両方で有意であった ($F_1(1,29)=15.32$, $MSE=423.72$, $p<.001$, $\eta_p^2=0.35$; $F_2(1,60)=11.46$, $MSE=578.99$, $p<.01$, $\eta_p^2=0.16$)。また、親密度の主効果も両分析で有意であった ($F_1(1,29)=193.80$, $MSE=355.66$, $p<.001$, $\eta_p^2=0.87$; $F_2(1,60)=76.39$, $MSE=578.99$, $p<.001$, $\eta_p^2=0.56$)。出現頻度と親密度の交互作用は、両分析で有意でなかった ($F_1(1,29)=2.21$, $MSE=318.14$; $F_2(1,60)=0.97$, $MSE=578.99$)。これらの結果は、漢字語においては出現頻度と親密度ともに高条件

の方が低条件よりも反応時間が速かったことを反映している。

次に、カタカナ語の反応時間の分析では、出現頻度の主効果が実験参加者分析と項目分析のいずれでも有意でなかった ($F_1(1,29)=0.02$, $MSE=1201.56$; $F_2(1,60)=0.10$, $MSE=1887.88$)。一方、親密度の主効果は両分析で有意であった ($F_1(1,29)=55.31$, $MSE=655.67$, $p<.001$, $\eta_p^2=0.66$; $F_2(1,60)=14.30$, $MSE=1887.88$, $p<.001$, $\eta_p^2=0.19$)。出現頻度と親密度の交互作用は、実験参加者分析で有意傾向であった ($F_1(1,29)=3.39$, $MSE=535.78$, $p<.10$, $\eta_p^2=0.10$) が、項目分析で有意ではなかつ

た ($F_2(1,60)=0.34, MSE=1887.88$).

実験参加者分析のみにおいてだが、交互作用が有意傾向であったことから、実験参加者分析における単純主効果検定を行った。その結果、高頻度語条件において親密語の効果が有意であり ($F_1(1,29)=47.71, MSE=569.40, p<.001$)、低頻度語条件においても親密度の効果が有意であった ($F_1(1,29)=17.56, MSE=622.05, p<.001$)。一方、親密度の高・低条件における出現頻度の効果は有意でなかった (全ての $F<1.2$)。実験参加者分析のみにおいて交互作用が観察されたことは、カタカナ語における出現頻度効果は親密度の高低に関わら

ず反映されず、親密度効果は両出現頻度条件において、高親密度条件の方が低親密度条件よりも反応時間が有意に速く、高頻度条件における親密度効果 (43ms) の方が、低頻度条件における親密度効果 (27ms) よりも大きい傾向があることを示唆している。

また、両実験の誤反応率について、実験参加者分析における漢字語の両要因の高・低条件の平均誤反応率を Figure3 に、カタカナ語の両要因の高・低条件の平均誤反応率を Figure4 に示す。いずれも縦軸は誤反応率 (%), 横軸は出現頻度と親密度の各条件を、エラーバーは標準誤差を示してい

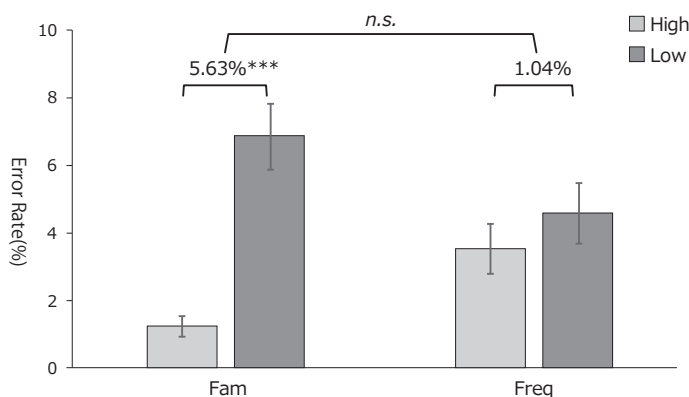


Figure 3. Mean error rate in percentage for familiarity(Fam) and frequency(Freq) for each(high and low) conditions in Kanji words experiment. Asterisks indicate significant differences (** $p < .001$, *n.s.* means non significant). Error bars reflect within-subjects SEM.

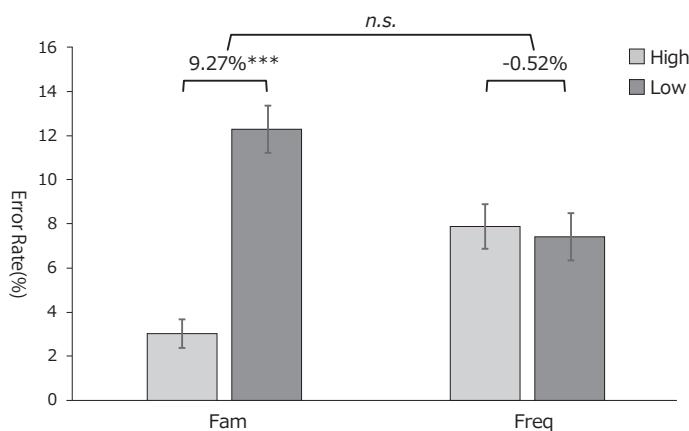


Figure 4. Mean error rate in percentage for familiarity(Fam) and frequency(Freq) for each(high and low) conditions in Katakana words experiment. Asterisks indicate significant differences (** $p < .001$, *n.s.* means non significant). Error bars reflect within-subjects SEM.

る。

漢字語の誤反応率の分析では、出現頻度の主効果が実験参加者分析と項目分析のいずれでも有意でなかった ($F_1(1,29)=2.16, MSE=15.04; F_2(1,60)=0.63, MSE=33.60$)。一方、親密度の主効果は両分析で有意であった ($F_1(1,29)=19.74, MSE=48.09, p<.001, \eta_p^2=0.41; F_2(1,60)=14.30, MSE=33.60, p<.001, \eta_p^2=0.18$)。出現頻度と親密度の交互作用は、両分析で有意でなかった ($F_1(1,29)=0.44, MSE=26.54; F_2(1,60)=0.13, MSE=33.60$)。これらの結果は、高親密度条件の方が低親密度条件よりも誤反応率が低かったことを表している。

また、カタカナ語の誤反応率の分析では、漢字語と同様に出現頻度の主効果が実験参加者分析と項目分析のいずれでも有意でなかった ($F_1(1,29)=0.24, MSE=34.40; F_2(1,60)=0.02, MSE=120.66$)。一方、親密度の主効果は両分析で有意であった ($F_1(1,29)=88.04, MSE=29.29, p<.001, \eta_p^2=0.75; F_2(1,60)=11.14, MSE=120.66, p<.01, \eta_p^2=0.16$)。出現頻度と親密度の交互作用は、両分析で有意でなかった ($F_1(1,29)=0.10, MSE=30.54; F_2(1,60)=0.01, MSE=120.66$)。これらの結果も漢字語の分析と同様に、高親密度条件の方が低親密度条件よりも誤反応率が低かったことを表している。

考察

前述のように、ことばを読むというプロセスの理解や認知モデルの構築において、出現頻度と親密度は中心的な役割を果たしている。さらに、出現頻度と親密度の間には高い相関関係が認められているが、親密度の方がより重要な要因であることが示唆されている (e.g., 天野・近藤, 2000; Gernsbacher, 1984)。しかし、Hino & Lupker (submitted) は、漢字語とカタカナ語における両要因間の相関が異なる可能性を指摘している。このことが正しければ、ことばを読む際、表記によって出現頻度と親密度それぞれの影響が異なる可能性がある。

そこで、本研究では、漢字語とカタカナ語における出現頻度効果と親密度効果について、両要因を高低に操作した語彙判断課題を行い、その影響

について検討した。実験の結果、漢字語では出現頻度効果と親密度効果がともに観察されたが、交互作用は有意でなかった。一方、カタカナ語では親密度効果のみが観察され、交互作用が実験参加者分析において有意傾向であった。この結果は、天野・近藤 (2000a) が提案する親密度のみを重視する考えに反して、表記ごとに両効果の影響は異なり、出現頻度も同時に検討する必要があることを示唆していた。このことは、漢字語とカタカナ語における音韻—形態間の一貫性の程度の違いに起因する可能性が考えられる。

Hino & Lupker (submitted) は、親密度は単語を視覚的に読む経験ばかりでなく、その単語を聴覚的に聞き取る経験にも影響される可能性を示唆している。音韻—形態対応の一貫性が高いカタカナ語は、その単語を聞き取った際に正しい綴りの情報を検索することが容易であり、綴りの学習も進行しやすい。このことから、形態情報と音韻情報の両方向からの学習が促進される。しかし、新聞や本などに代表される視覚的な情報では漢字語が使用される頻度が高いと推測されることから、カタカナ語の親密度は専ら聞き取り経験による学習によって決定されると仮定すれば、出現頻度に関わらず親密度が高くなる可能性が考えられる。その結果、両要因間の相関関係が弱くなったと思われる。このことが正しければ、カタカナ語では出現頻度効果より親密度効果の方が大きくなると予測される。

一方、音韻—形態対応の一貫性が低い漢字語の場合、その単語を聞き取った際に正しい綴りの検索に失敗する可能性が高くなる。その結果、単語の聞き取り経験があまり有効に機能しないことになる。このことから、漢字語における親密度は出現頻度を示される視覚的な読みの経験に基づくことになり、さらに、それは日常的に使用される頻度が高いと推測される。その結果、漢字語では出現頻度と親密度との間の相関関係が高くなったものと思われる。このことが正しければ、漢字語では出現頻度効果と親密度効果が同程度で観察されると予測される。

漢字語における語彙判断課題の結果は、この予測を支持するものであった。実験の結果、出現頻

度の効果と親密度の効果がともに観察され、この間の交互作用は観察されなかった。出現頻度効果が観察されたことは、音韻—形態間の一貫性が低いため音韻情報からの影響が少なく、その結果、視覚的な情報である出現頻度が影響した結果と考えられる。また、親密度効果が観察されたことは、漢字語の親密度は出現頻度と同様に、読みの経験を反映すると考えられることから、同様に観察されたものと思われる。さらに、交互作用は観察されなかったが、両効果の数値には多少の違いがあった。このことは、漢字語では高親密度語の方が低親密度語よりも音韻—形態間の一貫性が高いことから説明できる (e.g., Hino & Lupker, submitted)。高親密度語は、視覚的な情報である出現頻度による効果に加え、音韻—形態間の一貫性が高いことから生じる音韻情報の影響が加算されたものと思われる。一方、低親密度語は、音韻—形態間の一貫性が低いことにより音韻情報が有効に機能しない。この違いによって、親密度効果は出現頻度効果よりも大きな数値になったのかもしれない。これらの結果は、漢字語では出現頻度と親密度ともに語の読みの成績に影響を及ぼすことを示唆していた。

同様に、カタカナ語の結果も予測を支持するものであった。カタカナ語では、出現頻度の効果 (1ms) は観察されなかったが、親密度の効果 (35ms) とともに、実験参加者分析で交互作用が観察された。出現頻度効果が観察されなかったことは、カタカナ語の学習は視覚的な経験よりもむしろ、聴覚的な聞き取り経験に基づいたため、視覚的な情報である出現頻度は観察されなかったものと思われる。一方、親密度効果が観察されたことは、カタカナ語の親密度は聞き取り経験の貢献度が非常に高く、音韻—形態間の一貫性が高いことによって生じたものと思われる。また、高頻度語における親密度効果 (43ms) の方が低頻度語における親密度効果 (27ms) よりも大きい傾向が観察された交互作用は、高頻度語における形態情報の影響が低頻度語よりも大きかったため観察されたと推測される。

それでは、この結果は単語認知モデルを検討する上ではどのように考えられるだろうか。単語を

読む際、音韻—形態間の一貫性だけでなく、形態—音韻間の一貫性も影響を及ぼすことが知られている (e.g., Fushimi, Ijuin, Patterson & Tatsumi, 1999)。このことから考えると、形態情報と音韻情報の間の相互作用を説明できる PDP モデルなどニューラルネットを仮定するモデルの方が妥当性は高いと思われる。聴覚的な情報に関しては、音韻—形態間の一貫性をユニット間の結合強度で示すことが可能であり、また視覚的な情報に関しても形態—音韻間の一貫性を結合強度に変換することができる。最終的に、形態情報と音韻情報の間の相互作用という観点から出現頻度と親密度がそれぞれ計算されると仮定するならば、本実験の結果を上手く説明することが出来るものと思われる。

ここまで、本研究の結果について、漢字語とカタカナ語における音韻—形態間の一貫性の違いに起因するものとして解釈を行ってきた。しかし、本実験の結果は直接的に音韻—形態間の一貫性によって検討しているわけではないため、その因果関係を立証することは難しいと思われる。さらに、本実験は出現頻度と親密度を直交に操作しているが、本来は両要因の間では相関が認められている。この観点から、今回使用した刺激は一般的な単語ではなく、むしろ特異的な単語であるのかもしれない。今後、これらの問題について更なる検討が必要だと考えられる。

本研究では、漢字語とカタカナ語における出現頻度効果と親密度効果は、それぞれの表記が持つ音韻—形態間の一貫性の違いに起因する学習経験の違いによって異なる可能性が示唆された。実験の結果は、いずれの表記においても親密度効果が大きく観察されており、このことは天野・近藤 (2000a) の結果とは矛盾しない。彼らは、出現頻度は言語処理過程において間接的にしか関連しておらず、一方、親密度は読者の言語知識や経験に基づいた直接的な指標であると述べている (e.g., 天野・近藤 2000a, pp.62)。この言語知識や経験は、視覚的情報に限らず、聴覚的な情報も含むことを考えると、本実験の結果と整合性があるものと考えられる。しかし、親密度の影響の方が大きいことから、心理学実験において親密度のみを操作・統制することは本実験の結果から明らかに問題で

あると考えられる。このような問題は、失語症の研究において、主に親密度を要因として検討されている現状にも大きく関わってくるものと思われる。また、影響を及ぼす箇所は異なるが、いずれの表記においても出現頻度による影響が観察されていることから、単語を用いる実験を行う際には出現頻度の影響も十分に考慮する必要があると思われる。

引用文献

- 天野 成昭・近藤 公久 (1999). NTT データベースシリーズ日本語の語彙特性 (第1巻, 親密度), 東京:三省堂
- 天野 成昭・近藤 公久 (2000a). NTT データベースシリーズ日本語の語彙特性 (第7巻, 頻度), 東京:三省堂
- 天野 成昭・近藤 公久 (2000b). NTT データベースシリーズ「日本語の語彙特性について」, 音声研究, 4, 44-50.
- 天野 成昭・近藤 公久 (2003a). NTT データベースシリーズ日本語の語彙特性第1期 CD-ROM 版, 東京:三省堂
- 天野 成昭・近藤 公久 (2003b). NTT データベースシリーズ日本語の語彙特性第2期 CD-ROM 版, 東京:三省堂
- Balota, D.A. & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 340-357.
- Becker, C.A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology; Human Perception and Performance*, 5, 252-259.
- Becker, C.A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition*, 8, 493-512.
- Carroll, J.B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. New York: Houghton Mifflin.
- Connine, C.M., Mullennix, J., Shernoff, E. & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 1084-1096.
- Forster, K.I. (1976). Accessing the mental lexicon. In R.J. Wales and E. Walk (Eds.), *New approaches to language mechanisms* (pp.257-287). Amsterdam: North-Holland.
- Fushimi, T., Ijuin, M., Patterson, K., & Tatsumi I.M. (1999). Consistency, Frequency, and lexicality effects in naming Japanese kanji. *Journal of Experimental Psychology; Human Perception and Performance*, 25, 382-407.
- Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, 113, 256-281.
- Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory & Language*, 24, 631-645
- Hino, Y., Kusunose, Y., Miyamura, S., & Lupker, S.J. (2017). Phonological-orthographic consistency for Japanese words and its impact on visual and auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 43, 126-146.
- Hino, Y., & Lupker, S.J. (1998). The effects of word frequency for Japanese Kana and Kanji words in naming and lexical decision: Can the dual-route model save the lexical-selection account? *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1431-1453.
- Hino, Y. & Lupker, S.J. (2000). The effects of word frequency and spelling-to-sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 166-183.
- Hino, Y. & Lupker, S.J. (submitted). The impact of phonological-orthographic consistency on orthographic familiarity ratings and lexical decision performance for Japanese word. Manuscript submitted for publication.
- 広瀬 雄彦 (1984). 漢字および仮名单語の意味的处理に及ぼす表記頻度の効果, *心理学研究*, 55, 173-176.
- 広瀬 雄彦 (2007). 日本語表記の心理学—単語認知

における表記と頻度一, 京都: 北王子書房
国立国語研究所 (1970). 電子計算機による新聞の
語彙調査, 東京: 秀英出版
国立国語研究所 (1993). 分類語彙表 (フロッピー
版), 東京: 秀英出版

wordbook of 30,000 words. New York: Columbia
University, Teacher's College Press.

(2018.9.24 受稿, 2018.11.21 受理)

Kucera, H., & Francis, W.N. (1967). *Computational
analysis of present-day American English*. Providence,
R.I.: Brown University Press.

Monsell, S., Doyle, M.C. & Haggard, P.N. (1989). Effects
of frequency on visual word recognition tasks: Where
are they? *Journal of Experimental Psychology: General*,
118, 43-71.

McClelland, J.L. & Rumelhart, D.E. (1981). An
interactive activation model of context effects in
letter perception: Part 1. An account of basic findings.
Psychological Review, 88, 375-407.

Morton, J. (1969). Interaction of information in word
recognition. *Psychological Review*, 76, 165-178.

Paap, K.R., McDonald, J.E., Schvaneveldt, R.W. & Noel,
R.W. (1987). Frequency and pronounceability in
visually presented naming and lexical decision tasks. In
M. Coltheart (Ed.), *Attention and performance XII*
(pp.221-243). Hillsdale, NJ: Erlbaum.

Paap, K.R., Newsome, S.L., McDonald, J.E. &
Schvaneveldt, R.W. (1982). An activation-verification
model for letter and word recognition: The word-
superiority effect. *Psychological Review*, 89, 573-594.

Plaut, D.C. (1997). Structure and function in the lexical
system: Insights from distributed models of word
reading and lexical decision. *Language and Cognitive
Processes*, 12, 765-805.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S. &
Patterson, K. (1996). Understanding normal and
impaired word reading: Computational principles in
quasi-regular domains. *Psychological Review*, 103,
56-115.

Rubenstein, H., Garfield, L., & Millikan, J.A. (1970).
Homographic entries in the internal lexicon. *Journal of
Verbal Learning and Verbal Behavior*, 9, 487-494.

Seidenberg, M.S. & McClelland, J.L. (1989). A
distributed, developmental model of word recognition
and naming. *Psychological Review*, 96, 521-568.

Thorndike, E., & Lorge, I. (1944). *The teacher's*

研究ノート

報告・資料

