

ベトナム人日本語学習者による漢字語の翻訳における 音韻類似性と使用頻度の影響

Hoang Thi Lan Phuong (名古屋大学大学院生)

玉岡賀津雄 (名古屋大学)

于劭贇 (名古屋大学大学院生)

1. 研究の目的

語彙の書字, 音韻, 概念, 統語情報へのアクセスは, 2 言語間で言語非選択的 (language non-selective) であるといわれる (Costa, Caramazza, & Sebastián-Gallés, 2000; De Groot, Delmaar, & Lupker, 2000; Dijkstra & van Heuven, 1998, 2002; Green, 1998; Levelt, Roelofs, & Meyer, 1999; van Heuven, Dijkstra, & Grainger, 1998; van Heuven, Schriefers, Dijkstra, & Hagoort, 2008 など)。中国人日本語学習者の場合, 日中 2 言語間で書字が極めて類似しているため, 書字を介して日本語の語彙の意味を理解すると考えられる。一方, ベトナム人日本語学習者の場合は, 日越両言語間で漢語を共有しているものの, ベトナム語はアルファベット表記, 日本語は仮名・漢字表記であるため, ベトナム語と日本語で音韻的に類似している場合は, 音を介して日本語の語彙が理解されるのではないかと予想される。また, 語彙処理において, 最も強く影響するのは使用頻度であるといわれている。使用頻度の高い語彙は, 低い語彙と比べて想起されやすい (Grainger and Jacobs, 1996; McClelland and Rumelhart, 1981; Morton, 1969; Norris, 1986; Solomon and Postman, 1952; 日本語では, 玉岡・高橋, 1999 など)。そこで, 日越間の語彙使用頻度を含んで, 特に音韻類似性の効果を翻訳課題 (oral translation task) の実験で検討した。

2. 日越 2 言語間の音韻類似性データベース

旧・日本語能力試験出題基準〔改訂版〕(国際交流基金・日本国際協会, 2007) の 4 級から 2 級までの同形の漢字 2 字語のリスト (朴・熊・玉岡, 2014a, 2014b; 熊・玉岡, 2014; 于・玉岡, 2015 のデータベースを使用) を基礎資料として, 日越両言語に共通して使用される漢語 2 字語のベトナム語と日本語の発音を訓令式ローマ字表記で表記した。そして, 日越 2 言語間の音韻類似性について, 音韻的距離 (phonological distance, Buchta & Hahsler, 2016) と音素類似性 (phonemic similarity, 于, 玉岡, 2015) の 2 つの指標で計算した。日越で共通に使用されている漢字 2 字語 2,058 語中の 1,485 語で 72.16% であり, 語レベルの音韻的距離の平均は 6.06 で, 標準偏差は 2.52 であった。語レベルの音素類似性の平均は 0.5 で, 標準偏差は 0.2 であった。音韻的距離と音素類似性のピアソン積率相関係数は非常に高く, $N=1,485, r=-0.92, p<.001$ であった。両指標は極めて類似した指標であるため, 本研究では音韻的距離を使った。

3. 刺激語

日越2言語間の音韻類似性データベース (kanjigodb.herokuapp.com/) から日本語能力試験の4, 5級と2, 3級の50語ずつを抽出した。合計100語を翻訳課題実験の刺激語として使用した。刺激語の日本語の語彙使用頻度は、「2,136字の日本語常用漢字 Web 検索データベース (www.kanjidatabase.com, Tamaoka, Makika Sanders, Verdonschot, 2017) に掲載された毎日新聞の使用頻度を用い、ベトナム語の語彙使用頻度はベトナム新聞コーパスの Vietnam Lexicography Centre (vietlex.com/, Vietlex, 2000-2017) を用いて調べた。そして、100語の語彙使用頻度の平均と標準偏差を参照して、日本語は高(32語)・中(36語)・低(32語)に、ベトナム語は高(30語)・中(40語)・低(30語)の3群に分けた。

4. 日本語学習者

ホーチミン市師範大学外国語学部日本語分野のベトナム語を母語とする日本語学習者38名(男性:9名, 女性:29名)対象に実験を行った。この内、日本語能力試験のN3を持っている学生は5名, N2を持っている学生は26名, N1を持っている学生は7名であった。平均年齢は24歳3ヶ月(標準偏差は2歳2ヶ月), 日本語学習歴の平均は5年と0カ月(標準偏差は1年6ヶ月)であった。また、日本で滞在した経験があるのは7名(長期が1年3ヶ月で4名, 短期が6ヶ月で3名)であった。

5. 語彙テスト

被験者の日本語の語彙の習熟度を測定するために、非漢字圏日本語学習者のための語彙テスト(大和・玉岡・茅本, 2016)を実施した。このテストは、和語と漢語および機能語の3語種または名詞、動詞、形容詞の3つの品詞で設定されており、各12問ずつの合計36問である。ターゲット語の語彙レベルは、旧日本語能力試験で4級から2級であり、本調査の日本語学習者の語彙力を測定するのに適したテストである。テスト形式は、短文の空所に適切な語を入れる四者択一形式で、1問1点の36点満点である。フィリピン人日本語学習者103名にこのテストを実施した結果、クロンバックの信頼度係数は0.74で、ある程度高い信頼性が報告されている(大和・玉岡・茅本, 2016)。本実験では、ベトナム人日本語学習者38名にこの語彙テストを実施した。その結果、最低が28点、最高が満点の36点で、平均は33.53点、標準偏差1.65点であった。満点が36点であるため、本調査の日本語学習者の語彙力は高く、差が小さいことが分かる。得点の分布は、歪度が-1.24、尖度が2.08であり、得点が高いほうに偏り、低いほうに裾が広がっていることが分かる。この語彙テストで、平均が34点以上を上位群(24名)、未満を下位群(14名)とした。

6. 手法

実験用ソフトのE-prime 2.0 professional を使用して実験プログラムを作成した。コンピュータ(Dellのラップトップコンピュータ)の中央に凝視点「*」を600ミリ秒提示してから、100語

の刺激語を（横書き）ランダムに提示した。漢字 2 字からなる語を視覚的に提示して、ベトナム人日本語学習者がマイクに向かって、できるだけ迅速かつ正確にベトナム語に翻訳するように教示した。5 秒を越えた刺激語は誤りとされ、次の課題に行く。本番前に 8 語の試行練習を行った。反応から次の刺激提示までの間隔は 3 秒であった。翻訳が正解かどうかは、実験者がその場で判断し入力した。この課題では、日本語の視覚提示からベトナム語の訳を発音するまでの翻訳潜時（translation latency time）と正答（accuracy）を測定した。

7. 決定木分析の結果

マイクに反応しなかった項目と、300 ミリ秒（ms）以内の潜時は誤りと判断した。また、日本語で読んでしまった場合（3,800 項目の中 129 項目で、全体の 3.39%）は分析から除外した。本実験で有効であった 3,671 項目のデータの正答率は 91.26%で、標準偏差は 28.25%であった。正答率が 90%以上になっていたため、翻訳潜時に焦点を当てて分析した。なお、日本語の漢字 2 字語をベトナム語に訳するための翻訳潜時（ $N=3,671$ 項目）の平均は 1,490 ms で、標準偏差は 759ms であった。SPSS Statistic 22.0 上で起動する決定木（回帰木）で解析した。決定木分析では、翻訳翻訳潜時を予測する変数として 5 つを設定した。それらは、（1）ベトナム人日本語学習者の語彙能力（上位群と下位群）、（2）日本語の語彙使用頻度（高・中・低）、（3）ベトナム語の語彙使用頻度（高・中・低）、（4）日越 2 言語間の音韻類似性（高・中・低）、（5）語彙の日本語能力試験でのレベル（日本語能力試験 4, 5 級レベルと 2, 3 級レベル）、である。日本語をベトナム語に翻訳するにあたり、やはり日本語の語彙使用頻度が強い予測変数となった [$F(2, 3668)=50.37, p<.001$]。日本語の語彙使用頻度が高い場合（ $M=1,387\text{ms}, SD=687\text{ms}$ ）、中程度の場合（ $M=1,468\text{ms}, SD=732\text{ms}$ ）、低い場合（ $M=1,626\text{ms}, SD=833\text{ms}$ ）の順に処理速度が落ちた（図 1 を参照）。

さらに、日本語の語彙使用頻度の各グループについてそれぞれ異なる要因が影響した。日本語の語彙使用頻度が高い場合には、日越 2 言語の音韻類似性が促進的に影響した [$F(1, 1147)=19.67, p<.001$]。音韻類似性が高・中の場合（ $M=1,328\text{ms}, SD=673\text{ms}$ ）は、低い場合（ $M=1,538\text{ms}, SD=710\text{ms}$ ）よりも翻訳潜時が短かった。さらに、音韻類似性が高・中の場合には、ベトナム語の語彙使用頻度の影響がみられた [$F(1, 874)=14.43, p<.001$]。ベトナム語の語彙使用頻度が高い場合（ $M=1,236\text{ms}, SD=583\text{ms}$ ）は、中・低の場合（ $M=1,408\text{ms}, SD=733\text{ms}$ ）よりも翻訳潜時が短かった。つまり、日本語の語彙使用頻度が高い場合の日越 2 言語の音韻類似性が高く、ベトナム語の語彙使用頻度が高い漢字 2 字語は、それらの語彙に対する翻訳潜時が最も短かった。

次に、日本語の語彙使用頻度が中程度の場合は、学習者の語彙力で翻訳潜時に違いがみられた [$F(1, 1353)=23.52, p<.001$]。学習者の上位群（ $M=1,397\text{ms}, SD=661\text{ms}$ ）のほうが下位群（ $M=1,595\text{ms}, SD=827\text{ms}$ ）よりも翻訳潜時が短かった。上位群と下位群共に、語彙レベル（難易度）の要因が影響した。上位群の場合には [$F(1, 863)=7.62, p<.01$]、語彙レベルの 4, 5 級の翻訳潜時が、2, 3 級よりも速かった。下位群の場合 [$F(1, 670)=9.29, p<.01$] も、4, 5 級のほうが 2, 3 級より翻訳潜時が短かった。やはり、簡単な語彙は、難しい語彙より短時間で翻訳される。

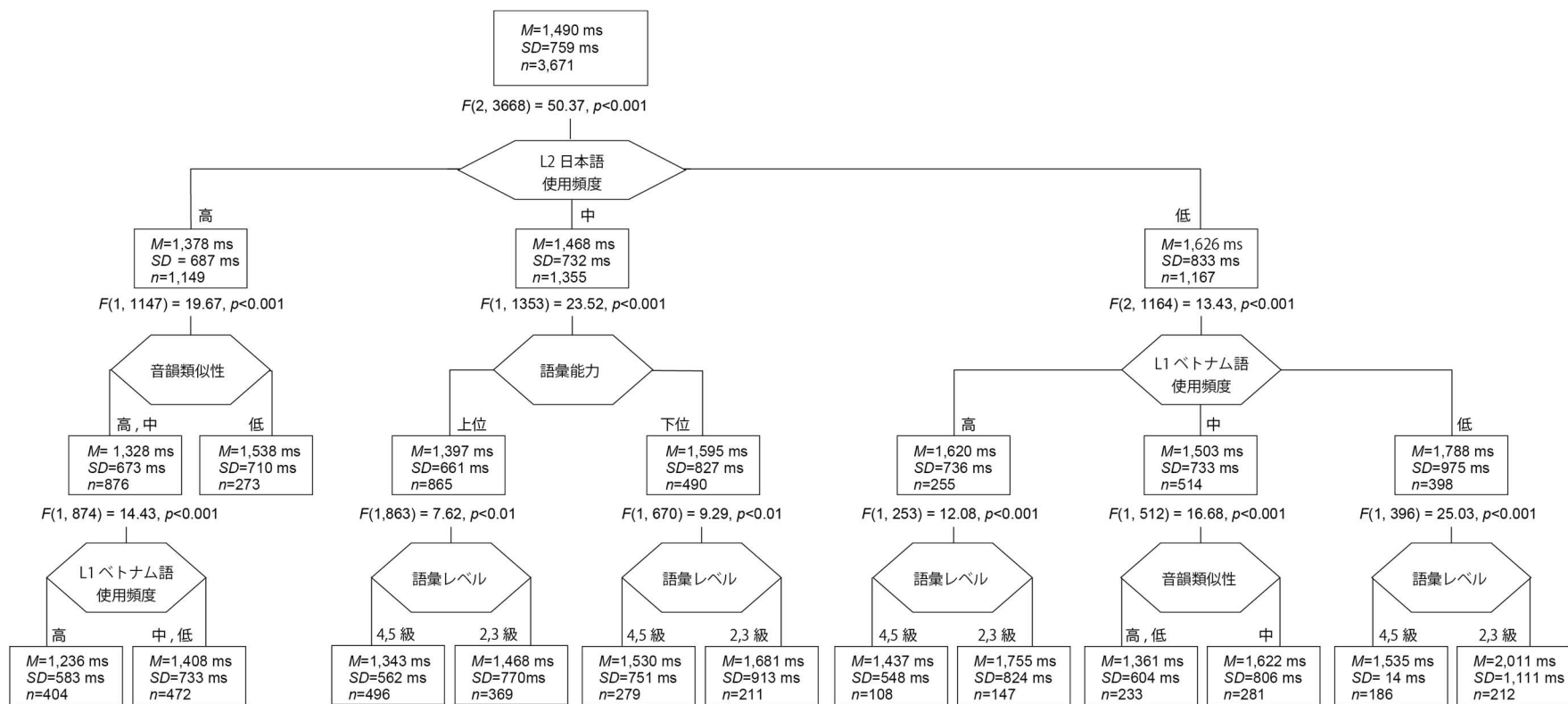


図 1. 日本語がからベトナム語への翻訳潜時を予測する決定木分析の結果 (樹形図)

最後に、日本語の語彙使用頻度が低い場合には、ベトナム語の語彙使用頻度の効果がみられた $[F(2, 1164)=13.34, p<.001]$ 。ベトナム語の語彙使用頻度が中程度の語彙 ($M=1, 503\text{ms}$, $SD=733\text{ms}$) は、高い語彙 ($M=1, 620\text{ms}$, $SD=736\text{ms}$) と低い語彙 ($M=1, 788\text{ms}$, $SD=975\text{ms}$) よりも翻訳潜時が短かった。おそらく、中程度の語彙に日越2言語の音韻類似性の促進効果があったのではないかと考えられる $[F(1, 512)=16.68, p<.001]$ 。ただし、音韻類似性が高・低の語彙 ($M=1, 361\text{ms}$, $SD=604\text{ms}$) が中程度の語彙 ($M=1, 622\text{ms}$, $SD=806\text{ms}$) よりも翻訳潜時が短かった。また、ベトナム語の語彙使用頻度が高い場合は、語彙レベル(難易度)が影響した $[F(1,253)=12.08, p<.001]$ 。さらに、ベトナム語の語彙使用頻度が低い場合も、語彙レベル(難易度)が影響した $[F(1,369)=25.03, p<.001]$ 。それらの場合は、日本語能力試験4,5級レベルの簡単な語彙の翻訳潜時が、難しい2,3級レベルの語彙よりも短かった。

8. 考察

本研究では、ベトナム人日本語学習者の38名に翻訳課題を実施して、5つの要因で学習対象の日本語から母語のベトナム語への翻訳に要する時間(翻訳潜時)を予測する決定木分析を行った。以下に、それらを要約する。

8-1. 日越の語彙使用頻度の影響

ベトナム人日本語学習者が、日本語の漢字2字語を翻訳する速度に最も強く影響したのは日本語の語彙使用頻度であった。つまり、日本語で頻繁に使用される語彙は、その語彙を知覚しやすく、母語のベトナム語に翻訳するための意味・概念へのアクセスも短時間で行われたと考えられる。一方、ベトナム語の語彙使用頻度は、日本語の語彙使用頻度が高い場合と低い場合に影響した。図1に表示したように、日本語の語彙使用頻度が高い場合には、ベトナム語の語彙使用頻度も翻訳速度に促進的に影響した。たとえば、漢字2字語の「教員」は、ベトナム語では *giao vien* である。この種の漢字語は、日越両言語の語彙使用頻度が高いため、翻訳潜時がさらに短くなった。また、日本語の語彙使用頻度が低い場合にも、ベトナム語の語彙使用頻度が影響した。たとえば、漢字2字語の「会話」は、ベトナム語では *hội thoại* である。この語は、日越両言語で語彙使用頻度が低い。両方の使用頻度が低いことから、この種の漢字語の翻訳潜時の平均は1,788msと長くなった。日本語の漢字語を処理には、日本語の語彙使用頻度が強く影響するが、さらにベトナム語の語彙使用頻度も独自に影響した。

8-2. 日越間の音韻類似性の影響

漢字2字語の処理における日越両言語間の音韻類似性の効果も有意な予測変数となった。図1に示したように音韻類似性の効果は2か所でみられた。まず、日本語の語彙使用頻度が高い語彙について、音韻類似性が高い語彙と中程度の語彙が低い語彙よりも翻訳速度が速かった。次に、日本語の語彙使用頻度が低い場合で、ベトナム語の語彙使用頻度が中程度の語彙にも影響した。ただし、この場合の音韻類似性の効果は、高い語彙と低い語彙が中程度の語彙より翻訳潜時が短かった。原則として、日越で発音が大きく異なる語彙については、翻訳潜時が長くなる傾向がうかがえるが、この点は今後の検討課題である。

8-3. ベトナム人日本語学習者の語彙能力の影響

実験に参加したベトナム人日本語学習者は、日本語能力試験の N3 から N1 までに相当する。非漢字圏のための語彙テスト（大和・玉岡・茅本, 2016）で、上位群と下位群に分けて比較した。基本的に、語彙力が上位群のほうが下位群よりも翻訳速度が速かったが、決定木分析では、日本語の語彙使用頻度が中程度の場合のみに影響がみられた。日本語の語彙使用頻度が中程度の語彙は、学習者の語彙力が強く反映されるようである。語彙力は日越両言語間の音韻類似性とベトナム語の語彙使用頻度が同じくらいの効果であった。

8-4. 語彙の難易度のレベル（2, 3 級レベルと 4, 5 級レベル）の影響

日本語能力試験の 2, 3 級と 4, 5 級の 2 つのグループに分けた語彙の難易度のレベルの効果は弱い予測変数になった。効果は限定的で、日本語の語彙使用頻度が中程度で、学習者の語彙力の上位群と下位群に影響した。また、日本語の語彙使用頻度が低く、ベトナム語の語彙使用頻度が高い場合と低い語彙に促進効果がみられた。基本的に、日本語能力試験における 4, 5 級の語彙のほうが 2, 3 級の語彙よりも翻訳速度が速かった。

参考文献（一部）

- Buchta, C., & Hahsler, M.(2016). Cba: Clustering for business analytics. R package version 0.2-17. <https://cran.r-project.org/package=cba> .
- Colomé, A. (2001). Lexical Activation in Bilinguals' Speech Production: Language-Specific or Language-Independent? *Journal of Memory and Language*. **45(4)**. 721-736.
- Caramazza, C. & Sebastián-Gallés, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*. **26(5)**. 1283-96.
- De Groot, A. M. B., Delmaar, P. & Lupker, S. J. (2000). The processing of interlexical homographs in a bilingual and a monolingual task: Support for non-selective access to bilingual memory. *Quarterly Journal of Experimental Psychology*, **53**, 397-428.
- Dijkstra, A., & van Heuven, W. J. B., (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, **5 (3)**, 175–197.
- Grainger, J. & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, **103**, 518-565.
- Green, D. W. (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition*, **1**, 67-81.
- 国際交流基金・日本国際教育協会（2007）『日本語能力試験出題基準（改訂版）』（第 4 版）東京：凡人社。
- 玉岡賀津雄・高橋登（1999）「漢字二字熟語の書字行動における語彙使用頻度および書字的複雑性
の影響」『心理学研究』70, 45-50.
- 于劭贇・玉岡賀津雄（2015）「日韓中同形二字漢字語の品詞性ウェブ検索エンジン」『ことばの科学』**29**, 43-61.