

中国人日本語学習者は 日本語の漢字の書き取りが正しくできるのか？¹

張 婧禕（名古屋大学大学院人文学研究科・助教）
玉岡 賀津雄（名古屋大学大学院人文学研究科・教授）
初 相娟（天津外国語大学・副教授）

要 旨

日中の漢字語には書字的に類似した同形語が多く存在する。そのため、漢字知識を有している中国人日本語学習者にとって、同形語を書字的に理解できても、音韻的に理解するのは難しいと予想される。そこで、本研究では、中国人日本語学習者を対象に、日本語の語彙力を測るテストと2字漢字語の平仮名提示での書き取りテストを実施して、中国人日本語学習者が、音韻表記から書字への転換を必要とする漢字語の書き取りをどのように達成しているかを検討した。そのために、(1) 日本語の使用頻度、(2) 中国語での存在有無、(3) 日中音韻的類似性、(4) 学習者による日本語の語彙力の4つの要因（独立変数）から、漢字語の書き取りの正誤（従属変数）を予測する決定木（分類木）分析を行った。その結果、全体の正答率は58.60%となり、予想以上に低かった。書き取りに影響した最も強い要因は、L2 日本語での使用頻度で、次にL1 の中国語での存在有無となった。その後で、日中音韻的類似性および日本語の語彙力も関与していた。漢字語が平仮名で音韻的に提示された場合、中国人日本語学習者は、L2 日本語で慣れ親しんだ使用頻度の高い語であることが前提となり、それがL1 中国語に存在するかどうかを参照することで漢字語の書字が促進され、L2 日本語の語彙力やL1 とL2 両言語での音韻的類似性も弱い影響があることを実証した。

キーワード：メンタルレキシコン、漢字知識、中国人日本語学習者、決定木分析

1 はじめに

人間の脳には、語彙に関わる情報が記憶されていると想定され、これらの語彙

¹ 本研究は2017年7月22日（土）に一橋大学で行われた「中国語話者のための日本語教育研究会・第39回大会」で発表した内容をもとにして、加筆したものである。

の記憶の総体はメンタルレキシコン (mental lexicon) と呼ばれる。成人の母語話者では, stop, stops, stopped などの派生語をそれぞれ別の語として数えて約5万語あるいはそれ以上の語彙がメンタルレキシコンに記録されており, そのうち, 約3分の2が理解や産出のために使われていると言われている (Aitchison, 1987)。また, メンタルレキシコンには, 各語の複数の特性が記載されており, それらは表象群と呼ばれる。各語の表象群は書字, 概念, 音韻および統語の4種類からなる (玉岡, 2013, 2017)。そして, 2つの言語間で文字表記が類似している場合には, 語彙処理において促進効果がみられ, これは「表記親近性効果 (script familiarity effect)」と呼ばれている。たとえば, アルファベット表記のL1 インドネシア語話者がL2 英語を学習した場合は, たとえ両言語に同形語 (あるいは同根語) がほとんど存在しなくても, L2 英語の書字的な処理が迅速かつ正確に行えることが報告されている (Djojomihardjo, et al., 1994)。また, 日中両言語では, 漢字を文字表記として使用しており, 両言語で多くの漢字および漢字語が類似している。そのため, 日本語能力が同じ条件で比較すると, L1 中国語話者はL1 英語話者およびL1 韓国語話者に比べ, L2 日本語の漢字語の書字を介した意味理解が迅速かつ正確に行えることもよく知られている (石田, 1986; 小森, 2005; 玉岡, 1994, 1997, 2000; 大和・玉岡, 2013; 張婧禱, 2017 など)。つまり, 中国人日本語学習者はL1 中国語の漢字知識を活用し, L2 日本語の漢字表記の語の意味理解に援用できる。

日中両言語の書字的類似性による意味理解が促進されることは多くの先行研究で示されている。一方, L1 の中国語とL2 の日本語で漢字語の発音はかなり異なるので, 両言語間での音韻表象群の結合が弱いと思われる。そのため, 音韻的類似性の影響はあまりないと予想され, そのような結果が報告されている (邱学瑾, 2002a, 2002b; 松下ほか, 2004 など)。しかし, 日本語の漢字語を音声的に聞いて漢字で書く場合を考えると, 日中の音韻的類似性が影響するのではないかと予想される。そこで, L2 の日本語の書字に何が関与するかを明らかにするために, 中国語を母語とする日本語学習者を対象に, 以下の4つの要因の影響を検討することにした。第1に, 日本語での使用頻度である。これは, 学習者が日本語の学習において漢字語に触れる頻度を反映していると思われる。第2に, 漢字語の中国語での存在の有無である。これは, 学習者がL1 の中国語を参照しながら日本語の漢字語を想起していると予想されるので, その程度を反映していると考えら

れる。第3に、日本語の漢字語が、日中両言語において音韻的にどのくらい類似しているかである。音韻的類似性は日本語の漢字語の理解には影響しないと言われているが（即ち、日中両言語における音韻的表象の結合関係が希薄）、平仮名から漢字語を書く場合には、音韻的な類似性が影響すると思われる。そして、第4に、学習者の日本語の語彙力である。日本語の語彙の豊富な学習者は、多様な漢字語を音による手掛かりにして書けると予想される。本研究では、以上の4つの要因が日本語の漢字語の漢字の書き取りにどう影響するかを検討する。

2 調査被験者

本研究は、日本語を専攻する中国人2年生80名を被験者とした。女性66名、男性14名であった。平均年齢は20歳0カ月、標準偏差は6カ月であった。最年少は18歳5カ月で、最年長は22歳0カ月であった。また、調査実施時点では、これらの日本語学習者は日本語能力試験で2級から1級合格の間くらいのレベルであり、中上級学習者に相当する。このような被験者に日本語の語彙能力テストおよび日本語の2字漢字語の書き取りテストを同時に行った。

3 調査内容

3.1 日本語の語彙能力の測定

日本語の語彙力を測定するため、宮岡ほか(2011)が開発したテストを借用し、80名の中国人日本語学習者に実施した。この語彙テストは四者択一の形で、名詞、動詞、形容詞、機能語の4つの下位分類からなる。各下位分類につき12問である。1問1点であるため、満点は48点となる。テストの問題例として、たとえば、「彼女はどんなに大変なときでも、()ひとつ言わずに病人の世話をしている。」という文を完成するために、A「語句」、B「苦難」、C「不評」、D「愚痴」の4つの選択肢から1つ正しいものを選ぶ問題がある。ここでは、D「愚痴」が正答である。宮岡ほか(2011)は、中国で281名の調査対象者にこの語彙テストを行った。クロンバックの信頼度係数(α)は0.74で、内的な一貫性は高かった。

この語彙テストはこれまでに、中国語や韓国語を母語とする日本語学習者を対象に、日本語能力を弁別する目的で、多数の研究に利用され、総じて1,846名の中国人および韓国人日本語学習者に実施された。いずれも高いクロンバックの信

信頼度係数が示され、信頼性の高いテストであることが実証されている。たとえば、木山ほか (2011)、斉藤ほか (2012)、張婧禕・玉岡 (2017) の研究でこの語彙テストが実施されているが、クロンバックの信頼度係数は、それぞれ $\alpha=0.89$ ($N=224$)、 $\alpha=0.85$ ($N=113$) および $\alpha=0.89$ ($N=143$) である。本研究では、この信頼性の高い語彙テストを日本語学習者に実施し、得点に基づいてグループ分けを行った。

3.2 日本語の漢字語の書き取りの測定テスト

本研究で実施する漢字語の書き取りテストは、宮岡ほか (2009) が開発したものである。テスト形式は、「わたしは あした ひこうきで とうきょうへ (しゅっぱつ) する。」または「これは (けつえき) の ながれを よくするくすりです。」のように、すべて平仮名表記の単文を使って被験者に提示し、括弧内の平仮名を漢字に変換してもらうものである。ここでは、それぞれ「出発」と「血液」が正解である。この漢字書き取りテストは、分析対象の 48 語と研究の分析対象とされないダミー語の 20 語の合計 68 語によって構成されている。

調査対象の 48 語は表 1 に示した。すべて日本語の新・旧常用漢字のいずれにも含まれ、『日本語能力試験出題基準』(2002, 改訂版) の 2~4 級から抽出した馴染みのある漢字である。これらの 24 語のターゲット漢字は、2 字漢字語の漢字の前後で分け、2 字漢字語の書き取り調査語を作成した。さらに、毎日新聞記事のコーパスの使用頻度を使って、これらの調査語を高低の 2 つのグループに分けた。本研究で使った日本語の使用頻度は、CD-毎日新聞データ集本社版(日外アソシエーツ, 1998-2015) の 1998 年から 2015 年までの 18 年分の新聞記事テキストデータを利用した。それにより、見出しと本文を抽出して、漢字語の頻度を算出した。このコーパスの形態素の総数(種類)は 663,243 であり、句読点などの記号を含む延べ頻度総計は 470,155,446 である。また、記号を除いた頻度総計は 398,406,147 で、約 4 億の形態素数である。このテキストデータを使って、48 語の調査語の使用頻度を表 1 に示した。1,000 回を基準として、1,000 回以下を低使用頻度とし、1,000 回以上を高使用頻度とした。その後、両使用頻度の違いを独立したサンプルの t 検定で比較した。その結果、高使用頻度 ($M=12,423$, $SD=17,000$) と低使用頻度 ($M=430$, $SD=264$) に有意な違いがみられた [$t(46)=3.46$, $p<.001$]。したがって、この 48 語における使用頻度の高低は有意に異なることが

確認でき、調査語における1つの特徴を示した。たとえば、24語のターゲット漢字のうち、2字漢字語の1番目の漢字に位置する「出」であれば、高使用頻度から「出発」(頻度=12,796)と低使用頻度から「出社」(頻度=688)のような対、2番目の漢字の「空」であれば、高使用頻度から「上空」(頻度=4,768)と低使用頻度から「天空」(頻度=177)のような対、というように24対を作った。このようにして、調査語の2字漢字語の48語を、12語のターゲット漢字を2字漢字語の

表1 漢字テストの調査語を含む詳細および特徴

ターゲット漢字	日本語能力試験配当級	高使用頻度の2字漢字語				低使用頻度の2字漢字語					
		調査語	語彙使用頻度	画面数	中国語での存在有無	日音韻的類似性指数	調査語	語彙使用頻度	画面数	中国語での存在有無	日音韻的類似性指数
出	4	出発	12,796	14	有	0.45	出社	688	12	無	0.62
急	3	急死	1,225	15	無	0.25	急病	286	19	有	0.13
食	4	食事	11,003	17	無	0.25	食通	67	19	無	0.39
映	3	映画	37,275	17	無	0.37	映写	169	14	無	0.33
安	3	安心	7,039	10	有	1.00	安住	558	13	無	0.79
試	3	試作	1,195	20	無	0.54	試用	71	18	有	0.69
探	2	探算	4,301	25	無	0.76	採集	761	23	有	0.33
吸	2	吸収	5,558	10	有	0.38	吸引	377	10	有	0.40
危	2	危険	14,254	17	有	0.34	危害	595	16	有	0.53
簡	2	簡単	4,154	27	有	0.62	簡便	107	27	有	0.57
血	2	血液	5,675	17	有	0.49	血清	359	17	有	0.31
消	2	消化	2,008	14	有	0.45	消失	435	15	有	0.54
図	3	地図	5,152	13	有	0.50	海図	226	16	無	0.58
空	3	上空	4,768	11	有	0.25	天空	177	12	有	0.43
金	4	年金	17,181	14	有	0.62	入金	923	10	無	0.50
室	3	教室	10,533	20	有	0.41	洋室	69	18	有	0.43
学	4	医学	5,820	16	有	0.48	勉学	360	18	無	0.29
行	4	旅行	20,593	16	有	0.20	歩行	867	14	有	0.00
因	2	原因	34,289	16	有	0.54	敗因	861	17	有	0.73
痛	2	苦痛	2,340	20	有	0.64	腹痛	651	25	有	0.48
念	2	残念	4,577	18	無	0.62	無念	475	20	無	0.54
報	2	情報	78,459	23	有	0.17	誤報	430	26	有	0.17
劇	2	演劇	6,887	29	無	0.20	観劇	339	33	有	0.29
難	2	盗難	1,260	29	無	0.67	受難	476	26	有	0.79
		平均	12,422.58	17.83		0.47	平均	430.29	18.25		0.45

前後の位置関係で、使用頻度の高低を統制した（12 語× 2つの 2 字漢字語の前後位置× 2 種の使用頻度）。本研究では、2 つの漢字を組み合わせて作られている高低使用頻度の 48 語（各 24 語）に、ターゲット漢字が含まれた場合に、それらの漢字が書けるかどうかを調べることを目的としている。なお、48 語の 4 つの特性の（1）視覚的複雑性（画数）、（2）ターゲット漢字の日本語能力試験配当級、（3）中国語での存在有無および（4）日中の音韻的類似性指数は使用頻度別である。はじめの（1）と（2）の特性は統制条件であり、（3）と（4）の特性は、ターゲット漢字の書き取りの正誤（従属変数）を予測する独立変数である。

3.2.1 視覚的複雑性（画数）

高低使用頻度の漢字の視覚的複雑性の指標として画数を使った。高低使用頻度で分類した 2 字漢字語の画数を合計して、独立したサンプルの t 検定を使って比較した。その結果、高使用頻度の漢字語の画数 ($M=17.83$, $SD=5.45$) と低使用頻度の漢字語の画数 ($M=18.25$, $SD=5.86$) に有意な違いはなかった [$t(46)=-0.26$, ns]。同様に、ターゲット以外の漢字の画数についても、独立したサンプルの t 検定で高使用頻度の漢字語に含まれる漢字の画数 ($M=8.46$, $SD=3.04$) と低使用頻度の漢字語に含まれる漢字の画数 ($M=8.88$, $SD=3.66$) に有意な違いはなかった [$t(46)=-0.43$, ns]。漢字語の高低使用頻度条件で、48 語の調査語の画数は統制されていることが確認できた。

3.2.2 ターゲット漢字の日本語能力試験配当級

ターゲットの 24 漢字の難易度は、『日本語能力試験出題基準』（2002, 改訂版）の語彙配当級を基準として調べた。その結果、2 級が 12 語、3 級が 7 語、4 級が 5 語であり、1 級レベルの漢字は含まれていない。また、これらの漢字はすべて中国語に存在しており、中国語を母語とする日本語学習者にこれらの漢字が書けないということはないと思われる。

3.2.3 中国語での存在有無

漢字語の 48 語の中国語での存在有無は、高低使用頻度で統制した。48 語（高低各 24 語）の調査語のうち、中国語に存在する漢字語は高低使用頻度でそれぞ

れ 16 語，中国語に存在しない漢字語はそれぞれ 8 語である。したがって，中国語での漢字語の存在の有無は，使用頻度の高低で同じ数である。

3.2.4 日中の音韻類似性

48 の漢字語に含まれる漢字はすべて日中両言語で使われている。そのため，日中で，各語を構成する 2 つの漢字の発音を比較することで音韻類似性を判定することができる。そのために，日中両言語の発音をある程度共通した音素で統一して表記して，比較する必要がある。そこで，日本語に似た音を基準として，中国語の発音を表 2 のように変換して表記した。

表 2 本研究で行った中国語の音素変換表記

中国語の音素表記	変換後の音素表記	中国語の音素表記	変換後の音素表記
/j/	/z/	/sh/	/sy/
/z/	/j/	/ch/	/cy/
/ju/	/zv/	/zh/	/jy/
/x/,/s/	/s/	/l/	/r/
/q/	/t/	/r/	/l/

このように中国語の音素表記を変換した後で，日中で類似した音素を 1 とし，類似していない音素を 0 とし，以下の公式で両言語の音韻的類似度を計算した。

$$\frac{\text{漢字}_1 \text{の類似音素数}}{\text{漢字}_1 \text{の総音素数}} + \frac{\text{漢字}_2 \text{の類似音素数}}{\text{漢字}_2 \text{の総音素数}} = \text{音韻的類似性}$$

計算の結果は，表 1 に報告した。日本語での使用頻度の高低で独立したサンプルの t 検定を行った結果，高使用頻度の音韻的類似性 ($M=0.47$) と低使用頻度の音韻的類似性 ($M=0.45$) に有意な違いはなく，高低使用頻度で音韻的類似性は統制されている [$t(46)=0.25, ns$]。さらに，全体の音韻的類似性の平均である 0.46 を境界値として，その値を超えた場合は音韻的類似性の高い漢字語，この境界値以下の場合は音韻的類似性の低い漢字語として，独立変数として使った。

4 分析および結果報告

4.1 2種テストの実施結果および語彙力による2群分け

80名の被験者に実施した日本語の語彙テストおよび漢字語の書き取りテストの結果は表3に示した。日本語の語彙テストの平均は24.30点で標準偏差は6.37点であった。48点満点のうち、最高得点は41点で最低得点は10点であった。12点満点の各下位分類では、形容詞の平均は5.23点で標準偏差は1.94であった。動詞の平均は6.41点で標準偏差は2.29点となった。名詞の平均は7.26点で標準偏差は1.90点であった。機能語の平均は5.40点で標準偏差は1.99点となった。

表3 2種テストおよび各下位分類の平均および信頼度係数

テストおよび 下位分類	満点	<i>M</i>	<i>SD</i>	Max	Min	信頼度係数 (α)
日本語の語彙テスト	48	24.30	6.37	41	10	0.78
形容詞	12	5.23	1.94	9	1	
動詞	12	6.41	2.29	11	2	
名詞	12	7.26	1.90	11	2	
機能語	12	5.40	1.99	11	2	
漢字語のテスト	48	25.55	7.02	41	11	0.87
高使用頻度	24	16.84	2.51	23	10	
低使用頻度	24	8.71	5.14	20	0	

注: $N=80$. M は平均, SD は標準偏差を示す。

一方、漢字語の書き取りテストでは、48点満点のうち、平均が25.55点で、標準偏差が7.02点であった。最高得点は41点で、最低得点は11点となった。さらに、24点満点の高低使用度の下位分類で得点をみると、高使用頻度の漢字語の平均が16.84点、標準偏差が2.51点であった。それに対し、低使用頻度の漢字語の平均が8.71点、標準偏差が5.14点となった。80名の日本語学習者のクロンバックの信頼度係数(α)は、語彙テストが0.78、漢字語の書き取りが0.87であり、いずれもかなり高い数値であり、信頼できるテストであることが示された。語彙テストの平均($M=24.30$)を使って、80名の被験者を2群に分けた。25点以上の34名を上位群として、24点以下の46名を下位群とした。これを学習者の語彙力として、一つの要因(独立変数)とした。

4.2 決定木分析（分類木分析）の結果

中国人日本語学習者の日中間の漢字語のネットワークを明らかにするため、以上で説明した（１）日本語での使用頻度の高低、（２）中国語での存在有無、（３）日中音韻的類似性の高低、（４）学習者による日本語の語彙力、の４つの要因（独立変数）で漢字の書き取りの正誤（従属変数）を予測するための決定木（特に、分類木）分析を行った。分類木分析は質的データを扱う予測手法で、複数の独立変数のなかから従属変数を有意に予測した変数を選び、予測力の強い順に、交互作用を含み込んで階層的に分析する手法である。この解析法を使うと、予測結果が樹形図で描かれる。樹形図に描かれない要因は、従属変数を有意に予測しない変数と判断される。この手法は、複数の要因を階層的に可視化することができるので、これまでに、第二言語習得における背景諸要因の検討や語用論に関する研究（Tamaoka, et al., 2010 ; 張婧禕, 2017 など）に利用されてきた。そこで、本研究で想定したこの４つの影響要因はすべて図１の樹形図に階層的に描かれているので、漢字語の書き取りの正誤を有意に予測する変数であることが分かる。

まず、漢字語の書き取りの正誤を最も強く予測したのは「日本語での使用頻度の高低」で、樹形図の最上部に描かれた。漢字語の書き取りの正誤の親ノード（ノード０）から子ノードの「高使用頻度」（ノード１）と「低使用頻度」（ノード２）に枝が分かれた $[\chi^2(1) = 304.88, p < .001]$ 。「高使用頻度」の漢字語（ $M=72.4\%$ ）は「低使用頻度」の漢字語（ $M=44.7\%$ ）よりも正答率が有意に高かった。日本語における頻度の高さが、漢字語の書き取りに最も影響することが分かった。「日本語での使用頻度の高低」の次に、「中国語での存在の有無」は２番目に強い影響要因となった。さらに、日本語での高低使用頻度別でノード３と４[高使用頻度： $\chi^2(1) = 103.02, p < .001$]およびノード５と６[低使用頻度： $\chi^2(1) = 155.35, p < .001$]に枝が分かれた。使用頻度の高い漢字語の場合、予想通り、ノード３の「中国語に存在する」漢字語（ $M=79.8\%$ ）はノード４の「中国語に存在しない」漢字語（ $M=57.8\%$ ）より有意に正しく書けることが示された。一方、日本語で使用頻度の低い漢字語の場合は予想に反して、ノード６の「中国語に存在する」の漢字語（ $M=34.7\%$ ）と比べ、ノード５の「中国語に存在しない」の漢字語（ $M=64.7\%$ ）のほうが有意に正しく書かれた。ここで、ひとまず第２の要因の影響からみると、中国人日本語学習者にとって、日本語での使用頻度は最も強い影響要因であるが、

ターゲット漢字の書き取りの正誤

ノード0	%	N
カテゴリ	41.40	1,591
誤答	58.60	2,249
正答	100.00	3,840
合計		

日本語での使用頻度の高低
 $\chi^2(1) = 304.88, p < .001$.

高使用頻度

ノード1	%	n
カテゴリ	27.60	529
誤答	72.40	1,391
正答	50.00	1,920
合計		

中国語での存在の有無
 $\chi^2(1) = 103.02, p < .001$.

中国語で存在有

ノード3	%	n
カテゴリ	20.20	259
誤答	79.80	1,021
正答	33.30	1,280
合計		

音韻的類似性の高低
 $\chi^2(1) = 47.44, p < .001$.

高音韻的類似性

ノード7	%	n
カテゴリ	12.50	80
誤答	87.50	560
正答	16.70	640
合計		

低音韻的類似性

ノード8	%	n
カテゴリ	28.00	179
誤答	72.00	461
正答	16.70	640
合計		

上位群

ノード9	%	n
カテゴリ	35.30	96
誤答	64.70	176
正答	7.10	272
合計		

下位群

ノード10	%	n
カテゴリ	47.30	174
誤答	52.70	194
正答	9.60	368
合計		

中国語で存在無

ノード4	%	n
カテゴリ	42.20	270
誤答	57.80	370
正答	16.70	640
合計		

日本語学習者の語彙力
 $\chi^2(1) = 9.22, p < .001$.

高音韻的類似性

ノード11	%	n
カテゴリ	28.40	91
誤答	71.60	229
正答	8.30	320
合計		

低音韻的類似性

ノード12	%	n
カテゴリ	42.20	135
誤答	57.80	185
正答	8.30	320
合計		

高音韻的類似性

ノード13	%	n
カテゴリ	59.40	323
誤答	40.60	221
正答	14.20	544
合計		

低音韻的類似性

ノード14	%	n
カテゴリ	69.70	513
誤答	30.30	223
正答	19.20	736
合計		

低使用頻度

ノード2	%	n
カテゴリ	55.30	1,062
誤答	44.70	888
正答	50.00	1,920
合計		

中国語での存在の有無
 $\chi^2(1) = 155.35, p < .001$.

中国語で存在無

ノード5	%	n
カテゴリ	35.30	226
誤答	64.70	414
正答	16.70	640
合計		

音韻的類似性の高低
 $\chi^2(1) = 13.24, p < .001$.

高音韻的類似性

ノード6	%	n
カテゴリ	65.30	836
誤答	34.70	444
正答	33.30	1,280
合計		

日本語学習者の語彙力
 $\chi^2(1) = 14.72, p < .001$.

上位群

ノード13	%	n
カテゴリ	59.40	323
誤答	40.60	221
正答	14.20	544
合計		

下位群

ノード14	%	n
カテゴリ	69.70	513
誤答	30.30	223
正答	19.20	736
合計		

図1 漢字の書き取りの正誤を4つの要因で予測する分類木分析の結果

唯一の要因とはならず、中国語での存在の有無も次に強く影響した。そのため、正答率の高い順に並べると、高使用頻度で中国語に存在する語が最も正しく書かれた。また、日本語での低使用頻度で中国語に存在しない語と日本語での高使用頻度で中国語に存在しない語が中間に位置し、日本語での低使用頻度で中国語に存在する語が最も書けないという結果になり、第2の要因での正答率の特徴がみられる。

さらに、第3の要因は、日中両言語の音韻的類似性と中国人日本語学習者の語彙力の2つが互いに絡み合って漢字の書き取りの正誤を有意に予測していた。日本語での高使用頻度で中国語に存在する漢字語 $\chi^2(1) = 47.44, p < .001$ および日本語での低使用頻度で中国語に存在しない漢字語 $\chi^2(1) = 13.24, p < .001$ は、日中の音韻的類似性が有意な要因となり、さらにノード3から「高音韻的類似性」のノード7 ($M=87.5\%$) と「低音韻的類似性」のノード8 ($M=72\%$)、ノード5から「高音韻的類似性」のノード11 ($M=71.6\%$) と「低音韻的類似性」のノード12 ($M=57.8\%$) に枝が分かれた。平均正答率からみれば、日中両言語の音韻的類似性が高いほど学習者による漢字語の書き取りの正答率が高かった。しかし、高使用頻度で中国語に存在しない漢字語 $\chi^2(1) = 9.22, p < .001$ および低使用頻度で中国語に存在する漢字語 $\chi^2(1) = 14.72, p < .001$ では、学習者の語彙力が有意な要因であり、ノード4から「上位群」のノード9 ($M=64.7\%$) と「下位群」のノード10 ($M=52.7\%$)、ノード6から「上位群」のノード13 ($M=40.6\%$) と「下位群」のノード14 ($M=30.3\%$) に分かれた。語彙力で下位群より上位群の学習者のほうが正しく漢字語を書き取ることができた。第3の要因において、日本語の語彙力と日中音韻的類似性の2要因は漢字の書き取りに影響するものの、樹形図から分かるように、これらは最も重要な要因ではない。

5 総合考察

本研究で使用したのは日本語の漢字表記語の書き取りテストである。そのため、中国語の母語で漢字を使用しているので、中国人日本語学習者にとっては簡単なテストだと思われる。ところが、中国人日本語学習者の全体の平均正答率は58.60%となり、予想以上に低かった。これは、日本語の漢字語に含まれた漢字そのものが書けないというより、中国人学習者にとって、2つの漢字から構成され

る L2 日本語の漢字語の発音と書字とが対応して学習されていないことが原因と考えられる。言い換えれば、L2 日本語の漢字語の発音がメンタルレキシコンに音韻表象として記憶されていない（学習されていない）ために、L1 中国語と L2 日本語で漢字を共有していても（書字表象群が L1 と L2 で対応していても）、L2 日本語での漢字語の音韻表象と書字表象の結合が確立されていないことが原因であろう。本研究では、4つの変数を想定して、L2 日本語の漢字語の書き取りの正確さを、分類木分析を使って検証した。その結果は以下の3点にまとめられよう。

第1に、ターゲット漢字を含む漢字語の使用頻度が第1要因となり、漢字語の処理に最も強く影響した。高使用頻度の語による平均正答率（72.40%）は低使用頻度（44.70%）より 27.70%高かった。L2 日本語の漢字語の書字（漢字の書き取り）においても「語彙使用頻度効果（word frequency effect）」が顕著にみられることが示された。実際、個々の語の表象には特定の閾値（threshold）があり、その閾値が高いほど、個々の語の表象が活性化され、より迅速に語彙処理できると指摘されている（玉岡, 2013）。L2 日本語での使用頻度が高いことは、中国人日本語学習者が頻繁にこれらの語彙に接していることを意味し、より頻繁に目にするほど迅速に知覚されることが検証された。

第2に、第2要因として強く影響したのは日本語の漢字語が L1 中国語に存在するかどうかであった。日本語の語彙使用頻度の高低別で、中国語での存在有無によって学習者の漢字語の習得が異なっていた。L1 中国語と L2 日本語で共通の漢字語は、L1 中国語の存在により L2 日本語の漢字語がより確実に学習されるので、発音から漢字を書く場合にも、より正しく書けることになるようである。とりわけ、日中両言語で同じ漢字を共有するために、両言語間の書字表象がより強く結合し、L2 日本語で漢字の書き取りにも強く影響したと考えられる。これは、2つの言語間で文字表記を共有する場合の処理の促進効果による「表記親近性効果（script familiarity effect）」（石田, 1986; 小森, 2005; 玉岡, 1994, 1997, 2000; 大和・玉岡, 2013; 張婧禕, 2017; Djojomihardjo, et al., 1994 など）であると考えられる。しかし、L2 日本語で使用頻度が低い漢字語では、表記親近性効果が弱くなり、ほかの手掛かりを使って適切な漢字を見つけだそうとするようである。

第3に、日中音韻の類似性および学習者による語彙力は第3要因として、日本語での使用頻度と中国語での語彙の存在有無の2条件と絡み合って影響した。ま

ず、高使用頻度で中国語に存在する語と低使用頻度で中国語に存在しない語は音韻的類似性からの影響を受けており、全員による平均正答率は72.23%であった。そのうち、最も正確に漢字が書けたのは、音韻的類似性が高く、L1 中国語に存在する高使用頻度の語(87.50%)である。つまり、書字と音韻の両方で両言語の語彙特性を共有している語は極めて正確に想起され、正答率も最も高くなった。しかし、漢字語をあまり見かけず、日本語に定着していない語(71.6%)に対しても、L1 中国語の音韻を検索して漢字を見つけようとする方略が機能するようである。なお、日本語では一つの漢字の音韻表象に対して複数の漢字単位の書字表象群が対応し結合しているので、低使用頻度の場合、逆に日中の音韻類似性が同音の異なる漢字を想起させて、誤った漢字を書くこともあった。

本研究の学習者による誤答をみると、3つの誤答パターンに分けられる。(1) 日本語の音節はもともと少なく、同音語になる可能性が高い。そのため、同音語または同音字を間違って書く。たとえば、「簡便」を「勘弁」、「試用」を「使用」、「誤報」を「語法」または「後報」である。(2) 繁体字を簡体字で誤答した。これは、L1 中国語とL2 日本語の漢字が完全に区別されていないことを示している。(3) まったく書けない。これら3つの誤用パターンのうち、(1)のパターンによる誤りが圧倒的に多かった。つまり、L2 日本語の漢字語を、L1 中国語の類似した発音の漢字で書いてしまうことである。

一方、学習者による日本語の語彙力はL2 日本語の高使用頻度でL1 中国語に存在しない語とL2 日本語の低使用頻度でL1 中国語に存在する語に影響し、学習者の日本語の語彙力が高いほど、正答率が高くなり、日本語の漢字語の語彙処理を促進する傾向がみられた。さらに、語彙力が上位群の学習者は、L2 日本語の高使用頻度でL1 中国語に存在しない語(64.7%)はL2 日本語の低使用頻度でL1 中国語に存在する語(40.6%)よりも正答率が高かった。つまり、日本語の語彙力は、日本語での語彙使用頻度が高い場合によく機能することが分かる。これは、L1 中国語の漢字語の知識がすぐに援用できないために、既習のL2 日本語の漢字語の知識が漢字の書き取りの決め手となるからであろう。

以上のように、中国人日本語学習者は、日中の漢字表記を共有しているので、L2 日本語の漢字語の漢字の学習では、日本語で頻繁にみられる漢字語をL1 中国語と照らして、日中のメンタルレキシコンにおける漢字の書字表象群のネットワ

ークをまず確認して、さらに弱いながらも音韻表象群の類似性を探り、L2 日本語で学習した語彙力を援用するというアプローチを採っていることが確認できた。

参考文献

- 石田敏子 (1986) 「英語・中国語・韓国語圏別日本語学力の分析」『日本語教育』 58, 163-194.
- 木山幸子・玉岡賀津雄・趙萍 (2011) 「外国語としての日本語 (JFL) の語用論的能力に関わる基礎的言語知識—中国語を母語とする日本語学習者を例に—」『言語教育評価研究』 2, 2-14.
- 邱学瑾 (2002a) 「台湾人日本語学習者における日本語漢字熟語の処理過程—一日・中 2 言語間の同根語と非同根語の比較—」『広島大学大学院教育学研究科紀要第二部』 51, 357-365.
- 邱学瑾 (2002b) 「漢字圏・非漢字圏日本語学習者における漢字熟語の処理過程—意味判断課題を用いた形態・音韻処理の検討—」『教育心理学研究』 50, 412-420.
- 国際交流基金・(財) 日本国際教育支援協会 (2002) 『日本語能力試験出題基準 (改訂版)』, 凡人社.
- 小森和子 (2005) 「第二言語としての日本語の文章理解における第一言語の単語認知処理方略の転移—視覚入力と聴覚入力の相違を中心に—」『横浜国立大学留学生センター紀要』 12, 17-39.
- 斉藤信浩・玉岡賀津雄・母育新 (2012) 「中国人日本語学習者の文章および文レベルの理解における語彙と文法能力の影響」『ことばの科学』 25, 5-20.
- 玉岡賀津雄 (1994) 「仮名と漢字による語彙処理のメカニズム—日本語学習者の学習歴と言語背景による影響—」『松山大学総合研究所所報』 15, 1-101.
- 玉岡賀津雄 (1997) 「中国語と英語を母語とする日本語学習者の漢字および仮名表記語彙の処理方略」『言語文化研究』 17, 65-77.
- 玉岡賀津雄 (2000) 「中国語系および英語系日本語学習者の母語の表記形態が日本語の音韻処理に及ぼす影響」『読書科学』 44, 83-94.
- 玉岡賀津雄・宮岡弥生・松下達彦 (2002) 「日本語学習者の心的辞書 (mental lexicon) の構造—中国語を母語とする超上級日本語学習者の漢字熟語の処理を例に

- 一」『平成 14 年度 日本語教育学会・中国地区研究集会予稿集』(pp.1-8), 日本語教育学会.
- 玉岡賀津雄 (2013) 「メンタルレキシコンと語彙処理—レフェルトの WEAVER++ モデル—」『レキシコンフォーラム』 6, 327-345.
- 玉岡賀津雄 (2017, 印刷中) 「実験的手法を用いた語彙習得研究」『第二言語としての日本語の習得研究』 20.
- 張婧禱・玉岡賀津雄 (2017) 「中国人日本語学習者による NV 型複合名詞の理解」『小出記念日本語教育研究会』 25, 35-50.
- 張婧禱 (2017) 「中国人日本語学習者の漢語同形語習得—同形類義語 (Overlap 語) を中心に—」『愛知工業大学研究報告』 52,6-13.
- 松下達彦・Marcus Taft・玉岡賀津雄 (2004) 「中国語『単語』を知っていることは日本語漢字語の発音学習に役立つか」『中国語・日本語学論文集 (平井勝利教授退官記念)』 578-590.
- 宮岡弥生・玉岡賀津雄・林炫情・池映任 (2009) 「韓国語を母語とする日本語学習者による漢字の書き取りに関する研究—学習者の語彙力と漢字が含まれる単語の使用頻度の影響—」『日本語科学』 25, 119-130.
- 宮岡弥生・玉岡賀津雄・酒井弘 (2011) 「日本語語彙テストの開発と信頼性—中国語を母語とする日本語学習者のデータによるテスト評価—」『広島経済大学研究論集』 34(1). 1-18.
- 大和祐子・玉岡賀津雄 (2013) 「中国語母語話者と韓国語母語話者の日本語テキストの読み処理における言語的類似性の影響」『小出記念日本語教育研究会論文集』 21, 61-73.
- Aitchison, J. (1987). *Words in the mind: An introduction to the mental lexicon* (1st edition 1987). Oxford and New York: Basil Blackwell.
- Djojomihardjo, M., Koda, K., & Moates, D. R. (1994). Development of L2 word recognition. In Q. Jing, H. Zhang & D. Peng (Eds.) *Information processing of Chinese language* (pp. 153-161), Beijing: Beijing Normal University Publishing.
- Tamaoka, K., Lim, H., Miyaoka, Y., & Kiyama, S. (2010). Effects of gender-identity and gender-congruence on levels of politeness among young Japanese and Koreans. *Journal of Asian Pacific Communication*, 20, 23-45.