コーパスを用いた日本語学習者に誤発音されうる語に関する研究

大井川　朋　彦[*1]

A Corpus-Based Study on Words Which Can Be Mispronounced by Learners of Japanese

Tomohiko OOIGAWA[*1]

In order to contribute to the development of Japanese language education, the present research examined words which can be mispronounced by learners of Japanese and classified the error patterns using the International Corpus of Japanese as a Second Language (Sakoda, 2020). The study analyzed about-30-minute interview data from each of the 974 participants from twelve different L1 groups. The examples of the mispronounced nouns and the patterns were *chuugoku* 'China' (/čuHgo/ for /čuHgo**ku**/: /ku/-deletion) for the Chinese group, *tomodachi* 'friend' (/tomoda**Q**či/ for /tomodači/: /Q/-insertion) for the English group, *paatii* 'party' (/paH**či**H/ for /paH**ti**H/: /ti/-substitution to /či/) for the French group, *sensei* 'teacher' (/seNse/ for /seNse**H**/: /H/-deletion) for the German group, *ryokou* 'travel' (/ryoko/ for /ryoko**H**/: /H/-deletion) for the Hungarian group, *koukou* 'high school' (/koko/ for /ko**H**ko**H**/: /H/-deletion) for the Indonesian group, *nihon* 'Japan' (/nioN/ for /ni**h**oN/: /h/-deletion) for the Korean group, *eiga* 'movie' (/eHg**o**/ for /eHg**a**/: /a/-substitution to /o/) for the Russian group, *kazoku* 'family' (/ka**s**oku/ for /ka**z**oku/: /z/-substitution to /s/) for the Spanish group, *kodomo* 'child' (/ko**r**omo/ for /ko**d**omo/: /d/-substitution to /r/) for the Thai and Vietnamese groups, *sensei* 'teacher' (/s**a**NseH/ for /s**e**NseH/: /e/-substitution to /a/) for the Turkish group. The findings also provide opportunities to reexamine the phonological system of each L1.

1．導入

1.1　目的と研究課題

様々な背景を持つ外国人が日本の国内外で日本語を学んでいる。日本語学習者の抱える問題を明らかにすることは日本語教育発展の一助となりうると考える。そのため本研究では、コーパスを用いて、国内外に住む日本語学習者に比較的多く使用され且つ誤発音されうる語を検討し、その誤発音パターン及びその要因を考察した。著者の知る限り、日本語学習者の発音に関する研究は多く存在するが、「語」に注目した研究は少ない。学習者が誤って発音しやすい分節音や超分節的特徴を検討することも有意義であるが、実際の教材などで研究成果を活用する場合、具体的な語である方が導入しやすいと考えられる。例えば、本稿で示しているデータでも分かるように、ベトナム語やタイ語を母語とする日本語学習者は「子供」を「ころも」と発音することがある。日本語の/d/が/r/に置換される現象であると解釈できるが、その情報だけでは、日本語の教科書などを作成する際に、その例文に学習者が使用しうる実用的な

＊1　日本大学国際関係学部国際教養学科　助教　Assistant Professor, Department of International Liberal Arts, College of International Relations, Nihon University

語である「子供」を採用するか否かを検討することはないかもしれない。日本語の/d/が/r/に置換されうるという情報に加え、特に自然会話においては「子供」という語でその現象が現れうる、という情報の方が、教科書の例文などを考えるときに大いに役立つと考える。つまり、発音指導の観点から「子供」という語を用いた例文を使用したり、反対に誤発音を避けるために使用しない、などの応用例が考えられる。加えて、近年では、母語話者側の教育という視点も存在する(Derwing & Munro, 2015, pp.145-147)。つまり、聞き取り指導などにより母語話者の方が学習者の発音の傾向に慣れることにより、聞き取り能力が向上したり、自信がつくという効果が得られるのではないかということである。これにより、母語話者と学習者間でより円滑な意思疎通がなされると考えられる。例えば、日本に住み、日常的に日本語を仕事などで使用する外国人と日本人との会話がよりスムーズになることが想像できる。そういった聞き取り指導においても、/d/が/r/に置換されるという情報のみより、特に「子供」が「ころも」と発音されうるという具体的な情報も含めて提示した方が、より役立つと考えられる。ここまで、学習者側と母語話者側の日本語教育について述べたが、本研究の成果はその両者に貢献しうると考える。

誤発音は実際の教育現場や職場などで生じていると思われるが、その大半は記録されることがないと想像できる。コーパスを用いることにより誤発音データを計量的に分析できるという意義も大きい。さらに、本研究の研究成果は音韻研究全般の一助になる可能性もある。本研究では母語話者も含め13言語の母語の協力者群のデータを分析した。そのため、学習者の母語の音韻体系の再確認の機会となったり、日本語母語話者の新しい発音の傾向に関するデータを提供する可能性もある。さらに、比較的新しいコーパスを使用するという方法により、今まで注目されなかった新しい現象などが提供される可能性も想定される。

本研究の研究課題は「日本語学習者に比較的多く使用され且つ誤発音されうる語は何か」、「その誤発音のパターンや要因は何か」、「どの程度の割合で誤発音されるのか」である。実際の教育現場では様々な背景の学習者が存在することが想像できるため、本研究ではできる限り多くの母語話者群のデータを分析対象とした。

本稿は第1節「導入」にて本研究の目的、研究課題、使用したコーパスや先行研究の紹介を行い、第2節「分析」では具体的な分析方法の説明とその結果の報告を行う。第3節「考察」では結果を考察し、第4節「結論」では、研究課題を振り返りつつ、結論を示し、第5節「今後の課題」では、今後取り組むべき研究に関して言及する。

1.2 コーパス

本研究において使用したコーパスは『多言語母語の日本語学習者横断コーパス：I-JAS（データバージョン2022.05)』（迫田，2020，迫田ら，2020）である。同コーパスには母語話者を含め国内外の22の母語の日本語学習者群のデータが収録されている（表1参照)。そのうち、計量的に分析するにあたり十分な人数であると考えられる50名以上の群を分析対象とした。

同コーパスのデータの収集方法（タスク）は「ストーリーテリング」、「対話」、「ロールプレイ」、「絵描写」、「ストーリーライティング」に分かれており、最も自然会話に近く、語彙の偏りが少ないと考えられる対話タスクのデータを分析対象とした。対話タスクは30分程度の会話を調査者と研究協力者（学習者）が行うというものであった。

尚、データは第一次から第五次データまであり、2020年3月に第五次データまでが公開された。コーパスを構築するには、音声データ収集後から文字起こしなどを経て、公開に至ることを考慮すると、本研究で使用するものは、新型コロナウイルス感染症（COVID-19）の本格的な流行前のものであると考えられる。流行前と後では使用される語彙に違いが出る可能性もあるため、本研究の結果は将来的にコロナ後のデータと比較することを視野に入れている。

表１：協力者の人数

	母語	海外	国内	合計
1	中国語	200	73	273
2	韓国語	100	11	111
3	英語	100	7	107
4	ベトナム語	50	16	66
5	スペイン語	50	7	57
6	ロシア語	50	4	54
7	タイ語	50	3	53
8	インドネシア語	50	2	52
9	フランス語	50	1	51
10	ドイツ語	50	0	50
11	トルコ語	50	0	50
12	ハンガリー語	50	0	50
13	ポルトガル語	0	12	12
14	タガログ語	0	3	3
15	ビサヤ語	0	3	3
16	モンゴル語	0	3	3
17	クメール語	0	1	1
18	ネパール語	0	1	1
19	ヒンディー語	0	1	1
20	フィリピン語	0	1	1
21	ペルシア語	0	1	1
22	日本語	0	50	50

1.3　先行研究

同コーパスを使用し、学習者の発音に注目した先行研究としては、平田（2019，2020）及び大井川（2022）が挙げられる。平田（2019）では中国語、平田（2000）では中国語及び韓国語を母語とする学習者の促音に着目し、「ピクニック」がテーマのストーリーテリング１のデータの「ピクニック」「サンドイッチ」「バスケット」の３語の分析が行われた。大井川（2022）では国内学習者の「対話」タスクのデータに限定し、特に協力者数の多い中国語、ベトナム語、韓国語、スペイン語、ポルトガル語を母語とする学習者の言い間違い（誤発音と同義）を分析した。その結果、共通して「ヒト」（「人」または「一」）に誤発音が認められ、韓国語母語話者群以外では「中国」、「ベトナム」、「ペルー」、「サンパウロ」に見られるように、その母語との関係が示唆される国名や地名の誤発音が確認された。本研究では、大井川（2022）の研究手法を踏襲しつつ、さらに分析対象を広げ、母語話者との比較を行いつつ、国内外の12の母語群の学習者の対話タスクの全てのデータを分析対象とした。

2．分析

2.1　分析方法

『多言語母語の日本語学習者横断コーパス：I-JAS』（迫田，2020，迫田ら，2020及び『書き起こし規程集』参照）の「対話」のデータにおいて、「Gタグ」が付与された名詞を分析対象とし、その中から誤発音されたと考えられるものを分析した。1.2で示したように、50名以上の12の母語話者群（中国語、韓国語、英語、ベトナム語、スペイン語、ロシア語、タイ語、インドネシア語、フランス語、ドイツ語、トルコ語、ハンガリー語）及び日本語母語話者群の「対話」タスクのデータを使用した。

対話タスクは予め設定された「日本について」などの質問項目に関して質問するインタビューであり、それを元に30分程度の会話を調査者と研究協力者（学習者）が行うというものであった。質問項目の詳細については、迫田ら（2020，pp.37-38）を参照されたい。その他のタスクは、登場人物名やキーワードが書かれているが、吹き出しはなく漫画のようにコマ割りされたイラストを見ながら行う「ストーリーテリング」、背景や役割が明確な「ロールプレイ」、様々な人物や物が１つの絵に描かれているイラストを見ながら行う「絵描写」、書くタスクの「ストーリーライティング」であった。対話タスクは最も自然会話に近く、特定の語が不自然に繰り返し使用されたり、特定の人物名などが連呼されるなどの分析に影響する現象が起きづらいと考え、このタスクのデータを分析対象とした。

録音された音声データは書き起こしの研修の受講と練習期間を経た日本語教育経験者や日本語教育を学んだことのある大学生・大学院生により書き起こされた（迫田ら，2020，pp.52-53）。同コーパスは本来音声学や音韻論の研究のためにデザインされたものではないため、音声記号などは使用されていない。しかしながら、仮名を用いた可能

な限り発音に忠実な書き起こしがなされている。また、同コーパスでは、誤用の認定は非常に困難であり、研究内容によってもその判定基準が異なるためという理由で、誤用および正用の表示は行われていない。しかしながら、より正確な自動形態素解析を行うために様々なタグが付与されている。そのうちGタグは語や活用や発音の誤りに付与されるものであり、「基本的に、発音や活用の間違いで、形態素解析がうまくできないために、タグをつけるものである。そのため、日本語として解析できるものには付与しない方針である。」（『書き起こし規程集』p.12）とされる。そこで、本研究ではGタグを全ての誤発音や不自然な発音を示すものとは捉えないが、日本語母語話者が日本語として解析できない可能性のある重要な誤発音を含みうるものを示すものとして使用することとした。

表２はGタグが付与された語の例を示している。国立国語研究所で開発されたコーパス検索アプリケーション『中納言』（2.7.0）を用いてI-JASのデータを検索することができる。例えば、トルコ語母語話者の「対話」のみを選択し、「書字形出現形」の「先生」という語を短単位検索し、データをダウンロードし、整理してその結果の一部を表示すると表２のようになる。各トークンに関して、「キー（解析向けに加工した文字列）」には自動形態素解析向けにどのように加工されたかが示されており、「発音形出現形」には模範的な発音がカタカナで表記されており、「キー（発話通りの文字列）」には実際の発音がGタグとともに記録されていることが分かる。つまり、「先生」という語を「さんせい」と発音した場合と「先生」（せんせい）と発音した場合が示されている。

表２：Gタグ付与例（一部）

キー（解析向けに加工した文字列）	発音形出現形	キー（発話通りの文字列）
先生	センセー	[さんせい=G]
先生	センセー	[さんせい=G]
先生	センセー	[さんせい=G]
先生	センセー	先生
先生	センセー	先生
先生	センセー	先生

また、中納言では特定のタグが付与された語のみを検索することも可能である。表３は学習者の12群と母語話者群のデータのうち、Gタグが付与された語のトークン数（延べ語数）を品詞別に分けた結果の一部を示している。Gタグが最も多く付与された品詞が名詞であったため、本研究では名詞に着目する。加えて、本研究は誤発音に注目しているため、そして、Gタグは語や活用や発音の誤りに付与されるものであるため、名詞に限定することにより活用の誤りのデータを除外することができるとも考えた。

表３：品詞別Gタグ付与延べ語数（一部）

品詞	延べ語数
名詞	27,741
動詞	8,362
副詞	2,893
形容詞	2,754

表４は本研究の分析対象をまとめたものである。I-JASの12の母語話者群と日本語母語話者群の対話タスクのデータのうち、Gタグが付与された名詞に着目した。

表４：分析対象のまとめ

コーパス	I-JAS
協力者群	50名以上の13群
タスク	対話
タグ	G
品詞	名詞

2.2　Gタグの総数

表５は各群のGタグが付与された名詞の延べ語数（Gタグ数）を示している。群により協力者数が異なるため、１人あたりのGタグ数（Gタグ数/協力者数）も算出した。ベトナム語群が最も多く、母語話者群が最も少なく、学習者群の中ではハンガリー語群が最も１人あたりのGタグ数が少ないことが分かる。

表５：各群の名詞のＧタグ付与数

母語	Gタグ数	協力者数	1人あたり
ベトナム語	2,527	66	38.3
タイ語	1,981	53	37.4
インドネシア語	1,788	52	34.4
スペイン語	1,717	57	30.1
韓国語	3,251	111	29.3
ドイツ語	1,373	50	27.5
フランス語	1,395	51	27.4
ロシア語	1,396	54	25.9
中国語	7,012	273	25.7
トルコ語	1,181	50	23.6
英語	2,360	107	22.1
ハンガリー語	836	50	16.7
日本語	243	50	4.9

2.3　Ｇタグが付与された語

表６は各群における、Ｇタグが付与された名詞のうち最も多く付与された上位第５位までの語を示している。「/」はＧタグ数が同数であることを示す。

母語が異なっても共通して見られる語が存在する。第１位のものでは、ベトナム語とタイ語群の「子供」、ドイツ語とトルコ語群の「先生」が共通している。その他、数詞や「高校」、「学校」、「時間」、「友達」、「全部」などの語が複数の母語話者群で見られる。国内の学習者のデータに限定して分析した大井川（2020）の結果とも共通している傾向として、「一」や「人」やその母語との関係が示唆される国名や地名も見られる。

次節からは表６の一覧のうち、太字で下線が引かれている名詞を特に分析対象語として着目する。尚、この一覧において「一」、「四」、「十」といった数詞も多くそして共通して見られるが、本研究の分析対象からは除外した。数詞は「一つ（ひとつ）」、「一回（いっかい）」、「一度（いちど）」のように後続する語により模範的な発音が異なるため、発音が一通りの他の名詞と同様に１つの名詞として数え、比較することが妥当ではないと判断したことに加え、得られたデータが誤発音なのか、語の読み間違えなのかの判断が難しいと想定されたからである。数詞に関しては詳しく分析するために別の研究機会を設けたいと考えている。

表６：各群のＧタグ付与名詞の一覧

母語	1位	2位	3位	4位	5位
ベトナム語	**子供**	十	勉強	四	ホーチミン/生活
タイ語	**子供**	時間	高校/授業	一	こと
インドネシア語	**高校**	人	景色/友達	四/伝統	誕生
スペイン語	十	**家族**	全部	プレゼント	四
韓国語	**日本**	学校	一	全部	高校
ドイツ語	**先生**	十	時間	オーストリア	ウィーン/人
フランス語	**パーティー**	十	一	人	他
ロシア語	**映画**	パーティー	四	英語/学校	公園/一/高校
中国語	一	十	**中国**	勉強	時間
トルコ語	**先生**	ツーリスト	イスタンブール	一	ケーキ/人
英語	**友達**	十/先生	一	人/勉強	四
ハンガリー語	**旅行**	ワイン	ストーリー	十	トルコ
日本語	**本当**	もの	あと/物	一/一応/一日/感じ	子供

2.4　各群の分析対象語の誤発音

表７から表19に各群の分析対象語の誤発音分析の結果をまとめた。表の左上の「」内には分析対象語を、その下には具体的な発音を、「延べ語数」はその発音が何回行われたかを、「人数」はその発音を何人の協力者が行ったかを示した。具体的な発音は表２（2.1）で示した「発話通りの文字列」に基づく。尚、人数が１のものは、偶然の産物やその協力者特有のものである可能性を考慮して、「その他」とした。

誤発音であると考えられる最も延べ語数と人数が多いパターンは、中国語群では「中国」を「ちゅうご」、英語群では「友達」を「ともだっち」、フランス語群では「パーティー」を「パーチー」、ドイツ語群では「先生」を「せんせ」、ハ

ンガリー語群では「旅行」を「りょこ」、インドネシア語群では「高校（生）」を「ここ（せい）」、韓国語群では「日本」を「におん」、ロシア語群では「映画」を「えいご」、スペイン語群では「家族」を「かそく」、タイ語とベトナム語群では「子供」を「ころも」、トルコ語群では「先生」を「さんせい」であった。尚、日本語母語話者の「本当」の「ほんと」という発音に関しては誤発音として捉えるべきなのか新しい発音の傾向として考えるべきなのかは議論の余地がある。

今回の検索方法において「中国」、「高校」、「日本」という語では「中国人」や「中国語」、「高校生」、「日本語」などの語も含まれるため、それぞれにおいて、「ちゅうごじん」、「ここせい」、「におんご」といった誤発音と考えられる例が見られる。インドネシア語群では実際には「ここせい」の例が最多であったが、実質的には同じ現象と考えられるため、「高校」が「ここ」と発音されたものと同等と解釈した。

表７：中国語群の誤発音分析

「中国」	延べ語数	人数
ちゅうご	16	11
ちゅごく	7	3
ちゅう	3	2
ちゅうく	3	3
ちゅうごう	3	3
ちゅうごじん	3	2
ちゅうごっく	2	2
ちゅうおう	2	2
ちゅうごーご	2	2
その他	49	
合計	90	

中国語群の「中国」、インドネシア語群の「高校」、韓国語群の「日本」は多様な結果となった一方、他の群の分析対象語のその他以外の発音は、多くても３種類であった。中国語群に関しては、協力者の人数が多いことも影響していると考えられるが、半数以上がその他（49/90）、つまり１名のみが行った発音であったという結果であった。

表８：英語群の誤発音分析

「友達」	延べ語数	人数
ともだっち	14	10
ともらち	9	5
ともらっち	6	4
その他	11	
合計	40	

表９：フランス語群の誤発音分析

「パーティー」	延べ語数	人数
パーチー	22	13
パーチ	6	2
パチ	3	2
その他	4	
合計	35	

表10：ドイツ語群の誤発音分析

「先生」	延べ語数	人数
せんせ	29	7
その他	4	
合計	33	

表11：ハンガリー語群の誤発音分析

「旅行」	延べ語数	人数
りょこ	8	6
りょうこう	2	2
その他	8	
合計	18	

表12：インドネシア語群の誤発音分析

「高校」	延べ語数	人数
ここせい	25	12
ここ	17	9
ここせー	6	3
ここー	5	4
こっこ	4	2
ここせ	3	3
こうこ	2	2
こっこせい	2	2
その他	29	
合計	93	

表13：韓国語群の誤発音分析

「日本」	延べ語数	人数
におん	40	19
におんご	14	8
でぃおん	12	4
でぃおんご	4	2
でぃおんー	3	2
におんごー	3	2
でぃほんご	3	2
ぎおん	2	2
におんー	2	2
によん	2	2
その他	19	
合計	104	

表14：ロシア語群の誤発音分析

「映画」	延べ語数	人数
えいご	18	8
えんが	3	2
その他	8	
合計	29	

表15：スペイン語群の誤発音分析

「家族」	延べ語数	人数
かそく	34	16
かそくー	3	3
その他	7	
合計	44	

表16：タイ語群の誤発音分析

「子供」	延べ語数	人数
ころも	36	14
ほども	3	2
その他	6	
合計	45	

表17：ベトナム語群の誤発音分析

「子供」	延べ語数	人数
ころも	79	29
ころもー	9	5
その他	7	
合計	95	

表18：トルコ語群の誤発音分析

「先生」	延べ語数	人数
さんせい	75	21
せんせ	2	2
その他	2	
合計	79	

表19：日本語群の分析結果

「本当」	延べ語数	人数
ほんと	11	3
その他	4	
合計	15	

2.5　各群の誤発音の割合

表20は各群において分析対象語で最も多く見られた誤発音を何名の協力者が行ったかと、その延べ語数と、両者の割合を示したものである。人数の割合は各群の総協力者数を分母とし、延べ語数の割合は各分析語のGタグが付与されていないものも含めた総延べ語数を分母として算出した。

人数で比較すると中国語群（「中国」→「ちゅうご」）の11名（4.0%）が最も割合が低く、ベトナム語群（「子供」→「ころも」）の29名（43.9%）が最も割合が高い。延べ語数で見ると中国語群（「中国」→「ちゅうご」）の16（1.2%）が最も割合が低く、スペイン語群（「家族」→「かそく」）の34（35.1%）が最も割合が高い。尚、「*」で示したが、インドネシア語群では「ここ」ではなく「ここせい」のデータを使用した（詳細に関しては2.4を参照）。中国語群の誤発音の割合が低い結果となった理由としては、この分析対象語の発音

表20：誤発音の割合

母語	語	誤発音	人数	延べ語数
中国語	中国	ちゅうご	11 (4.0%)	16 (1.2%)
英語	友達	ともだっち	10 (9.3%)	14 (2.7%)
フランス語	パーティー	パーチー	13 (25.5%)	22 (31.0%)
ドイツ語	先生	せんせ	7 (14.0%)	29 (9.8%)
ハンガリー語	旅行	りょこ	6 (12.0%)	8 (14.5%)
インドネシア語	高校	ここせい*	12 (23.1%)	25 (10.0%)
韓国語	日本	におん	19 (17.1%)	40 (2.5%)
ロシア語	映画	えいご	8 (14.8%)	18 (20.9%)
スペイン語	家族	かそく	16 (28.1%)	34 (35.1%)
タイ語	子供	ころも	14 (26.4%)	36 (20.9%)
トルコ語	先生	さんせい	21 (42.0%)	75 (16.0%)
ベトナム語	子供	ころも	29 (43.9%)	79 (33.5%)

が他のものと比べて多様な結果であったことが挙げられる。

3. 考察

3.1 誤発音パターン

各群の分析対象語における最も多く見られた誤発音についてその要因を考察する（表21参照）。日本語の音素表記はVance（2008）を参考にして行った。

中国語群における「中国」/čuHgo**ku**/→/čuHgo/のパターンに関しては、Lin（2007）などを参照してみたが、中国語において/ku/が脱落するような例は見つけることはできなかった。他の理由として挙げることができるのは、中国語の語長の影響である。例えば、普通話では中国はZhōngguó（[tʂuŋ.kwo]）のように２音節語として発音される一方、「ちゅうごく/čuH.go.ku/」は３音節である。そのことが影響し、/ku/が脱落したと考えられる。

英語群において「友達」/tomodači/→/tomoda**Q**či/という促音化が見られた。英語でも日本語の促音のように長い子音を短い子音と区別することはあるが、それは語境界（例：why tie及びwhite tie）や形態素境界（例：unknown）に限定される（Ladefoged & Johnson, 2015, p.261）。日本語ほど体系的且つ頻繁に子音の長さの区別を行わないため、「ともだっち」という発音が出現すると考えられる。加えて、そもそも破擦音そのものが重子音と考えることもできるため、長めに発音される傾向がある可能性もある。

フランス語群において「パーティー」/paH**ti**H/→/paH**či**H/というパターンが見られた。フランス語では変種により詳細は異なるが、/t/などが後続音の影響により[tʃ]などの破擦音として実現しうる（Dawson et al., 2016, p.154, Pustka & Vordermayer, 2016, p.198）。したがって、ここで見られる子音の置換はその音韻規則の影響で起こったと考えられる。

ドイツ語、ハンガリー語、インドネシア語の３群では、それぞれ「先生」/seNse**H**/→/seNse/、「旅行」/ryoko**H**/→/ryoko/、「高校」/ko**H**ko**H**/→/koko/というパターンが見られ、共通して長母音の短母音化として分析できる。インドネシア語群では２箇所見られるが、３群ともに共通して語末の長母音が短母音化された。

ドイツ語には、Kohler（1990）を参照する限り、Betten [ɛ] とbäten [ɛː] のように短・長母音の区別が存在する。ハンガリー語においても、Szende（1994）を参照する限り、ok [o] とtó [oː] のように短・長母音の区別が存在する。よって、母語にこの区別がないということが要因とは言えない。

両者ともに共通して長母音の短母音化が２音節語の最後の音節で起きている。加えて、ハンガリー語では強勢は形態素の最初の音節に常に置かれるとされる（Szende, 1994）。もし、短・長母音の区別が強勢のある音節に限定されるのであれば、/ˈryoko/と発音される要因となりうる。ドイツ語でも短・長母音の区別が強勢のある音節に限定され、/ˈseNse/と強勢が置かれやすい理由が存在するのであれば、説明がつく。しかしながらこれらの推察を裏付ける文献やデータは現在のと

ころ見つかっていないため、更なる詳しい考察は今後の課題とする。

インドネシア語には短・長母音の区別が存在せず、二重母音は存在するものの、/ai/、/oi/、/au/の3種類のみで、/ou/が存在しない（Soderberg & Olson, 2008）。そのため、/koHkoH/の代わりに/koukou/として発音する可能性も低いと考えられる。

韓国語群における「日本」/ni**h**oN/→/nioN/のパターンに関しては、韓国語の母音間の/h/の異音規則が影響していると考えられる。/h/は後続母音により調音場所が変化し母音間で有声音化しうるとされる（Shin, Kiaer & Cha, 2013, pp.73-77）。この影響により/h/が脱落したかのように聞こえると考えられる。

ロシア語群における「映画」/eHg**a**/→/eHg**o**/のパターンに関しては、ロシア語の母音の弱化が影響している可能性がある。ロシア語では変種や音韻的条件により異なるが、強勢のない音節では/a/は[ʌ]や[ɐ]などとして実現しうる（Yanushevskaya & Bunčić, 2015）。よって、「えいが」の/ga/に強勢が置かれなかった場合、「えいご」のように聞こえる可能性があると考えられる。

スペイン語群における「家族」/ka**z**oku/→/ka**s**oku/のパターンに関しては、スペイン語では一般的に[z]が発音されないからであると考えられる。スペイン語では/s/と/z/の音韻対立がなく、/s/の主な異音は[s]である（Hualde, 2005, pp.153-156）。そのため、日本語を話すときも/z/が/s/に置換されることがあると考えられる。

ベトナム語群においては、「子供」/ko**d**omo/→/ko**r**omo/というパターンが見られた。ベトナム語では[ɗ]（入破音）が使用される（Kirby, 2011）が、そのことが要因であると考えられる。金村＆松田（2020, p.52）には、「友達」（とも**ら**ち）や「灯台」（とう**ら**い）といった誤発音が報告されている。その理由として日本語のダ行音はベトナム語の入破音に比べて軽い感じの音に聞こえ、それを真似して出した音声が日本語のラ行音のようになるのではないかと考察している（金村＆松田, 2020, pp.53）。さらに、いずれの例も母音間の/d/であるため、それも1つの要因である可能性も考えうる。語頭や撥音/N/の直後の/d/に関する検討は今後の課題とする。

タイ語群においても、「子供」/ko**d**omo/→/ko**r**omo/のパターンが観察されたが、その理由に関しては現在のところ不明である。Tingsabadh & Abramson（1993）を参照する限りにおいては、タイ語には/d/と/r/という2つの音素が存在し、[d]や[ɾ]といった音声が区別されている。そのため、日本語を話しているとき/d/が/r/になる要因は明らかではない。更なる詳しい考察は今後の課題とする。

トルコ語群において、「先生」/s**e**NseH/→/s**a**NseH/というパターンが見られた。トルコ語では/e/は/m/や/n/の前で[æ]として実現する（Göksel & Kerslake, 2005, p.10）。よって、その規則が影響し母音の置換を引き起こしたと考えられる。

3.2　誤発音の影響

日本語教育発展の一助を目指す以上は、この結果がどの程度有益な情報かを判断する必要がある。本研究で着目した誤発音の例が実際のコミュニケーションにどれほどの影響を与えるかに関する本格的な検討は今後の課題とする。加えて、教育現場で発音矯正に関して、何をどこまで行うべきかは学習者の目標や職場のニーズなどによるところが大きいと考えられるため、本稿ではどういった語と混同されうるかという視点でのみ簡素に考察する。

ちゅうご（中国）、ともだっち（友達）、パーチー（パーティー）、せんせ（先生）、におん（日本）に関しては特に混同されそうな語が想起されないため、不自然な発音であると思われる程度で済む可能性がある。特に「せんせ」は本研究で扱ったデータからは認められなかったが、特に呼びかけるときに母語話者の発音としても観察されうる。

文脈により判断に困らないことも想像できるが、りょこ（旅行）は「よこ（横）」、ここ（高校）は「此処」、かそく（家族）は「加速」、ころも（子供）は「衣」、さんせい（先生）は「賛成」

表21：誤発音パターンのまとめ

母語	語	誤発音	パターン	
中国語	中国	ちゅうご	語長の短縮	/čuHgoku/
英語	友達	ともだっち	促音化	/tomodači/
フランス語	パーティー	パーチー	子音の置換	/paHtiH/
ドイツ語	先生	せんせ	長母音の短母音化	/seNseH/
ハンガリー語	旅行	りょこ	長母音の短母音化	/ryokoH/
インドネシア語	高校	ここ	長母音の短母音化	/koHkoH/
韓国語	日本	におん	子音の脱落	/nihoN/
ロシア語	映画	えいご	母音の置換	/eHga/
スペイン語	家族	かそく	子音の置換	/kazoku/
タイ語 ベトナム語	子供	ころも	子音の置換	/kodomo/
トルコ語	先生	さんせい	母音の置換	/seNseH/

と混同される可能性がある。ロシア語群に関しては表６でも分かるように、「映画」と「英語」の両方にＧタグが付与されている。加えて、詳細を確認してみると「えいご→えいが」というパターンも見られる。この２語は頻度も親密度も高いと推測でき、「映画／英語が好きです」などの文で可能であるため、混同される可能性が高いと考えられる。

4．結論

本研究では、日本語教育発展の一助とするため、『多言語母語の日本語学習者横断コーパス：I-JAS』（迫田，2020）を用いて、日本語学習者に比較的多く使用され且つ誤発音されうる語を検討し、その誤発音パターン及びその要因を考察した。本研究の研究課題である「日本語学習者に比較的多く使用され且つ誤発音されうる語は何か」、「その誤発音のパターンや要因は何か」、「どの程度の割合で誤発音されるのか」に対する、本研究の分析結果から現時点で導くことができる結論を以下に示す。

日本語学習者に比較的多く使用され且つ誤発音されうる語は、「中国」、「友達」、「パーティー」、「先生」、「旅行」、「高校」、「日本」、「映画」、「家族」、「子供」である。これらの語で見られた最も多い誤発音のパターンは、中国語母語話者においては「ちゅうごく→ちゅうご」（語長の短縮）、英語母語話者においては「ともだち→ともだっち」（促音化）、フランス語母語話者においては「パーティー→パーチー」（子音の置換）、ドイツ語母語話者においては「せんせい→せんせ」（長母音の短母音化）、ハンガリー語母語話者においては「りょこう→りょこ」（長母音の短母音化）、インドネシア語母語話者においては「こうこう→ここ」（長母音の短母音化）、韓国語母語話者においては「にほん→におん」（子音の脱落）、ロシア語母語話者においては「えいが→えいご」（母音の置換）、スペイン語母語話者においては「かぞく→かそく」（子音の置換）、タイ語やベトナム語母語話者においては「こども→ころも」（子音の置換）、トルコ語母語話者においては「せんせい→さんせい」（母音の置換）であった。要因に関してはほとんどの場合、母語の音韻体系の影響と考えられるが、ドイツ語、ハンガリー語、タイ語母語話者における誤発音パターンの要因に関しては現時点では明確ではない。上記に示した誤発音の割合は学習者の母語により異なるが、人数で算出すると全体の4.0%から43.9%、延べ語数で算出すると全体の1.2%から35.1%であった。

5．今後の課題

母語群によっては、誤発音の要因が明確でないものが見られた。他の誤発音の例と比較を行ったり、さらに様々な文献を調べ、これらの誤発音の要因に関して考察していく必要がある。

今回の研究では、Ｇタグを手掛かりに、仮名で書き取られたデータを用いて誤発音分析を行った。同コーパスでは音声データも公開されているため、音響分析によるより詳細な音声の特定も可能である。例えば、ベトナム語やタイ語群で観察された「子供」→「ころも」の例において、「ろ」の子音は［ɾ］のような音声なのか、［l］のような音声なのか、などの疑問が残っている。

「一」、「四」、「十」といった数詞に関して本研究では詳しい分析を行わなかった。理由としては、後続する語により模範的な発音が異なり、発音が一通りの他の名詞と同様に1つの名詞として数え、比較することが妥当ではないと判断したことなどが挙げられる。しかしながら、どの母語話者群においても何らかの数詞にGタグが付与されていた（表６参照）。このことから日本語教育発展に寄与できる可能性があると思われるため、数詞に限定した研究が必要であると考える。

今回の研究では日本語学習者の誤発音、つまり模範とされる発音から逸脱した発音に着目した。反対に、誤聴取、つまり話者の意図とは違う発音として学習者が聞き取ってしまう現象に関する研究も日本語教育発展の一助となりうると考えるため、今後の課題とする。

謝辞

本研究は令和３年度日本大学国際関係学部共同研究費（課題名：「コロナ禍における滞日外国人の生活世界の再構築」）の助成を受けたものである。また、様々な質問に親身に且つ丁寧にご回答してくださったI-JAS事務局の皆様に深く感謝申し上げる。

参考文献

Dawson, Alain, Auran, Cyril, Bouzon, Caroline, Delrue, Laurence, Loock, Rudy, O'Connor, Kathleen M., & Patin, Cédric. (2016). French in Nord (Nord-Pas-de-Calais): A speaker from La Madeleine. In S. Detey, J. Durand, B. Laks & C. Lyche (Eds.), Varieties of Spoken French (pp.148-158): Oxford University Press.

Derwing, Tracey M., & Munro, Murray J. (2015). Pronunciation Fundamentals: Evidence-based perspectives for L2 teaching and research: John Benjamins.

Göksel, Aslı, & Kerslake, Celia. (2005). Turkish: A Comprehensive Grammar: Routledge.

Hulade, José Ignacio. (2005). The Sounds of Spanish: Cambridge University Press.

Kirby, James P. (2011). Vietnamese (Hanoi Vietnamese). Journal of the International Phonetic Association, 41(3), 381-392.

Kohler, Klaus. (1990). German. Journal of the International Phonetic Association, 20(1), 48-50.

Ladefoged, Peter, & Johnson, Keith. (2015). A Course in Phonetics (7th ed.): Cengage Learning.

Lin, Yen-Hwei. (2007). The Sounds of Chinese: Cambridge University Press.

Pustka, Elissa, & Vordermayer, Martin. (2016). French in Haute-Savoie (Rhône-Alpes): A speaker from Passy. In S. Detey, J. Durand, B. Laks & C. Lyche (Eds.), Varieties of Spoken French (pp.192-201): Oxford University Press.

Sakoda, Kumiko. (2020). The background of compilation of I-JAS (Written in Japanese). In K. Sakoda, S. Ishikawa & J. Lee (Eds.), Introduction to the I-JAS: Application for research and teaching (pp.2-13): Kurosio Publishers.

Shin, Jiyoung, Kiaer, Jieun, & Cha, Jaeeun. (2013). The Sounds of Korean: Cambridge University Press.

Soderberg, Craig D., & Olson, Kenneth S. (2008). Indonesian. Journal of the International Phonetic Association, 38(2), 209-213.

Szende, Tamás. (1994). Hungarian. Journal of the International Phonetic Association, 24(2), 91-94.

Tingsabadh, M. R. Kalaya, & Abramson, Arthur S. (1993). Thai. Journal of the International Phonetic Association, 23(1), 24-28.

Vance, Timothy J. (2008). The Sounds of Japanese: Cambridge University Press.

Yanushevskaya, Irena, & Bunčić, Daniel. (2015). Russian. Journal of the International Phonetic Association, 45(2), 221-228.

金村久美，& 松田真希子．(2020)．ベトナム人に日本語を教えるための発音ふしぎ大百科：ひつじ書房．

大井川朋彦．(2022)．コーパスを用いた滞日日本語学習者の言い間違いに関する予備研究．第344回日本音声学会研究例会・日本音響学会音声コミュニケーション研究会（共同開催），オンライン開催．

迫田久美子．(2020)．I-JAS誕生の経緯．迫田久美子，石川慎一郎 & 李在鎬（編著），日本語学習者コーパスI-JAS入門：研究・教育にどう使うか（pp. 2-13）：くろしお出版．

迫田久美子，石川慎一郎，& 李在鎬（編著）．(2020)．日本語学習者コーパスI-JAS入門：研究・教育にどう使うか：くろしお出版（「I-JAS」書き起こし規程集_202003［https://www2.ninjal.ac.jp/jll/lsaj/wp-content/uploads/2020/03/92142a8bd185b2288d707c6f5a7bbc37-1.pdf］）．

平田秀．(2019)．中国人学習者の促音の誤用についての一考察：コーパスI-JAS を用いて．日本語教育方法研究会誌，26(1)，14-15.

平田秀．(2020)．中国人学習者・韓国人学習者の促音の誤用についての一考察：コーパスI-JASを用いて．武蔵野大学教養教育リサーチセンター紀要，10，63-70.