

日本語学習者の談話モードと使用表現

小 森 早 江 子*

1. はじめに

小森 (2008) では、日本語話しことばデータを収集し、特に受益表現である「てもらう」「てくれる」に注目して分析をおこなった。本稿では、小森 (2008) で紹介した話しことばデータを対象に、学習者と母語話者の談話モードによる使用表現の違いについて比較検討し、学習者の言語使用の特徴を明らかにする¹。

本研究の目的は、絵を見ながらの「対話」と、人に出来事を伝える「報告」という異なる談話モードで使われる表現の共通点と相違点を比較することによって、日本語学習者と母語話者の話しことばの特徴を明らかにすることである。具体的には、話しことばデータの「対話」と「報告」のそれぞれでおこなわれた発話を対象に、ngram というプログラムを使って共起表現を抽出し比較分析する。また、コーパス中に特定の表現がどの程度散らばっているかを観察するために学習者と母語話者のデータにおける出現頻度から分散率を求めて検証する。

2. 先行研究

コーパスを使った日本語学習者と母語話者の言語使用に関する研究として、杉浦・朴 (2003) や山内 (2004) などがある。杉浦・朴 (2003) は、日本語学習者の作文コーパスとそのコーパスの誤用部分の母語話者による書き換え文とを比較分析している。学習者が書いた作文データと母語話者の修正文データの中で使用された定型表現について n-gram による処理をおこない比較した。連続する言語単位を抽出する n-gram により共起表現を抽出して、学習者と母語話者が使用した定型表現の数を調べた結果、学習者よりも母語話者は定型表現をより多く使用することが明らかになったと報告している。

一方、山内 (2004) は、話しことばの研究として、日本語学習者の Oral Proficiency

*人文学部教授—第二言語習得

1 本稿は、2008 年 7 月 11 日に釜山外国語大学で開催された日本語教育学世界大会 2008 での口頭発表をもとに、大幅な加筆、修正をおこなったものである。

Interview (OPI) データである KY コーパスを対象に n-gram の分析をおこなって、OPI のレベル別に出現頻度の高い語彙を調べた。その結果、中級学習者の表現には、「たいです」「いいです」などが多いが、上級学習者は、「たいです」より「思います」が高頻度で使用され、超上級学習者になるとそれまで使用が見られなかった「思うんです」という「オモウ」＋「ノダ」の組み合わせが使われるようになると報告している。つまり、n-gram による分析によって、それぞれのレベルを特徴づける表現を探し出すことができることを示している。山内の研究は、学習者の話しことばコーパスを使った n-gram の分析として先駆的であるが、分析対象としたデータはインタビューデータであり、あらたまった表現が多いと推測され、この使用傾向に一般性があるか、それともスタイルによる特徴であるかは、そのほかのスタイルのデータで確認する必要がある。さらに、言語表現の特徴が特定の個人に偏っていないかどうかを確かめるために分散率を算出し、ある表現がコーパス全体の中でどの程度均一に散らばって不偏的に使われているのかを調べることによって、言語表現の特徴の一般性を明らかにすることができる。

3. 話しことばデータ

本稿の分析には中国語を母語とする日本語学習者 12 人と日本語母語話者 12 人の話しことばデータを使用した。被験者は、11 種類のストーリーのある一連の絵を見て「対話」と「報告」という 2 種類の異なるモードで発話をおこなった²。データの規模は、学習者の「対話」部分の書き起こしで 15,375 字、「報告」部分が 16,938 字で、母語話者はそれぞれ 11,579 字と 14,128 字で、総計 58,020 字のデータを本研究の分析対象とした。

「対話」モードは、たとえば、資料 1 の絵の場合、被験者に登場人物であるカヨコとケンジになって、2 人の会話を考えて話してもらった。資料 1 の絵はカヨコとケンジの会話場面である。発話例 1（母語話者）と発話例 2（学習者）を以下に示す。

発話例 1 対話モード（母語話者 NS12）

- カヨコ： お昼たべようっか。
 ケンジ： スガキヤにする，マックにする？
 ケンジ： あ，マックすごい並んでるね。
 カヨコ： スガキヤもすごい並んでるよ。
 ケンジ： 寒いけど，まあ，外で食べようっか。

2 絵の内容と実例については、小森（2008）を参照のこと。

発話例 2 対話モード (学習者 NNS07)

- カヨコ： 今日ご飯どうする。
ケンジ： スガキヤで食べようか。マックの方がいいかな。
ケンジ： あー、混んでるマックは。どうしよう。
カヨコ： じゃあ、スガキヤに行こうか。
カヨコ： あー、ここも混んでる。どうしよう。
ケンジ： サンドイッチとか買って、ベンチで食べようか。

一方、「報告」モードは、同じ場面の人物になって、その日にあった出来事として家族に教えるように話してもらった。資料 1 の絵では、発話例 3 (母語話者) と発話例 4 (学習者) のようにカヨコになって、発話してもらった。

発話例 3 報告モード (母語話者 NS12)

- カヨコ： 昨日大学で昼ごはん食べようと思ったんだよね。大学にスガキヤとマックがあって、どっちで食べようかなーって思ってた、マックに行ったら、すごい人が並んで、これは食べれないなって思ったから、スガキヤに行ってみたら、スガキヤもすごい並んで、ちょっとこれは授業に間に合わなくなっちゃうから、お弁当とか売ってるところで、お弁当買って食べたんだよね。すごい寒かったんだけど、外は。

発話例 4 報告モード (学習者 NNS07)

- カヨコ： 昨日、ケンジと一緒に昼ごはん食べようと思って、先にマックに行ったの。でも、とても混んでいて、スガキヤに行った。でも、スガキヤもとても混んでいて、しょうがなくて、2人でベンチで食べてしまった。とてもかわいそうだと思いました。

4. n-gram による分析

分析は、被験者ごとに「対話」と「報告」のそれぞれの発話を書き起こしたファイルを対象に、文字単位での n-gram による共起表現を抽出した³。

山内 (2004) の分析方法に従って 4 文字以上の文字列について母語話者と学習者それぞれ

3 分析には、ngram version 0.6.1 by Fujiwara Shigeru を使用した。

について表現を抽出した。母語話者が多く使用した表現は表 1 に示すように、「と思って」「と思った」など動詞「思う」の活用形, 「ってくれ」「てくれて」など補助動詞「～てくれる」

表 1 母語話者が多く使用した表現

表現	頻度
うと思っ	6
ものすごく	6
んだって	6
っていう	5
ってくれ	5
てくれて	5
てたんだ	5
と思った	5
と思って	5
なかった	5
めちゃめちゃ	5
ようと思って	5

表 2 学習者が多く使用した表現

表現	頻度
そうですか	8
って言った	8
ましょうか	8
いですね	7
ちゃって	7
たんだけど	7
ったから	6
はすごく	6
あたらしい	6
いですか	5
かわいい	5
すごーく	5
そしたら	5
たしたち	5
っていた	5
のところ	5
ようかな	5
わたしは	5
を持って	5
待ってい	5
いいですよ	5
っちゃった	5
って言って	5
てもらって	5
ようと思っ	5
作ってくれ	5
大丈夫だよ	5
うしようかな	5
どうしようかな	5

の活用形など 12 種類の表現が抽出できた⁴。また、学習者が多く使用した表現は表 2 に示すように、29 種類の表現が抽出できた。

表 1 と表 2 は、個々の日本語母語話者と学習者が多く使用した表現である。つぎに、個々の被験者が使用した表現ではなく、2 人以上の被験者に共通して使用が見られた表現を見てみる。母語話者に共通して頻度 5 以上使用された文字列は、表 3 のように頻度が多い順に「たんだけど」「ちょっと」「ちゃった」「ちゃって」の 4 種類であった。ほかに「めちゃめちゃ」「ようと思って」など頻度 5 以上使用された文字列が 12 種類あったが、特定の母語話者個人に使われているだけで、複数の母語話者に共通して使われているものではなかった。

2 人以上の学習者に共通して多く使用された文字列は、表 4 のように「なかった」「いいです」「いました」「ましょう」「ちゃった」「どうしよう」「と思って」「ってくれ」「てくれた」「作ってく」「に行った」「てもらっ」の 12 種類であった。ほかに「たんだけど」「そうですか」「って言って」など頻度 5 以上使用された文字列が 29 種類あったが、特定の学習者に限定して使われているのみであった。

表 3 2 人以上の母語話者が共通に使用した表現

表現	人数	平均頻度
たんだけど	11	8.5
ちょっと	4	8.5
ちゃった	5	5.8
ちゃって	4	5.3

表 4 2 人以上の学習者が共通に使用した表現

表現	人数	平均頻度
なかった	4	7.5
いいです	3	7.3
いました	2	7.0
ましょう	3	6.7
ちゃった	7	6.3
どうしよう	5	5.6
と思って	2	5.5
ってくれ	4	5.3
てくれた	3	5.3
作ってく	2	5.0
に行った	2	5.0
てもらっ	2	5.0

つまり、学習者 2 人以上に共通して繰り返し使われる表現が 12 種類と多かったのに対して、母語話者では 4 種類のみであった。また、特定の個人に限定して多用される表現は母語話者の 12 種類に対して、学習者は 29 種類と多く、学習者ごとに特定の限られた表現を繰り返

4 今回のデータは、全員が同じ絵を見て絵の中の人物になって発話するという状況であったため、絵に登場した人物や物の名前（「お母さん」「るみこちゃん」「コピーカード」「スカート」など）が母語話者にも学習者にも多く使われていた。ここでは、「お母さん」などの人物名を含む表現や「スカート」などの場面に依存する内容語は分析の対象から除いた。

返し使う傾向があるのではないかと予測できる。

表 1 から表 4 までの結果をまとめると、表 5 のようになる。学習者は、学習者ごとに特定の限られた表現を繰り返し使う傾向があるのではないかと推測できる。

表 5 母語話者と学習者が多用する表現のまとめ

多用する表現	母語話者	学習者
特定の個人に限定	12	29
2 人以上に共通	4	12

5. 多用表現「ただけど」「ちゃった」「ちゃって」の分散率

つぎに、データ全体にわたる表現の散らばり具合を見るために、ここでは、母語話者にも学習者にも使われた「ただけど」「ちゃった」「ちゃって」について、被験者ごとの出現頻度の分散率を算出した⁵。まず、「ただけど」については、母語話者の 12 人中 11 人が使用していた。一方、学習者で使用したのは 1 人だけだった。表 6 に示すように、母語話者の分散率は 0.88、学習者は 0.00 となり、母語話者のデータの中においては均一に散らばっていることがわかった。また「ただけど」は母語話者も学習者も「報告」の場面で集中して使用されていた。「ただけど」が、状況説明の「のだ」を含み、ほかの人に出来事を説明する談話状況でよく使用される表現であるためと考えられる。

表 6 「ただけど」の分散率

分散率	母語話者	学習者
ただけど（報告）	0.88	0.00

山内（2004）は、超級学習者は「んですけ」「んですね」「んですよ」「んですよね」を多く使う特徴があると指摘している。つまり文法項目では「のだ＋です＋けど」という組み合わせで話しているといえる。今回分析対象とした話しことばデータは、絵に描かれた状況を家族にその日にあった出来事として話すという場面設定がされていたため、丁寧体の「デス」ではなく、普通体の「ダ」が使用された。このようなスタイルの違いを考慮に入れると、今回の学習者は OPI によるレベル分けでは超級レベルの話しことばの特徴を持っているとみなすことができる。

5 分散率の計算方法については、応用言語学事典（p.635）を参照した。

つぎに、「ちゃった」については、母語話者、学習者ともに「対話」でも「報告」でも多く使われていた。母語話者は「対話」では12人中10人、「報告」では8人が使用し、分散率は、表7に示すように、それぞれ0.83, 0.85と高く、ほぼ均等に使われていたことがわかった。一方、学習者の場合も、母語話者と比べればやや少ないものの、「対話」では8人、「報告」では9人が使用しており、分散率も、それぞれ0.73と0.78であった。学習者の「ちゃった」の使用は母語話者よりは少ないが、「対話」や「報告」の場面でよく使用されるものであることがわかる。

表7 「ちゃった」の分散率

分散率	母語話者	学習者
ちゃった（対話）	0.83	0.73
ちゃった（報告）	0.85	0.78

「ちゃって」は「ちゃった」という言い切りの表現ではなく、「～ちゃって、～」というように、おもに文中で使用され、前項を受けて話を続ける表現である。「ちゃって」は、母語話者も学習者も「報告」でのみ使用が見られた。母語話者は12人中11人、学習者は12人中3人だけで、分散率は、表8に示すように、それぞれ0.85と0.41であった。母語話者は「ちゃって」を使って状況を報告するのに対して、学習者は短い文に区切って表現するため、「ちゃった」は使うが、連用形「ちゃって」の使用は限定的ではないかと考えられる。

表8 「ちゃって」の分散率

分散率	母語話者	学習者
ちゃって（報告）	0.85	0.41

6. 談話モードの比較分析

談話モードによる比較に関して、学習者は「作ってく」「ちゃった」「つてくれ」「スカートを」という表現を「対話」でも「報告」でも使った。一方、母語話者は「ちゃった」については「対話」でも「報告」でも使用しているが、そのほかの表現はすべて「報告」でのみ使用しており、「対話」で共通する表現の使用はほとんどなかったことが特徴的である。

これは、学習者は対話モードでも報告モードでも同じ表現を使う傾向があり、母語話者は、モードによって表現を使い分ける傾向があると考えられる。今回のデータだけでは、被験者数が少ないため断言はできないが、学習者が一つの事柄を表すのに、モードの違いによって

表現を選ぶことができず、同じ表現を繰り返すためではないかと推測される。

7. まとめ

全体として母語話者は「対話」に共起表現はほとんどないが、「報告」では共通に使われる共起表現があり、モードの違いが見られた。なぜ母語話者の「対話」モードで共起表現が観察できなかったかについては、個別に詳しくデータを見てみる必要がある。「報告」モードについては、母語話者の共起表現が観察されたわけであるが、これは、杉浦・朴（2003）の研究結果でも母語話者が定型表現を多用することが指摘されており、母語話者は、ある状況を描写するのにある程度共通する表現を使うのではないかと考えられる。一方、学習者は母語話者と違い、モードによる使用表現の違いはなく、「対話」でも「報告」でも同じ表現を使用する傾向があった。今後、さらに詳しく多くのデータを分析し、このような傾向の一般性を検証する必要がある。

本研究の結果をまとめると以下のような仮説が導き出される。

- (1) 日本語母語話者には共通する表現があるが、学習者には共通する表現が少ない。
- (2) 母語話者はモードによって表現を使い分けるが、学習者はモードによって表現を使い分けていない。

表 9 日本語母語話者と学習者の違い

	共通する表現	モードによる使い分け
母音話者	多い	使い分ける
学習者	少ない	使い分けていない

参考文献

- 小池生夫（編）（2003）『応用言語学事典』研究社。
- 小森早江子（2008）「日本語学習者の受益表現の習得について」『中部大学人文学研究論集』第 19 号，pp. 191-206.
- 杉浦正利・朴秀智（2003）「日本語学習者作文コーパスにおける形態素レベルでの共起表現について」『日本語学習者語彙辞典編纂に向けた電子化コーパス利用によるコロケーション研究』pp. 1-10. 平成 13 年度一平成 15 年度科学研究費補助金（基盤研究（B）（2））中間報告論文集，研究課題番号 13480069.

山内博之（2004）「語彙習得研究の方法―茶筌とN グラム統計―」『第二言語としての日本語の習得研究』
第7号，pp. 141-162.

（2008 年 10 月 30 日受理）

資料1：カヨコとケンジ

