

日本語時制・アスペクト表現の習得分析に向けたアノテーション設計

—I-JAS と教室データを用いて—

Designing an Annotation Framework for the Analysis of Japanese Tense-Aspect Acquisition: A Comparison of I-JAS and Classroom Data

Todor Vasilev

Kobe University Graduate School of Intercultural Studies

要旨

本研究は、日本語学習者の時制・アスペクト表現の習得をより詳細に分析するため、I-JAS および教室で収集した学習者データを用いて、形式と意味機能、時間関係に対応づけるアノテーション枠組みを提案するものである。従来の形式情報のみでは、～ているが表す進行・結果状態・習慣・反復などの意味機能や、主節と従属節の時間的基準点に基づく相対時制を十分に記述することが難しい。そこで、形式、アスペクト的意味、時制タイプ、時間的基準点、事象間の時間関係、誤用の特徴を多層的にアノテーションした。分析の結果、教室データでは進行用法および絶対時制への集中が見られた一方、I-JAS ではより多様な意味機能と時間関係が観察された。また、学習段階の上昇に伴い、～ているの意味機能が多様化する傾向も確認された。以上から、形式と文脈上の機能を接続するアノテーションは、異なる学習者データ間の比較および時制・アスペクト習得過程の記述に有効であることを示す。

1. はじめに

日本語の時制・アスペクト表現は、学習者にとって習得が容易ではない文法領域の一つである。特に、ル形・タ形や「～ている」などの形式は、単純に現在・過去と対応するものではなく、文脈に応じて進行、結果状態、習慣・反復などの異なる意味を表す。また、複文では発話時だけでなく、主節の事象を基準として時制形式が選択される場合もあり、形式のみから学習者の時間的解釈を捉えることは難しい。

学習者コーパスは、このような時制・アスペクト表現の使用実態を分析する上で有用である。一方、既存の形態情報だけでは、同一形式が文脈上どのようなアスペクト的意味を表しているのか、また複数の事象がどのような時間関係にあるのかを十分に記述できない場合がある。そのため、学習者が「どの形式を使用したか」に加えて、「その形式によって何を表しているか」を記述する必要がある。

そこで本研究では、I-JAS および教室で収集した学習者データを対象として、時制・アスペクト表現を形式、意味機能、時間関係などの複数の観点から記述するアノテーション枠組みを検討する。さらに、両データへの適用を通して、学習者の形式選択と時間的・アスペクト的解釈を接続して分析する方法の可能性を示す。

2. 先行研究と研究目的

日本語学習者の時制・アスペクト習得については、これまで語彙的アスペクト、学習者の母語、習熟度などとの関係から多くの研究が行われてきた。Shirai and Kurono (1998) は、日本語学習者による時制・アスペクト形式の習得を分析し、その使用が動詞の語彙的アスペクトと密接に関係することを示している。また、Sugaya and Shirai (2007) は、「～ている」が表す進行および結果状態の習得を対象とし、その習得には語彙的アスペクトだけでなく、学習者の母語、習熟度、課題など複数の要因が関与することを指摘している。これらの研究から、時制・アスペクト形式の習得は、特定の形式を使用できるか否かだけではなく、その形式と動詞の性質および文脈上の意味との対応から捉える必要があると考えられる。

本研究のアノテーション設計では、これらの第二言語習得研究に加え、時制・アスペクト研究における理論的枠組みを参照する。具体的には、Comrie (1976) のアスペクト論に基づき、事象の時間的な内部構造とアスペクト的意味を区別する。また、Comrie (1985) の絶対時制・相対時制の区別を用いて、発話時だけでなく、主節事象など他の時間的基準点との関係から時制形式を分析する。さらに、Vendler (1957) の state、activity、accomplishment、achievement という語彙的アスペクトの分類を参照し、動詞の意味的性質と「～ている」などの形式が示す意味機能との関係を検討する。

一方、学習者コーパスでは、学習者が実際に使用した形式やその文脈を大量のデータから観察することが可能であるが、形態情報のみから、個々の形式が進行・結果状態・習慣などのどの意味機能を担っているのか、あるいは複数の事象がどのような時間関係にあるのかを直接把握することは難しい。特に、同一の「～ている」であっても文脈によって意味機能が異なり、ル形・タ形についても絶対時制と相対時制を区別する必要がある。そのため、形式情報と文脈上の時間的・アスペクト的解釈を対応づける分析枠組みが求められる。

以上を踏まえ、本研究では、I-JAS の学習者産出を形式・意味・時間関係の対応という観点から再分析し、既存情報から把握できる内容と、追加的な記述が必要となる情報を整理する。その上で、同一のアノテーションを教室で収集した学習者データにも適用し、両者の比較を通してその有効性を検討する。最終的には、学習者の形式選択と時間的・アスペクト的解釈を接続して記述できる分析枠組みを構築することを目的とする。

3. 研究方法

3.1 分析対象

本研究では、I-JAS（多言語母語の日本語学習者横断コーパス）と、授業内で収集した教室データの二種類を分析対象とした。I-JAS からは、ル形・タ形、「～ている」「～ていた」など、時制・アスペクトに関わる述語表現を含む用例を抽出し、母語、習熟度、学習歴、日本滞在経験などの学習者情報を参照した。

教室データは、多言語背景を持つ日本語学習者 33 名の授業内における文法課題、作文、産出課題から収集した。母語は、中国語・タイ語・ベトナム語の孤立語群 8 名、モンゴル語・カザフ語・トルコ語の膠着語群 12 名、ロシア語・英語・スペイン語の屈折語群 13 名である。これらのデータから同様に時制・アスペクト表現を抽出し、I-JAS と共通の基準で分析した。

3.2 アノテーション項目

抽出した用例に対して、表 1 に示す四つの層を中心にアノテーションを行った。形式そのものに加えて、文脈上の意味機能と時間関係を記述することにより、学習者が使用した形式とその時間的・アスペクト的解釈との対応を分析する。

表 1 時制・アスペクト表現のアノテーション項目

層	主なアノテーション項目
形式層	ル形、タ形、～ている、～ていた
アスペクト層	進行、結果状態、習慣・反復、完了
時間関係層	発話時基準、事象時基準、相対時制
誤用層	形式選択、意味機能、時間関係、複合的誤用

さらに、語彙的アスペクトを参照しながら各形式の意味機能を判定し、習熟度との対応から意味機能の発達的な広がりについても検討した。すなわち、本研究では単に「何を使用したか」を数えるのではなく、**「その形式が文脈において何を表しているか」**を分析単位に含めた。

4. 分析結果

4.1 アスペクトの意味と時制タイプ

追加アノテーションを適用した結果、I-JAS と教室データでは、時制・アスペクト表現の分布に異なる傾向が確認された。まず、アスペクト的意味について、I-JAS では進行が 47%、結果状態が 27%、習慣・反復が 23%、その他が 3%であった。一方、教室データでは進行が 68%を占め、結果状態は 11%、習慣・反復は 20%、その他は 1%であった。すなわち、両データとも進行用法が最も多いという点では共通しているものの、教室データでは進行への集中がより顕著であるのに対し、I-JAS では結果状態や習慣・反復を含む複数の意味機能に比較的広く分布していた。

時制タイプについても同様の差が確認された。I-JAS では絶対時制が 64%、相対時制が 36%であったのに対し、教室データでは絶対時制が 89%を占め、相対時制は 11%にとどまった。この結果から、教室データでは発話時を基準とする比較的直接的な時間表現が中心となっている一方、I-JAS では、主節事象など発話時以外の時間的基準点との関係によって解釈される表現も一定数使用されていることが分かる。

この差には、両データに含まれる学習者の習熟段階および学習環境の違いが関係していると考えられる。教室データの対象学習者は概ね A2 程度であり、調査時点では『まるごと』第 3 冊までを中心に学習していた。そのため、進行を表す基本的な「～ている」は比較的定着している一方、結果状態を含む多様なアスペクト用法や相対時制については、まだ十分に学習していない、あるいは導入されて間もない学習者も含まれる。したがって、I-JAS と教室データの分布差を単純なコーパス間の頻度差として捉えるのではなく、習熟度や既習項目、教材進度などを考慮して解釈する必要がある。

また、I-JAS の個別用例を見ると、形式情報のみでは捉えにくい意味機能の違いがより明確になる。例えば、英語母語話者による「ここにベンチに座っている二人がいますが、どちらも女性のほうが泣いているみたいで……」という産出では、「座っている」と「泣いて

いる」の双方に同一の「～ている」形式が用いられている。しかし、前者は座るという動作の後に成立した状態の継続を表す結果状態として、後者は基準時に動作が継続している進行として解釈される。この二つを単に「～ている」の使用例として処理した場合、この意味機能上の差は明示されない。

このように、形式層に加えてアスペクト的意味と時間関係を付与することによって、「～ている」の使用頻度だけでなく、学習者が同一形式をどのような意味機能で使っているのかを区別することが可能となる。さらに、ル形・タ形についても、絶対時制と相対時制、および時間的基準点を区別することで、形式の選択と実際の時間的解釈との対応をより詳細に記述できる。

4.2 誤用の特徴

誤用については、単純な正用・誤用の二分法ではなく、どのレベルで目標表現との不一致が生じているかという観点から、形式選択、アスペクト的意味、時間関係・相対時制、複合的誤用の四つに分類した。これにより、学習者が時制・アスペクト形式を適切に形成できなかった場合と、形式自体は成立しているものの、その文脈に必要な意味機能や時間関係との対応に問題が生じている場合とを区別した。

暫定集計では、教室データにおいて形式選択に関する誤用が48%と最も多く、I-JASでは31%であった。一方、I-JASではアスペクト的意味に関する誤用が27%、時間関係・相対時制に関する誤用が29%を占めており、特定のタイプに集中するのではなく、複数の層に比較的広く分布する傾向が確認された。これに対して、教室データではアスペクト的意味が24%、時間関係・相対時制が20%であり、形式選択の占める割合が相対的に高かった。また、複数の層にまたがる複合的誤用は、I-JASで13%、教室データで8%であった。

教室データに形式選択の誤用が多いことについても、前節で述べた学習者の習熟段階との関連が考えられる。A2程度の段階では、複数の時制・アスペクト形式を文脈に応じて選択すること自体が習得上の課題となっている可能性がある。一方、I-JASではより多様な習熟度の学習者が含まれるため、形式そのものの選択だけでなく、アスペクト的意味や複数事象間の時間関係に関わる問題も広く観察されたと考えられる。ただし、この点については習熟度を統制した上で、さらに検証する必要がある。

具体例として、教室データでは「島へ行ったとき、船で行きました」という産出が見られた。この文では、「行った」「行きました」というタ形自体は形態的に正しく形成されており、単純な活用上の誤りとはいえない。しかし、意図された文脈では、船による移動は「島へ行く」という事象の過程に含まれるため、「島へ行くとき、船で行きました」とすることが適切である。すなわち、問題となるのはタ形そのものではなく、主節事象を時間的基準点とした場合の従属節における形式選択である。

本研究のアノテーションでは、この例を相対時制における形式選択の誤用として記述する。これにより、単に当該用例に「誤用」というタグを付与するだけでなく、誤用が形態形成、アスペクト的意味、あるいは事象間の時間関係のどの層で生じているのかを区別することができる。このような誤用情報を形式・意味・時間関係のアノテーションと組み合わせることで、学習者がどの形式で誤ったかだけでなく、時制・アスペクト体系のどの対応関係に習得上の問題が生じているかを分析することが可能となる。

4.3 発達段階と母語類型

「～ている」の意味機能を学習段階別に見ると、初級では進行が 76%を占めたが、初中級では 58%、中級以上では 39%へと低下した。一方、結果状態、習慣・反復、完了・経験等の割合は学習段階の上昇に伴って増加した。この結果は、「～ている」の発達を単純な使用頻度の増加ではなく、同一形式によって表現できる意味機能の拡張として捉える必要性を示している。

また、教室データを母語類型別に見ると、孤立語群では進行が 74%と最も高く、膠着語群では 61%である一方、結果状態は 17%と他群より高かった。屈折語群は進行 69%、結果状態 10%と、その中間的な分布を示した。ただし、各群の人数および母語構成が異なるため、これらを直接的な類型効果とみなすことはできず、現段階では今後検証すべき傾向として位置づける。

6. まとめと今後の課題

本研究では、I-JAS と教室データを用いて、日本語学習者の時制・アスペクト表現を形式、意味機能、時間関係の対応から分析するためのアノテーション枠組みを検討した。その結果、既存の形態情報にアスペクト的意味、時制タイプ、時間的基準点、事象間の時間関係などを付与することで、同一形式の機能差や相対時制、誤用の特徴、さらに学習段階に伴う意味機能の広がりをもより詳細に記述できることが示された。また、I-JAS と教室データに共通のアノテーションを適用することで、異なる習熟段階や学習環境における時制・アスペクト使用を比較する可能性も示された。

今後は、アノテーションの判断基準とタグ体系をさらに精緻化し、複数のアノテーターによる一致度を検証する必要がある。また、I-JAS の分析対象を拡大するとともに、教室データを継続的・縦断的に収集し、習熟度や母語背景を統制した分析を進める必要がある。最終的には、他の学習者コーパスにも適用可能な時制・アスペクトのアノテーション体系へと発展させることを目指す

参考文献

- Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.
- Shirai, Y., & Kurono, A. (1998). The acquisition of tense-aspect marking in Japanese as a second language. *Language Learning*, 48(2), 245–279.
- Sugaya, N., & Shirai, Y. (2007). The acquisition of progressive and resultative meanings of the imperfective aspect marker by L2 learners of Japanese: Transfer, universals, or multiple factors? *Studies in Second Language Acquisition*, 29(1), 1–38.
- Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 66(2), 143–160.
- 迫田久美子・小西円・佐々木藍子・須賀和香子・細井陽子 (2016) 「日本語学習者コーパス I-JAS」『国立国語研究所プロジェクトレビュー』6(3), 93–110.