

対話コーパスの文字情報に基づいた聞き手反応の抽出と分類

† 東京工業大学大学院社会理工学研究科 Tokyo Institute of Technology

‡ 東洋英和女学院大学人間科学部 Toyo Eiwa University

歌代 崇史† (Takafumi Utashiro) utashiro.t.aa@m.titech.ac.jp

柳沢 昌義‡ (Masayoshi Yanagisawa) my@toyoeiwa.ac.jp

赤堀 侃司† (Akahori Kanji) akahori@ak.cradle.titech.ac.jp

1. はじめに

本研究では、聞き手反応を「発話権を持たない者（聞き手）が発話権を持つ者（話し手）に送る、談話を形成する機能を持った音声もしくは視覚の表現」と定義する。また、聞き手反応に発話権が交代する機能も認めるので、発話権が交代した発話の冒頭に聞き手反応が存在し得ると考える。

聞き手反応の言語学的研究は、水谷（1983）に見られるように80年代から盛んになされており、聞き手反応の形態や機能、談話中の出現位置に関して多くの研究がある。日本語教育においても、第二言語学習者に対する聞き手反応の教育は、重要視されているが、教材が不足している学習項目である(才田 2001)。そこで、Utashiro & Kawai (2007)では、聞き手反応の広範的形態分類と性質記述に基づき、日本語学習者に対する聞き手反応のシラバスと教授方法が提案され、混合学習環境における教授の有効性が示された。しかし、Utashiro & Kawai (2007)がシラバス開発に用いた聞き手反応の広範的形態分類には実証的手続きが不足しているため、分類の網羅性は明らかではなく、特定の聞き手反応の頻度順序も不明のままである。

2. 目的

そこで、本研究では、(1) 日本語母語話者の対話コーパスから文字化された音声聞き手反応を抜き出し、機能分類を行う。(2) 表記のバリエーションをグループ化し、それぞれの代表表記の出現回数と頻度順序、総数に対する割合を明らかにする。

3. 方法

本研究で用いたコーパスは、北九州市立大学国際環境工学部情報メディア工学科上村研究室が開発した「インタビュー形式による日本語会話データベース」（上村コーパス）である。

対話テキストから、聞き手反応の抽出と機能分類を効率的に行うため検索支援ツール（RT-Finder）を開発し、コーパスから一括して聞き手反応を抽出することを可能にした。しかし、「うん」や「えー」、「そうですねー」などの表現は、話し手が使用するフィラーとも表記形態が同一のため、一度RT-Finderによって抽出した表現を、再度、発話権の有無と文脈により、著者が聞き手反応とフィラーに分け、聞き手反応のみを複数人で機能分類した。聞き手反応の機能はメイナード(1993)を援用し、以下の5分類とした。「続けてというシグナル」、「内容理解を示す」、「感情を強く出す」、「話し手の判断を支持する表現／相手の意見、考え方に賛成の意思を表示する」、「情報の追加、訂正、要求などをする」である。

4. 結果と課題

コーパスに含まれる母語話者同士の全50会話を分析した結果、総計7715の聞き手反応が抽出された。こ

これらの表記のバリエーションは836あり、「ああ、そうですか」や「あー、そうですか」などの同等表現と考えられる異なった表記を統一するため、代表表記を決定し、表記バリエーションをまとめた。代表表記の総数は438であった。7715の全聞き手反応の中で、50回以上出現した聞き手反応（代表表記）は18種類であり、総出現数の86%を超えている事が明らかとなった（表1、2）。今後は、どのような聞き手反応を日本語学習者に教授すべきか、教育的観点から日本語教師による聞き手反応の重要度判定を行う。さらに、学習者のインタビューコーパスから聞き手反応の抽出と機能分類を行い、母語話者との分布差を検討する予定である。

本調査結果は日本語教育において、どのような聞き手反応を教えるべきか判断基準を提供し、聞き手反応のシラバス設計、教材開発に有益な資料を提供するものである。また、コミュニケーション上重要視されつつも、語学教育において教えられることが少なかった聞き手反応の教育の素地を作る上で、本データは資料的価値がある。

表1 出現回数が50回以上の音声聞き手反応のグループ化と出現回数

この表は、聞き手反応の表記バリエーションをグループ化し、代表表記を決め、それぞれの頻度を高いものから順に示したものである。

(全種類の聞き手反応の総出現数：7715)

聞き手反応（代表表記）	表記のバリエーション	出現数
うん	うん	2069
はい	はい	1198
あーそうですか	あーそうですか／ああそうですか／ああーそうですかあ／ああー、そうですか／あ、そうですか	726
んー	んー／んん／んんー／うーん／ううーん	592
あー	あー／あ／ああー／ああ	406
うんうん	うんうん／うん、うん／うん。うん。	362
そうですね	そうですね／そうですねー／そうですねえ	284
そうですか	そうですか／そうですかー／そうですかあ	235
えーえー	えーえー／ええ、ええ／ええー、ええー	143
はー	はー／はあ／はあー	87
なるほどね	なるほどね／なるほどねえ／なるほどねー	82
うんうんうん	うんうんうん	78
あーなるほどね	あーなるほどね／ああーなるほどね／ああ、なるほどね	75
えーえーえー	えーえーえー／ええ、ええ、ええ	64
あーそうなんですか	あーそうなんですか／ああ、そうなんですか／あ、そうなんですか	63
はいはい	はいはい	59
繰り返し	ゆったりと／安心する／違ふと／映画を見て／覚えながらも／国際コミュニケーション／始まりますね	59
あーそう	あーそう／あー、そう／ああ、そう／ああー、そう	56
ふーん	ふーん	56

表2 高頻度聞き手反応の出現回数と総出現数に対する割合

この表は、頻度の高い聞き手反応を、50回以上と100回以上出現した聞き手反応の種類数、合計出現回数、全種類の総出現回数に対する割合を示している。

(全種類の総出現回数：7715)

出現頻度	種類数	出現回数	総出現数に対する割合 (%)
50回以上出現した聞き手反応	18	6635	86.0
100回以上出現した聞き手反応	9	6015	77.9

参考文献

メイナード・K・泉子（1993）『会話分析』くろしお出版

- 水谷信子 (1983) 「あいづちと応答」 『話しことばの表現』 筑摩書房 37-44.
- 才田いずみ (2001) 「あいづちと割り込み—異文化コミュニケーション上の問題が生じ易い場面・状況をめぐって—」 『双方向通信による遠隔日本語学習支援システムの研究』 平成11～12年度科学研究費補助金 基盤研究(C) 研究成果報告書 9-21.
- Utashiro, T. & Kawai, G. (2007). CALL courseware for learning reactive tokens in face-to-face dialogs. *Proceedings of Interspeech 2007, Antwerp, Belgium*. 2301-2304.

使用したコーパス

「インタビュー形式による日本語会話データベース」 『平成8 - 10年度文部省科学 研究費補助特定領域研究「人文科学とコンピュータ」公募研究（「日本語会話データベースの構築と談話分析」研究代表者 上村隆一）の成果による』
<http://www.env.kitakyu-u.ac.jp/corpus/>