

## 動画からのうなずきの半自動検出と談話研究への応用

大塚容子\*・宇佐美まゆみ\*\*・伊藤敏\*\*\*

\*岐阜聖徳学園大学外国語学部

\*\*国立国語研究所

\*\*\*岐阜聖徳学園大学経済情報学部

[ykotsuka@gifu.shotoku.ac.jp](mailto:ykotsuka@gifu.shotoku.ac.jp)

[usamima@ninjal.ac.jp](mailto:usamima@ninjal.ac.jp)

[itous@gifu.shotoku.ac.jp](mailto:itous@gifu.shotoku.ac.jp)

### 1 はじめに

談話の研究は録画技術の発展に伴い、音声だけではなく、非言語行動も含めて行われるようになってきている。とりわけ、日本語の聞き手の言語行動の代表として挙げられるあいづちは、非言語行動を含めて研究されることが多い(メイナード(1993)、Kita and Ide(2007)など)。

本稿では、まず、自然会話における非言語行動を含めたあいづちの研究(大塚(2016))を紹介すると共にその限界を示す。そして、動画からうなずきを半自動で検出する試み(伊藤・大塚・鷺野(2018))について説明し、その試みの談話研究への応用について検討する。

### 2 大塚(2016)の目視によるあいづちの研究

大塚(2016)は、男性の日本語母語話者2人による30分の初対面会話におけるあいづち行動を調査した。本会話は、通常のビデオカメラによって録画されたもので、会話参加者の正面の顔を把握することはできないが、2人の会話参加者の顔の動きはほぼ目視で確認することができる。

あいづちは、堀口(1997:42)の定義に

倣い、「話し手が発話権を行使している間に、聞き手が話し手から送られた情報を共有したことを伝える表現」とする。堀口(1997)では音声表現のみを考察の対象としているが、音声表現と同様の機能をもつと考えられる聞き手の非言語行動も考察の対象とした。そして、あいづちを①非言語行動のみによるあいづち、②音声のみによるあいづち、③非言語行動と音声表現が同時に現れるあいづち(同時あいづち)、の3種類に分類した。非言語行動によるあいづちは、頭を上から下に動かす動作と、下から上に動かす動作の両者を含む。頭を横に動かす動作、聞き返しとして機能する動作、笑い、微笑みなどは含まない。

2人の会話参加者の3種類のあいづちの頻度を調査したところ、いずれの会話参加者も音声のみによるあいづちの頻度は低かった。1人の会話参加者は、非言語行動のみのあいづちが総あいづち数の83%を占めており、もう1人の会話参加者は非言語行動のみのあいづちが総あいづち数の52%を占め、同時あいづちは42%を占めていた。この調査により、非言語行動によるあいづちが自然会話において多数用いられていることがわかる。

非言語によるあいづち（非言語のみと同  
時あいづちの両者を含む）がどのような  
ときに頻繁に打たれるのかを調べてみると、  
非言語のあいづちは自分が相手に発話を促  
し、その相手が話し手となったときの発話  
中であることが多いことがわかった。

さらに、聞き手が音声によるあいづちを  
打つタイミングについて調査してみると、  
直前の話し手が視線を聞き手に向けるとい  
った、何らかの非言語的なサインを送って  
いることが多いようである。

非言語行動のあいづちが話し手の発話中  
に打たれることが多いこと、聞き手の音声  
によるあいづちは話し手からの何らかの非  
言語的なサインによって誘発されることが  
多いことを踏まえると、画像を見ながら話  
し手、聞き手の双方の動きを正確に把握す  
るには限界がある。特に、通常のビデオ  
カメラでの映像の場合、会話参加者の正面  
の画像が得られにくいので、どのような顔  
の動きの変化があったのかを正確に記述す  
るのは極めて難しい。

### 3 伊藤・大塚・鷺野 (2018) の半自 動の顔の動きの検出の試み

伊藤・大塚・鷺野 (2018) は、大塚 (2016)  
の問題点を解決する方法として考えられた  
試みである。ここでは、非言語行動のあいづ  
ちの半自動検出に向けて、2 つの提案がさ  
れている。

1 つは検出する身体部位についてである。  
目視で非言語行動のあいづちを観察する場  
合、横顔であっても後ろ向きであっても、頭  
部の動きであいづちを打っているかどうか  
を推測できるが、頭部の動きを自動で検出

するのは難易度が高い。そこで提案された  
のが顔の検出である。

2 つ目は、顔を検出するための録画方法  
である。顔を検出するためには、顔の正面動  
画が必要になる。会話は通常、会話参加者が  
向かい合って行われる。このような状況で  
会話参加者の正面動画を得る方法として提  
案されたのが、全天球 360 度カメラによる  
録画方法である。会話参加者の正面にカメ  
ラを置き、会話参加者にはカメラを見て会  
話するように要請する。録画した動画を平  
面に展開することによって、会話参加者の  
正面の顔の動画を取り出すことが可能にな  
るのである。

顔の動きから非言語行動のあいづちを検  
出するツールとして Dlib を用いた。半自動  
検出では、まず顔を検出して、顔のパーツの  
座標を取得する。非言語行動のあいづちを  
顔の上から下への動き、あるいは下から上  
への動きと定義していることから、顔の構  
成要素のうち、鼻の頂点座標の垂直方向へ  
の変動と、顔の向き (pitch) の時間変化に  
よって非言語行動のあいづちの自動検出が  
可能になる。

2 人による会話の 1 人のみを録画し、鼻  
の位置と pitch の時間変化を検出したとこ  
ろ、両者がほぼ一致した。また、目視による  
非言語行動のあいづちと自動検出された非  
言語行動のあいづちの生起時間もほぼ一致  
することが確かめられた。

図 1 は、全天球 360 度カメラで録画した  
男性と女性による会話 (約 17 秒) の鼻の動  
きと pitch を示したものである。両者が一  
致していることがわかる。

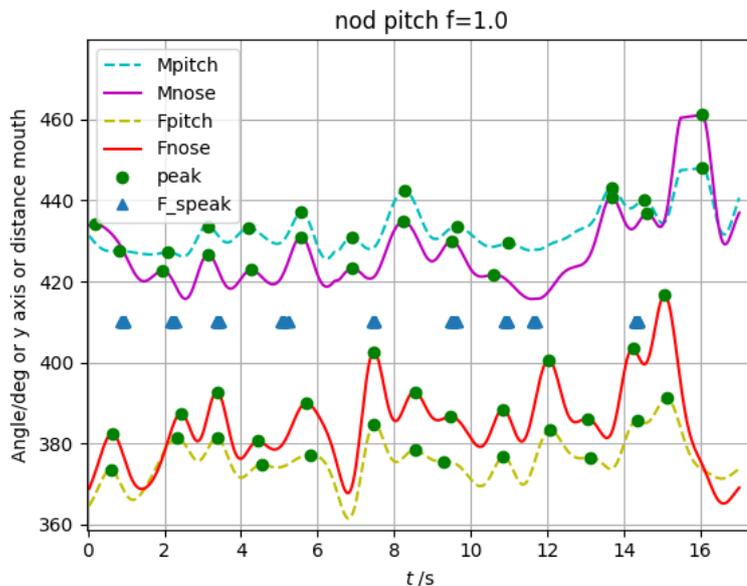


図1 2人会話における鼻の位置と pitch

#### 4 BTSJ による文字化資料と非言語行動のあいづち

図1で示した約17秒の会話を宇佐美(2012)に倣い、文字化すると以下のようになる。Mが男性話者、Fが女性話者であることを表わす。

ライン番号	発話文番号	発話文終了	話者	発話内容
1	1	*	M	はい。
2	2	*	F	はい。
3	3	*	M	まあ、こんなので、あのう。
4	4	*	F	そうですね。
5	5-1	/	M	どういふふうになるか,,
6	6	*	F	はい。
7	5-2	/	M	わかりませんが,,
8	7	*	F	はい。
9	5-3	/	M	あのう、いわゆる、ノンバーバルコミュニケーションの,,
10	8	*	F	はい。
11	5-4	/	M	研究を、あの、やるときに,,
12	9	*	F	はい。
13	5-5	/	M	ちょっと助ければ,,

14	10	*	F	はい。
15	5-6	/	M	いいかなと,
16	11	*	F	はい。
17	5-7	*	M	思いますね。
18	12	*	F	大変助かると思います。
19	13-1	/	M	またやっていたくと,
20	14	*	F	はい、ありがとうございます。
21	13-2	*	M	はい。

伊藤・大塚・鷺野 (2018) ではライン番号 3 が分析の開始時点である。上記の会話では男性が主要な話し手、女性が聞き手になっている。聞き手のあいづち行動に着目するために、図 1 には女性の発話場所が▲でマークされている。このマークは 9 か所あり、ライン番号 9 以降の F の発話回数と同時である。そして、F の「はい」というあいづち発話と非言語行動としてのあいづちが連動していることがわかる。伊藤・大塚・鷺野 (2018) のシステムでは、非言語行動としてのあいづちの発生時間が秒単位で計測可能である。

また、全天球 360 度カメラを使用することにより、会話参加者の正面の顔の動きが観察可能であることから、話し手がどのようなタイミングで、どのような方法で聞き手にあいづち (非言語行動を含む) を促しているのかも観察可能となる。

## 参考文献

- 伊藤敏・大塚容子・鷺野嘉映 (2018) 「動画から顔の動きを抽出する試み—対話解析・修学行動評価への適用を目指して—」教育システム情報学会第 43 回全国大会 2018 年 9 月 6 日於北星学園大学
- 宇佐美まゆみ (2011) 『基本的な文字化の原

- 則 (Basic Transcription System for Japanese:BTSJ) 2011 年版』
- 宇佐美まゆみ (2012) 「BTSJ 文字化入力支援・自動集計・複数ファイル自動集計システムセット (2012 年改訂版)」『人間の相互作用研究のための多言語会話コーパスの構築とその語用論的分析方法の開発』平成 20-22 年度科学研究費補助金基盤研究 B (課題番号 20320072) 研究成果
- 大塚容子 (2016) 「初対面の 2 人会話におけるあいづち行動—非言語行動を含めて—」『岐阜聖徳学園大学紀要<外国語学部編>』第 55 集、71-83 頁
- メイナード、泉子・K (1993) 『会話分析』くろしお出版
- Kita, Sotaro and Ide, Sachiko (2007) Nodding, *aizuchi*, and final particles in Japanese conversation reflects the ideology of communication and social relationship. *Journal of Pragmatics* 39, 1242-1254.