

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 20 日現在

機関番号：62618

研究種目：若手研究(B)

研究期間：2012～2013

課題番号：24700109

研究課題名(和文)音響的特徴に基づく話者交替に関する発話単位の認定基準の構築

研究課題名(英文)Construction of criteria for annotation of utterance-units based on acoustic features

研究代表者

石本 祐一 (Ishimoto, Yuichi)

大学共同利用機関法人人間文化研究機構国立国語研究所・研究情報資料センター・特任助教

研究者番号：50409786

交付決定額(研究期間全体)：(直接経費) 2,100,000円、(間接経費) 630,000円

研究成果の概要(和文)：話者交替に関する発話単位の会話分析の専門家でなくとも付与できるようにする客観的基準を構築するため、発話の区切りを示す音声の音響特徴量について分析を行った。その結果、発話全体に渡る音声の基本周波数が発話末へ向けて下降し、どの発話でも発話末ではほぼ同じ値になることが明らかになり、この変化が発話末を示す手がかりとなることを示した。また、発話末に置かれる統語要素である発話末要素で特に大きく韻律が変化し、ヒトがその韻律から発話末を予測できることから、音響特徴量から発話末を予測できることがわかった。

研究成果の概要(英文)：To construct subjective criteria for annotation of utterance-units related to turn-taking, I analyzed the prosodic behavior of spontaneous utterances. Furthermore, to clarify usefulness of prosodic features for a definition of the unit, I confirmed the influence on cognition of the end-of-utterance for hearers by perceptual experiments using utterances altered a part close to the end. In the results, the fundamental frequency (F0) range of the utterance was fixed to a certain width, and the F0s decline toward the end of the utterance in that range. The F0 declination is reset at the strong syntactic boundary in the utterance, and the final lowering occurs only at the end of the utterance. In addition, the conspicuous changes in the prosodic features occur mostly at the final accentual phrase with utterance-final elements. In the perceptual experiments, hearers can predict whether there is a following utterance-final element, and they feel that the end of the utterance arrives soon.

研究分野：総合領域

科研費の分科・細目：情報学、メディア情報学・データベース

キーワード：韻律 自発発話 話者交替 発話末予測 コーパス

1. 研究開始当初の背景

(1) 人文科学系分野のインタラクション研究において、人間同士のコミュニケーションを理解するために音声対話から人間の社会的・認知的メカニズムを明らかにしようとする試みが行われている。また、情報工学系分野では、音声認識技術や言語処理技術を組み合わせてコンピュータに実装することで、音声インターフェースによる対話型システムの実用化を目指した研究が進められている。このような幅広い領域からの要請に応じて、これらの研究の基礎とすべく音声対話を収録したコーパスが様々な機関で作成され公開・利用されている。しかし、対話が独話と異なる点でもっとも象徴的な話者の交替について、分野間で統一された基準を持たないため、これら対話コーパスに話者交替に関するラベルが付与されることはほとんどない。

(2) インタラクション研究の一分野である会話分析では、Sacks らにより提案された「ターン構成単位」(Turn Constructional Unit; TCU) が用いられている。TCU は人間が話者交替で用いている発話単位として認められているものの、会話分析の専門家は自己の母語話者としての直感を頼りに TCU を認定しており、TCU の識別方法は体系的に示されていない。一方、コンピュータによる音声対話システムにおいては、話者交替が起こりうる箇所として無音区間の長さを基に判断することが多い。無音区間による単位は間休止単位 (Inter-pausal Unit; IPU) と呼ばれ、一定時間以上の無音で区切るため客観的基準に基づいた認定をすることができ、その簡便さからラベリングも容易である。しかし、IPU には言いよどみや長い促音などが生じた場合でも単位を区切ってしまうという問題があり、話者交替の単位として適切とはいえない。TCU のような発話単位を客観的な基準から認定できれば、話者交替の発話単位のラベリングが容易となる。

2. 研究の目的

(1) 対話コーパスに必要とされる発話ラベルを付与するために、話者交替の起こりうる箇所の識別に有用な音響的特徴を明らかにすることを目的とする。

(2) これまで会話分析の専門家の直感によってなされていた話者交替についての発話単位 (TCU) の認定を、話者交替の起こりうる箇所を示す音響的情報から得られる客観的な基準から試みる。

3. 研究の方法

(1) 自発性の高い発話の全体的な音響的特徴を調べるために、「日本語話し言葉コーパス」の学会講演と模擬講演における独話と

インタビュー形式の対話の韻律変化について分析する。具体的には、イントネーション句 (Intonation Phrase; IP) を単位として発話中の基本周波数 (F0) の変動を追い、従来から発話の特徴とされている F0 declination (発話に要する時間の関数として単純に F0 が低下する現象) や final lowering (平叙文末尾で F0 が局所的に下降し発話の終了を示す現象) が長く複雑な自発発話においても生じるかどうかを、発話の長さや統語構造との関係も考慮に入れて調べる。

(2) 日常的な対話音声について「千葉大 3 人会話コーパス」を基に、話者交替が起こるときに生じる話し手発話の韻律変化を分析する。話者交替が起こりうる発話において、アクセント句 (Accental Phrase; AP) を単位として F0、パワー、モーラ時間長 (発話速度) を抽出し、話し手の発話の終わりを聞き手に予測させる韻律情報を調べる。また、日本語においては TCU 末尾に助動詞「です・ます」や終助詞「ね・よ」などの「発話末要素」と呼ばれる統語要素が存在し、この発話末要素が聞き手に発話の終わりを知らせる機能を有していると考えられていることから、前述の韻律情報と発話末要素との関係も考慮する。

(3) 話し手の発話の韻律変化が聞き手の発話末認知に及ぼす影響を調べる。聞き手が発話末予測に発話末直前 (たとえば、発話最後の 1 モーラ) 以外の情報を利用しているなら、発話末直前の音声がなくとも発話末の位置をある程度正確に推測できることになる。また、発話末付近の韻律が通常と異なる場合は発話末の予測が困難になることが考えられる。そこで、発話末付近の音声の一部を削除した刺激や F0 やパワーを操作した合成音声による刺激を用いた聴取実験を行い、音響的特徴が聞き手の発話末認知にどのような影響を及ぼすか調べる。

4. 研究成果

(1) 自発性の高い独話・対話において IP 単位の F0 最大値・最小値の発話中の変動を調べた。

① 独話については、

- 発話内で徐々に下降する (図 1)
- 発話の長さ (IP 数) に関わらず、IP はほぼ一定の高さで始まり一定の高さで終わるが、発話が長い場合は若干高い F0 で発話を開始する (図 2)
- 発話末に著しい F0 下降がみられる (図 1)

という傾向がみられた。

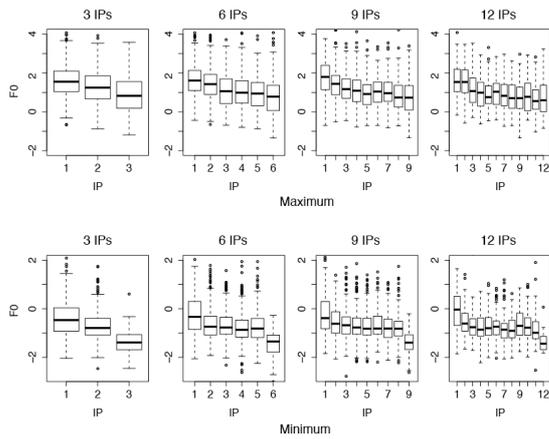


図1 独話における発話内の IP の F0 (Z 値) の推移

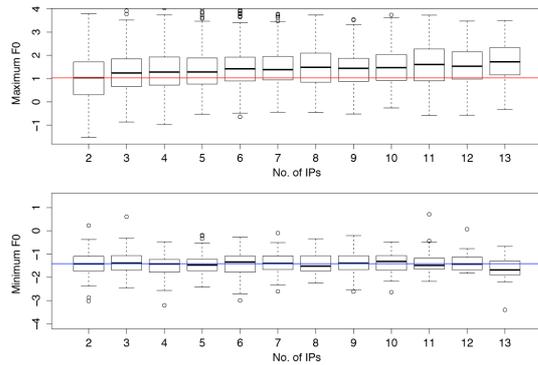


図2 独話における発話冒頭・末尾の F0:
(上) 節単位 2 番目の IP の F0 最大値
(下) 節単位末尾の IP の F0 最小値

また、特に発話中に強い統語的境界が存在する場合（二つ以上の節で発話が構成される場合）は

- 発話中の節境界で IP の F0 下降傾向が途切れてリセットされる（図3）
- 強い統語境界では F0 最小値は発話末のレベルにまで達せず、final lowering に相当する著しい F0 下降はみられない（図3）

という現象がみられることがわかった。

以上をまとめると、従来言われていたアクセント句のダウンステップ以外に発話の冒頭から末尾に向けて IP 単位でも徐々に基本周波数が下降していること、発話中に強い統語境界がある場合に下降傾向のリセットが生じること、それにより長い発話であっても発話末だけに final lowering と呼ばれる基本周波数の著しい下降が生じることが観察された。

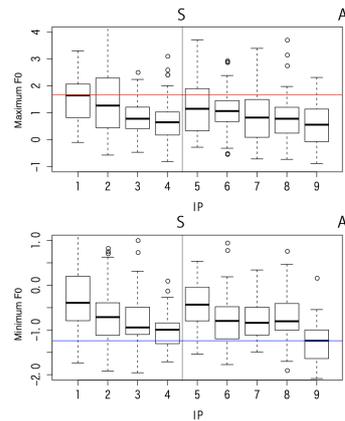


図3 発話内部に強い統語境界を持つ独話音声の F0 推移

② 対話についても同様の分析を行ったところ、独話においてみられた現象と類似した強い統語境界での F0 下降のリセットが観察された。しかし、独話においてははっきりとみられていた後発話末尾の F0 の急激な下降が明確には生じていなかった。（図4）

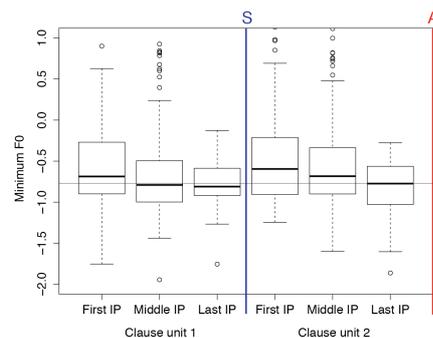


図4 発話内部に強い統語境界を持つ対話音声の F0 最小値の推移

これにより、対話音声においては final lowering が現れないことが考えられる。

(2) 話者交替が起こりうる発話に対し、AP 単位で見ると

- 最終 AP で基本周波数が最も低下して基底値に達する
- パワーは最終 AP で急激に低下する
- 発話速度は発話末に向けて速くなっていくが最終 AP では逆に遅くなる

という現象が観察された。また、これらの特徴は発話末要素と呼ばれる助動詞「です・ます」や終助詞「ね・よ」といった統語要素の存在によって強く現れ、韻律変化が発話末要素の出現を特徴づけているという知見を得た。

(3) 韻律情報が発話の聞き手の発話末認知に与える影響を調べるため、韻律情報を操作した音声を用いた聴取実験を行った。

① 図5に示す発話末付近から雑音により一部の音声情報を除去した音声を聴かせ、被

験者に発話が終わると感じたところでボタンを押させる聴取実験を行った。

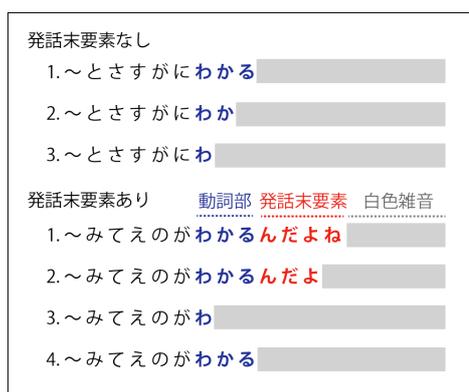


図5 発話末付近を除去した音声刺激の概要

その結果、発話末要素がない場合は発話の最後まで聴き取ることができないときには音声の途絶から 200ms 程度遅らせて反応しており、1~2 モーラ程度の時間差を以て発話末がくると判断していると考えられる。一方、発話末要素がある場合は最終モーラが聴き取れなくとも最後まで聴いた場合と比べて反応時間に差はなかった。(図6)

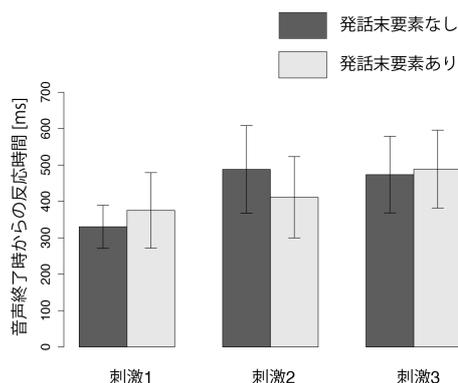


図6 発話末除去音声に対する反応時間

さらに、元々発話末要素が存在しない音声と発話末要素を除去した音声に対する反応時間を見ると、発話末要素が存在した音声では反応が有意に遅れていた。(図7)

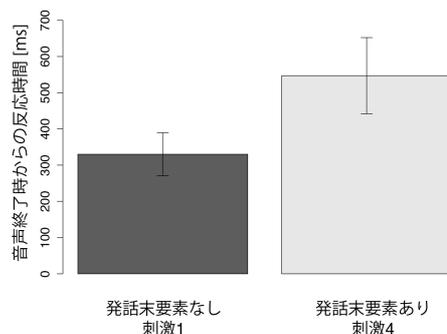


図7 発話末要素除去音声に対する反応時間

これらはともに言い切りの形で発話が終わ

っており、言語的には発話末要素が後続するか否かは判断できない刺激である。以上から、発話末要素が出現する前にヒトはその発話に発話末要素が付属することを知覚できることがわかった。すなわち、韻律等の音響的情報が発話末要素の後続を示していると考えられる。

② 最終 AP に相当する区間の F0 とパワー、発話全体の F0 を上昇(増加)・下降(減少)させた合成音声を作成し、被験者に発話が終わると感じたところでボタンを押させる聴取実験を行った。その結果、発話末要素が存在する場合は最終 AP の F0 やパワーの変化が通常と異なると発話末の予測に乱れが生じる傾向が観察された。一方、発話末要素がない発話に対しては反応時間の差が見られず、発話末要素がない場合に発話末をヒトに予測させる音響的特徴はこの実験で操作した F0・パワー以外のものである可能性がある。

(4) 以上の結果から、音響的特徴量、とくに F0 等の韻律情報が発話末の接近や発話末要素の存在を際立たせ発話末の予測に利用できる可能性が示唆された。しかし、発話末要素の存在しない発話における発話末予測の手がかりがまだ不明であること、韻律情報の変化の度合いと発話末との関係の解明がまだ不十分であることから、発話単位認定のための客観的基準の構築にはいたらなかった。今後の課題としては、さらに分析を加えて F0・パワー・発話速度以外の特徴量も考慮することで、発話末を予測できるような音響特徴の解明とその利用方法の開発が挙げられる。

5. 主な発表論文等

[学会発表] (計 12 件)

1. Yuichi Ishimoto, Tomoyuki Tsuchiya, Hanae Koiso, Yasuharu Den, Towards Automatic Transformation Between Different Transcription Conventions: Prediction of Intonation Markers from Linguistic and Acoustic Features, LREC2014, 2014 年 5 月 28 日, Harpa Conference Centre, Iceland.
2. 石本祐一, 榎本美香, 韻律情報を活用した発話単位の認定 -自動ラベリングに向けて-, 日本音響学会 2014 年春季研究発表会, 2014 年 3 月 10 日, 日本大学.
3. 石本祐一, 小磯花絵, 独話音声と対話音声の発話末の F0 変化, 第 5 回コーパス日本語学ワークショップ, 2014 年 3 月 6 日, 国立国語研究所.
4. 石本祐一, 榎本美香, 話者移行に関わる発話末付近の韻律変化の実験的検証, 日本音響学会 2013 年秋季研究発表会, 2013 年 9 月 27 日, 豊橋技術科学大学.
5. 石本祐一, 小磯花絵, 日本語話し言葉コーパスを用いた対話音声のイントネーシ

- ョン句の分析, 第4回コーパス日本語学ワークショップ, 2013年9月5日, 国立国語研究所.
6. 石本祐一, 土屋智行, 小磯花絵, 伝康晴, 会話コーパスの転記方式の相互変換 - 言語・音響特徴を用いた会話分析方式の音調マーカールの導出-, 第4回コーパス日本語学ワークショップ, 2013年9月5日, 国立国語研究所.
 7. Yuichi Ishimoto, Mika Enomoto, Hitoshi Iida, Prosodic Changes Pre-announcing a Syntactic Completion Point in Japanese Utterance, Interspeech2013, 2013年8月27日, Lyon Convention Centre, France.
 8. 石本祐一, 榎本美香, 発話完結可能点を示す発話末要素の後続を予測させる韻律情報, 日本音響学会 2013年春季研究発表会, 2013年3月15日, 東京工科大学.
 9. 石本祐一, 小磯花絵, 自発発話におけるイントネーション句単位の F0 変動の特徴, 第3回コーパス日本語学ワークショップ, 2013年3月1日, 国立国語研究所.
 10. Yuichi Ishimoto, Hanae Koiso, Prosodic Features of Utterances in the Corpus of Spontaneous Japanese: Intonation Phrase-based Approach, Oriental COCOSDA 2012, 2012年12月12日, University of Macau, China.
 11. 石本祐一, 榎本美香, 話者移行手がかかりとしての発話末の音調変化, 日本音響学会 2012年秋季研究発表会, 2012年9月19日, 信州大学.
 12. 石本祐一, 小磯花絵, 日本語話し言葉コーパスを用いた統語境界におけるイントネーション句変動の分析, 第2回コーパス日本語学ワークショップ, 2012年9月7日, 国立国語研究所.

6. 研究組織

(1) 研究代表者

石本 祐一 (ISHIMOTO, Yuichi)

国立国語研究所・研究情報資料センター・特任助教

研究者番号: 50409786