

パラ言語情報に含まれたユーザの曖昧な意図を抽出する客観的数値化法の提案

○稲岡 翔馬
Shoma Inaoka

信州大学大学院

Email: 13fm204g@shinshu-u.ac.jp

小松 孝徳
Takanori Komatsu

明治大学

Email: tkomat@meiji.ac.jp

松村 嘉之
Yoshiyuki Matsumura

信州大学大学院

Email: matsumu@shinshu-u.ac.jp

本研究では、パラ言語情報にユーザが言葉で表現したくても表現できない曖昧な意図が込められていると考え、同じく曖昧な意図を表現できるオノマトペの数値化の先行研究の考えをもとに、パラ言語情報の時系列特徴量に着目して、ユーザの曖昧な意図を抽出できるような手法の提案を行うことを目的とする。具体的には、先行研究で用いられた形容詞対に合わせた意図を発話に込めた実験を行い、意図と発話時間の関係を検討していく。

1. はじめに

日常のコミュニケーションは発話による音声を通して行われるものが大部分を占めているといわれている[1]。こういったコミュニケーションの中でやり取りされる情報は、文字に書き起こすことが可能な言語情報と、声の大きさや高さ、発話時間など、文字に書き起こすことができないパラ言語情報の二種類に分けることができる。パラ言語情報には、発話における意図やイメージ・ニュアンスといった明確に分類することが困難なアナログ的な情報が込められていると言われており、それらはピッチやパワーといった音響的特徴量や発話時間などの変化として観察することができる。近年、パラ言語情報からユーザの感情や態度といったデジタル的に扱うことが可能な情報を抽出しようという試みが数多く行われている [1, 2]。しかし、パラ言語情報の特徴量から、ユーザの「微妙」なニュアンスや意図といったアナログ的な情報を抽出しようという試みは行われていないのが現状である。そこで本研究においては、パラ言語情報に含まれる時系列的特徴量に注目した上で、ユーザの言葉で表現しきれない曖昧な意図を抽出する手法の提案を目的とする。具体的には、パラ言語情報のピッチ、パワーおよび発話時間といった時系列的特徴量と、発話意図との関係を精査することで、パラ言語情報の特徴量からユーザの意図をアナログ量として抽出する手法を構築する。

る方法としてオノマトペに着目した。「オノマトペ (Onomatopoeias)」とは物体の音の響きや状態を感覚的に表現した言葉であるため、一般語彙と比べ臨場感に溢れ、繊細な表現を可能とするという特徴を持つ。また、オノマトペには「表現したくても、しきれない曖昧な印象」が込められていると言われており、小松ら[3]はそれらの印象を客観的に扱うことを目的として、オノマトペの印象を、{キレ・俊敏さ、柔らかさ、躍動感}といった3属性で表現する方法を提案した。そこで、我々はパラ言語情報もオノマトペと同じように「ユーザの言葉で表現したいけれども、うまく表現できない曖昧な意図」を多分に含んでいると考え、パラ言語情報に込められたユーザの意図は、{キレ・俊敏さ、柔らかさ、躍動感}という3属性によって表現されると仮定し、パラ言語情報に含まれるユーザの曖昧な意図を抽出し数値化を行う。具体的には、図1に示すように、パラ言語情報の時系列的特徴量をオノマトペの印象の3属性である{キレ・俊敏さ、柔らかさ、躍動感}に変換することで発話に込められた曖昧な意図を抽出する。



図1 提案する手法の概要

2. 提案する手法

本研究では、パラ言語情報と同じく曖昧な意図を表現す

3. 実験

3.1 実験概要

発話に込められた意図とパラ言語情報の特徴量である発話時間とにどのような関係があるのかを検討するため、発話に意図を込めた実験を行った。

3.2 実験方法

本実験には、22歳～24歳の3人の大学生（男性：3名）が参加した。参加者にはヘッドセットを着用させ、マイクと口の距離を一定に保ち、ノートPCの正面に座らせ、あらかじめ実験者が指示した意図を込めて“あ～”と発声して音声解析ソフトに入力させた。また、実験者が指示した意図は、図2に示したような、小松ら[3]のオノマトペの印象の3属性を構成する43種類の形容詞対の内、各属性から1組ずつ{鋭い-鈍い} {柔らかい-堅い} {激しい-穏やかな}の3つを選び、音声解析ソフトwavesuferを用い、発話時間を抽出した。

キレ・俊敏さ	鋭い-鈍い	速い-遅い	はつきりした-ぼんやりした	鮮やかな-ぼけた
柔らかさ	柔らかい-堅い	暖かい-冷たい	丸い-四角い	ゆるんだ-緊張した
躍動感	激しい-穏やかな	明るい-暗い	大きい-小さい	興奮した-落ち着いた

図2 43種類の形容詞対の例

3.3 実験結果

参加者3人の意図に対する発話時間を表1に示す。

参加者	キレ・俊敏さ		柔らかさ		躍動感	
	鋭い	鈍い	柔らかい	堅い	激しい	穏やかな
A	0.15[s]	0.6[s]	1.25[s]	0.55[s]	0.76[s]	1.05[s]
B	0.15[s]	0.75[s]	0.94[s]	0.1[s]	2.45[s]	0.2[s]
C	0.2[s]	0.31[s]	0.7[s]	0.4[s]	0.5[s]	0.55[s]

表1 各意図に対する発話時間

表1より、意図に対する発話時間の影響を検討するために、各属性の形容詞間で対応のあるt検定を行った。その結果、柔らかい意図を込めて発話した際の発話時間(0.96[s])と堅い意図を込めて発話した際の発話時間(0.35[s])との間には有意差が観察され($t(2)=3.79$ $p=0.032$)、鋭い(0.167[s])と鈍い(0.55[s])、との間には有意傾向が観察された($t(2)=-2.97$ $p=0.058$)。激しい(1.24[s])と穏やかな(0.6[s])との間に有意差は観察されなかった ($t(2)=0.79$ $p=0.25$)。

3.4 考察

実験結果より、{柔らかい-堅い}間の発話時間に有意差が観察されたことから、{柔らかい}意図を表現すると発話時間が長くなるといえる。また{柔らかい-堅い}は「柔らかさ」の属性ベクトルに分類されることから、発話時間と「柔らかさ」の属性値との間に正の相関の対応関係を得た。しかし、今回の実験では実験参加者の数が3人と少なく、用いた形容詞対が各属性1組ずつと少なかったため、今後参加者と用いる形容詞対を増やして実験を行う必要がある。また、今回の実験では、発話時間にのみ注目したが、今後はピッチ、パワーといった音響的特徴量にも着目し分析をしていく必要がある。

4. おわりに

本研究では、パラ言語情報に込められたユーザの意図が、{キレ・俊敏さ、柔らかさ、躍動感}という3属性によって表現されるという仮定に基づき、3属性を構成する形容詞対の意図を込めた発話に意図を込めた実験を行った。そして、パラ言語情報の発話時間に着目し、分析を行った結果、発話時間の長さがオノマトペの「柔らかさ」の属性値に影響を及ぼすことが明らかになった。しかし、実験参加者が少なかったため、今後実験参加者を増やして検討していく必要がある。本研究を進めることでユーザの曖昧な意図を抽出することができれば、パラ言語情報を読み取って、ユーザの意図に合わせた情報を返すといった音声認識インタフェースの開発などが期待される。

参考文献

- [1] 藤江真也, 「パラ言語情報を利用した音声対話システムに関する研究」, 早稲田大学大学院工学系博士論文, 2005
- [2] 田島智宣, 小倉加奈代, 西本一志, 「聞き手によるパラ言語調整が対面対話に及ぼす影響の検討」, 情報処理学会, インタラクシオン 2012 論文集, pp.935-940, 2012.
- [3] 小松孝徳, 清河幸子, 秋山広美, 「オノマトペから感じる印象を表現する属性とその客観的数値化」, HAI シンポジウム 2009, 2009

¹ <http://www.speech.kth.se/wavesurfer/>