

非タスク指向対話における言語的協同の分析

An Analysis of Linguistic Alignment in Non-task-oriented dialogue

楊 潔¹ 浅井 拓也¹ 菊池 英明¹

Yang Jie¹, Asai Takuya¹, and Kikuchi Hideaki¹

¹ 早稲田大学人間科学研究科

¹ Graduate School of Human Sciences

Abstract: Interlocutors tend to express themselves in similar ways and this is termed alignment. According to previous researches, alignment plays an important role in a success dialogue, including the comprehension and production of utterance. In this study, we analyzed linguistic alignment at morpheme level. Based on our definition and judgement of aligned-utterance, we abstracted three feature values using 205 samples from non-task-oriented dialogue corpus. As a result, the length of aligned-utterance, the parts of speech of aligned-utterance and aligned position are analyzed as feature values. We expect this can be used in dialogue system for the automotive production of aligned-utterance.

1 はじめに

自然対話において、発話者は例えば語彙の繰り返しやイントネーションの同調など、相手に似せた方法で対話に参加する傾向がある。そのような現象は言語的協同(Linguistic Alignment)と呼ばれる[1]。言語的協同によって、対話者はより効率的に意思疎通をし、対話の進行に重要な役割を果たすため[2]、心理言語学・計算言語学・HAI など幅広い分野で研究がなされている。従来の研究では、ナビゲーションゲーム[2]や意思決定[3]などタスク指向対話を対象として数多く分析されているが、言語的協同は自然対話においてよく観察されたため、非タスク指向対話を対象とする研究も重要である。また、HAI 分野における言語的協同の研究も一問一答の形式に止まっている。そこで、本研究では協同発話を自動的に生成する対話システムを開発するための前段階として、人間同士の非タスク指向対話における協同発話生成のメカニズムを解明することを目的とする。

2 方法

2.1 新入生対話コーパス(FDC)

本稿では自然な非タスク指向対話解析用のコーパスとして新入生対話コーパス(FDC)を使用する。上記コーパスは著者らの研究室において作成され、毎年研究室に配属された学生 10 人程度の総当たりの発話を収録したものである。収録は 4 月および 10 月

の 2 回行われ、本年で 6 年分の対話が収録されている。

1 つの対話は 5 分とした。収録データに対し発話区間ラベリングと書き起こし作業が行われている。発話区間ラベリングでは機械的に 200msec 以上の無音区間で発話区間を区切った。

本稿ではそのうち 38 個の対話(時間にして約 230 分)を使用した。

2.2 協同発話の判定基準

本研究では、言語的協同のうち特に形態論レベルの協同を対象とする。協同の効果をもたらす簡単且つ汎用性が高いデザインを目指すため、形態的協同を「対話参加者が共感を示すために、相手の発話の全部或いは一部を繰り返す発話」と定義する。繰り返された発話は先行発話、繰り返す発話は協同発話と呼ぶ。また、対話システムへの応用可能性を前提として、協同発話の判定基準を以下の通りに定める。

(1)協同発話は隣のターンに限る。

(2)語彙に変更がない。

ただし、助詞のような機能語の言い換えがあっても「協同発話」と認める。

(3)情報付与を行っていない。

ただし、同じ話者による発話は複数ある場合に、他の文に新しい情報が加えられても構わない。

また、以下のような場合は協同のような「共感性」の意味が含まれていないため、先行発話が繰り返されても対象外とする。

(1)挨拶

(2)言葉の意味の確認

この中に、聞き返しや言い間違いの指摘などの発話が含まれる。

(3)「質疑—応答」の隣接ペア

2.3 協同に関わる特徴量

協同発話の定義と判定基準に従って、本研究では学生同士の非タスク指向対話コーパスから協同発話205例を抽出して形態素解析を行い、協同発話の長さ、協同発話の品詞と協同の位置を解析した。協同発話の長さは形態素数によって決めた。また、協同発話は2形態素以上の場合に、対話システムをデザインする際の便利さを考慮に入れ、一つ目の語彙の品詞情報だけを分析した。さらに、協同の位置を表す指標は先行発話における繰り返された部分の開始位置と協同発話の長さにした。

3 結果

3.1 協同発話の長さ

形態素解析の結果、協同発話の長さは1~7形態素の範囲に分布している。特に1~3形態素の協同発話は全体の90%以上に占め、対話者は先行発話の短い部分を繰り返すという特徴が観察された。

3.2 協同発話の品詞

品詞情報を分析した結果、全205例のうち186例が内容語であり、全体のほぼ91%に占めている。その中に、特に名詞から繰り返す対話例が多く、全体の半分以上になるとわかっている。

3.3 協同の位置

協同発話は先行発話のどの部分から繰り返され始めるかを図1の散布図で示した(先行発話が1形態素のみで構成されるデータを除外し、全188例を対象とする)。

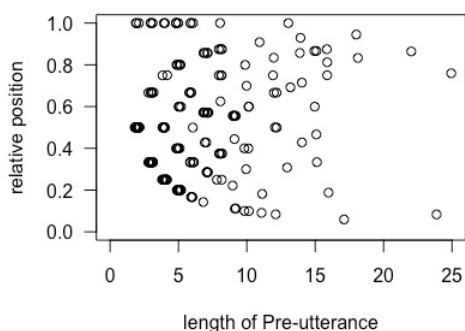


図1: 協同の位置

横軸は先行発話の形態素数であり、縦軸は繰り返され始めた部分の相対位置(先行発話の形態素数により正規化済)である。図1から、先行発話は長ければ長いほど、後ろの部分が繰り返されやすいという傾向が観察された。

先行発話の長さと協同の位置の関係性をより明らかにするために、相対位置の平均値を取って線形回帰分析を行なった。なお、信頼性を高めるために一事例しかないデータ(先行発話の形態素数が17,22,24,25)を除外した。

線形回帰の結果を図2に示している。

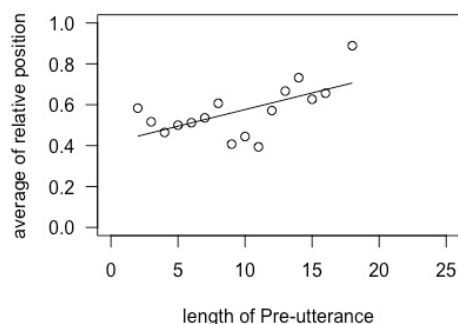


図2: 相対位置の平均値に対する線形回帰

4 おわりに

本研究では、非タスク指向対話を対象として、自然対話における形態的協同の特徴を「協同発話の長さ」、「協同発話の品詞」と「協同の位置」という三つの側面から考察し、協同位置の数値化を提案した。それによって、協同発話生成のメカニズムを明らかにし、協同発話を自動的に生成するモデルを構築するための示唆を得た。また、本研究で利用された学生同士の非タスク指向対話は自然対話なので、汎用性の高い対話システムの開発にも期待できる。それぞれの特徴量の優先順位や対話システムへの評価は今後の課題とする。

参考文献

- [1] Doyle, G., and Frank, M. C.. Investigating the sources of linguistic alignment in conversation, In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Vol.1, pp. 526-536, (2016)
- [2] Martin J. Pickering, Simon Garrod. Toward a mechanistic psychology of dialogue, Behavioral and brain sciences, Vol.27, No. 2, pp. 169-190, (2004)
- [3] Riccardo Fusaroli, Joanna Rączaszek-Leonardi, and Kristian Tylén. Dialog as interpersonal synergy. New Ideas in Psychology, No. 32, pp. 147-157, (2014)