

会話コーパスのアンノテーション手法の開発と対話の認知・伝達モデルの構築



千葉大学・文学部・教授
原 康晴

科学研究費助成事業 (科研費)

対話における発話単位とその機能の認定に関する研究 (2008-2010 基盤研究 (B))

発話単位アンノテーションに基づく対話の認知・伝達融合モデルの構築 (2011-2013 基盤研究 (B))

発話連鎖アンノテーションに基づく対話過程のモデル化 (2014-2016 基盤研究 (B))

2011-2014 国立国語研究所 共同研究 (独創・発展型)「多様な様式を網羅した会話コーパスの共有化」

2012-2013 国立情報学研究所 共同研究 (戦略研究公募型)「実場面インタラクション理解のための非談話行動アンノテーション手法の開発と談話・非談話行動の連鎖分析」

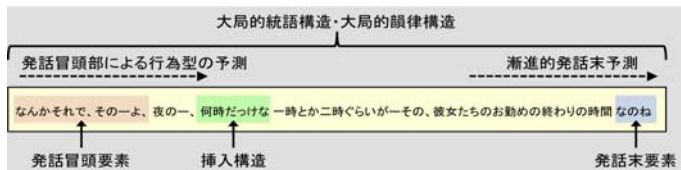


図1 会話のための文法 自然言語の文法には、会話を円滑に進めるための要素が組み込まれている

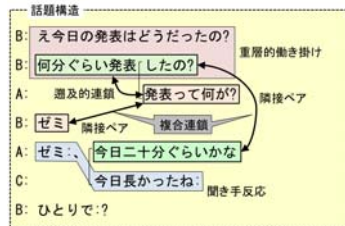


図2 発話連鎖 会話は、基本的な発話連鎖を「部品」として組み合わせることで展開



図3 非談話行動 食事動作等の非談話行動も微細に記述し、会話との関係进行分析 NII Today No. 62, p. 9の図を改変



図4 今後の展開 日常生活場面の会話コーパスの構築と分析へ

会話研究の推進には、会話のための文法や会話の構造を適切に記述する手法と、それに基づく認知・伝達融合モデルが必要

会話研究に必要な研究用付加情報(アンノテーション)の仕様の策定と、良質なアンノテーションを備えた会話コーパスの整備が不可欠

会話コーパスのアンノテーションに基づく研究を展開

- ① 発話の単位と機能のアンノテーションに関する仕様の策定
- ② 実会話コーパスにアンノテーションを付与
- ③ これらのアンノテーションデータに基づき、実時間的に構成される発話の文法に関する認知・伝達融合モデルを構築
- ④ 単一の発話や音声言語行動を超え、発話連鎖や非言語・非談話行動に関するアンノテーション手法についても検討

基盤となるアンノテーションデータは、様々な研究で有効に利用されており、今後Webページを通じてより広く公開する予定