

NICT 声優対話コーパス v1.1

国立研究開発法人情報通信研究機構 先進的音声技術研究室

2016/06/09

1. 概要:

声優2名による掛け合い収録を行ったものである。台本の製作にあたっては、[京都観光案内対話コーパス（京都観光案内対話データ対話データベース）](#)から、対話のやり取りが活発な21対話を抽出して書き起こしを行い、台本を製作した。声優には、(1)互いに相手の発話に重ならず、(2)対話として自然なスタイルで台本を読むように指示を行った。発話ごとに人手で切り出し、無音部分を除いてある。

- 用途: 対話調の音声合成、ロボット・音声対話システムの評価など
- 話者: 声優
- 話者数: 1名
- 発話内容: 掛け合い対話: 14179発話(約433分)、ロボット対話評価用: 224発話、対話システム評価用: 112発話
- 音声データファイル形式: Wav
- 音声データ形式: 48kHz/16-bit/little endian
- ラベルデータ: 書き起こし、自動生成した読み
- データサイズ(GB): 4.1

2. 引用

本コーパスを使用した場合は、以下の文献を引用してください。

K. Sugiura, Y. Shiga, H. Kawai, T. Misu and C. Hori: "A Cloud Robotics Approach towards Dialogue-Oriented Robot Speech," Advanced Robotics, Vol. 29, Issue 7, pp. 449-456, 2015.

3. 詳細

3.1 タスク

- TEL や SPD から始まるもの: 観光ガイドと旅行者による対面対話を収録した台本をもとに、声優に掛け合い対話を行わせたもの

- ROBOTPROMPT: ロボット評価用 システム発話
- HANNAPROMPT: 観光案内対話システム評価用 システム発話

3.2 マイク

スタンド型のマイクを使用して収録した。

3.3 発話ファイル

以下の3つ組からなる。Xは後述するファイル命名規則に従う。

1. X.wav: 音声ファイル

例: F128_TEL20080205-02G_00100_V01_T01.wav

2. X.txt: X.wav をラベラーが書き起こしたもの。台本とほぼ同じであるが、完全に同一ではない。声優が台本と異なる発話を行うことがあるため(省略、言い間違いなど)。

例: F128_TEL20080205-02G_00100_V01_T01.txt > よろしくお願いたします。

3. X.kna: 自動生成した読み。アクセント核等が誤っている可能性があることに注意。

例: F128_TEL20080205-02G_00100_V01_T01.kna > ヨロシク オネガイイ
タシマ' ス

3.4 ファイル命名規則

話者名_台本名_発話 ID_スタイル_テイク.(wav|txt|kna)

- 話者名 = “F128”(固定値)
- 台本名 = 掛け合い対話用ラベル or ROBOTPROMPT or HANNAPROMPT
- 掛け合い対話ラベル: 「(TEL|SPD)」 「日付」 「セッション」 「役割」 を結合したもの

- セッション: 01 から 06 まで。台本製作の過程で削除されたセッションもある。
- 役割 = G または U (G は観光ガイド、U は観光客)
- 発話 ID = 5 桁の数字
- スタイル = “V01”(固定値)
- テイク = “T01”(固定値)

例: - 掛け合い音声の場合 : F128_TEL20080222-05G_00001_V01_T01.wav
- それ以外 : F128_ROBOTPROMPT_00049_V01_T01.wav

3.5 ディレクトリ構造

ファイルは以下のディレクトリに格納される。

F128/台本名(英大文字)/スタイル/テイク/発話 ID の上三桁の数字/