

日本語方言音声コーパス

1. 日本語方言音声データの概要

日本語方言音声コーパスの構築に使用した方言音声データは科研費基盤研究A（展開）「日本語方言音声の地域差及び方言音声コーパスの設計・構築に関する研究」の援助によるものである。方言音声データはテキスト読み上げと自然談話によって構成されており、DAT に 16bit、48kHz で収録された。日本の主要方言地域をカバーするためにデータの収録地は青森、山形、千葉、愛知、富山、奈良、鳥取、香川、福岡の計 9 地点を選んだ。話者については 60 歳以上の男性および女性とし、その方言の地方で育ち、そこに長い間住んでいる人を選んだ。また方言の地方、話者の数、年齢、性別、発声内容、独話か対話か、録音受け入れ等の問題に付いて考慮し録音は以下の点に留意して行った。

- 1) 音声の内容はテキスト読み上げ及び自然談話を含む。
- 2) 談話の長さは 1 分以上、3 – 5 分が望ましい。
- 3) 方言毎に話者は 10 – 20 人、男女が大体均等になるようにする。
- 4) 録音された音声の書き起こし許可を話者から得る。
- 5) 書き起こしの際、強調部分の記述は特に行わない。
- 6) 話者は 55 才以上、典型的な方言を集める為には 60 才以上が望ましい
- 7) 名前、年齢、性別、生まれた場所、育った場所等のインフォーマントのデータを集める。
- 8) 記録された音声データの使用承諾書を依頼する。
- 9) 録音は出来るだけ静かな部屋で行うことが望ましい。

2. 方言音声コーパスの内容

2. 1 テキスト読み上げ発話

テキスト読み上げ発話とは表 1 に示すテキストを方言話者に提示し、方言話者に日常話しているように発話してもらったものを収録した音声データである。共通語のテキストの方言による音声表現なのでテキスト通りの発話がなされているとは限らない。

表 1 テキスト読み上げ発話の発話内容

発話番号

近隣の人との挨拶における発話

- 1 おはようございます
- 2 ご精がでますね
- 3 ごめんください

- 4 どうぞお上がりください
- 5 こんばんは
- 6 ありがとうございます
- 7 どうもすみません
- 8 さようなら

電話での対応

- 9 もしもし、佐藤さんですか。
- 10 いいえ、違います。山野です。

母音の発音

- 11 あいうえお

数の数え方

- 12 1、2、3、4、5、6、7、8、9、10
- 13 ひ、ふ、み、よ、いつ、む、なな、や、ここのつ、と
- 14 2、4、6、8、10

電話番号の読み上げ

- 15 0467-41-2436
- 16 0298-53-5206
- 17 06-6850-5131
- 18 0426-77-2136
- 19 075-753-7799
- 20 099-285-7557

2. 2 自然談話

方言話者間、または調査者と方言話者で自由に話した発話を収録したものを自然談話と呼ぶ。内容は方言話者の幼少期の話や戦時中の体験談、最近の話など多岐にわたり、内容も収録時間も異なっている。

2. 3 収録データの内訳

収録した音声データの内訳を示す。各音声データは話者1～3名ごとにDATに量子化精度16bit、サンプリング周波数48kHzで収録された。

2. 3. 1 話者数

表2に収録話者数を、表3に収録時間を示す。音声データの総収録時間は約145時間程になった。

表2 方言音声データの話者数（名）

地域	男性	女性	合計
青森県	6	7	13
山形県	6	6	12
千葉県	4	9	13
愛知県	5	6	11
富山県	6	17	23
奈良県	12	9	21
鳥取県	6	9	15
香川県	8	7	15
福岡県	4	6	10
合計	57	76	133

表3 方言音声データの収録時間（分）

地域	収録時間
青森県	924.2
山形県	571.4
千葉県	1059.5
愛知県	549.5
富山県	1395.3
奈良県	2328.9
鳥取県	903.2
香川県	1025
福岡県	550
合計	8757

2. 3. 2 年齢別内訳

全体では71～75歳の話者が最も多く、次いで75～80歳の話者が多くなっている。また60歳以下の話者や90歳以上の高齢の話者もみられる。

地域別で見ると比較的幅広い年齢分布をとっているのが青森県、奈良県、富山県である。特に青森県では91歳という今回収録した中では最も高齢の女性の収録も行われている。また富山県は女性が多く、奈良県は男性が多い。逆に年齢分布が狭いのは千葉県、愛知県で70歳代に話者が集中している。山形県は比較的若い年齢の話者を中心に収録されていて75歳以下の話者で構成されている。香川県、鳥取県では70歳台に話者がやや集中し

ている。

3. 日本語方言音声コーパスの構築

第2章で示した方言音声データから方言音声コーパスを構築する。コーパス中の音声データは 16bit、16kHz の wav ファイルで話者毎にディレクトリに分けられ CD-ROM に格納される。音声コーパスは以下の流れで構築される。

- 1) 音声データの編集
- 2) テキスト書き起こし
- 3) 方言音声コーパスの作成

3. 1 音声データの編集

音声データは二つの部分からなり、一つは共通語のテキストの読み上げ表現であるテキスト読み上げ、もう一つが方言話者間、もしくは収録者と方言話者で自由に話す自由談話である。方言音声データは話者毎に DAT テープに 16bit、48kHz で収録されているのでまず DAT から音声データをそのまま取り込み、必要な部分毎に切り出し最後に 16kHz にリサンプリングを行う。しかし方言音声データの収録は、一般の家庭や公民館、寺社などで行われているため、方言話者と調査者だけでなくそれ以外の様々な雑音も収録されている。その為以下のようにして音声データの編集を行った。

3. 1. 1 テキスト読み上げ

テキスト読み上げ発話については同じ発話を 1～3 回収録しているため、以下の条件を設定し、最も良い状態の音声データについて編集を行った。

- 1) 方言話者の発話に言い誤り、言い詰まりなどが無く良好に収録されているもの
- 2) 調査者の音声が入っていないもの
- 3) その他雑音等が入っていないもの

なお各テキスト読み上げ発話の切り出しの際は前後に 100～300msec の無音区間を設けている。

3. 1. 2 自然談話

自然談話については方言発話者の自然な音声についての分析を行うため、一つのまとまりとして扱い、セグメンテーションは行わない。また談話の連続性を保つため、テキスト読み上げ発話の編集時のような雑音回避は行わない。

3. 2 テキスト書き起こし

発話内容を聞き取り、その内容をテキストに書き取る。以下に自然談話の書き起こしテキストの例を示す。

自然談話書き起こしテキストの例（山形県男性二人の対話）

A) 方言話者

B) 方言話者

A) トコロデ アノー ナンダ ユーリョージンジャノ サンパイ アノケン
ところで あのー なんだ 優良神社の 参拝 あの件

B) アー ア
あー あ

A) アレー オレ ズット ヨサン ミタケドノ キョネンマデ ゴマンエンナケドモ
あれー 俺 ずっと 予算 見たんだけど 去年まで 五万円だったんだけど

B) ウン
うん

A) コドシカラ ロクマンエンニ ナッタケカラ
今年から 六万円に なったから

B) コトシモ アップ ナッタデアナ
今年もアップしてあったか

A) ナッタケカラ アレダ アノー ログニンイッデモ オレ ケッセキスルー
なったから あれだ あのー 六人で行っても 俺 欠席する

B) ドウシテヨ
どうして

A) オレ ケッセキ スッカラー ログニンデ イッテモラエ シテ
俺 欠席 するから 六人で 行ってもらえ そして

B) ダメンダゼー ウン
だめだよー うん

A) タノム アー イッテ
頼む あー いいって

B) ダメヤー ヤッパリ ワー イカネーバヤ ンダバ オレダッテ ズット
だめだー やっぱりあなたも行かなければ そしたら 俺だって ずっと

A) ソレワヤ ホレ ログニンノウヅダハゲ ンダダ
それは ほら 六人の内だったから そうだそうだ

B) イッターンダモシ ンダダ
行ったんだし そうだそうだ

3. 3 方言音声コーパスの作成

16kHz にリサンプリングした wav ファイルを発話者毎のディレクトリに分け、書き起こしテキストを付けて CD に書き込む事で方言音声コーパスが完成する。構築した日本語方言音声コーパスのファイル名とディレクトリ構造は以下の様になる。

1) テキスト読み上げ発話

ディレクトリ：県名/wav/Text/性別+テープ番号/

(例) Toyama/wav/Text/f01/、Toyama/wav/Text/f02/

音声ファイル名：県名_性別+テープ番号_発話番号.wav

(例) 富山県、女性、テープ番号 1、発話番号 1 → Toyama_f01_01.wav

2) 自然談話

ディレクトリ：県名/wav/Danwa/

(例) Toyama/wav/Danwa/

音声ファイル名：県名_性別+テープ番号_発話番号(+L 又は+R).wav

(例) 富山県、女性、テープ番号 1、発話番号 2 1 (独話) → Toyama_f01_21.wav

(例) 富山県、テープ番号 2、発話番号 2 2 (対話)

L チャンネル男性、R チャンネル女性の対話で L チャンネルの方の男性話者→
Toyama_mf02_22L.wav

L チャンネル男性、R チャンネル女性の対話で R チャンネルの方の女性話者→
Toyama_mf02_22R.wav

自然談話が独話でステレオ録音の場合、L と R で明瞭に聞こえるほうを使用した。

3) 書き起こしテキスト

ディレクトリ：県名/Text/(Danwa/または Text/)

(例) Toyama/Text/Text/f01/、Toyama/Text/Danwa/

ファイル名：音声ファイル名.txt

(例) 富山県、女性、テープ番号 1、発話番号 1 → Toyama_f01_01.txt

4) その他

1本のテープで2回テキスト読み上げを行っている場合、1回目と2回目を a、b で区別した。

(例) 奈良県、テープ番号 8、1回目男性、2回目女性

Nara/wav/m08a/ (1回目)

Nara_m08a_01.wav

Nara_m08a_02.wav

.

.

Nara/wav/f08b/ (2回目)

Nara_f08b_01.wav

Nara_f08b_02.wav

.

.

自然談話が対話でステレオ録音の L と R の区別がつかない場合、強制的に一方を L、もう一方を R とした。

愛知県の女性データ 1 名 (テープ番号 2) は、ノイズのためファイルから除いた。

4. まとめ

青森、山形、千葉、愛知、富山、奈良、鳥取、香川、福岡の 9 地点計 133 名からなる日本語方言音声コーパスを構築した。日本語方言音声コーパスの構築は予想以上の大量の音声データ、収録環境による音声データ中の外部雑音等に苦戦の連続であった。なお、本解説は参考文献[6]の一部を加筆・修正したものである。

謝辞

本研究の一部は科研費基盤研究 A (展開)「日本語方言音声の地域差および方言音声コーパスの設計・構築に関する研究」の援助を受けた。

参考文献

- [1] 平山輝男 “日本の方言” 講談社(1998)
- [2] 板橋秀一 “音声コーパス” 情報処理学会誌 38 巻、11 号 pp. 1012-1018 (1997)
- [3] 清水信昭、照沼篤、山本幹雄、板橋秀一 “日本語方言音声コーパスの構築と分析” 日本音響学会春季研究発表会講演論文集、2-6-2, pp. 285-286 (2001.3)
- [4] 嶋田幹貴、橘田裕也、山本幹雄、板橋秀一 “日本語方言音声コーパスの構築と分析” 日本音響学会春季研究発表会講演論文集、2-10-15, pp. 293-294 (2002.3)
- [5] 照沼篤 “日本語方言音声コーパスの構築と分析” 筑波大学第三学群情報学類卒業論文 (2000.2)
- [6] 橘田裕也 “日本語方言音声コーパスの構築とその分析” 筑波大学第三学群情報学類卒業研究論文(2002.2)

研究組織 (所属等は研究実施当時のもの)

研究代表者

板橋秀一 (筑波大学 電子・情報工学系 教授)

研究分担者

土岐 哲 (大阪大学 文学部 教授)

真田信治 (大阪大学 文学部 教授)

工藤真由美 (大阪大学 文学部 教授)

木部暢子 (鹿児島大学 法文学部 教授)

篠崎晃一 (東京都立大学 人文学部 助教授)

中井精一 (富山大学 人文学部 助教授)

桑原尚夫 (帝京科学大学 電子情報科学科 教授)

田中和世 ((独) 産業技術総合研究所 知能システム部門
音声聴覚情報研究グループ長)

壇辻正剛 (京都大学 総合情報メディアセンター 教授)

田中裕司 (筑波大学 現代語・現代文化学系 助手)

石川 泰 (三菱電機 (株) 情報技術総合研究所
音声・言語インタフェース技術部 チームリーダー)