

論文 / 著書情報
Article / Book Information

論題(和文)	読み上げ音声に対する話し言葉音声の言語的特徴の分析
Title(English)	
著者(和文)	中村匡伸, 岩野公司, 古井貞熙
Authors(English)	Masanobu Nakamura, Koji Iwano, SADAOKI FURUI
出典(和文)	日本音響学会2005年秋季講演論文集, Vol. , No. 3-6-10, pp. 347-348
Citation(English)	, Vol. , No. 3-6-10, pp. 347-348
発行日 / Pub. date	2005, 9

読み上げ音声に対する話し言葉音声の言語的特徴の分析*

◎中村匡伸, 岩野公司, 古井貞熙 (東工大)

1 はじめに

話し言葉音声の音響的および言語的特徴の分析は、話し言葉音声の認識性能の向上や、音声合成の品質向上に役立つと考えられ、非常に重要である。我々はすでに、日本語話し言葉コーパス(以下 CSJ と呼ぶ)に収録されている発話タイプの異なる 4 種類の音声データ(自発度の低い順に朗読音声, 学会講演音声, 模擬講演音声, 対話音声)において、各音素の音響的特徴に関する比較を行った [1]。その結果、自発度が高くなると音素空間が縮小し、音素間のマハラノビス距離の平均値が小さくなるという傾向が明らかになった。

本稿では読み上げ音声と話し言葉音声の言語的特徴の違いを分析する。特に、読み上げ音声と話し言葉音声の品詞の統計的特徴の違いを明らかにする。発話タイプの異なる 5 種類のテキストを用い、品詞の出現頻度、および品詞の言語モデルの作成と評価を行う。

2 テキストデータ

分析には、毎日新聞、ニュース解説、学会講演音声、模擬講演音声、対話音声という発話タイプの異なる 5 種類のデータを用いる。毎日新聞のデータ(以下 M と呼ぶ)としては、1995 年から 2001 年の毎日新聞コーパスから無作為に抽出したテキストを用いる。ニュース解説、学会講演音声、模擬講演音声、対話音声のデータ(以下それぞれ B, A, S, D と呼ぶ)としては、CSJ に収録されている書き起こしテキストを性別非依存で無作為に抽出したものをを用いる。テキストは、全ての発話タイプにおいて茶筌(Ver 2.3.3)を用いて形態素解析を行った。解析結果の表記誤りに関しては茶碗(Ver 2.06)を用いて修正した。

3 実験条件

言語モデルの学習に用いたテキストの総単語数と語彙数を表 1 に示す。言語モデルのエントリに登録する際の単語の表記は「漢字表記+カタカナ読み+品詞」とし、0.2 秒以上 0.5 秒未満の無音区間、または読点を<sp>、0.5 秒以上の無音区間、または句点を<sil>とした。また本実験において用いた品詞は 15 種類とした。言語モデルの学習・評価には、Palmkit(Ver 1.0.30)[2]を用いた。

4 品詞の出現頻度

図 1 に各発話タイプにおける品詞の出現頻度の割合を示す。図 1 より、読み上げ音声(書き言葉)と話し言葉音声には品詞の出現頻度に大きな違いがあることが分かる。毎日新聞においては名詞の出現頻度が極めて大きくなり、対話音声においてはフィラーの出現頻度が大きくなる傾向がある。図 2 に毎日新聞と対話音声における品詞の 2-gram の出現頻度の相対値を示す。左図は毎日新聞、右図は対話音声の出現

Table 1 言語モデルの学習に用いたテキストの総単語数と語彙数

	総単語数	語彙数
毎日新聞 (M)	3,153,052	30,000
ニュース解説 (B)	933,980	17,884
学会講演音声 (A)	3,003,512	27,151
模擬講演音声 (S)	3,005,201	30,000
対話音声 (D)	106,595	5,414

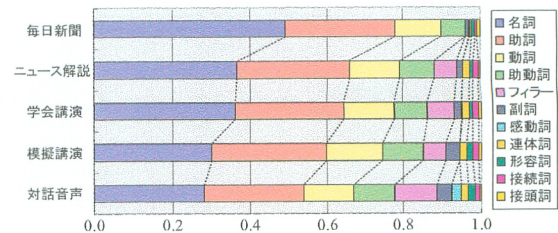


Fig. 1 品詞の出現頻度の割合

頻度を表す。ニュース解説、学会講演、模擬講演における品詞の 2-gram の出現頻度は、ほとんど対話音声と同じである。毎日新聞で出ている 3 つのピークは、大きい順に「名詞-助詞」、「名詞-名詞」、「助詞-名詞」である。対話音声においては「名詞-助詞」の出現頻度は毎日新聞と同様にみられるが、「名詞-名詞」、「助詞-名詞」の出現頻度は小さくなっている。毎日新聞における「名詞-名詞」の出現頻度が高い理由としては、複合名詞の頻出などが考えられる。また、毎日新聞と対話音声の 2-gram の全体的な分布を見てみると、毎日新聞では特定の 2-gram の出現頻度が大きく偏りが大きくなっているが、対話音声では比較的偏りが小さいことが分かる。

5 品詞の言語モデル

4 節により、各発話タイプに品詞の出現頻度による特徴が見られることが明らかになった。本節では品詞の言語モデルの評価を行う。品詞の言語モデルは 3-gram で、平滑化にはバックオフスムージング(Witten-Bell discount)を用いた。表 2 に、発話タイプごとのテストセットパープレキシティ(以下 PP と呼ぶ)を示す。列方向に評価される言語モデルを表し、行方向に使用した評価テキストを表す。各発話タイプの評価テキストの総単語数は約 50,000 語とした。表

Table 2 品詞の言語モデルの PP

評価テキスト	品詞の言語モデル				
	M	B	A	S	D
M	3.29	3.78	3.85	4.15	4.35
B	7.49	4.00	4.13	4.36	4.68
A	9.13	4.98	4.72	4.97	5.14
S	8.12	5.34	5.05	4.98	5.22
D	15.04	9.04	7.17	7.03	5.87

* Analysis of linguistic characteristics of spontaneous speech in comparison with read speech. by NAKAMURA, Masanobu, IWANO, Koji, FURUI, Sadaoki (Tokyo Institute of Technology)

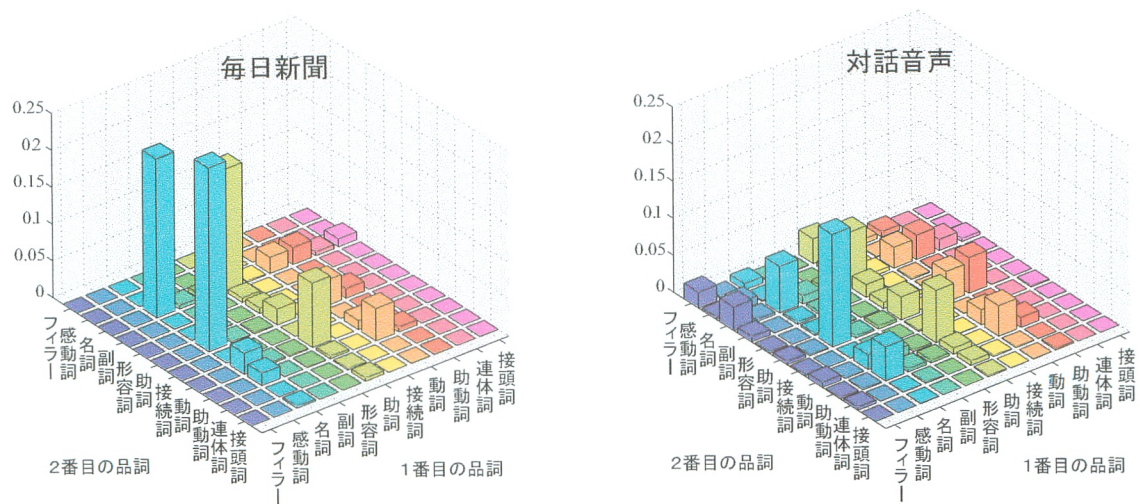


Fig. 2 品詞の 2-gram の出現頻度の相対値

2より, 毎日新聞の言語モデルを各発話タイプのテキストで評価すると, タスクによる PP のばらつきが大きくなること分かる. それに対して対話音声の言語モデルを評価すると, PP のばらつきは比較的小さくなること分かる. 毎日新聞では出現する品詞の N-gram が偏るが, 対話音声ではその傾向が小さいためと考えられる.

6 PP 行列の可視化

5節で算出した PP の行列は, 各発話タイプ間の品詞言語モデルの距離を近似的に表していると考えられる. そこで, (1) 式を用いて表 2 の PP 行列の距離行列化を行い, その距離行列を多次元尺度構成法 (Multidimensional Scaling)[3] により可視化を行う. ここで (1) 式の a_{ij} は PP 行列 A の i 行 j 列の要素, d_{ij} は距離行列 D の i 行 j 列の要素, T は発話タイプの種類 ($T = 5$) である.

$$d_{ij} = \sum_{i=1}^T \sum_{j=1}^T \frac{(a_{ij} + a_{ji}) - (a_{ii} + a_{jj})}{2} \quad (1)$$

図 3 に, 各発話タイプの品詞の言語モデルの距離関係を示す. 図 4 には比較のため, 単語の言語モデルの距離関係を表す. ただし単語の言語モデルの PP 行列は, 各発話タイプで語彙数を 5,400 語に統一して言語モデルの学習を行い, 品詞の言語モデルの評価で用いたテキストと同じ評価テキストを用いて算出した. 図 3, 4 より, 毎日新聞と対話音声との距離は 5 種類の品詞の言語モデルの中で最も大きく, ニュース解説, 学会講演, 模擬講演の 3 つの品詞の言語モデルがその中間に位置しているということが分かる. また品詞の言語モデルと比較して, 単語の言語モデルでは, ニュース解説, 学会講演, 模擬講演の間の距離が大きくなっていることが分かる. 図 3, 4 の横軸は, タスクの自発度に関連した特徴を表していると考えられる.

7 まとめ

本稿では, 毎日新聞, ニュース解説, 学会講演音声, 模擬講演音声, 対話音声という発話タイプの異なる

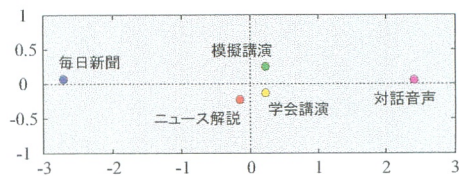


Fig. 3 発話タイプ間の品詞の言語モデルの距離関係

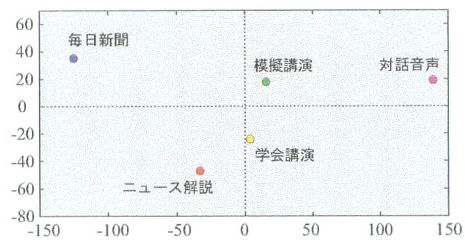


Fig. 4 発話タイプ間の単語の言語モデルの距離関係

データを用いて読み上げ音声と話し言葉音声の言語的特徴の比較を行った. 各発話タイプでの品詞の統計的特徴の違いを調べたところ, 毎日新聞においては, 特定の N-gram の出現頻度が高く, 分布に偏りがあることが明らかになった. 発話タイプ間の品詞の言語モデルの距離関係を調べたところ, 毎日新聞と対話音声との距離が最も大きく, ニュース解説, 学会講演, 模擬講演が, それらの中間に位置することが明らかになった.

今後の課題として, 本実験で明らかになった話し言葉音声の品詞の統計的特徴を利用し, あらゆるタスクに頑健な言語モデル作成の検討を行いたい.

参考文献

- [1] 中村匡伸, 岩野公司, 古井貞照, 日本音響学会 2005 年春季講演論文集, 2-1-4, pp.230-231 (2005-3).
- [2] <http://palmkit.sourceforge.net/>
- [3] 田中良久, 計量心理学, 東京大学出版会, 1969.