

方言音声認識のための話し言葉言語モデル構築

○ 平山 直樹 森 信介 奥乃 博 (京都大学 大学院情報学研究科)

1. はじめに

書き言葉の音声認識：よく研究されている [Anusuya 09]
話し言葉の音声認識：**まだ課題が多い**

話し言葉に影響する個人性

個人性の大きさ

性別 年齢 身分 **地域** ...

方言：本研究のターゲット

方言の特徴

本発表で扱う範囲

語彙

文法

言語的特徴

アクセント

音素の変化

音響的特徴

2. 関連研究

音響特徴による方言同定 [Ching 94; Miller 96; Lyu 06]

- ⊗ 音声コーパス収集の課題：話者の少ない方言は難しい
- ⊗ 語彙変化への対応不可

語彙変換規則による方言機械翻訳 [Zhang 98]

- ⊗ 人手で規則を列挙→拡張性に乏しい

書き言葉→話し言葉変換 [Akita 10; Neubig 12]

- ⊗ 統計ベースの変換
- ⊗ 共通語に限る

本手法：

共通語→方言変換 (Akita らの手法の発展) したコーパスを用いて音声認識

→ 語順変化なしと仮定して変換モデルを単純化

3. 日本語方言音声認識

共通語言語コーパス：BCCWJ [Maekawa 08] など ... 大量の文章が利用可能

対訳コーパス：[国立国語研究所 01-08] ... 少量ならばある

対訳コーパスから得た規則で
共通語言語コーパスを変換

↓

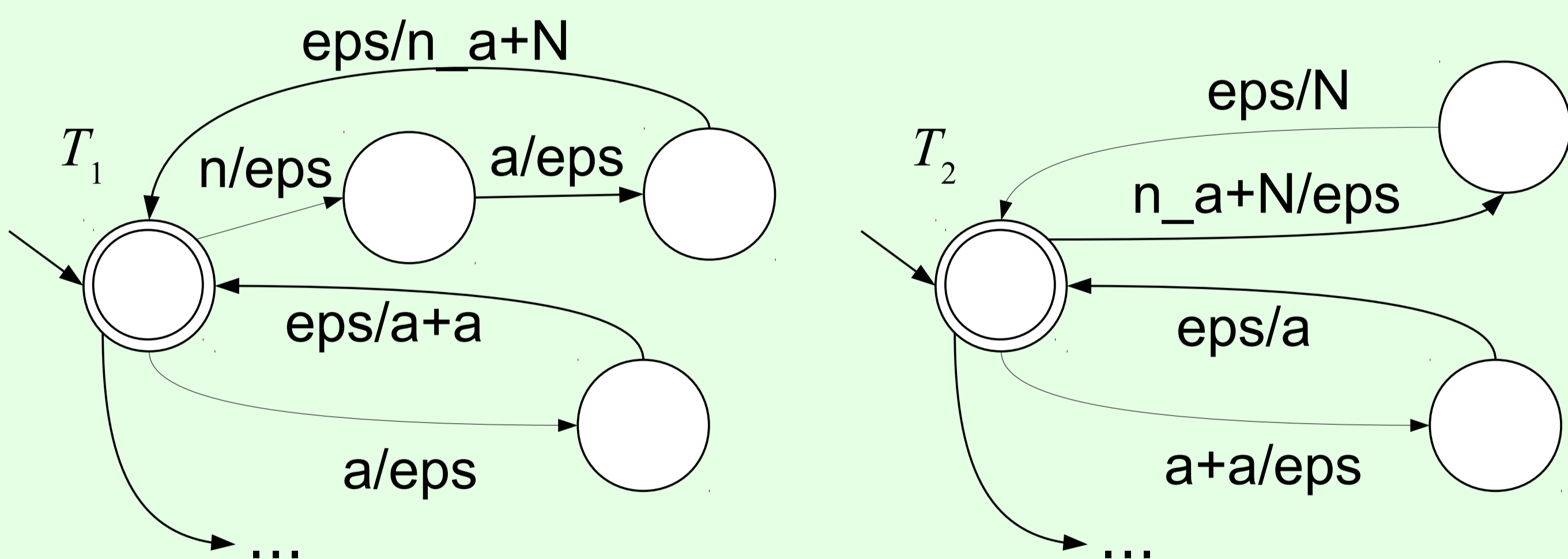
大量の疑似方言コーパスを
シミュレート

手法 1. 音素列 (読み) 変換器

(a) 対訳コーパスから各共通語単語に対する方言単語を抽出
(音素列に対応する DP マッチングの結果を利用)

(b) 重み付き有限状態トランスデューサ (WFST) で
n-gram モデルに基づく統計的規則を学習

共通語: a n a t a w a d o k o n i s u N d e i r u n o
関西弁: a N t a d o k o s u N d e r u N



2. コーパス変換

言語コーパスに方言単語を付与

→ 集計して共通語単語が与えられたもとの条件付き確率を計算

あなたはどこにいっているの

音素列
変換器

あんたどこにいってるん

$P(\text{あんた} | \text{あなた}) = 0.3$
 $P(\text{あなた} | \text{あなた}) = 0.2$

⋮
 $P(\text{NULL} | \text{い}) = 0.4$
 $P(\text{い} | \text{い}) = 0.2$

⋮

各対応単語組をシンボルとする n-gram モデルを学習

4. 実験

近畿方言 (関西弁) テキスト読み上げで単語認識精度 [%] 評価

共通語言語コーパス：BCCWJ Core 53,899 文；約 116 万語

対訳コーパス：大阪・京都・兵庫...計約 2.5 万語

共通語単語に対する、方言単語の条件付き確率の線形補間も実験

言語モデル	話者 1	話者 2	話者 3	話者 4	平均
共通語	53.5	43.4	54.8	43.3	48.8
方言 25%	60.6	50.1	63.6	46.7	55.2
方言 50%	61.4	50.6	63.4	47.6	55.8
方言 75%	61.2	52.2	63.7	47.6	56.2
方言 100%	60.1	49.4	63.9	49.4	55.7

※ 話し言葉起因の表記ゆれチェック→さらに 2pt. 向上

※ 言語コーパスを Yahoo! 知恵袋に変更→さらに 2.5pt. 向上

5. 今後の課題

- 語順変化への対応
→ 単語単位以外の変換方法
- 表記ゆれチェックの自動化
- 複数方言の音声認識
→ 事前情報無しでできれば理想
- 音響モデル
→ 話者間の認識精度の差を縮小したい

6. 結論

疑似方言言語コーパスによる方言音声認識

- 対訳コーパスから抽出した単語対応を利用して共通語コーパスを変換
- 平均 7.4 pt. の単語認識精度向上