

研究タイトル：

話し言葉を対象とした音声言語処理の研究



氏名： 太田 健吾 / OHTA Kengo E-mail: kengo@anan-nct.ac.jp

職名： 助教 学位： 修士(工学)

所属学会・協会： 情報処理学会, 人工知能学会, 日本音響学会

キーワード： 音声認識, 音声対話, e-Learning

技術相談
提供可能技術：
・話し言葉を対象とした音声認識
・自然で親しみやすい音声対話
・音声言語処理を用いた e-Learning

研究内容： 話し言葉を対象とした音声認識のための言語モデル

確率モデルによる音声認識には、音響モデルと言語モデルが必要である。現在の音声認識システムでは、大規模な言語データに基づいて、単語の繋がりが易さや文法的規則を自動学習する統計的言語モデルが非常に重要な役割を果たしている。このため、認識したい対象音声の言語的性質を十分に反映した大規模な言語データを獲得することが、実用レベルの認識精度を達成する必要条件となっている。具体的には、①話題(内容語による表現)と②発話スタイル(機能語による表現)の2点が対象音声と一致した言語データを獲得することが重要である。従って、ニュースキャスターによる番組原稿の読み上げ音声のような、書き言葉スタイルの音声を認識する際には、ニュース関連の話題を含み、かつ、書き言葉スタイルで記述された言語データを新聞や Web から獲得することで、実用的な認識精度を達成できた。

しかし、話し言葉スタイルの音声の認識に用いる言語データは、下図に示すような話し言葉の特徴を含んでいる必要があり、このような言語データが入手できることは実用上極めて稀である。たとえば、国会答弁の音声を認識する際には、衆議院等で公開されている国会会議録のデータが利用できるが、これは国会答弁に関連した話題を多く含む反面、速記者の手によって書き言葉スタイルに整形されているため、下図に示すような話し言葉の特徴を反映した言語モデルを学習することができない。

<話し言葉の特徴>

- (a) 間投詞の問題 : 間投詞(“あの一”、“えっと”といった場繋ぎ語)が発生する。
- (b) ポーズの問題 : 言語的な区切りとは異なった位置に無音(ポーズ)が発生する。
- (c) 言い回しの問題 : 話し言葉特有の言い回しが発生する。

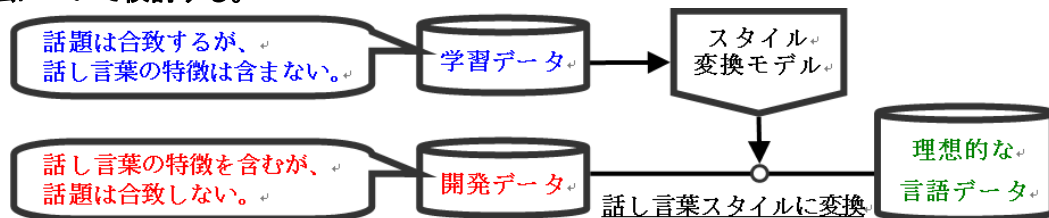
国会会議録 : それでは審議を始めさせていただきます。

C.S.J : えっと、それでは発表を始めさせていただきますけども。

認識したい音声 : えっと、それでは審議を始めさせていただきますけども。

↑ (a) 間投詞の問題
↑ (b) ポーズの問題
↑ (c) 言い回しの問題

そこで本研究では、機械学習の手法を用いて、認識対象と類似な Web 等に存在する既存の書き言葉の言語データから必要な情報だけを抽出し、それらと話し言葉への自動生成法を組み合わせる理想的な話し言葉の言語データを自動生成する手法について検討する。



提供可能な設備・機器：

名称・型番(メーカー)	
話し言葉を対象とした音声認識システム	