音声言語情報処理 18－4
(1997. 10. 24)

発話単位の分割または接合による言語処理単位への変換

竹澤 寿幸　森元 逞
ATR 音声翻訳通信研究所
〒619-02 京都府相楽郡精華町光台 2-2
E-mail: {takezawa, morimoto}@itl.atr.co.jp

あらまし　自然で自発的な発話を対象とする音声翻訳ないし音声対話システムへの入力としての発話単位は文に限定できない。一方、言語翻訳処理における処理単位は文である。話し言葉における文に関して、計算機処理から見て十分な知見は得られていないので、文の代わりに「言語処理単位」と呼ぶことにする。まず、一つの発話単位を複数の言語処理単位に分割したり、複数の発話単位をまとめて一つの言語処理単位に接合する必要があることを、通訳者を介した会話音声データを使って示す。次に、ポーズと細分化された品詞の N-gram を使って、発話単位から言語処理単位に変換できる見通しが得られたことを予備実験により示す。

Transformation into Language Processing Units by Dividing or Connecting Utterance Units

Toshiyuki TAKEZAWA　Tsuyoshi MORIMOTO
ATR Interpreting Telecommunications Research Laboratories
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto-fu 619-02, Japan
E-mail: {takezawa, morimoto}@itl.atr.co.jp

Abstract: The utterance units that serve as input to speech translation and/or spoken dialogue systems that handle spontaneous speech are not always sentences. However, the processing units of language translation are sentences. Since we do not have enough knowledge about the sentences of spoken languages, we use the term "language processing units" instead of sentences. First, using conventionally interpreted dialogue data, we show that utterance units sometimes need to be divided into several language processing units, and sometimes need to be connected to make up a single language processing unit. Next, we propose a method of transforming from utterance units to language processing units based on pause information and the N-gram of fine-grained part-of-speech subcategories. We have conducted preliminary experiments and have confirmed that our method yields good results.

1　まえがき

自然で自発的な発話を対象とする音声翻訳ないし音声対話システムの構築を目指している。読み上げ文を対象とする音声認識研究においては文が処理単位となっている。また、従来の音声翻訳ないし音声対話システムへの入力は、文節区切りのようなゆっくり丁寧に発話された文を単位とする音声であった [10]。しかしながら、自然で自発的な発話を対象とする音声翻訳ないし音声対話システムへの入力としての発話単位は文に限定できない。

一方、言語翻訳処理における処理単位は文である。書き言葉を対象とする自然言語処理システムにおける処理単位も一般に文である。話し言葉を対象とする言語翻訳処理における処理単位も文である [3]。音声対話システムにおける問題解決器のための解釈の処理単位も暗黙の内に文ないし文相当のものを想定していると考えられる。

ところで、本稿では文の定義の議論はしない。例えば、文献 [7] 等に文に関する説明がある。また、話し言葉における文は、無音と韻律に代表される表層のレベル、構造のレベル、意味のレベルで特徴付けられると言われるが、計算機処理から見て十分な知見は得られていない [4]。そこで、本稿では文という術語は使わず、翻訳や解釈のための自然言語処理単位という観点から「言語処理単位」と呼ぶことにする。

まず、2. で一つの発話単位を複数の言語処理単位に分割したり、複数の発話単位をまとめて一つの言語

処理単位に接合する必要があることを、通訳者を介した会話音声データを使って示す。次に、3. でポーズと細分化された品詞の N-gram を使って、発話単位から言語処理単位に変換できる見通しが得られたことを予備実験により示す。最後に 4. で全体をまとめ、今後の展望を述べる。

2 発話単位から言語処理単位への変換の必要性

2.1 音声翻訳システムへの入力としての発話単位

音声翻訳研究のために、日本語話者と英語話者の、通訳者を介した対話を収集し、データベース化している [9]。通訳の質を高めるために、日英方向と英日方向の 2 名の通訳者を介した。「近未来の音声翻訳システム」のための基礎資料を目指しているため、通訳者は 1 回の発話毎に逐次的に通訳を行なう。通訳者が正確に伝えられるために、1 回の発話は 10 秒以内とした。また、相手の話している間に割り込むことは禁止した。ホテル担当者やホテル滞在者であるという設定資料を用意し、それをもとに模擬対話を行なっている。役割や設定をいろいろ変化させた上で、多くの話者に演じてもらい、多様な音声言語現象を収録した。

このようにして集めた会話音声データにおいて、通訳者に渡す単位を音声翻訳システムへの入力としての発話単位とみなすことは妥当である。

2.2 発話単位の分割による言語処理単位への変換

句や節を単位として漸進的に翻訳処理を行なう研究 [8] も開始されているが、現段階では、言語翻訳は文を単位とすることが妥当である [3]。応答の「はい」等を含む感動詞はそれだけで文を構成すると言われている [7]。しかしながら、我々のデータベースでは「はい。それで結構です。」と書き起こされていたり、「はい、それで結構です。」と書き起こされていたりする。そこで、言語処理単位として、まず句点で区切られた節境界に注目する。

まず、一つの発話単位を複数の言語処理単位へ分割する必要がある例を示す。ポーズの長さの情報を [] 内に記す。

(1) お待たせいたしました。[440ms] シングル一泊一万円のお部屋でしたね。

(2) お部屋を調べます。[170ms] しばらくお待ちください。

例 (1) (2) 共に一つの発話単位が二つの言語処理単位で構成されている。300ms を閾値としてポーズ単位 [13] に分割すれば、例 (1) は言語処理単位に分割することができる。しかし、例 (2) は一つのポーズ単位に二つの言語処理単位が含まれることになる。つまり、ポーズ情報のみでは言語処理単位に分割することはできない。したがって、別の手法で発話単位を言語処理単位に分割する必要がある。

2.3 発話単位の接合による言語処理単位への変換

次に、複数の発話単位を一つの言語処理単位に接合する必要がある例を示す。

(3) (a) カードの番号が、[680ms] 五二七九。
(b) 三九二零。
(c) 二四六九。
(d) 零零九八 [410ms] でございますね。

例 (3) はカードの番号を確認する際に、数字の桁毎に区切って通訳者に渡した事例である。(a) (b) (c) (d) は別の発話単位となっており、それぞれに対して通訳が個別に挿入されている。現在の我々のデータベースでは、発話の最後は必ず句点で書き起こす決まりになっているため、(a) (b) (c) (d) の最後に句点が置かれている。

文献 [12] で我々が「箇条発話」と名付けたものは、話者ないし項目の数と内容により、同じ発話単位となったり、別の発話単位に分けられたりする。通訳の単位としては無理に接合しなくとも良いが、何らかの問題解決のための解釈の単位としては接合した方が良い可能性もある。接続することを明示するためには、例えば、例 (3') のように書き起こせば良い。

(3') (a') カードの番号が、[680ms] 五二七九、
(b') 三九二零、
(c') 二四六九、
(d') 零零九八 [410ms] でございますね。

発話の最後に読点を挿入するか、あるいは句点を挿入しないことで、発話単位としては終了していたとしても、言語処理単位としては継続することを表現できる。音声翻訳ないし音声対話システムの枠組みでも、言語処理単位が終了していない発話の最後には読点を挿入するか、あるいは句点を挿入しなければ、発話単位の接合による言語処理単位への変換をインタフェースとして実現することができる。

(4) お待たせいたしました。[1600ms] 洋室は、[1130ms] 一泊二食付き、[450ms] 二万円で、補助ベッドが入ります。

既に述べたように、我々のデータベースでは通訳者に渡す単位を発話単位として音声波形ファイルを切り出している。例 (4) は一つの発話単位として切り出

表 1: 分割または接合による言語処理単位への変換

	閾値より長い無音あり	閾値より長い無音なし
単語・品詞並びの統計モデルが成功する	(A) 句点挿入	(B) 句点挿入
単語・品詞並びの統計モデルが失敗する	(C) 読点挿入	(D) 挿入なし

され、データベース化されている。しかしながら、音声翻訳システムにおいて音声の終端検出を自動的に行なうと、この発話単位は細かく分割される可能性がある。例えば、1秒より長い無音区間を終端とみなせば、例 (4) は三つの発話単位から構成される。その場合、「洋室は、」という発話単位はその次の発話単位と接合した方が良い可能性がある。将来、機械による同時通訳が可能となれば、接合する必要はないが、現時点では接合する必要がある。

(5) シングルの [390ms] シャワー付きのお部屋が・・・

300ms を閾値としてポーズ単位 [13] に分割すると、例 (5) は翻訳のための言語処理単位よりも小さい単位に分割され過ぎてしまう事例である。

3 発話単位から言語処理単位への変換手法

3.1 単語・品詞並びを使った句点相当の節境界検出

英語の自然で自発的な発話を対象とする音声認識結果を構文解析する研究 [6] において、ロバストなパーザの探索空間を削減するために、節境界情報を利用する試みが検討されている。そこでは、確定した音声認識結果全体を入力とし、現在位置の前後2単語 (合計4単語) までの範囲を参照して、その位置の節境界らしさの値を求め、それが閾値を越えれば節境界とみなしている。

そこで、単語・品詞並びを使った句点相当の節境界検出手法を検討する。先行研究 [6] で提案されている推定式を次に示す。• の位置が句点相当の節境界の位置である。その前に二つの単語 w_1w_2 があり、その後に二つの単語 w_3w_4 がある。

$$\tilde{F}([w_1w_2\bullet w_3w_4]) = \frac{C([w_1w_2\bullet]) + C([w_2\bullet w_3]) + C([\bullet w_3w_4])}{C([w_1w_2]) + C([w_2w_3]) + C([w_3w_4])} \quad (1)$$

ここで、$C([w_iw_j\bullet])$ はバイグラム $[w_iw_j]$ の右に句点相当の節境界が現れる回数であり、$C([w_iw_j])$ はバイグラム $[w_iw_j]$ が訓練セットに現れる総数である。他の記号も同様である。

訓練データ中にバイグラムと一緒に現れる句点相当の節境界の回数から求められる個別の頻度 $F([w_1w_2\bullet])$, $F([w_2\bullet w_3])$, $F([\bullet w_3w_4])$ の平均や線形結合よりも効果的であったと報告されている [6]。理由は、十分信頼できる情報を含んでいない低い出現頻度のバイグラムを、他の要因と同じように使わないためである。なお、$F([w_iw_j\bullet])$ は次式で表される。

$$F([w_iw_j\bullet]) = \frac{C([w_iw_j\bullet])}{C([w_iw_j])} \quad (2)$$

$F([w_i\bullet w_j])$ と $F([\bullet w_iw_j])$ も同様に求められる。

$\tilde{F}$ の値が閾値より大きければそこを句点相当の節境界とする。閾値の値は人手で設定し、訓練セットに対して最良の性能が得られるように調整する。

文献 [6] では単語のみ検討しているが、我々は日本語を対象とし、次の3通りの組合わせを調べる。

1. 品詞のみ
2. 品詞・活用形・活用型 (プレターミナル [13] と同等)
3. 表層表現・品詞・活用形・活用型 (単語と同等)

さらに、それぞれに対して、現在位置の前後2単語 (合計4単語) の範囲を参照する場合 (式 (1)) と、前2単語と後1単語の合計3単語の範囲を参照する場合 (次式 (3)) を検討する。

$$\tilde{F}([w_1w_2\bullet w_3]) = \frac{C([w_1w_2\bullet]) + C([w_2\bullet w_3])}{C([w_1w_2]) + C([w_2w_3])} \quad (3)$$

3.2 分割または接合による言語処理単位への変換

先行研究 [6] では長い発話単位を分割することのみを検討していた。我々は分割のみならず接合による言語処理単位への変換も検討する。単語・品詞並びを使った句点相当の節境界検出のための統計モデルとポーズ情報を組み合わせた手法を考える。

表1において、(A) と (B) には句点「。」を挿入する。(D) には何も挿入しない。(C) は箇条発話が相当し、扱いが難しいが、原則的には読点「、」を挿入する。

ATR 音声言語データベース (SLDB) [9] の 618 会話から音声認識と言語翻訳を接続する評価実験用のホテル予約9会話を選択した。日本語話者が客の役割を務めているものが4会話、日本語話者がホテル担当者を務めているものが5会話である。そのホテル予約9会話には 166 ターン (発話権の交代)、216 発話あった。内容を確認したところ、体言止めの箇条発話は含まれていたが、発話単位を接合して言語処理単位に変換す

表 2: 粒度および参照する範囲の違いの比較

	条件	品詞のみ		品詞・活用形・活用型		単語	
	閾値	再現率	適合率	再現率	適合率	再現率	適合率
前後 2 単語	0.10	87.9%	24.8%	96.7%	32.4%	96.7%	31.9%
前 2 単語と後 1 単語	0.10	86.2%	26.7%	96.7%	39.9%	92.7%	41.6%

る必要のある事例はその 9 会話には含まれていなかった。つまり、例 (3) ないし (3') のような発話の仕方はまれである。

一方、その 166 ターン、216 発話の書き起こしテキストに含まれる句点の数は 289 個であった。発話の最後は必ず句点で書き起こす決まりになっているので、ターンの途中にある句点の数は 123 個、発話の途中にある句点の数は 73 個である。つまり、通訳者へ渡す発話単位を分割する必要のある事例の頻度は多い。

発話単位の最後が言語処理単位の最後になっているのか、それとも、言語処理単位としてはまだ継続するのかに関して、単語・品詞並びの統計モデルおよび韻律的な特徴等により判断すれば、発話単位の接合による言語処理単位への変換をインタフェースとして実現できる。

しかしながら、発話単位を接合する必要のある事例は少ないので、以下では、まず、発話単位を分割する手法について予備実験を行ないながら検討する。

3.3 発話単位の分割に関する予備実験

3.3.1 準備

評価実験用のホテル予約 9 会話以外の 609 会話を訓練に用いた。訓練は発話権の交代 (ターン) を単位として行なった。箇条発話は話者ないし項目の数と内容により同じ発話単位となったり、別の発話単位となったりするため、その影響を除くことを意図した。ターンの始めには開始記号を挿入し、ターンの終りには終了記号を挿入した。発話単位の開始と終了の情報は使わなかった。書き起こしテキストの句点をそのまま句点相当の正しい節境界とみなした。

3.3.2 書き起こしテキストを用いた実験結果

書き起こしテキストを用いた予備実験を行なった。句点と読点を除いた形態素列を入力とした。訓練時と同様に、発話単位の情報は使わず、発話権の交代 (ターン) 毎に一つの入力単位とした。ターンの途中にある句点 123 個が評価対象となる。書き起こしテキストの句点を正解として、再現率と適合率を求め、評価する。その際、結果を三つに分類する。

表 3: 最適な閾値に基づく再現率と適合率

	条件	品詞・活用形・活用型	
	閾値	再現率	適合率
前後 2 単語	0.37	80.5%	64.7%
前 2 単語と後 1 単語	0.43	88.6%	65.7%

1. 句点相当の節境界で成功する: 正解 [○]
2. 句点相当の節境界で失敗する: 誤り [×]
3. 句点相当の節境界ではない場所で成功する: 湧き出し誤り [※]

$$\text{再現率} = \frac{○}{○ + ×} \quad (4)$$

$$\text{適合率} = \frac{○}{○ + ※} \quad (5)$$

まず、閾値を 0.10 にそろえ、粒度および参照する範囲の違いの比較・検討を行なった。結果を表 2に示す。

粒度の違いについては、品詞・活用形・活用型の場合が最も良い結果となった。文献 [13] においても、品詞では粒度が荒らすぎ、単語では被覆率の観点で良くなかったため、妥当な結果と考えられる。また、参照する範囲については、前 2 単語と後 1 単語の方が前後 2 単語 (合計 4 単語) よりも良かった。

そこで、品詞・活用形・活用型の並びに関して、前後 2 単語を参照する場合と、前 2 単語と後 1 単語を参照する場合について、さらに最適な閾値を探してみた。結果を表 3に示す。

やはり、前 2 単語と後 1 単語の品詞・活用形・活用型の並びを利用した場合が最も良い。誤りおよび湧き出し誤りの内容を次に示す。あらかじめ要約すると、その分析内容も、前 2 単語と後 1 単語の範囲を見れば十分であることを示唆している。

3.3.3 誤りの分析

閾値を 0.43 として、前 2 単語と後 1 単語の品詞・活用形・活用型の並びを利用した場合の誤りは 14 件あった。その内容を分析する。

発話の途中の感動詞の直後が2件あった。発話の途中の感動詞の直後は読点で書き起こされることが多いためである。対策としては、感動詞の直後に接続助詞が続かない限り[1]句点相当の節境界とするというヒューリスティックスが考えられる。次に例を示す。行の先頭の「×」記号は誤り例を意味する。「+」記号は単語の区切り位置を示す。[]記号の中にポーズの長さや発話単位等の情報を加えた。「•」記号が現在位置を示す。

× 様 + ありがとうございました [60ms] • また

接尾辞の直後が5件あった。そのうちの3件は別の発話単位となっている。同じ発話単位に含まれるものは2件あり、そこには285msと350msのポーズがあった。例を示す。

× 千 + 円 [発話単位終了] • 和室
× 鈴木 + 様 [285ms] • それでは

名詞類の直後が2件あった。そのうち1件は別の発話単位となっている。同じ発話単位に含まれる1件については、615msのポーズが挿入されていた。例を示す。

× 零 + 零 [発話単位終了] • ご
× ご + 滞在 [615ms] • 零

接続助詞の直後が5件あった。1秒程度以上の長いポーズが挿入されるか、発話単位が終わらない限り、接続助詞の直後は読点で書き起こされているためと考えられる。そのうち4件は別の発話単位となっている。同じ発話単位に含まれる1件については990msのポーズが挿入されていた。例を示す。

× す + が [発話単位終了] • 予約
× た + もんですから [990ms] • あ

箇条発話の扱いを除けば、若干のヒューリスティックスを導入したり、ポーズとの関係を調べることで対処可能なものである。

3.3.4 涌き出し誤りの分析

閾値を0.43として、前2単語と後1単語の品詞・活用形・活用型の並びを利用した場合の涌き出し誤りは57件あった。その内容を分析する。

発話の先頭の感動詞の直後が45件あった。発話の先頭の感動詞の直後は句点で書き起こされていることが多いためである。これらの事例は句点とみなしても構わない。次に例を示す。行の先頭の「※」記号は涌き出し誤り例を意味する。他の記号は同様である。

※ + はい [640ms] • いつ
※ + はい [110ms] • そう

終助詞の直後の涌き出し誤りが7件あった。これらもすべて句点とみなしても構わない。例を示す。

※ す + か [590ms] • じゃあ

その他の事例が5件あった。すべて頻度のまれな個別的な事例であった。対策としては、助動詞終止形と終助詞の間や、感動詞の直後に接続助詞が続く場合は句点相当の節境界とはしない等のヒューリスティックスが考えられる。例を示す。

※ し + た • っけ
※ 大変 + 申し訳ございません • が

若干のヒューリスティックスを導入することで対処可能な事例を除けば、ほとんどすべてが句点相当の節境界とみなして構わないものであった。

3.3.5 ヒューリスティックス導入の効果

涌き出し誤り57件のうち、発話の先頭の感動詞の直後45件と終助詞の直後7件の合計52件については、句点相当の節境界とみなして良い。そこで、それらはすべて句点を正解とみなし、かつ、妥当なヒューリスティックスを導入して、再現率と適合率を求めた。結果を表4に示す。再現率、適合率ともに改善できた。箇条発話の扱いが今後の課題である。

3.4 音声認識結果への適用実験

文献[11]の音声認識器の結果を用いて、句点相当の節境界を検出する実験を行なった。書き起こしテキストによる評価実験を行なったホテル予約9会話を対象とした。書き起こしテキストを用いた予備実験では発話権の交代(ターン)毎に一つの入力単位としたが、音声認識結果を対象とする場合は発話単位を一つの入力単位とした。第1位候補に対する例を示す。

- 書き起こし お待たせいたしました。申し訳ございません。シングルは満室となっております。
- 認識結果 お+待/た/し/いた+し+ま+し+た○/申し訳ございません○/十/五/満室/に+な+っ+て+お+り+ま+す○

認識結果の「/」記号は音声認識で使っている単語辞書の区切りを表す。認識結果の「+」記号はデータ

[1]「涌き出し誤りの分析」に例を示す。

表 4: ヒューリスティックス導入の効果

	条件			品詞・活用形・活用型	
	閾値	句点の追加	ヒューリスティックス	再現率	適合率
前2単語と後1単語	0.43	なし	なし	88.6%	65.7%
前2単語と後1単語	0.43	あり	なし	92.0%	97.0%
前2単語と後1単語	0.43	あり	あり	97.7%	99.4%

ベースの形態素辞書の区切りを表す。「○」は検出できた句点相当の節境界のうち、正解とみなせるものを示す。

- 書き起こし [んー] ちょっと高いですね。もっと安い部屋は無いですか。
- 認識結果 二※ / ちょっと / 高 / い / で+す+ね○ / オー / で+す※ / いや / な / い / で+す+か○

書き起こしの [んー] は間投詞を表す。「※」は涌き出し誤りを示す。音声認識で使っている単語辞書では、話し言葉の文末表現に相当するものを一つの長い単位で扱うことが多いため、文末表現の位置に誤認識が少ない。定量的な評価は今後の課題である。

4 むすび

細分化された品詞並びの統計モデルとポーズ情報を用いる句点相当の節境界を検出する手法を提案し、音声認識結果に適用する実験を行なったところ、良好な結果を得た。箇条発話の扱いが今後の課題である。パワーの変化や韻律情報、さらに音韻の継続時間長等を組み合わせることも今後の課題である。次発話予測の研究 [5] で用いている発話タイプと関連付ける研究にも発展させていく。その際、韻律情報を用いた発話タイプの識別 [1, 2] も考慮する予定である。

また、音声認識過程で構文規則を利用する研究 [13] と組み合わせれば、統語構造の情報を併用することも可能である。そこで用いている部分木 [13] と同時通訳方式の実現に向けた処理単位 [8] との関係も調べる予定である。なお、言語翻訳知識を利用して音声認識候補からもっともらしい部分を見つける研究 [14] も行なわれているが、そこでの入力となる言語翻訳単位はあくまで文である。したがって、文献 [14] と組み合わせる場合であっても、本稿で提案した処理とメカニズムは必須となる。

謝辞

実験に協力いただいた大槻直子、林輝昭 両氏に感謝いたします。

参考文献

[1] 藤尾茂, ニックキャンベル, 樋口宜男: "韻律を用いた発話タイプの識別", 日本音響学会平成 7 年度秋季研究発表会講演論文集, 1-1-1, Vol. I, pp. 199–200 (1995-09).

[2] 藤尾茂, ニックキャンベル, 樋口宜男: "韻律を用いたテキスト非限定型発話アクト識別方法", 日本音響学会平成 8 年度春季研究発表会講演論文集, 1-4-14, Vol. I, pp. 245–246 (1996-03).

[3] 古瀬蔵, 美馬秀樹, 山本和英, Michael Paul, 飯田仁: "多言語話し言葉翻訳に関する変換主導翻訳システムの評価", 言語処理学会第 3 回年次大会発表論文集, pp. 39–42 (1997-03).

[4] 石崎雅人, 飯田仁: "音声翻訳システムにおける文", 日本語学, Vol. 15, pp. 71–79 (1996-08).

[5] 巌寺俊哲, 竹澤寿幸, 石崎雅人, 森元逞: "次発話予測による音声認識結果の再順序付け", 情報処理学会第 53 回 (平成 8 年後期) 全国大会, 7N-5, Vol. 2, pp. 355–356 (1996-09).

[6] Alon Lavie: "*GLR*: A Robust Grammar-Focused Parser for Spontaneously Spoken Languages*," School of Computer Science, Carnegie Mellon University, CMU-CS-96-126 (May 1996).

[7] 益岡隆志, 田窪行則: 基礎日本語文法 — 改訂版 —, くろしお出版, 東京 (1992).

[8] 美馬秀樹, 古瀬蔵, 飯田仁: "同時通訳システムの実現に向けた漸進的翻訳処理", 情報処理学会第 54 回 (平成 9 年前期) 全国大会, 4B-6, Vol. 2, pp. 11–12 (1997-03).

[9] T. Morimoto, N. Uratani, T. Takezawa, O. Furuse, Y. Sobashima, H. Iida, A. Nakamura, Y. Sagisaka, N. Higuchi, and Y. Yamazaki: "A speech and language database for speech translation research," *Proc. of ICSLP '94*, pp. 1791–1794 (1994-09).

[10] 森元逞, 田代敏久, 竹澤寿幸, 永田昌明, 谷戸文廣, 浦谷則好, 鈴木雅実, 菊井玄一郎: "音声翻訳実験システム (ASURA) のシステム構成と性能評価", 情報処理学会論文誌, Vol. 37, No. 9, pp. 1726–1735 (1996-09).

[11] 清水徹, 山本博史, 政瀧浩和, 松永昭一, 匂坂芳典: "大語い連続音声認識のための単語仮説数削減", 信学論 *D-II*, Vol. J79-D-II, No. 12, pp. 2117–2124 (1996-12).

[12] 竹沢寿幸, 田代敏久, 森元逞: "音声言語データベースを用いた自然発話の言語現象の調査", 人工知能学会研究会資料, SIG-SLUD-9403-3, pp. 13–20 (1995-02).

[13] 竹澤寿幸, 森元逞: "部分木に基づく構文規則と前終端記号バイグラムを併用する対話音声認識手法", 信学論 *D-II*, Vol. J79-D-II, No. 12, pp. 2078–2085 (1996-12).

[14] 脇田由実, 河井淳, 飯田仁: "意味的類似性を用いた音声認識正解部分の特定法と音声翻訳手法への応用", 人工知能学会研究会資料, SIG-SLUD-9603-2, pp. 7–12 (1997-01).