音声対話における実例に基づく未知語属性推定

高橋 康博　堂坂 浩二　相川 清明

NTT コミュニケーション科学基礎研究所
〒 243-0198 神奈川県厚木市森の里若宮 3-1
tkhs@atom.brl.ntt.co.jp

あらまし

音声対話システムがユーザとの対話を通して未知語知識を獲得することにより効率的な対話を実現することを狙いとして, ユーザ発話中の未知語の属性推定方法を提案する．この方法では, システム発話と直後のユーザ発話の理解結果からなる実例を蓄積し, 現在のユーザ発話の理解結果に未知語が現れた場合は, その理解結果に直前のシステム発話を加えたデータと最も類似度の高い実例を選び, それに基づいて未知語の属性推定を行う．評価実験は, 気象情報案内を行う音声対話システムとユーザとの対話から実例を収集し, ユーザ発話の理解結果の一部を未知語と仮定して未知語を含むユーザ発話を作ることにより行った．実験の結果, 69 % の属性推定率を得た.

キーワード: 音声対話システム, 未知語, 属性推定, 知識獲得, 効率的対話

An Example-Based Method for Estimating Attributes of Out-of-Vocabulary Words for Spoken Dialogue Systems

Yasuhiro Takahashi　Kohji Dohsaka　Kiyoaki Aikawa

NTT Communication Science Laboratories
3-1 Morinosato-wakamiya, Atsugi, Kanagawa, 243-0198 Japan
tkhs@atom.brl.ntt.co.jp

Abstract

This paper proposes a new example-based method for estimating attributes of out-of-vocaburaly (OOV) words in user utterances through natural conversation without explicit Q/A. This method is advantageous because there are many cases in which the system can continue a dialogue if only attributes of OOV words are acquired. We evaluated the proposed method using examples collected from dialogues between the weather information system and the user, and making the user utterance including OOV words by assuming a part of the system's understanding of the user utterance to be OOV words. The proposed method achieved attribute estimation accuracy of 69 %.

Key Words: spoken dialogue system, out-of-vocabulary word, attribute estimation, knowledge acquisition, efficient dialogue

1 はじめに

音声対話システムとは, 人とコンピュータが音声対話を通して, 決められた仕事 (タスク) を遂行するシステムである[1]. 音声対話システムが保有する知識には制限があるが, ユーザはシステムの知識の範囲を知っているとは限らない. したがって, ユーザがシステムにとって未知の語を含む発話をすることは避けられない. 未知語の処理を行わないシステムでは, 未知語を既知語に置きかえて認識するために, タスクを完了できないだけでなく, 対話が著しく非効率になる. したがって, 効率的な対話によりタスクを完了するシステムの実現には, ユーザ発話に含まれる未知語を検出し, タスク完了に未知語知識獲得が必要かどうかを判断し, 必要に応じて未知語知識獲得を行うことが重要となる. 未知語知識獲得とは, 未知語とシステムの知識を適切に関連付けることである.

本稿では, 未知語知識として未知語の属性に注目する. システムが未知語の属性を獲得するためには, 未知語の属性についてユーザに質問するという方法が考えられる. しかし, ユーザ発話に未知語が検出される度に質問していたのでは, 無駄な質問が増加し, 対話が非効率になってしまう. そこで, 本稿では, 明示的な対話をすることなく未知語の属性を推定する方法を提案する. 未知語の属性推定結果を利用することにより, システムは, タスク完了の必要に応じて未知語の属性についてユーザに質問することができるので, 無駄な質問を減らすことができる. 未知語の属性推定を行うためには, ユーザ発話に含まれる未知語を検出する機能が必要となるが, 本稿では, 未知語検出機能をもつと仮定した音声対話システムを考える.

ユーザが未知語を含む発話をした場合でも, 未知語の属性推定により, 効率的な対話を実現できる場合がある. 例えば, ビデオ録画予約システムにおいて, 次のような対話が可能になる.

ユーザ	:	「ザ・サタデーを予約して」
		(ザ・サタデーは未知語)
属性推定	:	『ざさたでー』＝番組
システム	:	「『ざさたでー』は番組名ですね」
ユーザ	:	「はい, そうです」
システム	:	「何時からの番組ですか」

従来では, ザ・サタデーという未知語を既知語に置きかえて認識するため, システムはユーザに何度も番組名を質問することになり, 非効率な対話を続けていた. しかし, 上の例では, システムが『ざさたでー』の属性を番組であると推定し, 番組はタスクの完了に必要な属性なので, 『ざさたでー』の属性をユーザに確認している. そして, システムは, 番組名について何度もユーザに質問せず, 番組の始まる時間をユーザに質問することにより, ユーザが予約したい番組の候補を絞り込もうとしている.

本稿では, 実例に基づいて未知語の属性推定を行う方法について述べる. この方法は, システム発話と直後のユーザ発話からなる実例をシステムが蓄積し, ユーザが未知語を含む発話をした場合は, そのユーザ発話と直前のシステム発話からなる組に最も類似する実例を選び, その実例に基づいて未知語の属性推定を行うことを基本とする.

自然言語処理において, 実例に基づく方法の適用例がいくつか知られている. 例えば, [3] は機械翻訳への適用例であり, [4] は語彙知識獲得への適用例である.

未知語知識獲得の研究として [5] がある. [5] では, 離散発声された単語列を入力としてシステムは行動を出力し, その行動に対するユーザの反応を基に, 単語を学習する方法を提案している. しかし, 我々の目指している未知語知識獲得は, 自然な対話によるものであり, ユーザに離散発声を強いるようなものではない.

音声対話における未知語の属性推定を行っている研究として [2] がある. [2] では, 未知語を含むユーザ発話の未知語部分以外の理解結果だけから未知語の属性を推定している. [2] の方法は, 単一の文から得られる文脈だけを利用しているという問題があるが, 本稿で提案する方法は, 単一の文だけでなく未知語が検出される直前のユーザ発話の理解状態やシステム発話内容を利用している.

実例に基づく未知語属性推定の従来方法として [1] がある. [1] における類似度は, ユーザが未知語を含む発話をした直前のシステム発話と未知語の前後の句に基づいている. しかし, 本稿における類似度は, [1] で扱われている情報に加えて, システム発話が行われたときのスロット状態を考慮したものである. また, 実例に基づく方法では, 音声認識誤りを含む実例を利用することにより属性推定精度が低下するという問題があるが, [1] の方法では, このような実例の質の問題を扱っていない. 本稿では, 実際のユーザ発話の書き起こしを利用することにより, 実例が音声認識誤りを含むかどうかを考慮した属性推定方法を提案する.

2 実例に基づく未知語属性推定

本稿における実例について説明する. ユーザとシステムの対話において, システムは, 現在のユーザ発話の理解状態 (スロット状態) に応じて発話し, そのシステム発話に対するユーザ発話によりスロット状態を変化

[1] 本稿で対象とする音声対話システムは, ユーザ発話の理解状態をスロットと値の組のリストで表現するものとする.

させ, 変化したスロット状態に応じてまた発話をする. このときの

スロット状態 → システム発話 → スロット状態変化

というユーザ発話の理解状態の変化とシステム発話からなる 1 組のデータを局所的対話遷移と呼ぶ. そして, スロット状態変化を引き起こした実際のユーザ発話の書き起こしと局所的対話遷移からなるデータを実例とする.

ビデオ録画予約システムを例にとり, 提案方法を適用してみる. システムは実例を蓄積しているとし, 現在のユーザとシステムの対話における局所的対話遷移が次のようになっているとする (例 1).

スロット状態 = (日付 : 今日)
システム発話 = 「番組名は何ですか」
スロット状態変化 = (人：松嶋たか子, 未知語)

すなわち, 日付スロットの値が「今日」であるとき, システムが「番組名は何ですか」と発話し, そのシステム発話に対するユーザ発話を理解した結果, 松嶋たか子という人名が理解され, その後に未知語が検出されたとする. このときシステムは, 例 1 と最も類似度の高い実例を蓄積された実例の中から探す. 具体的には, 次の条件を同時に満たす実例を探す.

- 日付スロットが値をもっている
- システムは番組名について質問している
- ユーザ発話を理解した結果, 人名が理解された後に未知語が検出されている
- スロット状態変化が音声認識誤りによるものでない

スロット状態変化が音声認識誤りによるものでないこと (スロット状態変化の信頼度) は, 実例の中のスロット状態変化と書き起こしを比較することにより判断する. 蓄積された実例の中から次のような実例が選ばれたとする.

スロット状態 = (日付 : 明日)
システム発話 = **「番組名は何ですか」**
スロット状態変化 = **(人：森田毅，番組：笑道)**
書き起こし = **「森田毅が出演する笑道です」**

このとき, 例 1 に含まれる未知語は, 位置的に実例の中の「笑道」に対応するので, 未知語の属性は, 「笑道」と同じ「番組」であると推定する.

3 類似度と信頼度の定義

提案方法を実現するためには, 次の類似度と信頼度の定義が必要である.

- 局所的対話遷移間の類似度 Sim
- スロット状態変化の信頼度 Rel

この節では, これらの定義の詳細について述べる.

3.1 局所的対話遷移間の類似度

局所的対話遷移は, スロット状態, システム発話, スロット状態変化からなるデータなので, これら 3 つの要素を順に並べたリストで局所的対話遷移を表すことにする.

2 つの局所的対話遷移 $A1$, $A2$ に対し, それらの類似度を各要素間の類似度の和とする. すなわち, スロット状態間の類似度, システム発話間の類似度, スロット状態変化間の類似度を計算する関数をそれぞれ Sim_1, Sim_2, Sim_3 とし, リスト L の第 i 成分を L_i と表すとき, 局所的対話遷移間の類似度 Sim を次のように定義する.

$$Sim(A1, A2) = \sum_{i=1}^{3} Sim_i(A1_i, A2_i).$$

以下で, Sim_1, Sim_2, Sim_3 の定義を述べる.

3.1.1 スロット状態間の類似度

ユーザ発話の理解状態を表すスロット状態は, スロットと値の組のリストである. 例えば, ユーザの予約したい番組が「今日の森田毅が出演する番組」であるとシステムが理解した場合は, (日付：今日, 人：森田毅) と表し, 「明日の番組」であると理解した場合は, (日付：明日) のように表す. スロット状態間の類似度を計算するためにスロット状態間で比較する情報は, どのスロットが値をもっているかということのみとし, スロットがもっている値は比較しない. 例えば, 上の 2 つのスロット状態間の類似度を計算するとき, 日付スロットが共に値をもっていることは考慮するが, 日付スロットの値が一方は「今日」で他方は「明日」であることは考慮しない. これは, ユーザ発話を単語の列ではなく, 属性の列として扱うためである.

以下では, 値をもっているスロットの集合でスロット状態を表すことにする. 2 つのスロット状態 $SL1$, $SL2$ に対し, $SL1$, $SL2$ の両方に含まれるスロットの個数を $SL1$, $SL2$ の類似度とする. すなわち, Sim_1 を次のように定義する.

$$Sim_1(SL1, SL2) = |SL1 \cap SL2|.$$

例えば, (日付：今日, 人：森田毅) と (日付：明日) の類似度は, 日付スロットが共に値をもっているので 1 となる.

3.1.2 システム発話間の類似度

システム発話として扱うのは, ユーザに対する質問またはユーザに対する確認のみとする. システム発話間の類似度を計算するために比較する情報は, まず, 質問と確認の区別である. この区別が一致しないシステム発話間の類似度は 0 とする.

比較するシステム発話が共に質問の場合は, 質問の焦点が当たっているスロット (焦点スロット) と焦点スロットが発話される順番について比較し, さらに, システム発話に含まれる語を値としてもつスロット (非焦点スロット) と非焦点スロットが発話される順番について比較する.

比較するシステム発話が共に確認の場合は, 確認の焦点が当たっているスロット (焦点スロット) と焦点スロットが発話される順番について比較する.

システム発話を次のように表すことにする.

(焦点スロットのリスト, 非焦点スロットのリスト)

ただし, システム発話が確認の場合は, 非焦点スロットのリストは空リストと考える. 例えば, 「いつの番組ですか」という質問と「森田毅が出演するいつの番組ですか」という質問は, それぞれ, ((日付), nil), ((日付), (人)) と表す.

システム発話間の類似度は, 焦点スロット同士, 非焦点スロット同士を比較し, 一致している個数を基にする. ただし, 同一のシステム発話間の類似度が最大となるようにするため, 一方のシステム発話内容が他方のシステム発話内容の拡張になっている場合は, 拡張になっている部分の個数に応じて, 類似度を低くする.

システム発話間の類似度を定義するため, 2 つの補助関数 Aux_1, Aux_2 を準備する. Aux_1 はスロットの 2 つのリストに対し, 先頭同士, 2 番目同士と順に比べて等しい個数を数える関数である. ただし, 途中で等しくない組合せが現れるまでとする. Aux_2 はスロットの 2 つのリストに対し, 一方が他方の拡張になっていれば, 2 つのリストの長さの差の絶対値を計算する関数である.

Aux_1 により焦点スロットが一致している個数と非焦点スロットが一致している個数を数え, 一方のシステム発話内容が他方のシステム発話内容の拡張になっている場合は, Aux_2 により拡張になっている部分の個数を数える.

2 つのシステム発話で, 共に質問または共に確認である $SY1$, $SY2$ に対し, Sim_2 を次のように定義する.

$$\begin{aligned} Sim_2(SY1, SY2) &= \sum_{i=1}^{2} s_i \times Aux_1(SY1_i, SY2_i) \\ &- s_3 \times \sum_{j=1}^{2} Aux_2(SY1_j, SY2_j). \end{aligned}$$

ただし, s_1, s_2, s_3 は定数で, それぞれ, 焦点スロット, 非焦点スロット, 拡張の度合を類似度に強調させる程度を表す. また, Sim_2 の計算結果が負になった場合は, Sim_2 の値を 0 とする.

例えば, ((日付), nil) と ((日付), (人)) の類似度は, 焦点スロットが一致しているが, 一方のシステム発話だけが非焦点スロットに人スロットを含むので, $s_1 - s_3$ となる.

3.1.3 スロット状態変化間の類似度

ユーザ発話の理解状態の変化を表すスロット状態変化は, システム発話直後のユーザ発話を理解することにより値の変化したスロットを変化した順に並べたリストである. 例えば, 日付スロットが値をもっているスロット状態におけるシステム発話に対し, 直後のユーザ発話を理解した結果, 人スロットが森田毅という値をもち, その後, 番組スロットが笑道という値をもった場合, (人, 番組) と表す. 人スロットのみが値をもった場合は, (人) と表す. これは, ユーザ発話を単語の列ではなく, 属性の列として扱うためである.

スロット状態変化間の類似度を計算するためにスロット状態変化間で比較する情報は, システム発話直後のユーザ発話を理解することにより値の変化したスロットと変化した順番である.

ユーザ発話中の未知語は, スロット状態変化の中の未知語スロットで表すことにする. Aux_1 や Aux_2 により, 未知語スロットが他のスロットと比較された場合は, どのスロットとも等しいと判定されることにする. これは, 未知語スロットがどのスロットにもなり得ると考えるためである.

スロット状態変化間の類似度は, 値の変化したスロットと変化した順番を比較し, 一致している個数を基にする. ただし, 同一のスロット状態変化間の類似度が最大となるようにするために, 一方のスロット状態変化が他方のスロット状態変化の拡張になっている場合は, 拡張になっている部分の個数に応じて, 類似度を低くする. Aux_1 により値の変化したスロットと変化した順番が一致している個数を数え, 一方のスロット状態変化が他方のスロット状態変化の拡張になっている場合は, Aux_2 により拡張になっている部分の個数を数える.

2 つのスロット状態変化 $CH1, CH2$ に対し, 類似度 Sim_3 を次のように定義する.

$$\begin{aligned} Sim_3(CH1, CH2) &= c_1 \times Aux_1(CH1, CH2) \\ &- c_2 \times Aux_2(CH1, CH2). \end{aligned}$$

ただし, c_1, c_2 は定数で, それぞれ, 値の変化したスロット, 拡張の度合を類似度に強調させる程度を表す. また, Sim_3 の計算結果が負になった場合は, Sim_3 の値を 0 とする.

例えば, (人, 番組) と (人) の類似度は, 人スロットが一致しているが, 一方のスロット状態変化だけが番組スロットを含むので, $c_1 - c_2$ となる.

3.2 スロット状態変化の信頼度

実例の中のスロット状態変化の信頼度は, スロット状態変化が音声認識誤りを含むかどうかを表すものである. スロット状態変化の信頼度を計算するために, スロット状態変化とスロット状態変化を引き起こした実際のユーザ発話の書き起こしを比較する.

書き起こしとして扱う情報は, 書き起こしから得られる実際のユーザ発話が完全にシステムに理解されたと仮定したときに, 値の変化するスロットと変化する順番とする. そこで, 値の変化するスロットを変化する順に並べたリストで書き起こしを表すことにする.

実例の中のスロット状態変化 CH の信頼度 Rel は, スロット状態変化を引き起こした実際のユーザ発話の書き起こし TR を使って次のように計算する. まず, 現在のスロット状態変化に未知語スロットが現れたとする. このとき, CH は, 未知語スロットに位置的に対応するスロット $slot_x$ を含む. そこで, $slot_x$ が TR に含まれるかどうかを判定する. もし含まれなければ信頼度 0 とする. もし含まれれば, CH と TR の中にある $slot_x$ の直前のスロット同士, 直後のスロット同士を比べる. 両方とも等しくない場合は信頼度 1, 片方だけ等しい場合は信頼度 2, 両方とも等しい場合は信頼度 3 とする.

4 属性推定手続き

実例は, 局所的対話遷移と書き起こしからなるデータなので, これら 2 つの要素を順に並べたリストで実例を表すことにする. スロット状態変化に未知語スロットを含む局所的対話遷移 A と実例 B に対し, それらの類似度 SIM を次のように定義する.

$$SIM(A, B) = \frac{Sim(A, B_1) + Rel((B_1)_3, B_2)}{Sim(A, A) + Rel(A_3, A_3)}.$$

現在のスロット状態変化に未知語スロットが現れたとき, 本稿で提案する未知語属性推定の手続きは次のように述べられる.

1. 現在の局所的対話遷移と蓄積されている各実例との類似度を SIM により計算する.
2. 1 で計算された値が最も大きい実例を選ぶ.
3. 2 で選ばれた実例が 1 個であれば, その実例により属性を推定する. 複数個であれば, 4 へ.
4. 選ばれた各実例により未知語の属性を推定し, 多数決により属性を選ぶ. それでも属性が一意に決まらない場合は, 5 へ.
5. 現在のスロット状態変化に既に現れている属性は未知語の属性ではないという条件の下で属性を選ぶ. それでも属性が一意に決まらない場合は, 候補となっている属性からランダムに選ぶ.

5 実験

我々のグループで開発された天気情報案内を行う音声対話システム飛遊夢 (ヒューム) を使い, 提案方法による属性推定実験をした. 飛遊夢のもつ属性の知識は,「時間」,「場所」,「警報種別」,「情報種別」の 4 つからなる.

実例を作成するために, グループ員 10 人がシステムと対話し, 14 対話, 402 発話収集した. この対話によるシステムログと実際のユーザ発話内容を記録した音声ファイルを基に, 96 個の実例を作成した.

飛遊夢は, ユーザ発話に含まれる未知語を検出する機能をもたないので, 作成した実例を使って未知語を含む局所的対話遷移を作成した. 作成方法は, 実例の中のスロット状態変化に含まれる語の一部を未知語と仮定するという方法である. 未知語と仮定された語の属性が推定されたとき, 属性推定正解とする.

3.1 節の類似度の定義の中のパラメータについては, 予備実験の結果から, システム発話間の類似度において, 焦点スロットを強調することにした. 具体的には, s_1 を 2, s_2, c_1 を 1, s_3, c_2 を 1/2 とした.

実例 96 個のうち, システム発話が質問であるものは 52 個, 確認であるものは 44 個であった. システムが質問した属性は,「時間」(25 回),「場所」(24 回),「警報種別」(3 回) の 3 種類であり, 1 回のシステム発話で 1 つの属性についての質問をした. システムが確認した属性は,「時間」,「場所」,「警報種別」,「情報種別」の 4 種類であり, 1 回のシステム発話で,「時間, 場所, 情報種別」(33 回),「場所, 情報種別」(6 回),「警報種別, 情報種別」(2 回),「時間, 場所」(1 回),「情報種

表 1: スロット状態変化に未知語を 1 語含む場合の属性推定結果 (正解数/問題数 (割合))

未知語位置	1 位に正解	1 位に正解 & $SIM \geq 0.8$	1 位に正解 & $SIM \geq 0.9$
第 1 番目	66/96 (0.69)	59/82 (0.72)	54/70 (0.77)
第 2 番目	18/38 (0.47)	10/14 (0.71)	3/6 (0.50)
第 3 番目	4/7 (0.57)	データ無し	データ無し
未知語位置	2 位までに正解	2 位までに正解 & $SIM \geq 0.8$	2 位までに正解 & $SIM \geq 0.9$
第 1 番目	78/96 (0.81)	61/72 (0.85)	27/32 (0.84)
第 2 番目	25/38 (0.66)	2/4 (0.50)	データ無し
第 3 番目	4/7 (0.57)	データ無し	データ無し

別」(2 回) の確認をした. ユーザ発話により値の変化した属性は, 1 回のシステム発話に対し最大 3 個であり, 各属性がスロット状態変化の第 1 番目に現れた回数は, 時間, 場所, 情報種別, 警報種別の順に, 22 回, 66 回, 0 回, 8 回, 第 2 番目には, 1 回, 27 回, 1 回, 9 回, 第 3 番目には, 0 回, 4 回, 0 回, 3 回であった.

実験の手順は次の通りである.

1. 96 個の実例を, 実例 10 個からなる 9 つのグループ $G_1 \ldots G_9$ と実例 6 個からなる 1 つのグループ G_{10} にランダムに分割
2. G_1 に含まれる実例から未知語を含む局所的対話遷移を作成
3. G_1 以外のグループに含まれる実例全体を蓄積された実例として, 2 で作成した未知語を含む局所的対話遷移に対して, 提案方法による未知語属性推定
4. G_2 から G_{10} までのそれぞれのグループに対し, 2, 3 の操作を繰り返す

この実験により, 表 1 のような結果が得られた. 表 1 中の「未知語位置」とは, スロット状態変化における未知語スロットの位置である. 「1 位に正解」とは, 類似度 SIM が最大となる実例による属性推定結果である. 「1 位に正解 & $SIM \geq 0.8$」とは, 類似度 SIM が最大かつ 0.8 以上となる実例による属性推定結果である. 「2 位までに正解」とは, 類似度 SIM が 2 位までの実例による属性推定結果である. 「2 位までに正解 & $SIM \geq 0.8$」とは, 類似度 SIM が 2 位までかつ 2 位の SIM が 0.8 以上となる実例による属性推定結果である.

スロット状態変化の第 1 番目に未知語が含まれる場合, 類似度最大となる実例による属性推定正解率は 69 % である. さらに, 類似度が 0.9 以上の場合は, 属性推定正解率が 77 % になる.

比較として, スロット状態変化の第 1 番目に未知語が含まれる場合について, 簡単なヒューリスティックスによる未知語の属性推定を行った. このヒューリスティックスは, システム発話が質問である場合, スロット状態変化の第 1 番目にある未知語の属性は, システムが質問している属性であると推定するものである. また, システム発話が確認の場合は, 第 1 番目に確認された属性を未知語の属性として選ぶ. ただし, 選ばれた属性が既にスロット状態変化に現れている場合は, 選ばれた属性以外の属性をランダムに選ぶ. このヒューリスティックスによる属性推定正解率は 46 % である.

6 おわりに

本稿では, 効率的な対話によりタスクを遂行するための未知語知識獲得の第一段階として, 未知語の属性推定に注目した. そして, 実例に基づく未知語属性推定方法を提案し, この方法が簡単なヒューリスティックスよりも高い属性推定正解率を得ることを確認した. 今後は, 大量の実例により, 今回と同様の実験を行うとともに, より多くの属性をもつ音声対話システムへの応用を考えていきたい.

謝辞 日頃より御指導頂くメディア情報研究部 萩田紀博部長, 有益な示唆を頂くマルチモーダル対話研究グループの諸氏に感謝致します.

参考文献

[1] 荒木雅弘, 堂下修司. 対話事例ベースによる発話内容の推定および未知語の解析. 情報処理学会第 49 回全国大会発表論文集 Vol. 3, pp. 155–156, 1994.

[2] 伊藤克亘, 速水悟, 田中穂積. 音声対話システムにおける未知語の扱い. 人工知能学会研究会資料 SIG-SLUD-9201-1 (4/15), pp. 1–9, 1992.

[3] 佐藤理史. MBT1 : 実例に基づく訳語選択. 人工知能学会誌 Vol. 6 No. 4, pp. 129–136, 1991.

[4] C. Cardie. A case-based approach to knowledge acquisition for domain-specific sentence analysis. AAAI-93, pp. 798–803, 1993.

[5] A. L. Gorin, S. E. Levinson, and A. N. Gertner. Adaptive acquisition of spoken language. ICASSP-91, pp. 805–808, 1991.