

자동통역에서의 대화체 기계번역을 위한 문맥의 구축과 이용¹

이재원^{○*}, 서정연^{*}, 김길창^{*}
한국과학기술원 전산학과, 인공지능연구센터^{*}

Context Construction and It's Using for Dialogue Machine Translation in Automatic Interpreting Telephony

Jae-won Lee^{○*}, Jungyun Seo^{*}, Kim Gil Chang^{*}
Department of Computer Science and CAIR, KAIST^{*}

요약

자동통역에서의 대화체 기계번역은 일반적인 문어체 문장 번역과는 다른 몇가지 특징을 고려하여야 한다. 첫째, 자동통역에서의 기계번역은 음성인식의 결과를 번역하는 부분으로, 하나의 문장이 아닌 다중 입력을 받아 이 중 가장 올바른 문장을 번역하여야 한다. 둘째, 대화체 문장에서는 일반적으로 생략을 포함하는 단편적인 발화나 대응어의 사용빈도가 많다. 그러나, 이러한 현상은 언어마다 다소 다르게 사용되기 때문에, 이들에 대한 올바른 해석을 한 후 번역하는 것이 필요하다. 대화체 기계번역이 이러한 문제점들을 해결하기 위해서는 문맥정보를 필요로 한다. 대화는 상호간 밀접한 관련성을 가지고 진행되기 때문이다. 이에 본 논문에서는 담화분석을 통해 올바르게 문맥을 구축하고, 이 정보를 이용하여 앞에서 언급한 문제점들을 해결하기 위한 방법론에 대해 논하고자 한다.

1. 서론

자동통역은 화자(speaker)의 음성 대화(spoken dialogue)를 입력으로 받아 청자(hearer)가 사용하는 언어로 음성 대화를 생성해 내는 시스템을 말한다. 자동통역 시스템은 일반적으로, 음성인식과 기계번역, 음성합성의 3부분으로 구성된다. 이 중 기계번역 부분은 음성인식 결과인 대화체 문장을 원하는 언어로 번역해내는 역할을 담당하게 된다. 즉, 언어처리를 담당하는 부분이다.

자동통역에서의 기계번역 시스템은 일반적인 문어체 기계번역 시스템보다 더 많은 기능들이 요구된다.[1,2,3] 일반적인 언어처리 기능 외에도 문맥(context)이나 대화모델(dialogue model) 같은 부가적인 지식과 처리를 필요로 한다. 다음은 자동통역을 위한 기계번역 시스템에서 요구되는 사항들을 기술한 것이다.

첫째, 다중입력(multiple input)을 처리할 수 있어야 한다. 음성인식 결과는 일반적으로 하나의 올바른 문장이 아닌 확률값을 가진 여러 후보문장을 제시하게 된다. 그러므로, 기계번역 시스템에서는 여러 언어적인 정보를 이용하여 이들 중 올바른 문장을 선택하여 번역할 수 있어야 한다.

둘째, 대화체에서는 문법적으로 완전한 문장이 아닌 일부분이 생략된 문장이나 단편적인 문장이 자주 사용된다. 대화는 화자간의 상호작용(interaction)에 의해 이루어지기 때문에 문맥이나 상황에 따라 많은 부분이 생략되어 나타난다. 특히, 문장간 생략현상이 빈번하며, 대응어 또한 빈번히 사용된다. 그러므로, 문맥을 이용한 정확한 분석이 요구된다.

셋째, 처리속도 문제이다. 자동통역 시스템은 화자간의 원활한 대화가 가능하도록 실시간 내에 처리되어야 한다. 따라서, 기계번역 시스템은 번역에 필요한 범위내에서 가능한한 단순하고 효율적인 시스템이 되어야 한다.

위의 사항들에서 알 수 있는 바와 같이, 대화체 기계번역에서는 하나의 문장에 대한 분석 외에도 전체 대화에 대한 분석 능력이 필요로 한다. 즉, 문맥이나 상황정보 등을 이용하여 예측성을 줄이고 생략된 부분들을 복구해낼 수 있어야 한다. 본 논문에서는 대화 모델링을 통한 대화분석 결과로 문맥 트리를 구축하고 이들이 다중입력이나 생략현상 처리에 어떻게 사용될 수 있는지에 관하여 기술하였다. 이는 호텔예약 대화를 대상으로 한 한 영 대화체 기계번역 시스템¹[1,2] 중 담화분석 모듈에 대한 것이다.

¹본 연구는 한국통신의 정기기초 과제 "자동통역 전화개발을 위한 대화체 기계번역에 관한 연구"의 지원을 받은 것입니다.

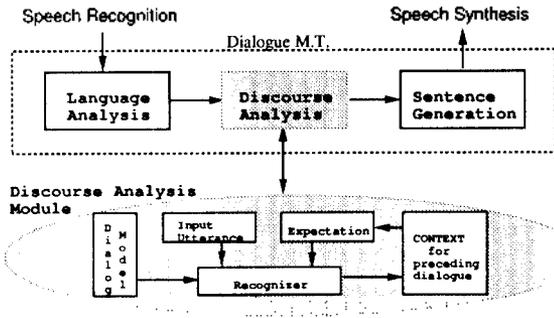


그림 1: 시스템 개요

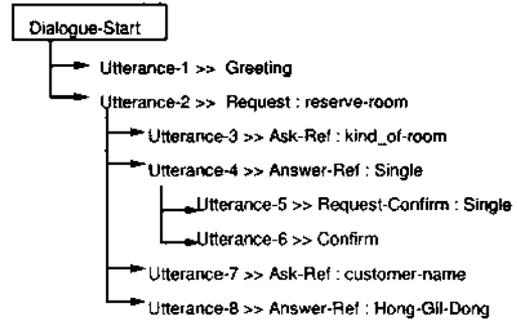


그림 2: 대화구조

2. 시스템 개요

2.1 시스템 구성

본 대화제 기계번역 시스템은 언어해석(language analysis) 모듈과 담화해석(discourse analysis) 모듈, 생성(generation) 모듈의 3부분으로 구성되어 있다. 본 논문은 주로 담화해석 모듈에 대해 기술한 것으로, 이는 언어해석 과정을 통해 얻어진 의미표현에 대해 문맥정보를 이용하여 담화해석을 하게 된다. 담화해석 결과는 다시 문맥에 반영되어 이후의 대화를 해석하는데 사용된다.

[그림 1]은 전체적인 번역시스템의 구성과 담화해석 모듈에 대한 처리 과정을 보인 것이다. 담화해석은 이전 대화에 대한 문맥으로부터 얻어지는 다음 대화에 대한 예측(expectation) 정보를 이용하여 가능한 후보 중 올바른 문장의 의미를 인식하게 된다. 이때 인식기는 문맥트리와 호텔예약 영역에 대한 영역 계획 지식, 대화 자체에 대한 지식인 담화계획 등을 사용한다. 이러한 과정에서 다중입력 문제, 생략과 대응어 처리등이 고려된다.

2.2 대화 예문

담화해석모듈의 이해를 돕기 위해 수집된 호텔예약 대화 중 일부를 발췌하여 분석결과와 함께 보였다. [그림 2]는 대화예문에 대한 담화구조를 도식화한 것이다. 담화해석 과정에 대한 자세한 내용은 3장의 대화모델(dialogue model)에서 기술하였다.

- (1) [호텔] 네, 한국호텔입니다.
- (2) [고객] 5월 4일로 예약 좀 하고 싶은데요.
- (3) [호텔] 어떤 방이 필요하신데요?
- (4) [고객] 싱글룸이요.
- (5) [호텔] 싱글룸이요?
- (6) [고객] 예.
- (7) [호텔] 성함이 어떻게 되십니까?
- (8) [고객] 홍길동입니다.

3. 대화모델

본 절에서는 입력된 발화가 이전대화(preceding dialogue)와의 관련성(relationship)을 추론하기 위한 방법에 대해 기술하였다. 이 과정을 통해 얻어진 대화구조(dialogue structure)는 예측성을 해소하거나 생략과 대응어를 처리하는데 매우 유용하게 쓰이게 된다. 이러한 분석 과정은 그동안 많이 연구되어 온 대화이해(dialogue understanding) 모델 [5,10]과도 유사한 것이다. 그러나, 기존의 대화이해 모델은 계산적인 측면과 견고성 측면에서 볼 때 실질적으로 기계번역에 적용되기는 어렵다고 본다. 본 논문에서는 기계번역의 입장에서 요구되는 범위내에서 가능한 단순화시킨 대화모델을 사용하였다.

3.1 담화계획

대화에서 각각의 발화는 화자가 전달하고자 하는 내용(content) 외에 어떠한 행위를 포함하고 있다고 볼 수 있다. 대화는 이러한 행위의 상호작용(interaction) 관점에서 설명될 수 있다. 그러므로, 이러한 행위는 각각의 발화간의 관계(relation), 즉 대화 구조를 반영하는 중요한 단서가 될 수 있다. 본 논문에서는 이를 담화목적이라 하며, 다음은 그 중 일부를 보인 것이다.

- Request-Action : 화자가 청자에게 어떤 행위를 요구하는 발화이다. 요구된 행위는 일반적으로 화자가 원하는 영역목적에 관련된 것이다. 이에 대한 응답으로 청자는 요구된 행위를 수행하게 된다.
[고객] 방을 예약하고 싶은데요.
- Ask-Information : 청자에게 어떠한 정보를 요구하는 행위이다. 요구된 정보는 일반적으로 영역목적을 수행하는데 필요한 정보가 된다. 청자는 이 발화에 대해 요구된 정보를 제공하게 된다. Ask-Ref와 Ask-If로 세분화될 수 있다.
[고객] 싱글룸의 가격이 얼마입니까?
- Request-Confirm : 상대방의 발화내용을 확인하고자 할 때 사용된다. 그러므로, 확인하고자 하는 내용은 이전 발화의

았을 때, 그는 먼저 주어진 발화를 이해하려고 하게 된다. 만일 청자가 정확히 화자의 발화를 이해하지 못했다면 그는 그 내용을 확인하거나 명확히 하기 위한 질문을 하게 될 것이다. 다음으로 질문을 정확히 이해했다고 하더라도 질문이 잘못된 내용을 포함하거나 질문에 답하기 위해서 부가적인 정보를 필요로 하는 경우 이를 해결하기 위한 대화(sub-dialogue)가 내포되어 나타날 수 있다. 이들 조건이 모두 만족되었을 때 비로소 응답을 할 수 있는 것이다. 그림 3은 이러한 대화의 진행과정을 화자의 담화목적과 관련하여 모델링 한 것이다. 즉, 대화에서 화자간의 상호작용을 모델링한 것으로 발화가 내포하는 화자의 담화목적들을 추론하는데 사용된다.

3.3 영역계획

목적 지향적(goal-oriented)인 성격의 대화에서 해당 대화 영역(domain)에 대한 지식은 발화의 의미를 정확히 이해하고, 발화간의 관계를 파악하는데 유용하게 사용될 수 있다. 대화구조는 영역계획이 표현하는 구조(structure)를 반영하기 때문이다. 본 논문에서는 호텔예약과 관련한 지식을 계획(plan) 표현 메카니즘을 이용하여 구축하였다.

영역지식은 해당 영역목적들을 이루기 위한 과정 즉, 동작/상태 개념들간의 관계와 각 동작/상태 개념들이 가지는 의미제약(semantic constraint)을 포함하고 있다. 그림 4은 객실예약의 한 과정으로 고객이 원하는 방이 비워져 있는지 검사하는 행위가 필요하다는 것과 이때 요구되는 파라미터로 객실의 종류, 사용기간 등이 있음을 보여주고 있다.

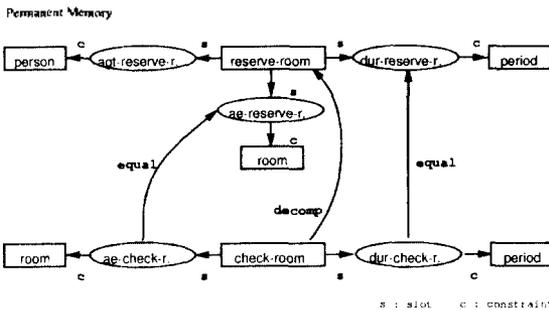


그림 4: 영역지식의 일부

3.4 추론 알고리즘

[그림 5]는 목적 지향적인 대화를 대상으로 하는 담화분석 알고리즘을 보인 것이다. 담화분석에는 계획지식(영역계획, 담화계획)과 대화 흐름도(dialogue flow diagram), 스택, 문맥트리(context tree)가 사용된다. 입력으로 언어해석 모듈의 결과인 의미해석 결과를 받아 담화처리가 끝나면, 그 결과를 스택과 문맥트리에 반영하게 된다. 스택은 두께가 사용된다. 하나는 대화 흐름도 상에서

```

INPUT :  1. SEM-REP
        2. STATE-STACK
        3. DISC-STACK
        4. PLAN-LIBRARY
        5. CONTEXT-TREE

OUTPUT : 1. Discourse and Domain Goal
        2. STATE-STACK
        3. DISC-STACK
        4. CONTEXT-TREE

PROCEDURE :
1. transform(SEM-REP, SPEECH-ACT)
2. Current-State <- top(STATE-STACK)
3. DISC-GOAL <- possible Discourse Goal
   in Current State
   3.1 decomp(DISC-GOAL) = SPEECH-ACT
   3.2 test constr(DISC-GOAL)
   3.3 execute effect(DISC-GOAL)
4. If Step 3 success, the process Next Dialogue
   else GOTO Step 3
    
```

그림 5: 추론 알고리즘

현재의 상태번호(state number in dialogue flow diagram)를 저장하는데 쓰이며, 다른 하나는 이전의 대화 중 아직 진행중에 있는 담화목적들을 유지하고 있다. 문맥트리(context tree)는 지금까지 진행되어 온 대화에 대한 해석 결과를 나타낸다. 추론과정은 다음과 같다.

먼저, 의미해석기로부터 입력문장에 대해 의미구조 형태로 입력을 받는다. 입력된 의미해석 결과는 단계 1에서 화행표현으로 변환된다. 화행변환은 문장유형과 동사, 양상(modal) 등을 고려하여 처리된다. 단계 2에서는 현재의 상태번호를 스택(STATE-STACK)으로부터 가져오게 된다. 단계 3에서는 현재의 상태에서 가능한 담화목적(DISC-GOAL)을 대화흐름도에 따라 우선순위에 의해 차례로 검사하면서 만족되는 담화목적들을 찾게 된다. 단계 3.1에서 3.3은 이 과정을 기술하고 있다. 첫째, 입력문장의 화행(SPEECH-ACT)이 담화계획의 부분동작(Decomp)과 일치해야 한다. 둘째, 담화계획의 조건부(Constraints)를 만족해야 한다. 일반적으로 이 단계에서는 해당 발화의 영역계획(domain plan) 또는 문맥(Context)과의 관련성을 검사하게 된다. 조건부가 만족될 경우 단계 3.3에서 해당 담화계획의 결과(Effects)를 수행하게 된다. 이 단계에서는 해석 결과에 따른 현재 상태의 변화를 스택에 반영하고 추론된 화자의 의도를 문맥에 저장하게 된다. 단계 3가 성공적으로 수행되면 다음 대화문장에 대한 처리를 진행하게 되며, 만일 실패하게 될 경우에는 단계 3로 돌아가 현재의 상태에서 가능한 다음의 담화목적들을 가져와서 단계 3이 성공할 때까지 같은 과정을 반복하게 된다.

위에서 언급한 바와 같이 담화계획의 조건(constraints)부에서는 일반적으로 대화 내용이 화자가 이루고자 하는 영역 목적과 계

획 또는 문맥과 어떠한 관련성을 가지는지를 검사하게 된다. 화자의 영역목적과 계획은 이 단계에서 추론된다. 이때, 사용되는 영역 계획 추론 규칙은 다음과 같다.

규칙 1 Header-Action Rule

발화 자체가 화자가 의도하는 목적을 나타내는 경우이다. 즉, 발화내용이 계획의 Header와 일치하는 경우이다. 다음은 그 예로 객실예약과 예약확인을 나타내는 발화이다.

- “싱글룸으로 방 하나를 예약하고 싶는데요.”
- “싱글룸에 대한 예약을 확인하고 싶습니다.”

규칙 2 Precondition-Action Rule

발화내용이 계획의 전제조건과 일치할 경우이다.

- “오늘 빈방이 있습니까?”
- 방을 예약하려면 원하는 방이 비워져 있어야 한다. 따라서 빈방이 있는지 질문했다면, 지금 화자는 객실예약을 원한다고 볼 수 있다.

규칙 3 Decomposition-Action Rule

발화내용이 계획의 부분동작과 일치하는 경우이다.

- “어떤 방을 원하십니까?”
- “얼마나 머무르실 예정입니까?”

위의 발화는 객실예약을 위한 부분동작으로 이러한 발화가 입력되면 이는 객실예약 대화의 일부임을 알 수 있다.

4. 문맥의 구축과 이용

분석과정을 살펴보기 위해 앞에서 보인 대화예문의 일부를 대상으로 담화분석을 통해 문맥트리(context tree)를 구축해가는 과정을 보이고 생략을 포함하는 발화가 어떻게 해석되는지를 보이고자 한다. 편의상, 앞의 대화 예문 중 일부분을 다시 보았다.

- (1) [고객] 5월 4일로 예약 좀 하고 싶는데요.
- (2) [호텔] 어떤 방이 필요하신데요?
- (3) [고객] 싱글룸이요.
- (4) [호텔] 싱글룸이요?
- (5) [고객] 예.

첫번째 (1)번 발화는 대화의 시작부분이므로 초기상태인 state-1에서 시작된다. state-1에서 가능한 담화목적으로는 request-action과 request-information, give-information이 있다. 먼저 request-action이 검사된다. 화행이 request 유형이므로 단계 3.1은 성공하게 된다. 다음으로 3.2에서 request-action에 대한 담화계획의 조건부인 want(Speaker,Action), step(Action,Plan)을 검사하게 된다. 고객이 예약을 원하므로 처음 조건은 만족된다. 다음으로 step(Action,Plan)은 요구된 행위가 영역목적과 부합하

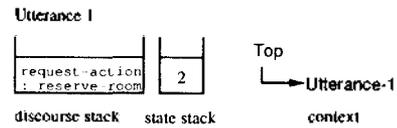


그림 6: 담화분석 결과-1

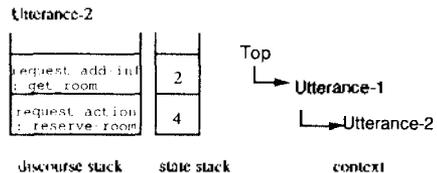


그림 7: 담화분석 결과-2

는지를 검사한다. 그러나, 현재는 대화의 초기 상태로 구축되어 있는 영역계획(Plan)이 존재하지 않는다. 따라서, 요구된 행위와 관련된 영역계획인 reserve-room을 추론한 후 성공적으로 끝나게 된다. 결과로, 단계 3.3에서 담화목적의 담화스택(DISC-STACK)에 저장되고 변경된 state-2번이 스택에 저장된다. 담화스택에 담화 목적을 저장하는 이유는 아직 요구된 행위인 호텔예약 행위가 아직 해결되지 않았기 때문에 이후에 나오는 대화를 처리하기 위해 필요한 것이다. [그림 6]은 담화분석 결과를 보인 것이다.

발화 (2) “[호텔] 어떤 방이 필요하신데요?”에 대해서는 state-2에서 시작된다. 현재의 상태에서 가능한 담화목적의 같은 방법에 의해 우선순위에 의해 비교된다. 즉, request-confirm, request-information 순으로 비교된다. 단계 3에서 먼저 request-confirm이 검사된다. 그러나, 이 발화는 (1)번 발화에 대해 확인을 원하는 것이 아니므로, 현재 화자는 (1)번 발화를 정확히 인지했다고 가정하고 현재의 상태는 상태 3번으로 바뀌게 된다. 다음으로 request-information이 검사된다. 이 과정에서 예약을 하는데 필요한 정보인 객실종류에 대한 정보를 요구하는 발화로 인식된다. 발화 (2)와 (3)은 내포된 sub-dialogue이므로 단계 3.3에서는 현재의 상태 state 3을 스택에 유지하면서, 새로운 sub-dialogue를 위한 정보, 즉 해당 담화목적(request-information)과 state-2를 스택에 입력한다(push operation). 이는 객실-예약에 대한 요구행위가 완료되지 않은 상황에서 발화 (2)와 (3)이 삽입된 경우이므로 (3)번 발화 이후에는 계속해서 호텔예약을 위한 대화가 진행될 수 있기 때문이다. [그림 7]은 발화(2)까지의 담화분석 결과이다.

이후의 대화에 대해서도 같은 방법에 의해 처리되며 [그림 8]은 발화(1)-(5)에 대한 분석결과를 보인 것이다.

발화에 포함되어 있는 생략이나 대응어 처리에 대한 자세한 알고리즘의 기술은 본 논문의 범위를 벗어나는 부분으로 자세한 언급은 없었다. 그러나, 기본적으로 위의 담화분석을 통해 얻어진 대화구조는 생략이나 대응어를 처리하는데 매우 중요하게 사용될 수 있

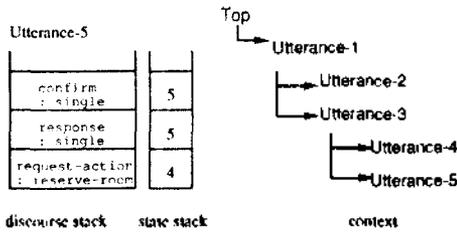


그림 8: 담화분석 결과-3

고, 예측정보를 통해 다중입력 처리 등에 이용될 수 있을 것이다.

5. 결론

본 논문에서는 자동통역에서의 대화체 기계번역에서 발생하는 여러가지 문제점들을 나열하고, 이들 문제점을 해결하기 위한 방법으로 담화분석을 통해 문맥을 구축하기 위한 알고리즘과 이물이 어떻게 사용되어질 수 있는지에 대해 언급하였다. 담화분석은 기본적으로 인공지능분야에서 많이 연구되어 온 계획에 기반한 대화이해 기법들을 사용하였다. 그러나, 이들 모델은 실질적으로 기계번역에 적용되기에는 많은 어려운 점이 있다. 이에 기계번역에서 요구되는 점들을 분석한 후 가능한 단순하고 효율적인 대화모델을 구축하고자 하였다. 결과로 얻어지는 문맥이라는 다중입력이나 생략, 대응어들을 처리하는데 유용하리라 생각한다.

대화는 상호 밀접하게 연관성을 가지고 진행되기 때문에 담화분석이 잘못되어 문맥이 잘못 구축될 경우 이후의 대화에 대해 잘못된 분석을 하기 쉽다. 특히, 자동통역의 경우 실시간 시스템인 점을 고려할 때 이러한 점은 매우 중요하다. 따라서, 모델을 견고하게 구축하고, 이후의 대화에 따라 잘못된 문맥을 수정해 나갈 수 있는 방법은 연구 등이 향후과제로 남아있다.

참고 문헌

[1] 서경연, 조경미, 김길상, "한-영 대화체 기계번역 시스템", 제 11회 음성통신 및 신호처리 워크샵 논문집, 한국음향학회, 1994

[2] 한국과학기술원, "대화체 기계번역에 관한 연구", 한국통신학회 최종 보고서, 1993.

[3] Jungyun Seo, Jae-won Lee, Jae-Hoon Kim, Jeong-Mi Cho, Chang-Hyun Kim, and Kim, Gil Chang, "Dialogue Machine Translation Using a Dialogue Model", *Proceedings of The First China-Korea Joint Symposium on Machine Translation*, Yanbian Univ. China, 1994

[4] Yasuharu ASANO, Keikichi HIROSE, "Dialogue Processing System for Speech Response with High Adapt-

ability to Dialogue Topics", *IEICE transactions on information and systems*, Vol.E76-D No.1, pp.95-105, 1993

[5] Sandra Caberry, "A Pragmatics-Based Approach To Ellipsis Resolution", *Computational Linguistics*, Vol.15, No.2, pp. 75-96, 1989.

[6] Jun-ichi Tsujii, Makoto Nagao, "Dialogue Translation vs. Text Translation - Interpretation Based Approach -", *Coling 88*, Vol.II., pp.688-693, 1988.

[7] H.Kitano, "Phi DM Dialog : An Experimental Speech To Speech Dialog Translation System", *IEEE Computer*, pp. 36 - 50, 1991.

[8] H.Kitano, "Parallel Incremental Sentence Production for a Model of Simultaneous Interpretation", *Current Research in NLG*, Academic Press, 1990

[9] Masako KUME, Gayle K. SATO, Kei YOSHIMOTO, "A Descriptive Framework for Translating Speaker's Meaning", *The 4th European Association for Computational Linguistics*, 1989.

[10] Diane J. Litman & James F. Allen, "A Plan Recognition Model for Subdialogues in Conversations", *Cognitive Science*, pp. 163-200, 1987.

[11] Yasuhiro Sobashima, "Local Context Analysis of Dialogues Using Bilingual Corpus", *International Symposium on Spoken Dialogue*, pp.61-64, 1993.