

論文96-33B-10-11

# 문맥 및 종결어미의 서법정보를 이용한 대화문의 화수력 분석

## (An Analysis of Illocutionary Force Types in a Dialogue, Based on the Context and Modal Information in the Ending of a Word)

金永吉\*, 崔炳旭\*

(Kim Young Kil and Choi Byung Uk)

### 요 약

본 논문에서는 문맥을 비롯하여 동사의 종결어미에 나타나는 서법정보를 이용한 화수력 분석 알고리즘을 제안한다. 한국어는 타언어와는 달리 동사의 어미에 나타나는 서법의 유형에 따라 화자의 발화의도를 나타내는 화수력의 변화가 심하게 나타난다. 따라서 대화문상의 발화문 화행분석을 위한 화수력 정보 자동추출에 있어 동사의 종결어미에 나타나는 서법정보는 중요한 단서가 된다. 본 논문에서는 실제 대화 데이터를 분석하여 화수력의 종류를 분류하며 수동적인 화수력 태깅 작업을 수행하여 대화문상에서의 화수력 출현빈도를 제시한다. 아울러 분석된 화수력과 종결어미의 관계를 바탕으로 화수력 추출 알고리즘을 제안한다. 그리고 본 알고리즘을 이용하여 실제 대화를 실험 데이터로 사용하여 그 효율성을 보인다.

### Abstract

This paper proposes an algorithm for analyzing illocutionary force type(IFT)s in a dialogue, based on the context and modal information in the ending of a word. In Korean, the variation of an illocutionary force type that represents a speaker's intention frequently occurs at the ending of a word, according to the type of modal information. And in an analysis of speech acts, the modal information that is included in the ending of a word may be an important clue to the automatic extraction of illocutionary force types. In this paper, we analyze real dialogue data, classify the types of illocutionary forces, perform the manual tagging of IFTs and show the frequency of each IFT's occurrence. And we also propose an algorithm to extract IFTs, based on the relationship between the analyzed IFTs and the endings of a word. And we use this proposed algorithm to make an experiment on dialogue data and show its efficiency

### I. 서 론

현재 대화체 문장을 그 처리대상으로 하고 자연언어 인터페이스를 이용하는 각종 응용 시스템으로서 대표적인 것이 사용자와의 대화를 통하여 데이터베이스 정보를 검색하는 시스템, CAI 시스템, 업무안내 등을 위한 지능적 대화시스템 등이 있다. 그리고 현재 기존의

문어체 중심적인 해석 방식에서 벗어난 대화체 음성언어 번역 시스템 등에 관한 연구가 활발히 진행중에 있다.

이와 같은 자연언어에 의한 대화체 응용 시스템은 사용자의 입력 발화문에 대한 구문적 및 의미적 해석 방식을 통한 사용자의 전달내용 분석외에도 사용자가 정보를 요구하는지, 정보를 제공하는지 등에 관한 화행론적인 발화의도를 인식할 수 있어야 한다. 따라서 발화문에 대한 전통적인 구문적, 의미적 분석 방법만으로는 입력 발화문에 담긴 화자의 화행론적 정보를 해석

\* 正會員, 漢陽大學校 電子通信工學科

(Dept. of Elec. Communication Eng., Hanyang Univ.)

接受日字:1996年7月2日, 수정완료일:1996年10月2日

할 수 없으며 화자의 발화의도를 분석할 수 있는 화행 분석이 선행되어야 한다<sup>[1, 2]</sup>.

그리고 대화체 음성언어 번역에 있어서도 대화자들의 대화 내용을 추적하면서 대화구조를 모델링하는 대화분석 과정이 필수적이다. 이와 같은 대화 모델링을 위한 중요한 정보로서 각 발화문의 화수력(illocutionary force) 정보가 사용될 수 있다. 여기서 화수력이란 화자의 발화문에 나타나는 명제적인 내용외의 화자의 발화의도를 나타내는 언어적 수행력 또는 연표내적 힘이라고 정의할 수 있다<sup>[3]</sup>. 화수력 정보는 대화구조 분석외에도 음성인식 결과 나타나는 구문적, 의미적으로 감지할 수 없는 인식 오류를 다음 발화문의 예측을 통해 오류를 정정할 수 있는 데이터로 사용될 수 있다.

화행분석을 통한 화수력 정보추출에 관한 연구는 다른 선진외국에 비해 국내에서는 매우 미약한 실정이다. 영어는 어휘위주의 화수력 분석이 활발이 진행되어 각 대화식 응용 시스템 등에 적용되고 있는 단계이며 일본어 또한 간단한 구문적 패턴을 이용한 화수력 분석 방식 등이 음성언어 번역이나 기타 대화식 응용 시스템 등에 사용되고 있다. 그러나 한국어에 대한 체계적인 화수력 분석 방식에 관한 연구는 거의 이루어지고 있지 않은 실정이다. 한국어는 영어나 일본어와 같은 타언어와는 달리 동사의 어미에 나타나는 서법의 유형에 따라 화자의 발화의도를 나타내는 화수력의 변화가 심하게 나타나는 특징을 가지고 있다. 즉 한국어 입력 발화문의 화행분석을 위한 화수력 정보 추출에 있어 타언어와는 달리 동사의 종결어미에 나타나는 서법정보가 중요한 단서가 될 수 있다.

따라서 본 논문에서는 우선 실제적인 대화 데이터의 분석을 통하여 계산학적으로 분석이 가능한 화수력의 종류들을 설정한다. 그리고 정의된 각 화수력과 동사의 종결어미에 나타나는 서법 정보와의 의존관계에 대한 분석 데이터를 기반으로 화수력 추출 알고리즘을 제시한다. 그리고 화수력 자동 추출을 위한 시범 시스템을 구성하여 실제적인 실험대화를 대상으로 본 논문에서 제시한 화수력 추출 알고리즘의 효율성을 보인다.

## II. 화행 및 화수력 분석

오스틴(Austin) 이후 발전되어 온 화행이론은 문장이 가지는 고정된 의미가 아닌 대화를 통해 나타나는

화자의 발화의도에 관한 연구이다. 일반적으로 화행이란 언어적인 속성과 행위의 속성을 갖춘 것이며 화자가 어떤 내용(Proposition)과 청자에게 미치고 싶은 의도를 나타내는 화수력을 청자에게 발화로서 전달하고 그것을 인지하는 단계이다. 이 때 화행을 화자가 청자에게 내용을 전달하는 행위와 화수력을 전달하는 행위로 구분하여 2단계적으로 분석할 수 있다<sup>[3]</sup>.

그림 1은 대화를 통한 DB 정보 검색 시스템에서의 화행의 2단계적인 분석 방식을 이용한 사용자의 발화의도 해석 과정의 일례를 보이고 있다. 즉 시스템은 사용자가 입력하는 발화문을 구문 및 의미분석을 통하여 그 명제 내용을 분석하고 이 후 대화분석을 통하여 발화문에 담긴 화수력 정보가 행위요구임을 판단함으로써 사용자의 요구를 인식하고 있다. 이때 MotivatebyRequest는 화자의 요구로 청자를 동기유발시키는 내용을 나타내는 논리 술어식이다<sup>[4]</sup>. 이와 같이 대화문에 대한 화행분석은 그 명제적인 내용 뿐만 아니라 명제적인 내용에 담긴 화수력 정보의 분석이 필수적으로 요구된다.

```

System> 어떤 정보를 원하십니까?
User> 컴퓨터에 관련된 책을 찾아 주세요.

< 구문 및 의미해석 정보 >
(CAT VP)
(HEAD (SYN (TENSE PRES) (VFORM DECLARATIVE) (CHANGE TERM) (WISH +) (HUMBLE +))
(SEM (DEF (CONT (REF 컴퓨터[SCI])) (PRED 관련되다)) (ACC 책[BOO]) (PRED 찾다))

< 대화해석 정보 >
Illocutionary Force : 행위요구(Act_Request)
Propositional Contents : 찾다( (컴퓨터/CONT, 관련되다/PRED)/DEF, 책/ACC )
Speech_Act : MotivatebyRequest(Speaker, Hearer, Act(Hearer, 찾다( (컴퓨터/CONT, 관련되다/PRED)/DEF, 책/ACC )))
    
```

그림 1. 데이터베이스 검색 시스템의 화행분석  
Fig. 1. An Analysis of Speech Acts in a DB Retrieval System.

### 1. 관련 연구

Austin, Searle 등은 주로 동사의 유형에 따른 화수행위의 분류를 시도하였다<sup>[5]</sup>. 그리고 Allen은 입력 발화문에 존재하는 화자의 발화의도를 나타낼 수 있는 어휘정보를 이용하여 화수력을 분석하였으며 간접화행

분석 방식을 제시한 바도 있다<sup>16)</sup>. Carberry는 계획 인식을 기반으로 하는 대화 시스템의 사용자 입력 발화문에 관한 의도 분석을 위하여 17가지 종류의 대화 목표(Discourse Goal)를 설정하였지만 입력 발화문의 해석정보를 이용한 대화목표 추출을 위한 구체적인 방법을 제시하고 있지 않다<sup>17)</sup>.

일본의 ATR에서는 9가지의 화수력 정보를 설정하고 실제 대화문들을 수동적으로 태깅하여 통계적인 데이터로 추출하였다<sup>18)</sup>. 그리고 이 통계정보를 음성인식에 있어서 다음 발화문의 예측을 위한 데이터로 사용하여 음성인식 오류 정정에 있어서의 효율성을 증명한 바 있다. 그러나 화수력 태깅을 위하여 간단한 표층적 어휘정보를 이용한 규칙기반의 추출 방법이 적용되었으며 거의 전문가의 주관적인 수작업에 의존하였다. 따라서 대량의 데이터에 대한 작업이 불가능하고 태깅된 데이터 또한 객관성을 지니지 못하는 단점이 있었다.

국내에서는 주로 언어학적인 측면에서 상황의미론에 의한 화행분석 등이 진행되고 있으며 화수행위 분류, 간접 화행 분석 등에 관한 연구가 수행되고 있다. 그리고 자연언어 인터페이스를 이용한 음성언어 대화시스템 등에서 키워드를 이용한 사용자의 발화 의도 분석 등이 시도된 바 있지만 이는 적용대상에 한정적이고 주관적인 데이터 작업을 필요로 하여 그 확장성에 문제가 있다. 논문 [1]에서 제시한 호텔예약을 위한 대화시스템에서는 화자의 전달내용 분석을 위해서 입력 발화문의 해석결과 얻어지는 구문 및 의미표현식을 이용하였으며, 화수력 분석을 위해서 이전 발화문의 화수력과 입력 발화문의 구문정보를 이용한 바 있다. 그러나 입력 발화문의 서술부의 구문정보로서 의문형과 평서형의 정보만을 이용함으로써 다양한 종류의 화수력 분석에 있어 어려움이 있었다.

2. 화수력의 분류

Searle이 믿음, 심리상태, 통사구조적 특성 등 12가지의 다양한 기준에 의하여 화수행위들을 5가지로 분류한 바 있으며 그외 여러가지 분류 방법들이 제안되어 왔다<sup>15)</sup>. 그러나 이러한 분류 기준을 시스템내에 계산학적으로 형식화하여 화행분석에 적용하기에는 무리가 따른다.

따라서 본 논문에서는 실제 대화문을 분석하여 표 1에서와 같이 화수력을 그 표층적인 어휘정보들을 기준으로 크게 표현, 요구, 제공, 기타 4가지로 분류하였다.

표현은 대화문에 나타나는 관용적인 표현으로 정의하며 대화개시, 대화종결 및 대화일상 표현으로 분류하였다. 그리고 요구는 화자(Speaker)가 청자(Hearer)에게 정보나 행위의 제공을 부탁하는 발화의도를 나타내며 3가지의 정보요구와 1가지의 행위요구로 구성한다.

정보요구중에서 진위정보요구와 의문정보요구는 의문사의 존재 여부로 구분하며 확인정보요구는 이전 발화문에 나타난 구문을 다시 언급하여 그 확인을 요구하는 발화의도로 정의한다. 제공은 화자가 청자에게 정보나 행위를 제공하거나 행위제공을 약속하는 화자의 발화의도를 나타내며 5가지의 제공형태로 구분한다. 그리고 놀라움의 표시나 발화중단으로 인한 단편적 발화문 등을 기타로 분류하였다. 이러한 화수력 분류는 각 응용 시스템의 적용 분야에 따라 더욱 세분화될 수 있다.

표 1. 화수력의 분류 및 그 일례

Table 1. A Classification of Illocutionary Force Types and Their Examples.

| 대분류                | 소분류                                 | 설명 및 일례  |
|--------------------|-------------------------------------|--|
| 표현<br>(Expression) | 대화개시표현<br>(Opening Expression)      | 대화의 개시를 나타내는 화자의 의도<br>안녕하세요, 여보세요, 수고하십니다 ...             |
|                    | 대화종결표현<br>(Closing Expression)      | 대화의 종결을 나타내는 화자의 의도<br>수고하세요, 안녕히 계세요, ...                 |
|                    | 대화인상표현<br>(Dialogue Expression)     | 감사, 미안함 등을 나타내는 화자의 의도<br>감사합니다. 고맙습니다. 죄송합니다, ...         |
| 요구<br>(Request)    | 진위정보요구<br>(InformationRequest-yn)   | 정보의 진위를 묻는 화자의 의도<br>A> 그때 되지도 많이 있습니까?                    |
|                    | 의문정보요구<br>(InformationRequest-wh)   | 정보의 내용을 묻는 화자의 의도<br>A> 너 시간 언제 되니?                        |
|                    | 확인정보요구<br>(InformationRequest-conf) | 정보의 확인을 묻는 화자의 의도<br>A> 저 그림 학생회관에서 만나자 학생회관서.<br>B> 학생회관? |
|                    | 행위요구<br>(ActRequest)                | 행위의 요구를 나타내는 화자의 의도<br>A> 저너도 할걸 저녁때쯤 만나자                  |
| 제공<br>(Offer)      | 정보제공<br>(InformationOffer)          | 정보를 제공하려는 화자의 의도<br>A> 객실 예약되었습니다.                         |
|                    | 응답제공<br>(ResponseOffer)             | 정보요구의 응답을 나타내는 화자의 의도<br>A> 예 전화했어?<br>B> 그냥 심심해서 걸었어.     |
|                    | 행위제안<br>(ActPropose)                | 행위를 제안하는 화자의 의도<br>A> 3월 10일자 예약을 좀 변경했으면 합니다.             |
|                    | 약속제공<br>(PromiseOffer)              | 행위의 약속을 나타내는 화자의 의도<br>A> 내가 끝나고 전화할게.                     |
|                    | 확인제공<br>(ConfirmationOffer)         | 정보의 확인을 나타내는 화자의 의도<br>A> 그림 그때 보자<br>B> 예 알겠어요            |
| 기타<br>(Etc)        | 기타<br>(Etc)                         | 놀라움, 망중단을 나타내는 화자의 의도<br>A> 하하하 ..                         |

3. 화수력의 출현 빈도

이러한 화수력의 출현 빈도를 조사하고 분류의 타당성을 검증하기 위하여 다양한 형태의 대화문을 수집하여 대화 말뭉치를 구성하였다. 그리고 화수력 자동 추출 알고리즘 개발을 위한 데이터 수집을 위해 수동적으로 본 대화 말뭉치들에 대해 화수력 태깅을 수행하였다. 표 2에서와 같이 화수력 분석을 위한 총 대화는 총 80대화이며 대화 당사자의 높임 관계와 대화내용을 기준으로 각 20대화씩 구분하였다. 대화 당사자들의 높임 관계는 대등, 주종 및 상호존중 관계로 구분하였으며 이는 각 발화문에 나타나는 다양한 형태의 종결어미를 분석하기 위함이다.

표 2. 화수력 태깅을 위한 대화 말뭉치 구성  
Table 2. A Construction of Dialogue Corpus for Tagging IFTs.

| 대화 유형   | 대화 관계 | 대화수 | 화수력 수 | 대화내용          |
|---------|-------|-----|-------|---------------|
| 일상 대화 1 | 대등    | 20  | 620   | 일반대화          |
| 일상 대화 2 | 주종    | 20  | 385   | 일반대화          |
| 일상 대화 3 | 상호 존중 | 20  | 478   | 일반대화          |
| 업무 대화   | 상호 존중 | 20  | 536   | 호텔예약을 위한 전화대화 |

다음 표 3은 화수력 태깅에 의한 각 화수력의 출현 빈도를 나타내고 있다. 분석 결과에 의하면 정보제공의 형태가 가장 빈번하며 응답제공, 확인제공, 정보요구도 자주 나타나는 화수력임을 알 수 있다. 본 대화 말뭉치가 전화 대화를 중심으로 구성되었기 때문에 확인제공의 형태가 예상외로 빈번히 나타남을 알 수 있다.

Ⅲ. 종결어미의 서법정보를 이용한 화수력 분석

한국어는 영어나 일본어와 같은 타언어와는 달리 서법의 유형에 따라 화자의 발화의도를 나타내는 화수력의 변화가 심하게 나타난다. 그림 2는 동사의 종결어미 부분에 나타나는 서법의 유형에 있어 차이가 있는 평서문과 의문문에 대한 구분, 의미 및 화행분석 과정을 보여 주고 있다. 두 발화문은 그 명제내용은 “비가 오다”로 동일하지만 화자의 발화의도를 나타내는 화수력에 있어 1)은 정보제공, 2)는 의문정보요구로서 차이가 난다.

이러한 종결어미의 서법정보에 의한 화수력의 차이

로 인해 시스템이 두 발화문을 분석하고 이해하는 논리표현식의 형태는 크게 차이가 있다.

표 3. 화수력의 출현 빈도  
Table 3. A Frequency of IFT's Occurence.

| 대화유형<br>화수력 | 업무대화<br>(상호존중) |        | 일반대화<br>(상호존중) |      | 일반대화<br>(주종) |      | 일반대화<br>(대등) |      | 전체대화 |      |
|-------------|----------------|--------|----------------|------|--------------|------|--------------|------|------|------|
|             | 빈도             | 비율     | 빈도             | 비율   | 빈도           | 비율   | 빈도           | 비율   | 빈도   | 비율   |
| 대화개시표현      | 19             | 3.1(%) | 19             | 4.9  | 15           | 3.4  | 16           | 3.2  | 69   | 3.4  |
| 대화종결표현      | 9              | 1.5    | 7              | 2.4  | 6            | 1.5  | 15           | 2.8  | 37   | 1.9  |
| 대화일상표현      | 31             | 5.0    | 7              | 1.8  | 6            | 1.0  | 2            | 0.9  | 46   | 2.3  |
| 전위정보요구      | 70             | 11.3   | 38             | 10.2 | 53           | 12.1 | 43           | 8.7  | 204  | 10.7 |
| 의문정보요구      | 67             | 10.8   | 33             | 8.4  | 49           | 9.2  | 55           | 8.7  | 204  | 9.4  |
| 확인정보요구      | 20             | 3.2    | 7              | 2.0  | 10           | 1.5  | 15           | 2.3  | 52   | 2.1  |
| 행위요구        | 30             | 4.8    | 15             | 5.5  | 47           | 9.0  | 24           | 5.2  | 116  | 6.2  |
| 정보제공        | 117            | 18.9   | 115            | 29.0 | 129          | 23.4 | 175          | 31.9 | 536  | 26.2 |
| 확인제공        | 113            | 18.2   | 59             | 15.7 | 65           | 13.7 | 88           | 15.5 | 325  | 16.1 |
| 행위제안        | 16             | 2.6    | 11             | 3.5  | 7            | 1.9  | 4            | 2.3  | 38   | 2.5  |
| 약속제공        | 19             | 3.0    | 13             | 1.6  | 6            | 1.2  | 14           | 2.8  | 52   | 2.3  |
| 응답제공        | 100            | 16.1   | 57             | 14.4 | 79           | 16.6 | 77           | 13.7 | 313  | 15.4 |
| 기타          | 9              | 1.5    | 4              | 0.9  | 6            | 1.7  | 8            | 2.0  | 27   | 1.5  |
| 총계          | 620            |        | 385            |      | 478          |      | 536          |      | 2019 |      |

이 때 ConvinceByInform은 정보를 제공하여 화자가 청자에게 그 내용을 확신시키는 논리 술어식이다<sup>14</sup>. 이와 같이 한국어는 종결어미의 서법 유형의 차이에 의하여 화자의 발화문에 대한 정신적 태도를 나타내는 화수력이 크게 달라짐을 알 수 있다.

1. 서법을 이용한 화수력 분석

화수력 분석의 중요한 단서가 되는 종결어미가 우리 말에서 차지하는 비중은 다른 문법 요소에 비해 아주 크다<sup>19,10</sup>. 즉 종결어미에는 문장을 종결시킨다는 기능적인 특수성 때문에 화자와 청자사이에 성립하는 사회적 등분인 준비법과 기능적 관계인 문체법이 성립하며 화자와 문장의 내용 사이에 성립하는 서법의 범주가 체계적으로 파악된다. 존대법과 서법에 대한 다양한 분류 방법이 존재하지만 본 논문에서는 표 4와 같이 존대법과 7가지 종류의 종결어미 서법에 따라 종결어미를 분류하였다. 표 4에 나타난 종결어미들은 각 높임법의 서법에 해당하는 국어사전에 수록된 종결어미들중 그 대표적인 것들을 나타내고 있다.

1) 비가 온다.  
2) 비가 오지?

< 구문 및 의미 분석 >  
1) (CAT VP)  
(HEAD (SYN (TENSE PRES) (VFORM PREDIC) (CHANGE TERM))  
(SEM (PRED 오) (SUBJ 비)))

2) (CAT VP)  
(HEAD (SYN (TENSE PRES) (VFORM IRROCATIVE) (CHANGE TERM))  
(SEM (PRED 오) (SUBJ 비)))

< 대화 분석 >  
1) Propositional Contents : 오다( 비/Subj )  
Illocutionary Force : 정보제공 (InformationOffer)  
SpeechAct : ConvinccByInform(Speaker, Hearer, 오다( 비 /Subj ))

2) Propositional Contents : 오다( 비/Subj )  
Illocutionary Force : 진의정보요구 (InformationRequest-yn)  
SpeechAct : MotivatebyRequest(Speaker, Hearer, ConvinccyInformIf(Hearer, Speaker, 오다( 비/Subj )))

그림 2. 화행분석의 일례  
Fig. 2. An Example of Speech Act Analysis.

표 4. 상대 높임법과 서법에 따른 현재형 종결어미

Table 4. The present endings of a word, according to the honorific expression and modal information.

| 서법 | 구분 | 종결어미             | 구분 | 종결어미                 |
|----|----|------------------|----|----------------------|
| 평서 | 낮춤 | -는다, -니다, -네     | 높임 | -오, -(으)소, -비니다      |
| 의문 | 낮춤 | -느냐, -니, -는가, -나 | 높임 | -오, -(으)소, -비니까      |
| 감탄 | 낮춤 | -누구나 -군, -누구면    | 높임 | -(는)구로, (는)군요, 비니다그러 |
| 명령 | 낮춤 | -(어)라, -(아)라, -게 | 높임 | -오, -(으)소, -(비)시오    |
| 청유 | 낮춤 | -자, -세           | 높임 | -(비)시다               |
| 허락 | 낮춤 | -렴, -으렴, -게나     | 높임 | -구로, (으)십시오          |
| 약속 | 낮춤 | -마, -오마, -미세     | 높임 | -리다, -오리다            |

2. 서법과 화수력

본 논문에서는 표 2와 같이 구성된 대화 말뭉치들의 각 발화문들을 수동적으로 분석하여 표 1에서 정의한 화수력들과 표 4에 나타난 종결어미들과의 상호관계를 조사하였다. 분석 결과, 화수력이 서법에 따른 종결어미에 많은 영향을 받는 것은 사실이었지만 발화문에 나타나는 어휘들의 종류와 이전 화수력과의 관계도 화

수력의 유형에 영향을 미치고 있음을 알 수 있었다. 다음 표 5는 서법 유형에 따른 종결어미, 어휘 양태 및 이전 화수력과의 관계에 따른 각 화수력들의 그 출현 특성들을 보이고 있다.

표 5. 화수력의 출현 특성  
Table 5. A Characteristics of IFT's Occurrence.

| 화수력    | 출현 특성      | 화수력  | 출현 특성        |
|--------|------------|------|--------------|
| 대화개시표현 | 어휘중심적      | 정보제공 | 종결어미 서법      |
| 대화종결표현 | 어휘중심적      | 응답제공 | 종결어미 서법 + 문맥 |
| 대화일상표현 | 어휘중심적      | 행위제안 | 종결어미 서법 + 어휘 |
| 진위정보요구 | 종결어미 서법    | 약속제공 | 종결어미 서법      |
| 의문정보요구 | 종결어미 서법+어휘 | 확인제공 | 문맥 + 어휘      |
| 확인정보요구 | 문맥 +어휘     | 기타   | 어휘중심적        |
| 행위요구   | 종결어미 서법    |      |              |

화수력은 크게 종결어미, 어휘, 문맥 중심적인 화수력으로 구분할 수 있었다. 어미 중심적인 화수력은 종결어미의 종류에 따라 화수력이 변화하는 것으로 진위정보요구, 정보제공, 행위요구 및 약속제공 등이 이에 해당한다. 어휘 중심적인 화수력은 관용적인 표현을 나타내는 대화표현과 말중단 등의 기타 등이 포함된다. 그리고 이전 화수력의 종류와 발화문내에 사용된 어휘들의 영향을 받는 문맥 중심적인 화수력으로는 확인정보요구, 응답제공 및 확인제공을 들 수 있고 의문정보요구와 행위제안은 어미와 어휘 중심적인 화수력임을 알 수 있다.

표 6은 실제 대화문의 분석을 통하여 나타난 종결어미 중심적인 화수력들인 진위정보요구, 정보제공, 행위요구 및 약속제공의 발화문상에 나타나는 종결어미의 종류를 보이고 있다. 실제 대화문에 나타나는 종결어미들은 표 4에서 제시한 국어 사전에 수록된 종결어미와 거의 일치하였으며 후어말 어미 “요”와 함께 나타나는 경우도 빈번함을 알 수 있었다.

진위정보요구에 나타나는 서법 유형은 의문형 종결어미와 의문형 서술격 조사이며 정보제공에서는 평서형 및 감탄형 종결어미와 서술격 조사가 나타난다. 본 논문에서는 서술격 조사 “이다”가 비록 종결어미는 아니지만 용언처럼 활용을 하면서 종결어미와 그 기능이 동일하고 서법의 유형도 나타나므로 종결어미의 서법 유형에 포함시킨다. 그리고 행위요구에서는 명령형, 청

유형 및 허락형이 나타나며 약속제공은 약속형 어미가 나타난다.

표 6. 화수력과 종결어미 서법과의 관계  
Table 6. A Relationship Between IFT and modal information in an ending of word.

| 화수력    | 종결어미 서법 유형 | 종결어미의 일례                    |
|--------|------------|-----------------------------|
| 진위정보요구 | 의문형 종결어미   | 습니까, 나요, 어요, 냐, 크까, ...     |
|        | 의문형 서술격 조사 | 이세요, 이에요, 어시죠, 이냐, 인가, ...  |
| 정보제공   | 평서형 종결어미   | 구요, 거든요, 어요, 네요, 니다, ...    |
|        | 평서형 서술격 조사 | 입니다, 이에요, 이거든요, 안테, 이다, ... |
|        | 감탄형 종결어미   | 구나, 니네, 더군, 더네, 더라, ...     |
|        | 감탄형 서술격 조사 | 어군, 어구나, 이라, 어군요, ...       |
| 행위요구   | 명령형 종결어미   | 어라, 어, 세요, 십시오, 시고요, ...    |
|        | 청유형 종결어미   | 자, 지, 르래, 야지, 죠, ...        |
|        | 허락형 종결어미   | 럼, 게, ...                   |
| 약속제공   | 약속형 종결어미   | 을께, 지요, 르께, 르께요, 지, ...     |

또한 의문정보요구는 진위정보요구에서 나타나는 어미와 같은 종류이지만 발화문의 구성성분중 의문사가 포함되어 있다는 면에서 그 차이가 있다. 따라서 의문사를 크게 의문체언, 의문관형사, 의문부사어 및 의문용언으로 분류하여 의문사표를 미리 구성함으로써 화수력 추출에 있어 진위정보요구와 의문정보요구의 구분은 가능하다.

3. 화수력 분석 알고리즘

본 논문에서는 이상과 같은 대화분석을 통한 화수력의 출현 특성을 이용하여 그림 3과 그림 4에서와 같이 2단계적인 화수력 추출 알고리즘을 제안한다. 본 알고리즘에서의 화수력 추출 과정은 종결어미나 서술격 조사가 존재하는 경우와 문맥현상으로 인한 단편적 발화문의 경우로 크게 나누어 진행한다. 그림 3에서와 같이 종결어미가 존재하는 경우는 입력 발화문의 종결어미 서법을 분석하여 각 서법의 특성에 맞는 처리를 수행한다. 이때 서법은 7가지로 구분되어 처리되며 서법에 따른 화수력 분석은 표 6의 화수력과 종결어미 서법과의 관계를 기준으로 한다.

그림 3의 서법정보를 이용한 알고리즘에서는 우선 화수력의 유형인 IFT를 NIL로 초기화시키고 입력 발화문이 판독적인 대화표현인가를 검사한다. 그리고 동사어미 사전을 참조하여 종결어미에 나타나는 서법의

유형을 판별하고 각 서법에 해당하는 처리부로 분기한다. 서법의 유형이 평서법 또는 감탄법인 경우, 보조용언의 판별은 보조적 연결어미 “고”와 희망을 나타내는 보조 형용사 “싶다” 등을 판단한다. 이와 같은 서법정보 이용은 화수력 추출 후보의 수를 효과적으로 제한함으로써 보다 정확한 화수력 추출이 가능하다.

```

Get Illocutionary_Force_Type1()
{
    화수력 유형(Illocutionary Force Type) := NIL; // 화수력 초기화
    if( 대화표현 어휘 존재 ) IFT := 대화표현;

    switch( 서법의 종류 ) {
        case 의문법 :
            if (의문사가 존재 || 선택 의문형 판별) then IFT := 의문정보요구;
            else if( 확인의문 판단 ) then IFT := 확인정보요구;
            else IFT := 진위정보요구;
            break;
        case 평서법 :
            감탄법 :
                if(이전 IFT가 정보요구이고 첫번째 발화문이다)
                    then IFT := 응답제공;
                if( 보조용언 판별 ) then IFT := 행위제안;

            if( IFT가 NIL이다 ) IFT := 정보제공;
            break;
        case 약속법 :
            if(이전 IFT가 정보요구이고 첫번째 발화문이다)
                then IFT := 응답제공;
            else IFT := 약속제공;
            break;
        case 청유법 :
        case 허락법 :
        case 명령법 :
            if(이전 IFT가 정보요구이고 첫번째 발화문이다)
                then IFT := 응답제공;
            else IFT := 행위요구;
            break;
        default : break;
    }
}
    
```

그림 3. 서법정보를 이용한 화수력 추출 알고리즘  
Fig. 3. An Algorithm for Extracting IFT, Based on Modal Information.

그림 4에서의 단편적 발화문에 관한 화수력 추출의 경우에는 이전 발화문에 나타나는 화수력 유형이나 현 발화문 구성성분의 어휘 종류 등을 기준으로 처리한다. 확인제공 어휘는 “네”, “맞습니다”, “그래요” 등과 같이 이전의 발화문의 정보를 긍정적으로 인정하여 확인하는 발화의도를 나타낸다. 이때 이전의 상대 발화문 화수력이 정보요구인 경우는 확인제공 어휘는 확인제공

이 아니라 응답제공으로 해석되기 때문에 대화 이력메모리상의 이전 화수력 정보들을 고려해야 한다.

의문형 심별이 존재하고 현재의 단편적 발화문이 이전 발화문의 일부 구성성분인 경우 이때의 화수력은 확인 정보요구이다. 확인 정보요구는 단일 어휘의 구성 성분으로도 가능하며 구를 이루어 반복하는 경우도 존재한다. 확인 정보요구가 아닌 경우는 의문형 서술격 조사가 생략된 진위정보요구로 판단한다. 이전 상대 발화문이 의문정보요구이고 심별이 의문형이 아닌 경우는 단편적인 발화문을 이용한 응답제공으로 판단하고 자신의 이전 발화문이 정보제공이면 계속적인 부가 설명으로 간주하여 정보제공으로 분석한다.

```

Get_Illocutionary_Force_Type2( )
{
    화수력 유형(Illocutionary Force Type) := NIL; // 화수력 초기화
    if( (이전 발화문 화수력이 정보요구가 아니다) and (확인제공 어휘 존재) )
        then IFT := 확인제공;

    if( 이전 IFT가 정보요구이고 응답 발화문이 존재 ) IFT := 응답제공;

    if( 의문형 심별이 존재 ) {
        if( 단편적 발화문이 이전 발화문의 구성성분이다 )
            then IFT := 확인정보요구;
        else IFT := 진위정보요구;
    }

    if( 이전 발화문이 의문정보요구이고 의문형 심별이 아니다 )
        then IFT := 응답제공;
    else if ( 자신의 이전 발화문이 정보제공이나 응답제공이다 )
        then IFT := 정보제공;

    if( IFT가 NIL이다 ) IFT := 기타;
}
    
```

그림 4. 단편적 발화문에 관한 화수력 추출 알고리즘  
Fig. 4. An Algorithm for Extracting IFT of an Elliptical Utterance.

#### IV. 실험 및 평가

본 논문에서 제시한 문맥 및 종결어미의 서법정보를 이용한 화수력 추출 알고리즘에 의하여 대화문의 각 발화문에 대한 화수력 태깅을 실시하였다. 본 실험에서 사용한 실험 데이터는 표 3에서 사용한 화수력 태깅을 위한 대화 말뭉치 이외의 다른 대화들을 사용하였다.

이 실험 대화 데이터 또한 다양한 종결어미의 특성을 반영하기 위하여 대등, 주종 및 상호존중 일반대화과 상호존중 업무대화로 구분되는 각 4대화씩 모두 16대화이며 대화상에 나타난 전체 화수력의 개수는 총 392개이다.

그림 5는 화수력 자동태깅을 위한 시범 시스템의 구성도를 나타낸다. 시스템의 입력은 형태소 태깅된 데이터를 사용한다. 각 발화문 단위의 입력 데이터에 대해 우선 종결어미나 서술격 조사의 존재여부를 검사하여 종결어미의 서법정보를 이용한 처리부와 문맥을 이용한 단편적 발화문에 관한 처리부로 분기한다. 그리고 그림 3과 4에서 제안한 화수력 추출 알고리즘을 바탕으로 화수력 분석을 실시한다.

이때 어미사전, 의문사표 및 대화 이력 메모리의 정보가 참조된다. 어미사전은 종결어미의 서법 유형을 판별하기 위해, 그리고 의문사표는 입력 발화문이 정보요구의 형태일때 의문사의 존재 여부를 판별하기 위해 사용된다. 대화이력 메모리는 단편적 발화문과 응답제공 등의 화수력을 판별하는데 있어 이전 발화문의 화수력 정보를 참조하기 위해 사용된다. 또한 확인정보요구를 판별하기 위해 이전 발화문의 구성 항목들의 어휘들이 대화이력 메모리에서 참조된다. 본 시스템의 문맥 복원부는 이전 화수력 발화문이 의문정보요구이고 현재의 발화문이 단편적 발화문인 경우 등의 간단한 서술어 복원을 그 대상으로 한다.

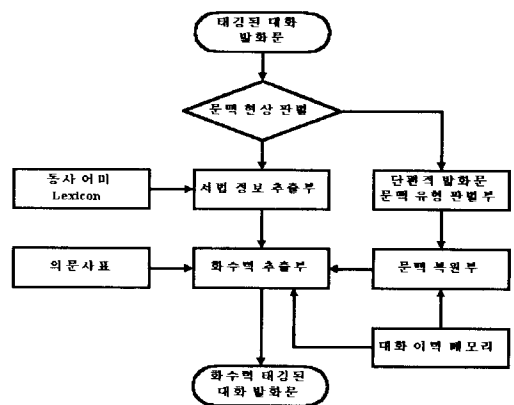


그림 5. 화수력 자동 태깅을 위한 시스템 구성도  
Fig. 5. A Configuration of System to Perform the Automatic Tagging of IFTs.

본 실험에서 사용한 대상 대화문은 품사태깅이 되어 있는 문장을 그 입력으로 하였다. 적용대상 문장이 대

화체 문장이므로 기본적인 형태소 해석을 수행한 후 수동적으로 품사 태깅을 실시하였다. 다음 표 7은 본 대화문의 품사태깅을 위한 품사태그 집합을 보여 주고 있다. 일반적으로 태깅을 위한 품사의 갯수는 시스템마다 다양한 차이를 보이고 있지만<sup>[11]</sup> 본 논문에서의 품사태그 집합은 종결어미, 후어말어미, 보조적 연결어미의 정보를 이용하기 위하여 어미 부분을 보다 상세히 분류하였다.

표 7. 대화문 품사태깅을 위한 태그 집합  
Table 7. Tag Set for Part-Of-Speech Tagging in a Dialogue.

| 품사    | 기호   | 일례            | 품사      | 기호     | 일례              |
|-------|------|---------------|---------|--------|-----------------|
| 명사    | Noun | 방, 약속, 책, ... | 전성어미    | JEomi  | ㅁ, ㄴ, 기, ...    |
| 대명사   | Pro  | 저, 너, 우리, ... | 선어말어미   | SEomi  | 시, 겠, 읊, ...    |
| 수사    | Num  | 1, 하나, 두, ... | 연결어미    | YEomi  | 고, 어서, ...      |
| 동사    | Verb | 만나다, 오다, ...  | 보조적연결어미 | BEomi  | 는게, 지, 고 ...    |
| 형용사   | Adj  | 빠르다, ...      | 종결어미    | EEomi  | 다, 냐, 는군 ...    |
| 보조동사  | AuxV | 하다, 드리다, ...  | 후어말어미   | HEomi  | 요               |
| 보조형용사 | AuxA | 있다, ...       | 관형사     | PreN   | 이, 저, 그, ...    |
| 부사    | Adv  | 좀, 빨리, ...    | 접사      | Suffix | 씨, 들, ...       |
| 감탄사   | Int  | 예, 어, 음, ...  | 심벌      | Symbol | ?, !, ,, ,, ... |
| 조사    | Josa | 에서, 을, ...    |         |        |                 |

그림 6은 실제 대화문의 입력 발화문에 대한 화수력 태깅 결과를 보여 준다. 대화문의 각 발화문은 DJ2-A11과 같이 구분되며 DJ2는 주중관계의 2번째 대화를 나타내고 A11은 대화자 A의 제 1의 첫번째 발화문임을 나타낸다. 각 발화문은 '/' 이후에 표 7의 태그집합들로 태그들이 부여되어 있으며 ( )속에서 수동적으로 화수력을 태깅한 결과들이 보이고 있다. 그리고 각 3번째 줄에 입력 발화문들에 대한 본 시범 시스템에 의한 화수력 분석 결과들이 보이고 있으며 이들은 수동적인 태깅 결과들과 모두 일치하고 있음을 알 수 있다.

표 8은 본 알고리즘에 의한 화수력 추출 실험 결과를 나타내고 있다. 본 시스템에 의해 추출된 태깅 결과들과 이전에 수동적으로 태깅한 결과들을 비교하여 화수력 추출 성공율을 산출하였다. 그 결과 전체적으로 94.4%라고 하는 좋은 결과를 얻을 수 있었다. 추출 결과, 어휘 의존적인 화수력인 대화표현과 기타 및 문맥 중심적인 화수력의 추출결과가 높은 성공율을 나타내고 있다.

|          |  |
|----------|--|
| DJ2-A11> | 호진이니?<br>호진/Noun 이니/Josa ?/Symbol<br>(InformationRequest-yn)<br>InformationRequest-yn  |
| DJ2-B11> | 예.<br>예/Int /Symbol (ResponseOffer)<br>ResponseOffer   |
| DJ2-A21> | 호진아, 요새 어떻게 지내니?<br>호진/Noun 아/Josa /Symbol 요새/Adv 어떻게/Adv 지내<br>/Verb 니/EEomi ?/Symbol<br>(InformationRequest-wh)<br>InformationRequest-wh |
| DJ2-B21> | 요즘 수업때문에 좀 바쁩니다.<br>요즘/Adv 수업/Noun 때문/Noun 예/Josa 좀/Adv 바쁘<br>/Verb 합니다/EEomi /Symbol<br>(ResponseOffer)<br>ResponseOffer                  |
| DJ2-B22> | 형은 어떠세요?<br>형/Noun 은/Josa 어떻/Adj 세요/EEomi ?/Symbol<br>(InformationRequest-wh)<br>InformationRequest-wh                                     |
| DJ2-A31> | 쩔 나도 역시 바빠<br>쩔/Int 나/Pro 도/Josa 역시/Adv 바쁘/Verb 아/EEomi<br>(ResponseOffer)<br>ResponseOffer  |
| DJ2-A32> | 언제 날 잡아서 저녁이나 함께 하자.<br>언제/Adv 날/Noun 잡/Verb 아서/YEomi 저녁/Noun 이나<br>/Josa 함께/Adv 하/Verb<br>자/EEomi /Symbol (ActRequest)<br>ActRequest      |

그림 6. 화수력 분석의 일례  
Fig. 6. An Example of IFT Analysis.

표 8. 화수력 분석 결과  
Table 8. A Result of IFT Analysis.

| 화수력    | 빈도 | 성공 | 실패 | 성공율(%) | 화수력  | 빈도  | 성공  | 실패 | 성공율   |
|--------|----|----|----|--------|------|-----|-----|----|-------|
| 대화개시표현 | 11 | 11 | 0  | 100.0  | 정보제공 | 102 | 98  | 4  | 96.1  |
| 대화중절표현 | 10 | 9  | 1  | 90.0   | 응답제공 | 66  | 62  | 4  | 93.9  |
| 대화일상표현 | 12 | 12 | 0  | 100.0  | 행위제안 | 11  | 10  | 1  | 90.9  |
| 진위정보요구 | 42 | 40 | 2  | 95.2   | 약속제공 | 7   | 6   | 1  | 85.7  |
| 의문정보요구 | 27 | 23 | 4  | 85.2   | 확인제공 | 61  | 59  | 2  | 96.7  |
| 확인정보요구 | 7  | 7  | 0  | 100.0  | 기타   | 8   | 8   | 0  | 100.0 |
| 행위요구   | 28 | 25 | 3  | 89.3   | 총계   | 392 | 370 | 22 | 94.4  |

그러나 종결어미의 서법정보를 이용하는 의문정보요구는 의문사의 어휘를 판단하는데 있어서의 모호성으로 인하여 그 추출 성공율이 다소 떨어지며 약속제공과 행위요구에서는 “지요”와 같은 종결어미가 다수의 서법 형태로의 분석이 가능하기 때문에 발생하는 애매성으로 인하여 추출 실패하는 경우가 발생하였다. 이와 같은 서법의 다중성으로 인한 오류는 동사의 실제적인



주체를 판단함으로써 해결할 수 있으리라 판단된다.

## V. 결 론

본 논문에서는 입력 발화문의 화행분석을 위하여 문맥 및 어휘 정보외에 동사의 종결어미에 나타나는 서법정보를 이용한 화수력 분석 알고리즘을 제안하였다. 그리고 실제 대화 데이터를 분석하여 화수력을 분류하였으며 분류된 화수력의 출현 빈도를 제시하였다. 화수력 태깅한 데이터를 분석하여 화수력과 종결어미 서법과의 상호 관계를 보였고 이를 바탕으로 화수력 분석 알고리즘을 제시하였다. 본 화수력 분석 알고리즘은 종결어미나 서술적 조사가 존재하는 경우와 존재하지 않는 단편적 발화문의 경우로 구분하여 처리한다.

본 논문에서는 기존의 수동적이고 주관적인 화수력 분석 방식에 대해 계산학적으로 구현 가능한 방법을 제시했다는 점에서 그 의의가 있으며 적용대상 영역에 의존적이지 않다는 장점도 지닌다. 그리고 본 논문에서 제시한 화수력 추출 방식이 수동적인 분석방식과 비교하여 그 결과가 큰 차이가 없음을 실험을 통하여 검증할 수 있었다.

본 논문에서 제안하는 화수력 분석 방식은 향후 음성언어 번역 시스템에서 대화 모델링을 위한 데이터로 사용될 뿐만 아니라 다음 발화문의 예측정보로서 오인식 후보의 선별에 도움이 될 수 있으며, 자연언어 인터페이스를 이용한 지능적 대화시스템에 있어서는 사용자 의도 분석을 위한 중요한 데이터로 사용될 수 있으리라 생각된다.

그러나 본 논문의 적용 대상이 일반 정형적 문장과는 다른 대화체 문장이므로 도치, 생략, 발중단, 머뭇거림 등의 다양한 언어 현상을 처리할 수 있어야 한다. 그리고 복수개의 서법 유형을 지니는 종결어미의 애매성 해소도 지속적으로 고려되어야 할 사항이며 간접화행에 대한 분석 방법도 연구되어야 한다. 그리고 보다

다양한 많은 대화 데이터를 분석하여 더욱 신뢰성 있는 통계 자료가 구축되어야 한다.

## 참 고 문 헌

- [1] 김영길, 최병욱, "대화정보를 이용한 계획인식 기반형 자연언어 대화이해 시스템의 설계 및 구현", 대한전자공학회 논문지, vol. 33-B, no. 3, pp 159-168, Mar. 1996
- [2] 豊田順一, 北橋忠宏, 溝口理一郎, "會話音聲の認識・理解", 電子情報通信學會誌, vol. 73 no. 12 pp. 1299-1303, Dec. 1990.
- [3] 윤홍섭, "상황의미론에 의한 화행분석", 화용론 논집, pp 17-27, 1991, 8
- [4] James F Allen, Natural Language Understanding, 2nd Edition, The Benjamin/Cummings Publishing Company, Inc, 1995.
- [5] 장석진, 話用과 文法, 탑출판사, 1993
- [6] James F Allen, "Using Structural Constraints For Speech Act Interpretation", Speech and Natural Language, pp. 385-401, 1989.
- [7] Sandra Carberry, Plan Recognition in Natural Language Dialogue, A Bradford Book, 1990.
- [8] Masaaki Nagata, Tsuyoshi Morimoto, "An Information-Theoretic Model of Discourse for Next Utterance Type Prediction", Transactions of Information Processing Society of Japan, Vol. 35 No. 6, June, 1994.
- [9] 한 길, 국어종결어미연구, 강원대학교 출판부, 1991
- [10] 남기심, 고희근, 표준국어문법론, 탑출판사, 1993
- [11] 이운재, 한국어 문서 태깅 시스템의 설계 및 구현, 한국과학기술원 전산학과 석사학위 논문, 1993

## 저 자 소 개

金永吉(正會員) 第33卷 第3號 參照

현재 한양대학교 전자통신공학과  
박사과정

崔炳旭(正會員) 第32卷 第2號 參照

현재 한양대학교 전자통신학과 교수