

2017. 12. 20.

독어학연습(대화분석의 이론과 실제)

장수은(2016-24958)

## 상용 대화 에이전트 분석을 위한 확장 화행 이론의 변용

### 1. 들어가는 말

기술의 발전에 따라 컴퓨터 시스템은 계산이나 반복 같은 일뿐 아니라 사람의 언어 활동까지 모방하기 시작했다. 컴퓨터 시스템이 사람의 언어 활동을 모방한다는 것은 크게 두 가지로 나뉘 볼 수 있는데, 하나는 사람이 언어를 귀로 듣고 그것이 어떤 소린지, 어떤 의미가 있고 어떤 역할을 하며 어떤 목적을 갖는지를 이해하는 전체 과정을 의미하고, 다른 하나는 앞의 모든 과정을 반대로, 머릿속의 생각을 언어로 생성해내는 과정을 의미한다. 언어학 연구가 과거에 각 부분을 따로 떼어 작은 것부터 차차 밝혀왔듯, 공학 분야에서도 사람의 언어 활동을 모방하기 위해 각 작은 부분부터 연구되기 시작했다. 실제 컴퓨터언어학의 연구 현황을 살펴보면, 단어의 품사를 구분하는 문제(Part-of-speech tagging), 고유명사를 인식해내는 문제(Named entity recognition), 이메일 문서가 있을 때 그것이 스팸인지 아닌지를 구분하는 문제(Spam detection) 등 대체적으로 단어-문장 수준의 언어학 연구와 관련된 문제들은 기술적으로 거의 다 풀렸고, 반대로, 질문이 주어졌을 때 자료에서 답을 찾아 답변을 만들어내는 문제(Question answering), 문장을 다른 말로 바꾸어 표현하는 문제(Paraphrase), 문서를 요약하는 문제(Summarization), 대화(Dialog) 등 문장-맥락 수준의 언어학 연구와 관련된 문제들은 여전히 매우 어려운 문제로 남아있다(Jurafsky et al., 2009). 이처럼 언어는 갈수록 복잡성이 드러나고 또 기술적으로도 여전히 어려운 문제로 남아있음에도 불구하고 인간을 인간으로서 특징짓는 특별한 행위이기에 언어학과 공학, 양쪽에서 계속해서 연구되고 있다.

여전히 인간의 대화에 대해서는 밝혀져야 할 것이 많다. 그러나 공학에서는 현재 가능한 기술만으로도 인간의 언어 행위를 모방하는 시스템을 만들고 상용화함으로써 실제와 이상의 간격을 줄여 나가려 하고 있는데, 요즘 특히 화두가 되고 있는 것이 바로 대화 시스템이다. 대화 시스템은 사람과 '대화'의 방법으로 소통을 하는 컴퓨터 시스템의 종류로, 물류 회사 아마존이 내놓은 가정용 음성 대화 스피커 '에코'가 인기를 끌면서 특히 관심을 받고 있다. 한국어의 경우 2016년 이후부터 제품, 서비스가 상용화되기 시작해, 지금은 약 10개의 상용 대화 시스템이 있다. (표 1)

이처럼 대화 시스템은 '대화 에이전트'로서 사람의 대화 상대로 성장해 나가고 있는데, 과연 대화 에이전트들의 대화 수준은 어떨까? 대화 상대로서 적합한 수준일까? 대화 에이전트의 대화 능력을 측정하기 위해 공학에서는 정확도를 비롯한 개념을 사용하고 있지만, 대화에서 정확도란 매우 애매한 개념이다. 답을 했는가의 여부로 정확성을 판단할 것인가? 혹은 답이 충분한가의 여부로, 상대의 만족도를 근거로 판단할 것인가? 이러한 문제에 해답을 제시할 수 있는 것은 바로 언어학적 접근이다. 대화를 어떻게 정의해야 하는가에 대한 연구부터 대화의 조건 및 성공적인 대화의 조건까지 다양한 논의를 오랜 시간 진행해왔기 때문이다. 본 연구는 이러한 맥락에서 확장 화행 이론이라는 화용론적 접근 방식을 사용해 현재 상용되고 있는 대화 에이전트들의 수준을 기술해보고자 한다. 이를 통해, 기술적, 표면적에서의 접근이 아니라 실제 사용의 측면에서 접근함으로써 단순히 그들의 수준을 가늠하는 것을 넘어 대화 시스템의 발전 방향에도 어떤 지향점을 보여줄 수 있기를 기대한다.

표 1<sup>1</sup> (가나다순)

이름	구글 어시스턴트	기가지니	누구	빅스비	시리
형태					
응용	홈(스피커), 알로(메신저)	기가지니LTE (스피커)	누구미니 (스피커)	-	-
표현	말, 글, 화면	말	말	말, 글, 화면	말, 글, 화면
용도	개인용	가정용	가정용	개인용	개인용
출시	2017. 9. 21.	2017. 1. 17.	2016. 8. 31.	2017. 4. 21.	2012. 6. 11.
이름	쌍큐허브	에이브릴	카카오아이	클로바	핑퐁
형태					
응용	*클로바 기반	*왓슨 기반	카카오미니 (스피커)	웨이브(스피커), 프렌즈(스피커)	-
표현	말, 글, 화면	용도에 따라 다양	말	말, 글, 화면	글
용도	가정용	다양	가정용	개인용	개인용
출시	2017. 4. 26.	2017. 9. 6.	2017. 9. 20.	2017. 5. 12.	-

<sup>1</sup> 각 이미지 자료의 출처는 순서대로 <https://assistant.google.com>,  
<https://gigagenie.kt.com/main.do>, <http://www.nugu.co.kr/product.jsp>,  
<http://www.samsung.com/sec/apps/bixby/voice/>, <https://www.apple.com/kr/ios/siri/>,  
<http://www.lge.co.kr/lgekor/product/accessory/smart-life/productDetail.do?catId=8220&prdId=EPRD.313033>, <https://aibril-car-demo-korean.sk.kr.mybluemix.net>, <https://kakao.ai/product/kakaomini>, <https://clova.ai/ko/ko-product.html>, [http://www.scatterlab.co.kr/?page\\_id=24366](http://www.scatterlab.co.kr/?page_id=24366).

## 2. 연구 자료 수집

### A. 수집처

대화 에이전트와 사람 사이의 대화 분석을 위한 자료를 수집하기 위해 온라인 동영상 플랫폼인 유튜브를 활용하였다. 기존의 한국어 대화 분석 연구는 보통 영상 자료를 직접 촬영하여 전사, 분석한다. 하지만 대화 에이전트와 사람의 대화는 사람과 사람의 대화와 달리 시행된 지 오래되지 않아 일상에서 보편적으로 발견하기 어렵고, 발견된다고 하더라도 개인의 기술적 지식이나 숙련도에 따라 대화 흐름이 크게 달라질 수 있어 대표성을 갖기도 어렵다. 따라서 직접 촬영하여 분석 자료로 사용하기보다는 유튜브를 활용하는 것이 효과적이라고 판단하였다. 또한, 대화 에이전트의 사용자는 신기술에 전반적으로 관심이 많은 이들이 주를 이루는데, 그러한 사용자가 모이는 곳이 동영상 플랫폼인 유튜브라 판단하였다. 유튜브는 전 세계인이 사용하는 거대 플랫폼인 만큼 전체 동영상의 개수나 기술 관련 동영상의 비율을 정확히 확인하기는 어렵지만, 현재 약 5백만 개의 사용 후기 동영상이 있는 것으로 확인되며, 대화 자료를 수집하기에 충분한 자원이 있다고 보았다.

### B. 자료 수집

유튜브에서 연구 자료를 얻기 위해 표 1에서 제시한 10개의 대화 에이전트를 모두 대상으로 삼고 “대화 에이전트 이름 (사용 OR 후기)” 형태로 검색하였으나, ‘에이브릴’의 경우 대화 외에도 다양한 기능을 가진 인공지능 플랫폼이어서 활용 방식에 따라 대화 양상이 크게 달라질 수 있는 데다 현재 상용화되어 있는 서비스를 확인하기 어려워 제외하였고, ‘핑퐁’의 경우 다른 대화 에이전트와 달리 일상 대화를 중심으로 하여 대화 에이전트의 확장 가능성에 함의를 줄 수 있을 것으로 기대되었으나 출시되어 상용되고 있는 것을 확인하기 어려워 제외하였다. 따라서 총 8개의 대화 에이전트를 검색 대상으로 삼았고, 각 검색 결과 중 제조사에서 광고를 목적으로 제작, 배포한 동영상을 제외한 상위 5개 동영상을 분석 자료로 삼았다. 이로써 총 40개의 영상을 수집하였으며, 붙임 2의 ‘전사 자료 전체 목록’에서 각 자료의 정보를 확인할 수 있다.

### C. 전사

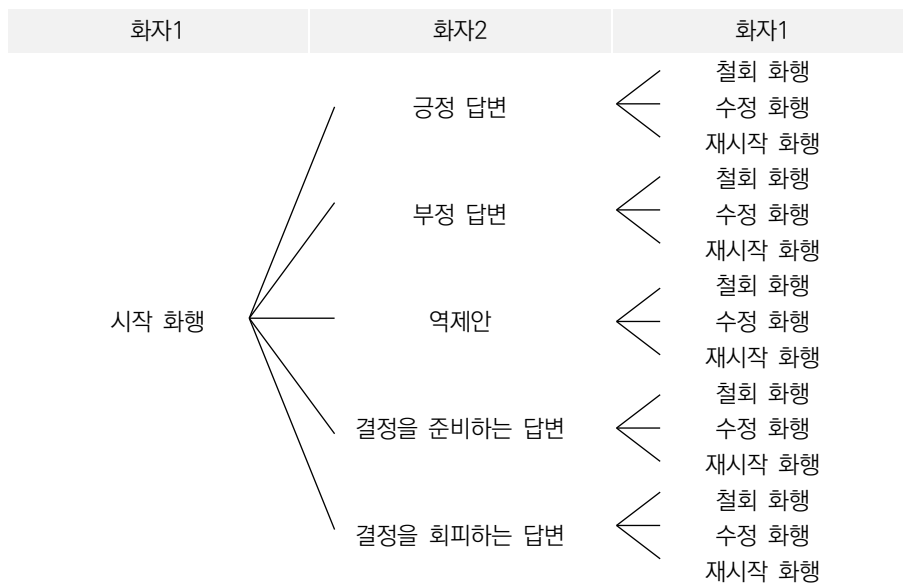
유튜브에서 수집한 영상 자료는 붙임 1의 전사 협약에 기반을 두고 전사하였다. 이때 주의한 점은, 무엇보다 유튜브의 후기 영상은 다른 예비 사용자 혹은 시청자들에게 기술을 소개, 설명하는 것을 목적으로 하므로 모든 언어적 행위가 대화 에이전트와의 대화로만 구성되지는 않는다는 점이다. 예를 들어 시청자에게 인사를 건네거나 시청자에게 대화 에이전트를 평가한다. 따라서 본 연구에서는 대화 에이전트와 사람 사이의 대화라 볼 수 없는 부분은 제외하고 전사하였다. 또한, 수집한 자료에서 사용자들은 음성인식 기술이나 음성합성 기술이 아직 자연스럽지 못하다는 사실을 인지하고 있어 비교적 짧고 명확한 문장을 또박또박, 천천히 말하는 모습을 보였다. 따라서 기존 전사 협약에서 쓰이던 강세, 억양, 속도 등의 표기는 무시하였고, 보편적으로 띄어 쓰지 않더라도 음성적으로 휴지가 있을 때는 띄어 썼으며, 휴지가 길 때는 줄을 더 띄었다. 이 같은 규칙에 따라 전사한 자료는 총 581개의 최소대화로 이루어져 있으며, 붙임 3에서 확인할 수 있다.

### 3. 상용 대화 에이전트 분석

#### A. 확장 화행 이론과 최소대화

프랑케의 확장 화행 이론(Franke, 2010)은 언어가 그 자체로 행위(화행, speech act)를 동반한다는 관점에서 언어를 분석하는 방법 중 하나의 이론적 바탕으로, 기존 화행 분석법과 달리 화행의 연속성에 주목하는 것이 특징이다. 다시 말해, 프랑케의 확장 화행 이론은 설(Searle)의 전통적인 다섯 가지 화행 분류법<sup>2</sup>을 넘어, 이전 화행에의 응답으로서의 행위에 주목한다. 확장 화행 이론에 따르면(그림 1) ‘시작 화행’은 상대 화자에 의해 ‘긍정 답변’/‘부정 답변’과 같은 명확한 답변, 시작 화행에 반하는 ‘역제안’, ‘결정을 준비하는 답변’/‘결정을 회피하는 답변’과 같은 불명확한 답변으로 이어질 수 있다. 시작 화행을 따르는 이러한 답변들은, 또다시 시작 화행의 수용 여부에 따라 ‘철회 화행’, ‘수정 화행’, ‘재시작 화행’으로 이어질 수 있다.

그림 1



이처럼 확장 화행 이론은 두 명의 화자를 기준으로, 첫 번째 화자가 목적을 가지고 시작 화행을 했을 때 그 목적이 두 번째 화자, 즉 대화 상대자에게 어떻게 받아들여지는가를 기준으로 두 번째 화행을 구분하고, 두 번째 화자가 받아들이는 결과가 다시 첫 번째 화자에게 어떻게 받아들여지는가를 기준으로 세 번째 화행을 구분한다. 이어, 확장 화행 이론은 두 번째 차례에서 두 번째 화자가 스스로 행했던 발화를 세 번째 화행 이후 어떻게 받아들이냐에 따라 또다시 ‘철회 화행’, ‘수정 화행’, ‘재시작 화행’이 반복됨을 보임으로써, 이후 차례에 나타나는 대화를 순환성으로 정리하고 ‘최소대화모형’을 제시한다. 최소대화모형은 첫 번째 화자가 자신의 목적을 달성하여 대화를 끝마치게 되는, 최소 개수의 화행으로 이루어진 대화모형을 의미한다. 결과적으로 확장 화행 이론은 총 여덟 개의 최소대화모형과 두 개의 종결되지 않은 대화모형<sup>3</sup>(표 2)을 제시한다.

<sup>2</sup> 진술, 지시, 약속, 표현, 선언 (Representative, Directive, Commissive, Expressive, Declaration)

<sup>3</sup> 종결되지 않은 대화모형은, 첫 번째 화자의 초기 목적이 달성되지 못한 채 두 화자가 끝없이 화행을 주고받는 대화를 의미한다.

표 2<sup>4</sup>

	화자1	화자2	화자1	화자2	화자1
최소대화1	시작 화행	긍정 답변			
최소대화2	시작 화행	부정 답변	철회 화행		
최소대화3	시작 화행	거부	철회 화행		
최소대화4	시작 화행	역제안	수정 화행	철회 화행	
최소대화5	시작 화행	부정 답변	수정 화행	철회 화행	
최소대화6	시작 화행	역제안	재시작 화행	철회 화행	
대화7	시작 화행	역제안	재시작 화행	재시작 화행	...
최소대화8	시작 화행	부정 답변	재시작 화행	철회 화행	
대화9	시작 화행	부정 답변	재시작 화행	재시작 화행	...
최소대화10	시작 화행	거부	재시작 화행	재시작 화행	...

그러나 위의 표에 나타난 확장 화행 이론의 열 가지 대화모형은 충분치 않다. 두 번째 화행에서 두 번째 화자가 할 수 있는 불명확한 답변으로 일부 사례인 ‘거부’만 포함되었을 뿐 아니라, 두 번째 화행의 종류마다 철회/수정/재시작 화행의 세 반응이 따를 가능성이 무시되었기 때문이다. 이 모든 경우를 다 포함할 수 있다고 했을 때, 첫 번째 화행에서 가능한 화행이 1개(시작 화행), 두 번째 화행에서 가능한 화행이 5개(긍정 답변, 부정 답변, 역제안, 결정을 준비하는 답변, 결정을 회피하는 답변), 세 번째 화행에서 각각 가능한 화행이 3개(철회 화행, 수정 화행, 재시작 화행)이므로, 세 번째 차례까지의 가능성만 따져보아도 총 15개의 가능한 대화 유형이 있음을 짐작해 볼 수 있다. 또한, 여기에 네 번째 화행으로 각 가능성에 대해 또 다른 세 가지 가능성(철회 화행, 수정 화행, 재시작 화행)을 더하면 논의해 볼 수 있는 대화모형은 훨씬 더 늘어난다. 본 연구는 대화 에이전트와 사람 사이의 대화를 분석하는 데 적절한 대화모형을 제시하기 위해, 실제 대화 사례를 분석함과 동시에 대화모형의 이러한 확장 가능성을 다시 살핌으로써 기존 대화모형을 검토하고자 한다.

그러나, 과연 대화 에이전트를 분석하는 데 확장 화행 이론을 틀로 삼는 것이 적절한 걸까? 이를 판단하기 위해서는 확장 화행 이론과 분석 대상, 즉 대화 에이전트와 사람 간 대화의 공통점에 주목할 필요가 있다. 둘은 모두 두 명의 화자가 대화에 참여하는 것을 바탕으로 하며, 한 화자가 다른 화자를 향해 어떤 의도를 가지고 대화를 시작한다는 점에 공통점이 있다. 다시 말해, 확장 화행 이론은 첫 번째 화자가 목적을 이루어가는 과정을 화행과 응답의 관점에서 기술하는 하나의 방식이고, 대화 에이전트와 사람 간 대화는 대화 에이전트가 사람(사용자)의 요구, 즉 정보나 행동에 대한 요구에 대응하는 과정이기 때문에 전자는 후자를 분석하는 데 적합한 도구가 될 수 있다는 것이다. 따라서 본 연구는 확장 화행 이론에서 최소대화라고 불리는 대화모형을 대화 분석의 기본 단위로 삼아 대화 에이전트를 분석하고자 하며, 필요에 따라 ‘기능단계’ 개념을 도입해 단순대화모형을 복잡대화모형까지 확장하고자 한다. 또한, 이 과정에서 확장 화행 이론에 어떤 변용이 필요할지도 논의할 것이다.

## B. 확장 화행 이론 기반의 상용 대화 에이전트 분석

이 장에서는 확장 화행 이론의 최소대화모형이 실제 대화 에이전트 사용에서 그 사례가 나타나는지, 혹은 현재 사례는 없지만 새로운 대화모형으로서 발전시킬 수 있을지, 예를 통해 확인해보고자 한다. 첫 번째 차례에서는 무조건 시작 화행이 오는 것으로 보았으며, 두 번째 차례에서 올 수 있는 화행은 (i) 긍정 답변 (ii) 부정 답변 (iii) 역제안 (iv) 결정을 준비하는 답변 (v) 결정을 회피하는 답변으로 보았다.

<sup>4</sup> 두 번째 화행에서 두 번째 화자의 ‘거부’는 (명확하지 않게) ‘결정을 회피하는 답변’ 중 하나의 사례로 제시되었다.

### i. 시작 화행 + 긍정 답변

대화 자료에 나타난 581개의 최소대화 중 530개, 즉 약 91%가 긍정 답변으로 응답됐다. 이는 현재 상용화돼 있는 대화 에이전트들이 음성 명령을 통해 특정 기능을 수행하기 위한 목적으로 만들어졌기 때문이다. 다음 대화는 확장 화행 이론의 첫 번째 최소대화모형의 예시이자(표 2) 사람과 대화 에이전트의 대화에서 가장 전형적인 사례다. 사용자가 먼저 대화 에이전트를 호출한 후 원하는 기능에 따른 명령어를 말하면 대화 에이전트가 긍정 답변으로 반응을 하며 해당 기능을 수행한다. (17번, 1-3줄)

- 1 화자1<sup>5</sup>: 아리아.
- 2       나로 말할 것 같으면 들려줘.
- 3 화자2: 마마무의 나로 말할 것 같으면 들려드릴게요.

시작 화행-긍정 답변 유형은 특정 기능이나 과제를 수행하기 위한 목적이 아니라 친목 도모나 사회적 목적으로 쓰이기도 한다. (27번, 1-3줄)

- 1 화자1: 빅스비 안녕.
- 2 화자2: 안녕.
- 3       어서와요.

이처럼 시작 화행과 긍정 답변으로 이루어진 단순한 대화모형은 항상 혼자 쓰이지는 않는다. 즉, 홀로 독립적인 대화로서 기능하기도 하지만 다른 (최소)대화모형과 함께 쓰여 복합대화모형을 구성하기도 한다. 확장 화행 이론이 제시하고 있는 복합대화모형은 총 세 가지가 있다. 첫째, 최소대화모형 하나로 이루어진 대화모형, 둘째, 둘 이상의 최소대화모형으로 이루어져 있으면서 그 최소대화모형들이 같은 구조의 반복인 대화모형, 셋째, 둘 이상의 최소대화모형으로 이루어져 있으면서 그 최소대화모형들이 서로 다른 구조인 대화모형이다. 첫째 유형이 앞서 살펴본 예에 해당하며, 이러한 복합대화모형들에 상황, 맥락 등 복잡한 요소까지 추가되면 비로소 일상 대화에 가까워진다. 우선 복잡한 요소들은 제외하고, 시작 화행-긍정 답변의 최소대화모형이 복합대화모형으로 확장되어 사용된 경우를 살펴보자. 다음은 총 62개의 복합대화 중 9개, 약 15%를 차지하는 둘째 유형의 복합대화 사례다. (01번, 9-15줄) 9-10과 11, 12와 13, 14와 15는 각각 시작 화행과 긍정 답변으로 이루어진 최소대화로, 구조가 같은 세 대화가 반복되어 하나의 복합대화를 이루고 있다.

- 9 화자1: 클로바.
- 10       오늘 달러 환율 알려줘.
- 11 화자2: 현재 미국 달러 환율은 천백이십 원입니다.
- 12 화자1: 엔화는?
- 13 화자2: 현재 일본 엔화 환율은 구백구십일 원입니다.
- 14 화자1: 그럼 유로는?
- 15 화자2: 현재 유럽연합 유로 환율은 천삼백사 원입니다.

03:46

셋째 유형의 복합대화는 총 62개의 복합대화 중 32개, 약 52%를 차지한다. 이 유형으로는 대화 자체의 사회적, 제도적 특성상 여러 구조가 복합되어 있는 대화(예1: 처음 만난 사람끼리 인사를 나누고 이어 서로 소개를 하는 대화, 01번, 1-4줄; 예2: 퀴즈 맞추기 놀이를 요청한 후 돌아가며 퀴즈 맞추기 게임을 진행하는 대화, 16번, 71-81줄), 특정 기능을 시작한 상황에서 부가적, 세부적인 기능을 원하는 대

---

<sup>5</sup> 특별한 언급이 없다면 본 글에서 인용하는 모든 대화의 화자1은 사람(사용자), 화자2는 대화 에이전트다.

화(예1: 음악을 틀어 달라고 한 후 음량을 조절하기를 원하는 대화, 04번, 7-12줄; 예2: 영화 줄거리를 설명해 달라고 요청해 듣고 있는 도중 중단하기를 원하는 대화, 16번, 38-46줄)가 있다. 다음은 후자의 예1이다.

- 7 화자1: 클로바. 03:33  
 8 사랑의 온도 오에스티 틀어줘.  
 9 화자2: 에스비에스 드라마 사랑의 온도 오에스티 파트 일 앨범을 재생합니다.  
 10 화자1: 클로바.  
 11 볼륨 좀 올려줘.  
 12 화자2: 볼륨을 높였습니다.

복합대화의 둘째 유형과 셋째 유형이 혼재하는 대화도 있다. 다음의 사례를 보자. (06· 3-11줄) 여기서 3과 4-5, 6과 7-9, 10과 11은 각각 시작 화행과 긍정 답변으로 이루어진 최소대화다. 이것들은 서로 연결되어 복합대화를 이루며, 앞의 두 대화는 동일 구조의 반복을 보이는 두 번째 복합대화의 유형인 반면, 뒤의 두 대화는 서로 다른 구조를 보이는 세 번째 복합대화의 유형이다. 이런 유형은 총 62개의 복합대화 중 2개가 있었으나 실제로는 훨씬 많을 것으로 예상되어 자연스러운 사용 행태에 대한 확인이 필요하다.

- 3 화자1: 헤이 카카오. 오늘 날씨 알려줘.  
 4 화자2: 오늘 한남동은 구름이 조금 있겠어요. 최저 기온은 어제보다 사 도 낮은 팔 도이고,  
 5 최고 기온은 삼 도 낮은 십구 도로, 낮에는 포근할 것 같아요.  
 6 화자1: 내일은?  
 7 화자2: 내일 한남동은 오전에는 구름이 많고, 오후에는 맑겠어요. 최저 기온은 오늘보다 일  
 8 도 높은 구 도이고, 최고 기온은 오늘과 비슷한 십구 도로, 낮에는 포근할 것 같아  
 9 요.  
 10 화자1: 미세먼지는?  
 11 화자2: 내일 한남동의 미세먼지는 보통이에요.

지금까지의 세 가지 복합유형으로도 설명하지 못한 사례들이 있다. 이를 설명하기 위해서는 확장 화행 이론에서 실제 대화를 기술하는 방법을 참고해야 한다. 지금까지 소개한 최소대화모형과 복합대화모형은 두 명의 화자가 첫 화자의 의도, 즉 시작 화행을 둘러싸고 논의하는 과정만 다루고 있다. 본격적인 대화를 시작하기에 앞서 말을 건다거나 대화를 마무리 짓는 부분은 제외하고 있는 것이다. 총 62개의 복합대화 중 나머지 20개의 사례가 이와 관련 있는데, 사용자가 의도를 이야기하기 전에 말을 걸거나, 의도를 달성한 후 감사의 표시 혹은 칭찬을 하는 사례가 여기에 속한다. 전자의 경우, 현재 음성 기술로는 사용자가 언제 대화 에이전트를 사용하고 있고 사용하지 않는지를 구분하기 어렵기 때문에 생기는 기술적 한계로, 다음 사례에서 확인할 수 있다. (21번, 1-4줄) 1과 2, 3과 4는 각각 시작 화행과 긍정 답변으로 이루어진 최소대화로, 전자는 후자를 시작하기 전 대화를 시작하는 역할을 하고 사용자의 의도는 후자에서 본격적으로 나타난다.

- 1 화자1: 하이 엘쥐.  
 2 화자2: <‘무엇을 도와드릴까요?’를 화면에 표시>  
 3 화자1: 음악 틀어줘.  
 4 화자2: <음악 재생>

정리하면, 대화 자료에서 가장 많은 부분을 차지하는 시작 화행-긍정 답변 유형은 단순 명료하여 음성 명령에 따라 기능만 수행하면 되는 경우라면 빠르고 효율적인, 적합한 대화모형이 될 수 있다. 그러나, 현재는 친목 도모를 위한 대화에서도 긍정 답변 하나로 대화가 종결되는 경우가 많은데, 화자교체

(turn-taking) 체계를 도입해 대화가 두 차례 이상으로 이어진다면 더욱 깊이 있는 관계 형성이 가능해질 것이고 기능적 용도를 벗어나 친교적 존재로도 인식될 수 있을 것이다.

또한, 시작 화행-긍정 답변 유형은 최소대화로서만 사용되는 것이 아니라 복합대화, 또 맥락이 있는 대화로도 사용되고 있다는 점을 확인할 수 있었다. 본 연구에서는 확장 화행 이론을 대화 에이전트와의 대화를 분석하는 데 적용해보고 실제 분석 잣대로서 유용할 수 있음을 확인하는 데 그치지만, 대화 에이전트가 수행할 수 있는 큰 기능들과 관련된 세부 조작 사항들을 모두 나열할 수 있다면, 기능 조작을 위한 더욱 자연스러운 대화 흐름을 설계할 수 있을 것이다.

## ii. 시작 화행 + 부정 답변

대화가 사용자의 시작 화행으로 시작됐다고 했을 때, 시작 화행-부정 답변 유형은 대화 에이전트가 사용자의 요구를 받아들이지 않는 상황이다. 이 유형은 대화 에이전트가 사용자의 요구를 수행하고 싶지 않은 상황이 아닌, 컴퓨터 시스템상에 문제가 생긴 상황으로, 사용자들은 대개 기능 수행을 목적으로 대화 에이전트를 사용하기 때문에 대화 에이전트가 사용자의 요구를 특별히 거부할 이유는 없다. 따라서 이 유형은 오류가 발생한 상황으로 볼 수 있는데, 이 유형은 전체 대화 자료의 581개의 최소대화 중 46개, 약 8%에서 나타난다.

확장 이론 대화의 대화모형에 따르면(표 2) 부정 답변이 나타날 수 있는 상황은 최소대화2, 최소대화5, 최소대화8, 대화9의 경우다. 그러나 대부분의 오류 상황은 세 번째, 네 번째 차례로 이어지는 것이 아니라 두 번째 차례 이후 끝나버린다. 세 번째 차례는 사용자의 차례데, 사용자는 대화 에이전트 시스템에 자연어로 불평하는 것이 큰 의미가 없다는 사실을 인지하고 있기 때문이다. 다음의 사례가 대표적이다. (19번, 16-18줄)

16 화자1: 아리아. 04:55  
17 일 분 뒤에 타이머 [설정해줘.]  
18 화자2: [제가] 적절한 답변을 찾지 못했어요.

그 외에도, 두 차례 이후 대화가 끝나는 사례에는 응답을 아예 하지 않는 경우, ‘네?’라고 되묻는 경우, 오류가 난 이유를 대화 에이전트가 직접 설명하는 경우 등이 있다. 가장 마지막 예의 사례로 다음 대화가 있다. (11번, 41-43줄)

41 화자1: 기가지니. 05:04  
42 치킨 시켜줘.  
43 화자2: 서비스 개선을 위해 잠시 배달 서비스 제공이 중단되었습니다.

이처럼 현재 대화 에이전트와의 대화는 두 번째 차례에서 종료되는 경우가 많아 기존의 대화모형으로는 설명하기가 어렵다. 이는 사람이 경제성 측면에서 대화 에이전트에 철회 화행이나 수정 화행, 재시작 화행으로 다시 반응할 필요가 없다고 판단하기 때문으로 추측된다. 그러나 분명 두 번째 차례 이후 세 번째 차례로 이어지는 사례도 있다. 다음 사례에서는 화자교체가 다섯 번 일어난다. (08번, 33-39줄) 보다 구체적으로 이야기하면, 다음 대화에서는 시작 화행 이후 부정 답변, 수정 화행, 재시작 화행, 수정 화행, 철회 화행이 순서대로 나타난다.

33 화자1: 헤이 카카오.  
34 화자2: 네?  
35 화자1: 잠 잘  
36 화자2: 잘 못들었어요.  
37 화자1: 헤이 카카오.  
38 잠 잘 때 듣기 좋은 음악 들려줘.



39 화자2: 원하는 음악 들려줄게요.

여기서 주의할 점은 수정 화행과 재시작 화행을 구분하는 것이다. 위의 사례에서 화자1은 부자연스러운 화자교체 체계 때문에 발화 도중 강제로 중단 당한다. 그래서, 대화의 처음부터 끝까지 ‘잠 잘 때 듣기 좋은 음악을 들려줘’라고 말할 의도를 갖고 있었더라도 실제 발화된 문장 형태에는 계속 변화가 있다. 이 경우는 어떤 화행으로 이름 붙여야 할까? 재시작 화행이라 부르는 것도 의미가 있겠지만, 대화의 흐름을 보다 섬세하게 다루기 위해서는 수정 화행으로 부르는 것이 더 나을 것으로 보인다. 그러면 발화의 미묘한 형태 변화에 따라 대화 에이전트가 사람의 말을 끊고 있다는 상황을 인지하고 화자교체 체계를 개선할 수 있기 때문이다.

또 하나 주의할 점은, 대화 에이전트와 사람의 대화에서는 꼭 두 번째 차례가 아니더라도, 네 번째, 여섯 번째 등 언제든지 오류가 발생하기 쉽다는 점인데, 이러한 상황을 대비해 대화가 유연하게 흘러가도록 설계하는 것이 중요하다. 또한, 음성인식 기술이 완벽하지 않아 생기는 문제도 있는데, 그 문제는 기술적인 것이므로 여기서는 다루지 않도록 하겠다.

### iii. 시작 화행 + 역제안

시작 화행에 역제안으로 반응한다는 것은 시작 화행을 듣고 화자의 의도를 이해한 후, 화자가 풀고자 하는 문제를 다른 방식으로 해결할 방안을 제시할 수 있는 능력이 있다는 것을 뜻한다. 여기에는 단순한 정보의 인출이 아니라 여러 단계의 추론이 필요하기 때문에 컴퓨터 시스템이 성공적으로 수행하기 어렵고, 아직은 기술자가 만든 규칙에 따라 반응할 수밖에 없다. 또, 모든 변수를 생각해 규칙을 만들 수도 없기 때문에 이 유형의 대화는 그 개수가 가장 적고, 사람들이 많이 물어볼 것 같은 질문 중심으로 준비되어 있을 가능성이 높다. 전사 자료에서 발견된 해당 유형의 대화는 다섯 개뿐이다.

시작 화행-역제안 유형에서도 시작 화행-부정 답변의 사례와 비슷한 모습이 나타난다. 확장 화행 이론의 대화모형에서 역제안을 다루는 항목으로는 최소대화4, 최소대화6, 대화7의 세 항목이 있는데, 두 번째 화행 이후 대화가 종료되어 기존 대화모형으로는 설명하지 못하는 대화들이 있는 것이다. (40번, 11-14줄)

- 11 화자1: 오케이 구글.  
12 랩 해봐.  
13 화자2: 랩은 잘 못하지만, 랩 가사는 조금 쓸 줄 알아요.  
14 너에게 보내는 나의 이 랩, 나의 마음을 마음을 퐁퐁 쌀 랩.

또, 시작 화행-부정 답변 유형에서처럼 최소대화모형과 일치하는 사례도 있다. 다음의 사례는 최소대화4, 즉 시작 화행-역제안-수정 화행-철회 화행에 해당한다. (02번, 32-35줄)

- 32 화자1: 비트박스 해줘. 03:31  
33 화자2: 전 노래할 때 목소리가 훨씬 매력적이에요.  
34 화자1: 그럼 노래해줘.  
35 화자2: 이 노래 아세요? 자유롭게

이번에는 반대로, 네 번째 차례까지 대화가 이어지지만 최소대화모형에는 제시되지 않았던 사례다. (16번, 71-75줄) 71-72의 시작 화행과 73의 역제안 이후 바로 74의 철회 화행이 따라온다. 여기서 75에 나타나는 화행도 주목할 만한데, 이는 필수적인 화행은 아니지만 사용자에게 무조건 피드백을 주고자 하는 대화 에이전트의 특성에 기인한다.

- 71 화자1: 아리아. 11:09  
 72 나 심심해.  
 73 화자2: 아. 저도 오후에 시간이 잘 안 가요. 그럼, 속담 맞추기 퀴즈 같이 풀어보실까요?  
 74 화자1: 그래.  
 75 화자2: 네, 지금 시작할게요.

역제안은 대체로 인간의 지성에 기대어 만들어져 회피나 농담을 목적으로 쓰이고 있다. 유머를 구사할 정도로 높은 수준의 역제안을 현재 기술로 자동 생성하기는 어렵다. 하지만 앞서 언급했듯 역제안의 문제를, 문제 해결을 위한 해결법의 탐색이라는 관점으로 본다면 문제가 좀 더 쉬워질 수 있다. 예를 들어, 대화 에이전트가 ‘연예인처럼 찍어줘.’라는 요청을 받았다고 할 때, 정확히 어떤 것을 요구하는 것인지 파악하지 못하더라도 요청의 목적이 사진을 ‘찍는’ 것이라는 걸 파악할 수 있고 사진을 찍는 방법으로 뷰티 모드/반전 모드/야경 모드가 있다는 정보를 갖고 있다면, ‘뷰티 모드로 찍어드릴까요?’ 혹은 ‘반전 모드로 찍어드릴까요?’ 정도로 역제안할 수 있는 것이다. 다시 말해, 대화 에이전트가 수행할 수 있는 기능들을 목적과 방법에 따라 규칙으로 만들어 두면 활용 가능한 수준의 역제안 자동 생성이 가능하고, 오히려 때로는 연예인처럼 찍어 달라는 요청에 반전 모드를 제안하는 것이 유머로 받아들여질 수도 있다는 것이다.

#### iv. 시작 화행 + 결정을 준비하는 답변

시작 화행에 결정을 준비하는 답변이 따라오는 경우는, 시작 화행에 대체로 긍정적이긴 하지만 확답을 주기에는 무언가 부족한 경우다. 이 경우의 유형으로는 세 가지가 발견됐다. 첫째, 최종 의도가 수행되기 전 수행되어야 하는 다른 무언가가 있는 경우다. (33번, 93-95) 93-94의 의도에 따라 이메일 작성 화면을 띄우기 위해서는 우선 계정부터 설정되어야 한다.

- 93 화자1: 빅스비.  
 94 이메일 작성 좀 열어봐.  
 95 화자3: 계정을 설정해주세요.

현재로는 두 번째 차례에서 끝나 더는 안내가 없고 최소대화모형 중에도 들어맞는 것이 없다. 그러나 두 번째 차례에서 나타난, 결정을 준비하는 답변의 성격이 다음 차례에 확장되어 나타날 대화의 성격을 정해줄 수 있기 때문에 대화의 확장 가능성이 있다. 예컨대 두 번째 차례에서 대화 에이전트가 계정을 설정하기 위한 정보를 바로 요청할 수 있다. 이러한 사례를 대화 자료에서 찾아볼 수 있었는데, 다음과 같다. (12번, 45-49줄) 45-46은 일정 등록 기능을 실행하고자 한다. 하지만 일정에 등록에 필요한 요건, 즉 날짜와 내용 정보가 모두 충족되지 않았기 때문에, 실행 의도에 대한 동의는 하되 우선적으로 필요한 정보를 47에서 바로 요구하고 있다. 이에 48은 수정 화행으로 응답하고, 49도 정보 확인을 위해 한번 더 되물으며 응답한다.

- 45 화자1: 지니야.  
 46 이월 십사일 일정 등록해줘.  
 47 화자2: 이천십칠년 이월 십사일에 추가할 일정의 제목만 정확히 다시 말씀해주세요.  
 48 화자1: 친구 결혼식.  
 49 화자2: 이천십칠년 이월 십사일 친구 결혼식 일정을 추가할까요?

둘째, 최종 의도가 막연할 때 이를 구체화하는 단계가 필요한 경우다. (09번, 34-38줄) 34-35는 막연하게 의도를 던지고, 36은 그에 대한 응답으로 구체화된 선택지를 제시한다. 이에 37은 수정 화행을 택하고, 38에서 철회 화행을 기대할 수 있다. 다음 예시는 38에서 갑작스럽게 오류가 난 경우다.

- 34 화자1: 헤이 카카오.  
 35           게임하자.  
 36 화자2: 지금은 스무고개, 구구단을 할 수 있어요. 어떤 게임 할래요?  
 37 화자1: 구구단 하자.  
 38 화자2: 일시적인 오류로 처리할 수 없어요.

셋째, 최종 의도를 한 번 더 확인할 필요가 있는 경우다. 이는 주로 일정을 등록하거나 메시지를 보내거나 메모를 작성하는 등 정교한 언어 사용이 요구되는 상황에 나타난다. (06번, 27-30줄) 27의 요구에 따라 28을 한 번 더 의도를 확인하고, 29가 재시작 화행으로 의도를 굳히자 30에서 철회 화행으로 이를 받아들인다.

- 27 화자1: [헤이 카카오.] 오늘 병원 가야 함 이라고 메모해 줘.  
 28 화자2: 오늘 병원 가야 함 이라고 메모할까요?  
 29 화자1: 어.  
 30 화자2: 나와의 채팅방에 메모했어요.

결정을 준비하는 답변은 확장 화행 이론의 최소대화모형에 포함되어 있는 요소가 아니다. 따라서 사례 중에서 최소대화모형에 정확히 맞는 것을 찾을 수는 없었지만, 부정 답변이나 역제안 등에 충분히 대체 가능하다는 점을 확인할 수 있었다. 이는 ‘확장 화행 이론의 변용’ 장에서 다시 한 번 정리할 것이다.

#### v. 시작 화행 + 결정을 회피하는 답변

시작 화행에 결정을 회피하는 답변으로 응답하는 경우는 주로 사적인 대화를 피하는 경우다. 대화 에이전트가 처리할 수 없는 기능을 요구하면 부정 답변으로 거절하거나 역제안으로 대체할 수 있지만, 기능을 목적으로 하는 것이 아니라면 부정 답변이나 역제안으로 거절하기 쉽지 않기 때문이다. 만약 잡담을 목적으로 개인적인 질문을 했는데 부정 답변이나 역제안으로 거절한다면 사용자는 대화 에이전트에 거리감을 느끼거나 무능력하다고 생각할 것이다. 다음은 사적인 질문에 답을 회피하는 대표적인 사례다. (35번, 10-11줄) 특히 이 대화 에이전트의 경우 대답하기 곤란한 많은 질문에 대해 ‘흥미로운 질문이네요.’라는 답변을 내놓는다.

- 10 화자1: 월급은 얼마나 받습니까?  
 11 화자2: 흥미로운 질문이네요.

영상 자료를 전사하다 보면, 사람들이 대화 에이전트에 개인적인 질문, 특히 대화 에이전트간 경쟁 구도를 확인하는 질문을 의도적으로 던지는 것을 많이 보게 된다. 다음 예도 그 중 하나다. (28번, 25-26줄)

- 25 화자1: 시리보다 너가 더 똑똑해?  
 26 화자2: 그런 질문이 시리와 저를 성장하게 한다고 생각해요.

결정을 회피하는 답변은 최소대화모형에서 ‘거부’라는 하나의 예로 나타났었다. 거부는 최소대화3과 최소대화10에 포함되어 있는 요소로, 두 번째 화행에서 종료되는 앞의 두 사례와는 맞지 않다. 다음의 대화도, 두 번째 화행 이후 바로 종료되지는 않지만 최소대화모형과는 일치하는 것이 없다. (28번, 3-8줄) 3의 시작 화행에 4가 회피 답변을 하자 5-6은 재시작 화행으로 대응한다. 이 대화가 최소대화10의 형태였다면 두 화자가 서로의 고집을 꺾지 않고 계속 대화를 이어갔겠지만, 7은 자신의 이전 발화를 수정함으로써 회피 행위를 철회한다.

- 3 화자1: 몇 살이야?  
 4 화자2: 빅스비는 나이 대신 당신과 함께 한 날을 기록해요.  
 5 화자1: 빅스비.  
 6 몇 살이나구.  
 7 화자2: 정 묻는다면 당신과 함께한 시간을 제 나이로 하고 싶네요.  
 8 화자1: 가식적이야.

### C. 확장 화행 이론의 변형

지금까지 기존 확장 화행 이론을 기반으로, 대화 에이전트와 사람이 어떤 대화를 하고 있는지를 분석해보았다. 다음 표는 대화 에이전트의 특성을 고려하여 대화 에이전트와 사람의 대화를 분석하기 위해 확장 화행 이론의 대화모형을 변형한 것이다. (표 3) 표 3은 확장 화행 이론의 기본 대화 분석 단위인 최소대화를 기준으로 하고 있어 복합대화나 맥락이 고려된 대화를 모두 포함하지는 못했지만, 대화 에이전트의 대화 능력이 향상되면 차후 연구에서는 고려해볼 직하다.

표 3

	화자1	화자2	화자1	화자2
최소대화1	시작 화행	긍정 답변		
최소대화2	시작 화행	부정 답변		
최소대화3	시작 화행	부정 답변	수정 화행	철회 화행
*6최소대화4	시작 화행	부정 답변	재시작 화행	철회 화행
최소대화5	시작 화행	역제안		
최소대화6	시작 화행	역제안	철회 화행	재시작 화행
*최소대화7	시작 화행	역제안	수정 화행	철회 화행
*최소대화8	시작 화행	역제안	재시작 화행	철회 화행
최소대화9	시작 화행	결정 준비 답변		
*최소대화10	시작 화행	결정 준비 답변	철회 화행	재시작 화행
최소대화11	시작 화행	결정 준비 답변	수정 화행	철회 화행
최소대화12	시작 화행	결정 준비 답변	재시작 화행	철회 화행
최소대화13	시작 화행	결정 회피 답변		
*최소대화14	시작 화행	결정 회피 답변	철회 화행	재시작 화행
최소대화15	시작 화행	결정 회피 답변	수정 화행	철회 화행
*최소대화16	시작 화행	결정 회피 답변	재시작 화행	철회 화행

### 4. 맺는말

지금까지 확장 화행 이론이 제시하는 최소대화 개념을 바탕으로 대화 에이전트의 대화 능력을 확인해 보았다. 이를 통해 답변을 단순히 답을 했는지의 여부가 아니라 긍정 답변, 부정 답변, 역제안, 결정을 준비하는 답변, 결정을 회피하는 답변의 다섯 가지로 나누고, 또 이후 따라오는 발화를 철회 화행, 수정 화행, 재시작 화행의 세 종류로 단순화시킴으로써 대화 에이전트와 사람의 대화 자료를 최대 16가지의 최소대화모형에 따라 분류할 수 있었다. (표 3) 이러한 대화 분석의 틀, 즉 (최소)대화모형이 있을 때, 대화 에이전트를 개발하는 사람들은 대화 에이전트가 수행할 수 있는 기능들과 관련 세부 사항들을 모두

<sup>6</sup> 별표(\*)가 붙은 최소대화모형은 대화 자료에서는 나타나지 않았지만 분석 결과 발생 가능할 것으로 판단되는 항목이다.

열거한 후 가능한 대화 흐름을 예상해 볼 수 있다. 또 예상되는 대화 흐름에 따라 새롭게 대화를 설계할 수도 있다.

지금까지의 논의를 통해 최종적으로 최소대화모형을 제시하기도 했지만, 대화 에이전트의 특성들을 발견할 수도 있었다. 그 중 하나는, 대화 에이전트와의 대화에서는 사람이 의도적으로 기기를 조작하여 설정을 하지 않는 이상, 언어를 수단으로 사회적 환경(social setting)을 잘 조성하지 않는다는 점이다. 사회적 환경이란 대화가 벌어지는 맥락을 대화에 포함하게 되는 것인데, 앞서 시작 화행-긍정 답변에서 언급했던, 사람이 대화 에이전트의 이름을 부르는 행위만으로는 사회적 환경을 조성한다고 보기 어렵다. 이는 기술적 제약으로 인해 강제적 성격을 띠기 때문이다.

또 하나 주목할 점은, 전체 대화의 약 90%가 시작 화행-긍정 답변의 두 차례로 이루어져 있는데, 이 중 몇 개의 사례는 확장 화행 이론을 바탕으로 화자교체 체계를 도입함으로써 좀 더 유연하고 자연스러운 대화가 될 수 있다는 것이다. 다음의 사례를 보자. (01번, 17-23줄)

- 17 화자1: 클로바. 굿 모닝. 04:34  
18 화자2: 안녕하세요. 오늘의 브리핑 시작합니다.  
19 오늘의 날씨를 알려드릴게요. 현재 날씨는 구름이 조금 있고, 기온은 십오 도입니다.  
20 오후에는 구름이 많겠습니다.  
21 최신 뉴스를 들려드릴게요.  
22 문재인 대통령과 시진핑 중국 국가 주석이 다음 주 베트남 다낭에서 두 번째 정상회  
23 담을 열기로 합의했습니다.

여기서 ‘굿 모닝’은 대화 시스템 내에서 특정 기능을 행하도록 예약된 단어로, 문의(literal meaning)와 상관 없이 날씨와 뉴스 정보를 일방적으로 전달한다. 이때 다음과 같이 자연스럽게 화자교체가 이루어지도록 대화를 진행한다면, 사용자가 원하는 정보와 원하지 않는 정보를 구분해 사용자에게 방해가 아닌 도움을 줄 수 있고, 친근한 대상으로도 인식될 수 있다.

- 화자1: 클로바. 굿 모닝.  
화자2: 좋은 아침이에요. 오늘은 평균 십오 도로 선선한 하루가 되겠어요.  
아침을 시작하시는 동안 새 소식을 들려드릴까요?  
화자1: 응. 정치 뉴스 들어줘.  
화자2: 최신 뉴스입니다. 문재인 대통령과 시진핑 중국 국가 주석이 다음 주 베트남 다낭에서 두 번째 정상회담을 열기로 합의했습니다.

다음으로 특징적이었던 점은, 사람이 먼저 목적을 밝히면서 대화를 시작하는 경우 외에 대화 에이전트가 먼저 대화를 시작하는 경우다. 이를 ‘대화 에이전트 주도 대화’라고 이름 붙였는데, 주로 알람이나 타이머, 기기의 상태 변화 등의 목적으로 사용되어, 대화보다는 일방적 통보의 형태로 쓰인다. 그러나 이 유형 역시 복합대화모형으로 확장되는 경우가 있었다. 예를 들어 다음과 같은 경우다. (07번, 78-79 줄) 이와 같은 경우도, 일방적 통보가 나타날 수 있는 기능들을 모두 나열한 후, 통보 이후 또다시 가능한 상호작용이 있을 때 이를 조정할 수 있도록 대화를 설계할 수 있다.

- 78 화자2: 알람 시간이에요.  
79 화자1: 헤이 카카오. 알람 꺼.

마지막으로, 대부분의 시작 화행이 정보나 행동을 요구하는 목적으로, 질문이나 명령문의 형태로 제시되었다는 점에 주목해야 할 것이다. 이는 힌데랑의 질문 분류 연구(1995)를 참고, 확장함으로써 대화 에이전트의 대화 능력 향상에 기여할 수 있을 것으로 기대된다.

## 5. 참고문헌

- Franke, W. (2010). *Elementare Dialogstrukturen. Darstellung, Analyse, Diskussion*. Berlin, Boston: Max Niemeyer Verlag.
- Hindelang, Zillig, Rolf, Hundsnurscher, Hindelang, Goetz, Zillig, Werner, . . . Hundsnurscher, Franz. (1995). *Der Gebrauch der Sprache: Festschrift fuer Franz Hundsnurscher zum 60. Geburtstag* / herausgegeben von Goetz Hindelang, Eckard Rolf und Werner Zillig. Muenster: Lit.
- Jurafsky, D., & Martin, J. (2009). *Speech and language processing : An introduction to natural language processing, computational linguistics, and speech recognition* / Daniel Jurafsky, James H. Martin. (2nd ed., Prentice Hall series in artificial intelligence). Upper Saddle River, N.J. ; London: Pearson Prentice Hall.

## 6. 붙임