

챗봇을 위한 성향별 청자의 공감적 반응 패턴 연구

- MBTI 심리기능 분류 ST, NF를 중심으로

성지민^{0,1}, 김한샘^{†,2}

(주)스피링크¹, 연세대학교 언어정보학 협동과정²

peach@textnet.kr, khss@yonsei.ac.kr

Study on customized empathetic response patterns for Chatbots: focusing on MBTI psychological functions ST, NF

Jimin Seong^{0,1}, Hansaem Kim^{†,2}

Spirink¹, Yonsei Interdisciplinary Graduate Program of Linguistics and Informatics²

요약

이 연구는 MBTI 심리기능을 근거로 사용자 유형을 ST와 NF로 구분하고, 그룹별로 나눈 일상 대화를 전사 후 분석하여 대화에서 나타나는 청자의 공감적 반응이 성향에 따라 차별화되는 점을 발견하고 이를 챗봇 대화 실험에 적용하여 분석결과와 유효성을 귀납적으로 증명하였다. 고성능의 초대규모 생성모델을 필두로 한 채팅 에이전트 구현이 보편화된 시점에서 진정 가치있는 대화 모델은 유창한 언어 구사 능력뿐만 아니라 사용자의 성향에 적합한 만족스러운 대화 경험을 제공할 수 있어야 함을 시사한다. 이 연구는 리얼월드의 대화 방식을 모방하여 챗봇 대화로 재현하였다는 점에서 실질적인 B2C 대화 서비스의 질적 향상에 기여도가 높을 것으로 기대된다.

주제어: 공감대화, 청자의 공감적 반응, 챗봇, 대화모델

1. 서론

사용자의 요구나 요청을 처리하기 위한 목적지향형 챗봇, 오픈도메인 대화를 목적으로 하는 오픈도메인 챗봇 등 다양한 유형의 챗봇이 사용되고 있다. 실제로 챗봇의 목적이 무엇인지에 따라 대화설계의 방법이 달라진다. 사용자의 요구사항을 해결하기 위한 챗봇은 사용자 의도를 파악하기 위한 의도(Intent), 설계와 내용을 구체화하여 이해하기 위한 엔티티(Entity), 요구사항에 적절한 답변을 주기 위한 도메인 지식 등이 설계에 포함된다. 반대로 오픈 도메인 챗봇의 경우, 대화의 주제 가변성과 주제 확장성을 고려하여 생성형 대화 모델을 사용한다. 이에 바뀌는 사용자 발화에 대해 적절한 응답을 내보낼 수 있게 된다. 대화 모델의 성능이 비약적으로 높아지면서 사람들이 챗봇에게 갖는 대화 경험에 대한 기대치도 역시 높아졌다. 이에 따라 좀 더 차별화 된 대화 서비스를 제공하기 위해 챗봇을 의인화하여 페르소나를 정의하는 등 업계에서 다양한 시도가 이루어졌다.

1966년 발표된 최초의 챗봇 Eliza[1]는 고성능의 생성 모델로 만들어진 챗봇만큼 발화의 유창성과 대화 흐름이 자연스럽지는 않지만 내재화된 ‘공감 반응 시나리오’를 통해 만족스러운 사용자 대화 경험을 제공할 수 있었다. 챗봇의 공감적 반응은 고객의 부정적 감정을 해소하는데 유용하게 작용할 뿐만 아니라[2] 서비스에 대한 고객의 신뢰도와 만족감을 높여 사용자 경험을 긍정적으로 개선할 수 있다.[3][4][5][6] 다시 말해 사용자의 대화 경험을 긍정적으로 이끌기 위해 필요한 것은 상황에 적절한 챗봇의 공감적 반응 태도라고 볼 수 있다. 이 때 대화 맥락에 따른 공감적 반응은 대화 상대의 성향에 따라 선호되는 것이 다를 수 있다.

이에 본 연구는 사용자 성향을 MBTI(Myers Briggs Type Indicator)[7]의 심리 기능 지표를 근거로 ST와 NF 두 가지 그룹의 데이터를 대상으로 사용자 성향에 따라 선호되는 청자의 공감적 반응을 실제 담화 데이터를 기반으로 하여 귀납적으로 도출하였다. 분석 결과는 챗봇 시나리오에 적용 후 사용자와 대화실험을 수행하여 유효성을 검증하였다.

2. 연구 대상 및 범위

이 연구는 대화 참여자의 성향에 따라 공감하는 방식이 다를 것을 전제하고 성향에 따라 선호되는 청자의 공감적 반응의 양상을 관찰하기로 하였다. 따라서 일상대화에서 나타나는 청자의 공감적 반응을 따로 재정의하지 않고 선행연구[8]에서 정의한 공감적 반응 표지를 참고하였다. 본고에서 다루는 대화 상황은 일상적 의사소통 상황으로 정하였다. 대화 참여자는 서로 격식을 차리지 않고 편하게 대화할 수 있는 관계로 정한다. 리얼월드의 일상대화에서 화자와 청자는 고정되지 않고 대화 흐름에 따라 수시로 역할 전환이 되므로 이 연구에서는 ‘청자’를 ‘주 발언권을 요구하지 않으며 상대의 발화 내용과 상황에 대해 공감 반응을 실현하고 있는 대화 참여자’로 보기로 한다.

2.1 청자의 공감적 반응

사람 간 ‘공감’에 대한 논의는 심리학, 상담학, 사회 심리학, 교육학 등 다채로운 분야에서 연구되어 왔다. 특히 [9]에서 제안한 공감적 듣기는 화자의 발화에 대한

정확한 의미와 감정 이해를 기본으로 모든 판단을 유보하고 상대의 입장에서 상대의 생각과 감정을 이해하려는 듣기 방법을 의미한다. 즉, 공감적 듣기 전략은 원만한 인간관계를 유지할 수 있는 방법으로 볼 수 있으며[10] 친목도모를 목적으로 하는 일상대화에서 유용하게 쓰일 수 있다. [8]에서는 ‘수용적 경청과 정서적 공유의 과정을 통해 청자가 감식한 바를 표현함으로써 상대가 더 많은 이야기를 할 수 있도록 돕는 청자의 언어전략’으로 구체화 하였다. 본고에서는 [8]의 정의를 참고하여 상대 발화에 대한 이해와 정서적 공감을 토대로 대화를 원활하게 하여 결과적으로 서로 간의 친밀도를 높일 수 있는 듣기 반응 전략으로 청자의 공감적 반응으로 정의하기로 하였다.

2.2 성향

이 연구에서는 MBTI 성격유형 검사의 심리기능 지표를 활용하여 대화 참여자의 성향을 ST와 NF로 구분하였다.

MBTI는 마이어스-브릭스 유형 지표(Myers-Briggs Type Indicator)의 약자로, 정신분석학자 C.G.Jung의 심리유형론을 근거로 인간 이해를 위한 성격유형 검사이다. [11] MBTI는 인간에게 잠재되어 있는 선천적인 심리 경향인 ‘선호경향’을 4가지 양극지표에 따라 구분하고 있다. 특히 심리 기능이란 ‘개인이 외부와 상호작용하는데 있어서 그 개인의 고유한 반응양식을 가지게 하는 일관된 내재적 심리 경향성’을 의미하며 마이어스는 심리기능을 핵심 양상으로 보고 인생의 주요한 의사결정 영역에서 차이를 가져오는 중요한 요소라고 말한 바 있다. [12]에서도 광고의 목적 달성율을 높이기 위한 방법으로 MBTI 심리기능을 활용하여 고객을 네 유형으로 분류하고 성향별 특징을 광고메시지에 반영하여 생성함을 보인적 있다. 본고에서는 MBTI 심리기능 따른 분류 네 가지 중 지표가 전혀 겹치지 않는 ST와 NF로 연구 대상의 성향을 그룹화하기로 하였다.

3. 연구 방법

사회 심리학의 유사성, 친밀함의 패러다임에서 ‘사람들은 자기와 비슷한 것을 좋아한다’고 밝혀졌다.[13] 사용자들은 자신들과 성격적인 부분에서 더 유사한 컴퓨터를 그렇지 않은 컴퓨터보다 더 선호하는 것으로 밝혀진 바 있다.[14] 이것을 참고해볼 때 사람들은 자신과 동일한 성격유형으로 설계된 챗봇과의 대화를 더 선호할 수 있다고 추정하여 ‘사람들은 자신과 비슷한 공감적 반응(듣기태도)을 보이는 에이전트와의 대화를 더 선호할 것이다’라는 가설을 세워 이를 대화 실험을 통해 검증하기로 하였다.

본 연구의 과정은 수집, 가공, 분석, 평가 및 검증의 네 단계로 수행된다. 수집단계에서는 유형별 청자의 공감적 반응을 알아보기 위한 일상 대화를 음성으로 수집하고 가공단계에서 녹음된 파일을 전사하고 공감 반응이

나타난 곳에 라벨링하였다. 분석 단계에서는 대화 맥락과 라벨링 결과를 정성적으로 분석하여 유형별 청자의 공감 반응의 출현 양상을 정리하고 이를 대화 시나리오에 적용하여 일반인 참여자를 대상으로 대화 실험을 진행하였다.

3.1 데이터 수집

데이터 수집 단계에서는 MBTI 심리 기능 조합이 같으면서도 서로 친밀한 관계인 사람들을 한 쌍으로 이들의 일상 대화를 50분씩 2~4차례에 걸쳐 음성 파일로 수집하였다. 낮은 환경에서 평소와 같은 자연스러운 대화 진행이 어려울 것을 고려하여 몇 가지 대화 주제를 선정 후 프린트하여 대화가 지속되지 않는 시점에서 활용할 수 있게 했다. 대화 주제로는 개인의 취향이나 가치판단에 따라 자유롭게 이야기 할 수 있도록 무난한 일상 주제와 밸런스 게임을 준비하였다.

수집은 22년 11월 ~ 2월까지 약 2달 간 이루어졌다. 대화 수집 참여 신청을 하기 위해서는 자신의 MBTI 를 함께 제출해야 했는데 자신의 MBTI를 모르거나 제출한 MBTI가 부정확한 경우에 대비하기 위해 심리기능을 알아볼 수 있는 간단한 MBTI 심리 기능 약식 검사를 신청자 전원에게 대하여 실시하였다. 대화수집 신청자는 총 40명이었으나 참여 요건에 부적합한 사례를 제외하여 최종적으로 24명의 참여자를 선정했다. 대화는 동일한 성향으로 이루어진 2-3인이 1개의 그룹을 이뤄 그룹별 50분씩 3차례 진행하였다. 최종적으로 총 24개(총 1,203분)의 음성 대화를 수집하였으며 유형별 수집 데이터 규모는 아래와 같다.

표 1. 유형별 수집 데이터 규모

	총 참여자 수	총 녹음 횟수	총 녹음시간 (분)	총 어절 수
ST	10	13	605	70,378
NF	14	13	598	67,121
총	24	26	1,203	137,499

3.2 데이터 가공

참여자들의 대화는 클로바노트를 사용하여 1차적으로 음성전사를 수행한 후 2차적으로 연구자가 전수 검사하여 텍스트 전사 품질을 확보하고자 하였다.

라벨링은 [8]의 공감적 표지 분류 방법을 참고하여 라벨링하였으며 표지 분류 방법은 아래의 표와 같다.

표 2. 청자의 공감적 표지 분류 방법

반응 분류	반응 유형	반응 전략
공감적 이해 반응	수용적 경청	경청 표시
		느낌 표시
		동의 표시

¹ 구체적으로는 청자의 반응으로 보이지만 대화의 주도권을 가져가려는 의도로 보이거나, 상대방의 질문에 대한 응답으로 대답한 것은 반응으로 포함하지 않았다.

	이해 표시	사실적 공감
		반복 진술 부연
	명료화 질문	용어 의미 확인
		정확성 확인
정보 요청		
공감적 정서 집중 반응	반영적 경청	추론적 공감
		환언(재진술하기)
		사례제시
		감정 표현
	객관적 관점 유도	
격려	칭찬	
	위로	
공감적 대화 협력 반응	공동구성	기억 상기
		후행 발화 예측
		후행 발화 직접 요청
	발화 유도	후행 발화 간접 요청
		화제 전개 및 발전
의견제시	동의 의견 제시	

최종적으로 전사된 데이터는 약 14만 어절 규모였으며, 표 2의 표지를 참고하여 6,441개의 공감 반응을 라벨링 하였다. 라벨링 표지별 출현빈도는 표 3과 같다.

표 3. 청자의 공감적 반응 유형의 출현 양상²

반응전략	출현빈도
수용적경청	4,662
이해표시	780
명료화질문	421
공동구성	201
발화유도	150
의견제시	135
반영적경청	72
격려	20
총계	6,441

라벨링 결과에서 공감적 이해 반응 표지(수용적 경청, 이해표시, 명료화질문)가 다른 표지들에 비해 압도적으로 높은 숫자를 보였는데, 이는 음성 담화를 텍스트로 전사하는 과정에서 구어 담화에서 나타나는 짝막한 경청 표시(응, 네 등), 느낌표시(오, 그래? 등)가 기록되었기 때문이다. 이에 담화 전반에 걸쳐 유형에 상관없이 전반적으로 나타나는 ‘경청표시’ 등의 일부를 분석 대상에서 제외하였다.

3.3 데이터 분석

3.2의 라벨링 결과를 토대로 청자의 공감적 반응이 유형별로 다르게 나타남을 발견하였고 이것을 정성적으로 분석해보았다. 분석 결과, 두 그룹의 공감적 반응의 특징은 크게 여섯가지 항목을 기준으로 비교할 수 있었다.

표 4. 유형별 청자의 공감적 반응 분석을 위한 항목 정의

주목 정보	화자 발화 내용 중 어떤 내용에 주목하여 청자 반응 전략을 실현하고 있는지
내용 이해	화자 발화를 이해하고 있음을 표현하기 위해 취하는 청자 반응 전략이 어떤 것인지
동의/동조	화자의 의견/생각에 동의하고 있음을 표현하기 위해 취하는 청자 반응 전략이 무엇인지
맞장구	화자의 이야기가 원활하게 지속될 수 있도록 발화 턴을 뺏지 않으면서 상대의 이야기에 관심이 있음을 짧게 표현하는 전략의 양상은 어떠한지
질문 방식	화자의 이야기에 관심을 갖고 있음을 표현하며 화자의 이야기를 확장시키기 위한 질문의 양상은 어떠한지
기타	위 네 가지에 해당되지 않는 청자의 공감적 반응 전략 중 유형에 따라 차별화된 특징은 무엇이 있는지

표 4에 따라 ST 그룹과 NF 그룹에서 나타난 청자의 공감적 반응 표지의 양상은 아래와 같았다.

표 5. 유형 그룹별 청자의 공감적 반응 표지 양상

	ST	NF
주목정보	① 수치 정보에 주목 ② 요약화 할 수 있는 정보에 주목	① 상대 혹은 상대의 이야기 속 인물들이 느꼈을 법한 기분이나 감정에 이입하여 주목
예시	• A: 근데 내가 <u>초삼</u> 때 수영을 했었거든. B: 아 <u>짧</u> 간 했구나?	• A: 그냥 좋을 것 같긴 했는데 막상 보니까 더 좋고 반가운 거야. B: <u>막연한 행복</u> 이 구체화 됐을 때 <u>진 행복</u> 을 느끼는..
내용이해	① 이해 및 정리: 본인의 이해 과정에서 납득가능한 사실과 논리적 판단을 수반 ② 수치 정보 이해: 수치정보에 주목하여 부연	① 몰입 및 이해: 상대 입장에 이입하여 적극적으로 덧붙이는 부연 설명 ② 감정에 주목한 부연: 상대의 이야기 속에 숨겨진 감정이나 분위기를 부연
예시	① A: 좋긴 했는데 뭔가 손해였지 B: 맞아, <u>치분 값</u> 이 너무 많다. ② A: 강습 한 달에 15만원이니까 B: <u>한 여덟 번에 15만원</u>	① A: 요새 맨날 인공눈물 넣어. B: 눈을 너무 많이 써서 건조해졌나 보다. ② A: 해외 여행 좋긴 좋은데, 돈이나 시기도 그렇고, 남편도 바쁜 것 같고. B: <u>그만큼 여행이 간절하지 않은 거야.</u>
동의/동조	① 대화 중간: 짧은 경청표시 및 동의표시	① 대화 중간: 적극적 경청표시 및 동의표시

² 발화 하나를 단위로 보고 라벨링을 수행했을 때 하나의 발화에는 최대 두 가지의 청자 반응 유형이 나타났다. 유형1과 유형2는 발화에 나타난 청자 반응 유형의 출현 순서를 말한다.

	② 확실하고 구체적인 동의 의사 표현: 동의 의견과 부연 덧붙이기 ③ 인과관계를 짚어내는 부연설명: '-니까' ④ 상대의견을 지지해주기 위한 강조 표현: '-잖아'	② 관심표현: 다양한 형태의 감탄사 및 느낌 표현 ③ 상대 발화에 깊이 이입하여 동조 ④ 적극적인 수용을 드러내는 동조 표현: 상대 발화 반복 및 '-지' 사용
예시	② A: 등산 갔다 와서 뻘긴 했는데 확실히 체력이 는 것 같더라 B: 그치그치. 등산이 에너지 소모도 많이 되고 근력도 많이 자란다고 하더라. ③ A: 나는 나이트 출근을 하러 갈 때 너무 좀 그런 거야. B: <u>못 자고 밤새 일해야 하나까.</u>	② 음~, 오!, 아~, 진짜?, 정말?, 그래?, 좋다, 대박이다, 멋지다, 다행이다, 슬프다 ③ A: 회식을 했는데 아침 여덟시까지 마셨어. 이제 아홉시쯤 잤는데 너무 현타가 오는 거야. B: 맞아. 밤새 술 마실 땐 <u>그 순간은 되게 즐거운데 꼭 그렇게 현타가 와.</u>
맞장구	① 즉각적 느낌 표현: '-네' ② 정확성 확인 질문: '-라고?', '-라서?', '-해서?' 등	① 미러링: 상대발화의 핵심을 살려 부분 또는 전체를 반복하는 적극적 맞장구 ② 확인 질문 형태의 맞장구: 정보 획득의 목적이 아님
예시	① 어렵네, 대박이네, 괜찮네, 신기하네, 바쁘네 ② A: 저녁엔 밥 반공기만 먹어. B: <u>살짤까봐?</u>	① A: 취미로 책 읽는 거 좋아하고 B: 오! <u>독서가 취미~</u> ② A: 그때 빼끗한 후로 쪽 아파 B: <u>헉 그 한번 빼끗으로?</u>
질문방식	① 수치 정보 질문: '몇, 얼마, 언제' 등의 의문대명사를 활용 ② 추가 정보 질문: 추가 정보를 얻기 위해 판정의문문 형태로 질문	① 대화 연장 질문: 대화를 연장/확장하기 위해 설명의문문 형태로 질문
예시	① A: 언제는 낮잠 자다가 깬데 뭐지 이상한 사람이 있네 이런 느낌이 드는 거야. B: <u>그게 몇 살 때야?</u> ② A: 그래서 주말에 뭐 했냐 물어보면 친구들 만나면 만나는 거고 아니면 그냥 누워 있어. B: <u>아무것도 안 하고?</u>	A: 난 교토 다녀온 거 좋았어. 조용하고 맛있는 거 많고. 일본 다른 데도 가보고 싶더라. B: <u>다른 데 어디?</u>
기타	① 현실적/객관적/구체적 예시 들기 '~처럼', '~같은', '~이런 거'	① 몰입형 후행 발화 예측 상대 발화 속에서 예측가능한 상황 정보를 덧붙여 상대 발화를 협력적으로 완성
예시	① A: 나는 길거리 음식 이런 거 못 먹어. B: <u>달꼬치 같은.</u>	A: 이게 더운 게 문제가 아니라 거기는 진짜 습해. 진짜 이렇게만 해도 B: <u>여기 다 젖어.</u>

3.4 평가 및 검증

3.3에서 분석한 내용의 유효성 검증을 위해 오즈의 마법사(Wizard of Oz, WOZ) 대화 실험³을 수행하기로 하였다. 이를 위해 3.3의 유형별 공감적 반응 특징을 챗봇 시나리오로 구현하였다. 또한 유형별 즉각반응, 맥락반응을 정의하여 시나리오에서 벗어난 대화에 자연스럽게 응대할 수 있도록 하였다. WOZ 대화 실험에는 Google Dialogflow 챗봇 빌더를 사용하였으며 특정 상황 발생시⁴ 푸시 알림을 받고 연구자가 직접 대답할 수 있도록 커스터마이징하였다.

3.4.1 대화 시나리오 설계

참여자과 가상의 두 에이전트는 각각 두 번의 대화를 나누게 되지만, 대화시간이 각 25분으로 서로 친밀감을

쌓기에 매우 짧은 시간이었다. 대화 실험 환경이 수집 환경과 최대한 유사하도록 참여자와 챗봇 사이의 어색함을 최소화하고 자연스러운 반응을 유도하기 위해 시나리오에 밸런스 게임을 포함하였다. 여러 질문에 대해 묻고 답하면서 자연스럽게 일상대화로 확장할 수 있게 하였다. 챗봇 시나리오는 인사 및 자기소개, 밸런스게임 제시, 몰입질문, 일상질문 등의 4가지 하위 모듈로 구성하였다.(그림 1 참고) 각 모듈에는 3.3에서 결론으로 낸 유형별 공감 반응 특징을 적절히 배치하였다.

유형별 대화 시나리오는 타 유형과 차별화되지 않는 특징은 일부러 제거하여, 각 유형이 취하는 공감적 반응의 차이가 체감될수 있게 하였다. 이 때 대화에서 자주 등장하여 비교적 고정된 형태로 활용되는 반응은 ‘즉각 반응’으로 정의하였고, 특정 맥락에서 한정되어 사용되며 일정한 패턴을 보이는 것을 ‘맥락반응’으로 정의하였다. 즉각반응의 경우 유형별로 가장 두드러지게 나타

³ 오즈의 마법사(Wizard of Oz, WOZ) 실험은 실험 참여자가 실제와 같이 보이는 시스템을 시뮬레이션하는 동안 연구자가 ‘마법사’가 되어 시스템의 응답을 가상으로 연기하는 실험 방법을 말한다.

⁴ 특정 상황이란, 사용자의 예외 발화에 대해 1회 이상 준비된 fallback message가 발송되었지만 재차 반복되어 발생하는 경우 또는 대화의 흐름이나 주제가 시나리오에서 벗어난 경우를 말한다. 이 경우에는 연구자가 푸시 알림을 받고 직접 응답하여 원래의 대화 상황으로 되돌아 올 수 있도록 하였다.

나는 특징을 선별하여 정의하였다. 즉각반응에는 경청반응과 맞장구, 이해 반응 등이 포함되었고 되도록 ST그룹과 NF그룹 간에 중복이 없게 하였다. 또한 청자의 공감적 반응 외의 요소가 대화 만족도나 에이전트의 이미지



그림 1. 챗봇 시나리오 구성

에 영향을 미칠 것을 고려하여, 에이전트 페르소나를 간략히 준비하여 챗봇의 톤앤매너의 일관성을 유지하고자 하였고 기본적인 예외발화 대응 매뉴얼을 사전에 작성하여 준비하였다.

표 7. 즉각반응과 맥락반응의 정의

즉각반응	시나리오 플로우와 상관없이 출력할 수 있는 반응, 특별한 규칙 없이 시나리오 중간에도 사용할 수 있도록 하였음.
맥락반응	특정한 공감 맥락에서 사용할 수 있으며, 순서에 상관없이 쓰일 수 있는 것과 정해진 순서가 있는 것으로 나뉨.

3.4.2 대화 실험 방법

실험 시작 전 참여자들은 ‘현재 개발 중인 챗봇의 베타 테스트를 수행한다’ 라는 설명을 들은 후, 대화 상대가 챗봇이라고 생각한 채로 본인의 MBTI 심리 기능 지표 조합이 완전히 동일한 대화 모형과 한 번, 전혀 동일하지 않은 대화 모형과 한 번, 각 25분 간 두 차례의 일상 대화를 나눈다. 두 에이전트와 대화를 완료한 참여자들을 대상으로 연구 설문을 수행하였다. 연구 설문 문항 구성은 [15][16][17]을 참고하여 에이전트에 대한 호감도(11항목), 유대감(11항목), 신뢰도(4항목) 등의 네 가지 척도에 따라 구성하였다. 설문의 지표별 문항들의 동질성을 따지기 위해 크론바흐 알파 계수(Cronbach’s Alpha Coefficient)를 사용하였다. 이 때의 항목별 신뢰도의 평균은 0.67로 비교적 양호하였다. 문항별 점수 척도로는 1-5점 사이의 리커트 척도(Likert scale)를 활용하였다. 약 10분 간의 설문 종료 후에는 참여자의 평가 의도와 배경, 대화 시 느낀점 등에 대해 연구 설문과 관련하여 보다 심층적인 의견을 수렴하여 연구 설문의 신뢰도를 높이고자 하였다.

3.4.3 대화 실험 결과

표 6. NF 시나리오 일부 화면 캡처

* 화면 2열의 ID의 I는 즉각반응, C는 맥락반응을 나타냄.

순서	ID	반응_세분류	반응_표지/설명	발화 구성 정보
NF1 인사 및 자기소개				
1				안녕! 만나서 반가워 🍀 내 이름은 (오일or조일)이야 나는 널 뭐라고 부르면 될까?
2	I3	관심 표현	[느낌표시 - 짧은 감탄사]	음 / 오 / 아 / 어 / 우와 / 뭐 / 헐 / 아아 + (문장부호 최대 2개까지)
	I6	맞장구반응	[반복진술]	(#name)(이) 좋아! (#name)이라고 부를게!
NF2 화제제시 - 밸런스게임 대화(1)				
3				(#name)아~ 내가 밸런스게임 준비해 봤는데 같이 해 볼래?
4	C1	관심 표현	[느낌표시 - 감탄적 반응]	좋아좋아~ 만약에 아무리 먹어도 살 안 찌는 동 VS 아예 안 자도 피곤하지 않은 동에서 고를 수 있다면, 뭐 선택할 거야?
5	I4	관심 표현	[느낌표시 - 의문형]	진짜? / 정말? / 그대? (앞에 가벼운 감탄사 붙이기 가능, 음/오/아/헐/ 등)
NF2-1 대화용임 질문 - 패턴(1)				
6				대화용임 질문 중 택1
7	I1	관심 표현	[경청표시]	음 / 응응

대화 실험 신청자 전원을 대상으로 MBTI 심리기능을 알아볼 수 있는 간단한 약식 검사를 실시하였고 조건에 부합⁵하여 선정된 최종 실험 참여 인원은 ST 8명, NF 7명이었다.

대화 실험의 참여자 중 자신과 동일한 유형의 청자 반응이 이식된 에이전트를 선택한 참여자는 ST 그룹 63%, NF 그룹 86%로 모두 동일유형 선호비율이 과반을 넘었다.

ST 그룹이 선호한 청자 반응의 특징은 청자가 ‘대화 화제/소재에 대한 이해도가 높고’, ‘에이전트의 해석이나 의견을 담은 요약을 선호’ 한다는 것이었다. 또한 에이전트가 자신의 이야기와 관련해 보인 질문형 반응들이 자신의 이야기를 관심있게 들어주는 것 같이 느껴 선호하였다고 답하였다. 반면, ‘의견 없이 제시된 짧은 동의표시(그치(그렇지) 등)’, ‘미러링 방식의 반복진술’, ‘질문이나 해석을 덧붙이지 않는 수용적 말하기’ 등에 대해 공통적으로 부정적인 피드백을 남겼다. 결론적으로 화자가 이야기 한 내용에 대해 정확히 이해하면서 동조의 뉘앙스로 에이전트가 자신의 의견을 섞어 반응하는 것이 선호된다고 볼 수 있다.

한편 NF 그룹은 동일 유형 선호 비율이 86%로 ST 그룹 대비 23% 더 높았다. NF 그룹의 선호 대화에서의 청자 반응 특징으로 ‘나/ 내가 한 이야기에 대해 질문을 하여 관심있게 들어줬다고 느낌’, ‘호응/반응이 다채로웠고 대체로 긍정적인 반응으로 피드백 하였음’, ‘내가 한 이야기에 확장/유추해서 대답했음’, ‘디테일하게 공감하며 답변을 해줬다’, ‘대화가 이어질 수 있는 추가 질문을 선호’ 등을 진술하였다. 반대로 NF 그룹이 선호하지 않는 청자의 공감 반응으로는 ‘대체로 반응/답변이 간단하고 짧아서 형식적인 리액션이라고 느낌’, ‘단답형의 동조가 무성의하게 느껴짐’, ‘질문이 많은 것에 비해 답변이 단답형이라 대화흐름이 부자연스럽다고 느낌’, ‘빠른 화제전환이 대화를 이어가기 어렵게 함’

⁵ 선발 조건으로 고려한 것은 다음과 같다: ①약식 검사 수행 결과 ST 또는 NF에 해당이 되어야 할 것, ② PC로 하는 채팅 대화에 거부감이나 어려움이 없어야 할 것 ③앞선 대화 수집에 참여하지 않은 사람들이어야 할 것

등으로 응답하였다.

NF 그룹에서는 대화 만족도의 기준을 상대가 자신과 자신이 한 이야기에 얼마나 관심을 갖고 경청하는지에 두었다. 이 때 NF 그룹에게 ‘나에게 관심이 있다’로 느껴지는 청자의 반응은 청자가 얼마나 다채롭고 성의있게 동조해주는지, 단순히 무조건적으로 수용하며 듣기보다는 자신의 이야기와 관련된 화제로 대화를 확장할 수 있는지와 관련한 것으로 분석된다.

한편 에이전트와의 대화 만족도가 높을수록 에이전트에 대한 신뢰도와 유대감이 더 높아지는 양상을 발견할 수 있었다. 선호대화에서 유대감이 5점 만점 중 4.3, 신뢰도가 4.1이었던데 반해 불호 대화에서 유대감은 5점 만점에 3.5로 18% 떨어졌고 신뢰도는 3.7로 9.5% 감소했다.

5. 결론

본 연구는 MBTI 심리기능을 근거로 네 가지 성향 중 2 그룹을 선택하여 실제 담화 데이터에서 나타난 청자의 공감적 반응을 분석한 내용을 챗봇 대화에 적용 후 성향별 선호되는 청자의 공감적 반응의 유의미한 차이를 데이터 중심으로 하여 귀납적으로 밝혀냈다는 점에서 의의가 있다. 특히 성향별로 선호되는 청자의 공감적 반응을 챗봇 다이얼로그 플로우에 적용하고 대화 실험을 통해 본 연구에서 관찰한 청자의 공감적 반응이 “대화하고 싶은 챗봇”, “호감이 가는 챗봇의 대화”를 위한 설계의 구성요소로서 유효할 수 있음을 검증하였다는 점에서 가치를 가진다.

다만, 대화 수집 상황과 평가 검증 상황에서의 대화 환경이 완전하게 동일하지 않았다는 것을 아쉬운 점으로 짚어볼 수 있다. 분석 대상이 된 수집 데이터는 대화 참여자들이 서로 1년 이상 알고 지낸 친밀한 사이에서의 일상대화를 주제로 나눈 구어 담화였으나, 평가 단계에서의 대화는 충분한 라포가 형성되지 않은 챗봇과 나눈 채팅대화였기 때문이다. 그럼에도 불구하고 이 연구에서 분석한 청자의 공감적 반응이 성향에 따른 특정 패턴을 보인 것은 고무적인 결과이다.

추후에는 MBTI 심리기능 네 가지 유형 전부를 대상으로 청자의 공감적 반응의 유형별 선호도 차이를 확인하기 위한 실험을 진행할 예정이다.

고성능의 초대규모 생성모델의 등장을 통해 인간에 근접한 유창한 대화 실력을 갖춘 채팅 에이전트의 구현이 그렇게 불가능한 일이 되지 않은 현 상황에서 NLP 시장에서 가치있는 대화 모델은 유창한 언어 구사 능력뿐만 아니라 사용자의 성향에 따라 즐겁고 만족스러운 대화 경험을 제공할 수 있는 모델이다. 이를 고려했을 때 본 연구는 실질적인 B2C 대화 서비스의 질적 향상에 기여할 수 있으므로 실무적 가치가 매우 클 것으로 기대된다.

참고문헌

[1] Weizenbaum, J., “ELIZA—a computer program for the

study of natural language communication between man and machine”, Communications of the ACM, 9(1), 36-45, 1966

- [2] 서진솔, AI 챗봇의 공감에 쇼핑 소비자의 불만 개선에 미치는 영향, 성균관대학교 일반대학원, 석사학위논문, 2022
- [3] Ball, D., Coelho, P. S., & Vilares, M. J., “Service personalization and loyalty”, Journal of services marketing, 20(6), 391-403, 2006
- [4] Aggarwal, P., Castleberry, S.B., Ridnour, R., & Shepherd, C.D., “Salesperson empathy and listening: Impact on relationship outcomes”, Journal of Marketing Theory and Practice, 13(3), 16-31, 2005
- [5] Mingming Liu, Qicheng Ding, Yu Zhang, Guoguang Zhao, Changjian Hu, Jiangtao Gong, Penghui Xu, Yu Zhang, Liuxin Zhang, and Qianying Wang. “Cold Comfort Matters - How Channel-Wise Emotional Strategies Help in a Customer Service Chatbot”, the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20), Association for Computing Machinery, 1-7, 2020
- [6] 김소연, 챗봇의 특성이 이용의도에 미치는 영향: 공감적 특성을 중심으로, 성균관대학교 일반대학원, 석사학위논문, 2020
- [7] Myers IB., McCaulley MH, 김정택 외 역, MBTI 개발과 활용: Theory, Psychometrics Application, 어세스타, p.54, 2007
- [8] 이혜원, “한국인의 대화에 나타난 공감적 청자 반응 전략 연구”, 한글, (312), 5-35, 2016
- [9] Rogers, C., ‘The necessary and sufficient conditions of therapeutic personality change’, Journal of Consulting Psychology 21, pp. 95~103, 1957
- [10] 전은주, 말하기 듣기 교육론, 박이정, 1999
- [11] Myers IB., McCaulley MH, 김정택 외 역, MBTI 개발과 활용: Theory, Psychometrics Application, 어세스타, p.54, 2007
- [12] 성지민, 최윤중, 광도연, 김한샘, “성격유형별 문체 특성 기반 맞춤형 광고 메시지 자동생성 연구”, 제 34회 한글 및 한국어 정보처리 학술대회 논문집, pp.431-436, 2022
- [13] William Griffitt and Russell Veitch, “Preacquaintance attitude similarity and attraction revisited: Ten days in a fall-out shelter”. Sociometry 37, 2 (1974), p.163-173, 1974
- [14] Byron Reeves and Clifford Nass, The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. CLSI and Cambridge University Press, Cambridge, 1996
- [15] 진보래, “인간과 대화형 에이전트 간 관계 형성의 효과: 챗봇의 칭찬이나 비판에 사용자 친밀감이 미치는 영향”, 한국HCI학회 논문지, 14(3), p.13-20, 2019
- [16] 장지혜, 주다영, “인공지능 스피커의 정서별 감정발화에 따른 사용성 평가”. 한국HCI학회 학술대회, (), 705-712, 2019
- [17] 이한나. 인간-로봇 상호작용을 위한 사회적 대화 전략 모델과 사용자 평가. 국내석사학위논문 고려대학교 대학원, 2022