

언어학적 단서를 활용한 동화 텍스트 내 발화문의 화자 파악

민혜진 정진우 박종철*†

Korea Advanced Institute of Science
and Technology

Hye-Jin Min, Jin-Woo Chung, and Jong C. Park. 2013. Identification of Speakers in Fairytales with Linguistic Clues. *Language and Information 17.2*, 93–121. Identifying the speakers of individual utterances mentioned in textual stories is an important step towards developing applications that involve the use of unique characteristics of speakers in stories, such as robot storytelling and story-to-scene generation. Despite the usefulness, it is a challenging task because not only human entities but also animals and even inanimate objects can become speakers especially in fairytales so that the number of candidates is much more than that in other types of text. In addition, since the action of speaking is not always mentioned explicitly, it is necessary to infer the speaker from the implicitly mentioned speaking behaviors such as appearances or emotional expressions. In this paper, we investigate a method to exploit linguistic clues to identify the speakers of utterances from textual fairytale stories in Korean, especially in order to handle such challenging issues. Compared with the previous work, the present work takes into account additional linguistic features such as vocative roles and pairs of conversation participants, and proposes the use of discourse-level turn-taking behaviors between speakers to further reduce the number of possible candidate speakers. We describe a simple rule-based method to choose a speaker from candidates based on such linguistic features and turn-taking behaviors. (**Korea Advanced Institute of Science and Technology**)

Key words: speaker identification, turn-taking, robot storytelling, quoted speech, quote attribution

* 본 논문은 한국연구재단의 지원을 받아 수행된 기초연구 사업의 일환으로 수행된 연구임 (No. 2010-0012527).

† 대전광역시 유성구 대학로 291, 한국과학기술원 (KAIST) 전산학과, 전화번호 042-350-7741, E-mail: park@cs.kaist.ac.kr

1. 서론

최근 전산학, 인문학, 기계공학, 전자공학 등 다양한 학문 간의 융합에 바탕을 둔 로봇 기술의 발전으로 로봇의 하드웨어 구현 및 제어뿐만 아니라 특정 사용자층을 대상으로 하는 소프트웨어가 개발되고 있다(Fong *et al.*, 2003; Goodrich and Schultz, 2007; Ray and Siegart, 2008). 이 중 아동을 대상으로 하는 교육용 애플리케이션 분야에서는 로봇이 동화와 같은 이야기를 들려주는 ‘로봇 스토리텔링(Robot Storytelling)’에 대한 연구가 활발히 이루어지고 있다(Plaisant *et al.*, 2000; Mutlu *et al.*, 2006; Ribeiro and Costa, 2009; Gelin *et al.*, 2010; Fridin, 2014).

보다 흥미롭고 생동감 있는 스토리텔링을 위해서는 등장인물 간 대화나 독백에 해당하는 발화문 상의 감정을 파악하고 이를 기반으로 음성을 합성하는 것이 효과적이다(Lee and Park, 2009; Montañaño *et al.*, 2013). 또한, 등장인물의 성별이나 연령 등을 고려하여 인물별 TTS (Text-To-Speech) 음성을 적절히 할당하여 활용한다면 더욱 자연스러운 스토리텔링이 이루어질 수 있다(Min *et al.*, 2013). 이와 같이 발화문 상의 감정을 파악하고 등장인물별 TTS 음성 할당을 할 때 발화문의 화자를 파악하는 문제가 중요하게 다루어져야 한다. 예를 들어, 동화 ‘잭과 콩나무’의 일부분인 아래 (1)의 예제에서 등장인물 ‘잭의 어머니’가 발화한 (1b)와 (1e)에서는 화자의 감정상태인 ‘화남(Anger)’을 알 수 있는 명시적인 표현이 있다고 보기 어려운 반면, 내레이션 (1c)와 (1f)에서는 “어이가 없다”, “화를 내다” 등 비교적 명시적인 표현들이 나타나 있다고 볼 수 있다. 따라서 내레이션에 명시적으로 드러난 표현의 주제와 발화문의 화자와의 일치 여부를 알 수 있다면 이를 발화문의 감정파악에 활용할 수 있을 것이다(이호준, 박종철, 2011). 마찬가지로 발화문 (1b), (1e) 및 (1d)의 음성 합성 시, 각각 화자를 파악할 수 있다면 음성 합성 과정에서 이들을 구분하여 좀 더 생동감 있는 음성 발화문을 도출할 수 있다(Min *et al.*, 2013).

- (1) a. 잭은 젓소를 콩 한 주머니와 바꾸어버렸어요.
- b. “젓소를 겨우 콩 한 주머니와 바꾸었다고?” (화자: 잭의 어머니)
- c. 잭의 어머니는 어이가 없었어요.
- d. “이 콩은 요술콩이래요.” (화자: 잭)
- e. “잭, 너는 그말을 믿는단 말이나?” (화자: 잭의 어머니)
- f. 잭의 어머니는 화를 내며 창밖으로 콩을 내던졌어요.

예문 (1) 또는 뉴스나 설명문에서는 대부분의 경우 화자가 사람이지만(O’Keefe *et al.*, 2012), 동화에서는 사람 이외에도 다음 예문 (2)에서와 같이 동물이나 무생물도 화자가 될 수 있다는 점에서 다른 장르에 비해 화자 선택 과정에서 많은 애매성이 따른다. 예를 들어, 아래 예문에서 밑줄로 강조된 ‘잭’, ‘금하프’, ‘콩나무’, ‘거인’ 모두 발화문 (2b)의 화자를 지칭하는 표현(이후로 **화자 멘션**으로 표기한다.)의 후보가 될 수 있음을 확인할 수 있다.

- (2) a. 책은 금하프를 안고 콩나무를 향해 뛰었어요.
 b. “주인님, 도둑이에요.” (화자: 하프)
 c. 하프가 소리치자, 거인이 깨어났어요.

동화 텍스트에서 각 발화문에 대한 올바른 화자를 파악하기 위해 기존 연구에서는 주로 (2c)에서의 발화동사(speech verb)나 (3c)에서의 등장인물의 등장이나 퇴장에 해당하는 동사(‘들어오다’)의 주체를 화자 멘션으로 채택하는 방법을 사용하였다.

- (3) a. 엄지공주는 호두껍질로 만든 침대에서 자고 있었어요.
 b. “하하하, 예쁜 아가씨군.” (화자: 두꺼비)
 c. 무섭게 생긴 두꺼비가 몰래 들어와 엄지공주를 연못으로 데려갔어요.

그러나 다음의 예문에서와 같이 발화문 (4c)의 화자 파악 시, 발화동사나 등장동사에 대한 단서를 가지는 화자 후보가 여러 개 존재 하는 경우 (예의 경우, ‘구름’, ‘아버지’), 이 중 어떤 후보가 (4c)의 화자 멘션인지 파악하기는 어렵다.

- (4) a. 쥐 아버지는 곧 딸을 데리고 구름에게 부탁하러 갔어요.
 b. 구름에게 자기 딸을 신부로 맞아 주라고 하자, 구름이 말했어요.
 c. “아냐, 아냐. 나보다 훨씬 더 힘센 분이 계시는 걸.” (화자: 구름)
 d. 그 말을 듣고 아버지는 말했어요.
 e. “그 분이 누구십니까? 이름을 좀 가르쳐 주십시오.” (화자: 아버지)

그런데 위 예문에서 등장인물 ‘구름’과 ‘아버지’가 서로 대화를 하고 있고 이들이 대화를 주고 받는 턴테이킹(turn-taking) 행위를 하고 있다는 사실을 활용하면 연속된 발화문에서 두 발화문이 모두 동일한 화자가 될 경우를 배제시킬 수 있다.

본 논문에서는 동화 텍스트 상의 발화문의 화자를 파악하기 위하여 기존 연구에서 활용되었던 단서인 화자 후보들의 위치, 주어, 목적어 여부 및 발화동사, 등장/퇴장 동사의 존재 여부, 화자 후보들의 의미적 분류를 포함하여 대화 시 나타나는 턴테이킹(turn-taking) 행위에 기반한 단서인 등장인물의 호격 정보와 연속발화문의 대화참여자 쌍에 대하여 논의하고 이 단서들을 활용한 화자 파악 방법에 대하여 논의한다.

본 논문은 다음과 같은 순서로 기술된다. 2절에서는 화자 파악에 관련된 기존의 연구에 대하여 살펴본다. 3절에서는 화자 파악을 위한 단서들을 관련 예문을 통해 제시하고 4절에서는 이러한 단서를 기반으로 한 화자 파악 방법에 대하여 소개한다. 5절에서는 본 연구에서 제안한 방법론에 대한 평가 결과에 대하여 논의하고, 6절에서는 제안한 방법론의 의의와 한계에 대하여 토의하며, 7절의 결론 및 향후 계획으로 본 논문을 맺는다.

2. 관련연구

화자 파악에 대한 연구는 아동을 위한 동화나 소설 등을 대상으로 규칙기반 방법에 의해 먼저 연구되었다. Zhang *et al.* (2003)은 발화문을 기준으로 한 문장의 상대적 위치를 주요 단서로 고려하였다. 이 연구는 연구결과를 통해 각 동화의 특성이 정확도에 영향을 미친다는 점을 시사하고 있으나, 확장성(scalability) 및 적용성(applicability) 면에서 한계를 가진다. Glass & Bangay (2007)는 사전의 기계학습이나 배경정보 없이 주어진 텍스트로부터 파악되는 자질만을 활용하여 자질들의 효용성을 파악하였다. 이들은 화자 후보 파악 단계에서는 주어/목적어 여부, 품사 태그, 고유명사, 약어, 동사와의 거리를 고려하였고, 화자 파악 단계에서는 명칭(alias) 및 소유격, 전치사 등을 고려하여 각 자질 별 점수를 할당하였다. 이 연구에서 활용한 명칭 정보는 유용하지만 그 수가 너무 많아지면 정확도가 오히려 낮아지는 문제가 발생한다. Mamede & Chaleira (2004)는 휴리스틱(heuristic) 규칙을 활용하여 발화문을 식별하고 역시 규칙 기반으로 단일 명사구에 한해 화자를 파악하는 연구를 진행하였다. 이들은 학습 데이터 및 검증 데이터에 대해 각각 84.8%, 65.7%의 정확도를 달성하였다. 그러나 이 연구에서 제안된 방법론은 68%의 발화문에서 화자를 선정하는 규칙이 적용되지 않는 문제점을 보이며 34.3%, 10.6%의 낮은 재현율을 얻었다.

최근 소설이나 뉴스와 같은 도메인에서는 통계적인 방법론이 활발히 적용되고 있다. Elson & McKeown (2010)에서는 영어, 러시아어, 프랑스어로 작성된 19, 20세기 소설들로부터 발화문의 화자 후보를 선정하고 화자를 파악하였다. 이들은 발화문을 몇 가지 유형으로 분류하고 해당 유형을 식별할 수 있는 자질을 벡터화하여 기계 학습 기법을 적용하였다. 이들의 연구는 상대적 위치가 가장 가까운 후보를 화자로 선택하는 베이스라인(baseline)에 비하여 30% 이상 높은 평균 83%의 정확도를 달성하였다. O'Keefe *et al.* (2012)는 뉴스 도메인에서 Elson & McKeown의 연구에서와 마찬가지로 발화문을 여러 유형으로 분류하고, 화자 파악을 시퀀스 레이블링(sequence labeling) 문제로 정의하여 greedy, Viterbi 및 linear chain CRF와 같은 시퀀스 디코딩 모델을 제안하였다. 이에 화자/비화자 이진 클래스의 경우에는 시퀀스를 적용하지 않은 모델이, 화자 n명의 클래스의 경우에는 Viterbi 디코딩이 가장 높은 성능을 보였다.

본 연구는 동화 텍스트를 대상으로 한다는 점에서 기존의 규칙기반 방법론을 토대로 한 연구와 유사하나, 동식물이나 무생물까지 화자 후보로 고려하여 문제의 복잡도가 더 높으며 다양한 장르의 동화를 포함했다는 점에서 데이터가 보다 균형적이라는 특징이 있다. 본 연구의 선행연구인 민혜진 외(2012)에서는 화자 파악 문제에 관련하여 각 화자 후보를 화자 또는 비화자로 분류하는 문제로 간주하고 의사결정 트리 기반의 분류기(Decision Tree-based classifier)를 제안하였다. 분류기에 활용된 자질은 규칙을 기반으로 한 기존 연구에서 활용된 화자 후보의 상대적 위치, 격 정보, 발화동사 존재 및 데이터 분석을 통해 추가로 제안한 등장/퇴장을 의미하는 동사와 화자 후보의 의미적 분류였다. 이 연구에서는 화자 후보의 의미적 분류 및 발화동

사, 등장/퇴장 동사의 존재 여부를 고려한 분류기가 규칙 기반의 베이스라인에 비해 최대 49%의 정확도 향상을 달성하였음을 보였으며 오류분석 결과를 통해 등장인물 간의 관계 정보 및 연속 발화문 상의 대화 참여자 정보와 같은 추가적인 단서의 필요성을 제시하였다. 이에 본 연구에서는 선행연구의 오류분석 과정에서 논의된 연속 발화문 상의 대화 참여자 정보 및 선행연구에서 활용한 격 정보 가운데 호격 정보를 보다 정교하게 분석하여 이 단서들을 규칙 기반 분류기에 활용하였다.

3. 화자 파악을 위한 단서

3.1 개관

3.1.1 동화 상의 발화문의 화자를 가리키는 패턴/문장 구조. 동화나 소설과 같은 이야기 글은 이야기의 중심 내용을 기술하는 문장(내레이션으로 표기)과 이야기 내 등장인물의 독백 또는 대화가 나타나는 문장(발화문으로 표기)으로 구성된다. 등장인물 간의 대화가 이루어지는 부분에서는 독자의 이해를 돕기 위하여 예문 (1)~(4)에서와 같이 내레이션 문장에서 종종 발화문의 화자를 명시한다. 예문 (4b)나 (4d)에서와 같이 내레이션이 다음에 나타나는 발화문의 화자를 직접적으로 명시하는 경우가 있는가 하면, 예문 (3c)와 같이 등장인물이 장면에서 등장함을 묘사하여 간접적으로 명시하는 경우도 있다.

직접적으로 화자를 명시하는 경우에는 ‘말하다’와 같은 종류의 동사의 주체로 화자를 드러내게 된다. 간접적으로 화자를 명시하는 경우에는 발화를 암시하는 등장인물의 행동을 나타내는 동사, 예를 들면 장소를 나타내는 동사 ‘나타나다’, 또는 감정을 표현하는 동사 ‘화내다’ 등을 활용하고 서술어의 논항 중 하나로 화자를 드러내게 된다. 두 경우 모두에서 화자를 명시하는 내레이션 문장이 발화문과 멀리 떨어져 있다면 독자가 해당 화자를 파악하기 어려울 것이므로 예문 (1)~(4)에서와 같이 화자 멘션을 포함하는 내레이션 문장은 비교적 발화문과 가까이에 위치할 것이다. 만약, 등장인물 간 대화가 연속발화문과 같은 형태로 나타난다면 턴테이킹이 어느 시점에서 이루어질 수 있는데, 이때는 발화문 내에 턴테이킹을 가리키는 단서가 존재하여 독자로 하여금 각 발화문의 화자를 잘 파악할 수 있도록 글이 기술되어 있을 것이다.

본 논문에서는 위와 같은 가설을 토대로 동화 상의 발화문의 화자를 파악하기 위한 단서를 데이터로부터 분석하는 작업을 진행하였다. 3.1.2 절에서는 분석/검증 데이터에 대하여 기술하였고, 각 단서들은 3.2~3.7 절에 보다 자세히 기술하였다.

3.1.2 동화 데이터. 본 논문에서는 전래 동화, 외래 동화, 창작 동화의 세 가지 장르로 구성된 총 20 편의 동화를 데이터 집합으로 사용하였다. 또한, 동식물 및 무생물이 발화하는 빈도가 높은 동화들과 그렇지 않은 동화들을 분석 데이터 집합과 검증 데이터 집합에 균형있게 분포시켜 데이터 편향성을 최소화하였다. 분석 및 실험을 통한 검증에 사용된 동화목록은 표 1과 같다.

데이터 집합은 총 1,088 문장으로 이루어져 있으며 이 중 발화문은 인용부호 큰 따옴표(“ ”), 작은따옴표(‘ ’)가 문장의 앞과 뒤에 나타나는 문장으로 총 353 개이다

| 구분 | 장르 | 제목 |
|-------|-------|--|
| 분석데이터 | 전래 동화 | 콩쥐 팥쥐, <u>은혜 값은 가치</u> , <u>효성다한 호랑이</u> |
| | 외래 동화 | 잭과 콩나무, 피노키오, 헨젤과 그레텔, 쥐의 신랑감, 엄지공주 |
| | 창작 동화 | 찌찌와 뽀뽀의 여행, <u>타조의 세가지 보물</u> |
| 검증데이터 | 전래 동화 | 금도끼 은도끼, 무덤가에 핀 꽃, 빨간부채 파란부채 |
| | 외래 동화 | 알라딘의 램프, 개구리 왕자, <u>빨간모자</u> , <u>아기 돼지 삼형제</u> , <u>별거뻐은 임금님</u> |
| | 창작 동화 | <u>그루터기의 새싹</u> , <u>농부와 난장이들</u> |

[표 1] 동화목록 (밑줄: 동식물, 무생물의 발화빈도가 높은 동화)

(인용부호 쌍 내에 속하는 문장을 한 단위로 간주하였을 때 발화문은 327 개이다.)

정답 집합을 구축하기 위해 분석 데이터 집합과 검증 데이터 집합에 대해 각각 두 명의 연구원들이 주석(annotation)을 진행하였다. 주석 정보는 각 발화문의 화자와 그 화자가 나타나 있는 문장의 상대적 위치이다. 분석 데이터 집합에서는 화자와 상대적 위치가 각각 94.3%, 73.6%의 일치율(inter-annotator agreement)을 보였으며 검증 데이터 집합에서는 각각 89.1%, 72.6%의 일치율을 보였다. 일치율에 근거하여 화자 및 화자의 상대적 위치는 사람이 비교적 명료하게 파악할 수 있음을 있었다. 상대적 위치의 경우, 화자가 서로 다르게 태깅되면 상대적 위치가 일치할 가능성이 거의 희박하기 때문에 이를 배제하면 일치율이 각각 78.0%, 81.4%까지 높아짐을 볼 수 있었다.

3.2 상대적 위치

동화를 비롯한 소설 등의 문학작품에서는 드라마 대본이나 영화 시나리오와는 달리 발화문 전후의 내레이션 문장에서 발화문의 화자를 언급하는 경우가 많다. 예문 (1)에서와 같이 내레이션 문장에서 이전 발화문의 화자의 감정상태를 언급하면서 발화문에서는 나타나지 않았던 부가적인 설명을 덧붙이기도 한다. 이와 같은 동화 텍스트의 특성을 통해 해당 발화문을 중심으로 한 화자 멘션 후보들의 상대적인 위치가 화자 파악의 중요한 단서가 될 수 있다. 예문 (1)~(4)를 살펴보면, 각 화자 멘션은 발화문을 중심으로 한 문장 앞(-1) 또는 뒤(+1)에서 언급되고 있다. 그러나 예문 (5)의 발화문 (5e)에서와 같이 화자를 명시적으로 언급하지 않아 화자 멘션인 ‘바람’이 발화문에서 비교적 멀리 떨어진 (5a)에 위치하기도 한다.

- (5) a. 그래서 이번에는 바람에게 가서 부탁했어요. 그러자 바람이 말했어요.
 b. “아냐, 아냐. 나보다 훨씬 더 힘센 분이 계시는걸. ” (화자: 바람)
 c. 그 말을 듣고 아버지는 깜짝 놀라 물었어요.
 d. “그분은 도대체 누구입니까? ” (화자: 아버지)
 e. “벽님이지. 내가 아무리 힘껏 불어도 벽님은 끄덕없어. 벽님에겐 도저히 못 당해.” (화자: 바람)

f. 이번에는 벽을 찾아가 부탁했어요.

화자 멘션이 나타나는 상대적 위치의 분포를 파악하기 위하여 분석 데이터인 10 편의 동화(표 1 참조)에 대한 주석 정보를 분석하였다. 표 2와 표 3은 화자 멘션이 나타나는 상대적 위치를 분석한 것이다.

| 상대적 위치 | 주석자 1 | 주석자 2 |
|--------|-------|-------|
| -5 이상 | 10 | 9 |
| -4 | 6 | 5 |
| -3 | 7 | 5 |
| -2 | 20 | 13 |
| -1 | 76 | 73 |
| 0 | 5 | 1 |
| 1 | 44 | 55 |
| 2 | 0 | 5 |
| 3 | 0 | 1 |
| 4 | 0 | 1 |
| 5 이상 | 1 | 1 |
| 총합 | 169 | 169 |

[표 2] 화자 멘션의 상대적 위치

| | 4 | 1 | +4 | -4 | +1 | -1 |
|------|-------|-------|-------|-------|-------|-------|
| 주석자1 | 94.0% | 73.9% | 26.0% | 64.4% | 26.0% | 44.9% |
| 주석자2 | 93.4% | 76.3% | 36.6% | 56.8% | 32.5% | 43.3% |
| 평균 | 93.7% | 75.1% | 31.3% | 60.6% | 29.2% | 44.0% |

[표 3] 상대적 위치별 화자 멘션의 비중

표에 나타난 것과 같이, 발화문의 다음 문장보다 이전 문장에, 먼 문장보다 가까운 문장에 화자 멘션이 나타나는 경우가 많았다. 예를 들면 발화문에서 앞으로 4 번째 이내, 1 번째 이내 떨어져 있는 문장이 각각 평균적으로 전체 발화문의 60.6%, 44.0%를 차지하였고, 뒤로 4 번째 이내, 1 번째 이내 떨어진 문장은 각각 31.3%, 44.0%를 차지한 것을 확인할 수 있다. 또한, 발화문 앞 뒤로 4 번째 이내 떨어진 문장 내에 화자 멘션이 존재하는 경우가 전체 발화문의 93.7%를 차지하였고, 앞 뒤로 1 번째 이내 떨어진 문장 내에 존재하는 경우가 전체 발화문의 75.1%를 차지하였다. 분석한 결과를 종합하면, 발화문 바로 앞 문장에 존재하는 후보가 화자 멘션일 가능성이 가장 높음을 알 수 있다.

그러나 화자 멘션 위치가 6~7 개 이전 또는 4 개 이후에 나타나는 경우도 약 6% 정도 존재하는데 이는 발화문이 연속적으로 나타나는 경우로 발화문 내 존재하는

‘나’, ‘너’와 같은 지시체(referent)는 화자 멘션으로 간주하지 않았기 때문에 발생하는 것으로 보인다. 이와 같이 화자 멘션이 멀리 떨어져 있는 경우, 동화 스토리텔링의 주 사용층인 유아 및 아동의 인지적 특성에 맞는지에 대해서는 유아/아동의 각 연령대에 맞는 동화 집합을 재구성하여 주석을 진행한 후 추가 분석을 하는 것이 필요한데, 이는 추후 연구로 남겨둔다.

3.3 화자 멘션의 격 정보

화자 멘션의 상대적인 위치는 기준 발화문으로부터의 단순 거리만을 나타내는 단서이나, 화자 멘션 후보의 문장성분 및 격 정보는 문장에서 해당 후보가 어떠한 역할을 하는지를 가리키므로 보다 중요한 단서가 될 수 있다(Zhang *et al.*, 2003; Glass and Bangay, 2007). 예를 들어, 문장에서 주어 역할을 하는 (3c)의 ‘두꺼비’는 동일 문장에 있는 후보 ‘엄지공주’, ‘연못’ 보다 화자 멘션일 가능성이 높다. 또한, (1b), (1e)의 화자인 ‘잭의 어머니’의 경우 주제화 표지 ‘는’과 결합해 있는데, 이는 화자 멘션의 후보가 주제격 명사이면서 문장에서 주어 역할을 하는 경우 다른 후보(가령 ‘콩’)에 비해 실제 화자 멘션에 해당할 가능성이 높음을 보여준다.

그러나 주격 또는 주제화 표지가 붙어 있는 후보 외에 다른 문장성분을 가진 후보도 화자 멘션이 되는 경우가 있다. 예문 (6)에서는 목적격을 취하는 후보인 ‘할아버지’가 (6b)의 화자이며 예문 (7)에서는 소유격에 위치하는 후보가 화자 멘션임을 알 수 있다.

- (6) a. 잭은 길에서 이상한 할아버지를 만났어요.
 b. “꼬마야, 젓소를 이 콩과 바꾸지 않겠니?” (화자: 할아버지)
 c. “안돼요. 젓소를 팔아서 어머니 약을 사야 해요.” (화자: 잭)
 d. “이 콩은 요술콩이야. 이 콩을 심으면 너의 집이 행복해 질 거란다.” (화자: 잭)
- (7) a. “어머니가 돌아가셨으니 내 무슨 기쁨으로 살겠는가. 멀찌감치에서라도 어머니를 뵈 수 있다는 게 기뻐다네. 나는 어머니 생전에 효도 한 번 제대로 못 해 드렸는데 이런 불효자가 살아 무엇하겠어. 어머니를 위해 사냥할 때가 가장 행복했었지...” (화자: 호랑이)
 b. 호랑이의 얼굴에는 행복한 미소가 어려 있었어요.

예문 (6)에서 목적격에 위치하는 ‘할아버지’의 경우 문맥 상 처음 등장하여 같은 문장에 존재하는 주제화 표지가 붙어 있는 ‘잭’보다 더 강조되고 있다. 이 경우에서 보듯이 격 정보 단서만으로도 화자 파악이 어려우며, 해당 문장 내의 다른 단서도 함께 고려되어야 한다.

예문 (7b)와 같이 소유격에 위치하는 후보는 수식하는 명사의 종류에 따라 화자 멘션이 될 가능성이 달라진다. 즉, (7b)에서 수식을 받는 멘션이 소유격에 위치하는 멘션의 일부분을 가리키는 경우에는 말을 하는 주제일 가능성이 낮으므로 소유격에

위치하는 멘션 ‘호랑이’가 수식을 받는 멘션 ‘얼굴’에 비해 화자 멘션일 가능성이 높은 반면, (1b)나 (1e)의 ‘책의 어머니’와 같이 어머니의 속성을 보다 구체적으로 나타내주는 기능을 하는 멘션일 경우, 수식을 받는 멘션, ‘어머니’가 소유격에 위치하는 멘션 ‘책’에 비해 화자 멘션이 될 가능성이 높다. 따라서 본 연구에서는 소유격의 수식을 받는 명사가 화자가 될 가능성이 높은 사람, 동/식물, 무생물의 객체를 가리키는 명사이면 이를 화자 멘션 후보로 간주하고, 예문 (7b)의 얼굴과 같이 객체의 일부분을 가리키는 명사일 경우 소유격에 위치하는 명사를 후보로 간주하였다. 이외에 객체의 일부분을 가리키는 명사와 객체 자체를 가리키는 명사와의 관계가 소유격이 아닌 다른 격으로 표현될 수도 있으나(예: “공주님이 발을 동동 굴리며 속상해 하고 있을 때였어요.”), 이러한 소유관계를 파악하기 위해서는 보다 정교한 의미분석이 이루어져야 하는데, 이는 추후 연구로 남겨둔다.

3.4 발화동사

말을 하거나 소리를 내는 행위를 가리키는 발화동사(speech verb)는 동사의 주체가 되는 후보가 화자 멘션임을 명확하게 해주는 단서로 활용될 수 있다. 예를 들면, (2c)의 ‘소리치다’나 (8a)의 ‘말하다’ 동사는 그 주체인 ‘하프’와 ‘요정’이 각각 발화문 (2b)와 (8b)의 화자임을 가리키는 단서가 된다. 본 연구에서는 ‘말하다’, ‘묻다’, ‘대답하다’와 같은 대표적인 동사를 포함하여 세종사전의 용언의 의미 부류 체계¹에서 표 4와 같은 10개의 의미부류(sem_class)에 속하는 용언들을 대상으로 발화동사 사전을 구축하였다. 발화동사 사전 구축 시, 각 의미부류에 속하는 용언들 가운데 일부는 사전 항목에 포함시키지 않았다. 예를 들면, 심리행위에 속하는 ‘쓰이다’와 같은 동사는, “시험 준비하는 데 신경이 많이 쓰이는 모양이구나.”와 같은 문장에서와 같이 ‘신경’, ‘힘’, ‘머리’와 같은 특정 명사구가 논항(argument) 이어야만 심리행위를 의미한다. 이와 같은 동사도 항목에 포함시키기 위해서는 구문 단계의 분석이 필요하다. 본 연구에서는 ‘말’에 해당하는 명사(구)에 한하여 해당 문장에 나타나는 경우 이에 가장 근접한 동사를 파악하여 발화를 나타내는 서술어구를 파악하였고, 나머지 서술어구의 분석은 추후 연구로 남겨둔다.

- (8) a. 요정이 구름 속에서 나타나 말했어요.
 b. “너는 참 용감하구나. 앞으로 좋은 일이 생길 거야.” (화자: 요정)
 c. 요정은 말을 마치고 곧 사라졌어요.
 (중략)
 d. 문이 덜컥 열리더니 거인이 나타났어요. 거인은 암탉을 탁자 위에 놓고 말했어요.
 e. “알을 낳아라.” (화자: 거인)

¹ 세종전자사전, 21세기 세종계획, 국립국어원, <http://www.sejong.or.kr/>

| 의미 부류 (sem-class) | 개수 | 예 |
|-------------------|-----|-----------------|
| 질문 | 22 | 묻다, 되묻다, 자문하다 |
| 대답 | 14 | 대답하다, 답하다, 대꾸하다 |
| 대화 | 16 | 곶속말하다, 대화하다 |
| 명령 | 20 | 요구하다, 경고하다 |
| 보고제안 | 84 | 말하다, 참견하다 |
| 비난욕설 | 101 | 간죽대다, 다그치다 |
| 소리/소리내기 | 214 | 노래하다, 소리치다, 울다 |
| 소통행위 | 298 | 나불거리다, 부르다 |
| 언어행위 | 25 | 해명하다, 확인하다 |
| 심리행위 | 120 | 부끄러워하다, 분통터트리다 |

[표 4] 발화동사 사전 구성

3.5 등장/퇴장 동사

3.1.1 절에서 논의한 것과 같이 등장인물이 주어진 상황에 처음 나타나거나 사라짐을 의미하는 동사의 경우 화자 파악의 단서로 활용될 수 있다. 예를 들어, (3c)에서의 ‘들어오다’ 및 (8a)의 ‘나타나다’는 등장인물의 등장을 가리키므로 그것의 주체가 되는 ‘두꺼비’ 및 ‘요정’이 주변 발화문의 화자 멘션임을 암시한다. 또한, (6a)에서의 ‘만나다’는 그것의 목적격인 ‘할아버지’가 등장함을 가리키므로 바로 이어지는 발화문의 화자를 알려주는 단서가 될 수 있다. 한편, (8c)의 ‘사라지다’는 등장인물의 퇴장을 가리키며, 이는 동사의 주체가 되는 ‘요정’이 이후에 나타나는 발화문(예: 8e)의 화자가 될 가능성이 매우 낮다는 것을 암시하게 된다.

본 연구에서는 발화동사 사전 구축 방법과 유사하게 표 5와 같이 세종사전의 용언의 의미 부류 체계에서 11개의 의미부류(sem-class)에 속하는 용언들을 수집하고 이를 통해 등장인물의 등장/퇴장 사건을 구축하여 화자 파악의 단서로 활용하였다.

3.6 등장인물의 의미적 분류

앞서 제시한 화자 파악의 단서 외에 화자 멘션의 후보가 되는 명사 자체의 특성을 이용하여 그것을 후보에서 미리 제외할 수 있다면 화자 파악의 성능을 높이는 데 도움이 될 것이다. 이를 위해 발화문 주변의 명사들 중에 사람을 지칭하는 명사들만 화자 멘션의 후보로 선택하는 방법을 고려해 볼 수 있다. 그러나 동화에서는 사람 이외에 동물(예: 두꺼비(3c)), 식물(예: 완두콩), 그리고 무생물(예: 하프(2c))도 경우에 따라 화자가 될 수 있으므로, 동물이나 사물을 지칭하는 명사를 화자 멘션 후보에서 제외할 수 없다. 이와 반대로 텍스트에 언급된 모든 명사구를 화자 멘션의 후보에 포함시키는 것도 후보의 수를 너무 많아지게 하여 화자 선택을 어렵게 하므로 바람직하지 않다.

본 논문에서는 발화의 가능성이 있는 후보를 제외하지 않으면서도 후보의 수를 가급적 줄이기 위하여 동화 텍스트 내 해당 후보가 언급되는 빈도와 사람, 동물, 식물

| 범주 | 의미 부류 | 항목 수 | 예 |
|----------|-----------|------|-----------------------|
| 등장 (6 개) | 도착 | 50 | 갔다오다, 날아오다 |
| | 사동적 이동행위 | 15 | 데려오다 |
| | 이동행위 | 80 | 나타나다, 들어오다, 내려오다, 나오다 |
| | 전이 | 2 | 돌아오다 |
| | 접근 | 6 | 다가오다, 뒤쫓아오다, 만나다 |
| | 출생 | 13 | 태어나다 |
| 퇴장 (8 개) | 출발 | 11 | 가다, 도망가다 |
| | 사동적 이동행위 | 8 | 데려가다 |
| | 이동 / 이동행위 | 99 | 가다, 날아가다, 사라지다 |
| | 접근 | 7 | 다가가다, 쫓아가다 |
| | 탈퇴 | 19 | 도피하다, 탈퇴하다 |
| | 하차 | 2 | 내리다 |
| | 소멸 / 소멸행위 | 102 | 달아나다, 물러가다 |
| | 죽음 | 54 | 죽다, 숨지다 |

[표 5] 등장/퇴장 동사 사전 구성

등과 같은 후보 등장인물들의 의미적 부류 (sem_class) 를 활용하였다. 이를 위하여 세종사전의 체언의 어휘 의미 부류 체계 (이성현, 2005) 에서 화자가 될 가능성이 높은 사람 (예: 남성 / 여성인간, 연령인간, 비대칭적 친족), 동물 (예: 짐승, 새, 벌레), 식물 (예: 나무, 꽃), 무생물 (예: 해, 구름) 과 관련된 의미 부류에 속하는 체언들을 중심으로 동화 등장인물 사전을 구축하였다. 표 6은 동화 등장인물 사전에 포함되는 체언의 예를 보여준다.

세종사전에서 하나의 명사는 여러 의미 부류에 속할 수 있다. 예를 들면, ‘개’는 짐승이라는 부류와 부정적 속성을 가지는 인간이라는 부류에 모두 속한다. 본 논문에서는 동화에서 자주 나타나는 명사의 의미 부류에 우선순위를 부여하여 사전을 구축하였다. 즉, ‘개’는 짐승에 해당되는 것으로 분류하였다.

한편, 예문 (1)의 ‘잭’, 예문 (3)의 ‘엄지공주’와 같이 고유명사로 지칭되는 화자는 등장인물 사전에 포함되지 않으므로 해당 동화에 나타나는 빈도가 높지 않을 경우 후보에서 제외할 수 있다. 그러나 동화의 특성 상 이와 같이 고유명사로 표현되는 등장인물은 대체로 주인공이나 부주인으로 동화의 이야기 전개에서 매우 중요한 역할을 담당하며, ‘콩쥐 팥쥐’와 같이 동화 제목에 자주 나타나는 경향을 보인다. 본 논문에서는 이러한 특성을 감안하여 고유명사인 후보가 제목에 등장할 경우, 후보에서 제외시키지 않도록 하였다.

3.7 턴테이킹 (turn-taking)

앞 절에서는 각각의 발화문을 독립적으로 간주한 상태에서 그것의 화자에 대한 단서를 추출하는 방법에 대해 논의했으나, 본 절에서는 동화 텍스트에 나타나는 대부분의

| 범주 | 의미 부류 | 개수 | 예 |
|-----------|--|-------|--------------------|
| 사람 (45 개) | 능력속성인간, 속성인간, 시간속성인간, 신체속성인간, 정도속성인간, 정신속성인간, 인간 | 251 | 선비, 쌍둥이, 대머리 |
| | 남성인간, 여성인간 | 37 | 남자아이, 계집아이, 총각, 색시 |
| | 연령인간 | 105 | 갓난아기, 늙은이, 소녀, 꼬마 |
| | 긍정적도덕속성인간, 긍정적 속성인간, 긍정적신체속성인간, 긍정적정신속성인간, 부정적도덕속성인간, 부정적 속성인간, 부정적신체속성인간, 부정적정신속성인간 | 542 | 가난뱅이, 개구쟁이, 부자 |
| | 상상적 인간 | 96 | 귀신, 거인 |
| | 계급인간, 반사회적 직업인간, 사회계급인간, 서열인간, 소속인간, 역할인간, 위계인간, 일시적역할 인간, 일시적인간집단, 종교역할인간, 직업인간, 직위인간 | 1,374 | 농부, 주인, 후견인 |
| | 비대칭적 친족 | 368 | 새엄마, 아빠, 손녀 |
| | 호칭 | 44 | 아바마마, 여편네, 친구 |
| | 기타 (결과인간, 결과행위인간, 관계인간, 권리인간, 대칭적관계인간, 비대칭적관계인간, 범위인간집단, 상태인간, 수동적행위인간, 지역인간, 행위인간, 화시적인간) | 585 | 가해자, 불합격자 |
| | 동물 (8 개) | 동물 | 68 |
| 짐승 | | 135 | 구렁이, 개, 고양이 |
| 벌레 | | 116 | 풍뎡이, 개미 |
| 새 | | 93 | 갈매기, 딱따구리 |
| 물고기/생선 | | 150 | 물고기, 도미 |
| 상상적 동물 | | 15 | 용, 공룡, 괴물 |
| 비인간집단 | | 2 | 물고기무리, 양떼 |
| 식물 (4 개) | 식물/식물의 부분 | 205 | 나무, 포도덩굴, 콩나물 |
| | 꽃 | 86 | 튤립, 장미꽃 |
| | 비인간집단 | 1 | 수풀 |
| 무생물 | 무생물 | 164 | 바람, 해, 구름, 하프 |

[표 6] 동화 등장인물 사전 구성

발화문이 대화문에 속한다는 사실을 고려하여 대화를 주고 받는 행위, 즉 턴테이킹(turn-taking)을 이용하여 여러 발화문에 대한 화자의 단서를 동시에 고려하는 방법에 대해 논의한다.

턴테이킹(turn-taking)은 대화의 중요한 요소인 한번에 한 사람이 말하기, 그리

고 한 사람이 말을 멈춘 후 다른 사람이 다시 시작하는 간격을 일정하게 유지하기 위해 대화과정에서 화자 간에 나타나는 행위이다(Sacks, 1992). 만약 화자 A와 화자 B가 서로 대화를 하고 있다면 A와 B는 턴을 주고 받을 것이며, 현재 발화문의 화자가 A 라면 다음 발화문의 화자는 턴테이킹에 의해 B로 바뀌거나 여전히 A 일 가능성이 매우 높고 대화에 참여하지 않는 인물일 경우는 거의 희박하다고 볼 수 있다. 즉, 턴테이킹을 주고 받는 화자를 미리 알고 있다면 턴테이킹에 참여하는 발화문의 화자를 파악하는 데 도움이 될 것이다.

동화 텍스트에서 턴테이킹은 예문 (9)과 같이 발화문이 연속적으로 나타나는 경우에 가장 많이 등장한다. 그러나 동화를 비롯한 소설 등에서는 발화문 사이에 내레이션이 존재하기도 하므로, 예문 (10)와 같이 발화문이 연속적으로 나타나지 않더라도 정황상 서로 같은 대화에 참여하고 있다면 턴테이킹이 일어나고 있다고 봐야 한다.

- (9) a. “정말 너를 한 번 보면 잊어버리지 않겠구나. 그런데 넌 이름이 뭐니?”
 b. “응, 난 세상에서 제일 큰 새 타조야.”
 c. “너도 새였구나.”
 d. “응, 그런데 난 날지 않아. 몸이 너무 뚱뚱해서 이 날개로는 날 수가 없어.”
 e. “그럼 먹이는 어떻게 구해? 또 힘센 동물이 나타나면 어떻게 해?”
 f. “하하하, 나에겐 세 가지 보물이 있어서 걱정 없어.”
- (10) a. 그래서 구름에게 자기 딸을 신부로 맞아 주라고 하자, 구름이 말했어요.
 b. “아냐, 아냐. 나보다 훨씬 더 힘센 분이 계시는걸.”
 c. 그 말을 듣고 아버지는 물었어요.
 d. “그분이 누구십니까? 이름을 좀 가르쳐 주십시오.”
 e. 구름이 대답했어요.
 f. “바람님이지. 아무리 내가 뽀내도 바람님이 불면 곧 쫓겨가게 되는 걸.”
- (11) a. 피노키오는 인형극에 정신을 빼앗겼어요.
 b. “우와, 재미있다.”
 c. “어라, 나무인형이 말을 하네. 이 녀석을 구경꾼들에게 보여줘야지.”
 d. 극장주인은 피노키오를 무대에 세우려고 창고에 가두어버렸어요.
- (12) a. 그날 밤, 헨젤은 살그머니 뜰로 나가 달빛에 반짝이는 조약돌을 주웠어요.
 b. “이것만 있으면 다시 돌아올 수 있어.”
 c. 헨젤은 조약돌을 담은 주머니를 숨겨놓고 잠자리에 들었어요.
 d. 다음날 아침이 되었어요.

- e. “애들아 나무하러 가자꾸나.”
- f. 새어머니가 다정한 척 하며 말했어요.

한편, 예문 (9), (10)에 대비되는 구조도 존재한다. 예를 들어, 예문 (11)의 경우, 발화문이 연속적으로 나타남에도 불구하고 턴테이킹이 이루어지지 않는 경우를 보인다. (11c)의 발화문은 (11b)와는 무관하게 발화된 것이다. 예문 (12)는 예문 (10)에 상반되는 경우로 내레이션에 의해 단절된 발화문들 사이에 턴테이킹이 이루어지지 않는 경우이다. (12d)를 보면 동화의 시점과 장면이 변화되었음을 알 수 있으며, 이와 같은 문장에 의해 두 발화문에서 턴테이킹이 이루어지지 않음이 더욱 분명해진다. 표 7은 분석 데이터 셋의 전체 발화문이 위와 같은 네 가지 유형 중 각각 어디에 속하는지 파악하여 각 유형별 발화문의 개수를 나타낸 표이다.

| 유형 | 개수 | 비율 |
|-------------------|-----|-------|
| 연속발화 & 턴테이킹 | 76 | 43.7% |
| 연속발화 아님 & 턴테이킹 | 33 | 19.0% |
| 연속발화 & 턴테이킹 아님 | 3 | 1.7% |
| 연속발화 아님 & 턴테이킹 아님 | 62 | 35.6% |
| 총합 | 174 | 100% |

[표 7] 턴테이킹의 유형별 발화문 개수

본 논문에서는 연속된 발화문에 존재하는 턴테이킹만을 화자 파악에 활용하고자 한다. 이는 턴테이킹 정보가 매우 중요하여 화자 파악에 큰 기여를 하는 경우가 모두 연속된 발화문에서 나타났기 때문이다. 반면, 비연속적인 발화문의 경우, 내레이션 상에 해당 발화문의 화자를 파악할 수 있는 단서가 존재하는 경우가 많기 때문에 턴테이킹 여부에 크게 의존하지 않아도 된다. 본 논문에서는 또한 연속된 발화문인 경우, 모두 턴테이킹이 일어난다고 간주하고 턴테이킹 정보를 화자 파악에 활용하였다. 그 이유는 연속된 발화문이지만 턴테이킹이 아닌 구조는 약 2%로 그 비중이 매우 낮으며(표 7 참고), 턴테이킹이 아님을 밝혀내기 위한 연구, 즉 화자 파악뿐만 아니라 해당 동화와 관련된 배경지식의 활용 및 담화 분할(discourse segmentation)과 같은 연구가 추가로 진행되어야 하는데, 이는 본 논문의 범위를 벗어나기 때문이다.

3.7.1 발화문 내 호격 및 명사 정보를 활용한 다음 발화문의 화자 예측. Sacks *et al.* (1974)은 대화 시 참여자들의 턴테이킹 행동에 대하여 연구하여 다음과 같은 규칙(turn-taking rules)이 턴테이킹 행동을 지배한다고 하였다.

- (13) Turn-taking rules (Sacks *et al.*, 1974)
 - a. 현재 화자가 이번 턴에서 A를 다음 화자로 지목했다면, A는 다음에 반드시 발화해야 한다.
 - b. 현재 화자가 다음 화자를 지목하지 않았다면 누구든 다음 대화의 턴을 가질 수 있다.

- c. 아무도 다음 대화의 턴을 가지지 않았다면, 현재 화자가 그것을 가질 가능성이 크다.

규칙 (13a)에 따르면, 현 발화문에서 화자가 다음 화자를 지목한 경우, 다음 발화문의 화자는 지목된 화자이어야 한다. 예를 들어, 다음의 예문에서 발화문 (14c)의 화자는 발화문 (14b)의 화자 ‘제비’가 다음 화자로 지목한 ‘엄지공주’이다.

- (14) a. 추운 겨울이 가고, 따뜻한 봄이 되자 제비는 다시 날 수 있게 되었어요.
 b. “엄지공주님, 고마웠어요. 제가 꽃나라로 데려다 드릴게요.” (화자: 제비)
 c. “안돼요. 함께 가고 싶지만 저는 두더지의 색시가 되어야 한답니다.” (화자: 엄지공주)
 d. 제비는 혼자서 꽃나라로 떠났어요.

(14b)에서와 같이 현재 화자가 다음 화자를 지목하기 위해 활용하는 표현은 주로 상대방의 이름이나 2인칭 대명사 ‘너’, ‘당신’ 등이 포함되는 표현이며 해당 화자를 부르는 경우가 대부분이다. 즉, 턴테이킹이 이루어지는 발화문 연속쌍에서 현 발화문에 호격 정보가 포함되어 있으면 그것을 다음 발화문의 화자 파악에 활용할 수 있다.

데이터 분석 결과, 호격은 동화 텍스트에서 주로 다음과 같은 패턴으로 문장에 나타났으며, 본 연구에서는 이러한 패턴을 기반으로 호격 정보를 추출하여 화자 파악의 단서로 활용하였다.

명사(구) + 호격조사 (아/야) + 쉼표(,)
명사(구) + 쉼표(,)
 문장 + 쉼표(,) + 명사(구) + {문장부호 (./?!)}
명사(구) + {문장부호 (!)}

기존 연구(Zhang *et al.*, 2003)에서는 발화문 내에 나타나는 멘션은 해당 문장의 화자 멘션이 될 수 없다는 제약을 두었는데, 본 논문에서도 호격에 위치한 명사를 제외하고 발화문 내의 다른 명사들은 화자 멘션이 될 수 없다고 가정하였다. 예를 들어, 발화문 (14b)에 나타난 ‘꽃나라’ 및 (14c)에 나타난 ‘두더지’나 ‘색시’는 해당 발화문의 화자가 될 수 없음을 토대로 가능한 화자 후보의 수를 줄였다.

3.7.2 대화참여자 쌍을 활용한 연속발화문의 화자 예측. 앞 절에서는 호격 정보와 같이 다음 화자로 지목된 등장인물을 가리키는 단서를 활용하여 화자를 예측하였으나 발화문 상에 이 같은 단서가 존재하지 않는 경우도 있을 수 있다. 그러나 예문 (15)(예문 6과 동일)와 같이 연속된 발화문이 시작되기 이전에 나타난 내레이션 문장이 턴테이킹 행위를 하는 대화참여자 쌍을 암시하는 경우가 있는데, 이 대화참여자 쌍을 화자 파악에 활용할 수 있다.

- (15) a. 잭은 길에서 이상한 할아버지를 만났어요.
 b. “꼬마야, 젓소를 이 콩과 바꾸지 않겠니?” (화자: 할아버지)
 c. “안돼요. 젓소를 팔아서 어머니 약을 사야 해요.” (화자: 잭)
 d. “이 콩은 요술콩이야. 이 콩을 심으면 너의 집이 행복해 질 거란다.” (화자: 할아버지)

발화문 (15b), (15d)의 화자는 ‘할아버지’이고, 발화문 (15c)의 화자는 ‘잭’이다. 발화문 (15c)가 연속된 발화문 내에 존재하므로 화자 파악을 위한 단서들이 비교적 멀리 떨어지게 되어 예측의 확실성이 떨어질 수 있다. 그러나 내레이션 문장 (15a)로부터 턴테이킹의 참여자가 ‘잭’과 ‘할아버지’임을 알 수 있으므로 이를 활용하여 각 발화문의 화자 파악이 가능하다.

대화참여자 파악은 다음과 같은 과정을 거쳐 파악된다. 첫 번째 단계로 해당 내레이션 문장에 대화에 참여하는 행위를 가리키는 서술어가 존재하는지 파악한다. 본 연구에서는 이를 위해 세종사전의 용어의 의미부류 체계 가운데 소통행위(예: 인사하다), 대칭적행위(예: 만나다), 방향성행위(예: 맞이하하다), 접근(예: 다가가다)의 의미부류를 활용하여 동사사전을 구축하였다. 두 번째 단계에서는 세종사전으로부터 서술어가 취하는 명사의 격 정보를 추출하여 해당 내레이션 문장에 이 격 정보를 가진 명사가 존재하는지 파악하고, 존재한다면 이를 추출한다. 예를 들어, (15a)의 동사 ‘만나다’가 취하는 명사의 격 정보는 그림 1에서 보는 바와 같이 <frame> 태그에 있는 frame 정보 ‘X=N0-이 Y=N1-을 V’에서 논항정보 X와 Y를 통하여 파악할 수 있다.

4. 화자 파악 단서 분류 및 화자 파악 규칙

4.1 화자 파악 과정

화자 파악은 다음과 같은 세 단계에 의해 이루어진다. 먼저 전처리기에서 형태소 분석 및 품사 태깅을 거친 후, 원 문장과 형태소 분석된 결과를 바탕으로 내레이션 문장과 발화문을 파악한다(발화문 파악). 두 번째 단계로 각 발화문에 대하여 화자가 될 수 있는 후보 멘션들을 발화문을 중심으로 추출한다(화자 후보 파악). 마지막 단계로, 각 화자 멘션 후보에 대해 화자 파악 단서를 추출하고, 이 단서들을 기반으로 화자 파악 규칙을 적용하여 화자 멘션 후보들 가운데 하나의 화자 멘션을 선택한다(화자 파악). 두 번째 단계인 화자 후보 파악에서 너무 많은 후보가 선정되면 화자 파악의 정확도가 떨어질 가능성이 높다. 하지만 말을 할 수 있는 ‘사람(인간)’을 가리키는 멘션만을 후보로 선정하는 것은 서론에서 논의한 바와 같이, 동물, 무생물도 화자가 될 수 있는 동화 도메인에 적합하지 않는 방법이다. 따라서 본 논문에서는 형태소 분석 정보를 활용하여 명사(구)를 화자 멘션 후보로 삼고 표 8의 화자 후보 제거에 속하는 단서를 활용하여 후보를 줄이는 방법을 사용하였다. 이와는 반대로 3절에 논의한 단서들 가운데, 해당 후보가 화자일 가능성을 높여주는 역할을 하는 단서(표 8 화자 가능성 상승 단서 참조)도 존재하는데, 제거되고 남은 후보들 가운데


```

- <sense n="01">
- <sem_grp>
  <sem_class>대칭적행위</sem_class>
  <trans>see</trans>
</sem_grp>
- <frame_grp type="FTR">
  <frame>X=N0-이 Y=N1-을 V</frame>
  - <subsense>
    <sel_rst arg="X" tht="AGT">인간</sel_rst>
    <sel_rst arg="Y" tht="THM">인간</sel_rst>
    <eg>오늘 중로에서 12시에 영화를 만나기로 돼 있다.</eg>
    <eg>어제 길에서 우연히 철수를 만났다.</eg>
    <eg>어쩐지 그 사람은 만나기 싫어.</eg>
    <eg>영화가 싫다는 것을 억지로 만났지.</eg>
    <eg>영화가 미정미하고 함께 극장에서 막 나오는 것을 만났어.</eg>
  </subsense>
</frame_grp>
- <frame_grp type="FIN" cor="sym" n="01">
  <frame>X=N0-이 Y=N1-와 (서로) V</frame>
  - <subsense>
    <sel_rst arg="X" tht="AGT">인간</sel_rst>
    <sel_rst arg="Y" tht="COM">인간</sel_rst>
    <eg>철수는 영화와 10년 만에 서로 만났다.</eg>
    <eg>왜 이렇게 서로 만나기가 힘들니?</eg>
    <eg>나중에 그 카페에서 만나자.</eg>
  </subsense>
</frame_grp>
- <frame_grp type="FIN" cor="sym" n="01">
  <frame>X+Y=N0-이 (서로) V</frame>
  - <subsense>
    <sel_rst arg="X" tht="AGT">인간</sel_rst>
    <sel_rst arg="Y" tht="AGT">인간</sel_rst>
    <eg>철수와 영화는 10년 만에 서로 만났다.</eg>
    <eg>왜 이렇게 서로 만나기가 힘들니?</eg>
    <eg>나중에 그 카페에서 만나자.</eg>
  </subsense>
</frame_grp>
</sense>

```

[Figure 1] 세종사전 내의 동사 ‘만나다’ 항목의 프레임 구조 일부

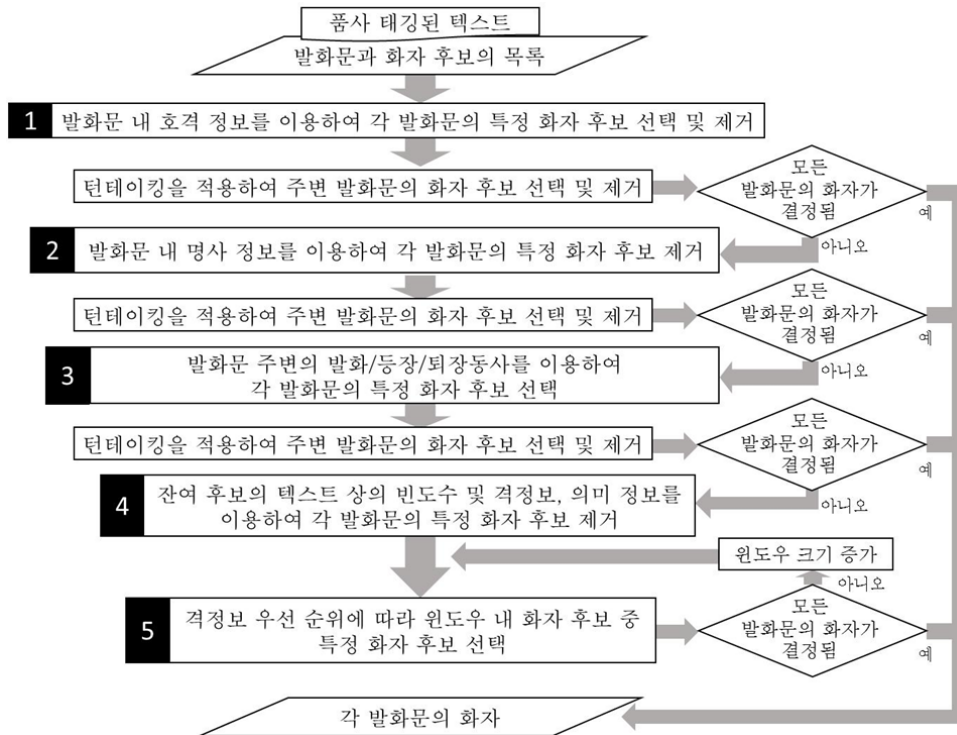
이 단서들을 활용하여 해당 발화문의 화자 멘션을 선택한다.

| 용도 | 단서 | 비고 |
|-----------|-------------|---------------------|
| 화자 후보 제거 | 상대적 위치 | |
| | 호격 정보 | 해당 발화문의 화자 배제 |
| | 퇴장동사 | |
| | 의미적 분류 | 등장인물 사전에 미존재 |
| | 화자 후보 빈도 | 빈도가 1 이하 |
| 화자 가능성 상승 | 호격 정보 | 주변 발화문의 화자 예측 |
| | 화자 멘션의 격 정보 | 주격, 주제격 > 목적격 > 소유격 |
| | 발화동사 | |
| | 등장동사 | |
| | 대화참여자 쌍 | |

[표 8] 화자 파악 단서 분류

4.2 화자 파악 규칙

화자 파악을 위하여 3절에서 제안한 단서들을 그림 2에 제시되어 있는 규칙에 따라 적용하여 화자의 후보를 줄여 나간다. 화자 파악 규칙 적용에 앞서 동화 테스트가 입력으로 주어지면 먼저 전처리기에서 형태소 분석 및 품사 태깅을 거친 후, 원 문장과 형태소 분석된 결과가 화자 파악 모듈의 입력으로 주어지게 된다. 본 연구에서는 KMA (Korean Morphological Analyzer)²를 활용하여 형태소 분석을 진행하였다.



[Figure 2] 화자 파악 규칙

앞서 제시한 단서 중 3.7.1절에 설명한 발화문 내 호격 및 명사 정보(그림 2의 1, 2번 단서)와 3.4절과 3.5절에서 설명한 발화문 주변 발화/등장/퇴장 동사 정보(그림 2의 3번 단서)는 화자를 암시하는 중요한 단서라고 판단하여, 이 단서들에 의해 특정 발화문의 화자가 결정될 경우에는 바로 이어서 해당 발화문에 대해 3.7절에서 설명한 턴테이킹을 적용하여 연속 발화문의 화자 후보를 함께 결정한다. 단, 3.4절에서 제시한 발화동사의 경우, 발화의 바로 이전 문장의 화자뿐만 아니라 바로 다음 문장의 화자도 암시할 수 있다는 애매성이 있으므로 이를 해소하는 과정을 병행한다. (16)에 제시된 예문(예문 4와 동일)은 이러한 과정이 필요한 경우를 보여 준다. (16d)에 언급된 발화동사 ‘말하다’의 주체는 그것의 주어인 ‘아버지’인데, 이전 문장 (16c)와 다음 문장 (16e)가 모두 발화문이므로, 해당 문장의 표면적인 단서만 이용해서는 ‘아버지’가 둘 중 어떤 발화문의 화자인지 결정하기가 어렵다. 그런데

² 강승식, 한국어 형태소 분석기와 한국어 분석 모듈, <http://nlp.kookmin.ac.kr/HAM/kor/index.html>

또 다른 발화동사가 포함된 (16b)의 경우, 그것의 이전 문장인 (16a)가 발화문이 아니므로 (16b)를 다음 발화문인 (16c)의 단서로 활용할 수 있다는 것을 알 수 있다. 즉, (16b)에 언급된 발화동사의 주체 ‘구름’이 (16c)의 화자라는 것을 결정하는 데 추가의 애매성은 없다. 그리고 (16c)의 화자가 결정되었으므로 그것이 ‘아버지’가 아니라는 사실로 미루어 (16d)의 ‘아버지’는 (16c)가 아닌 (16e)의 화자가 된다는 것을 알 수 있다. 이와 같이 애매성이 없는 경우에 대해서 먼저 발화동사 단서를 사용하여 화자를 파악하고, 이를 통해 다른 경우의 애매성을 순차적으로 해소한다.

- (16) a. 쥐 아버지는 곧 딸을 데리고 구름에게 부탁하러 갔어요.
 b. 구름에게 자기 딸을 신부로 맞아 주라고 하자, 구름이 말했어요.
 c. “아냐, 아냐. 나보다 훨씬 더 힘센 분이 계시는 걸.” (화자: 구름)
 d. 그 말을 듣고 아버지는 말했어요.
 e. “그 분이 누구십니까? 이름을 좀 가르쳐 주십시오.” (화자: 아버지)

그림 2의 1, 2, 3번 단서를 적용한 이후에도 화자가 결정되지 않은 발화문에 대해서는 텍스트 상의 어휘 빈도수 및 3.3절에서 설명한 격 정보를 이용하여 화자가 될 가능성이 적은 후보들을 먼저 제외시키고(그림 2의 4번), 3.2절에서 설명한 상대적 위치 단서를 바탕으로 해당 발화문의 특정 윈도우 내에 남은 후보들을 탐색하여 상대적 위치와 격 정보가 우선 순위에 있는 것을 선택한다(그림 2의 5번). 이때 활용하는 격 정보의 우선 순위는 주제격, 주격, 목적격, 소유격 순이다. 동일한 격 정보를 가지는 여러 후보가 발견될 경우 3.6절에서 제시한 명사의 의미 부류에 속하는 것을 선택한다. 해당 윈도우에서 후보가 발견되지 않을 경우 윈도우 크기를 늘려나가면서 다시 탐색한다. 만약 최종적으로 2개 이상의 후보가 남을 경우 그 중 하나를 임의로 선택한다.

그림 2의 화자 파악 규칙은 각 단서의 화자 파악에 대한 기여도와 단서 적용 순서 변화에 따른 성능 변화를 관찰하여 고안한 것이다. 예를 들어, 발화문 내 호격 정보(그림 2의 1번)와 명사 정보(그림 2의 2번) 단서의 경우 이를 먼저 적용하여 특정 화자 후보를 선택 및 제거해야 다음 단계에서 실제 화자가 아닌 멘션이 화자로 잘못 선택되는 경우를 최소화할 수 있었다. 이 두 단서의 경우 화자 후보를 선택 및 제거하는 과정에서 서로 영향을 미치지 않았고, 순서를 서로 바꾸어도 성능 상에 큰 변화가 없었다. 이는 발화문 내에서 호격으로 쓰이는 단어는 항상 명사이므로 호격 정보를 이용하여 제거되는 후보는 명사 정보를 이용하여도 동일하게 제거되기 때문인 것으로 보인다. 발화문 주변의 발화/등장/퇴장 동사 단서(그림 2의 3번)는 호격 및 명사 정보에 비해 올바른 화자를 결정하는 정확률은 낮으나 빈도수, 격 정보, 의미 정보 단서(그림 2의 4번)와 윈도우 내 격 정보 우선 순위 단서(그림 2의 5번)보다는 우선하여 배치해야 성능 저하가 최소화되는 것으로 확인되었다. 윈도우 내 격 정보 우선 순위 단서는 다른 단서에 비해 화자를 판별하는 성능이 상대적으로 가장 낮은 것으로 확인되어 규칙 순서상 마지막에 배치하였다.

하나의 발화문에 대하여 여러 단서가 동시에 등장하였을 때, 그림 2의 단서 우선 순위에 따라 화자가 결정되는 예는 아래와 같다.

- (17) a. 추운 겨울이 가고, 따뜻한 봄이 되자 제비는 다시 날 수 있게 되었어요.
 b. “엄지공주님, 고마웠어요. 제가 꽃나라로 데려다 드릴게요.” (화자: 제비)
 c. “안돼요. 함께 가고 싶지만 저는 두더지의 색시가 되어야 한답니다.” (화자: 엄지공주)
 d. 제비는 혼자서 꽃나라로 떠났다요.

(17)을 살펴보면 발화문 (17c)의 화자 멘션을 결정하는 데 있어 두 가지 단서가 존재하는 것을 알 수 있는데, 하나는 이전 발화문인 (17b)의 호격 ‘엄지공주’이고, 다른 하나는 다음 내레이션인 (17d)의 퇴장동사 ‘떠나다’이다. 그림 2에 제시된 단서 우선 순위에 따라 호격 정보를 먼저 적용하면 턴테이킹에 의해 (17c)의 화자 멘션은 ‘엄지공주’로 올바르게 선택되지만, 만약 호격 정보보다 퇴장 동사 단서를 먼저 적용하면 (17c)는 화자 멘션은 바로 다음 내레이션의 퇴장동사의 주체인 ‘제비’로 잘못 선택된다. 또한 만약 우선 순위가 가장 낮은 윈도우 내 격 정보(그림 2의 5번)를 가장 먼저 적용할 경우, (17b)의 화자 멘션으로 그것과 거리가 가장 가까우면서 주격에 위치하여 상대적 위치(-1) 및 격 정보 우선 순위가 가장 높은 (17a)의 ‘겨울’이 화자 멘션으로 잘못 선택될 것이다. (‘겨울’은 실제로 텍스트 내 빈도수 및 의미 정보 단서(그림 2의 4번)를 이용하여 후보에서 제거될 수 있는 어휘이다.) 이와 같이 데이터 분석 과정에서 단서 적용 순서에 따라 화자 파악의 성능이 달라지는 것을 관찰하였으며, 그림 2의 규칙은 실험에 사용한 동화 데이터에 대해 성능 저하가 최소화되도록 고안된 것이다.

5. 실험 및 평가

5.1 실험 준비

본 연구에서 제안한 방법의 검증을 위해서 3절에 제시한 단서들과 4절에 제시한 규칙을 이용하여 화자 파악 시스템을 구현하였고, 시스템이 동화 텍스트에 주어진 각각의 발화문에 대해 1개의 화자 멘션을 선택하게 하였으며, 각 발화문의 실제 화자 멘션과 시스템이 선택한 화자 멘션을 비교하여 정확도(accuracy)를 측정하였다. 정확도 계산 공식은 다음과 같다.³

$$\text{정확도} = \frac{\text{실제 화자와 시스템이 출력한 화자가 일치한 경우의 수}}{\text{동화 내 발화문의 총 개수}}$$

³ 본 연구에서 다루는 화자 파악의 경우, 텍스트 내에서 발화문을 찾는 것이 목적이 아니라 주어진 각각의 발화문에 대한 화자 멘션의 여러 후보 중 올바른 것을 선택하는 것이 목적이므로, 정확률(precision)과 재현율(recall)의 계산 결과가 항상 일치한다. 그러므로 본 연구에서는 이 두 용어 대신에 정확도(accuracy)라는 용어로 통일하여 사용한다.

그런데 실제 화자라고 주석 처리된 멘션과 시스템이 선택한 화자 멘션이 서로 형태만 다르면서 실제로 동일한 대상을 지칭(corefer)하는 경우가 있을 수 있으므로, 이를 정답으로 간주한 경우와 오답으로 간주한 경우를 구분하여 정확도를 따로 측정하였다. 이는 실제로 시스템 상에서 참조 현상 해소(coreference resolution) 과정을 적용하여 해결해야 하나, 본 연구에서는 이를 논외로 한다.

본 연구에서는 다음과 같은 어려움으로 인하여 기존 관련 연구의 방법론과의 성능 비교를 하지 않았다. 첫째, 규칙을 기반으로 한 기존 연구들의 경우, 해당 언어(예: 스페인어(Mamede and Chaleira, 2004), 영어(Glass and Bangay, 2007)) 사용의 특징을 이용하여 화자 멘션을 파악하는 규칙들이 상당수 있어 이를 본 연구에 직접 적용하기가 어려웠다. 둘째, 본 연구에서는 기존 연구에서 고려하지 않은 턴테이킹과 같은 언어학적 단서를 분석하는 목적으로 동화 데이터를 수집, 분석하였기 때문에 데이터의 규모가 작은 편이라 기계학습을 기반으로 한 기존 연구의 방법론을 적용하기에는 무리가 있었다.

5.2 평가 결과

각각의 동화 텍스트에 대해 총 문장 개수, 발화문 개수, 시스템의 출력이 정답과 일치한 경우의 수, 정확도를 정리한 결과는 표 9에 제시된 것과 같다(참조 인정 정확도 순으로 정렬).

6. 토의

표 9의 평가 결과에 따르면 동화마다 성능 상의 편차가 존재한다. 실험 결과를 분석한 결과, 낮은 성능을 보인 동화에서는 3절에서 제시한 단서들이 표면적으로 드러나지 않은 경우가 많았다. 이와 같이 표면적인 단서의 존재 여부는 동화의 줄거리나 유형, 장르적 특성보다는 작가의 문체에 따라 달라지는 것으로 보인다. 가령 ‘쥐의 신랑감’에서는 모든 발화문 주변 내레이션에 발화/등장/퇴장 동사가 사용되어 높은 성능이 나타난 반면, ‘은혜 깊은 까치’에는 이러한 동사가 단 한번도 사용되지 않아 낮은 성능이 나타났다. 이와 같이 작가가 내레이션 상에 명시적인 단서를 드러내지 않은 동화의 경우 화자를 유추하기 위한 의미/화용 수준의 분석이 필요하나, 이는 향후 연구로 남긴다. 구체적으로 오류의 유형을 분석한 결과, 이를 표 10과 같이 크게 세 가지로 분류할 수 있었다. 각각의 오류에 대한 구체적인 설명과 예문은 아래에 제시한 것과 같다.

단서 추출 오류

본 연구에서는 호격 정보를 추출하기 위하여 3.7.1 절에서 제시된 패턴을 활용하였는데, 이를 이용하였을 때 ‘나는 소리개야.’와 같은 예문에서 ‘소리개’가 호격인 것으로 잘못 분석되었다. 호격 정보 추출에 대한 재현율(recall)을 유지하기 위해서는 문미에 나오는 호격 정보도 중요하게 고려되어야 하는데, 문미에 나타나는 후보가 호격으로 사용되지 않은 경우가 있어 정확률(precision)이 낮아지게 된다. 정확률을

| 동화 제목 | 문장 총 개수 | 발화문 총 개수 | 일치 개수 (참조 불인정) | 일치 개수 (참조 인정) | 정확도 (참조 불인정) | 정확도 (참조 인정) |
|-------------|------------|-------------|-------------------|------------------|-----------------|----------------|
| 쥐의 신랑감 | 52 | 19 | 18 | 19 | 0.95 | 1.00 |
| 효성 다한 호랑이 | 62 | 13 | 9 | 13 | 0.69 | 1.00 |
| 농부와 난쟁이들 | 32 | 8 | 3 | 7 | 0.38 | 0.88 |
| 엄지공주 | 64 | 17 | 14 | 15 | 0.82 | 0.88 |
| 알라딘의 램프 | 81 | 21 | 17 | 18 | 0.81 | 0.86 |
| 헨젤과 그레텔 | 63 | 20 | 14 | 17 | 0.70 | 0.85 |
| 콩쥐 팥쥐 | 46 | 13 | 9 | 11 | 0.69 | 0.85 |
| 책과 콩나무 | 69 | 18 | 13 | 15 | 0.72 | 0.83 |
| 벌거벗은 임금님 | 64 | 24 | 20 | 20 | 0.83 | 0.83 |
| 아기 돼지 삼형제 | 61 | 21 | 13 | 17 | 0.62 | 0.81 |
| 그루터기의 새싹 | 46 | 14 | 11 | 11 | 0.79 | 0.79 |
| 빨간 모자 | 35 | 15 | 9 | 11 | 0.60 | 0.73 |
| 피노키오 | 68 | 18 | 13 | 13 | 0.72 | 0.72 |
| 무덤가에 핀 꽃 | 49 | 12 | 7 | 8 | 0.58 | 0.67 |
| 금도끼 은도끼 | 32 | 8 | 5 | 5 | 0.63 | 0.63 |
| 개구리 왕자 | 64 | 26 | 15 | 16 | 0.58 | 0.62 |
| 빨간 부채 파란 부채 | 77 | 21 | 12 | 13 | 0.57 | 0.62 |
| 은혜 값은 가치 | 79 | 20 | 9 | 12 | 0.45 | 0.60 |
| 타조의 세가지 보물 | 44 | 19 | 6 | 6 | 0.32 | 0.32 |
| 계 | 1088 | 327 | 217 | 247 | 0.66 | 0.76 |

[표 9] 화자 파악 시스템 평가 결과

| 오류 유형 | 발화문 개수 | 비율 |
|---------------|--------|------|
| 단서 암시 및 단서 부족 | 65 | 59% |
| 참조 표현 | 30 | 27% |
| 단서 추출 오류 | 15 | 14% |
| 계 | 110 | 100% |

[표 10] 화자 파악의 오류 유형

보장하기 위해서는 구문 단위의 정교한 분석이 이루어져야 하는데, 이는 추후 연구로 남겨둔다. 또한 형태소 분석 과정에서 발생한 오류로 인해 단서가 올바르게 추출되지 않아 오류로 이어진 경우도 이 분류에 포함된다.

참조 표현

주어진 발화문의 화자 후보를 선택하는 과정에서 서로 다른 멘션이 실제로 하나의 대상을 가리키고 있다는 정보를 파악하지 못하여 의도하지 않은 후보가 화자로 선택

되는 경우가 있었다. 가령 아래 예문 (18) (예문 3과 동일)에서 (18b)의 경우 실제 화자 멘션은 바로 다음 문장에 있는 ‘두꺼비’였으나, 시스템은 이를 ‘엄지공주’로 잘못 선택하였다. 만약 ‘아가씨’와 ‘엄지공주’가 실제로 동일한 대상을 가리키고 있다는 사실을 알 수 있다면 3.7.1 절에서 제시한 발화문 내 명사 정보 단서에 의해 (18b)의 화자 후보에서 ‘엄지공주’를 제외하여 또 다른 후보인 ‘두꺼비’를 화자로 선택할 수 있을 것이다.

- (18) a. 엄지공주는 호두껍질로 만든 침대에서 자고 있었어요.
 b. “하하하, 예쁜 아가씨군.” (화자: 두꺼비)
 c. 무섭게 생긴 두꺼비가 몰래 들어와 엄지공주를 연못으로 데려갔어요.

참조 표현을 파악하지 못하여 발생한 오류는 턴테이킹에 의해서 다른 발화문으로 전파되어 의도하지 않은 오류를 발생시키기도 한다. 가령 아래 예문 (19) (예문 6, 15와 동일)에서는 (19b)의 ‘꼬마’가 ‘잭’과 동일한 대상을 가리킨다는 정보가 없어 바로 이전 문장인 (19a)의 주제격 명사인 ‘잭’이 (19b)의 화자로 잘못 파악되었으며, 이는 턴테이킹을 통해 다음 발화문인 (19c)와 (19d)의 화자까지 잘못 선택되는 결과를 초래하였다. 그러므로 호격을 화자의 단서로 활용하는 과정에서 참조 표현을 올바르게 파악하는 것은 특히 중요하다고 볼 수 있다.

- (19) a. 잭은 길에서 이상한 할아버지를 만났어요.
 b. “꼬마야, 젓소를 이 콩과 바꾸지 않겠니?” (화자: 할아버지)
 c. “안 돼요. 젓소를 팔아서 어머니 약을 사야 해요.” (화자: 잭)
 d. “이 콩은 요술콩이야. 이 콩을 심으면 너의 집이 행복해 질 거란다.” (화자: 할아버지)
 e. 잭은 젓소를 콩 한 주머니와 바꾸어 버렸어요.

또한 시스템이 화자를 올바르게 선택하였으나 그것의 멘션이 주석자가 표시한 화자 멘션과 달라서 정답으로 인정되지 않은 경우가 있었다. 가령 아래 예문의 경우 시스템은 (20c)의 화자를 (20d)에 있는 ‘램프요정’으로 선택하였으나, 정답 데이터에는 (20a)의 ‘요정’이 실제 화자로 주석 처리되어 있어 정답으로 인정되지 않았다.

- (20) a. 마법사는 램프를 문질러 요정을 불러냈어요.
 b. “이제 내가 너의 주인이니 내 말을 들어야 한다. 당장 알라딘의 궁전을 아프리카로 옮겨라.” (화자: 마법사)
 c. “예, 주인님.” (화자: 램프요정)
 d. 램프요정은 알라딘의 궁전을 번쩍 들고 날아갔어요.

발화문 내 화자가 자기 자신을 지칭하는 표현

3.7.1 절에서 호격을 제외하고 발화문 내에 드러난 명사는 해당 발화문의 화자가 될

수 없다고 가정하였는데, 이에 대한 예외로 발화자가 스스로 자기 자신을 지칭하는 표현이 있었다. 이는 실제로 발화문 내 명사 정보 단서와는 구별되어야 하는 또 다른 중요한 단서이나, 본 연구에서는 이를 따로 처리하지 않아서 오류가 발생하였다. 예를 들어 발화문 (21b)는 화자인 소리개가 자기 자신을 상대방에서 소개하는 문장으로, ‘나는 X다’라는 문형에 해당한다. 이 때 X에 위치하는 명사 ‘소리개’를 화자 멘션으로 선택하면 해당 발화문의 화자를 쉽게 선택할 수 있다. 그러나 본 시스템에서는 X를 단순히 발화문 내 명사 정보로 간주하여 화자 멘션이 될 수 없다고 가정함으로써 다른 화자 멘션을 선택하는 오류가 발생하였다. 또 다른 예로 발화문 (22b)에서는 ‘X 살려’라는 관용표현이 쓰이고 있는데, 이 경우도 마찬가지로 해당 표현이 자기 자신을 지칭하는 표현임을 파악하여 X에 위치하는 명사 ‘거인’을 해당 발화문의 화자 멘션으로 선택해야 했으나, 시스템은 이를 처리하지 못해 오류를 발생시켰다. 이와 같이 자기 자신을 소개하거나 지칭하는 표현은 동화에서 다양한 형태로 나타날 수 있는데, 이를 수집하고 화자 파악의 단서로 활용하는 것은 추후 연구로 남긴다.

- (21) a. 소리개는 재빨리 날아가서는 반갑게 인사를 했어요.
 b. “안녕? 나는 소리개야. 그런데 넌 누구니?” (화자: 소리개)
 c. “소리개야, 이렇게 크고 멋진 나를 처음 본다고?”
- (22) a. 짝은 땅에 뛰어내리자마자, 도끼로 콩나무를 찍어 쓰러트렸어요.
 b. “오아, 거인 살려” (화자: 거인)
 c. 거인은 아래로 곤두박질 쳤어요.

단서 추출 오류와 참조 현상으로 인해 발생한 오류 이외에 텍스트 상에서 단서가 표면적으로 드러나 있지 않아 발생한 오류는 아래에 나열된 것과 같다.

복수명사가 지칭하는 화자들 간의 턴테이킹 및 동일 화자의 연속 발화

(23)과 같이 한 명의 화자가 연속 발화하는 경우, 본 연구에서 제안한 규칙으로 올바른 화자를 파악하기에는 한계가 있다. 현재 시스템에서는 (23a)의 호격 정보를 이용하여 (23b)의 화자가 ‘콩쥐’라는 사실은 알 수 있으나, 이후 바로 턴테이킹을 적용하여 다음 발화문인 (23c)의 화자 후보에서 ‘콩쥐’를 제외시키기 때문에 (23c)의 화자를 파악하는 과정에서 필연적으로 오류가 발생한다. 실제로 이야기의 내용 흐름상 (23b)와 (23c) 사이에는 시간적 또는 공간적 간격이 존재하므로 이 두 발화를 턴테이킹으로 보기는 어려우나, 이를 시스템이 표면적 단서만을 이용하여 파악하는 것은 쉽지 않았다. 그러나 3.7절에서 논의한 것과 같이 본 연구에서 사용한 데이터에서는 이와 같은 한 화자의 연속 발화가 드물게 나타났기 때문에, 이러한 경우를 제대로 처리하지 못한 것이 전체 성능에 크게 영향을 미치지 않았다.

- (23) a. “애, 콩쥐야. 넌 독에 물을 가득 채우고, 베를 다 짜거든 잔치에 오너라, 알겠니?”
 b. “네, 어머니.” (화자: 콩쥐)

- c. “이 많은 일을 다 어찌한담.” (화자: 콩쥐)
- d. 콩쥐는 독에 물을 붓기 시작했어요.

또한 연속 발화문의 화자가 동일한 멘션으로 주석 처리되어 표면적으로는 턴테이킹이 아닌 것처럼 보이나, 실제로 내용상 다른 화자를 가리켜 오류가 발생한 경우가 있었다. 이러한 오류는 표면적으로 드러나 있지는 않으나 복수명사가 암시적으로 지칭하는 여러 화자가 턴테이킹에 참여할 때 주로 발생하였다. 가령 정답 데이터에는 (24b)와 (24c)의 화자에 해당하는 멘션이 모두 (24a)의 ‘아이들’로 표시되어 있었으나, 이 두 발화문의 화자는 실제로 서로 다른 대상일 가능성이 높기 때문에 턴테이킹이라고 볼 수 있다. 그러나 현재 시스템에서는 이러한 복수명사가 지칭하는 대상을 파악하지 않고 있기 때문에 (24c)의 화자로 ‘아이들’이 아닌 다른 화자를 선택하여 결국 오류가 발생하게 된다.

- (24) a. 그런데 도중에 마차에 타고 있는 아이들을 만났어요.
- b. “장난감 나라로 가는 마차야.” (화자: 아이들 일부 A)
- c. “공부도 하지 않고 놀고 먹는 곳이라. 우리랑 같이 가지 않을래?” (화자: 아이들 일부 B)
- d. 피노키오는 또 아이들을 따라갔어요.

이와 유사하게 ‘형제’, ‘가족’ 등의 집합 명사가 가리키는 여러 대상이 턴테이킹을 하는 경우에도 오류가 발생하는 것을 확인할 수 있었다. 가령 아래의 예문에서 (25b)와 (25c)의 화자는 각각 (25a)의 ‘형제’를 구성하는 ‘아우’과 ‘형’라고 볼 수 있는데, (25c)의 경우 (25b)의 호격 정보를 이용하여 (25c)의 화자가 ‘형’이라는 것을 알 수 있는데 반해, (25b)의 경우 실제 동화 텍스트에서 ‘아우’라는 명사가 단 한 번도 언급되지 않아 멘션 기반으로 화자를 찾는 현재 시스템에서 이에 대한 올바른 화자를 찾는 것은 불가능하였다. 이러한 오류를 처리하려면 ‘형제’라는 집합 명사가 ‘형’과 ‘아우’로 구성되어 있음을 추론할 수 있어야 한다.

- (25) a. 꾀가 많고 거짓말을 잘하는 형제였지요.
- b. “형님, 빨리 가요.” (화자: 아우)
- c. “그래, 어서 가자.” (화자: 형)
- d. 형제는 성문 앞에서 크게 외쳤어요.

이야기의 내용 및 흐름에 의해 암시되는 화자

아래의 예문에서 발화문 (26c)의 실제 화자는 ‘선비’이나, 그것의 바로 이전 문장과 다음 문장에 ‘선비’에 대한 언급이 없고, 각각 주격과 주제격 명사로 ‘여인’이 두 번 언급되기 때문에 시스템은 (26c)의 화자를 ‘여인’이라고 판단하였다. 이와 같이 표면적 단서가 부족한 상황에서 (26c)의 주변 문장에 있는 ‘여인’을 화자 후보에서 배제하고 좀 더 멀리 위치한 (26a)의 ‘선비’를 화자로 선택하게 하려면, ‘선비’가 산을

돌아다니다 밤이 깊어 거처를 찾고 있다는 일련의 서사적 정보를 추적하는 방법과 이를 해당 발화문의 표면적으로 드러난 정보와 연결하는 좀 더 깊은 수준의 처리 과정이 필요하다.

- (26) a. 아무도 없으리라 생각되었지만 혹시나 한 젊은 선비는 문을 두드렸습니다.
 b. 그런데 뜻밖에도 안에서 하얀 옷을 입은 여인이 나오는 것이었습니다.
 c. “날이 저물어 하룻밤 쉬어 갈까 합니다.” (화자: 선비)
 d. 여인은 고개를 끄덕이더니 젊은 선비를 안으로 들여 보내 주었습니다.

위의 경우 이외에도 동화 내용의 흐름을 추적하고 있어야만 올바른 화자를 파악할 수 있는 경우는 아래의 예에서도 확인할 수 있다. 가령 (27b)에서는 ‘할머니’가 호격으로 파악되어 다음 발화문인 (27c)의 화자 또한 ‘할머니’로 잘못 선택되었다. 하지만 이야기의 흐름상 ‘빨간모자’는 ‘할머니’를 부른 것이 아니라 실제로 ‘늑대’를 부른 것이므로 (27c)의 화자 멘션을 ‘할머니’로 선택한 것은 오류에 해당한다. 이를 해결하려면 이야기 중간에 늑대가 할머니로 변장하는 과정에서 ‘늑대’ = ‘할머니’라는 정보를 추론하여 활용할 수 있어야 한다.

- (27) a. 얼마 후, 빨간모자가 문을 열고 들어왔어요.
 b. “할머니, 빨간모자 왔어요.” (화자: 빨간모자)
 c. “어서 온.” (화자: 늑대)
 d. 늑대는 벌떡 일어나 빨간모자를 한숨에 꿀꺽 삼켜버렸어요.

7. 결론

본 연구에서는 발화문 상의 감정 파악 및 등장인물별 적절한 TTS 음성의 할당에 활용 가능한 발화문 상의 화자 파악 문제를 해결하기 위한 언어학적 단서에 대하여 논의하고, 이를 기반으로 화자 파악 규칙을 제안하였다. 분석 결과, 기존 연구에서 주로 활용된 자질인 후보의 위치, 후보의 격 정보, 발화동사, 등장/퇴장 동사의 존재 여부 및 등장인물의 의미적 분류뿐만 아니라 본 논문에서 새롭게 제안한 턴테이킹도 화자 선택의 애매성을 해소하는 데 매우 유용한 단서임을 알 수 있었다. 다만 턴테이킹에 참여하는 화자가 한 번 잘못 파악되었을 경우에는 다른 화자까지 연쇄적으로 잘못 파악되는 오류가 발생할 수 있기 때문에 초기 후보 선택에 유의해야 한다. 특히 호칭과 같은 참조표현은 오류 유형 중에 큰 비중을 차지하므로 화자 파악 시스템의 안정적인 성능을 위해 관련 문제의 해결이 선행되어야 할 것이다. 그 외 표면적인 단서가 드러나지 않아 암시적으로만 화자를 파악할 수 있는 경우 본 연구에서 제안하는 방식으로만 접근하기에는 한계가 있다.

향후에는 발화문 상에서 화자 자신을 지칭하는 표현 및 가족 간의 호칭 정보가 발달한 한국어에서의 특성을 반영한 등장인물의 관계 정보를 활용한 알고리즘의 개발과 오류 분석을 통하여 화자 파악 성능을 보다 개선시키고자 한다. 또한 대규모의

동화 데이터를 수집하고 본 연구에서 발견한 단서를 기반으로 한 기계학습 모델을 화자 파악 문제에 적용하여 보다 강인한 시스템으로 발전시킬 계획이다.

참고문헌

- 민혜진 · 김상채 · 박종철. 2012. 로봇 동화 구연을 위한 동화 상 발화문의 화자 자동 파악, 제 24회 한글 및 한국어 정보처리 학술대회 논문집, 77-84, 한국정보과학회 & 한국인지과학회.
- 이성현. 2005. 전자사전 구축과 의미부류 - 세종 명사 의미부류 체계의 예, 한국사전학 5, 103-138.
- 이호준 · 박종철. 2011. 동화에서 나타나는 감정 발화문의 특성분석을 통한 감정 범주의 평가. HCI 학술대회 발표논문집, 475-477, 한국 HCI 학회.
- Elson, David. K. and Kathleen R. McKeown. 2010. Automatic attribution of quoted speech in literary narrative. In Maria Fox and David Poole (eds.), *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10)*, 1013-1019.
- Fong, Terrence, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems* 42, 143-166.
- Fridin, Marina. 2014. Storytelling by a kindergarten social assistive robot: A tool for constructive learning in preschool education. *Computers & Education* 70, f53-64.
- Galss, Kevin and Shanun Bangay. 2007. A naive, salience-based method for speaker identification in fiction books. In Jules-Raymond Tapamo and Fred Nicolls (eds.), *Proceedings of the 18th International Symposium of the Pattern Recognition Association of South Africa*, 1-6.
- Gelin, Rodolphe, Christophe d'Alessandro, Quoc Anh Le, Olivier Deroo, David Doukhan, Jean-Claude Martin, Catherine Pelachaud, Albert Rilliard, and Sophie Rosset. 2010. Towards a Storytelling Humanoid Robot. Paper presented for the AAAI 2010 Fall Symposia *Dialog with Robots*, Arlington, VA, November.
- Goodrich, Michael A. and Alan C. Schultz. 2007. Human-Robot Interaction: a survey. *Foundations and Trends in Human-Computer Interaction* 1, 203-275.
- Lee, Ho-Joon and Jong C. Park. 2009. Interpretation of user evaluation for emotional speech synthesis system. In Julie A. Jacko *et al.* (eds.), *Proceedings of the 13th International Conference on Human-Computer Interaction (HCI International 2009)*, 295-303.

- Mamede, Nuno and Pedro Chaleira. 2004. Character identification in children stories. In José Luis Vicedo, Patricio Martínez-Barco, Rafael Muñoz, and Maximiliano Saiz Noeda (eds.), *Advances in Natural Language Processing* (=LNCS 3230), 82-90.
- Min, Hye-Jin, Sang-Chae Kim, Joonyeob Kim, Jin-Woo Chung, and Jong C. Park. 2013. Speaker-TTS voice mapping towards natural and characteristic robot storytelling. In Hyouk Ryeol Choi (ed.), *Proceedings of the 22nd IEEE International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN)*, 793-800.
- Montaño, Raúl, Francesc Alías, and Joseph Ferrer. 2013. Prosodic analysis of storytelling discourse modes and narrative situations oriented to Text-to-Speech synthesis. In Antonio Bonafonte (ed.), *Proceedings of the 8th ISCA Speech Synthesis Workshop*, 171-175.
- Mutlu, Bilge, Jodi Forlizzi, and Jessica Hodgins. 2006. A storytelling robot: modeling and evaluation of human-like gaze behavior. In Giulio Sandini and Aude Billard (eds.), *Proceedings of the 6th IEEE-RAS International Conference on Humanoid Robots*, 518-523.
- O'Keefe, Tim, Silvia Pareti, James R. Curran, Irena Koprinska, and Matthew Honnibal. 2012. Sequence labelling approach to quote attribution. In Jun'ichi Tsujii, James Henderson, and Marius Pasca (eds.), *Proceedings of the 2012 Conference on Empirical Methods on Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012)*, 790-799.
- Plaisant, Catherine, Allson Druin, Corinna Lathan, Kapil Dakhane, Kris Edwards, Jack Maxwell Vice, and Jaime Montemayor. 2000. A storytelling robot for pediatric rehabilitation. In Marilyn Tremaine, Elliot Cole, and Elizabeth Mynatt (eds.), *Proceedings of the fourth international ACM conference on Assistive Technologies*, 50-55.
- Ray, Céline., Francesco Mondada, and Roland Siegwart. 2008. What do people expect from robots?, In Raja Chatila, Jean-Pierre Merlet, and Christian Laugier (eds.), *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008)*, 3816-3821.
- Ribeiro, C. R. and M.F.M. Costa, and C. Pereira-Coutinho. 2009. Robotics in child storytelling. In Manoj Patariya (ed.), *Proceedings of the 6th International Conference on Hands-on Science*, 198-205.
- Sacks, Harvey. 1992. *Lectures on Conversation*. Cambridge, MA: Blackwell.
- Zhang, Jason Y., Alan W. Black, and Richard Sproat. 2003. Identifying speakers in

children's stories for speech synthesis. In Herve Bourlard (ed.), *Proceedings of the 8th European Conference on Speech Communication and Technology (EUROSPEECH)*, 2041-2044.

접수 일자: 2013년 11월 15일

수정 일자: 2013년 12월 6일

게재 결정: 2013년 12월 16일