

영어 학습자를 위한 데이터 기반 맞춤형 단어 추천

김서영⁰¹ 정석훈^{†1}·최슬기^{†2} 이주훈¹ 김주호¹

¹ 한국과학기술원 전산학부

² 한동대학교 전산전자공학부

youthskim@kaist.ac.kr, sraccoon@kaist.ac.kr, 21700729@handong.edu,

juhoonlee@kaist.ac.kr, juhokim@kaist.ac.kr

Data-Driven and Personalized Vocabulary Recommendation for English Learners

Seoyoung Kim⁰¹ Seokhun Jeong^{†1}·Seulgi Choi^{†2} Juhoon Lee¹ Juho Kim¹

¹School of Computing, Korea Advanced Institute of Science and Technology

²School of Computer Science and Electrical Engineering, Handong University

요 약

영어 말하기 실력 증진을 위해서는 말하기 시 자유롭게 쓸 수 있는 영어 단어, 즉 산출적 어휘의 수를 늘리는 것이 필수적이다. 하지만 뜻을 아는지 여부만으로 해당 단어가 산출적 어휘 지식에 포함되는지 판단할 수 있는 것이 아니기에 학습자의 현재 산출적 어휘 지식을 파악하기 힘들다. 더 나아가 학습자는 어떠한 단어를 자신의 산출적 어휘 지식에 추가하면 좋을지 판단하기 힘들다. 따라서 본 논문에서는 학습자의 이전 영어 회화 데이터를 분석하여 학습자의 산출적 어휘 지식을 추정하고, 학습자의 산출적 어휘 지식에서 부족한 맞춤형 영어단어를 추천하는 시스템을 제안한다. 해당 시스템을 평가하기 위해 유저스터디를 진행한 결과, 본 시스템이 추천한 단어의 만족도가 기존의 산출적 어휘 지식 시험을 통한 단어 추천 방식보다 더 높았다.

1. 서 론

어휘지식의 크기는 제2외국어를 학습함에 있어 중요한 척도 중 하나로 여겨진다. 따라서, 어휘지식을 확장하는 것을 돕기 위한 많은 선행 연구들이 존재한다. 이러한 어휘지식은 외국어를 읽고 들을 때 필요한 수용적 어휘지식과 외국어를 쓰거나 말할 때 필요한 산출적 어휘지식으로 나뉜다. 영어 말하기 실력 증진을 위해서는 산출적 어휘지식, 그 중에서도 능동적으로 단어를 떠올려 사용할 수 있는 자유 산출적 어휘지식 향상이 필수적이다. 하지만, 시험을 위해 단어를 주는 순간 학습자가 능동적으로 해당 단어를 쓸 수 있는지를 알 수 없기에 학습자의 자유 산출적 어휘지식을 측정하기 어렵고 학습자 또한 수용적 어휘지식 위주의 영어 학습법에 익숙하여 본인의 산출적 어휘 지식과 거기서 부족한 단어가 무엇인지 잘 알지 못한다. 때문에 기존의 연구는 통제된 상황에서 어휘 지식을 측정하거나, 산출적 어휘지식을 몇 가지 단계로 나누어 대략

적으로 측정하는 데에 그친다.

본 연구에서는 자유 산출적 어휘지식을 보다 정확하게 측정하기 위해 학습자의 이전 영어 회화 데이터를 분석한다. 자유 산출적 어휘 지식을 측정함으로써 학습자에게 부족한 어휘를 알 수 있고, 학습자가 사용할 가능성이 큰 어휘를 알 수 있다는 점에 착안하여 학습자의 산출적 어휘지식을 확장시키기 위한 단어 추천 기법을 제시한다. 학습자 별 맞춤형 단어를 추천하기 위해 학습자의 반복적인 표현을 대체할 수 있는 다른 표현을 추천해주는 대체 추천과, 대화 주제에 관련된 다양한 표현을 추천해주는 확장 추천의 두 가지 방법을 제시하고, 해당 기술을 이용한 영어단어 추천 시스템의 효능을 유저스터디를 통해 검증하였다.

2. 방 법

본 연구에서는 산출적 어휘지식을 향상시키기 위해 학습자의 이전 영어 회화 데이터를 분석하고, 이를 바탕으로 맞춤형 단어를 추천하는 두가지의 추천 기법을 제안한다. 첫번째 추천 기법은 대체(Substitution)방식의 추천으로, 동의어 학습을 통해 학습자의 반복적인 표현을 대체함으로써 같은 의미를 다양한 단어로 표현하도록 돕는다. 이는 산출적 어휘 지식이 제한적인 외국어 학습자들이 동일한 단어를 반복적으로 사용하는 문제를 개선시키고자 고안되었다. 두번째 기법은 확장(Expansion)방식의 추천으로, 다양한 단어를 활용하여 특정 주제에 대해 자유롭게 의사를 표현할 수 있도록 대화 주제와

* This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government (MSIT) (No.2020-0-02237, Personalized Progress Analysis and Exercise Recommendation for Remote Language Learning Using AI and Big Data) (No.2021-0-01347, Video Interaction Technologies Using Object-Oriented Video Modeling)

† These authors contributed equally.

관련된 단어를 추천한다. 이 방식은 제2외국어를 배우는 대부분의 학습자가 특정 주제에 대한 다양한 뜻과 표현을 사용하는 것에 어려움을 겪기 때문에 의미적 표현의 확장을 도와주되, 학습자의 대화 내용과 관련된 단어를 추천해주고자 고안되었다.

2.1 산출적 어휘지식의 확장을 위한 맞춤형 단어 추천 시스템

2.1.1 대체(Substitution) 추천 파이프라인

학습자가 반복적으로 사용하는 단어를 적절한 단어로 추천하는 대체(Substitution)추천 파이프라인은 총 6단계로 구성된다(그림 1).



그림 1. 대체(Substitution) 추천방식의 파이프라인 구조

사용자의 대화 데이터를 분석하기에 앞서 어휘 대체 모델이 음성 인식 결과로 나온 학습자의 대화문을 더욱 잘 인식할 수 있도록 BERT를 기반으로 한 유창성 검사 모델(Disfluency Detection)[2]을 통해 화자가 잘못 발화한 내용을 교정해주었다. 그 다음, 사용된 단어들의 빈도수를 기준으로 자주 사용한 단어들을 추출한다. 그 후 화상영어 플랫폼을 통한 과거 일대일 수업들에서 학습자의 산출적 어휘 지식, 즉 학습자가 사용한 단어를 추출하고, 현재 세션의 대화 데이터를 기반으로 산출적 어휘 지식의 범위를 업데이트한다. 추출한 단어 리스트 중 해당 사용자의 이전 추천 데이터에 해당되는 단어들은 제외시킨 후, 어휘 대체(Lexical Substitution) 모델을 사용하여 대체가능한 동의어들을 찾는다[3]. 어휘 대체는 주어진 문장에서 문장의 문맥을 유지한 채 특정 단어를 대체할 수 있는 단어를 생성하는 것을 의미한다[3]. 본 논문에서는 빈칸채우기(Fill-in-the-blank) 모델과 같이 문맥만을 고려한 단어 생성이 아닌 문맥과 목표 단어를 동시에 고려한 대체어를 얻기 위해 빈칸에 알맞은 모든 단어를 추측하는 빈칸채우기(Fill-in-the-blank) 모델이 아닌 어휘 대체(Lexical Substitution) 모델을 사용하였고[3], 그 중에도 성능이 가장 높았던 XLNet을 기반으로 한 어휘 대체 모델을 사용했다[1]. 이 모델의 경우, 대화 문맥 데이터와 대체하고자 하는 목표 단어의 위치를 인풋으로 넣으면 대체 가능한 후보 단어들의 확률을 리스트 형태로 얻게 된다[1]. 따라서 대체하고자 하는 각 단어마다 후보 단어들을 10개씩 구한 뒤, 대체가능 확률의 우선순위를 고려하여 추천단어를 추출하였다. 추출된 단어는 크게 산출적 어휘지식의 포함여부, 단어의 난이도, 문맥과 목표를 동시에 고려했는지의 여부를 고려하여 필터링 된다.

2.1.2 확장(Expansion) 추천 파이프라인

대화 주제에 대한 의미적 확장을 돕는 확장(Expansion) 추천 파이프라인은 총 5단계로 구성된다(그림2).



그림2. 확장(Expansion) 추천방식의 파이프라인 구조

대체(Substitution) 추천 파이프라인과 같이 음성인식 결과를 후처리 한 뒤, 자연어처리 기술 중 하나인 대화 주제 추출(Topic Extraction)을 해주는 MonkeyLearn API를 활용하여 대화 기록 데이터에서 하나의 주제를 추출한다. 다음으로 대화 도중 사용된 단어들을 사용 빈도 순으로 나열한 후 앞서 찾은 주제와 관련있는 단어, 즉 대화의 ‘키워드’를 추린다. 이때 벡터 공간 모델, 즉 각 단어를 벡터로 표현하는 대수적 모델을 사용하여 단어간의 유사도 또는 관련성을 파악하였다. 추린 키워드마다 입력한 단어에 대해 가장 유사한 단어들을 추출한다. 추출된 유사어들 중, 학습자의 어휘 수준과 비슷한 난이도의 단어들을 골라 학습자 별로 제시한다. 표 1을 통해서 단어 추천 시스템의 예시를 확인할 수 있다.

표 1. 단어 추천 시스템의 예시

대체(Substitution) 추천 시스템 결과 예시			
대화 데이터에서 해당 단어가 사용된 문장		I <i>totally</i> agree with her and it is <i>totally</i> correct because I am ...	
학습자가 자주 반복한 단어	totally	추천된 단어	utterly
확장(Expansion) 추천 시스템 결과 예시			
대화 데이터의 주제		food	
대화 내 키워드	eat	추천된 단어	consume

2.2 추천 웹 시스템

웹 시스템은 참가자의 Ringle세션 정보를 입력받아 대화 스크립트를 가져오고, 산출적 어휘 지식 레벨을 측정하기 위한 단어 시험인 PVLТ 시험을 거친 후 스크립트를 분석해 추천 단어 리스트(그림3)를 보여준다.

3. 유저 스테디

본 연구에서 개발한 단어 추천 시스템의 추천 성능을 평가하기 위해, 영어가 제2외국어인 학습자를 대상으로 실험을 진행했다. 적절한 추천을 제공하기 위해 해당 학습자의 이전 영어 말하기 기록이 필요하기 때문에, 원어민 튜터와 1:1 화상통화 기반의 영어 학습 서비스인 Ringle의 이용자 중에서 16명을 선정해 Ringle의 1:1 화상 통화 세션에서 추출된 대본을 이용했다.

실험은 실험군 6명, 대조군 10명으로 나누어 진행했다. 실험군은 본 연구에서 개발한 단어 추천 방식을 이용하여, 대체 추천 방식과 확장 추천 방식으로 각 15개의 영어 단어를 추천하여 제공하였다. 대조군은 학습자의 이전 영어 대화 스크립트를 사용하지 않고 기존의 산출적 어휘지식을 평가하는 시험인 PVLТ와 단어 리스트인 BNC(British National Corpus) & COCA(Corpus of Contemporary American English)를 사용

했다. PVLTS시험으로 학습자의 단어 레벨을 평가한 뒤, 해당되는 레벨의 단어리스트 중 무작위로 30개의 단어를 선택하여 대조군 유저들에게 제공하였다. 단어 추천을 받은 후, 참가자들은 추천된 단어에 대한 생각을 물어보는 7-리커트 스케일로 구성된 설문에 응답하였다.



그림3. 단어 추천 결과 웹 페이지

4. 결과 및 분석

유저스터디 결과, 만족도(실험군 m=4.2, 대조군 m=3.9), 실용성(실험군 m=4.6, 대조군 m=4.2), 영어 실력 향상에 도움이 될지 여부(실험군 m=4.8, 대조군 m=4.4), 개인 맞춤형(실험군 m=3.6, 대조군 m=3.0), 추천 이유 파악 가능성(실험군 m=4.8, 대조군 m=2.7) 측면에서 본 연구에서 개발한 단어 추천 방식이 대조군보다 유용하다는 결과가 나왔다. 또한 30개의 단어 중 마음에 드는 추천을 원하는 만큼 고르도록 한 결과 실험군이 만족스러운 추천 비율이 더 높았다 (실험군 m=45%, 대조군 m=36%).

두 컨디션 모두 추천된 단어가 대부분 수용적 어휘 지식에 있는 단어라고 응답하였지만, 실험군의 경우 수용적 어휘 지식에 없는 단어가 상대적으로 더 많았다 (실험군 m=2.4, 대조군 m=1.8). 난이도의 적절성은 대조군과 실험군 모두 약간 쉽다는 쪽으로 결과가 나왔다 (실험군 m=3.2, 대조군 m=3.6). 반면 대조군에서는 산출적 어휘지식의 반영이 부족했던 반면 실험군에서는 과거 대화 기록을 이용했기 때문에 산출적 어휘 지식이 반영된 추천이 이루어졌다: “내가 자주 활용하는 어휘임에도 추천되어 있는 어휘가 존재하였다.” (P3, 대조군), “내가 익숙하지만 평소애 말할 때 잘 사용하지 않는 단어들을 추천해주셔서 좋았고...” (P12, 실험군).

하지만, 만족도가 대조군과 실험군 모두 다소 낮게 나타났다. 몇몇 참여자는 설문에서 직접적으로 그 이유로 추천된 단어의 수준이 낮아서라고 언급하였다. 이를 통해 학습자들이 수용적 어휘지식 확장을 위한 단어 추천에만 익숙하여 산출적 어휘지식의 수준을 실제 본인의 수준보다 높게 판단하는 경향이 있음을 알 수 있다. 특히, 대조군의 경우 산출적 어휘지식 측정을 위한 기존의 시험으로 레벨을 측정해 단어가 추천되었음에도 추천 수준이 낮다는 평가가 있었다. “어휘 수준 평가가 적절하지 못한 것 같고, 추천한 단어의 수준이 내가 인지한 수준보다 낮았다.” (P3, 대조군), “생각보다 생각해내기 쉬운 수준의 단어들이었다.” (P13, 실험군). 따라서 장기적으로 본 시스템을 사용하면 본인의 산출적 어휘 지식을 객관적으로 판단하는 데 유용할 수 있다. 마지막으로 만족스러운 추천 비율은 대체 추천 방식이 확장 추천 방식보다 높게 나왔다 (대체 m=60%, 확장 m=30%). 이는 기존의 단어 추천 방식이 동의어 추천에 더 초점이 맞춰져 있거나 동의어 사전이 존재하기에 학습자가 이러한 추천 방식에 더 익숙할 수 있기 때문으로 추정된다.

5. 결론 및 고찰

본 논문은 산출적 단어지식의 확장을 위한 맞춤형 단어 추천 시스템을 제시하였다. 유저스터디 결과, 대조군보다 본 연구에서 제안한 방식으로 추천된 단어에 대한 만족도가 더 높았다. 유저스터디 결과 본인의 산출적 어휘 지식을 실제 본인의 수준보다 높게 평가하는 추세가 나타났는데, 이처럼 본인의 산출적 어휘 지식 수준을 객관적으로 인지하지 못하는 경우 산출적 어휘 지식 확장 시 방해가 될 수 있다. 따라서, 추후 본인의 산출적 어휘 지식을 객관적으로 파악할 수 있도록 학습자의 산출적 어휘 지식을 시각화하는 연구를 진행할 예정이다. 또한, 향후 연구에서는 영어단어 추천에서 더 나아가 추천된 단어들이 학습자의 산출적 어휘 지식에 추가될 수 있도록 실제 영어 회화 상황에서 추천된 단어를 사용하도록 하는 학습시스템을 연구할 예정이다.

6. 참고문헌

[1] Arefyev, N., Sheludko, B., Podolskiy, A., & Panchenko, A. (2020). Always Keep your Target in Mind: Studying Semantics and Improving Performance of Neural Lexical Substitution. *Proceedings of the 28th International Conference on Computational Linguistics*.

[2] Jamshid Lou, P., & Johnson, M. (2020). Improving Disfluency Detection by Self-Training a Self-Attentive Model. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*.

[3] Zhou, W., Ge, T., Xu, K., Wei, F., & Zhou, M. (2019). BERT-based Lexical Substitution. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*.