

## CHILDES 코퍼스를 기반으로 한 아동의 영어 굴절형태소 발달 연구\*

민 명 숙<sup>†</sup>            전 종 섭<sup>†</sup>            이 선 영<sup>‡</sup>  
한국외국어대학교            사이버한국외국어대학교

본 연구의 목적은 선행 연구에서 보고된 영어 모국어 아동의 굴절형태소 습득 과정을 대규모 언어습득 데이터베이스를 활용하여 검증하는 것이다. 이를 위해, 우리는 CHILDES(Child Language Data Exchange System)[1] 데이터베이스에 등장하는 1-7세 영국 및 미국 아동 1,630명이 발화한 470만 어절 말뭉치를 대상으로 굴절형태소의 발달 과정을 분석하였다. 본 논문에서는 동사의 현재분사 *-ing*, 과거형 *-(e)d*, 형용사의 비교/최상급 *-er/est* 등의 형태소에 대해 어

---

\* 이 논문은 2012년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2012S1A5A2A03034786).

† 제1저자: 민명숙, 한국외국어대학교 대학원 국어국문학과, (130-791) 서울시 동대문구 이문로 107

E-mail: michelleclick@empal.com

‡ 제2저자: 전종섭, 한국외국어대학교 언어인지과학과, (449-791) 경기도 용인시 모현면 외대로 81

E-mail: jongsupjun@korea.com

‡ 교신저자: 이선영, 사이버한국외국어대학교 영어학부, (130-791) 서울시 동대문구 이문로 107, 연구 분야: 언어습득, 언어처리, 심리언어학, 통사론

E-mail: alohasylee@cufs.ac.kr

휘 유형(Type)과 사례(Token) 빈도, 전체 사례(Token)에 대한 유형(Type) 비율인 TTR(Type per Token Ratio), 어휘 다양성 척도인 Lexical Diversity(D) 값을 구하여 이를 국가 및 연령별로 비교, 분석하였다. 그 결과, 굴절형태소별로 연령과 D 값의 상관관계가 다르게 나타났다. 특히, 현재분사 *-ing*와 D 값 사이에는 주목할 만한 상관관계가 나타나지 않은 반면, 과거형 *-(e)d*의 경우 양의 상관관계 경향성이 보였고, 비교/최상급 *-er/-est*는 유의미한 상관관계를 보였다. 이는 현재진행형이 과거형보다 먼저 습득된다고 보고한 Brown(1973)의 견해를 지지한다. 다음으로, 과잉일반화에 따른 오류 표현이 2-3세 사이에 많이 나타나면서 U자형 발달 양상을 보였다. 과잉일반화도 현재분사보다 과거형에서 많이 나타났는데, 이것 또한 현재분사가 과거형보다 일찍 습득된다는 주장을 지지한다. 영국과 미국 아동의 연령별 굴절형태소 사용 양상을 비교한 결과, 미국 아동의 D 값이 영국 아동보다 높았다. 이는 미국 아동이 영국 아동보다 더 많은 어휘 유형에 대해 굴절형태소를 사용했음을 의미한다. 본 연구는 소수의 아동을 대상으로 수행된 선행 연구의 다양한 논점을 대규모 데이터베이스로 검증하고, CHILDES 코퍼스를 효율적으로 분석하는 연구 방법론을 제안했다는 점에서 의의가 있다.

주제어 : 언어습득, CHILDES, 굴절형태소, 과잉일반화

## 서 론

아동의 언어 발달을 측정하는 척도 중의 하나로 생산성(productivity)이 있다. 여기서 생산성이란 아동이 주변의 다양한 입력을 통해 습득한 언어 지식을 규칙화하는 문법적 생산성을 의미한다. 문법적 생산성은 영어에서 현재의 진행 상태를 나타내는 *-ing*, 과거시제 형태소 *-(e)d*, 현재시제 3인칭 단수 형태소 *-(e)s* 등과 같은 굴절형태소(Inflectional morpheme) 습득에 의해 측정될 수 있다. 영어 모국어 아동의 형태소 습득에 대한 대표적인 연구로 Brown(1973)을 들 수 있는데, Brown은 그의 연구에서 영어를 모국어로 배우는 아동 Eve, Adam, Sarah를 대상으로 통시적 연구를 진행하였고, 그 결과로 총 14개 문법적 형태소 간의 습득 순서를 제안하였다. 3명의 아동이 가장 먼저 습득한 굴절형태소는 동사의 현재분사형(예: *going*)이었으며 그 뒤를 이어 명사 복수형(예: *apples*), 불규칙 동사 과거형(예: *wore*), 소유격(예: *Daddy's*), 규칙 동사의 과거형(예: *walked*), 3인칭 단수 현재형(예: *sleeps*)을 습득하였다. 아동의 굴절형태소 습득 순서와 더불어 아동의 언어 발달 단계에서 보이는 과잉일반화(Overgeneralization) 경향은 언어 습득 연구 분야에서 많은 연구자들의 관심을 받았다. 과잉일반화란 아동이 부모와 같은 주변인으로부터 입력 받은 언어적 또는 문법적 양상을 의미 전달 시 적극 적용하는 창조적 과정에서 발생하는 문법적 오류를 의미한다. 이러한 과잉일반화 오류는 동사, 형용사, 명사에 걸쳐 불규칙 형태가 존재하는 영어의 습득 연구에서 적지 않게 확인되었다[2, 3]. 예를 들면, 아동이 언어 학습 초기에는 불규칙 동사의 과거형을 올바르게 사용하다가(예: *went*) 2-3세가 되면 규칙 동사를 과거형으로 만드는 형태소 *-(e)d*를 불규칙 동사의 기본형 또는 과거형에 붙여 사용하고(예: *\*goed, \*wentad*) 다시 4-5세가 되면 정확한 형태의 과거형을 발화하는 것이다(예: *went*). 이러한 U자형 곡선의 언어 발달 양상이 영어의 불규칙 동사 과거형 습득 과정에서 나타난다는 것이 여러 연구[4, 5]에서 확인되었다. 이렇게 영어 모국어 아동의 굴절형태소 습득과 관련된 발견들, 즉 굴절형태소 간 습득 순서와, 과잉일반화 현상 및 그에 따른 U자형 발달 양상 등은 그동안 많은 습득 연구에서 받아들여 온 것이다. 그런데, 이와 관련된 기존의 많은 연구들은 한정된 수의 아동이 발화한 데이터에 기반을 둔 것이다. Brown(1973)의 연구도 3명의 아동을 기반으로 하였다. 따라서 이러한 선행 연구 결과로 아동의 언어 발달

양상을 일반화하기에는 부족함이 있어, 더 많은 수의 아동을 대상으로 한 검증 필요성은 항상 존재해 왔다고 하겠다. 이는 또한, 언어 입력이 아동의 언어 발달에 끼치는 영향에 대해 검증할 필요성이 있음을 의미하기도 한다.

본 연구의 목적은 CHILDES[1] (<http://childes.psy.cmu.edu/>)라는 대규모 코퍼스에 포함된 1-7세 아동 전체를 대상으로 기존의 제한된 수의 아동에 바탕을 둔 기존 연구 결과를 재검증하는 데 있다. 코퍼스가 다양해지고 분석 도구가 발전하면서 대규모 코퍼스를 분석할 수 있는 기반이 마련되어 본 연구에서는 실제로 대규모 코퍼스를 재조직하여 사용할 수 있는 방법을 연구하고 이를 바탕으로 기존의 연구 결과들을 재검증하는 것이다. 특히 본 논문에서 다루는 것은 전체 CHILDES 데이터 기반의 아동 굴절형태소 발달 연구를 목적으로 하는 대규모 프로젝트의 초기 결과이다. 보다 자세한 분석에 앞서 전체적으로 그동안 연구되었던 굴절형태소의 발달 양상, 특히 영어 동사와 관련한 가장 대표적인 굴절형태소인 현재분사 *-ing*과 과거형 *-(e)d*, 그리고 형용사의 굴절과 관련된 비교급 *-er*과 최상급 *-est*에 대해 알아보았다. 먼저 각 형태소의 유형(Type)과 사례(Token)의 출현 빈도와 전체 사례에 대한 유형 빈도(Type per Token Ratio: TTR)를 국적(영국과 미국)과 연령에 따라 살펴보고, CHILDES 코퍼스에 있는 데이터의 길이가 서로 다르다는 것을 고려하여 어휘 다양성을 나타내는 D값을 측정하여 그동안 발견되었던 형태소 습득 순서 및 과잉일반화, 이와 관련된 아동의 U자형 언어 발달에 대해 재검증하였다. 또한, 미국과 영국 아동을 비교함으로써 굴절형태소 습득에 미치는 언어 환경의 영향을 알아보았다.

## 선행연구

### 형태소 발달 순서

영어를 모국어로 하는 아동의 영어 굴절형태소 습득 양상에 대한 많은 연구 중에 고전적인 연구로는 대표적 실험 연구인 Berko(1958)의 연구와 그 이후 코퍼스를 기반으로 한 Brown(1973)의 연구가 있다. Berko(1958)는 4-7세 사이의 미국 아동을

대상으로 WUG TEST<sup>1)</sup>라는 유도발화 과제를 사용하여 형태소 습득을 연구하였다. 이 연구를 통해서 Berko는 아동들이 형태소 사용에 있어서 규칙을 적용한다는 것을 보여주었다. 즉, 아동이 한 번도 들어보지 못한 동사의 과거형을 만들 때 *-(e)d*라는 형태소를 동사 원형에 붙이는 규칙을 적용한다는 것을 발견하였다(6). 또한 아동의 형태소 규칙 적용은 형태소마다 차이가 있음을 밝혔다. 연구에 포함된 현재분사(*-ing*), 규칙 동사 과거형(*-(e)d*), 3인칭 단수 현재(*-(e)s*), 소유격(*'s*)을 나타내는 형태소 중에서 현재분사가 가장 먼저 습득된다는 것이다. 다시 말해, 형태소 사이에 습득 순서가 존재한다는 것을 보여주었다. 이 실험 연구 결과를 바탕으로 Brown(1973)은 어린 아동을 대상으로 한 통시적 연구에서 아동의 자발적 발화 데이터를 사용하여 실제로 각각의 형태소가 언제 처음 출현하고 언제 습득되었는지에 대해 연구하였다. 그의 실험 연구는 3명의 아동 Eve, Adam, Sarah를 대상으로 하였는데 처음 연구를 시작할 당시 가장 연령이 낮은 Eve는 18개월이었고 Adam과 Sarah는 27개월이었다. 연령이 다른 세 아동의 발화를 평균 발화 길이(Mean Length of Utterance: MLU)를 기준으로 5단계 언어 발달 단계로 나눠 문법형태소 습득 순서를 확인하였다. 언어 발달 단계는 1단계는 MLU가 1.75 단어, 2단계는 2.25 단어, 3단계는 2.75단어, 4단계는 3.50 단어, 5단계는 4.00 단어를 기준으로 하였다. 3명의 아동이 연구에 참여했을 때는 2단계 수준의 언어를 사용하고 있었다. 특정 문법형태소가 필수적으로 나타나야 하는 문맥(Obligatory Context, 예) *'Yesterday I played...'*에서 90%이상 문법적으로 발화하면 습득한 것으로 보고 총 14개의 문법형태소 습득 순서를 제시하였다.

3명의 아동이 가장 먼저 습득한 형태소는 아래 (1)에서 보는 것과 같이 동사의 현재분사였으며, 그 뒤를 이어 전치사 *in*과 *on*, 명사 복수형, 불규칙 동사 과거형, 소유격, 축약불가능한 계사, 관사, 규칙 동사 과거, 3인칭 현재 단수형, 축약불가능한 조동사, 축약가능한 계사, 축약가능한 조동사 순으로 나타났다.

(1) Brown(1973)의 영어 모국어 아동의 습득 순서

1. 현재진행(present progressive)

1) 아동의 굴절형태소 발달 양상을 파악하기 위해 영어에 존재하지 않는 단어를 사용하여 목표형태소를 발화하게 하는 유도발화 실험방법이다.

- 2/3. 전치사(prepositions(in/on))
4. 복수형(plural)
5. 불규칙 과거형(irregular past tense)
6. 소유격(possessive)
7. 축약불가능한 계사(copula, uncontractible)
8. 관사(articles)
9. 규칙 과거형(regular past tense)
10. 3인칭 단수 규칙 현재형(third person present tense, regular)
11. 3인칭 단수 불규칙 현재형(third person present tense, irregular)
12. 축약불가능한 조동사(auxiliary, uncontractible)
13. 축약가능한 계사(copular, contractible)
14. 축약가능한 조동사(auxiliary, contractible)

Brown(1973)의 연구는 형태소 발달 연구에 대한 아동의 자발적 발화 자료를 바탕으로 한 고전적인 연구가 되어 이후 다양한 형태소 발달 연구에 인용되어 왔다. 그러나 실제 이 데이터는 단지 3명의 아동을 대상으로 한 것으로 여기서 발견한 결과를 얼마나 일반화할 수 있을지는 항상 의문이 되어 왔다고 할 수 있다. 이에 본 연구에서는 CHILDES 코퍼스를 대상으로 Brown에서 연구했던 형태소의 발달 양상을 다시 재검증하고자 하였다.

#### 과잉일반화

아동의 형태소 습득 연구와 관련하여 나타나는 큰 특징 중의 하나는 과잉일반화이다. 과잉일반화는 규칙을 적용할 수 있는 범위를 넘어서 적용하는 현상으로 예를 들면, 동사의 과거형을 만들 때 과거형태소 *-(e)d*를 규칙 동사가 아닌 불규칙 변형동사 *go*에 *-(e)d* 형태를 붙여(예: \**goed*) 결과적으로 비문법적 형태를 생성해내는 것을 의미한다. 이러한 현상은 아동이 단순히 부모의 발화를 그대로 모방하거나 입력 받은 그대로 발화하는 것이 아니라 실제로 아동의 머릿속에서 언어 입력을 바탕으로 규칙을 만들어 그것을 적용한다는 것을 보여주는 것으로 해석되었다. 이

는 아동의 언어 습득 이론에서 아동의 타고난 언어 습득 능력을 중시하는 생득주의적 접근에 의한 언어 습득 이론을 지지하는 현상으로 여겨졌으며, 부모 언어의 모방과 연습을 통한 아동의 언어 습득을 주장하는 행동주의적 언어 습득 이론에 대한 반증이 되는 것으로 해석되어 왔다. 이렇게 과잉일반화에 대한 연구는 두 가지 상반되는 습득 이론과 관련하여 많은 연구가 되어 왔다.

Marcus et al.(1992)은 83명 아동의 자발적 발화에서 나타난 11,521개의 불규칙 과거형을 분석하여 과잉일반화에 대해 살펴보았다. 이 연구에서 동사 과거형의 과잉일반화가 비록 높은 비율은 아니지만 실제로 아동의 자발적 발화에 존재함을 확인하였다. Marcus et al.(1992)은 특히 과거형 형태소의 과잉일반화 현상이 2세 이전에는 나타나지 않다가 2세부터 취학연령으로 가는 시기에 나타난다는 것을 확인하였는데, 이는 과잉일반화와 관련한 아동의 언어 발달 특징 중 하나인 U자형 발달 양상을 잘 보여주는 것이다. 이러한 과잉일반화 현상은 형태소에 따라 다르게 나타난다. 예를 들어, 현재분사는 (1)의 습득 순서에서 확인한 것과 같이 아동이 가장 빨리 습득하는 형태소이다. 이는 현재분사의 과잉일반화 예가 거의 발견되지 않아 발화 오류율이 그만큼 없음을 의미한다. 또한 Kuczaj(1977)도 15명 아동의 자발적 발화에서 현재분사와 관련한 과잉일반화 오류를 단 1회(예: *'we are not seeing her.'*) 발견하였다고 보고하였다. 이에 대해서 Kuczaj(1977)는 현재진행 형태소 *-ing*가 다른 동사나 명사의 굴절형태소와 달리 과잉일반화가 가장 적게 일어나는 것은 '불규칙' 진행 형태소가 따로 존재하지 않기 때문일 것이라 제안하였다. 이처럼 선행연구는 과잉일반화 현상이 형태소마다 달리 나타나는 것을 시사하고 있다. 본 연구에서는 대규모 CHILDES 데이터를 가지고 과잉일반화에 대한 U자형 발달 곡선과 형태소 간의 과잉일반화 경향 차이를 다시 검증하고자 하였다.

요약하자면, 영어 모국어 아동의 영어 굴절형태소 습득에 관련한 그동안의 연구는 대부분 소수 아동의 발화를 대상으로 한 것이어서 그 결과를 통해 아동의 언어 발달 양상을 일반화하기에는 부족함이 있다고 하겠다. 그리하여 본 연구에서는 선행 연구의 결과에서 밝힌 굴절형태소 발달 순서와 굴절형태소 발달에서 나타나는 과잉일반화 현상이 대규모의 CHILDES 데이터에서도 확인되는지 알아보았다. 또한 미국과 영국 아동의 데이터를 비교함으로써 같은 언어권이지만, 두 나라의 언어적 환경 차이가 굴절형태소 발달에 어떤 영향을 미치는지 알아보았다. 그리하여, 굴절

형태소 습득에서 환경의 영향 가능성에 대하여 기초자료를 제공하고자 하였다. 아직까지 형태소와 관련한 영국과 미국 아동의 발달 양상 비교는 거의 없었으므로 이를 비교하는 것은 앞으로의 언어 습득과 환경 영향에 대한 연구에 도움이 될 것이다.

이에, 본 연구에서 다루고자 하는 구체적인 연구 문제는 다음과 같다.

- i) 아동의 굴절형태소의 사용에 있어서 아동 연령이 증가하면서 형태소 사용이 다양한 동사에 적용되는가?
- ii) 아동의 굴절형태소의 사용에 있어서 과잉일반화가 확인되는가? 또한 과잉일반화에 따른 U자형 형태소 발달 양상이 보이는가?
- iii) 위의 두 질문과 관련하여 영국과 미국 아동의 발달 양상에 차이가 있는가? 있다면 그 차이는 엄마의 입력과 관련이 있는가?

본 연구에서는 아동의 연령에 따른 굴절형태소 발달을 측정하기 위해 각 형태소의 Type과 Token의 빈도를 구하고, TTR과 어휘 다양성을 나타내는 D값을 측정하였다. 전통적으로 어휘 다양성은 전체 Token에 대한 Type의 비율인 TTR로 측정되었는데[7] 이러한 TTR 측정의 문제점은 Token 수가 많으면 많을수록 TTR 수치가 낮아진다는 데 있다[8]. 연령이 증가할수록 Type과 Token도 증가하게 되는데, 연령별 TTR 차이를 비교해 보면 Type의 증가보다 Token의 증가가 훨씬 많아 결과적으로 TTR이 감소하게 된다. 이는 아동이 새로운 어휘를 배울 때 기존의 알고 있는 단어를 사용하여 문장을 길게 만들어내기 때문이다. 그리하여 TTR만으로 어휘 다양성을 판단할 경우 연령이 증가할수록 어휘 다양성이 낮아지는 결과를 낳게 된다. 아동의 언어 발달 측정에 있어, 이러한 TTR의 단점을 보완하기 위해서 텍스트의 길이에 영향을 받지 않는 측정 방법이 필요했으며, 여러 연구자[1, 9, 10]에 의해 Computerized Language ANalysis(CLAN) 프로그램의 어휘 다양성 측정 도구인 Vocd 명령이 개발되었다. Vocd 명령은 D값을 측정하는데 여기서 D는 다양성(Diversity)을 의미한다. D값은 데이터에서 무작위로 뽑은 자료를 대상으로 어휘 다양성을 측정한 값으로 D값이 높다는 것은 어휘 다양성이 높다는 것을 가리킨다. 이러한 D값 측정이 텍스트 길이의 영향권에서 완전하게 벗어난 것은 아니라는 연구 결과가 있



였지만[11] CLAN의 Vocd 명령을 통한 D값 산출은 어휘 다양성을 통해 아동의 언어 발달을 알아보는 여러 연구에서 많이 사용되는 값으로 가장 정확한 방법 중 하나이다[8]. 서로 다른 길이의 텍스트로 이루어진 대규모의 CHILDES 데이터에서 아동의 언어 발달을 측정하기 위해서는 D값을 통해 재검증하는 단계가 필요하다. 실제로 CHILDES 데이터에서 2-4세의 자료가 기타 연령대의 자료보다 훨씬 많아 D값 측정은 필수적이라 하겠다. D값의 상승은 아동 연령의 증가에 따라 각 굴절형태소의 적용 어휘가 점점 더 다양해진다는 것을 의미한다고 볼 수 있으며, 따라서 연령에 따른 D값 상승을 기대할 수 있다.

## 코퍼스 연구

### CHILDES 코퍼스

본 연구에서는 CHILDES 코퍼스의 데이터를 사용할 것이다. CHILDES 코퍼스는 주로 아동의 자발적 발화를 연구할 때 사용하는 자료로 언어습득 연구에서 가장 많이 사용되는 코퍼스 중의 하나이다. 이 코퍼스는 다양한 언어를 포함하고 있으나 대부분 영어 데이터로 이루어져 있으며 주로 영국과 미국 영어 자료이다. 주로 부모와 아동 사이에서 나누는 자발적인 대화 내용을 녹음한 것이며, 이를 CHAT Transcription and Cording Format(CHAT)으로 전사한 것이다. CHILDES 코퍼스는 아동의 언어 발달과 특히 아동이 부모로부터 받는 언어 입력 관련 연구에서도 많이 사용되어 왔다.

### CHILDES 코퍼스 정비

아동의 굴절형태소 사용 양상을 파악하기 위해 CHILDES 코퍼스 자료 중 1-7세에 해당하는 대규모 자료를 분석하기 쉽도록 재조직하였다. 우선 연구자의 이름으로 저장된 파일명을 연구자, 국적, 아동의 연령이 명시된 파일명으로 바꾸는 작업을 실시하여 파일명만으로도 그 파일을 구축한 연구자, 아동의 출신 국가 및 연령

을 파악할 수 있게 하였다. 또한 한 파일 안에 동일한 연령대의 여러 아동 발화가 코딩되어 있는 경우, 각각의 아동 발화를 CHI tier로 바꾸는 작업을 실시하였다. 그리고 한 파일 안에 다른 연령대의 아동들이 참여한 경우에도 마찬가지로 아동마다 파일을 만들어 파일명에 아동의 연령을 표시하였다. 각 아동의 연령 구분은 1년 단위로 하였는데, 예를 들어 1;1.1(년;월.일)이나 1;11.30은 모두 1세로 표시하였다. 이러한 작업을 통해, CLAN 프로그램에서 CHAT으로 전사된 CHILDES 코퍼스의 국적, 연령별 분석이 가능해졌다. 표 1을 보면 1-7세까지의 CHILDES 코퍼스는 영국의 경우 275명의 아동에게서 2,272의 파일이, 미국의 경우는 1,355명의 아동에게서 5,569개의 파일이 생성되어 전체 파일 규모는 총 7,841개에 달한다.

표 1. 1-7세 아동의 CHILDES 데이터 정보

구분	영국		미국	
	아동 수	파일 수	아동 수	파일 수
1세	62	203	203	960
2세	67	1,742	339	1,652
3세	69	194	302	1,124
4세	24	51	204	834
5세	28	57	134	665
6세	6	6	89	124
7세	19	19	84	210
합계	275	2,272	1,355	5,569

국적, 연령별 파일을 대상으로 CLAN에서 빈도 찾기 명령어인 FREQ 명령을 실행하여 아동이 발화한 어휘 Type과 Token을 확인한 결과, 영국 아동은 35,130 Type으로 1,937,624 Token을 발화한 것으로 나타났으며 미국 아동은 총 63,705 Type으로 2,771,312 Token을 발화한 것으로 나타났다. 또한 1-7세까지의 영국과 미국 파일을 통합하여 CLAN에서 FREQ 명령을 실행한 결과, 총 45,435개의 Type과 4,708,936개의 Token을 확인하였다. 본 연구는 약 470만 단어의 대규모 코퍼스

를 대상으로 하였다.

## 분석

CHILDES 데이터에서 영국과 미국 자료 중 1-7세로 분류된 파일에서 추출된 45,435개의 Type을 대상으로 EXCEL 프로그램을 사용하여 전형적인 굴절형태소를 기준으로 1차적으로 자동 분류를 실시하였다. 본 연구의 굴절형태소 유형은 Brown(1973)의 문법형태소 습득 순서상에 나타난 형태를 기준으로 하여 규칙 동사와 불규칙 동사 과거형(-ed), 3인칭 단수 현재형/복수형(-es) 동사 현재분사형(-ing), 비교급/최상급(-er, -est), 소유격 단수/복수('s, -s)로 분류하였다. 여기에 인칭대명사를 추가하여 총 6가지의 굴절형태소를 정하였다. Excel의 자동 분류가 정확한지 확인하기 위해 연구자들이 3번 이상의 교차 확인 작업을 수작업으로 진행하였다. 영어는 동사, 형용사, 명사의 불규칙 형태가 있는데 *wore*와 같은 불규칙 동사 과거형, *worst*와 같은 불규칙 형용사 최상급, *children*과 같은 불규칙 명사 복수형은 전형적인 굴절형태소로 분류하기가 힘들다. 그리하여 수작업을 통해 전체 Type에서 이를 불규칙 형태(Irregular(Irr.))로 분류한 다음 다시 동사(V), 명사(N), 형용사(A)로 분류하였다. 이를 1차적으로 분류한 것과 통합하여 각 굴절형태소별 Type 목록이 들어있는 파일을 만들었다. 그 다음 CLAN에서 국적, 연령별 파일을 대상으로 FREQ 명령을 실행하여 각 굴절형태소의 Type과 Token, TTR을 산출해 내었고 VocD 명령을 실행하여 D값을 산출해 내었다.

## 결과

코퍼스 분석 결과 총 Type 13,528개의 1,221,307개 Token의 굴절형태소를 추출해 냈다. 이 수치는 표 2의 수치와 일치하지 않는데 이는 1-7세의 전체 Type과 Token 분석 시와 마찬가지로 두 나라 아동 발화에서 Type이 중복되었기 때문이다.

가장 Type이 많은 굴절형태소는 3인칭 현재 단수형과 명사 복수형인 -(e)s로 영국은 총 4,245개의 Type과 114,524개의 Token, 미국은 3,905개의 Type과 172,631의 Token이 확인되었다. 그러나 -(e)s는 두 가지 굴절형태소를 합한 것이므로 단일 굴

표 2. 국적별 굴절형태소 TTR과 D값

굴절형태소	UK				USA			
	Type	Token	TTR	D	Type	Token	TTR	D
-ing	1,229	39,759	0.031	33.26	1,084	47,458	0.023	27.38
-(e)d_irr(V)	1,006	82,474	0.012	10.78	1,472	132,405	0.011	19.23
-er_-est_irr(A)	217	11,499	0.008	0.77	198	13,978	0.014	1.57
-(e)s_irr(N)	4,245	114,524	0.037	18.18	3,905	172,631	0.023	30.78
pronoun	52	165,778	0.000	2.64	51	321,019	0.000	1.82
-s'_'s	1,359	49,219	0.028	4.48	1,904	71,172	0.027	3.77
합계	8,108	463,253	0.019	11.685	8,614	758,663	0.016	14.09

절형태소로는 현재진행 형태소인 *-ing*의 TTR이 가장 높게 나타나 현재분사가 다른 굴절형태소에 비해 다양한 어휘에 사용되었음을 알 수 있었다. 이는 굴절형태소가 얼마나 다양한 어휘에 적용되었는지를 나타내는 D값에서도 확연히 드러났는데 영국은 33.26, 미국은 27.38로 나타나 굴절형태소 중에서도 D값이 가장 높게 나타났다. 이는 결과적으로 영국과 미국 아동들이 현재진행 형태소를 다른 굴절형태소들에 비해 더 자유롭게 사용하고 있음을 의미한다. 반면 가장 TTR이 낮은 굴절형태소는 대명사이지만 D값을 확인한 결과 비교급과 최상급 형태소 *-er*과 *-est*가 가장 낮게 나타나 대명사와 비교급과 최상급 형태소는 소수의 Type이 높은 빈도로 사용되었음이 확인되었다. 이는 (1)에서 살펴본 Brown(1973)의 습득 순서와도 거의 일치하는 결과이다. 그러나 이 결과는 형태상의 분류를 통해 얻은 것으로 향후 용례 분석을 통한 좀 더 상세한 분석이 필요하다.

다음은 각 형태소 별로 아동의 연령에 따른 Type과 Token, TTR, D값에 대한 분석 결과이다. 본 연구에서는 앞에서 언급하였듯이 다양한 굴절형태소 중에 동사의 대표적 굴절형태소인 현재분사 *-ing*와 과거형 *-(e)d*, 그리고, 형용사의 대표적 굴절형태소인 비교급 *-er*과 최상급 *-est* 분석에 집중하였다.

### 현재분사

Brown(1973)에 따르면, 영어에서 동사의 현재분사는 아동이 다른 형태소보다 가장 먼저 습득하는 문법형태소이다. 13,528개의 전체 굴절형태소 Type 중 현재분사로 분류된 *-ing*의 경우 영국은 1,997개의 Type과 39,759 Token, 미국은 2,881개의 Type과 47,458개의 Token이 확인되었다. 아래 표는 국적, 연령별로 현재분사가 사용된 어휘 Type과 Token, TTR 및 D값에 대한 분석 결과이다.

1-7세의 D값을 보면 영미 아동 모두 연령에 따라 차이가 별로 없는 것으로 나타났다. 연령과 D값의 상관관계를 알아보기 위해 Spearman 상관계수를 알아본 결과 모두 통계적으로 유의미하지 않은 것으로 나타났다(영국 아동:  $r = 0.025$ ,  $p > .05$ ; 미국 아동  $r = 0.385$ ,  $p > .05$ ). 이는 1세부터 *-ing* 형태소를 여러 동사에 걸쳐 활발하게 사용하고 있음을 의미한다고 하겠다.

현재진행형이 사용된 고빈도 어휘를 살펴보면, 영국 아동의 고빈도 현재분사는 *going, doing, coming* 순이었으며 미국은 *going, doing, making* 순이었다(빈도 상위 20위권 동사는 <부록 1>과 <부록 2> 참고). 동사의 진행형은 동명사와 동일한 형태를 취하므로 향후 용례 검토를 통한 세세한 재분류가 요구된다. 형태상 분류를 기준으로 아동과 엄마가 생산해 낸 *-ing*형의 어휘를 확인, 비교한 결과, 영국과 미국 모

표 3. 현재분사의 TTR과 D값

-ing	UK				USA			
	Type	Token	TTR	D	Type	Token	TTR	D
1세	97	719	0.135	13.83	290	3,378	0.086	27.60
2세	1,158	33,264	0.035	33.27	637	15,904	0.040	28.14
3세	256	3,071	0.083	20.18	558	10,924	0.051	23.93
4세	131	642	0.204	21.84	569	11,343	0.050	26.67
5세	154	1,009	0.153	27.07	383	3,740	0.102	28.95
6세	83	264	0.314	27.07	244	1,210	0.202	36.05
7세	118	790	0.149	22.48	200	959	0.209	27.23
Total	1,997	39,759	0.153	23.68	2,881	47,458	0.106	28.37

두에서 엄마 발화의 고빈도 50위권 어휘 약 80-90%가 아동의 50위권 고빈도 어휘에서도 나타난 것을 확인했다. 엄마의 발화에서 가장 빈도가 높은 어휘로는 *doing, making, coming, playing, eating, looking, having* 등이 있었다. 한편, 과잉일반화 예는 매우 드물게 나타났는데, 명사에 *-ing*을 붙여 사용한 *tomising, swording, appetizing* 각 1-2회, 형용사에 *-ing*을 붙여 사용한 *noticeabling* 1회만 확인되었다.

### 동사의 과거형

굴절형태소로 분류된 13,528개의 Type 중 동사 과거형으로 분류된 *-(e)d\_irr(V)* 굴절형태소를 확인한 결과, 영국 CHILDES 데이터에서는 1,886개의 Type과 82,474개의 Token, 미국 데이터에서는 3,746개의 Type과 132,405개의 Token이 확인되었다. 이를 국적, 연령별로 보면 표 4와 같다.

동사 과거형 형태소의 사용에 있어서도 연령이 증가함에 따라 대체적으로 D값이 증가하는 추이를 보인다. 이는 아동의 연령이 높아질수록 다양한 어휘에 굴절 형태소를 적용하고 있음을 의미한다. D값은 영국의 경우 4세 때까지 증가하다가 그 이후 감소하는 경향을 보였고 미국은 계속 증가하다가 4세와 7세 때 감소하는 경향을 보였다. 국적, 아동별 연령과 D값의 상관관계를 확인하기 위하여 Spearman

표 4. 규칙 동사와 불규칙 동사의 과거형 TTR과 D값

-(e)d _irr(V)	UK				USA			
	Type	Token	TTR	D	Type	Token	TTR	D
1세	94	1,800	0.052	5.24	223	5,145	0.043	16.34
2세	913	68,022	0.013	11.45	726	33,190	0.022	19.92
3세	262	6,474	0.040	12.63	757	33,020	0.023	21.24
4세	166	1,600	0.104	15.68	820	38,482	0.021	21.00
5세	181	2,327	0.078	14.82	547	13,106	0.042	22.60
6세	108	610	0.177	12.34	352	4,755	0.074	24.52
7세	162	1,641	0.099	13.39	321	4,707	0.068	20.05
Total	1,886	82,474	0.080	12.22	3,746	132,405	0.042	20.81

상관계수를 구하였는데, 각각 통계상으로 유의미한 값인  $r = 0.68$ 을 아주 근소한 차이로 벗어났다(영국 아동:  $r = 0.643, p > .05$ ; 미국 아동:  $r = 0.66, p > .05$ ). 이는 연령에 따라 과거형 형태소의 사용이 증가하는 경향성을 보였다고 할 수 있다. 또한, D값이 모든 연령대에서 미국 아동이 영국 아동보다 높게 나타나 미국 아동들이 영국 아동들에 비해 굴절형태소를 좀 더 다양한 어휘와 함께 사용하고 있음을 알 수 있다. 이러한 미국과 영국 아동의 차이는 앞서 살펴본 현재진행 형태소의 사용에서도 확인되었다. D값은 현재진행 형태소보다 과거형 형태소에서 더 낮게 나타나 아동들이 과거형태소를 현재진행 형태소보다 더 늦게 습득한다는 Brown(1973)의 결과와 일치한다고 할 수 있다.

국적, 연령별 동사 과거형 고빈도 어휘 50개를 확인한 결과, 두 나라의 모든 연령에서 사용된 빈도가 높은 동사는 대부분 불규칙 동사가 차지하고 있었다. 영국의 경우 전 연령에 걸쳐 *gone, got, put*이 빈도가 가장 높았으며 미국의 경우는 *put, was, got*이었다(빈도 상위 20위권 과거형 어휘 <부록 3>과 <부록 4> 참고). 그러나 *put*은 현재형과 과거형이 동일하므로 향후 용례 확인이 필요하며 *got*도 현재의

표 5. 1-7세 아동의 발화 빈도 상위 50위권의 과거형 어휘 유형

구분	UK	USA
규칙 동사	called, finished, happened, dropped, wanted, stopped,	called, finished, wanted, happened, supposed
불규칙 동사	got, gone, done, was, did, stuck, said, broken, went, had, found, been, broke, saw, lost, made, fell, left, heard, came, were, fallen, thought, seen, bought	was, got, did, said, had, went, were, made, ate, saw, done, came, fell, gone, broke, found, broken, took, was, forgot, stuck, told, thought, gave, lost
기본형-과거형 동일	put, come, let, shut, read, run, cut, hurt	put, come, let, read, cut, hurt, hit, run, set
조동사 과거형	could, didn't, might, would, got to, should, couldn't	could, didn't, would, should, might
형용사적 용법	bit, fit, wet, tired	bit, wet, fit, wet, tired, scared

의미로 쓰이므로 이 또한 용례 확인을 통해 분석할 필요가 있다. 영미 아동의 과거형 발화 빈도 상위 50위권의 어휘를 자세히 보면 표 5와 같다.

규칙 동사의 경우, 영국은 6개가 확인되었고 가장 빈도가 높은 규칙 동사 과거형은 *called*로 전체 과거형에서 15위를 차지하였다. 불규칙 동사 과거/과거분사형인 *got, gone, done, was, did, stuck* 등은 10위권 내에 포함되어 있었으며 50개의 고빈도 어휘에서 25개를 차지하였다. *put, come* 등과 같은 동사 기본형이 과거형 또는 과거분사형과 동일한 것은 총 8개가 있었으며 *put*과 *come*은 각각 2위와 4위를 차지하였다. 또한 조동사 과거형에는 *could, didn't* 등 7개가 있었고 나머지는 형용사로 사용됐음직한 *bit, fit, wet, tired*였다. 미국의 경우 규칙 동사 과거형은 5개가 확인되었고 가장 빈도가 높은 규칙 동사 과거형은 *called*로 29위에 있었다. 불규칙 동사 과거형은 *was, got, did* 등이 고빈도 어휘에 속해 있었으며 총 50개의 고빈도 어휘 중 25개를 차지하였다. *put, come* 등과 같은 동사 기본형이 과거형 또는 과거분사형과 동일한 것은 총 9개가 있었으며 *put*과 *come*은 각각 1위와 5위를 차지하였다. 또한 조동사 과거형에는 *could, didn't* 등 5개가 있었고 나머지는 형용사로 사용됐음직한 *bit, wet, fit, wet, tired, scared*였다. 이처럼, 아동의 동사 과거형 고빈도 어휘 목록에서 규칙 동사보다는 불규칙 동사의 활용이 더 많았는데 이는 빈도만 고려해 볼 때 불규칙 동사 과거형이 규칙 동사 과거형보다 빨리 습득된다는 Brown(1973)의 결과와도 일치한다.

다음으로, 과잉일반화와 관련하여 국적과 연령별로 어휘 빈도를 확인한 결과, 규칙 동사 과거형태소 *-(e)d*를 불규칙 동사에 붙여 과잉일반화한 어휘 중 동사 *go*와 *fall*의 과잉일반화율이 가장 높게 나타났다. 표 6에 *go*의 과잉일반화 유형 및 빈도 결과를 제시하였다.

표 6에서 보여주는 것처럼, 영국 아동이 1세 때 *\*goed*를 한 번 생산해낸 것을 제외하고는 영국과 미국 아동은 1세 때 정확하게 과거형을 사용하다가 2세 때 과잉일반화 하는 경향이 두드러지게 나타남을 확인하였다. 2세 영국 아동의 경우 과잉일반화 유형으로 기본형에 과거형태소 *-(e)d*를 붙인 형태인 *\*goed*, 과거분사에 *-(e)d*를 붙인 형태인 *\*goned*가 나타났으며 2세 미국 아동의 경우에는 과거형에 *-(e)d*를 붙인 형태인 *\*uented*도 생산되었다. 전반적으로 이러한 불규칙 동사 과거형의 과잉일반화 경향은 3세 때 그 유형이 줄기 시작해 4-5세가 되면서 점차 사라지는



표 6. 불규칙 동사 Go의 과잉일반화 유형 및 빈도

구분	Correct						Overgeneralization								Total	
	went		gone		Subtotal		*goed		*goned		*wented		subtotal			
	UK	USA	UK	USA	UK	USA	UK	USA	UK	USA	UK	USA	UK	USA	UK	USA
1세	-	29	580	239	580	268	1	-	-	-	-	-	1	-	581	268
2세	784	572	4,978	627	5,762	1,199	17	38	3	2	-	1	20	41	5,782	1,240
3세	73	675	192	142	265	817	3	52	-	-	-	-	3	52	268	869
4세	23	860	21	109	44	969	-	4	-	-	-	-	-	4	44	973
5세	33	286	22	30	55	316	-	-	-	-	-	-	-	-	55	316
6세	24	158	2	6	26	164	-	-	-	-	-	-	-	-	26	164
7세	100	74	15	22	115	96	-	-	-	-	-	-	-	-	115	96

것으로 나타났다. 또한 영국과 미국 아동 과잉일반화율을 비교해 보면, 미국 아동이 과잉일반화 오류를 더 많이 보인 것으로 나타났으며 이는 카이제곱 검정 결과, 통계적으로도 유의미한 것으로 확인되었다( $\chi^2(1, N = 10,797) = 101.468, p < .05$ ). 이러한 과잉일반화에 따른 U자형 발달 곡선이 그림 1에 잘 나타나 있다. 그림 1은 불규칙 동사 go의 정확도를 국적, 연령별로 비교한 것이다.

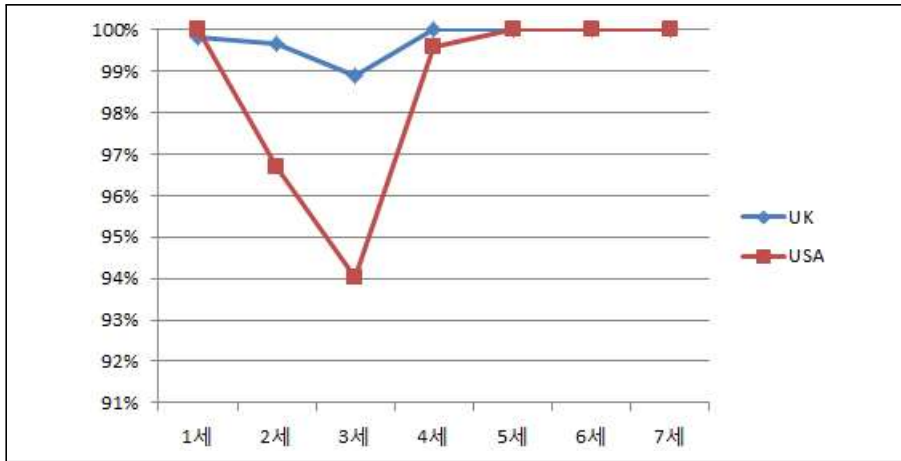


그림 1. 불규칙 동사 go의 과거형 정확도 비교

그림 1은 과잉일반화와 관련한 U자형 형태소 발달 양상을 잘 보여준다. 영국과 미국 아동을 비교해 보면, 1세 때에는 과잉일반화 오류를 보이지 않다가 2-3세 사이에 과잉일반화율이 증가하고 4세 때부터는 줄어들어 과잉일반화 오류가 사라지는 것을 볼 수 있다. 두 나라 아동 모두 U자형 발달 양상을 보이고 있는데 영국 아동의 과잉일반화율이 상대적으로 낮게 나타났다.

다음으로 *fall*의 과잉일반화 유형 및 빈도에 대해서 알아보았으며 그 결과는 표 7에 나와 있다.

표 7. 불규칙 동사 *fall*의 과잉일반화 유형 및 빈도

구분	Correct						Overgeneralization								total	
	fell		fallen		subtotal		*falled		*felled		*fallened		subtotal		UK	USA
	UK	USA	UK	USA	UK	USA	UK	USA	UK	USA	UK	USA	UK	USA		
1세	3	73	-	-	3	73	-	-	-	1	-	-	-	1	3	74
2세	315	462	290	2	605	464	57	94	6	5	4	-	67	99	672	563
3세	37	324	16	5	53	329	4	34	-	2	-	-	4	36	57	365
4세	15	264	2	1	17	265	1	8	-	1	-	-	1	9	18	274
5세	16	94	2	-	18	94	-	2	-	1	-	-	-	3	18	97
6세	7	32	-	1	7	33	1	-	-	-	-	-	1	-	8	33
7세	5	15	1	1	6	16	-	1	-	1	-	-	-	2	6	18

*fall*의 경향은 *go*와는 달리 영국 아동에게서 더 다양한 과잉일반화 유형이 생산되었다. 영국 1세 아동에게서는 어떠한 과잉일반화도 나타나지 않았으며 미국 1세 아동은 과거형에 *-(e)d*를 붙인 *\*felled*를 생산해 내었다. 2세 영국 아동의 경우 과잉일반화 유형은 기본형에 과거형태소 *-(e)d*를 붙인 형태인 *\*falled*, 과거형에 *-(e)d*를 붙인 *\*felled*, 과거분사에 *-(e)d*를 붙인 *\*fallened*가 나타났으며 2세 미국 아동의 경우는 과거분사에 *-(e)d*를 붙인 *\*fallened*은 나타나지 않았다. *go*와 마찬가지로 2세 때 영국과 미국 아동 모두에게서 과잉일반화 경향이 현저히 나타났으며 연령이 증가하면서 과잉일반화 경향은 점점 줄어들어 6-7세 때에는 거의 나타나지 않았다. 전반적으로 이러한 불규칙 동사 과거형의 과잉일반화 유형은 3세 때 줄기 시작해 4-5세가 되면서 점차 사라지는 것으로 나타났다. 또한 영국과 미국 아동의 과잉일

반화율을 비교해 본 결과, go에서와 마찬가지로 미국 아동이 과잉일반화 오류를 더 많이 보인 것으로 나타났다. 그러나 카이제곱 검정 결과, 이러한 차이가 통계적으로는 유의미하지는 않은 것으로 확인되었다( $\chi^2(1, N = 2,206) = 0.798, p > .05$ ). 그림 2는 불규칙 동사 fall의 정확도를 국적, 연령별로 비교한 것이다.

그림 2를 보면 그림 1에서처럼 영국과 미국의 아동 모두 U자형 발달 곡선을 보이며 미국보다 영국 아동의 정확도가 더 높게 나타났다. 그러나 go에서와 마찬가지로 fall의 정확도가 2-4세 사이에 낮게 나타났지만 go와는 달리 5세 이후에도 과잉일반화 오류를 보였다. 이는 *fell-felled-felled*라는 fall과 다른 동사의 규칙 변화의 형태에 영향을 받은 결과로 해석될 수 있고 또한 이러한 현상의 원인을 CHILDES 코퍼스 6-7세 자료의 규모가 상대적으로 작은 것에서 찾을 수 있을 것이다. 규모가 작은 데이터에서의 오류가 1회 나타난 경우, 정확도는 상대적으로 낮게 나타나게 된다. 즉 분석 대상 CHILDES 데이터가 적어 그림 2에서와 같이 6-7세의 정확도가 감소하는 결과를 낳은 것으로 해석할 수 있다.

그러나 표 2의 규칙 동사와 불규칙 동사의 과거시제 굴절형태소의 Type과 Token을 국적, 연령별로 비교해 보면 영국은 2세 아동의 자료가 압도적으로 많은 반면 미국은 2-4세의 코퍼스 규모가 큰 차이를 보이지 않고 있다. 그럼에도 불구하고

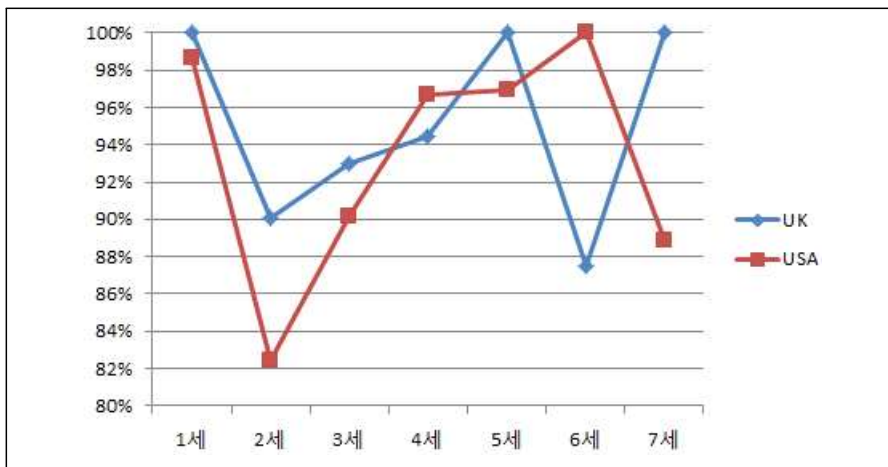


그림 2. 불규칙 동사 fall의 과거형 정확도 비교

미국 아동의 *go*와 *fall*의 과잉일반화 오류가 2세에 가장 많이 나타났다가 3세부터 점차 사라지는 U자형 발달 곡선을 보이므로 이는 과잉일반화 오류로 볼 수 있다.

### 비교급과 최상급

굴절형태소로 분류된 13,528개의 Type 중 비교급과 최상급으로 분류된 *-er*과 *-est*는 영국의 경우 199개 Type에 11,499 Token, 미국은 482개 Type에 13,978 Token이 확인되었다. 비교급과 최상급의 Type, Token, TTR과 D값이 국적, 연령별로 표 8에 제시되어 있다.

비교급과 최상급의 경우, D값이 연령에 따라 증가하다가 7세 때 소폭으로 낮아지는 추세를 보였다. 연령과 D값의 상관관계를 확인하기 위해 Spearman 상관계수를 구한 결과 두 나라 모두에서 상관관계가 통계적으로 유의미하게 나타났다(영국 아동:  $r = 0.779, p < .05$ ; 미국 아동:  $r = 0.776, p < .05$ ). 이는 연령의 증가에 따라 아동들의 비교급과 최상급 사용이 보다 다양한 어휘로 확장된다는 것을 의미한다. 또한, 영국과 미국 아동을 비교할 때, 앞서 살펴본 두 굴절형태소와는 달리 비교급과 최상급에서는 두 나라의 D값 차이가 확연히 드러나지 않았다. 1-3세의 경우 미국 아동이 영국 아동보다 더 다양한 어휘에 비교급과 최상급 형태소를 적용

표 8. 비교급과 최상급의 TTR과 D값

-er _est _irr(A)	UK				USA			
	Type	Token	TTR	D	Type	Token	TTR	D
1세	11	661	0.017	0.10	19	1,660	0.011	0.26
2세	76	9,940	0.008	0.74	70	4,242	0.017	0.92
3세	26	420	0.062	1.24	99	2,758	0.036	2.03
4세	21	126	0.167	2.89	122	3,349	0.036	2.70
5세	28	194	0.144	2.92	84	1,211	0.069	3.44
6세	20	57	0.351	5.82	47	376	0.125	4.06
7세	17	101	0.168	2.54	41	382	0.107	2.17
total	199	11,499	0.131	2.32	482	13,978	0.057	2.23

하는 것으로 나타났으나 4세, 6세와 7세에는 영국 아동의 D값이 더 높은 것으로 나타나 두 나라 간 별 차이가 없었다.

두 나라 모두에서 가장 빈도가 높은 단어는 *more*와 *better*였다. *more*의 빈도가 높은 이유는 'give me more'에서와 같이 명사로 쓰이기 때문이기도 할 것이다. 그 외에 영국 아동의 경우에는 *last*, *bigger*, *higher* 등이 많이 쓰였고, 미국 아동의 경우에는 *cleaner*, *higher*, *bigger*, *later* 등의 순으로 많이 쓰였다(상위 20위권 빈도 <부록 5>와 <부록 6> 참고).

과잉일반화와 관련하여, 아동이 발화한 비교급과 최상급 빈도를 확인한 결과 영국과 미국 아동에게서 모두 과잉일반화 오류를 발견하였는데, 그 중 가장 빈도가 높은 것은 원형 *little*에 비교급 접미사 *-er*과 최상급 접미사 *-est*를 붙인 *littler*와 *littlest*였다. 또한 영국과 미국 아동의 비교급에 대한 과잉일반화율을 비교해 본 결과, 미국 아동이 과잉일반화 오류를 더 많이 보인 것으로 나타났다. 그러나 그 차이는 카이제곱 검정 결과 통계적으로는 유의미하지 않은 것으로 확인되었다( $\chi^2$  (1,  $N = 52$ ) = 0.021,  $p > .05$ ). 불규칙 형용사 *little*의 비교급과 최상급 과잉일반화 유형과 각각의 빈도를 분석하였으며, 결과는 표 9에 국적, 연령별로 제시되어 있다.

표 9를 보면 영국과 미국 아동 1세에서는 *little*의 사용이 없었고 2세부터 나타났다. 2세 영국 아동의 경우 총 4번의 *little* 관련 발화 중 한 번의 과잉일반화가 일어

표 9. 불규칙 형용사 *little*의 비교급/최상급 과잉일반화 유형 및 빈도

구분	Correct		Overgeneralization				Total	
	less		littler		littlest		UK	USA
	UK	USA	UK	USA	UK	USA		
1세	0	0	0	0	0	0	0	0
2세	3	1	1	7	0	5	4	13
3세	0	3	2	6	1	9	3	18
4세	0	6	0	6	1	4	1	16
5세	0	7	0	4	1	0	1	11
6세	0	0	1	4	1	3	2	7
7세	0	1	0	0	0	0	0	1

났으나 미국 아동은 총 13번의 발화 중 한 번의 정확한 발화가 있었고 7번의 비교급 오류와 5번의 최상급 오류를 보였다. 그림 3은 *little*의 정확도를 국적과 연령에 따라 보여주고 있다.

비교급과 최상급의 과잉일반화 경향은 앞에서 살펴본 현재분사, 과거형과 마찬가지로 2-4세에 정확도가 떨어지다가 점점 올라가는 U자형 발달 형태를 보여주고 있다. 그러나 현재분사 및 과거형과는 달리 비교급과 최상급 오류가 5-6세에도 나타났는데 이는 *fall*의 경우와 마찬가지로 5-6세의 자료가 적어 그림 3에서와 같이 정확도가 떨어진 것으로 해석할 수 있다. 또한 실제로 *smaller* 대신에 *littler*가 영미 영어에서 드물지만 사용되고 있다는 점에서 엄마 입력의 가능성을 확인해 볼 필요가 있었다. 이를 위해 먼저 두 나라의 1-5세 파일을 대상으로 CLAN의 FREQ 명령을 실행하여 엄마 발화에서 *littler*와 *littlest*를 확인하였다. 그 결과, *littler*는 영국과 미국 각각 11회, 33회, *littlest*는 각각 6회, 19회 나타났다. 이 중에서 *littler*가 많이 발화된 미국 아동 데이터를 확인한 결과, 총 4개의 파일에서 아동과 엄마가 모두 *littler*를 발화한 것으로 나타났다. 이 4개의 파일에서 아동은 8회, 엄마는 17회 *littler*를 발화한 것으로 나타나 이는 엄마 입력의 영향 가능성을 확인해주는 결과라 할 수 있다.

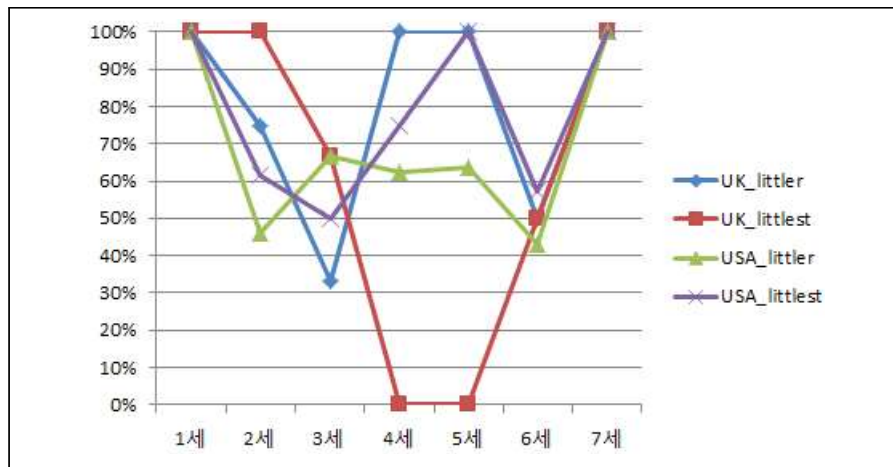


그림 3. 불규칙 형용사 little의 정확도 비교

## 논 의

본 연구에서는 아동의 굴절형태소 발달에 대하여 CHILDES 코퍼스의 1-7세 아동 발화를 분석하여 기존의 소수 아동 데이터에 기반을 둔 굴절형태소의 발달 순서 및 발달 양상을 검증하였다. 본 연구에서 제시한 구체적인 연구 문제와 관련하여 코퍼스 분석 결과는 다음과 같다.

첫째, 아동의 굴절형태소 사용에 있어, 아동의 연령 증가에 따라 각 형태소별 양상이 달리 나타남을 확인하였다. 즉, 현재분사 *-ing*의 경우는 연령과 D값의 상관관계가 유의미하지 않았고, 규칙 동사와 불규칙 동사의 과거형 *-(e)d*에서는 양의 상관관계 경향성을 발견하였으며, 비교급과 최상급 *-er/-est*에서는 유의미한 양의 상관관계를 보였다. 이 결과는 아동들이 1세 때부터 현재분사를 활발히 사용하고 있다는 것을 보여주는 것으로, 과거형은 아동의 연령 증가에 따라 천천히 다른 동사로 확장해 가는 경향성을 보여주는 것으로, 비교급과 최상급 *-er/-est*는 아동의 연령 증가에 따라 다른 형용사로 확장해 가는 것이 확연히 보여주는 것으로 해석될 수 있다. 이러한 결과는 각 형태소의 D값 비교에서도 나타나는데, 그 크기는 ‘현재진행 > 과거 > 비교급/최상급’의 순서로 나타났다. 다시 말하면, 아동의 발화에서 나타나는 굴절형태소의 어휘 다양성을 확인한 결과, 영국과 미국 모두 현재진행 굴절형태소 적용이 가장 다양했으며 그 다음으로 동사의 과거형, 비교급과 최상급이었다.

또한 규칙 동사와 불규칙 동사의 과거형 고빈도 어휘를 확인한 결과 고빈도 어휘 중 불규칙 동사의 과거형이 규칙 동사의 과거형보다 약 4배 많이 사용되었음을 발견하였다. 이러한 결과는 위의 (1)에서 보이는 Brown(1973)의 굴절형태소 습득 순서의 결과에서처럼 불규칙 동사 과거형이 규칙 동사 과거형보다 빨리 습득된다는 선행연구의 발견을 지지하는 결과이다.

둘째, 본 연구에서는 아동의 굴절형태소 사용에 있어서 과잉일반화 현상을 확인하였다. 또한 과잉일반화에 따른 U자형 형태소 발달 양상도 발견하였다. 이는 Brown(1973)과 Marcus et al.(1992) 외 여러 선행 연구에서 제안한 아동의 U자형 언어 발달이 대규모 CHILDES 데이터에서도 확인되었음을 의미한다. 과잉일반화 현상은 과거형에서 가장 많이 발견되었으며, 현재분사의 과잉일반화 현상은 거의 발

견되지 않았다. 이런 경향은 과잉일반화가 형태소마다 다르게 나타난다는 기존의 발견들[2, 4]의 결과와 일치하는 것이다. 과잉일반화 현상이 사라지는 시기는 동사마다 달랐는데, 예를 들면, 과거형 과잉일반화 오류인 \**fell-ed*는 \**go-ed*보다 늦게까지 남아 있었다. 이는 *felled*가 *fell*이라는 규칙변화(*fell-felled-felled*) 단어의 존재로 인하여 아동이 불규칙변화 동사인 *fall*(*fall-fell-fallen*)과 혼동하여 늦게까지 오류를 보인 것으로 해석될 수 있다. 또한 6-7세의 CHILDES 데이터가 상대적으로 적어 과잉일반화율이 더 높게 나타난 것으로 해석할 수도 있다. 마지막으로, 과잉일반화 오류는 영국 아동보다 미국 아동에게서 더 많이 일어났다. 이와 관련한 보다 자세한 논의는 아래에 있다.

셋째, 위에서 살펴본 굴절형태소의 발달 양상을 영국과 미국 아동을 대상으로 비교한 결과, 유사점과 차이점을 발견하였다. 먼저, 유사점으로는 두 나라 아동 모두 연령이 증가하면서, 굴절형태소를 다양한 어휘들에 확장하여 사용한 것과 과잉일반화의 감소에 있어서 두 나라 아동 모두 U자형의 발달 양상을 보인 것을 들 수 있다. 차이점으로는 D값이 미국 아동이 영국 아동보다 더 높게 나타나 미국 아동이 좀 더 다양한 어휘에 각각의 굴절형태소를 적용하고 있음을 들 수 있다. 그러나 과잉일반화 오류는 영국 아동보다 미국 아동에게서 더 많이 나타났는데 이러한 차이는 카이제곱 검정 결과에서도 나타났다. 또한 형용사 *little*의 비교급과 최상급 과잉일반화라 할 수 있는 *littler*에 대한 빈도를 조사한 결과, 영국 아동보다 미국 아동에게서 더 많이 나타났다. 그러나 카이제곱 검정 결과, 통계적으로는 유의미하지 않았다. 좀 더 정확한 분석을 위해 아동과 엄마의 발화에서 *littler*가 발화된 파일을 확인한 결과, 엄마의 발화에서 *littler*가 많이 발화된 경우 아동의 발화에서도 *littler*가 많이 나타났는데 이는 엄마의 입력 영향 가능성을 강하게 시사하는 결과이다.

## 결 론

본 연구에서는 CHILDES 데이터에 나타나는 1-7세의 영어 모국어 아동의 자발적 발화 자료를 분석하여 굴절형태소 습득 양상을 확인하였다. 아동의 연령이 증가함



에 따라 아동이 굴절형태소를 다양한 어휘에 확장 적용하고 있음을 발견하였고, U 자형의 발달 양상을 보이는 것을 검증하였다. 굴절형태소 사용 시, 아동이 규칙을 적용하고 있음을 보여주는 과잉일반화 현상을 재확인하였으며, D값을 확인한 결과, 여러 형태소의 습득 순서에서 현재분사 *-ing*가 과거형 *-(e)d*보다 먼저 습득되고, 불규칙 동사의 과거형이 규칙 동사의 과거형보다 더 먼저 습득된다는 Brown(1973)의 굴절형태소 습득 순서를 지지하는 결과를 발견하였다. 마지막으로, 영국과 미국 아동의 자료에서 굴절형태소 습득 시 부모의 입력 영향 가능성을 확인하였다. 본 연구의 결과는 약 470만 단어로 이루어진 방대한 코퍼스 분석을 통해 나온 결과라는 데에 큰 의의가 있다. 즉 이전의 비교적 소수 아동의 발화 데이터에 의존한 결과들에 비해 보다 믿을 만한 결과를 제공했다고 할 수 있다. 또한, 기존 연구에서는 많이 다루지 않았던, 같은 영어권이지만 다른 지역인 영국과 미국 아동의 굴절형태소 발달을 비교함으로써 아동의 굴절형태소 발달에 미치는 부모의 입력 영향 가능성을 제시하였다.

앞에서 언급했듯이, 본 연구 결과는 전체 CHILDES 데이터 기반의 아동의 굴절형태소 발달에 대한 진행 중인 연구의 초기 결과이다. 앞에서 언급한 바와 같이 각각의 형태소 분류의 세분화된 분석이 앞으로 더 진행되어야 한다. 방대한 규모의 CHILDES 코퍼스를 대상으로 세분화된 분석이 완료된다면 이는 이전에 이루어진 연구 결과에 대한 보다 확실한 검증 및 향후 아동의 언어 발달 연구에 중요한 기초 자료를 제공한다고 할 수 있다. 특히, 본 연구에서는 아동의 국적, 연령별로 데이터를 재조직하여 보다 효율적으로 대규모 코퍼스를 사용할 수 있게 한 것에 큰 의의가 있다고 하겠다. 또한 본 굴절형태소 발달 연구에 의한 새로운 결과를 이후 기존의 제2언어로서의 영어 굴절형태소 습득 연구 결과[12] 등과 비교함으로써, 그동안 많은 연구자들이 논쟁해 온 제1언어와 제2언어의 습득 이론에 통찰력을 제공할 수 있을 것이다.

## 참고문헌

- [1] MacWhinney, B. & Snow, C. E. (2000). The Child Language Data Exchange System:

An Update. *Journal of Child Language*, 17, 457-472.

CHILDES (<http://childes.psy.cmu.edu/>)

- [2] Kuczaj, Stan A. (1977). Why do children fail to overgeneralize the progressive inflection?, *Journal of Child Language*, 5, 167-171.
- [3] Maslen, Robert J C; Theakston, Anna L; Lieven, Elena V M; Tomasello, Michael (2004). A Dense Corpus Study of Past Tense and Plural Overgeneralization in English, *Journal of Speech, Language, and Hearing Research*, 47-6, 1319-1333.
- [4] Brown, Roger (1973). *A first Language-The early Stages*, Harvard University Press.
- [5] Marcus, Gary F.; Pinker, Steven; Ullman, Michael; Hollander, Michelle; Rosen, T. John; and Su, Fei (1992). Overgeneralization in Language Acquisition, *MONOGRAPHS OF THE SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT* Serial No. 228, Vol. 57.
- [6] Berko, Jean (1958). The child's learning of English morphology. *Word*, 14, 47-56.
- [7] Templin, M. C. (1957). *Certain language skills in children*. Minneapolis: University of Minnesota Press.
- [8] Johansson, Victoria (2008). Lexical diversity and lexical density in speech and writing: a developmental perspective, Lund University, Dept. of Linguistics and Phonetics, *Working Papers*, 53, 61-79.
- [9] Richards, Brian J. & David Malvern (1997). Quantifying lexical diversity in the study of language development. Reading: *Faculty of Education and Community Studies*.
- [10] Malvern, David; Brian Richards; Ngoni Chipeer & Pilar Duran (2004). *Lexical diversity and language development: quantification and assessment* New York: Palgrave Macmillan.
- [11] McCarthy, Philip M. & Jarvis S (2007). vocd: A theoretical and empirical evaluation, *Language Testing*, 24-4, 459-488.
- [12] Shin, Sarah J. & Milroy, Lesley (1999). Bilingual language acquisition by Korean schoolchildren in New York City, Bilingualism: *Language and Cognition*, 2, 147-167.

1차 원고접수 : 2013. 05. 11.

2차 원고접수 : 2013. 07. 18.

최종게재승인 : 2013. 08. 27.

(Abstract)

## A Study on the Development of English Inflectional Morphemes Based on the CHILDES Corpus

Myung Sook Min<sup>1)</sup>      Jongsup Jun<sup>1)</sup>      Sun-Young Lee<sup>2)</sup>

<sup>1)</sup>Hankuk University of Foreign Studies, <sup>2)</sup>Cyber Hankuk University of Foreign Studies

The goal of this paper is to test the findings about English-speaking children's acquisition of inflectional morphemes in the literature using a large-scale database. For this, we obtained a 4.7-million-word corpus from the CHILDES (Child Language Data Exchange System) database, and analyzed 1,630 British and American children's uses of English derivational morphemes up to age 7. We analyzed the type and token frequencies, type per token ratio (TTR), and the lexical diversity (D) for such inflectional morphemes as the present progressive *-ing*, the past tense *-(e)d*, the comparative and superlative *-er/-est* with reference to children's nationality and age groups. To sum up our findings, the correlations between the D value and children's age varied from morpheme to morpheme; e.g. we found no correlation for *-ing*, a marginal correlation for *-ed*, and a strong correlation for *-er/-est*. Our findings are consistent with Brown's (1973) classical observation that children learn progressive forms earlier than the past tense marker. In addition, overgeneralization errors were frequently found for *-ed*, but rarely for *-ing*, showing a U-shaped developmental pattern at ages 2-3. Finally, American children showed higher D scores than British children, which showed that American children used inflectional morphemes for more word types compared with British children. The present study has its significance in testing the earlier findings in the literature by setting up well-defined methodology for analyzing the entire CHILDES database.

*Key words* : Language acquisition, CHILDES, Inflectional morpheme, Overgeneralization

〈부록 1〉 영국 아동의 현재분사 발화 빈도 (상위 20위, 괄호 안은 빈도값)

구분	1세	2세	3세	4세	5세	6세	7세	종합 (1-7세)
1	doing (117)	going (6347)	going (714)	going (154)	going (145)	going (24)	going (86)	going (7545)
2	going (75)	coming (1646)	doing (232)	making (29)	playing (54)	playing (18)	playing (48)	doing (1893)
3	coming (56)	doing (1464)	having (104)	playing (28)	running (43)	walking (13)	running (41)	coming (1848)
4	sleeping (45)	go-ing (1242)	making (90)	coming (21)	kneeling (35)	running (12)	kneeling (33)	go-ing (1242)
5	crying (55)	making (694)	coming (88)	doing (18)	walking (34)	crying (11)	walking (32)	making (847)
6	shopping (26)	getting (670)	playing (86)	getting (17)	getting (30)	doing (11)	lying (29)	getting (823)
7	pushing (21)	having (623)	getting (77)	falling (16)	looking (28)	kneeling (11)	standing (26)	having (769)
8	swimming (20)	come-ing (602)	running (59)	putting (13)	doing (27)	sweeping (11)	doing (24)	playing (714)
9	eating (15)	eating (496)	putting (58)	eating (12)	coming (24)	sitting (9)	sweeping (23)	come-ing (602)
10	driving (13)	playing (470)	sitting (53)	looking (11)	sweeping (23)	watching (7)	jumping (22)	eating (583)
11	morning (13)	do-ing (457)	crying (49)	having (10)	putting (22)	holding (6)	pulling (22)	looking (514)
12	falling (12)	looking (418)	walking (41)	taking (9)	holding (21)	jumping (6)	putting (21)	sitting (490)
13	sitting (12)	sitting (387)	coloring (38)	trying (9)	making (21)	reading (5)	watching (21)	do-ing (457)
14	having (10)	hiding (356)	eating (37)	flying (8)	sitting (18)	standing (5)	holding (20)	hiding (407)
15	hiding (10)	building (345)	lying (37)	swimming (8)	standing (18)	getting (4)	getting (19)	crying (398)
16	looking (10)	shopping (336)	trying (35)	walking (8)	watching (18)	looking (4)	kicking (18)	morning (390)
17	playing (10)	morning (335)	hiding (33)	gluing (7)	lying (17)	lying (4)	looking (17)	putting (385)
18	jumping (9)	crying (279)	taking (31)	running (7)	jumping (16)	racing (4)	crying (14)	shopping (378)
19	racing (9)	putting (270)	kneeling (30)	saying (7)	pulling (16)	carrying (3)	carrying (12)	building (354)
20	barking (8)	driving (249)	raining (29)	holding (6)	trying (16)	coloring (3)	writing (12)	running (348)

〈부록 2〉 미국 아동의 현재분사 발화 빈도 (상위 20위, 괄호 안은 빈도값)

구분	1세	2세	3세	4세	5세	6세	7세	종합 (1-7세)
1	doing (562)	going (2848)	going (2640)	going (2144)	going (646)	going (203)	going (154)	going (8857)
2	going (222)	doing (1249)	doing (591)	doing (661)	doing (169)	making (44)	morning (57)	doing (3328)
3	eating (161)	making (520)	making (345)	coming (431)	playing (166)	doing (42)	doing (54)	making (1522)
4	coming (138)	eating (425)	playing (327)	getting (407)	making (149)	getting (39)	making (43)	coming (1450)
5	sleeping (111)	coming (419)	coming (279)	playing (395)	coming (141)	morning (37)	playing (25)	playing (1317)
6	sitting (89)	sleeping (358)	getting (278)	making (349)	getting (92)	playing (36)	eating (24)	eating (1205)
7	making (72)	looking (315)	eating (237)	eating (259)	trying (88)	swimming (33)	coming (23)	getting (1156)
8	crying (69)	playing (310)	trying (198)	talking (258)	talking (84)	talking (23)	talking (22)	looking (818)
9	playing (58)	getting (299)	sleeping (185)	morning (206)	eating (83)	looking (21)	asking (21)	sleeping (776)
10	looking (56)	putting (280)	looking (161)	trying (199)	morning (72)	coming (19)	writing (20)	talking (735)
11	swimming (55)	crying (268)	talking (161)	looking (186)	having (69)	eating (16)	reading (19)	trying (718)
12	holding (46)	sitting (249)	taking (137)	taking (182)	looking (68)	saying (16)	Thanks- giving (17)	taking (670)
13	having (42)	taking (234)	putting (135)	putting (151)	taking (56)	trying (16)	saying (16)	putting (663)
14	running (42)	swimming (209)	crying (120)	having (145)	building (54)	putting (14)	trying (15)	morning (662)
15	nursing (41)	trying (183)	swimming (119)	building (142)	calling (54)	taking (14)	getting (14)	sitting (617)
16	barking (40)	talking (154)	sitting (113)	sitting (117)	putting (45)	jumping (13)	taking (13)	crying (535)
17	writing (37)	working (153)	morning (112)	saying (111)	walking (44)	walking (13)	having (12)	having (518)
18	drinking (35)	morning (152)	walking (112)	running (106)	telling (41)	having (12)	looking (11)	swimming (491)
19	hiding (34)	having (139)	having (99)	flying (96)	sleeping (34)	building (11)	boring (10)	walking (426)
20	taking (34)	writing (139)	flying (82)	walking (92)	driving (33)	sleeping (11)	working (10)	running (372)

〈부록 3〉 영국 아동의 동사 과거형 발화 빈도 (상위 20위, 괄호 안은 빈도값)

구분	1세	2세	3세	4세	5세	6세	7세	종합 (1-7세)
1	gone (580)	got (9631)	got (1176)	got (259)	got (409)	got (115)	got (234)	got (11923)
2	stuck (163)	put (7063)	put (833)	put (168)	put (332)	put (79)	put (224)	put (8787)
3	fit (121)	gone (4978)	come (442)	was (81)	was (179)	was (70)	was (165)	gone (5810)
4	got (99)	come (4430)	bit (320)	finished (76)	could (75)	could (27)	went (100)	come (5116)
5	come (92)	done (2926)	did (283)	come (67)	had (68)	went (24)	had (60)	done (3196)
6	broken (91)	bit (2402)	was (204)	done (63)	did (64)	had (21)	could (58)	bit (2954)
7	put (88)	was (2159)	got_to (194)	did (57)	bit (62)	should (14)	did (48)	was (2863)
8	bit (78)	did (1747)	gone (192)	could (52)	done (47)	might (11)	bit (45)	did (2227)
9	happened (54)	stuck (1429)	let (172)	said (42)	come (40)	come (9)	come (36)	stuck (1654)
10	read (50)	let (1083)	finished (147)	bit (41)	should (37)	did (9)	couldn't (29)	let (1326)
11	done (48)	said (1012)	had (125)	got_to (40)	made (36)	saw (8)	would (25)	said (1177)
12	broke (38)	shut (995)	didn't (112)	didn't (34)	went (33)	fell (7)	saw (24)	broken (1163)
13	shut (25)	broken (994)	would (112)	would (31)	would (31)	made (7)	been (20)	could (1113)
14	wet (21)	called (985)	done (97)	let (27)	finished (30)	run (7)	should (20)	shut (1113)
15	did (19)	fit (863)	cut (85)	made (27)	been (29)	allowed (6)	called (19)	called (1111)
16	cut (18)	could (831)	said (77)	had (26)	didn't (29)	been (6)	hurt (19)	fit (1061)
17	finished (16)	went (784)	shut (77)	went (23)	couldn't (28)	bit (6)	didn't (18)	went (1037)
18	lost (14)	run (761)	went (73)	gone (21)	cut (28)	came (6)	might (18)	didn't (960)
19	run (12)	didn't (759)	called (70)	cut (18)	might (26)	didn't (6)	came (16)	had (909)
20	hurt (11)	read (735)	could (68)	been (17)	called (25)	gave (6)	gone (15)	read (876)

〈부록 4〉 미국 아동의 동사 과거형 발화 빈도 (상위 20위, 괄호 안은 빈도값)

구분	1세	2세	3세	4세	5세	6세	7세	종합 (1-7세)
1	put (687)	put (5055)	put (3615)	was (3277)	was (1365)	was (555)	was (730)	put (13651)
2	read (411)	got (2557)	was (2586)	put (2906)	put (959)	got (293)	ate (442)	was (10402)
3	did (289)	was (1771)	got (2434)	got (2754)	got (774)	put (225)	put (204)	got (9198)
4	come (264)	come (1650)	did (1901)	did (2276)	did (691)	did (209)	did (180)	did (7164)
5	got (251)	did (1618)	come (1651)	come (2102)	come (645)	had (204)	come (150)	come (6598)
6	gone (239)	read (1513)	could (1367)	could (1833)	said (520)	said (183)	had (150)	could (4325)
7	done (175)	let (1089)	didn't (1033)	didn't (1617)	didn't (474)	didn't (178)	said (146)	didn't (4122)
8	broken (166)	didn't (649)	read (915)	said (1295)	could (384)	were (172)	didn't (136)	let (3609)
9	cut (153)	gone (627)	let (786)	let (1214)	ate (341)	went (158)	got (135)	read (3542)
10	stuck (140)	done (595)	said (777)	had (957)	would (336)	come (136)	came (128)	said (3519)
11	wet (138)	went (572)	went (675)	went (860)	had (331)	could (134)	wasn't (116)	had (2799)
12	broke (131)	said (540)	would (669)	would (806)	let (298)	would (97)	could (96)	went (2654)
13	was (118)	hurt (488)	had (663)	were (626)	were (295)	ate (90)	read (82)	would (2289)
14	let (105)	made (464)	were (564)	saw (601)	went (286)	saw (79)	went (74)	were (1994)
15	hurt (94)	fell (462)	saw (461)	made (562)	made (201)	came (71)	called (68)	made (1849)
16	could (85)	had (450)	cut (452)	read (412)	came (164)	let (64)	would (67)	ate (1698)
17	came (76)	broken (446)	made (422)	cut (378)	read (163)	made (62)	made (66)	saw (1666)
18	fell (73)	could (426)	fell (324)	ate (356)	saw (161)	used (47)	were (59)	cut (1549)
19	made (72)	cut (426)	hurt (310)	came (314)	done (149)	read (46)	pushed (56)	done (1540)
20	tired (66)	broke (386)	found (300)	should (306)	should (123)	should (42)	stayed (56)	came (1317)

〈부록 5〉 영국 아동의 비교급/최상급 발화 빈도 (상위 20위, 괄호 안은 빈도값)

구분	1세	2세	3세	4세	5세	6세	7세	종합 (1-7세)
1	more (624)	more (7150)	more (230)	more (48)	more (71)	last (18)	more (21)	more (8146)
2	better (19)	better (1047)	bigger (49)	better (15)	better (25)	faster (7)	biggest (17)	better (1159)
3	last (5)	later (426)	better (44)	last (13)	last (18)	biggest (5)	faster (15)	later (447)
4	bigger (4)	bigger (302)	last (32)	bigger (12)	faster (16)	older (4)	last (11)	bigger (388)
5	higher (2)	last (281)	faster (14)	higher (7)	biggest (14)	most (3)	bigger (10)	last (378)
6	later (2)	faster (130)	later (12)	best (5)	bigger (9)	better (2)	better (7)	faster (186)
7	closer (1)	biggest (46)	bouncier (6)	faster (3)	best (8)	bigger (2)	best (4)	biggest (86)
8	faster (1)	higher (46)	best (5)	later (3)	smaller (5)	fastest (2)	most (3)	best (65)
9	further (1)	harder (45)	higher (3)	louder (3)	older (3)	more (2)	smaller (3)	higher (58)
10	lighter (1)	smaller (45)	taller (3)	softer (3)	sharper (3)	smaller (2)	further (2)	smaller (57)
11	tighter (1)	best (42)	biggest (2)	biggest (2)	darker (2)	best (1)	older (2)	harder (50)
12	n/a	nearer (37)	further (2)	cleaner (2)	fastest (2)	closer (1)	easier (1)	nearer (39)
13	n/a	cleaner (27)	harder (2)	harder (2)	later (2)	highest (1)	fastest (1)	cleaner (29)
14	n/a	longer (24)	littler (2)	easier (1)	warmer (2)	later (1)	later (1)	longer (27)
15	n/a	louder (19)	nearer (2)	fastest (1)	closer (1)	littler (1)	louder (1)	louder (25)
16	n/a	hotter (16)	smallest (2)	further (1)	easier (1)	littlest (1)	oldest (1)	most (21)
17	n/a	closer (15)	brighter (1)	littlest (1)	funner@n (1)	longer (1)	youngest (1)	closer (18)
18	n/a	taller (14)	darker (1)	middlesizest@n (1)	harder (1)	oldest (1)	n/a	taller (18)
19	n/a	most (13)	easier (1)	smaller (1)	littlest (1)	quicker (1)	n/a	older (17)
20	n/a	bestest (12)	farther (1)	smallest (1)	longer (1)	shorter (1)	n/a	hotter (16)



〈부록 6〉 미국 아동의 비교급/최상급 발화 빈도 (상위 20위, 괄호 안은 빈도값)

구분	1세	2세	3세	4세	5세	6세	7세	종합 (1-7세)
1	more (1475)	more (2911)	more (1423)	more (1594)	more (495)	more (163)	more (150)	more (8211)
2	better (47)	better (426)	better (414)	better (462)	better (162)	last (39)	better (95)	better (1641)
3	cleaner (39)	bigger (230)	bigger (205)	last (235)	last (106)	better (35)	last (35)	bigger (795)
4	higher (33)	later (149)	last (149)	bigger (229)	bigger (93)	bigger (17)	best (22)	last (669)
5	bigger (14)	last (97)	later (76)	best (96)	biggest (40)	best (12)	most (9)	later (337)
6	later (14)	best (86)	best (55)	later (69)	most (34)	most (10)	bigger (7)	best (297)
7	biggest (10)	faster (63)	faster (36)	faster (46)	best (24)	least (9)	easier (6)	faster (167)
8	last (8)	biggest (24)	biggest (32)	most (45)	later (24)	older (8)	lower (6)	biggest (143)
9	older (4)	higher (23)	older (24)	longer (43)	least (17)	higher (7)	upper (4)	higher (136)
10	harder (3)	smaller (19)	smaller (22)	harder (39)	faster (16)	faster (6)	busier (3)	most (130)
11	littlest (3)	louder (16)	closer (17)	higher (39)	higher (16)	worse (5)	greatest (3)	smaller (88)
12	best (2)	cleaner (15)	most (17)	smaller (38)	longer (14)	easier (4)	higher (3)	older (86)
13	most (2)	most (13)	taller (17)	biggest (32)	older (12)	greatest (4)	biggest (2)	longer (80)
14	easier (1)	older (12)	cleaner (16)	older (24)	stronger (11)	littler (4)	colder (2)	cleaner (76)
15	longer (1)	closer (11)	higher (15)	taller (21)	hotter (6)	biggest (3)	foggier (2)	harder (54)
16	louder (1)	further (11)	stronger (14)	closer (13)	quiester (6)	later (3)	later (2)	taller (49)
17	messier (1)	hotter (7)	longer (12)	least (13)	worst (6)	littlest (3)	least (2)	least (45)
18	smaller (1)	lower (7)	farther (10)	lower (11)	highest (5)	oldest (3)	longer (2)	closer (43)
19	whiter (1)	highest (6)	longest (10)	stronger (11)	prettiest (5)	smaller (3)	older (2)	louder (39)
20	n/a	littler (6)	littlest (9)	warmer (11)	richer (5)	farther (2)	prettier (2)	stronger (37)