

영미 아동의 비교급 사용과 습득: CHILDES 데이터베이스 사용기반 분석

연령별 비교급 사용 양상



한국언어정보학회 월례발표회 **2017. 3. 25**

한국외국어대학교 언어인지과학과

서지혜 (SUH.BWISE@GMAIL.COM)

지도: 전종섭 교수님

목 차



1. 서론

1.1 연구의 배경

1.2 연구의 주제 및 목적

2. 선행연구

3. 비교급 관련 변수

4. 연구 방법

5. 결과

6. 논의

참고문헌

1. INTRODUCTION



영어 비교급

handsomer, **more** handsome; commoner, **more** common

- 형용사 + *-er* [Synthetic, S-comparatives: **ER** 비교급]
John is **bigger** than Bill.
- *more* + 형용사 [Periphrastic, P-comparatives: **MORE** 비교급]
Mary is **more beautiful** than Sue.

(Otaki, 2010; Quirk et al. 1985)

1. INTRODUCTION



1. 연구의 배경

- **MORE** 비교급과 **ER** 비교급은 의미는 같지만 두 가지 다른 형태로 존재하는 문법 요소이다.
 - ⇒ 형태와 구문 습득을 살펴볼 수 있다.
- **CHILDES database**
 - ⇒ 아동이 초기부터 사용하는 어휘와 형태소의 습득과 처리를 살펴볼 수 있다.
- 아동이 직접 발화한 비교급 자료를 분석하여, 각 문장에 대한 언어적 요인과 빈도, 연령 요인을 종합적으로 고려할 필요가 있다.

Q: 영미 아동의 비교급 습득/사용은?

2. LITERATURE REVIEW



사용기반 이론 (Usage-Based theory)

<표 1> 사용기반 이론

구분	사용기반 이론
핵심사상	패턴 찾기
연구자	Tomasello
구조	형식-기능 연결
심성 어휘	항목 기반, 빈도에 따른 강약
시스템	습관화, 유추, 확장
인지	명시화(Salience)
사회 상호작용	입력과 상호작용(필수), 의도 파악
과정	점진적
아동/성인 문법	다름

2. LITERATURE REVIEW



아동의 비교급 선호도 – 실험 연구

- Graziano-King(1999): 성인 vs. 4세, 7세 아동; 4세: MORE > ER
- Bobaljik(2007): 4세, 7세 아동의 선호도 (*old, cold, short, long*)
4세 아동: 60% MORE 선호, 7세 아동: 12.5% ER 선호

아동의 비교급 습득 – 코퍼스 연구

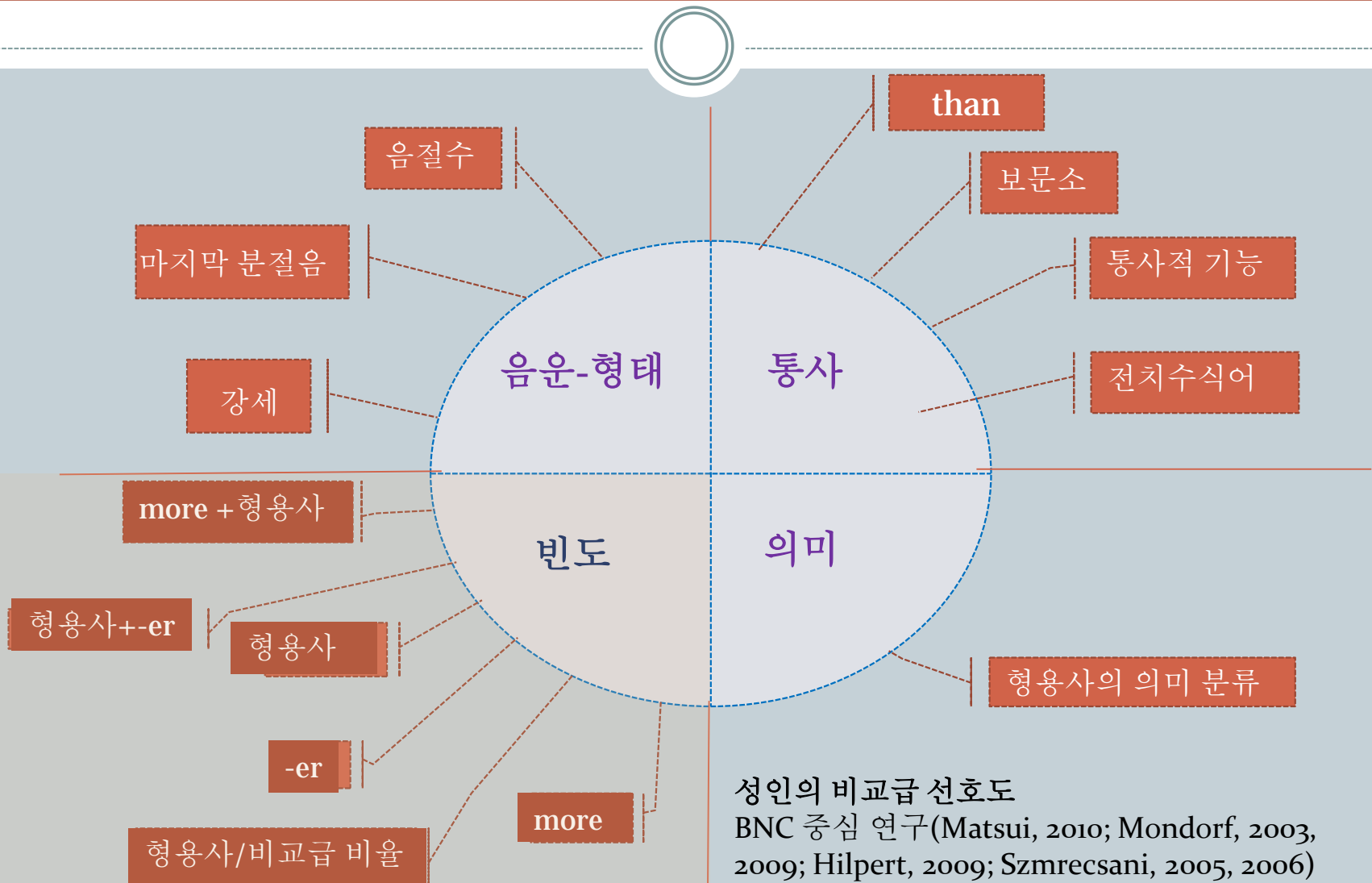
- Otaki(2010): 형용사 원형/ER 비교급 비율
 - MORE 비교급의 과잉일반화는 발화하지 않음; ER > MORE
- 서지혜, 전종섭, 이선영(2013): ER > MORE
 - ER (1세) 처음 사용 vs. MORE (2세) 처음 사용
 - Hohaus & Tiemann(2009)의 결과(ER 2;5, 2;10; MORE 3;5, 3;11)에 비해 빠름

RESEARCH QUESTION



- 실험연구에서 4세 아동은 MORE 비교급을 더 빈번하게 사용하였지만, ER 비교급을 더 정확하게 사용하였다 (Graziano-King, 1999). 반면에 코퍼스 연구에서 아동은 ER 비교급을 먼저 발화하고 MORE 비교급을 발화하는 경향을 보였다 (Hohaus & Tiemann, 2009; Suh et al., 2013).
- CHILDES 데이터베이스의 영국과 미국 아동의 발화 자료에서 4세와 7 아동의 비교급 사용 양상은 어떠할까?
- 아동의 연령별 비교급 사용이 다르다면, 어떤 언어 요인과 빈도 요인에서 비교급 사용의 차이가 나타나는가?

3. VARIABLES



성인의 비교급 선호도
 BNC 중심 연구(Matsui, 2010; Mondorf, 2003, 2009; Hilpert, 2009; Szmrecsani, 2005, 2006)



음운-형태적 변수

1) 음절 수 (Graziano-King & Cairns, 2005; Hilpert, 2008; Mondorf, 2003)

1, ≥ 2

- ① 1음절: big, bright, cheap, clean, little, nice, short, straight, strange
- ② 2음절: easy, cracky, curly, happy, healthy, special
- ③ 3음절 이상: beautiful, difficult, important, interesting

2) 강세 (Hilpert, 2008; Mondorf, 2003) one, (first, second)

- ① 단음절 big, brave, deaf, fast, rough, smooth
- ② 첫 음절 beautiful, broken, busy, creamy, difficult, hungry, special
- ③ 둘째 음절 alive, important



음운-형태적 변수

3) 마지막 분절음 (Hilpert, 2008; Lindquist, 2008; Mondorf, 2003)

/l/, /i/, 자음, 모음

- ① -l/-le: beautiful, full, little, purple, small, special, tall
- ② -y/-ly: easy, fluffy, funny, happy, noisy, pretty, scratchy, curly
- ③ 자음: big, hard, green, long, short, wide, fair, sour, broken, haunted, interesting, speckled, walking
- ④ 모음: blue, extra, high, yellow



통사적 변수

1) 보문소 (Hilpert, 2008; Szmrecsanyi, 2006) **x, PP, infinitive, clause**

① 아무것도 없는 경우

*CHI: the [: house] it feels bigger. (VanKleeck_walter)

② 전치사구(PP)

*CHI: and then when it gets easier for me. (MacWhinney_35a2)

③ to 부정사(infinitive)

*CHI: I'm bigger to ride on my bicycle. (Hall_WhitePro_zoe)

④ 절

*CHI: I want to be stronger than they are. (Garvey_valabe)



통사적 변수

2) 통사적 기능 (Hilpert, 2008; Booth & Waxman, 2009)

① 한정적 (**attributive**)

*CHI: my older brother used to be in this school.

Age 5 (Fletcher7)

② 서술적 (**predicative**)

*CHI: car bigger.

Age 1 (Manchester_Domin)

*CHI: that's straighter.

Age 2 (Manchester_Gail)



통사적 변수

- 3) *Than* (Arii, Syrett, & Goro, 2014; Mondorf, 2003) o, x
- ① CHI: faster than you. Age 2 (Manchester_Joel)
CHI: your eyes are bigger then your belly, aren't they?
Age 3 (Manchester_Nick)

- 4) 전치수식어 (Naigles 1990; Syrett & Lidz, 2010) o, x
- ① 숫자: one, two, nineteen, a month, a year, two times, two years, five inches, seven minutes, twenty feet
- ② any, anymore ③ some, some more ④ lots, a lot
- ⑤ little bit, bit, a bit, a tiny bit, just a bit, ~bit, little, a little, ~little
- ⑥ much, much too, really much, so much, how much
- ⑦ 부사: almost, ~enough, even, definitely, far, generally, just, nearly, really, slightly, so, still, ~still, too, very
- ⑧ 그 외: ~ago, all, few, funny, kinda, pretty big, this, too sort of, way



의미적 변수

1) 형용사의 의미 분류 (Blackwell, 1999, 2009; Dixon, 1980)

- ① 나이 old, young
- ② 색깔 black, blue, green, pink, red, white, yellow
- ③ 규모 big, deep, large, little, long, short, small, tall, thick, wide
- ④ 위치 close, high
- ⑤ 속도 fast, quick, slow
- ⑥ 가치 great, gross, fair, interesting, important, neat, nice, safe, special, strange
- ⑦ 인간적 속성 alive, bossy, busy, funny, happy, healthy, hungry, lucky, mean, pleasant, scary, silly, smart, strong, weird
- ⑧ 물리적 속성 beautiful, bright broken, clean, cold, cracky, creamy, crunchy, curly, dark, deem, dirty, dry, fat, foggy, fluffy, full, heavy, hot, light, loud, loose, noisy, pretty, rocky, scratchy, screwy, sharp, shiny, smooth, soft, solid, sour, speckled, straight, sweet, tight, ugly, walking, warm, yummy
- ⑨ 그 외 different, difficult, easy, extra, hard, haunted, rich, tough



빈도변수

- 1) 형용사 (Quirk et al., 1985) 고빈도: ER 비교급 ↑
- 2) *-er* (Clark, 1982)
- 3) *more* (Clark, 1982)
- 4) ER 비교급 (Szmrecsanyi, 2005, 2006)
- 5) MORE 비교급 (Szmrecsanyi, 2005, 2006)
- 6) 형용사/비교급 비율 (Mondorf, 2003; Otaki, 2010)

연령

1세, 2세, 3세, 4세, 5세, 6세, 7세

4. RESEARCH METHOD



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

- 형태소 태깅된 영국과 미국 - 영어 코퍼스

<표 2> 파일과 아동 수

연령	영국		미국	
	파일	아동	파일	아동
1	181	60	321	188
2	885	62	864	315
3	156	66	617	229
4	51	24	325	106
5	57	28	112	44
6	6	7	87	81
7	19	19	31	15
합계	1,355	-	2,357	-

<표 3> 평균발화길이(MLU) & 발화 문장 수

연령	영국		미국	
	MLU	발화	MLU	발화
1	1.802	32,257	1.886	38,093
2	2.615	223,330	3.353	167,236
3	3.565	19,262	4.557	104,774
4	4.171	6,877	4.841	116,423
5	4.695	8,621	5.469	18,056
6	5.230	1,910	5.936	11,147
7	5.148	5,238	4.483	4,007
평균/합계	3.889	297,495	4.361	459,736

연령별 교차분석 비교

<표 4> 연령별 교차분석 결과



연령	문장 수	χ^2	df	p	
3세 vs 6세	304	4.327	1	.038	✓
3세 vs 7세	284	1.138	1	.286	
4세 vs 6세	426	1.307	1	.309	
4세 vs 7세	405	0.117	1	.732	
3세 vs 5세 ~ 6세	426	3.382	1	.066	
3세 vs 5세 ~ 7세	468	4.084	1	.043	✓

<표 5> 3세와 5세 ~ 7세 아동의 비교급 발화

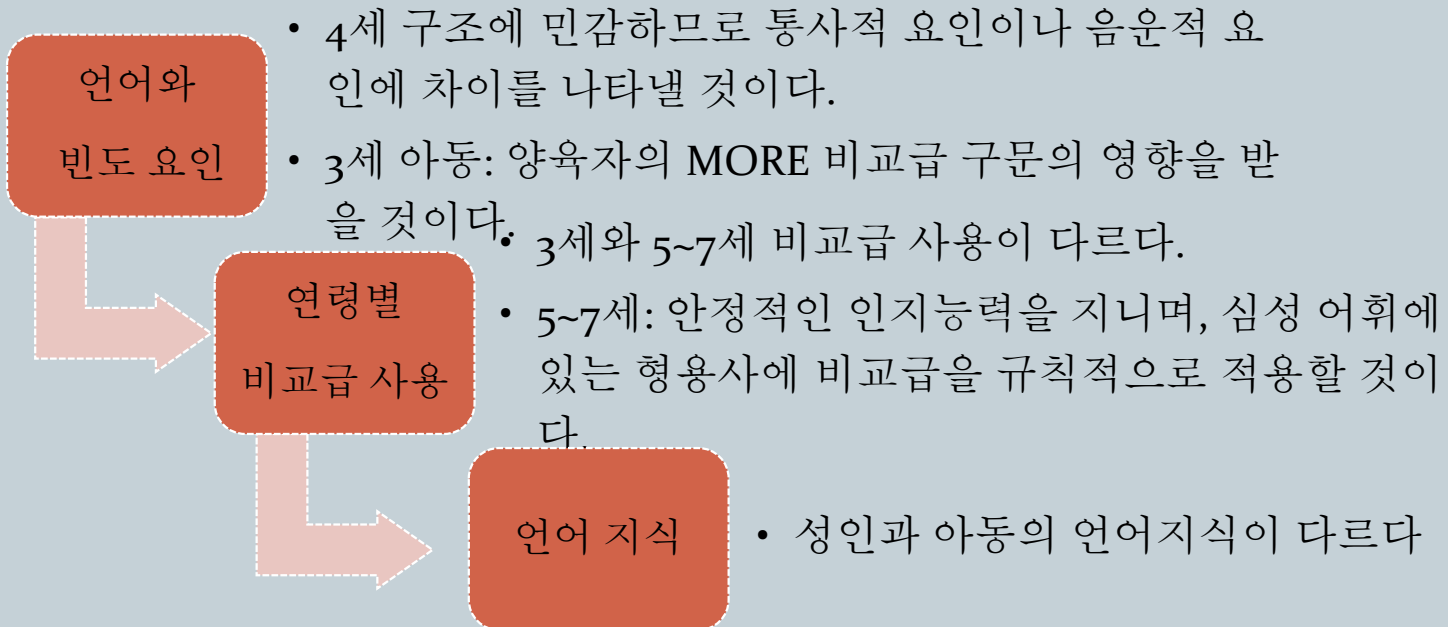
연령	3세	5세 ~ 7세	합 계
MORE 비교급	16	6	22
ER 비교급	220	226	446
합 계	226	242	468

RESEARCH QUESTIONS



- **CHILDES** 데이터베이스에서 **4세**와 **7세**를 대신하여, 비교급 사용에 차이를 나타내는 **3세**와 **5세~7세** 아동의 비교급 사용은 어떻게 다른가?
- **Graziano-King (1999)**의 연구에서 **4세** 아동이 **MORE** 비교급을 더 빈번하게 사용하지만 **ER** 비교급을 더 정확하게 사용한 것처럼, **3세** 아동도 **MORE** 비교급을 더 빈번하게 사용하지만 **MORE** 비교급보다 **ER** 비교급을 정확하게 발화하는가?
- 실험연구와 마찬가지로, **5세~7세** 아동의 **MORE** 비교급 사용이 **3세**보다 적은가?
- **3세** 아동과 **5세~7세** 아동의 비교급 사용이 다르다면, 어떤 언어 요인과 빈도 요인에서 차이가 나타나는가?

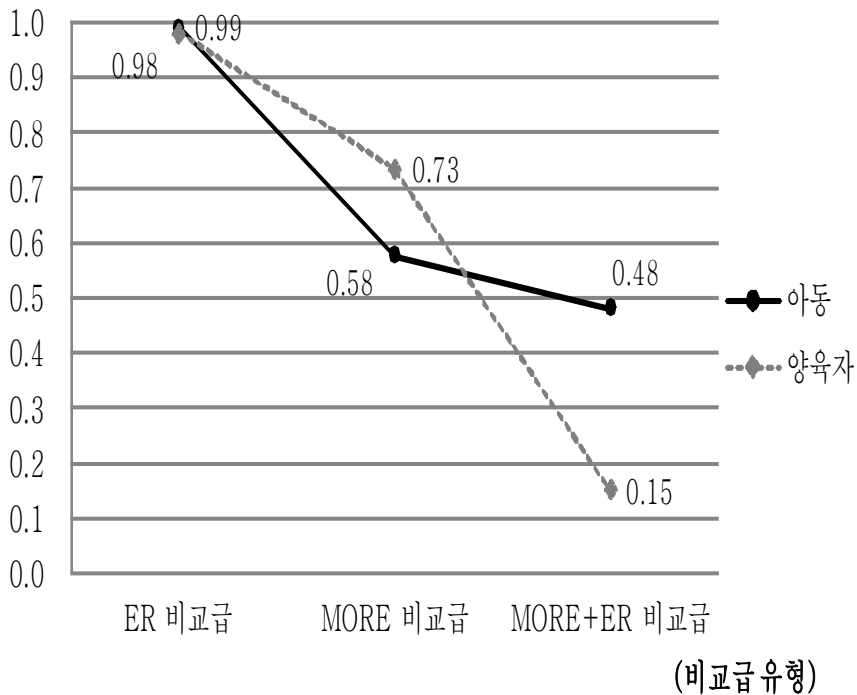
HYPOTHESES



상대빈도

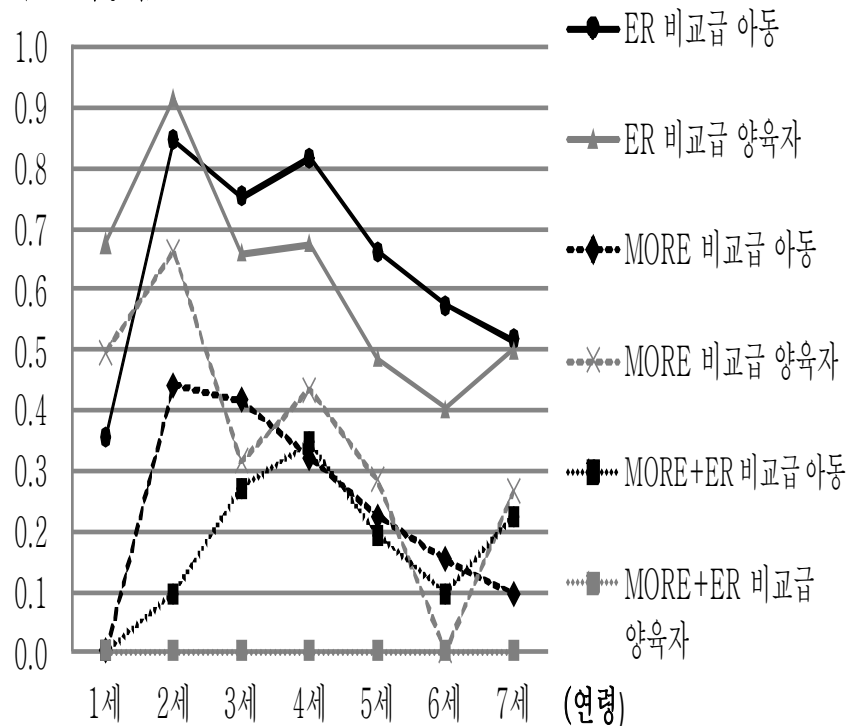


(로그 가중치)



<그림 1> 아동과 양육자의 유형별 비교급 발화

(로그 가중치)

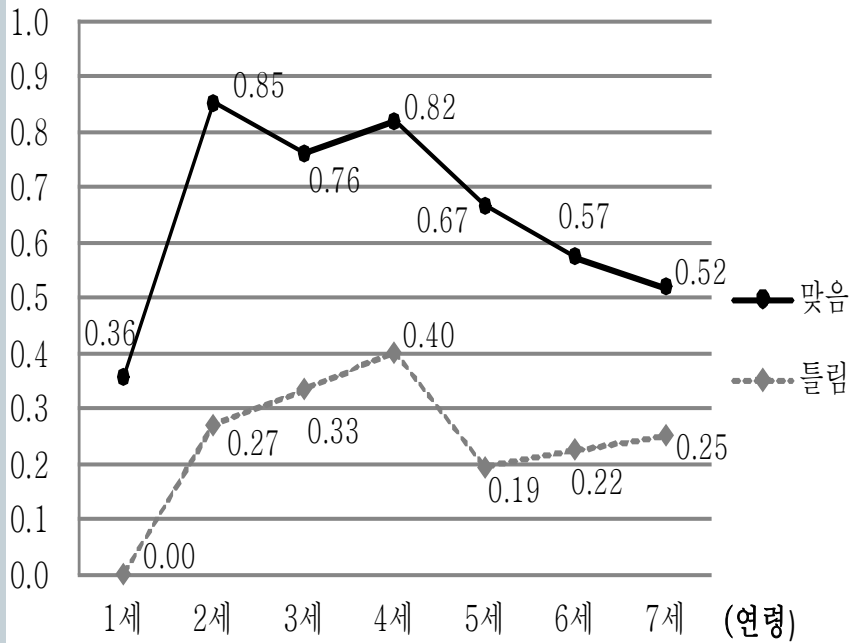


<그림 2> 아동과 양육자의 비교급 발화

상대빈도

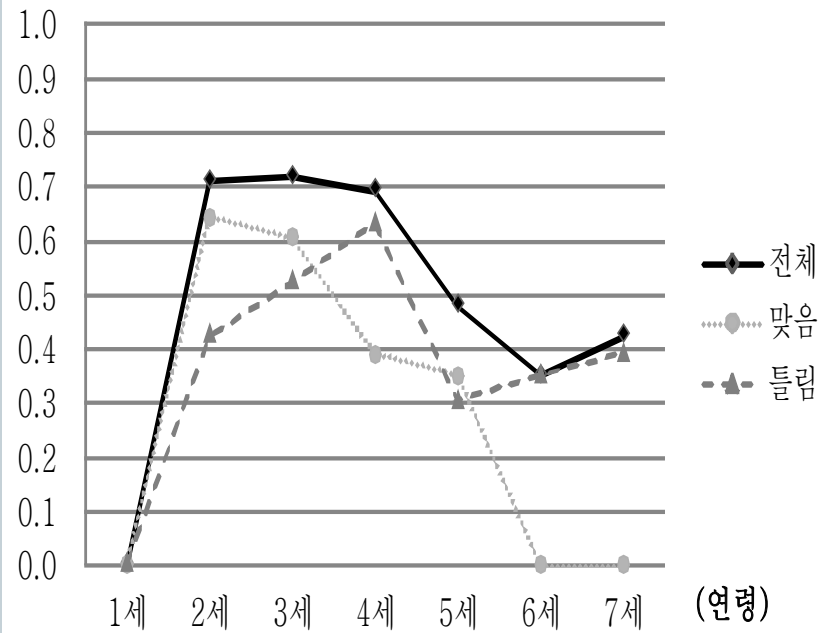


(로그가중치)



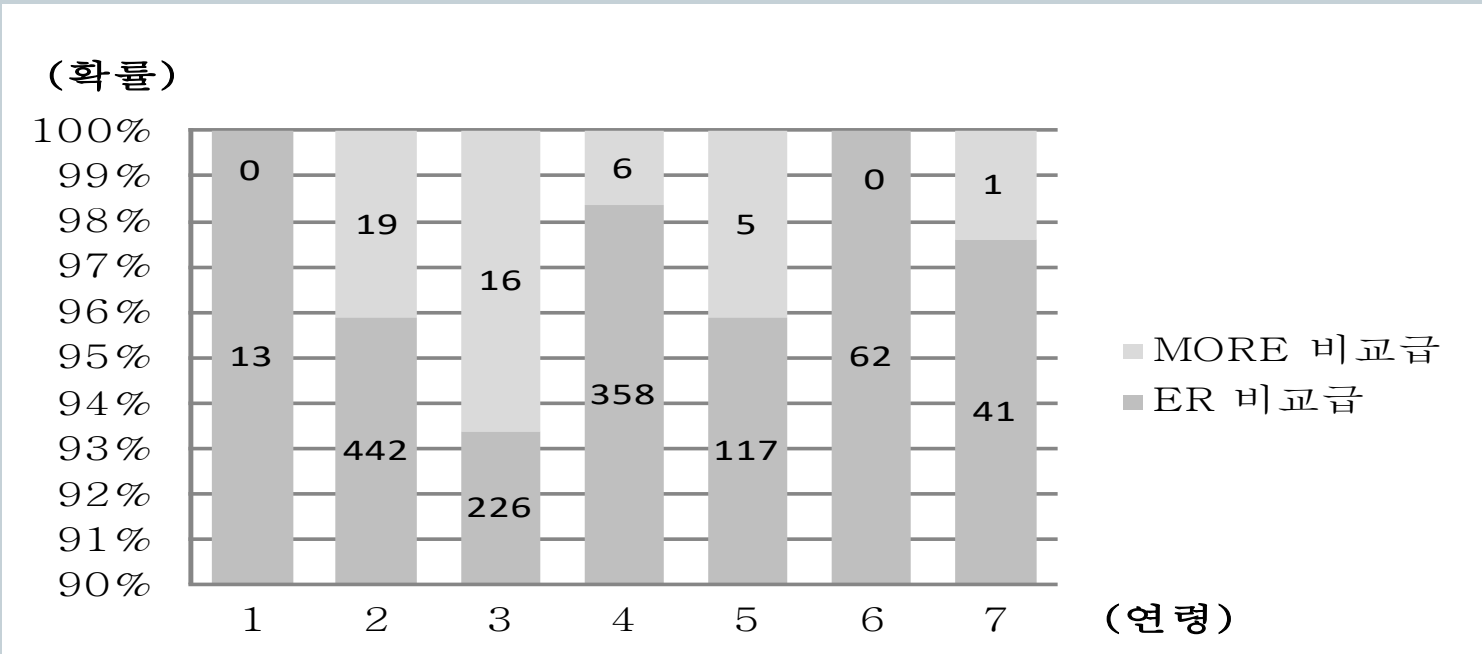
<그림 3> 아동의 비교급 발화

(로그가중치)



<그림 4> 아동의 More 비교급 발화

상대빈도



<그림 7> 아동의 연령별 올바른 비교급 사용 비율



- **CHILDES** 데이터베이스에서 4세와 7세를 대신하여, 비교급 사용에 차이를 나타내는 3세와 5세~7세 아동의 비교급 사용은 어떻게 다른가?
- **Graziano-King (1999)**의 연구에서 4세 아동이 **MORE** 비교급을 더 빈번하게 사용하지만 **ER** 비교급을 더 정확하게 사용한 것처럼, 3세 아동도 **MORE** 비교급을 더 빈번하게 사용하지만 **MORE** 비교급보다 **ER** 비교급을 정확하게 발화하는가?
- 실험연구와 마찬가지로 5세~7세 아동의 **MORE** 비교급 사용 비율이 줄어드는가?

사용 ⇒ 3세, 5~7세: **ER** 비교급 > **MORE** 비교급

⇒ **MORE** 비교급: 3세 > 5~7세 (20:10)

정확도 ⇒ **ER**: **MORE**=(3세)226:16; (5~7세)220:6

MORE 정확률=6.3%:2.4%

⇒ **MORE** 오류율 (3세: 5세~7세=1.6%:1.6%)

RESEARCH QUESTIONS



- 3세 아동과 5세 ~ 7세 아동의 비교급 사용이 다르다면, 어떤 언어 요인과 빈도 요인에서 비교급 사용의 차이가 나타나는가?

RESULTS

<표 6> 3세와 5세~7세의 언어변수

언어 변수	연령	χ^2	df	p		언어 변수	연령	χ^2	df	p
	3세	9.499	1	.002			3세	15.547	1	.000
음절 수	5세~7세	24.903	1	.000		통사적 기능	5세~7세	48.679	1	.000
	총	32.044	1	.000			총	68.125	1	.000
(F) 마지막 분절음	3세	2.082	3	.556		(T)	3세	2.382	1	.123
	5세~7세	46.531	3	.000	√	than 유무	5세~7세	3.914	1	.048
	총	36.894	3	.000			총	6.926	1	.008
	3세	9.499	1	.002		(P)	3세	4.213	1	.075
강세	5세~7세	24.903	1	.000		전치 수식어	5세~7세	6.964	1	.027
	총	32.044	1	.000			총	7.454	1	.006
	3세	6.640	3	.084		형용사의 의미 분류	3세	96.007	8	.000
보문소	5세~7세	2.259	3	.520			5세~7세	199.313	8	.000
	총	3.116	3	.374			총	330.099	8	.000

RESULTS

<표 7> 로짓 로그선형 회귀분석 결과

#	모형	G2	df	p	#	모형 (계속)	G2	df	p
14	TA_AP_moreAdj	20.234	24	.843	23	AP_Adjer	25.684	26	.481
15	AP_TP_Adjer	19.480	24	.726	24	TP_Adjer	32.776	26	.169
16	AT_PT_Adjer	19.506	24	.725	25	A_T_Adjer	19.506	27	.624
17	AP*R_TP	2.652	20	1.000	26	A_P_Adjer	25.979	27	.520
18	AP_T_Adjer	19.48	25	.774	27	AP*moreAdj_TP	.000	20	1.000
19	AT_P_Adjer	19.506	25	.772	28	AP_T_moreAdj	20.234	25	.734
20	TP_A_Adjer	19.506	25	.772	29	AT_P_moreAdj	24.177	25	.509
21	T_A_P_Adjer	19.505	25	.814	30	AP_moreAdj	27.056	26	.406
22	AT_Adjer	24.105	26	.570	31	TP_moreAdj	27.392	26	.389

RESULTS

<표 8> 로짓 로그선형 모수 추정값 {AT} {P} {Adjer}

상수 + 비교급 사용 + 비교급 사용*Adjer +
비교급 사용*전치수식어(P) + 비교급 사용*연령(A)*Than(T)

	모수	추정값	S. E.	Z	p	95% 신뢰구간	
						하한	상한
{비교급 사용}	{MORE}	-9.663	2.399	-4.028	.000	-14.365	-4.962
	{ER}	0 ^b					
{비교급 사용} {Adjer}	{MORE}{Adjer}	.003	.010	.306	.760	-.017	.023
	{ER}{Adjer}	.087	.026	3.325	.001	.036	.139
{비교급 사용} {연령}{Than}	{MORE}{3세}{than x}	9.096	1.952	4.659	.000	5.270	12.923
	{MORE}{5~7세}{than o}	-8.969	2880.0	-.003	.998	-5653.7	5635.7
	{MORE}{5~7세}{than x}	0.000					
{비교급 사용} {전치수식어}	{MORE}{전치수식어 o}	-1.535	.674	-2.28	.023	-2.856	-.215
	{ER}{전치수식어 x}	0 ^b					

DISCUSSION



- 연령별 비교급 사용에 영향을 준 언어 요인
30개월(2;6) - 문장구조 정보 이용하여 새로운 정보 해석
- 마지막 분절음: 연령별 차이 구체적 연관성은 발견되지 않음
- **Than**의 유무: 3세 아동- *than* 없는 MORE 비교급을 더 발화함;
문맥에서 비교 대상 파악 용이; 문맥, 형용사의 의미에 더 집중
- 전치수식어: MORE 비교급에 전치수식어는 덜 사용함;
MORE의 *more*와 충돌하기 때문일 것으로 예상;
- 연령별 비교급 사용에 영향을 준 빈도 요인
- 양육자의 ER 사용 빈도가 아동의 ER 발화에 미미하게 영향

DISCUSSION




➤ 비교급 습득

- 단음절 형용사: ER 비교급을 많이 사용
- 사용, 정확률: ER > MORE – 아동의 ER에 대한 지식이 더 확고함
- 3, 4세에 MORE 비교급 선호 (Graziano-King의 결과와 일치)

➤ 사용기반 이론

- 성인과 아동의 언어 지식이 다르다는 것을 보여줌
- 아동이 언어 요인 뿐 아니라 양육자의 빈도 요인의 영향을 받아 비교급을 사용하고 습득하는 사용기반 이론 (usage-based theory)의 증거를 제시함
- 아동이 양육자나 공동체와의 상호 작용을 통해 비교급 구문을 구체화하며 사용/습득함

References

- 
- Bobaljik, J. D. (2007). *On comparative suppletion*. (Unpublished master's thesis, Ms.). University of Connecticut, Storrs, CT.
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press.
- Curme, G. O. (1931). *Syntax*. Boston: D. C. Heath and Company.
- Graziano-King, J. (1999). *Acquisition of comparative forms in English*. (Unpublished doctoral dissertation). University of New York, NY.
- Hilpert, M. (2008). The English comparative - language structure and language use. *English Language and Linguistics*, 12 (3), 395-417.
- Hohaus, V., & Tiemann, S. (2009). "...this much is how much I'm taller than Joey..." A Corpus Study in the Acquisition of Comparison Constructions. In *Proceedings of AquisiLyon*, 90-93.
- Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. (2nd Ed.) Cambridge: Cambridge University Press.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.
- Matsui, C. (2010). *-er* type or *more* type adjectives of comparison in English. *Studia Neophilologica*, 82 (2), 188-202.
- Mondorf, B. (2003). Support for more-support. In: Günter Rohdenberg and Britta Mondorf (eds.). *Determinants of Grammatical Variation in English* (Topics in English Linguistics 43), pp. 251-304. Berlin: Moutonde Gruyter.
- Otaki, K. (2010). The Syntax-Lexicon Interface in the Acquisition of English Comparative Morphology. In *Proceedings of the 2009 Mind/ Context Divide Workshop*, pp.96-103.
- Suh, J-H., Jun, JS., & Lee, S-Y. (2013). English-speaking children's acquisition of comparatives: Based on the CHILDES database. *Korean Journal of Linguistics*, 38(2), 345-366.
- Szmrecsanyi, B. (2005). Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory*, 1(1), 113-150.
- Szmrecsanyi, B. (2006). Morphosyntactic Persistence in Spoken English: A Corpus Study at the Intersection of Variationist Sociolinguistics, Psycholinguistics and Discourse Analysis. Berlin and New York: Mouton de Gruyter.