Errors of redundancy in child English past tense formation

Johannes Hein[2], Imke Driemel[2], Fabienne Martin[2], Yining Nie[3] & Artemis Alexiadou[1,2]

[1] ZAS Berlin — [2] Humboldt University of Berlin — [3] San José State University SJSU

Introduction

- English-learning children occasionally produce overregularization errors (1a, b) or overtensing/doubling errors (1c) with irregular verbs (Kuczaj 1977, 1978; Stemberger 1982, 2007; Marcus et al. 1992; Maratsos 2000; Hattori 2003).

(1) a. *Distributive error* I **eated** an apple. b. *Redundant error* I **ated** an apple. c. *Periphrastic error* I **did ate** an apple.

- Distributive errors like (1a) have led researchers to propose that **children prefer a 1-to-1 mapping between form and meaning** (Slobin 1985, Brighton et al. 2005, van Hout 2008, Guasti et al. 2023). How do redundant/periphrastic (1b/c) errors fit in?
- Redundant/distributive errors are also reported for child French causatives and comparatives (Bezinska et al. 2008; Martin et al. 2022), child English comparatives (Hein et al. 2022) and child German/English negative indefinites (Hein et al. 2023, Driemel et al. 2023), but relative frequencies may be confounded by language-specific properties.

Goals: 1. Determine the error types' frequencies for English past tense errors.
2. Provide unified analysis for different error types across domains and languages that accounts for relative frequencies.

Corpus study

- Previous studies either compare different error types of a subset of verbs across limited corpora (Kuczaj 1977, Marcus et al. 1992) or focus on one error type across different verbs in a larger number of corpora (Stemberger 2007).
- We conducted a corpus study on **all typically developing children aged at least 1;01** from **39 North American English** and **17 British English** corpora available through the ChiLDES database (MacWhinney 2000)
- We ran a query for past tense forms of **37 irregular verbs** within the 100 most frequent verbs in English ChiLDES, including distributive and redundant error forms in various orthographic variants.
- We excluded the homographs *cut, read, let, put, fit, hit* and by accident also *buy/bought*.
- Hits were annotated for target (TAR) or error type (DIS, RED, PER_DO, PER_DID)
- Participles that are syncretic with the past tense were excluded.

Results

(2) *Overall error counts*

Type	N	%
TARGET	100,674	97.19
NON-TARGET	2,916	2.81
DISTRIBUTIVE	1,771	1.71
REDUNDANT	382	0.37
PERIPHRASTIC	416	0.40
did	365	0.35
do	51	0.05
OTHER	347	0.33

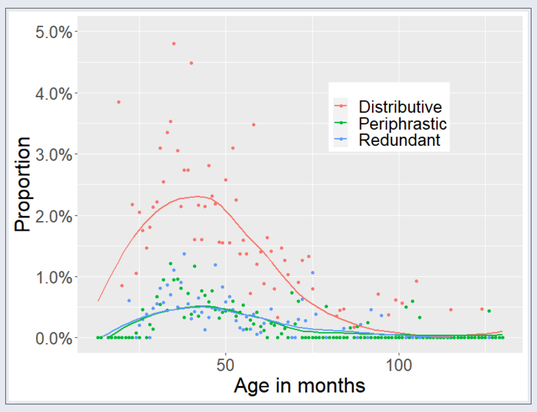

Fig. 1: Error rates over age

(3) *Examples of redundant errors*

a. so elephant **wented** [: went] [*] and got a ride . (Laura, 2;05, Braunwald)
b. he broke [*] [= actually says **broked**] it ? (Fraser, 2;06, MPI-EVA)
c. the workers **builted** [: built] [* m] it . (Stuart, 4;01, Belfast)

(4) *Examples of periphrastic errors*

a. I **did**n't **caught** it &-uh (.) one . (Sarah, 3;03, Brown)
b. **does** it **fell** [*] into the water ? (Lara, 2;11, Lara)
c. I **do made** [*] the shopping . (Becky, 2;09, Manchester)
d. (.) <why didn't he> [//] why **did** he **ate** [!] her ? (Geoffrey, 3;08, HSLLD)

(5) *Examples of distributive errors*

a. Bill **gived** [: gave] [*] me a ride in the motorcycle . (Peter, 2;05, Bloom)
b. it **falled** [: fell] [* +ed] in the briefcase . (Eve, 1;10, Brown)
c. he **runned** [: ran] . (Helen, 4;11, Gleason)

Children produce the correct forms before or alongside errors.

Selected References: Arregi & Pietraszko (2021). The Ups and Downs of Head Displacement. *Linguistic Inquiry 52*: 241–289. • Driemel, Hein, Bill, Gonzalez, Ilić, Jeretič & van Alem (2023). Negative concord and negative indefinites: Insights from commission errors. Ms., Humboldt University of Berlin. • Guasti, Alexiadou & Sauerland (2023). Undercompression errors as evidence for conceptual primitives. Ms., University of Milano-Bicocca, ZAS Berlin. • Hein, Driemel, Martin, Nie & Alexiadou (2022). Errors of Multiple Exponence in Child Language. *WCCFL 40 Proceedings*. • Kuczaj (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior 16*: 589–600. • MacWhinney (2000). *The CHILDES Project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates. • Martin, Nie, Alexiadou & Guasti (2022). Wearing Causation on Its Sleeve: Overt *cause* in Child French Causatives. *Proceedings of BUCLD 46*. Somerville, MA: Cascadilla Press, 497–510. • Slobin (1985). *The Cross-linguistic Studies of Language Acquisition. Vol. 2: Theoretical Issues*, 406–605. Hillsdale: Lawrence Erlbaum Associates. • Stemberger (2007). Children's overtensing errors: Phonological and lexical effects on syntax. *Journal of Memory and Language 57*: 49–64.

Acknowledgements: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (LeibnizDream, grant agreement No 856421).

Analysis

Generalized Head Movement (Arregi & Pietraszko 2021)

(6) $[_{YP}\ Y^{hm}_{[M:\,Y_m]}\ [_{XP}\ X_{[M:\,X_m]}\ \dots]] \rightarrow [_{YP}\ Y^{\cancel{hm}}_{[M:\,\_]}\ [_{XP}\ X_{[M:\,\_]}\ \dots]]$

$[_{Y_m}\ X_m\ Y_m\ ]$

(7) *GenHM in English verbs*
$[_{CP}\ C\ [_{TP}\ DP\ [_{T'}\ T^{hm}\ (Adv)\ [_{VP}\ V^*\ DP\ ]]]]$

$[_{T_m}\ V_m\ T_m\ ]$

Head Chain Pronunciation:
Delink all positions in a head chain except
a. the highest strong position, if any;
b. otherwise, the highest position.

Deriving children's redundant and distributive errors

Children occasionally ignore secondary features during Vocabulary Insertion.
This is an implementation of the bias for 1-to-1 mapping (Slobin 1985, Guasti et al. 2023).

(8) *Local errors*
$[_{T_m}\ \text{EAT}\ T^{[PST]}_m\ ]$

			#	location	type
a.	/ate/	/-∅/	0	—	target
b.	/ate/	/-ed/	1	T_m	redundant
c.	/eat/	/-ed/	2	V_m & T_m	distributive
d.	/eat/	/-∅/	1	V_m	omissive

(9) *Vocabulary Items in English past tense*
a. /eat/ ⇔ $[\sqrt{\text{EAT}}]$
b. /ate/ ⇔ $[\sqrt{\text{EAT}}]$ / __[PST]
c. /-ed/ ⇔ [PST]
d. /-∅/ ⇔ [PST] / __$[\{\sqrt{\text{EAT}}, \sqrt{\text{BRING}}, \dots\}]$

***Do*-support in Generalized Head Movement**

(10) *Split-by-Intervention* (Arregi & Pietraszko 2021, 261)
In a head chain terminating in V* such that a specifier marked [+P] intervenes between the top of the chain and V*, split the chain at V*.

(11) *Orphan Assignment* (Arregi & Pietraszko 2021, 261)
Assign [O] to morphological terminal X_m in a head chain that does not contain the syntactic terminal X.

(12) *Subject-Auxiliary Inversion in English*
$[_{CP}\ C\ [_{TP}\ DP_{[+P]}\ [_{T'}\ T\ [_{VP}\ V^*\ DP\ ]]]]$

$[_{C_m}\ [_{T_m}\ V^{[O]}_m\ T_m\ ]\ C_m\ ]$ $[_{C_m}\ [_{T_m}\ V_m\ T^{[O]}_m\ ]\ C^{[O]}_m\ ]$

(13) Do-*support in past tense*
$[[\ \text{EAT}^{[O]}\ T^{[PST]}_m\ ]\ C_m\ ]$ $[[\ \text{EAT}\ T^{[O,PST]}_m\ ]\ C^{[O]}_m\ ]$
/did/ /-∅/ /-∅/ /eat/ /-∅/

(14) *Vocabulary items for* do*-support*
a. /do/ ⇔ $[V_m, O]$
b. /did/ ⇔ $[V_m, O]$ / __[PST]
c. /-∅/ ⇔ [PST] / __$[\{\sqrt{\text{EAT}}, \sqrt{\text{DO}}, \dots\}]$
d. /-∅/ ⇔ $[C_m]$

(15) *Obliteration rule*
$T^{[O]}_m \rightarrow \varnothing$

Deriving children's periphrastic errors

Children occasionally fail to obliterate $T^{[O]}_m$, which can then condition stem allomorphy on V_m.

(16) *Periphrastic errors*
$[[\ \text{EAT}^{[O]}\ T^{[PST]}_m\ ]\ C_m\ ]\ [[\ \text{EAT}\ T^{[O,PST]}_m\ ]\ C^{[O]}_m\ ]$

							# O.	# S.	type	N
a.	/did/	/-∅/	/-∅/	/ate/	/-∅/	/-∅/	1	0	periphrastic	356
b.	/did/	/-∅/	/-∅/	/ate/	/-ed/	/-∅/	1	1	peri.-red.	1
c.	/did/	/-∅/	/-∅/	/eat/	/-ed/	/-∅/	1	2	peri.-dis.	8
d.	/did/	/-∅/	/-∅/	/eat/	/-∅/	/-∅/	1	1	target/peri.-omi.	n.a.
e.	/do/	/-∅/	/-∅/	/ate/	/-∅/	/-∅/	1	1	*do*-periphrastic	52

Frequencies

Given an error's probability of occurrence p (≤ 1), the probability of occurring twice is p^2 ($< p$).
⇒ Distributive errors should be rarer than redundant ones, but are 4.5 times more frequent!
(Arnon 2009 found a similar frequency distribution in English plural errors , e.g. *foots*∼*feets* 3:1.)

Consistency bias

A type of mistake tends to be made consistently within the domain of the M-value.

Errors like (16b–e) should be infrequent as they involve two distinct types of mistake. Among (16b–d), (16c) should be most frequent as it conforms to the Consistency bias.

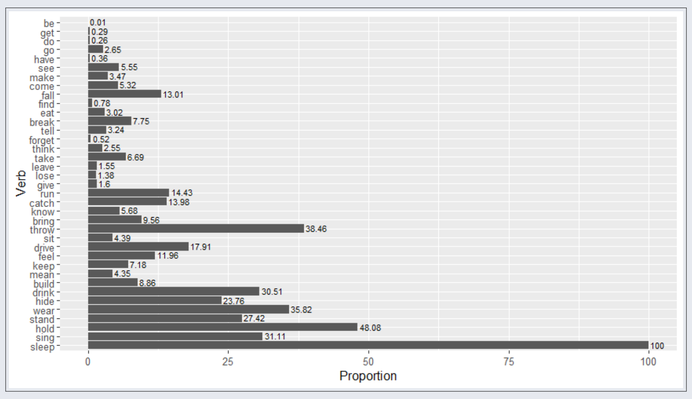

Fig. 2: Proportion of total errors by verb ordered by output frequency

For each lexical item:

- likelihood of neglecting a secondary feature, negatively correlates with stability of representation of that feature
- more frequent items have more stable representations

⇒ more errors with less frequent lexical items (Fig. 2)