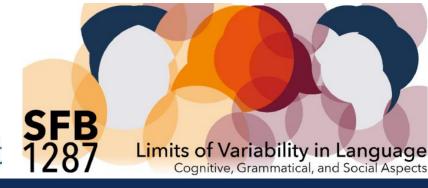


Background





Graded morphological generalization: A study of German inflection

Anna Jessen¹, Lara Schwarz², & Harald Clahsen¹ ¹University of Potsdam, ²Pennsylvania State University

ajessen@uni-potsdam.de

What are the sources of morphological productivity?

Associative generalization using lexical similarity with existing forms (e.g., Hahn

- & Nakisa, 2000)
- Grammatical generalization using morphological rules (e.g. Marantz, 2016)
- → A realistic model employs both mechanisms (Veríssimo & Clahsen, 2014), but how and to what extent?

Gradient Symbolic Computation (GSC, Goldrick et al., 2016):

- combines symbolic grammars with graded representations
- allows for violable constraints to be weighted

A GSC model reveals the relative contributions of lexical and grammatical constraints for a given set of data:

→ Compare a GSC model trained on a large set of corpus data with results from an elcited production experiment testing nonce-word generalization

Past participle formation in German

Regular (weak): no stem change in any form, preterit with –te, participle with –t: kaufen – kaufte – gekauft

Irregular (strong): preterit without affix, participle with –(e)n, unpredictable stem changes (167 verbs)

- 1. schreiben schrieb geschrieben (A-B-B) (n≈80)
- 2. singen sang gesungen (A-B-C) (n \approx 50)
- 3. $laufen lief gelaufen (A-B-A) (n\approx 30)$

(9 verbs have so-called mixed inflection: e.g. kennen - kannte - gekannt; were not tested.)

Procedure & materials

Nonce verbs for four conditions (24 items each):

- **Pure Irr**: VC cluster (i.e., rhyme) typical of irregulars (strong) (e.g., 'ind' \rightarrow finden)
- **Both**: VC cluster common to both weak and strong verbs ('erb' → sterben)
- **Pure Reg**: VC cluster only for weak verbs (e.g., 'ach' → machen, lachen)
- No-Rhyme: VC cluster does not appear in any German verb

Task: Fill in the participle form of a nonce verb presented in its infinitive form:

KRINGEN

Peter kringt täglich morgens seinen Zatt. Wie jeden Tag hat Peter auch gestern seinen *Zatt* _____.

(Peter krings every morning his Zatt. Like every day, yesterday Peter has ____ his Zatt.)

Constraints

Six constraints identified from grammars of German were applied to the dLex Corpus, a data base consisting of 100Mio. words (types : 2.3Mio.); relative weights (reflecting the contribution of each constraint to the corpus data) are shown:

1.1 **Plus** +t: The default ending for participles is +t:

19.9 **Parse**: maintain input-output correspondence of verb stem:

Rhyme: Adhere to an existing rhyme cluster pattern for the provided stem: **20.4**

*Change+t: Do not mix a stem change with a regular participle ending: 1.6

*Parse+en: Do not mix a maintained stem with an irregular participle 0.9 ending:

*IRRPre+t: Do not add a -t if the verb can be associated with other 1.6 verbs that have irregular preterit forms:

Participants of the elicited-production experiment

40 German Native Speakers (mean age 29.9, SD 11.2, range 20-68 years)

Results

Participants' answers vs. probabilities expected from GSC (corpus) model

Pure irregular	Answers (%)	Probability GSC	Pure regular	Answers (%)	Probability GSC
*ge-stem-t	57.5	0.21	ge-stem-t	85	0.88
*ge-stem-n	1.5	0.14	ge-stem-n	7.1	0.12
*ge-change-t	2.3	0.7	ge-change-t	1.6	0.00
*ge-change-n	38.6	0.58	ge-change-n	6.3	0.00

both	Answers (%)	Probability GSC	No-rhyme	Answers (%)	Probability GSC
ge-stem-t	68	0.21	ge-stem-t	89.3	0.88
ge-stem-n	2	0.14	ge-stem-n	4.2	0.12
ge-change-t	3.2	0.7	ge-change-t	2.3	0.00
ge-change-n	26.8	0.58	ge-change-n	4.2	0.00

Example:

*ge-stem-t:	ge-kring-t	*ge-change-t:	ge-krung-t
*ge-change-n:	ge-krung-en	*ge-stem-n:	ge-kring-en

Elicited production shows:

- regular -t most common response in all four conditions
- Pure Irreg attract most irregular –*n* responses

Comparison with GSC (corpus) model shows:

• GSC model's predicted probabilities for regular –t forms are lower than those of the participants' responses and those for irregular forms are higher.

Conclusions

- The model underestimates the role of regularizations: participants produced significantly more –t participles than predicted by the model
- The model overestimates the role of similarity: participants produced fewer irregular responses than predicted by the model
- Morphological productivity is mainly achieved by rule-based (grammatical) generalization and less so by lexical (similarity and frequency-based) association.

References

Goldrick, Putnam, & Schwarz (2016): Bilingualism: Language & Cognition, 19, 857–876. Hahn & Nakisa (2000): Cognitive Psychology, 41, 313–360.

Marantz (2016). Morphology. In S. Small, & G. Hickok (Eds.), Neurobiology of language (pp. 153-163).

Pinker & Ullman (2002): Trends in Cognitive Science, 6, 456-463

Veríssimo, J. & Clahsen (2014): Journal of Memory and Language, 76, 61-79