Conditioning, but on which distribution? Grammatical gender in German plural inflection

Kate McCurdy Adam Lopez Sharon Goldwater

University of Edinburgh

Summary

Grammatical gender is a consistent and informative cue to the plural class of German nouns. We find that neural encoder-decoder models learn to rely on this cue to predict plural class, but adult speakers are relatively insensitive to it. This suggests that the neural models are not an effective cognitive model of German plural formation.

Grammatical gender predicts plural class

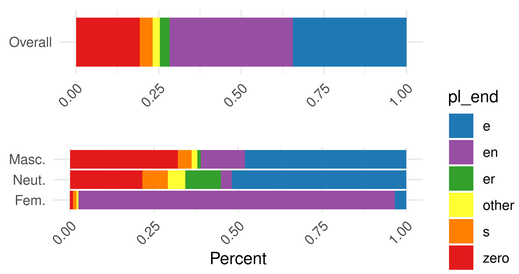

Figure 1: Plural suffix overall (upper) and by gender (lower) in the UniMorph corpus[3].

Plural form in German is realized by a range of suffixes (ignoring umlaut). The two most frequent suffixes show a gender-conditioned distribution:

- **-n** appears predominantly with feminine nouns, which take the definite article *die* in the singular.
  - *die Wahl* `the vote' → *Wahlen* `votes'
- **-e** appears predominantly with masculine nouns (singular definite article *der*) and neuter nouns (singular definite article *das*).
  - *das Mahl* `the time' → *Male* `times'

Hypothesis

Speakers and neural models will show comparable sensitivity to grammatical gender when producing plural inflections for unknown German nouns.

Stimuli

Non-word nouns, developed by [4]. By design, these stimuli lack strong phonological cues to plural class. They are all monosyllabic words ending in a consonant.

R		NR	
Bral	Pisch	Bnaupf	Plaupf
Kach	Pund	Bneik	Pleik
Klot	Raun	Bnöhk	Pläk
Mur	Spand	Fnahf	Pnähf
Nuhl	Spert	Fneik	Pröng
Pind	Vag	Fnöhk	Snauk

Table 1: Experimental stimuli [4]

Nouns with this phonological structure are generally nonfeminine and take **-e**, but still tend to take **-n** if feminine (Fig. 3).

Data collection

Speakers

- Type prompt: *Der Bral — Die ____*
- Random assignment to gender-counterbalanced lists
  - Participants see *Der*, *Die*, or *Das Bral*
  - Each participant sees 8 feminine, 8 masculine, and 8 neuter nouns
- 92 native German speakers included in analysis
  - 100 recruited through Prolific; 8 excluded for failing attention checks

Neural encoder-decoder (ED) model

- Architecture: MED [1], a biRNN proposed for cognitive modeling [2], implementation following [5].
- Data: 11,243 German nouns in UniMorph [3].
- Task: map input (singular form) to output (plural inflected form).
  - ⟨*f*⟩ w a h l → w a h l e n
  - ⟨*m*⟩ h u n d → h u n d e
  - ⟨*n*⟩ k i n d → k i n d e r
- Evaluation: test on same forms shown to speakers
  - ⟨*f*⟩ b r a l → ???
  - ⟨*m*⟩ b r a l → ???
  - ⟨*n*⟩ b r a l → ???

Neural model conditions on gender — speakers don't

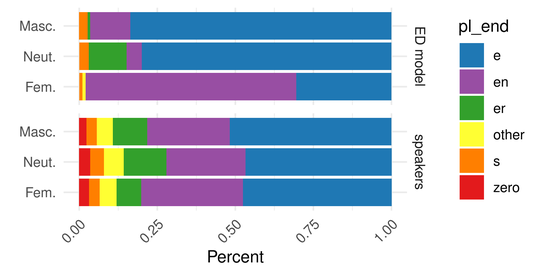

Figure 2: Plural suffix productions by gender, speakers (lower) vs. ED model (upper)

Speaker productions correlate better to distributions *without* gender

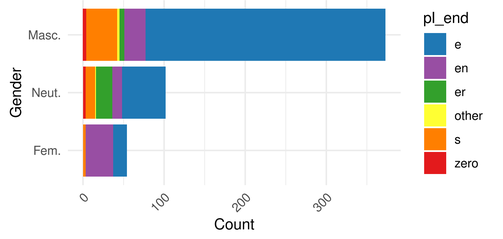

Figure 3: Plural suffix by gender for consonant-final monosyllabic words in UniMorph[3]

	Overall (Fig. 1)		Phon. cond. (Fig. 3)			
	Gender	No gender	Gender	No gender	Speakers	ED
Speakers	.49	**.67**	.70	**.78**		.49
	(.40, .56)	(.60, .72)	(.64, .75)	(.73, .82)		(.35, .61)
ED	**.62**	.41	**.71**	.47	.49	
	(.50, .71)	(.27, .54)	(.61, .78)	(.33, .59)	(.35, .61)	

Table 2: Correlations (Pearson's *r*, 95% confidence intervals in parentheses below) between item-level production percentages for speakers and ED model with
1) overall type frequency with and without gender conditioning (Fig. 1),
2) phonology-conditioned type frequency with and without gender (Fig. 3),
3) each other (correlation between item-level speaker and ED model productions).

References

[1] Kann & Schütze, ACL, 2016
[2] Kirov & Cotterell, TACL, 2018
[3] Kirov et al., LREC, 2016
[4] Marcus et al., Cog. Psych., 1995
[5] McCurdy et al., ACL, 2020