Topic drop in German: Grammar and usage

Giuseppe VaraschinAntonio Machicao y PriemerYaHU BerlinHU BerlinHU

Yanru Lu HU Berlin

{giuseppe.varaschin | machicao.y.priemer | yanru.lu} @ hu-berlin.de

1 Introduction

German declarative clauses are normally characterised by having one phrase of any type (NP, PP, VP, etc.) in the *pre-field*, i.e. in the front position of the sentence followed by the inflected verb in second position (1). But there is a subtype of declarative sentences in German (3a) that has a V1 structure, similar to German polar questions (2), but with an "assertive" meaning, like V2 declarative clauses. This construction has been called a.o. null topic, pronoun zap, pre-field ellipsis, topic drop (cf. Huang 1984, Fries 1988, Cardinaletti 1990, Wöllstein 2010, Müller 2014, Frick 2017, Schäfer 2021; a.o.). We use the last term in the current study.

A specific characteristic of this construction is that the pre-field (PF) has to remain empty (hence V1) (cf. (3a) vs. (3b)) and one (and only one) constituent has to be deleted. Furthermore, the constituent to be deleted, e.g. *den Aufsatz* 'the essay' in (3a), must have a prominent information structural status in the discourse (hence 'topic' drop). The context that licenses the construction needs to mark the deleted constituent as a topic (3).

- (1) [Den Aufsatz]_{PF} muss der Schüler morgen schreiben.
 the essay must the student tomorrow write
 'The student must write the essay tomorrow.'
- (2) []_{PF} Muss der Schüler den Aufsatz morgen schreiben? must the student the essay tomorrow write
 'Does the student have to write the essay tomorrow?'
- (3) What's going on with *the essay*?
 - a. [den Aufsatz]_{PF} muss der Schüler morgen schreiben. the essay must the student tomorrow write
 'The student must write the essay tomorrow.'
 - b. * [Morgen]_{PF} muss der Schüler schreiben.
 tomorrow must the student write
 INTENDED: 'The student must write the essay tomorrow.'

It has been mentioned in the literature (Fries 1988, Frick 2017; a.o.) that topic drop appears only in specific "text types" (e.g. private letters, spoken dialogues, etc.), but – as far as we know – this has not yet been empirically investigated, this being one of the goals of the present paper. Since clauses with and without topic drop ((1) vs. (3a)) represent truth-conditionally equivalent variants, we will show with a matched-guise study that this syntactic variation correlates with different social meanings, hence leading to a register-dependent variation, which explains the correlation of topic drop with specific text types postulated in the literature.

Furthermore, following recent studies of social meaning and register variation in HPSG (Green 1994, Wilcock 1999, Paolillo 2000, Bender 2001, Asadpour et al. 2022, Machicao y Priemer et al. 2022), we propose a formal analysis of topic drop in German (taking Müller 2014 as a starting point) that takes social meanings (as conventional implicatures, cf. Asadpour et al. 2022) into account in order to achieve more accurate predictions.

2 A matched-guise experiment

2.1 Hypotheses and predictions

To investigate the social meanings of topic drop in German, we conducted a matched-guise experiment (Lambert et al. 1975, Bender 2005, Campbell-Kibler 2007). Specifically, we hypothesize that listeners assign different characteristics to topic drop users as opposed to full form users. Thus, we expect a main effect of topic drop on the ratings of the characteristics of the speakers. Furthermore, we predict that the ratings for speakers who drop the subject are different from those who drop the object, i.e. there is an interaction between the variables topic drop and the topicalized argument.

2.2 Design and procedures

(4)

The experiment has a 2×2 within-subjects, within-items design with two independent variables with two levels each: TOPIC DROP (topic drop vs. full form) and the topicalized ARGUMENT (subject vs. object). The conditions are illustrated in (4). The materials consist of 8 items (each appears in the four different conditions) and 32 fillers. All items are in the form of written dialogues like (4)¹ and were presented to each participant in a fully randomized order. Participants were tasked to rate speaker B in each dialogue on a 6-point scale (1 = e.g. not friendly at all, 6 = e.g. very friendly) in terms of the following characteristics: *gebildet* 'educated', *arrogant* 'arrogant', *wortgewandt* 'articulate', *höflich* 'polite', *locker* 'relaxed', *pingelig* 'pedantic', *formell* 'formal', *freundlich* 'friendly'. The choice of these characteristics is based on those used in the matched-guise experiment in Beltrama (2018) and on results of prior qualitative interviews with a small group of native speakers addressing attitudes towards topic drop. The experiment was conducted online on the platform Ibex farm.

A: Hast du letzte Woche den Brief geschrieben?	
have 2sg.nom last week DEF.sg.m.Acc letter written	
'Did you write the letter last week?'	
B: a. Kann ihn morgen schreiben.	topic drop \times subject
can 3sg.m.Acc tomorrow write	
'(I) can write it tomorrow.'	
b. Ich kann ihn morgen schreiben.	full form $ imes$ subject
1sg.nom can 3sg.m.Acc tomorrow write	-
c. Kann ich morgen schreiben.	topic drop $ imes$ object
can 1sg.nom tomorrow write	
'I can write (it) tomorrow.'	
d. Den kann ich morgen schreiben.	full form $ imes$ object
DEM.3SG.M.ACC can 1SG.NOM tomorrow write	

23 self-reported German native speakers (17 female, 3 male, 1 diverse, 2 not specified) participated in the experiment. 21 are between the age 18 and 25, one is between 36 and 45 and another between 56 and 65.

2.3 Data analysis and results

The data is analyzed in R (R Core Team 2023) with cumulative link mixed models (CLMMs) for ordinal data (Christensen 2022). The model includes TOPIC DROP and ARGUMENT as main effects, the effect of their interaction, and PARTICIPANTS and ITEMS as random effects. The ratings on each scale are analyzed in a separate

¹As a reviewer points out, the use of written instead of spoken stimuli may have an influence on participants' perception of topic drop, as it is arguably a phenomenon associated with conceptually spoken language. As a first step, the dialogue form adopted here aims to simulate a spoken conversation as far as possible. Further studies using spoken stimuli are definitely worth carrying out. However, as several corpus studies show, topic drop is also widely used in the medium of informal written communication, e.g. text messages, chats, mails (Frick 2017, Schäfer 2021, a.o.). The stimuli we used are also compatible with such kinds of contexts.



Figure 1: Percentage of each rating by conditions for the polite scale.

univariate analysis. The results show a main effect of TOPIC DROP for the scales educated ($\chi^2 = 18.99$, p < 0.01), articulate ($\chi^2 = 22.12$, p < 0.01), polite ($\chi^2 = 15.30$, p < 0.01), pedantic ($\chi^2 = 4.71$, p = 0.03), formal ($\chi^2 = 20.51$, p < 0.01) and friendly ($\chi^2 = 246.38$, p < 0.01), but not for the scales arrogant and relaxed. Specifically, participants find speakers who use topic drop less educated, less articulate, less polite, less pedantic, less formal and less friendly compared to their counterparts who use the full form. Furthermore, we only find an interaction between TOPIC DROP and ARGUMENT for the scales polite ($\chi^2 = 7.66$, p < 0.01) and formal ($\chi^2 = 4.89$, p = 0.03). Participants rate speakers who use subject topic drop as even less polite and less formal than those who use object topic drop. Using the data for the polite scale as an example, Figure 1 illustrates the percentage of each rating by condition.

3 Analysis

(5)

We start by stating the purely grammatical licensing conditions for topic drop. Contrary to Cardinaletti (1990), we assume that subject and object topic drop have the same underlying structure, *pace* the difference in grammatical function. Given this, there are broadly two possible theoretical treatments. The first is to assume an empty category in the pre-field (either *pro* or *Op*) which binds another empty category inside the VP (Huang 1984, Cardinaletti 1990). The second is to assume a unary branching rule which, in itself, ensures the binding of a referential trace (Müller 2014: 101). The former approach would require some additional machinery to prevent empty elements referring to discourse topics from appearing outside of the pre-field. In order to avoid this, we assume the unary rule approach of Müller (2014). We also adopt the theory of information structure features in Paggio (2009), where attributes like TOPIC take indices as their values.



The NON-LOCAL FEATURE PRINCIPLE ensures that the mother node in a topic drop clause will have an empty slash value (Pollard & Sag 1994: 164). Since INDEX is specified as *ref*, topic drop with expletives is ruled out.²

²We can capture the commonalities between (5) and the HEAD-FILLER-SCHEMA used for normal V2 clauses with a sortal hierarchy of clausal types, where *topic-drop-phrase* and *head-filler-phrase* are subtypes of a general *slash-binding-phrase* (Müller 2014: 101).

However, this purely grammatical account does not explain the secondary inferences about the identity of the speaker which, as we saw in Section 2.1, German users systematically associate with topic drop structures – i.e. their *social meanings*. We represent these meanings – along with other use conditions – as constraints on the CONTEXT in which topic drop structures can be instantiated (Wilcock 1999, Paolillo 2000, Bender 2001, Asadpour et al. 2022). Since these constraints have the same projective properties as conventional implicatures, we represent them as parts of the set of propositions that form the value of the feature c(ONVENTIONAL) I(MPLICATURE) (McCready 2019, Salmon 2022, Asadpour et al. 2022). What distinguishes the social meanings from other conventional implicatures is that they always predicate some socially-relevant property of the context coordinates themselves, i.e. the values for C-INDEX features (SPEAKER, ADDRESSEE, etc.).³

In addition, most social meanings are associated with contrasting variants of a single abstract variable (Eckert 2005, 2008, Oushiro 2019, i.a.). Typically, each variant expresses a different partition of the (scalar) property corresponding to the social meaning: e.g. topic drop conveys low formality, whereas V2 with overt topic pronouns convey high formality. We represent this by treating social meanings essentially as gradable adjectives, taking a DEGR(EE) argument. For simplicity, we assign a binary value to this feature, indicating whether the degree to which context-coordinate possesses the property in question is *high* (9) or *low* (7). In a more realistic model, these values should be replaced by continuous values representing non-empty subsets of scales defined along a specific dimension, as in standard degree semantics (Kennedy 2001, i.a.). See Potts & Kawahara (2004: 261) and McCready (2019: 29) for a similar approach applied to the semantics of honorifics.

In principle, grammars can associate any social meaning to any type of sign. However, linguistic forms are generally not associated with a unique social meaning. Rather, the socially-relevant assumptions that a form conveys about the utterance participants are highly context-dependent, with each form being associated with an *indexical field* of related meanings (Eckert 2008). In our experiment, participants associated a set of (arguably related) properties with speakers of topic drop and declarative V2 clauses and not a single property.

We model this underspecification by stating our grammar in a way so that constraints only associate linguistic structures with abstract social meaning types. These types are only resolved to maximally specific social meanings in concrete communicative situations. We assume that the major distinction between different types of social meanings is between social meanings that concern the way the speaker presents themselves to the addressee (*speaker-sm*) and those that say something about the relation between speaker and addressee (*relational-sm*). The latter is arguably what is grammaticalized by T/V pronouns (Kaur & Yamada 2022).⁴



With (6) in mind, we propose the following schemas, incorporating the results of our experiment.



³The ascription of a social meaning can be embedded under a series of attitude predicates expressing that such ascription is relative to the speakers' beliefs about what the communicative norms in a particular linguistic community are (Green 1994, Asadpour et al. 2022). For reasons of space, we do not consider this possibility here and assume simpler social meaning structures. On the basis of these, we can define a register as a cluster of linguistic constraints whose social meanings are satisfied in recurrent situation types. ⁴We opted for a simple formulation of the hierarchy where each of the adjectives we tested reflects a property of the speaker (given the nature of the matched-guise task) and is also a maximal sort. But there are other alternatives. It is likely that *polite* is underspecified, in that it can be interpreted either as property of the speaker (e.g. as equivalent to *formal*) or as a relational property indicating (social or psychological) distance (McCready 2019: 28–29). Similarly, *friendly* is arguably related to something like *psych-prox*.



Notice that, unlike in standard accounts of CI projection, the schemas above don't have a strictly compositional nature, because the social meaning of the mother is not solely determined by the social meanings of the daughters (Pollard & Sag 1994: 333, Paolillo 2000: 242–243, Potts 2005: 68, i.a.). The construction itself contributes its own social meaning. Schema (7), for instance, ensures that the social meaning of topic drop sentences consists of whatever CIS may be gathered in the head daughter, plus a non-empty list of propositions describing situations where the speaker is specified as having a low degree for any of the *cognitive* traits in (6). Which subset of *cognitive*-typed predications is chosen depends on other features of the context.⁵

The analysis we propose here has consequences for the standard view about which linguistic variables are visible to sociolinguistic evaluation. Since topic drop is licensed by a phrasal construction, our results imply that speakers' evaluations are sensitive to abstract syntactic variables (see also Bender 2007, Robinson 2022). This is a departure from some sociolinguistic literature (Labov 2001, Meyerhoff & Walker 2013, Eckert & Labov 2017, i.a.), which claims that only phonological or lexical variables can be socially monitored. Our analysis is also incompatible with a model of grammar that only attaches social and other CI-meanings to vocabulary items or surface realizational patterns in PF (Adger 2006, Saab 2021). Rather, we require a more flexible architecture that can represent the social information speakers indexically associate with *any* linguistic structure. Furthermore, these social meanings have to be further arranged in a sortal hierarchy, which allows us to capture the fact that variants are often underspecified w.r.t. the conditions they can be felicitously used in.

References

- Adger, David. 2006. Combinatorial variability. Journal of Linguistics 42(3). 503–530. https://doi.org/10.1017/s002222670600418x.
- Asadpour, Hiwa, Shene Hassan & Manfred Sailer. 2022. Non-wh relatives in English and Kurdish: Constraints on grammar and use. In Stefan Müller & Elodie Winckel (eds.), Proceedings of the 29th International Conference on Head-Driven Phrase Structure Grammar, Nagoya University & Institute for Japanese Language and Linguistics, 6–26. Frankfurt am Main: University Library. https://doi.org/10.21248/hpsg.2022.1.
- Beltrama, Andrea. 2018. Precision and speaker qualities: The social meaning of pragmatic detail. *Linguistics Vanguard* 4(1). Article 20180003. https://doi.org/10.1515/lingvan-2018-0003.
- Bender, Emily M. 2001. Syntactic variation and linguistic competence: The case of AAVE copula abscence. Stanford, CA: Stanford University dissertation.
- Bender, Emily M. 2005. On the boundaries of linguistic competence: Matched-guise experiments as evidence of knowledge of grammar. *Lingua* 115(11). 1579–1598. https://doi.org/10.1016/j.lingua.2004. 07.005.
- Bender, Emily M. 2007. Socially meaningful syntactic variation in signbased grammar. *English Language & Linguistics* 11(2). 347–381. https: //doi.org/10.1017/S1360674307002286.
- Campbell-Kibler, Kathryn. 2007. Accent, (ING), and the social logic of listener perceptions. *American speech* 82(1). 32–64. https://doi.org/10. 1215/00031283-2007-002.

- Cardinaletti, Anna. 1990. Subject/object asymmetries in German nulltopic constructions and the status of specCP. In Joan Mascaró & Marina Nespor (eds.), *Grammar in progress*, 75–84. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110867848.75.
- Christensen, Rune Haubo Bojesen. 2022. ordinal: Regression models for ordinal data. https://cran.r-project.org/package=ordinal (12 June, 2024).
- Eckert, Penelope. 2005. Variation, convention, and social meaning. In Annual meeting of the Linguistic Society of America (LSA). Oakland CA, vol. 7.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of sociolinguistics* 12(4). 453–476. https://doi.org/10.1111/j.1467-9841.2008. 00374.x.
- Eckert, Penelope & William Labov. 2017. Phonetics, phonology and social meaning. *Journal of sociolinguistics* 21(4). 467–496. https://doi. org/https://doi.org/10.1111/josl.12244.
- Frick, Karina. 2017. Elliptische Strukturen in SMS: Eine korpusbasierte Untersuchung des Schweizerdeutschen. Berlin: De Gruyter. https://doi.org/ 10.1515/9783110517859.
- Fries, Norbert. 1988. Über das Null-Topik im Deutschen. Sprache und Pragmatik 3. 19–49.
- Green, Georgia M. 1994. The structure of context: The representation of pragmatic restrictions in HPSG. In James H. Yoon (ed.), *Proceedings* of the 5th Annual Conference of the Formal Linguistics Society of Mid-America, vol. 1/2, 215–232. Urbana, IL: Department of Linguistics, University of Illinois.

⁵The general principle of CI composition should state, thus, that the sum of the CI values of the daughters is only a *sublist* of the CI value of a phrase (see also Asadpour et al. 2022: 19). Without further constraints, this implies that there could be infinitely many models for each sign, which would vary only with respect to their CI values. Asadpour et al. (2022: 22) suggest a solution to this problem based on the idea of minimal exhaustive models: essentially, only signs that have a minimal number of elements in their CI lists would be considered models of the grammar. This seems to entail that only one new maximal social meaning could be introduced by each variant of topic drop in any given utterance, i.e. the *nelists* in the social meaning assignments in (7)-(9) could only be satisfied by singleton lists. This approach might be too restrictive because we want to leave open the possibility that an utterance of topic drop could convey that the speaker is presenting as, for instance, both unfriendly and informal. Another idea would be to constrain CI values to be a sublist of a closed list representing speakers' background knowledge, like the *stored persona* proposed by Taniguchi (2019: 24). In addition to ensuring a finite number of models, this would capture the effect that underspecified social meanings are typically resolved to those that confirm prior beliefs about the identity of the speaker (Podesva et al. 2015).

- Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15(4). 531–574.
- Kaur, Gurmeet & Akitaka Yamada. 2022. Honorific (mis)matches in allocutive languages with a focus on Japanese. *Glossa: A Journal of General Linguistics* 44(1). https://doi.org/10.16995/glossa.7669.
- Kennedy, Christopher. 2001. Polar opposition and the ontology of 'degrees'. Linguistics and Philosophy 24. 33–70. https://doi.org/10.1023/A: 1005668525906.
- Labov, William. 2001. Principles of linguistic change: Social factors. Vol. 2. Malden, MA: Blackwell.
- Lambert, Wallace E, Howard Giles & Omer Picard. 1975. Language attitudes in a French-American community. *Linguistics* 13(158). 127–152. https://doi.org/10.1515/ling.1975.13.158.127.
- Annual Meeting of the Linguistic Society of America (LSA). Oakland CA. 2005. Vol. 7. Linguistic Society of America (LSA).
- Machicao y Priemer, Antonio, Stefan Müller, Roland Schäfer & Felix Bildhauer. 2022. Towards a treatment of register phenomena in HPSG. In Stefan Müller & Elodie Winckel (eds.), Proceedings of the 29th International Conference on Head-Driven Phrase Structure Grammar, Nagoya University & Institute for Japanese Language and Linguistics. Frankfurt am Main: University Library. https://doi.org/10.21248/hpsg.2022.5.
- McCready, Elin. 2019. The semantics and pragmatics of honorification: Register and social meaning. Oxford: Oxford University Press.
- Meyerhoff, Miriam & James A. Walker. 2013. An existential problem: The sociolinguistic monitor and variation in existential constructions on Bequia. *Language in Society* 42(4). 407–428.
- Müller, Stefan. 2014. Elliptical constructions, multiple frontings, and surface-based syntax. In Paola Monachesi, Gerhard Jäger, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004*, *Nancy*, 91–109. Stanford, CA: CSLI Publications.
- Oushiro, Livia. 2019. A computational approach for modeling the indexical field. *Revista de Estudos Linguagem* 27(4). 1737–1786. https://doi. org/10.17851/2237-2083.0.0.1737-1786.
- Paggio, Patrizia. 2009. The information structure of Danish grammar constructions. Nordic Journal of Linguistics 32(1). 137–164. https://doi.org/ 10.1017/S0332586509002066.
- Paolillo, John C. 2000. Formalizing formality: An analysis of register variation in Sinhala. *Journal of Linguistics* 36(2). 215–259. https://doi.org/ 10.1017/S0022226700008148.
- Podesva, Robert J, Jermay Reynolds, Patrick Callier & Jessica Baptiste. 2015. Constraints on the social meaning of released /t/: a production

and perception study of US politicians. *Language Variation and Change* 27(1). 59–87. https://doi.org/10.1017/S0954394514000192.

- Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press & CSLI Publications.
- Potts, Christopher. 2005. The logic of conventional implicatures. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/ 9780199273829.001.0001.
- Potts, Christopher & Shigeto Kawahara. 2004. Japanese honorifics as emotive definite descriptions. In Robert B. Young (ed.), *Proceedings of the 14th Semantics and Linguistic Theory Conference (SALT)*, vol. 14, 253–270. Northwestern University: The Linguistic Society of America. https://doi.org/10.3765/salt.v14i0.2917.
- R Core Team. 2023. *R: A Language and Environment for statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https: //www.R-project.org.
- Robinson, Mary. 2022. Negative concord as a window into social perception of morphological and syntactic variables. New York, NY: NYU dissertation.
- Saab, Andrés. 2021. On the locus of expressivity: Deriving parallel meaning dimensions from architectural considerations. In Eleonora Orlando & Andrés Saab (eds.), *Slurs and expressivity: Semantics and beyond*, 17–44. Lanham, MA: Lexington Books.
- Salmon, William. 2022. Social markers and dimensions of meaning. Journal of Pragmatics 192. 98–115. https://doi.org/https://doi.org/10.1016/ j.pragma.2022.02.014.
- Schäfer, Lisa. 2021. Topic drop in German: Empirical support for an information-theoretic account to a long-known omission phenomenon. Zeitschrift für Sprachwissenschaft 40(2). 161–197. https:// doi.org/10.1515/zfs-2021-2024.
- Taniguchi, Ai. 2019. Social meaning at the semantics-sociolinguistics interface.
- Wilcock, Graham. 1999. Lexicalization of context. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional* aspects of linguistic explanation, 373–387. Stanford, CA: CSLI Publications.

Wöllstein, Angelika. 2010. Topologisches Satzmodell. Heidelberg: Winter.

- Yoon, James H. (ed.). 1994. Proceedings of the 5th Annual Conference of the Formal Linguistics Society of Mid-America. Vol. 1/2. Urbana, IL: Department of Linguistics, University of Illinois.
- Young, Robert B. (ed.). 2004. Proceedings of the 14th Semantics and Linguistic Theory Conference (SALT). Vol. 14. Northwestern University: The Linguistic Society of America.