On the use of the definite article with proper names in German

Edgar Onea[1], Melanie Loitzl[1], Anna Volodina[2]

[1] University of Graz, [2] IDS Mannheim

The possibility to use the definite article with proper names (PNs) denoting humans in German has been observed in the literature early on, however, research has mainly focused on oral German and regional and historical variation. The present study contributes to generating important empirical data regarding written German and suggests a novel theory involving covert adjectival features in the DP projected from the PNs.

Introduction PNs are a special form of nominal phrases that are considered by many to be rigid designators (cf. Kripke 1980). Being inertly definite, they do not require a definite article in many languages, as witnessed in the German example (1). However, some languages such as French, Dutch, and German require the use of a definite article before a PN if the PN occurs together with an adjective (cf. van Langendonck 2007) as in (2):

(1) Anna liest ein Buch.
Anna reads a book
'Anna is reading a book.'

(2) *(Die) kleine Anna liest ein Buch.
The small Anna reads a book
'Anna is reading a book.'

It has been suggested that in cases like (2), the definite article merely indicates the left edge of a nominal phrase (cf. Gunkel 2017, van Langendonck 2007). However, there are also cases such as (3) in which the definite article is optionally present in German (see Nübling et al. 2015, Werth 2014, 2020, among others), raising the question what factors determine its distribution:

(3) Anna sieht (den) Alfred.
Anna sees the.acc Alfred
'Anna can see Alfred.'

Sentences like (3) are associated with oral language, and their acceptability varies with regional preferences (cf. Werth 2014 and 2020, Nübling et al. 2015, Zifonun 2017, Bellmann 1990, Eichhoff 2000), emotive implications (cf. Bellmann 1990, Busley/Nübling 2021, Busley/Fitzinger 2021), and syntactical opacity (cf. Golato 2013, Betz 2015). Yet even in spoken varieties the semantic or pragmatic correlates of the use of the definite article appear to be widely unclear. However, for written language we are not aware of any corpus linguistic (or otherwise quantitative) analyses of this phenomenon in German. We address this research gap in this paper.

Method In our corpus study, we gathered data from the *Deutsches Referenzkorpus* (*DeReKo*) where we explicitly searched for bare PNs and for PNs accompanied by a definite article, both for male and female PNs. The selection of the PNs was based on a ranking of the most common PNs in Austria between 1984 and 2020 (cf. Statistik Austria 2020) as we expected these names to occur frequently in German texts in general; i.e. the data includes texts from all German varieties and the ranking by Statistik Austria 2020 ensured higher probability of name occurrence than random selection. Out of the total hits in *DeReKo*, we extracted a sample of 1200 hits, which we cleared errors (e.g. plural PNs, PNs that do not refer to people, but also cases in which there is a definite article with the name in query aimed for bare PNs). This left us with a total of 569 correct hits, which we annotated according to several criteria such as occurrence of a definite article, occurrence of a last name, gender, case, position in the sentence, occurrence of an attribute (before or after the PN) and optionality. Out of all the criteria we investigated, three are relevant for the current analysis: gender, case, optionality. Optionality represents the grammaticality of an example both with and without a definite article.

Results Based on the relevant data from the annotation, we conducted a statistical analysis in R using standard chi-square tests and generalised linear models. As witnessed by the association plots in Figure 1, one can see that in the general case, the presence of the definite article is way more likely in genitive, while in the optional cases, the picture radically changes and the definite article is only significantly overrepresented for accusative. Crucially all genitive data are non-optional because the genitive case needs to be marked either on the article or by a suffix on the proper name. When the suffix is present, thus, the article cannot appear and when the suffix is not present the article needs to appear.
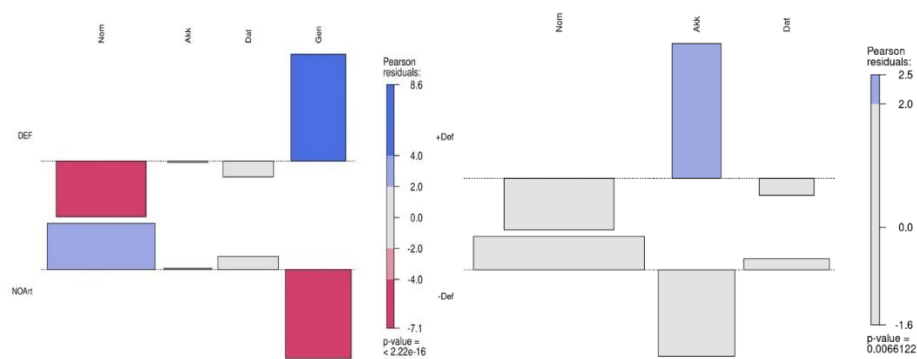

Figure 1 Asssociation plot all data (l), optional article (r)

Gender is important because there is wide syncretism in the case marking pattern of German. Chiefly, feminine inflection does not distinguish nominative and accusative on the article, while masculine does. Hence, if the functional explanation of increased article use in accusative was due to argument structure disambiguation (as suggested by parts of the literature), we would expect an interaction between gender and case in the optional cases. We have fitted a generalized linear mixed model using the formula ARTICLE ~CASE*GENDER and found no significant interaction but, as expected, a main effect of accusative. Hence, a case-disambiguating function is unlikely. Finally, we computed the probability of the optional use of the definite article for each case for the corpus using Bayesian simulation: Nom (0.0018), Acc (0.0048), Dat (0.0019), Gen (0). The methods converge on a higher likelihood of definite article + PN in accusative.

Discussion Our theoretical proposal is that when the definite article is present with a PN denoting a human there is always an adjectival feature in the projected DP, the article being necessary because the DP extends beyond the PN itself ([def [[adj] [PN]]). Put differently, the article is never actually optional: with adjective in the structure it is obligatory, without it is ungrammatical. Crucially, this adjectival feature can be implicit, thus yielding the impression of optionality. We assume that the kind of adjectival features that can be left implicit are usually not lexical in nature but grammatical correlates of typical high-transitivity features such as high-affectedness, [topic], [familiar], [person], etc. which have been associated with differential object marking in the literature (Hopper and Thompson 1980, Iemmolo 2010, Dalrymple & Nikolaeva 2011, Hill & Mardale 2021). This correctly predicts that it is precisely accusative that we find most cases of 'optional' definite articles with PNs, albeit the article on the PN is not itself a case of DOM. Moreover, since such features are generally discursive features it is not surprising that we find them more likely to be present at LF and easier to reconstruct in spoken language that generally relies more on ellipsis. This correctly predicts a higher likelihood of articles in spoken language. Yet, while the exact nature of these features cannot be reliably reconstructed based only on the data we currently have, we propose an experimental method to further investigate their

discursive and semantic correlates. Finally, our suggestion broadly agrees with the sorts of semantic correlates discussed in the literature on oral language but takes a broader theoretical perspective on the matter.

Literature

Bellmann, G. (1990): Pronomen und Korrektur. Zur Pragmalinguistik der persönlichen Referenzformen. Tübingen: Niemeyer. – **Betz**, E. (2015): Recipient design in reference choice. Negotiating knowledge, access, and sequential trajectories. In: Gesprächsforschung 16, S. 137–173. – **Busley**, S./**Fritzinger**, J. (2021): Neutrales Rufnamengenus zwischen Semantik und Pragmatik. *Linguistik online* 107.2, 125–144. – **Busley**, S./**Nübling**, D. (2021): Referring to women using feminine and neuter gender. Sociopragmatic gender assignment in German dialects. *Nordisk tidskrift för socioonomastik / Nordic Journal of Socio-Onomastics* 1, 33–59. – **Dalrymple**, M./**Nikolaeva**, I. (2011): *Objects and Information Structure*. Cambridge: Cambridge University Press. – **Eichhoff**, J. (2000): *Wortatlas der deutschen Umgangssprachen*. Bd. 4. Bern: Saur. – **Golato**, A. (2013): Reparaturen von Personenreferenzen. *Deutsche Sprache* 41.1, 31–51. – **Gunkel**, L. (2017): *Artikel und Demonstrative*. Hrsg. von L. Gunkel/A. Murelli/S. Schlotthauer/B. Wiese/G. Zifonun: Grammatik des Deutschen im europäischen Vergleich. Das Nominal. Bd. 1. Berlin/Boston: de Gruyter. 254–341. – **Hill**, V./**Mardale**, A. (2021): *The Diachrony of Differential Object Marking in Romanian*. Oxford: Oxford University Press. – **Hopper**, P./**Thompson**, S. (1980): *Transitivity in grammar and discourse*. Language 56. 251–299. – **Iemmolo**, G. (2010): Topicality and differential object marking. *Studies in Language* 34.2, 239–272. – **Kripke**, S. A. (1980/2001): *Naming and Necessity*. 12. Aufl. Massachusetts: Harvard University Press Cambridge. – **Langendonck**, W. van (2007): *Theory and typology of proper names*. Berlin/New York: de Gruyter. – **Nübling**, D./**Fahlbusch**, F./**Heuser**, R. (2015): *Namen. Eine Einführung in die Onomastik*. Tübingen: Narr Studienbücher. – **Statistik Austria** (2020): *Liste der 60 häufigsten Vornamen 1984-2020 in Österreich - Original-Schreibweise ohne Sonderzeichen*. – **Werth**, A. (2014): Die Funktionen des Artikels bei Personennamen im norddeutschen Sprachraum. Hrsg. von F. Debus/R. Heuser/D. Nübling: *Linguistik der Familiennamen*. Hildesheim/Zürich/New York: Georg Olms. 139–174. – **Werth**, A. (2020): *Morphosyntax und Pragmatik in Konkurrenz. Der Definitartikel bei Personennamen in den regionalen und historischen Varietäten des Deutschen*. Berlin/Boston: de Gruyter. – **Zifonun**, G. (2017): Substantive. Hrsg. von L. Gunkel/A. Murelli/S. Schlotthauer/B. Wiese/G. Zifonun: *Grammatik des Deutschen im europäischen Vergleich. Das Nominal*. Berlin/Boston: de Gruyter. 408–515.