

Sociohistorical linguistics is about to go medieval on early German

In this talk, I discuss a new means of executing the sociohistorical analysis of early medieval European languages, i.e., the literization approach of Somers (2024), and apply it to the early history of German. My empirical basis is the ninth-century bilingual Codex Sangallensis (G), which contains a Latin and German version of Tatian's second-century gospel harmony. To historical linguists of any bent, there seems to be little to recommend Codex G as a good data source. Its transmission history is complicated, and the people involved in its production are unknown. One might also mistrust the value of the manuscript's German data because the German text that occupies the right-hand column of the codex appears to be a close translation of the Latin text in the codex's lefthand column (fig. 1). This raises the possibility that the German data actually reflect Latin syntactic patterns. A better data source for variationists and generativists alike, both of whom aim to reconstruct historical grammars, would be personal letters and diaries, which seem to approximate spoken language and, thus, "cast light on the history of natural language" (Elspeß 2012: 156). Yet, ego-documents of this type are unattested in German's earliest period (700–1050), which is likely a reason why scholars have not analyzed medieval texts like G sociolinguistically, despite the variation exhibited across the corpus.

Instead, scholars have approached the variation by separating the data they believe reflect a genuine German competence from those that were influenced by confounding factors, e.g., a Latin source. In the case of G, this method has limited analyses to so-called 'contrasting tokens,' examples in which the German translation deviates from the Latin original (Fleischer 2006). Its underlying logic is as follows: tokens for which the possibility of foreign influence cannot be excluded risk obscuring native grammatical patterns and, thus, must be excised from datasets. In ex. 1, for instance, the definite determiners *thes* and *themo* ('the') are contrasting tokens in that they are present in the German translation but not in the Latin, which has no determiners. This difference suggests that the German-speaking translator added them because their native grammar required it. While the addition of determiners undoubtedly reflects a grammatical difference between Latin and German, questions persist, including why *thes* and *themo* were added here, but not before the definite nouns *tagun*, *Iudeno*, and *namen*. If *thes* and *themo* are an expression of a grammatical rule, why were they not added before each definite noun?

A literization analysis, in contrast, understands G as a linguistic and literary artifact produced in a particular sociocultural environment: German-speaking Carolingian Europe with its limited vernacular literization and preoccupation with the evangelization of Europe. Literization refers to the process of speakers transforming their almost exclusively oral multilectal vernacular into ad hoc written varieties, i.e., *scripti*, for defined purposes (Somers 2024). Here, the project was exegetical and entailed transferring the important, but difficult to read Codex Fuldensis (F) (fig. 2) into a document that elucidated the structure and meaning of F's Latin text for German-speaking learners of Latin. To this end, the translators relied on the classical interpretative tool of columnar translation, which functions similarly to a modern-day gloss (Dickey 2015). This perspective shifts how we analyze contrasting tokens; regarding the added determiners in ex. 1, the question becomes: How do they explicate the Latin text? I argue that the translators added definite determiners in positions that effected lexical cohesion across sentences specifically for the German reader, who is socioculturally and linguistically distinct from the classically oriented, literate Latin-speaking Christian of Late Antiquity, the original consumer of the Latin Bible. What is linguistically interesting about G is how its creators represented Latin's well-developed system of text cohesion with the grammatical tools that were available in a newly literate German.

Figures and examples:

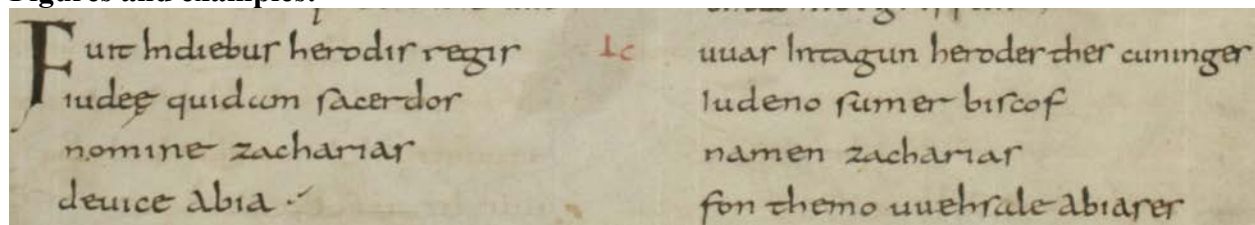


Figure 1: An early excerpt from the Codex Sangallensis 56 (G), 25: 29–32

29. Fuit in diebus Herodis regis Was in days Herod.GEN king.GEN	29. Uuas in tagun Herodes thes cuninges Was in days Herod.GEN the king.GEN
30. Iudee quidam sacerdos Judea.GEN certain priest	30. Iudeno sumer biscof Jews.GEN certain priest
31. nomine Zacharias name.ABL Zacharias	31. namen Zacharias name.DAT Zacharias
32. de vice Abia of office Abia.GEN	32. fon themo uuehsale Abiases of the office Abia.GEN

‘There was in the days of Herod the king a certain priest by the name of Z. of the office of Abia’

Example 1: Codex Sangallensis 56 (G), 25: 29–32 transcribed and glossed

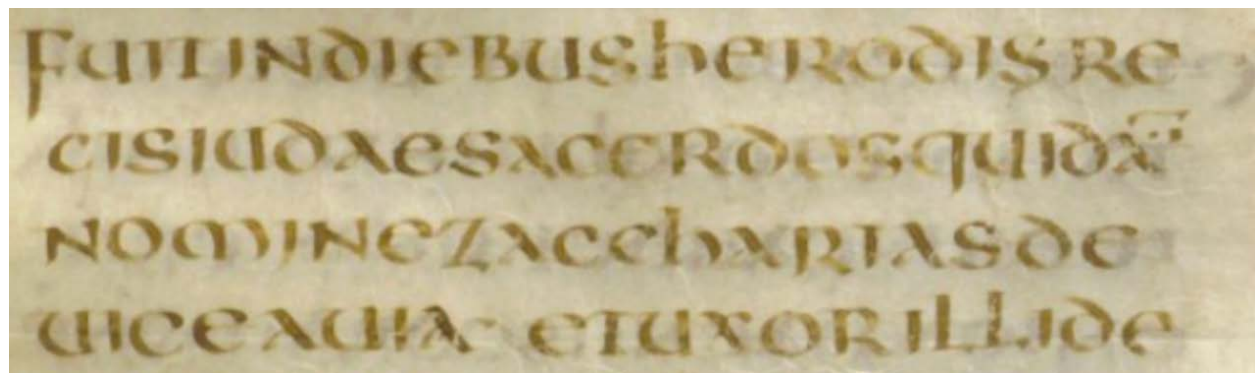


Figure 2: The corresponding excerpt to fig. 1 from the sixth-century Codex Fuldensis, 45:21r

Links to digitized manuscripts:

Codex [Sangallensis](#)

Codex [Fuldensis](#)

Works cited:

- Dickey, Eleanor. 2015. Columnar translation: An ancient interpretive tool that the Romans gave the Greeks. *The Classical Quarterly* 65.2 807–821.
- Elspaß, Stephen. 2012. The Use of Private Letters and Diaries in Sociolinguistic Investigation. In: *Handbook of historical sociolinguistics*, ed. by Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, 156–169. Malden, MA: Wiley-Blackwell.
- Fleischer, Jürg. 2006. “Zur Methodologie althochdeutscher Syntaxforschung.” *Beiträge zur Geschichte der deutschen Sprache und Literatur* 128 (1): 25–69.
- Somers, Katerina. 2024. *How to create an early German scriptus: The literization approach to historical German syntax*. Language Science Press (Open Germanic Linguistics): Berlin.