

Analysis on the relative frequency of the Saxon versus of-genitive over time

Mischa Koster S6463851

12-01-2026

Abstract

This study investigates diachronic changes in the use of the Saxon genitive (e.g., the king's butler) and the of-genitive (e.g., the butler of the king) in written English prose from the 19th century (1800–1849) to the 20th century (1950–1999), using texts from Project Gutenberg dating from the respective time periods. Normalized frequencies of both constructions will be compared across the two time periods. Based on other studies, the results are expected to show an increase in of-genitive usage, which would support the view that English continues to shift from synthetic to analytic structures. This research contributes to the understanding of grammatical change over time.

1 Introduction

Languages tend to change over time, a typological shift is common, and languages can undergo a process of grammatical simplification. An example of this change of language in English is the transition between the construction which is more widely used; a more synthetic system, which mainly depends on inflectional morphology, or a more analytic one, which depends on word order and function words. Although both constructions have co-existed for centuries in the English language, the relative preference of which construction is used is influenced by a complex bundle of factors, which depend on semantics, syntax and processing factors according to Kreyer (2003).

The selection between the inflectional Saxon Genitive (marked by 's) and the *of*-genitive form, is a clear example of a grammatical change, of which the relative frequency can be studied historically, which this project aims to research in or-

der to better understand grammatical change over time.

The research question is as follows: How does the relative frequency of the Saxon Genitive and the of-genitive in English texts differ between the texts from the 19th century (1800–1849) and the 20th century (1950–1999)? The hypothesis is that the of-genitive will be significantly more frequent in the 20th century texts compared to the 19th century, which indicates an increase in analyticity and grammatical change. The hypothesis is based on previous studies that indicate a comeback of the Saxon genitive in certain genres, but that still show a general preference for of-genitive constructions in formal, complex, and most other forms of prose texts.

2 Related Work

Several studies have examined the factors that determine the choice between which genitive construction is chosen, and where. In the study by Kreyer (2003), we learn that the *of*-construction at one point of time used to be rarely used, however that changed over time, as around the 14th century, the *of*-construction was the most used variant. The Saxon genitive would later make its comeback in Modern English. However in cases where the *of*- or Saxon genitive could both be used, which genitive is used is largely dependent on two main factors: processability (syntactic weight and complexity) and human involvement (animacy and semantic relation). Based on Kreyer's research on 698 instances from the British National Corpus (BNC), he discovered that objective, non-personal inanimate modifiers substantially favor the *of*-genitive, whilst proper and personal names or nouns, strongly preferred the Saxon genitive. This study provides the theory for understanding the variation, but does not focus on

the diachronic change itself.

Another study from Rosenbach (2008) also highlights the significance of animacy in grammatical variations. Rosenbach's work shows that animate possessors strongly prefer the Saxon genitive, while inanimate ones favor the of-genitive. The research provides insight into how animacy interacts with other factors like topicality and syntactic weight, and even how this occurs across languages. While the study does not provide much insight into diachronic development, it does make clear that in order to accurately measure the diachronic change regarding which genitive is more widely used, we must factor in context.

A third relevant study by Denison et al. (2010) researches the distribution of the so called "group genitive" (*a feature where the possessive 's attaches to the last word of the whole phrase, rather than to the head noun.*) Using the spoken component of the BNC, they show that postmodified possessors with possessive 's (e.g, the man in the corner's hat) are surprisingly rare, which aligns with our hypothesis, and that speakers often use avoidance strategies like the of-genitive where the postmodifier appears after the possessum. Their findings show that the Saxon genitive prefers to attach to the head noun, and that processing ease plays a key role in the actual usage, similar to what (Kreyer, 2003) discovered. Their finding that speakers avoid postmodified Saxon genitive constructions in spoken English, supports the broader trend in a shift towards analyticity in the English language.

3 Data

Data Collection The study will use texts from Project Gutenberg, a digital library of public-domain works in order to research the relative frequencies of both the Saxon and the of-genitive. To study diachronic changes, two time periods will be compared: The 19th century period, of which we will use works from 1800-1849, and secondly the 20th century period, which will use works from 1950-1999. A random sample of 20 prose texts (both fiction and non-fiction) will be selected from each period. The chosen texts will be limited to those originally written in English to avoid possible issues that would stem from translations. The independent variables will be the time periods, being 1800-1849, and 1950-1999. The dependent variables will be the genitive construction choices,

measured as the raw frequency of Saxon and of-genitives per text, and the normalized frequency per 10.000 words ($((\text{raw frequency} / \text{total words in text}) * 10.000)$) for both genitive construction types, together with the relative frequency ($\text{frequency of Saxon genitives} / (\text{Saxon} + \text{of-genitives}) * 100$), in order to further highlight preferred genitive constructions.

Data Extraction Each text will be processed using corpus analysis tool AntConc (Anthony, 2024) to identify and count genitive constructions. Both the Saxon genitive and the of-genitive will be extracted. Only transformable constructions, where both forms are theoretically interchangeable in expressing the exact same semantic relationship (e.g, *the king's butler / the butler of the king*), will be included, following the methodology used by Kreyer (2003). Non-transformable constructions will be excluded, such as:

- Post-genitives (e.g, *a friend of hers*)
- Local genitives (e.g, *at John's*)
- Independent genitives (e.g, *John's* in *faster than John's*)
- Quantitatively and qualitatively partitive constructions (e.g, *a bowl of soup*)
- Constructions of material (e.g, *a chair of wood*)

To ensure accuracy, 20 percent of the derived results will be randomly manually checked, which will allow us to adjust AntConc if needed to maximize correct sampling. This does offer a possible limitation regarding a possible generalization issue.

Pre-processing We will clean the texts by removing metadata (titles, chapter headings, Project Gutenberg headers) and convert them to plain text. Texts will be tokenized by AntConc which splits the text into words. OCR errors in 19th-century texts will be noted as a limitation.

Table 1 provides a summary of the data that will be used in this study.

Parameter	19th Century	20th Century
Time period	1800–1849	1950–1999
Texts per period	20	20
Language	English	English
Total words*	~500,000	~500,000

Table 1: Corpus design and parameters. *Approximately

4 Predicted Results

We predict an increase in the use of the of-genitive in the 20th-century texts, due to the increase in syntactic complexity, which has been linked with greater of-genitive usage (Kreyer, 2003). This would support a trend towards more analytical English diachronically.

The following table summarizes the expected normalized frequencies:

Period	Saxon Gen*.	Of-Gen*.
19th century	~15	~20
20th century	~10	~25

Table 2: *Predicted frequencies of genitive constructions per 10.000 words across the 19th and 20th century, stemming from 1800-1849 and 1950-1999 respectively.

Discussion If the predicted results are confirmed, they would align with the hypothesis that the English language has continued to change toward more analytic structures. The decrease in the frequency of the Saxon genitive (from 15 to 10 per 10.000 words) and the increase in of-genitive frequency (from 20 to 25 per 10.000 words) would imply that the English language shifts away from inflectional morphology. However, these results may vary depending on the sampled genres, as literature that would contain more animacy (e.g., biographies) would show an increase in the Saxon genitive according to Rosenbach (2008).

5 Conclusion

This study aimed at investigating the diachronic changes in genitive construction usage by comparing the relative frequency of the Saxon genitive versus the of-genitive across the 19th and 20th century, in English prose. Using the framework from (Kreyer, 2003), it tested the hypothesis that English has become increasingly more analytical (*dependent on separate words rather than in-*

flection to express grammatical relationships) over time.

The predicted results support this hypothesis, suggesting a substantial increase in usage of the of-genitive from the 19th to 20th century. If those results were confirmed through the corpus analysis, they would form empirical evidence for continuing grammatical change in the English language.

However several limitations should be noted, including: genre restrictions, as this study solely uses prose texts, and OCR quality in historical texts, as older texts from project Gutenberg could contain scanning errors which would affect the accuracy of the found results. Lastly, the sample size of 20 texts per period may pose an issue for generalization.

Future research could expand the analysis to a broader timeframe, further genres, increase sample size, or apply similar methodologies to other languages.

6 Github Link

<https://github.com/mischa188/school>

References

- Anthony, L. (2024). Antconc (version 4.3.1) [macos 10/11 intel].
- Denison, D., A. K. Scott, and K. Börjars (2010). The real distribution of the english “group genitive”. *Studies in Language. International Journal sponsored by the Foundation “Foundations of Language”* 34(3), 532–564.
- Kreyer, R. (2003). Genitive and of-construction in modern written english. processability and human involvement. *International Journal of Corpus Linguistics* 8(2), 169–207.
- Rosenbach, A. (2008). Animacy and grammatical variation—findings from english genitive variation. *Lingua* 118(2), 151–171. Animacy, Argument Structure, and Argument Encoding.