

Abstract. *shall* is among the least common modal auxiliaries in most varieties of present-day U.S. English but is by far the most common modal in sources of law like statutes and constitutions (Brooks 2023). In this paper, we offer empirical evidence that the prominence of *shall* in law is motivated by genre-specific convention rather than communicative content or writing process. In a production experiment building on the findings of Martínez et al. (2024), laypeople were tasked with writing about crimes as legislators and as tour guides for foreign tourists. We find that laypeople use *shall* more frequently when writing in the legal genre than they do when writing about the law in non-legal settings, even when the content described is conceptually equivalent. We argue that these results offer insight into the nature of performativity and genre conventions which exist at the functional level. Marked items like deontic *shall*, which are relatively uncommon in non-legal genres are able to serve specialized functions in law precisely because of their markedness.

Keywords. psycholinguistics; applied linguistics; modality; genre; law; legal linguistics; style

1. Introduction. The deontic usage of the modal *shall* is ubiquitous in legal English, but relatively rare in other registers (Witczak-Plisiecka 2007). Although the prominence of *shall* has been described as a matter of distributional fact, questions remain about the underlying motivation for the phenomenon.

- (1) *The Senate shall have the sole Power to try all Impeachments.* (U.S. Const. art. I § 3)
- (2) *No state shall enter into any Treaty, Alliance, or Confederation.* (U.S. Const. art. I § 3)

This paper presents empirical data to test whether this usage has arisen as an arbitrary convention maintained through legal education, or whether it has a genre-specific signaling function recognized by laypersons. Research on legal language (“legalese”) has provided a framework for distinguishing these alternatives. Martínez et al. (2022, 2024) found support for what they call the “magic spell hypothesis,” positing that legal language is written in a complex manner to signal the performativity of legal texts, over the “copy-and-edit” hypothesis, which posits that legalese is complex because it is the result of an iterative writing process whereby revisions are inserted into existing sentences rather than as separate ones. Although Martínez et al. (2022, 2024) focus on syntactic complexity, similar claims have been made about other linguistic properties of the law (Tiersma 2008; Kurzon 1986; Trosborg 1995). In particular, deontic *shall* has been identified by several linguists as a distinguishing property of legal English (e.g., Brooks 2023, Garzone 2013, Witczak-Plisiecka 2007, Gotti 2003), which is not easily explained by the “copy-and-edit” hypothesis. This past research on *shall* motivates the present study.

2. Possible drivers of the frequency of *shall* in law. Here, we focus on testing two competing explanations for the prominence of *shall* in legal texts.¹ One proposal is what we call the

* Thank you to Emily Goodwin, Josh Phillips, Cleo Condoravdi, and Kim Tien Nguyen for their helpful feedback on this project. Authors: Tilden “Tilly” Brooks, Stanford University Department of Linguistics, Yale Law School (teb44@stanford.edu) & Robert D. Hawkins, Stanford University Department of Linguistics (rdhawkins@stanford.edu).

“deonticity-expressing hypothesis,” which posits that *shall* is common in legalese because it frequently describes deontic concepts like duties and prohibitions. Crucially, this hypothesis proposes that *shall* usage is driven by the content of the law, not its performative nature. The magic spell hypothesis offered by Martínez et al. (2024) makes the opposite proposal: *shall* is used frequently in legalese because it signals the creation of legal obligation. Here, we offer support for the magic spell hypothesis over the deonticity-expressing hypothesis. When tasked with describing a set of legal requirements in legal and non-legal genres, experimental participants were more likely to use *shall* when writing in the legal genre. Thus, genre, not content, appears to underlie increased *shall* usage in law. This finding offers insight into the functional role of marked features in language usage.

3. Experimental research on complexity in law. Recent research on legal language has identified linguistic properties such as center-embedding and long-distance dependencies, not conceptual complexity, as driving processing difficulty (Martínez et al. 2022). Martínez et al. (2024) assess the linguistic complexity of legalese through two hypotheses, the magic spell hypothesis and the copy-and-edit hypothesis. The magic spell hypothesis posits legal language is written in a complex manner to lend legalese a “ritualistic, spell-like element” (Martínez et al. 2024: 2). On this view, linguistic complexity is taken to signal or facilitate performativity. On the other hand, the copy-and-edit hypothesis proposes legalese is complex because it is the result of an iterative writing process whereby revisions are inserted into existing sentences rather than as separate ones. Across two production experiments, participants without legal training were tasked with writing stories, laws, and unofficial descriptions about crimes from scratch and in a step-by-step writing and editing process. The authors found that laypeople use more center-embedded syntax when writing legal texts, but that content produced through a writing and editing process does not contain more center-embedded syntax than content produced all at once. These findings are consistent with the authors’ magic spell hypothesis but do not support the copy-and-edit hypothesis. At a theoretical level, these findings offer support for the view that the linguistic complexity associated with legal language functions as an indicator of the law’s performative effects.

4. Methods. Following the method developed by Martínez et al. (2024), we implemented an experiment which prompted participants to write visitors’ guides and laws describing certain crimes. In a within-subject design, participants were instructed to (1) pretend they were tour guides and write a summary of a specific criminal statute for visiting tourists; and (2) pretend they were lawmakers and write a draft law about a specific crime. As an additional manipulation participants were instructed to write from scratch (the “from scratch condition”) in some of their trials and in an editing process (the “editing condition”) in the other trials. In both processes, participants were instructed to write a law or guidebook description based on provided specifications. After completing their entry, participants moved on to the next page. In the from scratch condition, participants moved on to the next trial at this stage. In the editing condition, participants were informed that their supervisor (if in the tour guide condition) or their fellow lawmakers (if in the legal condition) requested that they revise their original draft and were given

¹ As discussed above, the copy-and-edit hypothesis was viewed as unlikely to serve as an explanation for *shall* usage in the law. However, for the sake of completeness in replicating the experimental design used in Martínez et al. (2024), a writing process manipulation was also retained and assessed.

additional specifications to incorporate into their draft. Participants then revised their original entry to reflect the added specifications.

Participants were adult native English speakers living in the United States recruited using Prolific. Each participant completed eight trials ($N = 67$).² No participant was presented with the same item more than once and was presented with each of the four conditions at least once. Each participant saw one item in the tour guide, editing condition, three items in the tour guide, from scratch condition, two items in the law, editing condition, and two items in the law, from scratch condition.

Before beginning the first trial, participants were prompted to commit to not use large language models to assist them in completing the task. In addition, participants were asked to confirm that they understood their role for each trial before completing it. Participants who completed all eight trials and were found not to have used large language models to assist them in completing the task were retained in the analysis. Thus, 67 out of 70 participants were retained.

5. Analysis plan. According to a pre-registered analysis plan, participant responses were extracted and then the effect of trial and genre manipulation on *shall* usage was analyzed. Because the outcome variable could be measured in two ways (i.e., as a count of the number of times a participant uses *shall* or as a binary measure of whether or not *shall* is used in a response at all), two models were used in the analysis: a random effects Poisson regression with frequency of *shall* usage in participant responses as the outcome variable and a random effects logistic regression using the presence or absence of *shall* as the outcome variable. Genre was the fixed effect of interest and random effects consisted of participant and item. In the case of trials in the editing condition, only the final response was considered.

Results were to be interpreted as offering support for the magic spell hypothesis if there was a significant effect of genre on *shall* usage, with *shall* being used more frequently in the legal condition than the tour guide condition. In the alternative, results were to be interpreted as evidence for the copy-and-edit hypothesis if there was a significant effect of writing process on *shall* usage, but relative similarity in *shall* usage in legal and tour guide trials. Finally, results were to be interpreted as evidence for the deonticity-expressing hypothesis if neither genre nor writing process were observed to have a significant effect on *shall* usage.

6. Results. The results reveal significant variation in *shall* usage along the lines of genre but little variation according to writing process. The modal was substantially more common in responses prepared in legal writing trials compared to responses generated during tour guide writing trials. Results of the Poisson regression model indicate that this difference is statistically significant ($p = 0.00195$). As with the Poisson model above, a logistic regression model treating *shall* as a binary outcome variable also indicated that genre has a significant effect on *shall* usage, with *shall* being less likely to appear in responses prepared for trials in the tour guide condition than those in the legal condition ($p < 0.001$). Although there was little variance by item, substantial variation by participant was observed in both models.

² Materials associated with the experiment are available on GitHub: <https://github.com/tillybrooks/245b-legalise>.

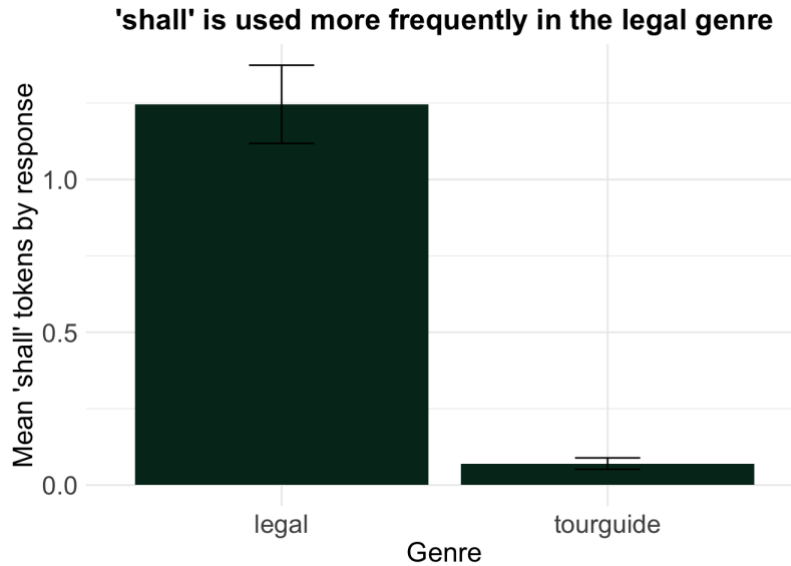


Figure 1. *shall* is used more frequently in the legal genre

Unlike genre, the models offered little evidence that the way responses were prepared had a significant effect on *shall* usage. Neither model indicated that there is a significant effect of writing process on the outcome variable.

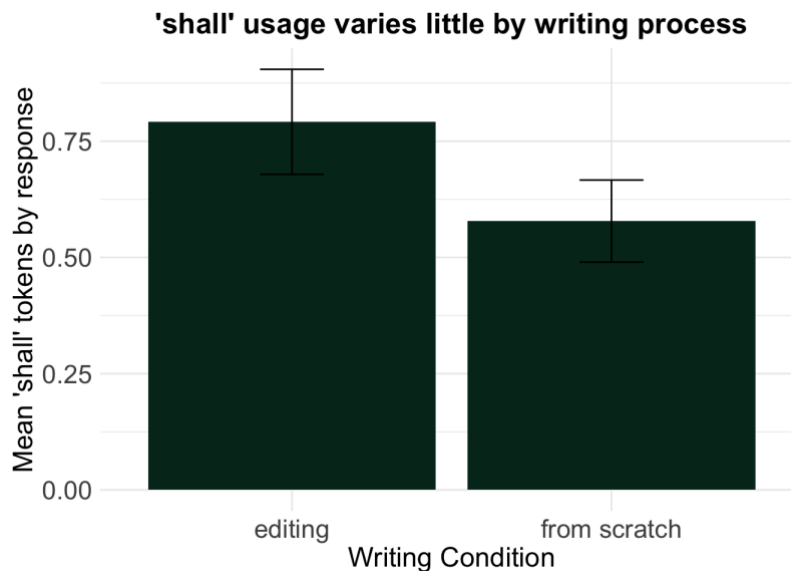


Figure 2. *shall* usage varies little by writing process

Looking at all four conditions together, average *shall* usage is higher in trials in the legal genre regardless of whether the text was written from scratch or in an iterative process. These results are consistent with Martínez et al.'s (2024) findings, which indicated that genre affects syntactic complexity much more than writing process.

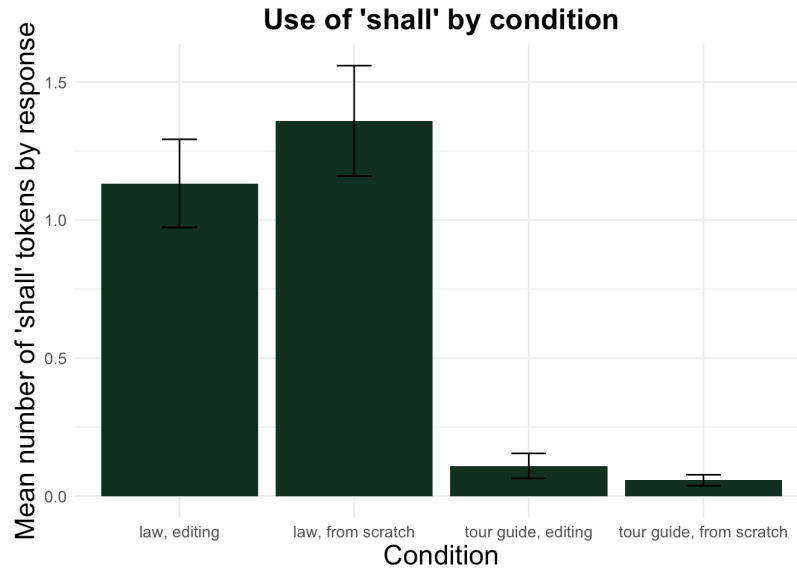


Figure 3. Use of *shall* by condition

These results are consistent with the predictions of the magic spell hypothesis (i.e., *shall* is more common in the legal genre and does not appear to be predicted by writing process) and inconsistent with those of the copy-and-edit (i.e., *shall* usage is predicted by writing process and not genre) and deonticity-expressing hypotheses (i.e., *shall* usage is not predicted by either writing process or genre, but rather content).

7. Discussion. We find that the magic spell hypothesis is the most likely explanation for the prominence of *shall* in law. While genre does appear to impact usage of the modal and writing process does not, the empirical evidence does not support either the copy-and-edit or deonticity-expressing hypotheses. Instead, the evidence indicates that it is the legal genre itself which motivates the prominence of *shall* in law.

We hypothesize that the best explanation for this finding is stylistic and pragmatic in nature. First, we propose that legal *shall* indexes authoritativeness. As a matter of sociolinguistic style, the usage of the modal *shall* in law allows for the signaling of the authority that is often linked to laws like the criminal statutes used in the experiment. This link is facilitated by the relative paucity of *shall* in non-legal English. Where *shall* is generally an infrequent modal in most varieties of U.S. English, its prominence in legal settings is somewhat marked. As Acton (2021: 110) observes, marked linguistic forms can be linked with specific stances, functions, identities, or characteristics. Thus, it may be the case that the infrequency of deontic *shall* in other contexts allows it to be easily recognizable as serving a particular function in the legal context. The stylistic contribution of *shall* also relates to the modal's pragmatic function in law as a performative. The law does not merely describe obligations; it creates them. Because the law operates in this way, it is important that the nature of the rights, duties, privileges, and powers delineated in the law is clear from how they are presented in legal language. Where *shall* is stylistically associated with authoritativeness, the modal is able to serve the function of indicating that the utterance in which it appears is a performative bearing strong legal force. Indeed, commentators have observed that *shall* is often used specifically to indicate unequivocal legal obligation (Williams 2009). That *shall* serves as an element of stylistic practice offers an explanation for why genre rather than content appears to condition *shall* usage in descriptions of legal obligations. It is desirable for a language user to signal authoritativeness when they are

producing a performative utterance that is intended to create legal obligations (as in statutes and other legal texts). When merely reporting legal obligations (as in visitors' guides), however, it is not similarly important for speakers to appear authoritative. Thus, like Martínez et al. (2024), we find that the evidence indicates that performativity underlies the phenomenon discussed in this paper.

8. Conclusion. In this paper, we empirically addressed the frequency of *shall* in law. Focusing on whether this phenomenon is the result of the communicative content of law or the legal genre itself, we find evidence that genre conditions differences in *shall* usage in writing produced by laypeople, even when the content remains consistent. We argue that the performative nature of the law is what motivates this genre effect. In addition, we argue that an indexical link between *shall* and legal authoritativeness allows for the modal to clarify the nature of the performative utterances within which it appears.

One limitation of the study is its focus on laypeople. Further empirical research such as experiments involving participants with legal training and perception studies may further illuminate the nature of legalese. In addition, the social meanings associated with *shall* and the other distinctive linguistic elements of law merit further inquiry.

Although the findings of this paper are constrained to the language of the law, we contend that the systematic study of genre-specific language usage has the potential to provide important insight into the nature of performativity and linguistic style. It is commonly observed that the language of the law is highly opaque. Recent research indicates that this opacity is linguistic, not conceptual, in nature (Martínez et al. 2022). The reliance on complex linguistic structures not strictly tied to the content of the law (e.g., center-embedded syntactic constructions, use of low frequency modals), then, is a topic worth exploring. The law presents a case in which the usual tendency toward efficiency (Gibson et al. 2019) appears to have been outweighed by some competing pressure. Further research on this topic may help linguists understand what those other pressures are.

References

- Acton, Eric. 2021. Pragmatics and the third wave. In Moore, Emma & Hall-Lew, Lauren & Podesva, Robert J. (eds.), *Social meaning and linguistic variation: Theorizing the third wave*. Cambridge University Press.
- Brooks, Tilden "Tilly" E. 2023. *Modal auxiliaries in legal United States English: The case of the United States Constitution*. New Haven, CT: Yale University Senior Essay.
- Garzone, Giuliana. 2013. Variation in the use of modality in legislative texts: Focus on *shall*. *Journal of Pragmatics* 57. 68–81. <https://doi.org/10.1016/j.pragma.2013.07.008>.
- Gibson, Edward & Futrell, Richard & Piantadosi, Steven P & Dautriche, Isabelle & Mahowald, Kyle & Bergen, Leon & Levy, Roger. 2019. How efficiency shapes human language. *Trends in Cognitive Sciences* 23(5), 389–407. <https://doi.org/10.1016/j.tics.2019.02.003>.
- Gotti, Maurizio. 2012. Shall and will in contemporary English: A comparison with past uses. In Facchinetti, Roberta & Palmer, Frank & Krug, Manfred (eds.), *Modality in Contemporary English*, 267–300. De Gruyter Mouton.
- Kurzban, Dennis. 1986. *It is Hereby Performed: Explorations in legal speech acts*. John Benjamins Publishing Company.
- Martínez, Eric & Mollica, Francis & Gibson, Edward. 2022. Poor writing, not specialized concepts, drives processing difficulty in legal language. *Cognition* 224. 105070. <https://doi.org/10.1016/j.cognition.2022.105070>.

- Martínez, Eric & Mollica, Francis & Gibson, Edward. 2024. Even laypeople use legalese. *Proceedings of the National Academy of Sciences* 121(35). e2405564121. <https://doi.org/10.1073/pnas.2405564121>.
- Tiersma, Peter. 2008. The nature of legal language. In Gibbons, John & Turell, M. Teresa (eds.), *Dimensions of Forensic Linguistics*, 7–25. John Benjamins Publishing Company.
- Trosborg, Anna. 1995. Statutes and contracts: An analysis of legal speech acts in the English language of the law. *Journal of Pragmatics* 23(1). 31–53. [https://doi.org/10.1016/0378-2166\(94\)00034-C](https://doi.org/10.1016/0378-2166(94)00034-C).
- Williams, Christopher. 2009. Legal English and the ‘modal revolution.’ In Salkie, Raphael & Busuttil, Pierre & van der Auwera, Johan (eds.), *Modality in English: Theory and Description*, 199–210. De Gruyter Mouton.
- Witczak-Plisiecka, Iwona. 2007. Linguistic Aspects of the deontic shall in the legal context. In Kredens, Krzysztof & Gozd-Roszkowski, Stanislaw (eds.), *Language and the Law: International Outlooks*, 181–197. Peter Lang.