Language Variation and Change (2026), **38**, 91–113
doi:10.1017/S0954394526100660

CAMBRIDGE UNIVERSITY PRESS

Unpredictable grammatical choices are not harder than predictable ones

Thomas Van Hoey[1], Matt Hunt Gardner[2], Ruiming Ma[3] and Benedikt Szmrecsanyi[3]

[1]Fonds voor Wetenschappelijk Onderzoek, Belgium and KU Leuven, Brussel, Belgium; [2]Queen Mary University of London, London, UK and [3]KU Leuven, Leuven, Belgium
Corresponding author: Matt Hunt Gardner; Email: matthuntgardner@gmail.com

Abstract
Recent variationist research indicates that grammatical intra-speaker variation (or: optionality) is unproblematic in speech production. Optionality contexts do not coincide with dysfluencies. In this study, we ask a theoretically significant follow-up question about optionality contexts that strongly cue grammatical variant choice: are optionality contexts in which all variants are probabilistically equally likely more problematic than those that strongly cue choice of grammatical variant? After all, unbiased (un-cued or freer) choices are sometimes theorized as being more difficult than biased (cued) ones. We empirically analyzed a subset of the SWITCHBOARD corpus of spoken American English on a turn-by-turn basis. The dataset covers 7,295 conversational turns containing 7,001 optionality contexts (spread over 20 grammatical alternation types), 2,970 filled pauses, and 41,297 unfilled pauses. Contrary to claims in the literature, weak probabilistic cueing does not trigger more production difficulties than strong probabilistic cueing. Unpredictable grammatical choices are not harder than predictable grammatical choices.

Keywords: alternation; dysfluency; disfluency; complexity; American English; optionality

Introduction

Our point of departure is recent literature suggesting that grammatical variation, or grammatical *optionality* (the term we will use in the remainder of this contribution, as it unambiguously[1] refers to intra-speaker variation) does not complicate language production because optionality contexts (or: variable contexts) do not attract dysfluencies in naturalistic discourse (Gardner et al., 2021; see also Szmrecsanyi et al., in press; Van Hoey et al., in press). Ma et al. (2025) additionally demonstrated that lack of dysfluency attraction persists even when we consider (a) the number of variants among which speakers can choose (alternations offering more choices do not attract more dysfluencies), and (b) the number of probabilistic constraints conditioning optionality (alternations conditioned by more constraints do not attract more dysfluencies). What

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (http://creativecommons.org/licenses/by-nc-sa/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is used to distribute the re-used or adapted article and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

we do not know at this point, however, is the extent to which probabilistic cueing of particular variants makes a difference. That is, are optionality contexts in which a particular variant is strongly predictable from the linguistic context "easier" than contexts in which all variants have a similar probability of occurring? This is the gap that the present study endeavors to fill.

We exemplify as follows. The retention/omission of English complementizer *that* is a well-known case of grammatical optionality in variationist linguistics (e.g., Jaeger, 2006; Tagliamonte & Smith, 2005):

(1) I don't think ____ it's a deterrent at all. I don't think ***that*** anyone that, uh, commits a murder actually thinks ***that*** they will ever even uh be punished let alone put to death for it. (Woman b. 1961, South Midland Dialect Area, SWITCHBOARD)

In Example (1), there are three optionality contexts where the speaker must make a linguistic choice between an overt *that* or a null complementizer (i.e., "zero") in a single turn. Here, *that* and zero are "alternate ways of saying 'the same' thing" (Labov, 1972:188); both function identically to introduce a complementizer phrase. They are functional equivalents. While we must eschew here a lengthy review of the pertinent literature (e.g., Labov, 1978; Lavandera, 1978; Sankoff, 1988) about whether linguistic variables can be extended from phonology to grammar[2] (we believe the answer is "yes!"), it is clear that Example (1) harbors a good deal of optionality. The question, however, is the following: is this optionality suboptimal in terms of language production? Many theorists believe so (see below). We turn the issue into an empirical question by marrying the variationist methodology to a corpus-based psycholinguistics research design, asking: do some optionality contexts trigger more speech dysfluencies than others?

The larger theoretical context of this study is a widespread, often implicit, and non-evidence-based suspicion in more theoretically oriented circles[3] that grammatical optionality and form-function asymmetry as in Example (1) is so odd that it must be theorized away. Berruto (2004:293-294) aptly described the situation as follows:

> [F]or most linguists, variation and variety are in fact a *crux*. At first sight, and often for some also in final analysis, linguistic variation, while empirically evident, represents an element of disturbance, something that seems to obscure the true perception of things, an obstacle to the theorizing and abstraction required for the scientific understanding of facts. This is so much so that the fundamental theoretical traditions in linguistics, from Saussure to Chomsky, to post-generativism, to many functionalists themselves (not Halliday, of course; but certainly Dik, or Givón up to a certain extent, to name but two) […] have more or less systematically sought to eliminate all elements of variation from the linguist's scope, positing only that which is constant, invariable, underlying the changing superficial realizations and independent from the speaker's actuation as a worthy object of study.

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

This is another way of saying that for many theorists, optionality and form-function asymmetry are synchronically abnormal (Goldberg, 1995:67; Haiman, 1980:516; see also Uhrig, 2015). It follows that when optionality does arise against allegedly all odds, it is typically assumed to be short-lived diachronically (see the extended discussion in De Smet, 2019 and references cited therein; e.g., Anttila, 1989; Dik, 1988; Geeraerts, 1997). In exactly this spirit, Goldberg (2019:26) wrote that "two species that share the same ecological niches cannot co-exist […] Darwin in fact long ago drew the analogy to language, noting that two words cannot remain in a long-term equilibrium if they are both associated with the same meaning." The reason, we submit, that optionality between synonymous expressions is thought to be short-lived diachronically is that it is allegedly suboptimal and difficult—otherwise, there wouldn't be evolutionary pressure to restore form-function symmetry and isomorphism. Note here that the assumption that optionality is suboptimal and difficult is not entirely implausible, given the psychological literature. Take, for example, Hick's Law (see Proctor & Schneider, 2018), according to which the time it takes to make a decision is proportional to the number of choices one has. We add in passing that prescriptive grammarians and language mavens are, of course, more often than not also strongly anti-variationist (see, e.g., Sundby, 1998:476). This sentiment has been referred to as "the doctrine of form-function symmetry" (Poplack & Dion, 2009:557) in the variationist literature.

It is fair to say that the idea that optionality is abnormal while form-function symmetry is a design feature of language is empirically problematic, given the sizable variationist literature on the existence, ubiquity, and systematicity of grammatical variation. It appears that the reasoning that we reviewed in the preceding paragraph essentially boils down to the view that optionality adds dysfunctional "complexity" to language. Complexity has been studied from various angles, which broadly fall into two groups (see Miestamo, 2008; Van Hoey et al., in press): measures of absolute complexity and measures of relative complexity. Measures of absolute complexity focus on the complexity of system-inherent structures: for example, counting the number of contrastive elements in a system (Nichols, 2013). Measures of relative complexity, by contrast, focus on user complexity and evaluate system-inherent properties as they relate to a language user (Kusters, 2003): for example, how hard or difficult it is to use a particular language or language variety compared to others.

Optionality in language, by definition, increases the *absolute* complexity of grammar, as the existence of multiple forms or patterns that encode the same meaning or grammatical function will inevitably yield a longer grammar compared to grammars that observe strict form-function symmetry. It is less clear, however, how optionality relates to *relative* complexity. For example, does having to choose between variants make producing an utterance harder for language users? Further, are all types of choices between variants equal in their effect (if any) on production difficulty?

We see no point in non-empirical theorizing about optionality, but we do wish to acknowledge there may exist empirically enlightened reasons for believing that grammatical optionality could be burdensome and cognitively problematic. First, as variationists well know, optionality is typically conditioned probabilistically by any number of contextual (language-internal) constraints (e.g., Bresnan et al., 2007 on the dative alternation, among many others). Thus, before they can make a choice between

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Table 1. Datasets and dependent variables under study in previous SWITCHBOARD research. Szmrecsanyi et al. (in press) offers a short synopsis of the studies cited here

Study	Dataset	Dependent variables (dysfluencies)	How tested
Gardner et al. (2021)	SWITCHBOARD subset	Filled pauses and unfilled pauses	Binary (presence/absence), per turn
Ma et al. (2025)	Full SWITCHBOARD	Unified measure	Continuous, per turn
Van Hoey et al. (in press)	Full SWITCHBOARD	Filled pauses and unfilled pauses	Continuous, per conversation
Current study	SWITCHBOARD subset	Unified measure	Continuous, per turn

variants in context, language users need to check the linguistic context for the various constraints that regulate the variation at hand. It is plausible that this extra cognitive work, regardless of how automatic it is, results in increased cognitive load. However reasonable this assumption may seem, Gardner et al. (2021) showed that in a subset of the SWITCHBOARD corpus, conversational turns with more optionality contexts do overall not attract more dysfluencies than turns with fewer optionality contexts. This suggests that optionality does not trigger production difficulties.

That said, Gardner et al.'s (2021) study can be accused of using a "sledgehammer-type method" (we thank an anonymous reviewer for this phrasing), because it does not factor in differences between different grammatical alternations or between individual optionality contexts. To address this shortcoming, subsequent work (see Table 1 for a summary) has investigated a number of follow-up questions: (a) Does the finding that optionality does not trigger dysfluencies hold when the whole SWITCHBOARD is taken into account? (b) Is dysfluency attraction or repellence perhaps a function of the number of variants among which people can choose, or of the number of probabilistic constraints that regulate grammatical alternations? (c) Do different types of alternation (insertion, permutation, or substitution, see De Troij, 2022) attract dysfluencies to different extents? Questions (a-b) were investigated in Ma et al. (2025); question (c) was studied in Van Hoey et al. (in press). The answer to the above questions (a–c) is "no" throughout: the full SWITCHBOARD corpus shows the same pattern as the SWITCHBOARD subset investigated in Gardner et al. (2021); it does not matter among how many variants speakers can choose or by how many probabilistic constraints particular alternations are conditioned; and there are no differences, in terms of dysfluency attraction, between different types of alternations. Thus, under no circumstances does grammatical optionality measurably trigger production difficulties. We add that the SWITCHBOARD studies reviewed above also differ in the way that the dependent variable (dysfluencies) is operationalized (see Table 1 for more details).

Crucially, however, there is one theoretically important issue that has remained uninvestigated so far (and this is the gap that the present study will fill): in certain instances, after considering all contextual constraints, the most appropriate option to choose may remain indeterminate because there is relatively equal contextual probability among choices. Some analysts predict that selecting one option may be more

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

cognitively demanding in this scenario compared to utterances for which, after all contextual constraints have been considered, one option has, say, 90% probability. Goldberg (2019:26), for instance, argued that free choices incur inefficiencies because such decisions take longer to make (see also Levshina & Lorenz, 2022:250-251). Simply put, some specific optionality contexts may present easy choices because one option is highly cued, while others may present harder choices because the choice is less predictable (no single option is highly cued). Consider again the English complementizer alternation as exemplified in (2) and (3):

(2) He's — he'll be five in in July and it's just hard to believe **that**, you know, time has- has flown by so quick, so, yeah
(Woman b. 1961, South Midlands Dialect Area, SWITCHBOARD)

(3) Um, I don't really think ____ we have anything
(Woman b. 1964, South Midlands Dialect Area, SWITCHBOARD)

According to our probabilistic modeling (see following sections for more detail) the predicted probability for the overt *that* complementizer in Example (2) is 50% (the null complementizer, or zero, has the same predicted probability in this specific context. Hence, it is not highly cued or predisposed toward one choice over the other), while the predicted probability for zero in Example (3) is 90% (thus, strongly cued). Relevant factors include, among other things, the length of the complement clause, which is long in Example (2) but short in Example (3). If it were true that free choices incur inefficiencies because such decisions take longer to make (Goldberg, 2019:26), then the optionality context in Example (2) should have been a "difficult" choice, while the optionality context in Example (3) should have been an "easy" choice.[4]

To summarize, it is not unreasonable to hypothesize that choosing between grammatical alternatives requires some cognitive effort and that some choices may require more cognitive effort (with concomitant production difficulties) than others. Below, we endeavor to test the above hypothesis in a corpus of naturalistic spoken data. We specifically synthesize methodologies developed by the authors over the past few years. The aim is to investigate the link between production difficulty/suboptimality and grammatical optionality using a corpus-based psycholinguistics research design in the spirit of, for example, Levy and Jaeger (2007). Specifically, we analyze a sub-section of the well-known SWITCHBOARD, a corpus of telephone conversations between speakers of American English. This sub-section is essentially identical to the sub-section investigated in Gardner et al. (2021); we do not cover the entire SWITCHBOARD, as in Ma et al. (2025), because the present study requires extensive manual annotation for probabilistic conditioning factors (see following sections for details). On a turn-by-turn basis, we check whether grammatical optionality contexts (i.e., variable contexts) correlate with two established symptoms of increased cognitive load during production: filled pauses (*um* and *uh*) and unfilled pauses (speech planning time). Such hesitation

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

phenomena have been used previously as metrics of relative cognitive effort (summarized by Berthold, 1998; Berthold & Jameson, 1999) and have been attested as more frequent in contexts independently judged to be more difficult, such as when utterances are longer or more syntactically complex (Christodoulides, 2016:211-212; Clark & Wasow, 1998; Cooper & Paccia-Cooper, 1980:79; Ferreira, 1991; Grosjean et al., 1979:68-72; Lickley, 2015:460-463; Oviatt, 1995:29-30; Shriberg, 1996), when the topic of conversation is unfamiliar (Merlo & Mansur, 2004; Smith & Clark, 1993:152-153), when the discursive task is more challenging (Abel, 2015; Freeman, 2015:20; Le Grézause, 2017:67-68; Oomen & Postma, 2001:1001-1002), or when lexical items are low frequency and/or have low contextual probability (Beattie & Butterworth, 1979:208; Tannenbaum & Williams, 1968; see also Tily et al., 2009).

In the corpus, we have annotated 20 grammatical alternations (i.e., linguistic variables) common across varieties of American English. Our dataset covers 7,295 turns containing 7,001 optionality contexts, 2,970 filled pauses, and 41,297 unfilled pauses (totaling 230 minutes of silence). To factor in probabilistic cueing, we annotated all optionality contexts in the dataset for known language-internal conditioning factors. Based on this, we subsequently used multivariate modeling to determine and then assign predicted probabilities to each optionality context.

Our analysis of the dataset is guided by the following research question: **do grammatical optionality contexts that strongly cue variant choice attract fewer production difficulties than grammatical optionality contexts where one variant is not strongly cued?** If the suspicion that unbiased choices are hard(er) is correct, then turns in which optionality contexts are highly cued (i.e., one variant is highly likely, as in Example [3] discussed above) will have fewer dysfluencies than turns with optionality contexts that are not highly cued (e.g., all variants are equally likely, see Example [2] discussed above). Analysis will show that no demonstrable difficulty is detectable in the data, even when probabilistic cueing is included in our modeling. These findings call into question the idea that unbiased (i.e., un- cued, or freer) choices are harder than biased (i.e., cued) choices.

Beyond the core research questions above, our large-scale and systematic analysis of a battery of grammatical alternations observed in thousands of optionality contexts (each subjected to probabilistic modeling) in a large speech corpus enables us to provide three secondary outputs that will be of interest to the variationist community:

- Information about the extent to which 20 different grammatical alternations are differentially modelable (i.e., the extent to which we can calculate good variationist models).
- Information about the extent to which optionality contexts tend to be cued in naturalistic speech (it turns out that most optionality contexts are indeed fairly cued—"free" optionality is rare).
- Information about the extent to which different alternations are more or less likely to attract dysfluencies than others.

This paper is structured as follows. First, we discuss our methodology. Then, we report the results. Finally, we discuss the findings and offer some concluding remarks.

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Table 2. Demographics of the SWITCHBOARD corpus

Dialect region	Age	Gender	Post-secondary education
South Midlands (156)	15-19 (23)	Men (280)	Some (468)
Western (82)	20-29 (185)	Women (240)	None (52)
North Midlands (75)	30-39 (146)		
Northern (73)	44-49 (108)		
Southern (55)	50-59 (54)		
New York City (31)	60+ (4)		
Mixed/Unknown (27)			
New England (21)			

Methods and data

Data

The SWITCHBOARD corpus of spoken American English (Godfrey et al., 1992) is a widely used spoken corpus that consists of 2,438 telephone conversations between 542 American English speakers who, in principle, are strangers to each other. The data were recorded by Texas Instruments between 1989 and 1990. Most recordings last five minutes, totaling 240 hours for the whole SWITCHBOARD corpus. Demographic information about participants' age (15-69 years old), dialect region, gender, and education level (Table 2) is available alongside audio files and time-aligned transcripts as part of this corpus's public distribution.

Because the variationist annotation that our analysis requires is *extremely labor-intensive* (see below), we restrict attention to a fairly homogeneous subset of SWITCHBOARD, comprising young (born in or after 1960) South Midland females ($n = 35$ participating in 296 different conversations). The homogeneity of this subset (which overlaps with the dataset studied in Gardner et al., 2021, albeit with substantially more annotation) minimizes potential language-external confounds (see Wieling et al., 2016). We use individual speaker turns as the unit of analysis in this study, which yields 7,295 data points (observations).

Speech dysfluencies and control variables

There is extensive previous literature on dysfluencies in SWITCHBOARD (Clark & Fox Tree, 2002; Gardner et al., 2021; Le Grézause, 2017; Schneider, 2016; Shriberg, 1996; Wieling et al., 2016). Here, we continue that line of research by combining overt hesitation markers (filled pauses) and speech planning time (unfilled pauses) into a single metric of "speech dysfluency," which we interpret as a diagnostic of the difficulty incurred to produce an utterance, that is, its *relative* complexity.

Filled pauses are defined here as all turn-internal uses of *um* and *uh*. This excludes other similar sounding tokens such as *um-hmm* or *uh-oh*. In other words, this assumes that all instances of *um* and *uh* are hesitation markers, rather than tools for discourse organization (Clark & Fox Tree, 2002). We also only consider turns greater than three words, which effectively excludes any use of *um* or *uh* as backchannels or failed

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

attempts at taking over the conversational floor. For turns that have at least one filled pause, we find that there are 2,970 filled pauses spread over 2,176 turns, with an average 1.36 filled pause per turn. We interpret unfilled pauses as speech planning opportunities. These were identified using the built-in "Sound: To TextGrid (silences)" script in *Praat* (Boersma & Weenink, 2023). The main function of this script is to detect silence intervals in audio streams. We define silence as any part of the audio stream below 50 dB and longer than 130 ms, consistent with Hieke et al. (1983) and Gardner et al. (2021). Turns that have unfilled pauses have on average 1.89 s of total turn-internal silence (ranging from 0.002 s to 13.27 s), totaling to more than 230 minutes even in our restricted sample.

Unlike some previous SWITCHBOARD-based dysfluency studies (see Table 1), we rely on a unitary measure of speech dysfluency—an operationalization introduced in Ma et al. (2025) (though see also Oviatt, 1995; Shriberg, 1996). Calculating the measure consists of three steps. First, the number of unfilled pauses per turn are counted rather than their durations being summed (as in Gardner et al., 2021). On average, there are 5.68 unfilled pauses per turn in turns that contain at least one unfilled pause. Second, filled pauses and unfilled pauses are min-max scaled so that they lie in the same interval [0, 1]. This transformation is necessary before combining them into a unitary variable because unfilled pauses outnumber filled pauses significantly ($n = 41{,}297$ versus $n = 2970$) and have a wider range ([1, 26] versus [0, 5], see Ma et al., 2025). Third, the min-max scaled filled and unfilled pauses are added together to produce a standardized continuous measure of dysfluency, which we then use as our dependent variable.

The major benefit of this unitary measure is that speech dysfluency can be modeled as a single dependent variable using (mixed-effects) linear regression instead of having to calculate two parallel models for each kind of dysfluency. After all, dysfluency triggered by cognitive overload may surface either as a filled or unfilled pause.[5]

To predict the dependent variable, our dysfluency prediction models contain three control variables: **speech rate, turn duration**, and **mean character length of all words in a turn**, which previous studies show to be significant predictors of dysfluency (see Note 5 and Gardner et al., 2021). These predictors were centered and scaled prior to regression analysis.

Our two test variables, linked to our two research questions, are the presence of optionality contexts and the extent to which these optionality contexts are cued for a particular variant.

Alternations and annotations

We annotated the SWITCHBOARD subset for 20 grammatical alternations (i.e., grammatical variables), as in previous work of ours (Table 1). These are the "usual suspects" in the literature on grammatical variation in American English (and beyond). They range from syntactic (the dative alternation, the genitive alternation) to lexicogrammatical (deontic modality). The 20 grammatical alternations under analysis are summarized in Table 3. We extracted all licit variants within each envelope of variation (for example, we considered no less than seven future temporal reference variants; see the supplementary materials at https://osf.io/53rbz for details). That said, for the sake of calculating predicted probabilities, multinomial alternations were re-factored

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Table 3. Summary of alternations with distribution of major variants considered for 35 young women from the South Midlands Dialect Area in SWITCHBOARD

Alternation[6]	Variants	*n*	Total *n*
1. **Indefinite pronouns with singular human reference**	-one -body	125 249	374
2. **Pronoun coordination and case**	matching non-matching	45 10	55
3. **Complementation: *that* vs. zero**	that zero	303 1037	1340
4. **Complementation: infinitive vs. gerund**	infinitive gerund	36 34	70
5. **Complementation: *that* vs. gerund**	that gerund	22 16	38
6. ***Try* complementation**	try to try and	5 62	67
7. **Verba dicendi complementation**	VD about VD of	251 118	369
8. **Particle placement**	joined split	103 216	319
9. **Dative alternation**	ditransitive prepositional	60 15	75
10. **Genitive alternation**	*of*-genitive *s*-genitive	47 39	86
11. **Restrictive relativizers**	that zero	180 126	306
12. **Non-restrictive relativizers**	that wh-	192 59	251
13. **Comparatives: synthetic vs. analytic**	synthetic analytic	33 10	43
14. **Plural existentials**	there is there are	45 17	62
15. **Future temporal reference**	will be going to/gonna	642 381	1023
16. **Deontic modality**	have to other	423 125	548
17. **Stative possession**	have have got	430 326	756
18. **Quotatives**	be like say other	54 62 65	181
19. **Negation type**	no not…any	122 150	272
20. **Contraction with negation**	auxiliary contraction *not*-contraction	311 455	766
TOTAL			7001

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

to binary or ternary variables as shown in Table 3 (see section on probabilistic cueing below).

Below we exemplify our annotation protocol using three alternations: *that* versus zero complementation; particle placement; and future temporal reference. A complete and fully referenced coding protocol covering all 20 alternations is provided as supplementary material at https://osf.io/53rbz. This supplementary material also includes clear inclusion/exclusion criteria for each alternation.

Language-internal constraints necessary for calculating cueing strength per optionality context were annotated as follows: per alternation we consulted up to the three most recent multivariate and/or probabilistic studies to identify known variants and constraints, which in most cases amounted to about five constraints per alternation. We note in passing that manual annotation for the 7,001 optionality contexts under study here took more than 200 person-hours.

Alternation #3—Complementation: that *versus zero*

(4) **Variant 1: *that*-complementation** (303 observations)

[…] I know **that**, for instance, in Saudi Arabia, women are just […] (Woman b. 1964, South Midlands Dialect Area, SWITCHBOARD)

(5) **Variant 2: zero-complementation** (1,037 observations)

[…] I didn't think ____ I had the desire to be a mother […] (Woman b. 1961, South Midlands Dialect Area, SWITCHBOARD)

Constraints annotated: the subject of the matrix clause (*I, you, they, we,* other); the matrix verb lemma (e.g., *think, know, say,* etc.); whether the matrix clause is *I think* (yes, no); length of the embedded clause; and the subject of the complement clause (*I, you, he, she, it, we, they*, other) (Szmrecsanyi & Kolbe-Hanna, 2019).

Alternation #8—Particle placement

(6) **Variant 1: joined, verb→particle→direct object** (103 observations)

[..] that way I could **load up the rest of the wives** and go to wherever the football game is (Woman b. 1965, South Midlands Dialect Area, SWITCHBOARD)

(7) **Variant 2: split, verb→direct object→particle** (216 observations)

if you do not **pick the trash up** […] (Woman b. 1964, South Midlands Dialect Area, SWITCHBOARD)

Constraints annotated: particle in question (e.g., *up, out, down,* etc.); the idiomaticity of the expression (idiomatic, compositional); the concreteness of the direct object

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

(concrete, abstract); animacy of the direct object (animate, inanimate); length in words of the direct object; complexity of the direct object (simple, intermediate, complex); presence of pronouns in the direct object (present, absent); and the definiteness of the direct object (definite, indefinite) (Lee & Mackenzie, 2023; Szmrecsanyi & Grafmiller, 2023; Szmrecsanyi et al., 2016).

Alternation #15—Future temporal reference

(8) **Variant 1: will** (642 observations)
[…] over a couple of generations it **will** all change […] (SWB 2453)

(9) **Variant 2: be going to/gonna** (381 observations)
[…] how can you even afford to- to even think about what-what you**'re going to** do with it […] (SWB 2490)

Constraints annotated: animacy of the subject (animate, inanimate); clause type (main, subordinate, apodosis, protasis); polarity (positive, negative); sentence type (affirmative, negative, interrogative); and subject (*I, you, he, she, it, we, they*, other) (Blondeau et al., 2014; Denis & Tagliamonte, 2017; Gardner, 2017).

Operationalizations

To determine the extent to which individual optionality contexts are cued, we calculated 20 conditional random forest models (Tagliamonte & Baayen, 2012:158-165) in *R* (R Core Team, 2024), one for each alternation under study.

The conditional random forest models (which we use as classifiers here, not as tools to calculate variable importance) were tuned, on an alternation-by-alternation basis, for the number of trees and number of variable splits using the tidymodels workflow (Kuhn & Wickham, 2020) and with randomForest (Liaw & Wiener, 2002) as the engine. Metrics were assessed with *yardstick* (Kuhn et al., 2024), with area under curve (AUC) as the main metric for tuning. We then obtained the predicted probability for each variant per optionality context. In other words, we calculated how probable it was that the actual observed variant would occur in its specific context based on the overall variation pattern in the dataset. These predicted probabilities lie in the interval [0%, 100%], with 0.5 (or 50%) as the midway point between two variants when the alternation is binary. Variants with strong probabilistic cueing have predicted probability values closer to 0% or 100%, while weak or no probabilistic cueing for binary alternations yields predicted probabilities close to 50% (equivalent to a 50/50 chance of either variant occurring). In other words, if the model predicts that a given variant, say the ditransitive dative variant, has a probability of 93% in a given context, then that means the model is relatively sure that this variant is favored. But if the probability is 55%, the model has a harder time predicting variant choice. The corresponding weak/no cueing mark for a ternary alternation (alternation #18 [Quotatives]) is 33%.

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

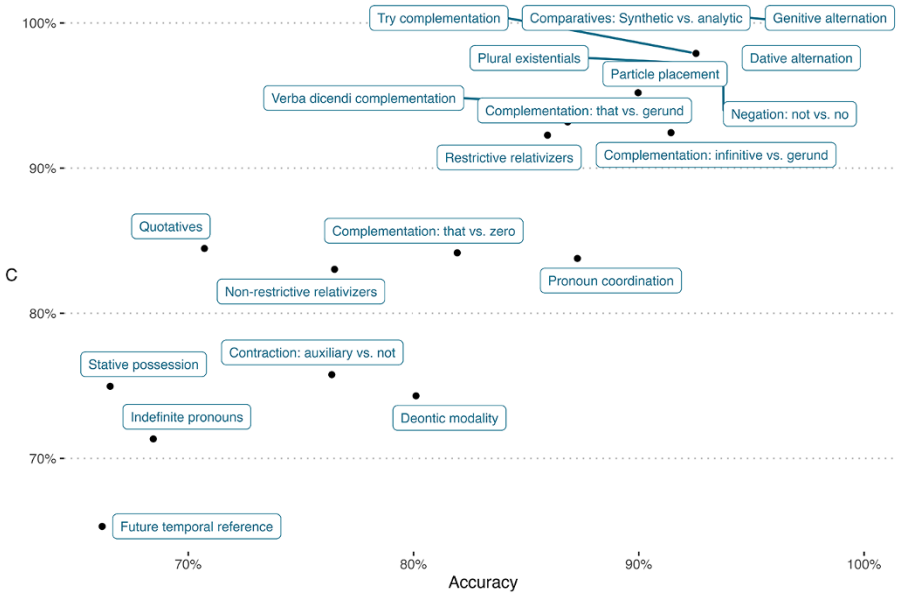


Figure 1. Conditional random forest metrics per alternation. Models were tuned for number of trees and iterations. Metrics include accuracy (x-axis) and concordance *C*-value (y-axis). For binary models, the *C*-value is the same as the area under curve AUC value; for the ternary alternation #18 (Quotatives), we made use of AUNP, that is, a macro-weighted multiclass metric for calculating the area under the curve for each class against the rest, using the a priori class distribution.

Because the strength of probabilistic cueing is a function of distance between predicted probabilities and midway points (50% in binary modeling or 33% in ternary modeling), we calculate the strength of probabilistic cueing as the absolute deviance from the midway point. This means that, in the case of binary alternations, a predicted probability of 93% has a deviance from the midway point of $|50\% - 93\%| = 43\%$. For the ternary alternation, a dominant variant with a predicted probability of 93% has a deviance of $|33\% - 93\%| = 60\%$.

Before continuing, it is instructive to examine the metrics produced by the random forest modeling, such as accuracy ("can the model correctly predict variant choice"), *n* correct predictions/total predictions, concordance *C* (or area under the ROC curve, "can they discriminate well [Levshina, 2015:259] between variants?"), and the distribution of deviance values.

Figure 1 plots accuracies of the 20 models against discriminative power (concordance *C*). The accuracy ranges from exceptionally good (alternation #9 [Dative alternation]: 97% correctly predicted, #10 [Genitive alternation]: 100%) to suboptimal (alternation #15 [Future temporal reference]: 66%, #17 [Stative possession]: 67%).[7] Most models adequately

discriminate ($C \geq 0.7$, except future temporal reference where $C = 0.65$), based on the scale proposed by Hosmer and Lemeshow (2000:162). The *C* value for the ternary alternation #18 (Quotatives) was obtained through an AUNP algorithm—"area

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

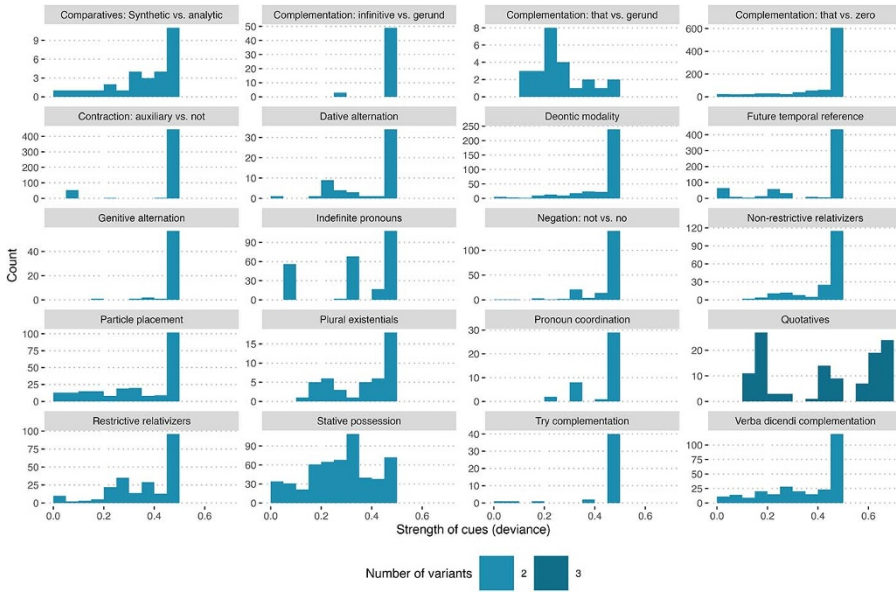


Figure 2. Histogram of deviance values per alternation. Notice that most alternations have a peak at the 0.5 mark.

under the ROC curve of each class against the rest using the a priori class distribution" (see Ferri et al., 2009:30). Discriminative power C is positively correlated with accuracy (Pearson $r = 0.92$). This plot puts the better-understood or better-modelable alternations in the top-right corner, regardless of how many observations we have, and the alternations that are notoriously harder to model in the bottom-left corner: for example, alternation #15 (Future temporal reference) (Blondeau et al., 2014; Denis & Tagliamonte, 2017; Gardner, 2017; Mikkelsen & Hartmann, 2022; Poplack & Tagliamonte, 2000).

The second visual checkpoint to judge the quality of the conditional random forests consists of inspecting the distribution of deviance per alternation (Fig. 2). Values close to |50%| (|66%| for quotatives) indicate that the models are very confident about predicted outcomes. By contrast, values that are closer to |0%| indicate that the variants are equally likely to be selected. What immediately stands out from Figure 2 is that most histograms are negatively skewed, that is, their tail is on the left side of the distribution (e.g., alternation #13 [Comparatives: synthetic versus analytic]), suggesting that most observations in the data are quite constrained. Furthermore, most values are clustered around the |50%| mark, which means that the typical optionality context is fairly strongly cued. There are, however, exceptions, such as alternation #5 (Complementation: *that* versus gerund) and alternation #17 (Stative possession). The distribution of the ternary alternation #18 (Quotatives) also shows a mixed pattern. These atypical distributions might be related to the nature of the constraints that govern them; these three alternations only have a few constraints as per the literature. More follow up work is needed here, in the spirit of Ma et al. (2025).

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Modeling

To address our research question, we use a three-step modeling pipeline:

Step 1: To set the stage, we calculate a comprehensive mixed-effects linear regression model with speech dysfluency as the dependent variable; number of optionality contexts per turn, turn duration, speech rate, and word length as fixed effect predictors; and speaker as a random effect. Turns included in the model = 7,295. We note that this comprehensive model considers (as a control group) many turns in which there are no grammatical optionality contexts.

Step 2: We focus on turns in which one and only one optionality context is observed ($n = 1749$, or 26% of the original dataset analyzed in Step 1). This baseline model is constructed with only the three control variables (turn duration, mean word length, and speech rate) as fixed effect predictors. Given that the number of optionality contexts is constant (= 1), unlike in the comprehensive model (Step 1), number of optionality contexts is not included as a predictor in this model. This mixed-effects linear regression model also includes speaker as a random effect.

Step 3: We enhance the Step 2 baseline model by adding probabilistic cueing as a fixed effect predictor and assess whether this adds explanatory power to the model.

Regression analysis models the relationship between the response (dependent) variable and one or more explanatory (independent) variables. In the case of two or more independent variables, as in the present study, we can estimate the effect of each individual independent variable while controlling for the other independent variables (Levshina, 2015:141). Step 1 sets the stage for addressing our research question in Steps 2 and 3. Recall that we are asking whether degree of probabilistic cueing predicts dysfluency (i.e., are freer choices harder to produce?) We restrict our attention to turns with only one variable context because turns with multiple optionality contexts (where each context is potentially cued to varying extents) would introduce hard-to-manage confounds. We compare the predictive power of the baseline model (Step 2), with just the known predictors of dysfluency, to that of the enhanced model (Step 3) that also includes a measure of the probabilistic cueing of the one optionality context in each turn. If the enhanced model is better at predicting dysfluencies, we can conclude that the degree of probabilistic cueing of an optionality context does influence how hard a turn is to produce.

The data and scripts used in the analysis are all available in the OSF repository at https://osf.io/53rbz/.

Results

Below we discuss outputs from the three-step modeling approach. Subsequently, we explore how different alternations are more or less likely to attract dysfluencies.

Step 1: The comprehensive model

The comprehensive model (Table 4) shows that all four predictors in the model are significant. The number of optionality contexts per turn has a *negative* effect on speech

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Table 4. Mixed-effects linear regression model with min-max scaled speech dysfluency as dependent variable and number of alternations per turn, turn duration, mean word length, and speech rate as fixed effects predictors and individual speaker as a random effect

Fixed effects	Coefficient	Std. Error	*t*	*p*-value
Intercept	0.20	0.01	15.10	< 0.001
n of optionality contexts	-0.01	0.00	-5.79	< 0.001
Turn duration	0.05	0.00	79.82	< 0.001
Mean word length	-0.04	0.00	-15.38	< 0.001
Speech rate	-0.04	0.00	-14.66	< 0.001
Random effects	Variance			
Speaker (intercept)	0.005			

$n_{observations} = 7295$. $n_{speakers} = 35$.
Marginal $R^2 = 0.466$, Conditional $R^2 = 0.552$, AIC = -5959.2, variation inflation factors < 1.42. All fixed effect predictors centered and scaled.

dysfluency. In other words, as the number of optionality contexts per turn increases, the number of dysfluencies decreases—more optionality makes speech *more* fluent, not less fluent. In any event, optionality does certainly overall not attract dysfluency, in line with, for example, Gardner et al. (2021) and Ma et al. (2025).

As for the control variables, mean word length and speech rate have a negative effect, as well. Turns with longer words or produced faster also coincide with fewer dysfluencies. Turn duration, however, has a positive effect, that is, longer turns have more dysfluencies, again in line with previous work. It is perhaps not surprising that longer turns offer more opportunity for speech dysfluency to occur (Oviatt, 1995:29-30). When people speak faster or produce longer words on average, however, there is less space for such opportunity (Clark et al., 2002; Engelhardt et al., 2013, 2019; Goldman-Eisler, 1968; Swerts, 1998). The *R*-squared values indicate good model performance. In sum, even when speech dysfluency is operationalized as a unitary measure and number of grammatical optionality contexts is used as a predictor, there is no significant positive correlation between optionality context and speech dysfluencies.

Step 2: The reduced baseline model

Table 5 shows that in a dataset covering only turns with one optionality context, the three control variables (turn duration, speech rate, and content complexity) behave similarly to how they behave in the comprehensive model. The *R*-squared values of the baseline model remain relatively high, though they are lower than in the comprehensive model. Note that in contrast to the comprehensive model's by-speaker intercept adjustments, this model uses by-item (i.e., by-alternation) intercepts. The reason we took this modeling route is that by-speaker variance was so small it led to singular fit. Of course, we recognize that in an ideal scenario, both speaker and alternation type should have been included in the random effects structure of this model. By only opting for by-alternation intercepts, we can explore whether particular alternations significantly attract or repel dysfluencies in the third model (see below).

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Table 5. Mixed-effects linear regression model with min-max scaled speech dysfluency as dependent variable; turn duration, mean word length, and speech rate as fixed effect predictors; and alternation type as a random effect

Fixed effects	Coefficient	Std. Error	*t*	*p*-value
Intercept	0.18	0.02	9.84	< 0.001
Turn duration	0.05	0.00	35.60	< 0.001
Mean word length	-0.08	0.01	-8.71	< 0.001
Speech rate	-0.06	0.01	-9.73	< 0.001
Random effects	Variance			
Alternation type (intercept)	0.0006			

$n_{observations} = 1749$. $n_{speakers} = 20$. Marginal $R^2 = 0.436$, Conditional $R^2 = 0.446$, AIC = -1026.7, variation inflation factors $<$ 1.10. All fixed effect predictors centered and scaled.

Table 6. Mixed-effects linear regression model with min-max scaled speech dysfluency as dependent variable, and turn duration, mean word length, speech rate, and turn mean deviance as predictors

Fixed effects	Coefficients	Std. error	*t*	*p*-value
Intercept	0.18	0.02	8.27	< 0.001
Turn duration	0.05	0.00	35.60	< 0.001
Mean word length	-0.08	0.01	-8.71	< 0.001
Speech rate	-0.06	0.01	-9.74	< 0.001
Deviance	-0.01	0.03	-0.19	0.852
Random effects	Variance			
Alternation type(intercept)	0.0006			

$n_{observations} = 1749$. $n_{speakers} = 20$. Marginal $R^2 = 0.436$, Conditional $R^2 = 0.446$, AIC = -1024.7, variation inflation factors $<$ 1.11. The model was run with by-alternation varying intercepts.

Step 3: The enhanced baseline model

In the enhanced baseline model, we recreate the model in Step 2 but add probabilistic cueing as an additional fixed effect predictor. The enhanced model is displayed in Table 6. Deviance (i.e., the extent to which the observed variant was cued in its optionality context) has a negative coefficient, suggesting more deviance (i.e., more cueing) coincides with less dysfluency; however, the coefficient, -0.006, (which represents the change in likelihood of the dependent variable when the predictor increases by one unit) is non-significantly different from 0 or no change in likelihood. The other control variables are virtually the same as in the preceding reduced model (Table 5). The inclusion of deviance does not improve the goodness of fit of the model. It is no more predictive than the model in Table 5. A model of the same data with a lower AIC is considered to be more predictive; however, an AIC of -1,024.7 (Table 6) and -1,026.7 (Table 5) are virtually the same (x^2 (1) $= 0.0349$, $p = 0.85$). This answers our central research question: consideration of probabilistic cueing does not buy us any explanatory mileage.

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

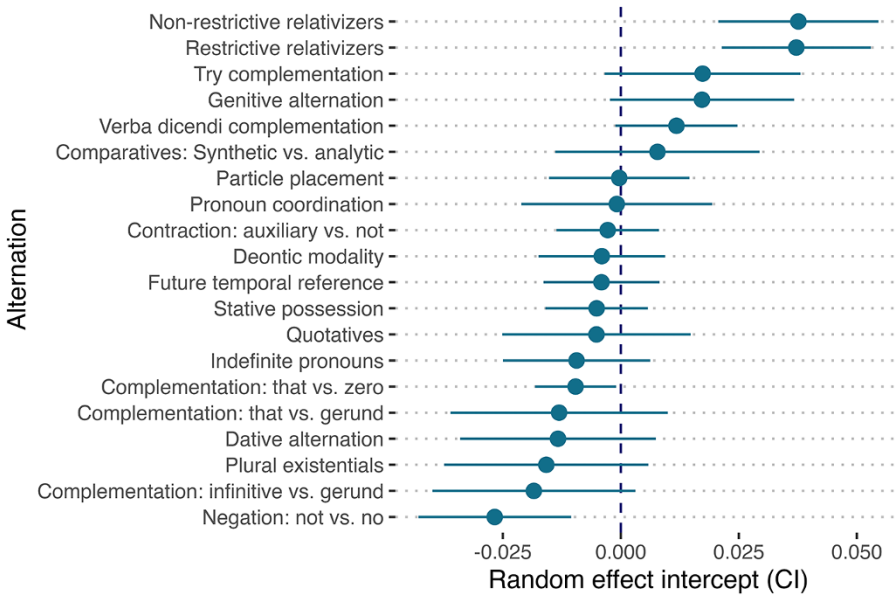


Figure 3. Estimates and confidence intervals (estimates ± standard error) for adjustments to intercept by grammatical alternation type in the enhanced baseline model. Response variable: number of dysfluencies by turn.

Are all alternations equal?

Table 4 shows that, overall, grammatical optionality does not attract production difficulties, but there are perhaps subtle differences between individual grammatical alternations, which may either attract or repel dysfluencies. Conveniently, the two reduced models (Tables 5-6) include alternation type as a random effect (but see our comment regarding the random effects structure above), such that different types of alternation are allowed adjusted intercepts in the model.

Figure 3 plots the intercept adjustments in the enhanced baseline model and thus generates a ranking that can be interpreted as follows: alternations whose estimates are located to the right of the dotted line are more likely to attract dysfluencies, all other things being equal, than alternations whose estimates are located to the left of the dotted line. We note that the intercept adjustments are relatively small and for all but four alternations the overall intercept lies within the 95% confidence interval (represented by the horizontal error bars), indicating the adjusted intercept cannot be statistically verified as different from the overall intercept.

However, both alternation #11 (Restricted relativizers) and #12 (Non-restrictive relativizers) coincide with a higher level of dysfluency than other alternations, while alternation #3 (*That* versus zero complementation) and #19 (Negation: *not* versus *no*) coincide with a lower level of dysfluency than other alternations—the implications of which are discussed in the next section.

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Discussion and conclusion

Consonant with previous work (see Table 1), our findings directly challenge the assumption that optionality is cognitively burdensome for speakers. Our empirical analysis reveals the opposite: grammatical optionality not only fails to induce production difficulties but may correlate with increased fluency, as measured by a significant reduction in speech dysfluency (as in Table 4). This result holds true even when we account for specific features that exacerbate cognitive load, such as speech rate, mean word length, and turn length. We hedge that subject to the limits of our dataset, optionality is not difficult for the particular demographic subset of speakers studied here (young South Midland females) and acknowledge that the link between variation, dysfluency, and socio-demographic differences warrants investigation in future research. That said, we note that research based on the entire SWITCHBOARD corpus (and thus including older and male speakers from all over the U.S.) likewise fails to find dysfluency attraction (Gardner & Szmrecsanyi, 2022; Ma et al., 2025).

Be that as it may, the finding that optionality does not trigger dysfluencies—and may even enhance fluency—is a bit surprising. The reason is, as we explain in the Introduction section, that grammatical optionality is typically conditioned probabilistically by contextual constraints, the processing of which ought to incur cognitive cost on the part of language users. Against this backdrop, we have argued that any additional cognitive inefficiency introduced by having to choose between grammatical alternatives is offset by a number of compensatory benefits, including (a) adjusting explicitness, (b) managing information density, (c) communicating efficiently, (d) establishing Easy First order, (e) achieving rhythmic well-formedness (eurythmicity), (f) domain minimization, and (g) stalling for planning time (see Ma et al., 2025 for a detailed discussion). These benefits, in conjunction with the absence of dysfluency attraction by optionality, we have interpreted elsewhere through the lens of a new Principle of Optionality: "Languages and language users favor the availability of different ways of saying the same thing" (see Ma et al., 2025; Szmrecsanyi et al., in press). The point is that optionality's persistence across linguistic systems and historical contexts suggests that optionality is integral to effective communication. Rather than being an aberration, optionality should be considered as a fundamental feature of linguistic systems, giving speakers flexibility in managing a range of communicative and cognitive demands.

The key finding of this study is the absence of any measurable effect of probabilistic cueing on production difficulty. As discussed in the Introduction section, some theorists (e.g., Goldberg, 2019:26) assume that "free" choices—where no variant is strongly preferred—impose greater cognitive demands than contexts in which one variant is heavily cued. Our analysis finds no support for this hypothesis. The comparison of our baseline and enhanced models (Tables 5-6) shows that how free or constrained a choice between variants is does not predict dysfluency. Even in contexts of low cueing (e.g., close to 50%/50% odds of either variant of a binary variable occurring), speakers do not exhibit more dysfluency, suggesting that decision-making in such scenarios is not inherently burdensome. Turns with freer choices are not "harder" (i.e., attracting more dysfluency) like longer utterances, as we find in our analysis, or highly syntactically complex utterances, as Shriberg (1994) reported for SWITCHBOARD. Taken

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

alongside the negative correlation reported in the model in Table 4, the comparison of the baseline and enhanced model suggests that the cognitive mechanisms underlying linguistic choice are robust, capable of handling complexity without significant detriment to fluency. In short, despite the fact that weakly cued optionality contexts are comparatively rare (see Fig. 2), there is nothing "wrong" with them from a production perspective. Speakers are not inconvenienced by freer choices; there are no "bad" optionality contexts.

Thus, on the whole optionality does not harm fluency. However, Figure 3 indicated that four grammatical alternations appear to deviate from the remaining 16: alternation #11 (Restricted relativizers) and #12 (Non-restrictive relativizers), which coincide with greater dysfluencies, and #3 (*That* versus zero complementation) and #19 (Negation: *not* versus *no*), which coincide with fewer dysfluencies. While each of the 20 alternations varies in salience and prescriptive attention, variation in restricted and unrestricted relativizers is particularly subject to heavy prescriptivist regulation (see, for example, Hinrichs et al., 2015), while #19 (Negation: *not* versus *no*) and #3 (*That* versus zero complementation) elicit more neutral opinions in (North) American English (Childs et al., 2018; Thompson & Mulac, 1991). This contrast suggests a more complex interplay between linguistic structure, prescriptive norms, and cognitive processing, and underscores the need for further investigation into how social and stylistic factors intersect with production fluency.

In conclusion, our study, which employs a battery of probabilistic modeling techniques, demonstrates that grammatical optionality is not a source of cognitive difficulty. In fact, variability is a cornerstone of linguistic competence. Theories that blindly assume difficulties or inefficiencies because of variation are rendered untenable by the evidence presented here. Instead, variation emerges as a robust functional feature of linguistic systems, one that enhances fluency and facilitates adaptive communication—regardless of how (un)predictable linguistic choices are.

Acknowledgements. Funding by the KU Leuven Research Council (grant # 3H220293) is gratefully acknowledged.

Competing interests. The authors declare none.

Data availability statement. Data and code can be found in the supplementary materials at https://osf.io/53rbz.

Notes

1. We note that one reviewer does not agree with this and acknowledge that optionality arguably can have a connotation of "process," which contrasts with the connotation of "product" for variation. However, we consider them as largely synonymous in this study.
2. By "grammar" and "grammatical" we essentially refer to morphosyntax in this paper, that is, all linguistic levels above phonology.
3. This does not necessarily mean generative linguists, who tend to be fairly relaxed about optionality since transformations, or their equivalents in non-transformational approaches, are thought to be able to trigger optionality.
4. Although the difficulty in categorizing choices may be more of an issue for the linguist than for the language user (Kroch, 1994; Tuggy, 1993).

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

5. For an investigation of what factors contribute to hesitation occurring as a filled versus unfilled pause, see Bernaisch (2022).
6. Three alternations included additional low-frequency variants that were omitted from the analysis. There was one token of *try -ing*, four tokens of a *wh-* word as a restricted relativizer, and 15 tokens of future temporal reference expressed using the simple present with or without a temporal adverb.
7. These numbers are in line with the literature: work on the dative alternation has achieved outstanding accuracies (e.g., Bresnan et al., 2007), while even very recent work on the future temporal reference alternation (e.g., Engel & Szmrecsanyi, 2022) struggles to obtain acceptable *C* values.

References

Abel, Jennifer Colleen. (2015). *The effect of task difficulty on speech convergence*. Doctoral dissertation, University of British Columbia.
Anttila, Raimo. (1989). *Historical and comparative linguistics*. Amsterdam: Benjamins.
Beattie, G. W. & Butterworth, B. L. (1979). Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. *Language and Speech* 22(3):201-211. https://doi.org/10.1177/002383097902200301
Bernaisch, Tobias. (2022). Comparing generalised linear mixed-effects models, generalised linear mixed-effects model trees and random forests: Filled and unfilled pauses in varieties of English. In O. Schützler & J. Schlüter (eds.), *Data and methods in corpus linguistics: Comparative approaches*. Cambridge: Cambridge University Press. 163-193.
Berruto, Gaetano. (2004). The problem of variation. *The Linguistic Review* 21:3-4. https://doi.org/10.1515/tlir.2004.21.3-4.293
Berthold, André. (1998). *Repräsentation und Verarbeitung sprachlicher Indikatoren für kognitive Ressourcenbeschränkungen*. Master thesis, Universität des Saarlandes.
Berthold, André & Jameson, Anthony. (1999). Interpreting symptoms of cognitive load in speech input. In J. Kay (ed.), *User modeling: Proceedings of the Seventh International Conference, UM99*. New York: Springer. 235-244.
Blondeau, Hélène, Dion, Nathalie, & Michel, Zoe Ziliak. (2014). Future temporal reference in the bilingual repertoire of Anglo-Montrealers. *International Journal of Bilingualism* 18(6):674-692. 10.1177/1367006912471090
Boersma, Paul & Weenink, David. (2023). *Praat: Doing phonetics by computer* (Version 6.4.01) www.praat.org. Accessed 30 November 2023.
Bresnan, Joan, Cueni, Anna, Nikitina, Tatiana, & Baayen, R. Harald. (2007). Predicting the dative alternation. In G. Boume, I. Krämer, & J. Zwarts (eds.), *Cognitive foundations of interpretation*. Amsterdam: Royal Netherlands Academy of Science. 69-94.
Childs, Claire, Harvey, Christopher, Corrigan, Karen P., & Tagliamonte, Sali A. (2018). Transatlantic perspectives on variation in negative expressions. *English Language and Linguistics* 24(1):23-47. https://doi.org/10.1017/S1360674318000199
Christodoulides, George. (2016). *Effects of cognitive load on speech production and perception.* Doctoral dissertation, Université catholique de Louvain.
Clark, Herbert H. & Fox Tree, Jean E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition* 84(1):73-111. https://doi.org/10.1016/S0010-0277(02)00017-3
Clark, Herbert H. & Wasow, Thomas. (1998). Repeating words in spontaneous speech. *Cognitive Psychology* 37(3):201-242. https://doi.org/10.1006/cogp.1998.0693
Cooper, William E. & Paccia-Cooper, Jeanne. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.
De Smet, Hendrik. (2019). The motivated unmotivated: Variation, function and context. In K. Bech & R. Möhlig-Falke (eds.), *Grammar – Discourse – Context*. Berlin: De Gruyter. 305-332. https://doi.org/10.1515/9783110682564-011
De Troij, Robbert. (2022). *Natiolectal variation in Dutch grammar: A data-driven approach*. Doctoral dissertation, Radboud University/KU Leuven.
Denis, Derek & Tagliamonte, Sali A. (2017). The changing future: Competition, specialization and reorganization in the contemporary English future temporal reference system. *English Language and Linguistics* 22(3):1-28. https://doi.org/10.1017/S1360674316000551
Dik, Simon. (1988). Isomorfisme Als Functioneel Verklaringsprincipe. *Glot* 11:87-106.

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Engel, Alexandra & Szmrecsanyi, Benedikt. (2022). Variable grammars are variable across registers: Future temporal reference in English. *Language Variation and Change* 34(3):355-378. https://doi.org/10.1017/S0954394522000163

Engelhardt, Paul E., McMullon, Mhairi Eg, & Corley, Martin. (2019). Individual differences in the production of disfluency: A latent variable analysis of memory ability and verbal intelligence. *Quarterly Journal of Experimental Psychology* 72(5):1084-1101. https://doi.org/10.1177/1747021818778752

Engelhardt, Paul E., Nigg, Joel T., & Ferreira, Fernanda. (2013). Is the fluency of language outputs related to individual differences in intelligence and executive function? *Acta Psychologica* 144(2):424-432. https://doi.org/10.1016/j.actpsy.2013.08.002

Ferreira, Fernanda. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language* 30(2):210-233. https://doi.org/10.1016/0749-596x(91)90004-4

Ferri, Cèsar, Hernández-Orallo, José, & Modroiu, Ramona. (2009). An experimental comparison of performance measures for classification. *Pattern Recognition Letters* 30(1):27-38. https://doi.org/10.1016/j.patrec.2008.08.010

Freeman, Valerie. (2015). *The Phonetics of stance-taking*. Doctoral dissertation, University of Washington.

Gardner, Matt Hunt. (2017). *Grammatical variation and change in industrial Cape Breton*. Doctoral dissertation, University of Toronto.

Gardner, Matt Hunt & Szmrecsanyi, Benedikt. (2022). Um, uh, *and variation in American English*. Paper presented at *International Conference on Methods in Dialectology (Methods XVII)*. Mainz: Johannes Gutenberg-University Mainz, August 1-5.

Gardner, Matt Hunt, Uffing, Eva, Van Vaeck, Nicholas, & Szmrecsanyi, Benedikt. (2021). Variation isn't that hard: Morphosyntactic choice does not predict production difficulty. *PLOS ONE* 16(6):e0252602. https://doi.org/10.1371/journal.pone.0252602

Geeraerts, Dirk. (1997). *Diachronic prototype semantics: A contribution to historical lexicology*. Oxford: Clarendon Press.

Godfrey, John J., Holliman, Edward C., & McDaniel, Jane. (1992). SWITCHBOARD: Telephone speech corpus for research and development. Paper presented at *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. San Francisco, California, March 23-26. https://doi.org/10.1109/ICASSP.1992.225858.

Goldberg, Adele E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. (2019). *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton, NJ: Princeton University Press.

Goldman-Eisler, Frieda. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.

Grosjean, François, Grosjean, Lysiane, & Lane, Harlan. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology* 11(1):58-81. https://doi.org/10.1016/0010-0285(79)90004-5

Haiman, John. (1980). The iconicity of grammar: Isomorphism and motivation. *Language* 56(3):515-540. https://doi.org/10.2307/414448

Hieke, Adolf E., Kowal, Sabine, & O'Connell, Daniel C. (1983). The trouble with 'articulatory' pauses. *Language and Speech* 26(3):203-214. https://doi.org/10.1177/002383098302600302

Hinrichs, Lars, Szmrecsanyi, Benedikt, & Bohmann, Axel. (2015). *Which*-hunting and the Standard English relative clause. *Language* 91(4):806-836. https://doi.org/10.1353/lan.2015.0062

Hosmer, David W. & Lemeshow, Stanley. (2000). *Applied logistic regression*. Hoboken, NJ: Wiley. https://doi.org/10.1002/0471722146

Jaeger, Florian. (2006). *Redundancy and syntactic reduction in spontaneous speech*. Doctoral dissertation, Stanford University.

Kroch, Anthony. (1994). Morphosyntactic variation. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (eds.), *CLS 30*. Chicago, IL: Chicago Linguistic Society. 180-201.

Kuhn, Max, Vaughan, Davis, & Hvitfeldt, Emil. (2024). *yardstick: Tidy characterizations of model performance* (Version 1.3.1). https://CRAN.R-project.org/package=yardstick. Accessed 21 March 2024.

Kuhn, Max & Wickham, Hadley. (2020). *Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles*. https://www.tidymodels.org. Accessed 21 November 2025.

Kusters, Wouter. (2003). *Linguistic complexity*. Utrecht: LOT.

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Labov, William. (1972). *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.
Labov, William. (1978). Where does the linguistic variable stop? A response to Batriz R. Lavandera. *Working Papers in Sociolinguistics* 44:1-22.
Lavandera, Beatriz. (1978). Where does the sociolinguistic variable stop? *Language in Society* 7:171-183. https://doi.org/10/d82gxm
Le Grézause, Esther. (2017). Um *and* uh, *and the expression of stance in conversational speech*. Doctoral dissertation, University of Washington.
Lee, Naomi & Mackenzie, Laurel. (2023). Social role effects on English particle verb variation fail to replicate. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 68(2):329-343. https://doi.org/10.1017/cnj.2023.13
Levshina, Natalia. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: Benjamins. https://doi.org/10.1075/z.195
Levshina, Natalia & Lorenz, David. (2022). Communicative efficiency and the principle of no synonymy: Predictability effects and the variation of *want to* and *wanna*. *Language and Cognition* 14(2):249-274. https://doi.org/10.1017/langcog.2022.7
Levy, Roger & Jaeger, T. Florian. (2007). Speakers optimize information density through syntactic reduction. In B. Schölkopf, J. Platt, & T. Hoffman (eds.), *Advances in neural information processing systems 19*. Cambridge, MA: MIT Press. 849-856.
Liaw, Andy & Wiener, Matthew. (2002). Classification and regression by Randomforest. *R News* 2:18-22.
Lickley, Robin J. (2015). Fluency and disfluency. In M. Redford (ed.), *The handbook of speech production*. Oxford: Wiley-Blackwell. 445-469.
Ma, Ruiming, Van Hoey, Thomas, & Szmrecsanyi, Benedikt. (2025). Isomorphism-inspired theorising about optionality and variation: No empirical support from English grammar. *English Language and Linguistics* 1-21. https://doi.org/10.1017/S1360674325000097
Merlo, Sandra & Mansur, Letícia Lessa. (2004). Descriptive discourse. *Journal of Communication Disorders* 37(6):489-503. https://doi.org/10.1016/j.jcomdis.2004.03.002
Miestamo, Matti. (2008). Grammatical complexity in a cross-linguistic perspective. In M. Miestamo, K. Sinnemäki, & F. Karlsson (eds.), *Language complexity: Typology, contact, change*. Amsterdam: Benjamins. 23-42.
Mikkelsen, Olaf & Hartmann, Stefan. (2022). Competing future constructions and the Complexity Principle: A contrastive outlook. In S. Flach & M. Hilpert (eds.), *Studies in corpus linguistics*. Amsterdam: Benjamins. 9-40. https://doi.org/10.1075/scl.105.01mik
Nichols, Johanna. (2013). The vertical archipelago: Adding the third dimension to linguistic geography. In P. Auer, M. Hilpert, A. Stukenbrock, & B. Szmrecsanyi (eds.), *Space in language and linguistics: Geographical, interactional, and cognitive perspectives*. Berlin: De Gruyter. 38-60.
Oomen, Claudy & Postma, Albert. (2001). Effects of divided attention on the production of filled pauses and repetitions. *Journal of Speech, Language, and Hearing Research* 44:997-1004. https://doi.org/10.1044/1092-4388(2001/078)
Oviatt, Sharon. (1995). Predicting spoken disfluencies during human-computer interaction. *Computer Speech & Language* 9(1):19-35. https://doi.org/10.1006/csla.1995.0002
Poplack, Shana & Dion, Nathalie. (2009). Prescription vs praxis: The evolution of future temporal reference in French. *Language* 85(3):557-587. https://doi.org/10.1353/lan.0.0149
Poplack, Shana & Tagliamonte, Sali A. (2000). The grammaticalization of *going to* in (African American) English. *Language Variation and Change* 11:315-342. https://doi.org/10.1017/S0954394599113048
Proctor, Robert W. & Schneider, Darryl W. (2018). Hick's law for choice reaction time: A review. *Quarterly Journal of Experimental Psychology* 71(6):1281-1299. https://doi.org/10.1080/17470218.2017.1322622
R Core Team. (2024). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/. Accessed 23 December 2024.
Sankoff, David. (1988). Sociolinguistics and syntactic variation. In F. Newmeyer (ed.), *Linguistics: The Cambridge survey*. Cambridge: Cambridge University Press. 140-161.
Schneider, U. (2016). Hesitation placement as evidence for chunking: A corpus-based study of spoken English. In H. Behrens & S. Pfänder (eds.), *Experience counts: Frequency effects in language*. Berlin: De Gruyter. 61-90. https://doi.org/10.1515/9783110346916-004
Shriberg, Elizabeth E. (1994). *Preliminaries to a theory of speech disfluencies*. Doctoral dissertation, University of California, Berkeley.

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press

Shriberg, Elizabeth E. (1996). Disfluencies in Switchboard. In H. Bunnell & R. Foulds (eds.), *International Conference on Spoken Language Processing (ICSLP 96)*. Wilmington, DW: Alfred I. duPont Institute. 11-14.
Smith, Vicki L. & Clark, Herbert H. (1993). On the course of answering questions. *Journal of Memory and Language* 32(1):25-38. https://doi.org/10.1006/jmla.1993.1002
Sundby, Bertil. (1998). Syntactic variation in the context of normative grammar. In M. Rydén, I. Tieken-Boon van Ostade, & M. Kytö (eds.), *A reader in Early Modern English*. Frankfurt am Main: Peter Lang. 475-484.
Swerts, Marc. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30(4):485-496. https://doi.org/10.1016/S0378-2166(98)00014-9
Szmrecsanyi, Benedikt, Gardner, Matt Hunt, Ruiming, Ma, & Van Hoey, Thomas. (in press). Empirical accountability meets theorizing about language variation. In P. Cukor-Avila & S. Tagliamonte (eds.), *Empirical accountability in variationist linguistics: Taking the next step*. Cambridge: Cambridge University Press.
Szmrecsanyi, Benedikt & Grafmiller, Jason. (2023). *Comparative variation analysis: Grammatical alternations in world Englishes*. Cambridge: Cambridge University Press.
Szmrecsanyi, Benedikt, Grafmiller, Jason, Heller, Benedikt, & Röthlisberger, Melanie. (2016). Around the world in three alternations: Modeling syntactic variation in varieties of English. *English World-Wide. A Journal of Varieties of English* 37(2):109-137. https://doi.org/10.1075/eww.37.2.01szm
Szmrecsanyi, Benedikt & Kolbe-Hanna, Daniela. (2019). New ways of analyzing dialect grammars: Complementizer omission in traditional British English dialects. In S. Grondelaers & R. van Hout (eds.), *New ways of analyzing syntactic variation*. Berlin: Mouton de Gruyter. 1-25.
Tagliamonte, Sali A. & Baayen, R. Harald. (2012). Models, forests, and trees of York English. *Language Variation and Change* 24(2):135-178. https://doi.org/10.1017/S0954394512000129
Tagliamonte, Sali A. & Smith, Jennifer. (2005). No momentary fancy! The zero "complementizer" in English dialects. *English Language and Linguistics* 9(2):289-309. https://doi.org/10.1017/s1360674305001644
Tannenbaum, Percy H. & Williams, Frederick. (1968). Generation of active and passive sentences as a function of subject or object focus. *Journal of Verbal Learning and Verbal Behavior* 7(1):246-250. https://doi.org/10.1016/S0022-5371(68)80197-5
Thompson, Sandra A. & Mulac, Anthony. (1991). The discourse conditions for the use of the complementizer that in conversational English. *Journal of Pragmatics* 15(3):237-251. https://doi.org/10.1016/0378-2166(91)90012-m
Tily, Harry, Gahl, Susanne, Arnon, Inbal, Snider, Neal, Kothari, Anubha, & Bresnan, Joan. (2009). Syntactic probabilities affect pronunciation variation in spontaneous speech. *Language and Cognition* 1(2):147-165. https://doi.org/10.1515/LANGCOG.2009.008
Tuggy, David. (1993). Ambiguity, polysemy, and vagueness. *Cognitive Linguistics* 4(3):273-290. https://doi.org/10.1515/cogl.1993.4.3.273
Uhrig, Peter. (2015). Why the Principle of No Synonymy is overrated. *Zeitschrift Für Anglistik Und Amerikanistik* 63(3):323-337. https://doi.org/10.1515/zaa-2015-0030
Van Hoey, Thomas, Szmrecsanyi, Benedikt, & Gardner, Matt Hunt. (in press). Choice and complexity: In naturally occurring data, absolute complexity does not necessarily trigger relative complexity. *Linguistic Typology at the Crossroads*. 1-29.
Wieling, Martijn, Grieve, Jack, Bouma, Gosse, Fruehwald, Josef, Coleman, John, & Liberman, Mark. (2016). Variation and change in the use of hesitation markers in Germanic languages. *Language Dynamics and Change* 6(2):199-234. https://doi.org/10.1163/22105832-00602001

Cite this article: Van Hoey, Thomas, Matt Hunt Gardner, Ruiming Ma & Benedikt Szmrecsanyi. (2026). Unpredictable grammatical choices are not harder than predictable ones. *Language Variation and Change* **38**:91–113. https://doi.org/10.1017/S0954394526100660

https://doi.org/10.1017/S0954394526100660 Published online by Cambridge University Press