Constructions 18 (2026)
doi 10.24338/cons-785

Preterite and present perfect: Towards modelling the connection between usage, (meta)linguistic awareness, norm orientation and constructional knowledge

Marianne Hundt, Bethany Dallas & Laetitia van Driessche

University of Zürich

Abstract

This study examines whether speakers' knowledge of regional variation should inform constructionist models of preterite-present perfect alternation. Combining a maze-task experiment and post-experiment survey, data on usage and metalinguistic knowledge were collected from speakers in England, Scotland, the US, and Canada. While regional preferences emerged for specific contexts, structural priming completely overrode regional differences. Importantly, participants' metalinguistic awareness reflected a cross-Atlantic split inconsistent with actual usage, and norm orientation did not affect constructional choices. These findings reveal mismatches between usage, perception, and normative beliefs. While usage data supports semantic distinctions between perfect subtypes, the disconnect between speakers' behaviour and their knowledge of regional variation raises questions whether sociolinguistic information merits inclusion in constructional templates.

1 Introduction

When English speakers talk about events in the past, they have a choice, in principle, between the preterite ((1) a.) and the present perfect ((1) b.), with the latter prototypically emphasising current relevance in addition to pastness:

(1) a. *Did* you *eat* breakfast this morning?
b. *Have* you *eaten* breakfast this morning?

The present perfect is more frequent in British than in American English (Elsness 2009; Hundt & Smith 2009; Yao 2024), to the extent that the (currently relatively stable) transatlantic difference has been considered a shibboleth (Biber et al. 1999: 463; Crystal [7]2002: 97; Schneider 2025: 51). The question is whether speakers of English are aware of such regional differences and whether this awareness plays a role for the choices they make when they use English. In addition to meta-linguistic awareness of regional differences, choices between constructional variants may also depend on speakers' norm orientation. Moreover, previous research has shown that speakers are more likely to choose a particular pattern if it has occurred in the immediately preceding context, a phenomenon referred to as 'structural priming' in psycholinguistics and 'persistence' in corpus linguistic re-search (e.g. Bock 1986; Pickering & Ferreira 2008; Szmrecsanyi

2005).[1] It could be argued that 'accommodation' in sociolinguistic research studies the same phenomenon, albeit from a slightly different perspective (see also Kootstra & Muys-ken 2019: 13).

Recent work in construction grammar has started to make a case for the inclusion of socio-pragmatic knowledge in the representation of what speakers know about constructions (e.g. Silvennoinen, 2025: 355). The purpose of our case study is to explore whether a connection can be made between patterns of regional variation emerging from usage, on the one hand, and metalinguistic awareness – what speakers think they know about such variation – as well as norm orientation, on the other hand. If such a connection exists, this knowledge should be included in the constructional template for the present perfect (PP) in English. Finally, corpus-based evidence has shown that persistence is an important factor in the choice of the preterite (SP) and PP constructions (Dallas & Hundt, submitted).

As corpus data do not provide information on speakers' metalinguistic knowledge nor their norm orientation, empirical evidence needs to come from controlled experimentation. This study breaks novel ground in combining psycholinguistic experimentation with questionnaires on metalinguistic awareness and norm orientation in World Englishes (WEs). Collecting data from participants in England and the US as well as from two neighbouring regions – Scotland and Canada – allows us to model the dynamics in usage and norm orientation beyond two major norm providing varieties. More specifically, our study is concerned with variable use of PP and SP in English English (EngE), US American English (USEng), Canadian English (CanE) and Scottish English (ScotE) and whether what speakers know about this variation merits inclusion in a constructionist modelling of the variation.

Section 2 briefly summarises the main findings of previous research into variable use of the SP and PP in English (2.1) and presents an earlier suggestion for modelling the perfect from a construction grammar point of view (2.2). It also outlines the main findings from research on metalinguistic awareness and norm orientation in WEs (2.3). On the basis of what we know from existing studies, we formulate our research questions (2.4). Section 3 provides details on the experimental design (3.1) and the survey used to elicit information on speakers' awareness of the target constructions' regional provenance and norm orientation (3.2) as well as data collection, namely the implementation of the research design and recruitment of participants (3.3). It also briefly introduces the statistical approach taken in the analysis of the data (3.4). The results of our study are presented in section 4 and the implications for the modelling of variation in the tense aspect constructions are discussed in section 5.

2 Background

2.1 Variable use of preterite and present perfect in English: What we know

Previous research has investigated longitudinal developments of the two constructions (e.g. Elsness 1997, 2014; Yao 2024), analysed corpus evidence on their use in varieties of English as a first, institutionalised second, and foreign language (e.g. Werner 2014; Fuchs 2016; Yao & Collins 2013; Davydova 2011), studied their distribution across different written and spoken registers (e.g. Biber

[1] This is different from the use of the term 'persistence' to refer to stable variation in Franco & Tagliamonte's (2020) study on the present perfect in Canadian English.

et al. 1999; Bao et al. 2018), and modelled the alternation against sociolinguistic variables (Franco & Tagliamonte 2022).

In general, the overall use of the SP and PP shows stable patterns of variation in the recent past (e.g. Hundt & Smith 2009; Yao 2024: 202). The PP is found significantly less frequently in American (AmE) than in British (BrE) English (see e.g. Hundt & Smith 2009; Bowie et al. 2013; Yao 2024), while BrE has been found to have the highest PP frequency amongst WEs (see among others Seoane & Suarez Gomez 2013: 9).[2] In both BrE and AmE, the PP has been found to be used more frequently in written language and formal registers (Elseness 1997), but other studies have found contradicting evidence that it is more frequent in spoken language (e.g. Werner 2014: 151). More in-depth analyses show that the PP is highly register dependent and appears more frequently in certain registers due to a specific function it fulfils in this specific register, such as the frequent "hot news" function the PP is used for in news reports (e.g. Yao 2024: 150). However, the diachronic proportional use of the PP and SP remains relatively stable both across varieties and across most registers (Hundt & Smith 2009: 57; Yao 2024: 202). Franco and Tagliamonte (2022: 802-803) found the same to be the case in vernacular Canadian English (CanE), with the use of the PP and the SP presenting stable variation and "continuing the longitudinal specialization of the preterit/present perfect contrast" (803). Against this background, interestingly, their preliminary evidence shows that at the micro-constructional level, the combined use of ever with a superlative might be a context in which the perfect is currently expanding its use (Franco & Tagliamonte 2022: 799-800). The (relative) frequency of the PP and its variation with the SP in CanE have not been examined extensively,[3] although it is to be expected that due to the proximity and influence from AmE, the SP:PP variation will fall significantly closer to the US end of the cross-Atlantic usage patterns. Empirical studies on the frequency and distribution in Scottish English are similarly lacking, as previous research has focussed on the types of alternative perfect constructions found in Scots and Scottish Standard English (e.g. Miller 2003, 2016) but has not examined them in corpora. An early study explicitly found the PP to be a more prominent feature of ScotE child language input and production than AmE input and production (Gathercole, 1986: 557). Children receiving ScotE input acquire the PP earlier, including the different meanings of the perfect, indicating that ScotE falls closer to BrE use. However, ScotE is also known to consistently use the SP in certain variable contexts, such as with the adverbials *just* or *there* as in example (2) (taken from Miller 2000: 339):

(2) a. Somebody *just* kicked your car there.
b. Somebody kicked your car *there*.

Previous research thus confirms that despite the stable variation within the varieties, there is regional variation in use and frequency of the PP and the SP on either side of the Atlantic.

[2] Fuchs (2016: 234) finds PP rates in some WEs that approximate those observed for BrE.
[3] An exception is Roy (2014), who examines CanE based on the Quebec English Corpus. As the data examined is limited and the analysis focusses on the percentage of PPs used with adverbials, his dissertation does not allow a comparison with the other varieties of overall perfect frequency and register distribution.

With variation across BrE and AmE being stable, it could be argued that what Vanneck (1958) labelled the 'colloquial preterite' of the type *Did you have breakfast yet?* might provide a good argument for including regional variation as part of speakers constructional knowledge. However, on the basis of anecdotal evidence, Vanneck (1958: 240) claims that Americans are unaware of the regional provenance of the colloquial preterite whereas speakers of BrE avoid it in their usage: "The slang phrase 'You've had it', 'He's had it' etc. was recently imported from England without any change of tense; whereas the American phrase 'You said it!" became "You've said it!' in England." This suggests that hearers of BrE might have a construction in their passive knowledge of English that includes information on its regional provencance. However, avoidance of a construction in a specific context does not automatically mean that speakers are aware of its regional or stylistic flavour, either.

2.2 The SP and the PP: A Construction Grammar account

To date, Michaelis (1993/1998) provides the most comprehensive construction grammar (CxG) account on the PP. CxG claims that, while all linguistic knowledge is saved as constructions, we do not necessarily store fully specified descriptions of each construction separately; instead, constructions are linked in a dynamic network of constructions, commonly referred to as the construct-i-con (see e.g. Goldberg 1995, 2019; Diessel 2023; Trousdale & Traugott 2026). This means that constructions with form/meaning overlap are linked, and a construction A can inherit overlapping information from a construction B, rather than both constructions being fully represented in the construct-i-con to increase storage efficiency and ease retrieval (Michaelis, 1998: 75). In the case of the PP, Michaelis argues that while a general perfect construction exists (which the future, past and present perfect inherit from), there are in fact three individual PP constructions in the construct-i-con that inherit from the general PP construction (1998: 76). Michaelis differentiates between the experiential or existential perfect (EPrP), the resultative perfect (RPrP), and the continuative perfect (CPrP) (1993: 216ff.). The EPrP refers to "an event that occurred one or more times prior to reference time" (Michaelis, 1998: xiii) such as in *I have eaten at that restaurant before*. In contrast, the RPrP "denotes an event that occurred prior to reference time and whose accessible consequences are present at reference time" (Michaelis 1998: xvi), such as I ha*ve done my homework*. The CPrP has a further different nuance in meaning, as it focuses on a state spanning a time period that continues until the reference time (Michaelis, 1998: xii), as in *It has been raining all morning*. Other researchers identify a fourth perfect type, the perfect of recent past (recP), which foregrounds the recency of an action relative to the speaking time, and which Michaelis subsumes under the RPrP category (1998: 157). This paper, in concordance with other researchers, considers this a fourthperfect category. We will return to this point below.[4]

[4] It should be noted that there is contention in the literature as to how (if at all) the PP can be categorised. Particularly, there is a split between monosemous accounts, which claim that only one general meaning underlies the interpretation oft the PP (see Werner 2014: 60-67 for an overview), and polysemous accounts, which claim there are multiple distinct semantic readings oft he PP (see Werner 2014: 67-72 for an overview). It is beyond the scope of this study to conduct an extensive critical review of these accounts; however, our approach aligns with the polysemous account stipulating four distinct readings of the PP, with an underlying overall meaning of current relevance, that is accepted by a significant number of researchers (see table 3.3.1 in Werner 2014: 72).

Michaelis considers the EPrP, the RPrP and the CPrP to be individual constructions, rather than one single perfect construction, as these three constructions not only have distinct senses but also show individual idiosyncratic constraints and collocational preferences that are not compatible with each other and do not allow for subsumption under one single construction (Michaelis 1993: 342). An example is the co-occurrence of the different perfect types with temporal adverbials. On the one hand, all three perfect types reject deictic punctual past time adverbials, and accept the adverbial *now*, features they inherit from a shared overarching general perfect construction. On the other hand, while all perfect types are also all compatible with the adverbial *since*, there is a distinct difference in meaning depending on the exact perfect type it combines with. This arises from the individual constraints of each perfect type (Michaelis 1993: 229 ff.). In addition, there are certain adverbials that only combine with one perfect type, an example being durational adverbials that are only compatible with the CPrP (Michaelis 1993: 229 ff.). Michaelis (1998: 210) provides a visualisation of the constructional network, which we replicate here with the relevant parts for our discussion highlighted in blue.

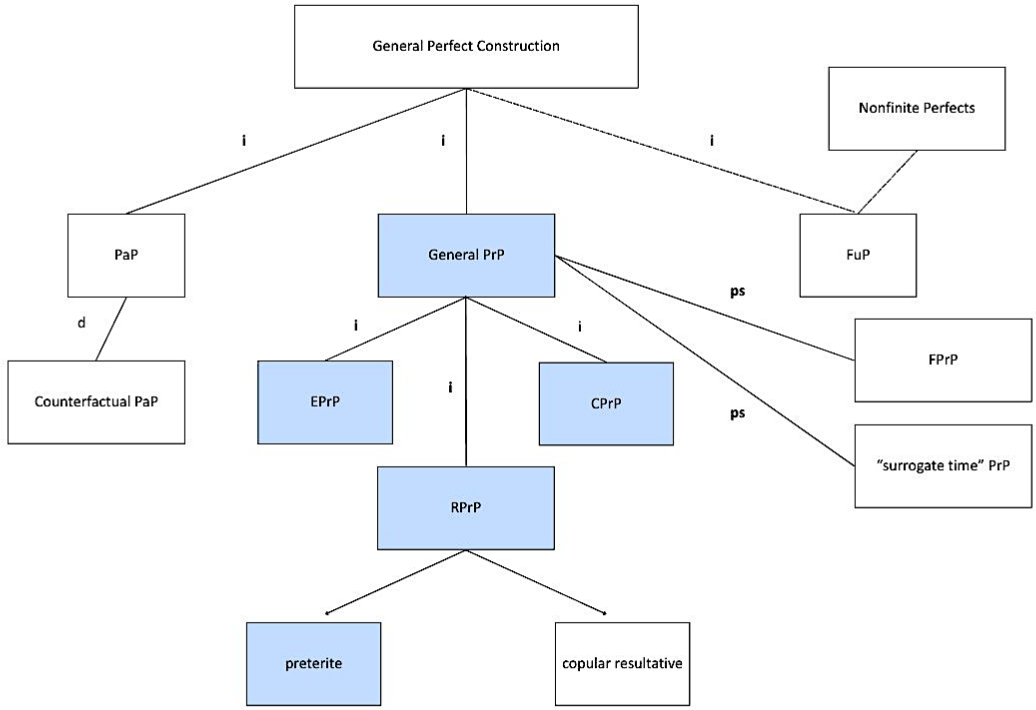

Figure 1: Michaelis' Perfect System (adapted from Michaelis, 1998: 210)

While Michaelis does not explicitly include the preterite/SP in the constructional network beyond an opposition link to the RPrP (see the arrow linking RPrP and preterite in Figure 1),[5] her examination of the PP:SP alternation aligns with other studies (e.g. Davydova 2011: 54ff.) that observe how certain perfect types are more susceptible to constructional variation due to their semantics. Particularly the EPrP is frequently found to alternate with the preterite, as it does not

[5] Michaelis claims that the preterite is used in cases where the resultative perfect cannot be used as the "event proposition denoted represents pragmatically presupposed material" Michaels (1998: 212).

evoke the current relevance of the PP in the same way as, for example, the RPrP does with its explicit focus on the current relevance or result of a past action (Klein 1994: 112).[6] Similarly, the CPrP emphasises the current relevance of the time span more explicitly than the EPrP, making it more likely to trigger the PP. This is also why we deem it prudent to examine the recP as a separate category for this paper: While Michaelis subsumes the recP under the RPrP, the most PP-friendly category, previous research has in fact found the recP to be among the perfect types most frequently substituted with the SP and other constructions (Klein 1994; Miller 2004; Davydova 2011). Davydova (2011) and Klein (1994) argue that it is the least prototypical of the perfect categories as, while it evokes recency, it does not explicitly evoke current relevance (Davydova 2011: 61-62).

While there is little empirical research on the different degrees of PP-friendliness (with the exception of Miller 2004), they seem to be relevant to different WEs. Not only varieties known for substituting the PP with the SP like AmE but also varieties expected to have higher PP frequency like BrE are hypothesised to be more likely to substitute the PP in the contexts with less prototypical current relevance, meaning the experiential perfect and the recent past. This is in line with the few empirical results available (Miller, 2004: 313-315). However, resultative perfective meaning is often expressed with the preterite depending on the exact context and variety, as the examples in (1) show. It is therefore relevant to keep Michaelis' general construction grammar account in mind, as well as the frequency of alternative constructions for the different perfect types, when examining the emergence and awareness of regional usage patterns in our data.

2.3 Metalinguistic awareness, norm orientation, the construct-i-con and priming

Research on metalinguistic awareness in cognitive sociolinguistics and perceptual dialectology shows that speakers possess varying levels of conscious and unconscious knowledge about linguistic variability. Experimental work (e.g. Squires 2016), for instance, reveals that individuals differ in their ability to detect grammatical variation. More specifically, studies in the field tap into participants' ability to notice, mention and mimic regionally-specific features (e.g. Preston 1996, 2016; Niedzielski 1999). These varying levels of awareness of variation correspond to what, in traditional sociolinguistic theory, is modelled in terms of the well-known tripartite distinction of 'indicators', 'markers' and 'stereotypes' (Labov 1972). These are characterised by sociolinguistic salience, that is the degree to which speakers are aware of their social relevance. Additionally, indicators in this view are not subject to style shifting but simply vary with social stratification. Markers, on the other hand, correlate with some social characteristic or identity; they show both social and stylistic stratification and overtly emblemize social values (as identity carriers). Stereotypes, finally, not only have well-known social meanings, but are generally stigmatized and often avoided. This relatively static view of the relation between linguistic variables and their social meaning has been reconceptualised in third-wave sociolinguistics, notably in work by Eckert (2008, 2012), who considers variables to have potential rather than fixed meanings:

[6] In contrast, the continuative perfect, which is used in reference to repeated or continued actions from a point in the past until now, is frequently substituted with the present in many non-standard and standard varieties of English such as Irish English (Davydova 2011: 59).

> The meanings of variables are not precise or fixed but rather constitute a field of potential meanings – an indexical field, or constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable. The field is fluid, and each new activation has the potential to change the field by building on ideological connections. Thus variation constitutes an indexical system that embeds ideology in language and that is in turn part and parcel of the construction of ideology (Eckert 2008: 454).

Crucially, Eckert's view on the meaning of variation shifts the focus from the empirical study of frequency distributions by sociolinguistic variables (age, gender, class, register, etc.) to the speaker as a meaning-making agent, and this more agentive view of how social meaning is situationally created presupposes both conscious knowledge of the differences and a certain degree of control over their use. Thus, speakers' choices give rise to distinct styles that serve as tools for constructing identity (see Drager 2015 for an exemplar-based account of this process). In her discussion of what the different levels of language contribute to social meaning making – from morphosyntax to voice pitch – she assigns morphosyntax to the relatively fixed or 'public' end of the continuum, a view that might be influenced by the fact that she juxtaposes the choice of standard and non-standard variants for morphosyntactic categories like negation and agreement. According to her, the non-standard choices

> … are highly enregistered, functioning as shibboleths, with quite fixed social meanings associated with external facts like class and particularly education. In situated discourse, orders of indexicality that emerge from these external facts are generally consciously and ideologically related to class, such as toughness and alienation (Drager, 2018: 190).

The situation might be different with less salient variation as the one between a preterite and a present perfect. However, most WEs research has focused on usage rather than what speakers explicitly know about variety-specific features, leaving a significant gap in understanding how metalinguistic awareness links to usage.

Norm orientation is central to understanding language use and identity construction within the WEs paradigm (e.g. Schneider 2007). Earlier work emphasised a shift from dependence on varieties like BrE or AmE as norm providing toward local, endonormative models in a process of linguistic and sociopolitical emancipation. More recently, norm orientation has come to be viewed as dynamic, context-dependent, and multi-normative, reflecting the complexity of WEs in their local and global ecologies (e.g. Deshors & Gilquin 2018), a view that is also more in line with the agentive, speaker-centred, situated view of the third wave in sociolinguistics. Empirical evidence of such shifting orientations remains scarce, as corpus-based studies capture usage but not speakers' beliefs or attitudes (cf. Hänsel & Meer 2023), but where such studies exist they do not link back to individual informants' usage.

What speakers know about language is also a concern of research in usage-based construction grammar. There is experimental evidence that speakers can access constructional meaning both on a conscious (e.g. Bencini & Goldberg 2000) and a subconscious level (Johnson & Goldberg 2013). The latter type of experiment is relevant in the context of the present study because it shows the effect of priming on participants' ability to decide whether a sequence of letters

was a word or a non-word: Their decision times were significantly shorter if they had been primed with a construction that was congruent with the target word's argument structure construction than in situations where the prime consisted of an incongruent argument structure construction. In other words, just like sociolinguistic awareness, constructional knowledge can be shown to operate at the level above and below consciousness. Goldberg (2019: 43-45) comments on the fact that argument structure constructions exhibit regionally-specific preference patterns and argues that these kinds of differences, in turn, shape speakers' knowledge of the constructions "in dynamic ways" (p. 45; see also Schmid 2020: 2).

Finally, priming has increasingly gained traction in cognitive sociolinguistics as a possible competing explanation for variation, challenging the concept of style construction (Eckert & Rickford 2001; Eckert 2018). A key assumption underlying style construction is that speakers consciously control their selection of variants and deploy them deliberately – a point that Sandow (2022) develops in his 'attention-to-self' reformulation of the concept. Sociolinguistic research has largely treated observed occurrences in data as independent, when in reality they rarely are. Incorporating recency or persistence into the analysis considerably complicates the picture. As Clark (2018: 731) notes, priming evidence "open[s] up an intriguing possibility: that previous work invoking style as an explanation could perhaps be equally well, if not better, explained by models of priming."

2.4 Research questions and predictions

With respect to usage, metalinguistic awareness and norm orientation, our study aims to answer the following research questions:

RQ 1: Does regional variability in the choice between a PP and and SP emerge from the forced choice task, is this variability sensitive to different types of perfect contexts, and to what extent are the patterns of regional variation subject to priming effects?

RQ 2: What do participants explicitly know about the regional distribution in the use of PP and SP across the four varieties?

RQ 3: What is the variety that participants overtly orient towards in different contexts and is this norm orientation relevant to their preferred usage patterns?

The overarching question is whether the findings from our experiment and survey data merit inclusion in a constructionist modelling of what speakers of EngE, USEng, ScotE and CanE know about the grammar of the PP and how it relates to the SP.

With respect to our more specific research questions, we make the following predictions:

Prediction 1:

- regional preferences will be sensitive to prototypicality effects, with recent past contexts more likely than the more prototypical PP constructions (i.e. resultatives and persistent situations) to show variability with the SP and therefore more likely to be subject to regional divergence;

PRediction 2:

- knowledge of differences in the use of the SP and the PP in neighbouring speech communities will be asymmetrical, with Scottish

participants being more aware of English preferences and Canadian informants more aware of US preferred choices than vice versa;
- participants who are aware of the regional 'flavour' of the preterite and the PP are less likely to converge on an external norm under structural priming conditions than speakers who do not associate the constructional choices with a particular region (e.g. Squires 2016);

Prediction 3:
- usage and norm orientation will diverge in cases where (a) speakers hold strong beliefs about their own variety, and (b) the pattern has attracted prescriptive comments (e.g. Niedzielski 1999).

3 Methodology

We believe that a combined experimentation- and survey-based approach is better suited to capturing the dynamic nature of norm orientation, metalinguistic awareness and usage across different WEs than separately conducted corpus-based research and attitude surveys. Moreover, it answers the call for more experimental approaches to the sociolinguistics of WEs (see Sharma 2025: 38).

3.1 Eliciting usage data: Task type and materials

We elicit usage data on the choice between a preterite and a present perfect embedded in a self-paced reading experiment. As self-paced reading might allow participants to simply click through the experiment without actually processing the stimuli, we combine it with a grammatical maze task (Forster et al. 2009), where participants read sentences word by word and choose between a correct word and a distractor at each step (see Figure 2). This design ensures deep linguistic processing. In the critical region of the test sentences, participants are presented with the choice between a preterite and a present perfect. The creation of distractors in a maze experiment used to be very labour- and thus cost-intensive. The advent of language models has greatly facilitated the design of such experiments. We created our distractors using Heuser's (2022) T-maze algorithm, which makes use of a transformer language model.

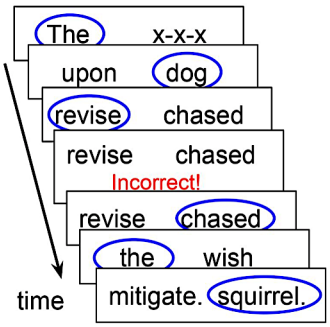

Figure 2: Frames presented as choices in a maze task (https://vboyce.github.io/Maze/intro.html)

The maze task that elicited variable use of preterite and present perfect was part of a larger experiment designed to tap into variable grammatical usage preferences for constructions at different levels of abstraction.

The following stimuli were used in the experiment, with those in (3) used for the baseline condition. Sentences in (4) form the SP-prime condition and those

in (5) the PP-prime condition. The presentation of the choices that correctly continued the sentence and the distractors was randomised across participants.

(3) a. In recent years markets for manufactured items *have become*|*became* fragmented.
b. In some instances their records *have been*|*were* lost.
c. In the past, age, gender and moral worthiness *have been*|*were* used as factors in decision-making.
d. Since I was about nine, *I've*|*had* this interest.

(4) a. For breakfast this morning I *ate* a grape. [preterite PRIME]
b. All afternoon, they *have* (not)|*missed* (not) a single ball. [TARGET][7]

(5) a. This is the highest he *has* ever *scored* in years. [perfect PRIME]
b. In my life I*'ve been*|*was* able to make two crucial choices with autonomy. [TARGET]

The baseline condition elicited data on four different types of potential PPs:[8] Recent past (example (3) a.), resultative (shown in (3) b.) and persistent situation (given in (3) d.); while *in the past* in (3) c. can combine with the PP (under a current relevance reading that implies a contrast past vs. now), the adverbial is more likely to combine with a SP than a PP. Also, as pointed out in section 2.2, the PP is more frequently replaced by a preterite in recent past contexts than in a resultative or a persistent situation (Klein 1994; Miller 2004; Davydova 2011). We would therefore be more likely to see a distinct preference for the SP if it is chosen in these contexts than in a recent past or experiential context.

In the priming condition, the SP prime and target both occurred in contexts that allowed for a persistent situation reading (see examples in (4)). The PP prime comes from an example with a current relevance reading, followed by a target with a persistent situation reading (see (5)).

3.2 Post-experiment survey: metalinguistic awareness and norm orientation

In order to collect information on the participants' knowledge of regional preferences in the choice between preterite and perfect, they were presented with a set of two sentences (see (6) and (7)) and asked to choose the one they thought was the preferred option in EngE, USEng, CanE and ScotE. A third alternative to the SP and the PP – 'no clear preference' – was also offered. Both sets of sentences are examples of the persistent situation subtype. Nevertheless, we model the results on metalinguistic awareness separately for examples (6) and (7). They illustrate different types of perfect context (resultative and experiential) and make use of different temporal adverbials, which might affect participants' perceptions of regional preferences.

(6) a. *Did* you *have* breakfast yet? [resultative]
b. *Have* you *had* breakfast yet?
(7) a. This is the best film I ever *saw*. [experiential]

[7] We are reporting the results of a proof-of-concept study that was designed to test the methodological approach. In a future version, we would avoid negation.
[8] The aim of the study is not to empirically test Michaelis' (1998) perfect system but to tap into the aspect of regional preferences in the potential variation between PP and SP. Hence, we did not focus on the four PP types that she introduces but chose the stimuli with the aspect of regional variation in mind.

b. This is the best film I'*ve* ever *seen*.

The questions on norm orientation took different contexts (formal vs. informal) and modalities (spoken vs. written) into account. Participants were asked to choose the variety that would be their preferred choice when (a) giving a presentation at work, (b) talking to a friend over coffee, (c) writing a business mail, (d) texting with a friend or member of their family. The choices were EngE, ScotE, CanE, USEng and 'other', with the variety labels presented in randomised order. In addition to the four contexts, participants were asked where they thought (a) *the best English* was spoken and (b) where *the most pleasant English* was used, with the four countries (England, Scotland, the US, Canada) and 'elsewhere' being offered as choices. For the 'other' choices, participants were asked to specify the variety/country.

A short questionnaire on participants' gender, age, country of origin, mobility and handedness concluded the data collection. These sociolinguistic variables are included in the multifactorial modelling as they could potentially impact the preference for the PP or SP, norm orientation and or participants' perception of regional preferrences in the use of PP and SP.

3.3 Implementation, data collection and participants

We implemented the experiment via Gorilla (https://www.gorilla.sc). Participants first completed a practice round with two sentences. The stimuli for the experiment were distributed across four blocks of 14 sentences each, with at least one filler sentence presented at the beginning of each block. In two of the blocks, the stimuli for the choice context were immediately preceded by a sentences that primed either the preterite (SP-prime) or the perfect (PP-prime). In the other two blocks, the stimuli were separated by four fillers to avoid immediate structural priming effects, seeing that persistence effects in corpus data can be shown to weaken across three immediately preceding verb phrases (see Dallas & Hundt, submitted). In other words, there were never more than two SP:PP stimuli per block. The order of the blocks was randomised across participants. Overall, they were asked to complete 56 sentences in the experiment, eight of them aiming to collect information on their usage of the SP and the PP. We randomised the position of the words (left-right) that participants were given as options to complete the sentences. They were asked to respond as quickly and accurately as possible and to complete the experiment in a quiet room in one sitting. After each correctly completed sentence, 'Well done!' was provided as feedback. Whenever a wrong continuation was chosen, the participants continued to the next sentence.

For the metalinguistic awareness survey, the order of the varieties was randomised across participants and sets of choice contexts. The two sets of choice contexts were presented as part of the larger survey that collected information on participants' metalinguistic awareness for other constructions (agreement with collective nouns and complementation of antagonistic verbs). Participants therefore did not see the two sets of examples in close proximity.

We recruited participants via Prolific (https://www.prolific.co), with a target of 150 per country and 600 overall; two participants had to be excluded because they did not properly work on the maze task (one from Canada and the other from Scotland) so that the final set of participants included in the study amounts to only 598. In order to keep the number of potential confounding factors as low as possible, we aimed to recruit relatively homogenous samples and decided to filter for first and primary language (English), country of birth (UK, US, Canada), nationality (English, Scottish, US American, Canadian), residence in the

respective countries, ethnicity (white) and age (18-70). Prolific also allows researchers to recruit a balanced sample as regards gender (approximately 50% male and female, see Table 1A in the Appendix). Following the recommendation in Boyce and Levy (2023: 14), we sampled participants with a bachelor's degree as a qualifications cutoff point as this has been shown to decrease the number of random answers in a maze task experiment. Filtering to participants without known neurodiversity was an additional measure taken to increase the number of correctly completed sentences.

Table 1 and Figure 3 provide an overview of the demographic background of the participants included in the study. The questionnaire had asked participants whether they had spent time living abroad in another English-speaking country. We operationalise this variable as a binary one here ('yes' or 'no'). While participants in Scotland turned out to be much more mobile than those in the US, with people from Canada and England falling in between, the age cohorts are relatively comparable across the four samples.

	non-mobile	mobile	Total
Canada	132	17	149
England	130	20	150
Scotland	103	46	149
USA	145	5	150
Total	510	88	598

Table 1: Participants' mobility by country

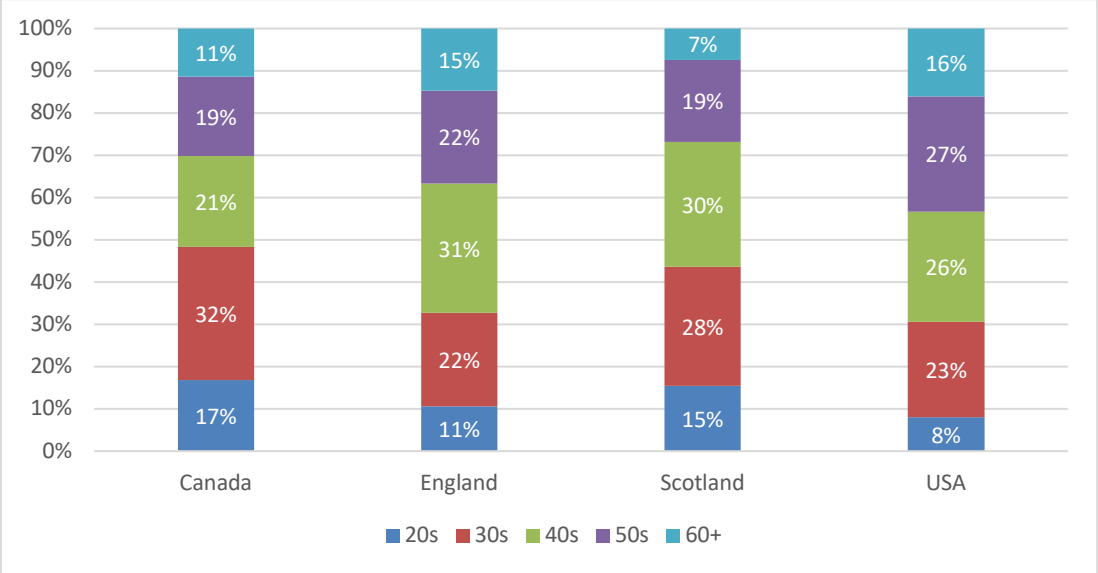

Figure 3: Participants' age by country

3.4 Analysis

In a first step, we take a mixed-modelling approach to participants' baseline use (without priming) across the four varieties. We also provide summary statistics for participants' metalinguistic awareness of regional preferences as well as information on their norm orientation in different situational contexts. In a second step, we take a mixed-modelling approach to see whether metalinguistic know-

ledge and norm orientation impact the choice between a past and a present perfect construction. It is possible to combine mixed modelling with a decision tree approach. The latter applies recursive partitioning to divide the data into increasingly homogenous subgroups at each binary split. Generalised Mixed Model (GLMM) trees or glmertrees thus model interaction between variables while also allowing for the inclusion of random effects such as participant ID (see Fokkema et al. 2018 and 2020). Moreover, the R-package (Fokkema & Zeileis 2019) provides a visualization of the tree that facilitates the interpretation of the results.

Table 2 lists the predictor variables included in the mixed modelling, their labels and the different levels they take on. The ID of the participants was included as a random effect (PARTICIPANT_ID).

Predictor	**Label**	**Levels**
Country of origin	SAMPLE	England, Scotland, USA, Canada
Construction sub-type	PPSUBTYPE	resultative (res), persistent situation (perS), recent past (recP), past
Norm orientation situational	SPEAK_INF, SPEAK_FOR, WRITE_INF, WRITE_FOR	endoNORM, exoENG, exoUS, other
Norm orientation global	BEST_E, PLEASANT_E	endoNORM, exoENG, exoUS, other
Gender	GENDER	male, female, other
Age	AGE	20s, 30s, 40s, 50s, 60 +
Mobility	MOBILITY	y, n
Metalinguistic Awareness	VARIETY	USEng, EngE, CanE, ScotE
Priming effect	PRIMINGTYPE	PP_prime, SP_prime

Table 2: Predictor variables to model variable use of the preterite and the perfect across English in England, Scotland, the US and Canada

4 Findings

4.1 Usage

Figure 4 shows the glmertree for the baseline condition with PPSUBTYPE, GENDER, AGE, MOBILITY and SAMPLE as fixed effects and PARTICIPANT_ID as a random effect.[9] Somers2 returns a good model fit (C = 0.781). The tree reveals that context (PP_TYPE) affects the choice more than the regional variety (SAMPLE). With a past-tense adverbial (in the past), all participants are more likely to choose an SP than a PP (see node 2). Similarly, with an adverb indicating recent past (*in recent years*) participants overwhelmingly choose a PP, regardless of their regional background (see node 4).

[9] See Table 2A in the Appendix.

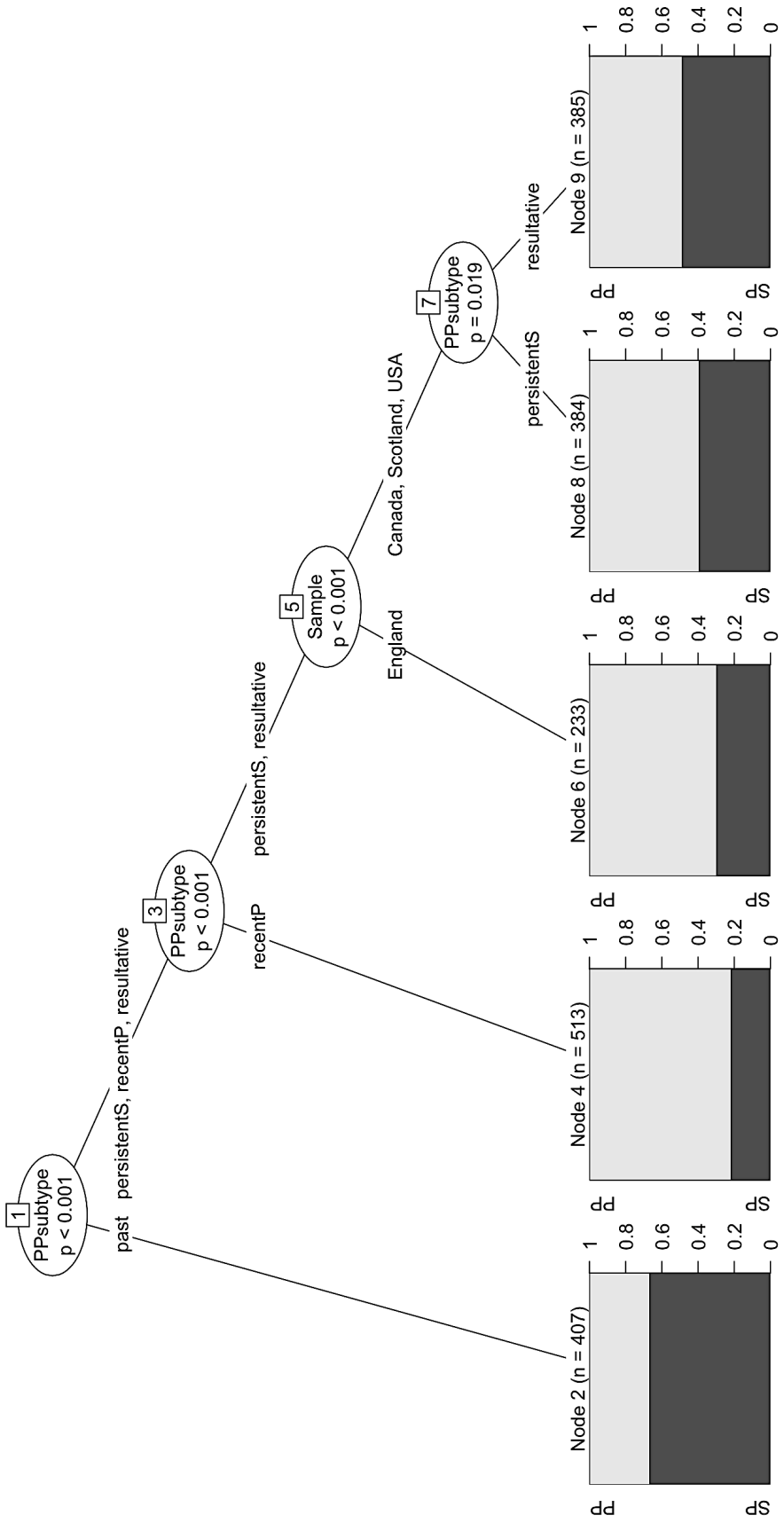

Figure 4: glmertree for usage and sociolinguistic predictor variables (baseline condition)

The regional split in the tree occurs with the SP:PP choice in resultative (res) and persistent situation (persistentS) contexts (see node 5): this is where participants from England (see node 6) differ significantly from those in the US, Canada and Scotland. The latter choose the SP more often than the former, particularly in resultative contexts (see node 9). With respect to the regional split, however, usage patterns in the experiment align with findings from corpus-based studies where they are available, i.e. for USEng and CanE (e.g. Biber et al. 1999, Hundt & Smith 2009, Fuchs 2016). The fact that usage in Scotland aligns more closely with that in the US and Canada might reflect the the long-term effect of earlier substrate influence on the variety. Additional corpus-based research is needed to support this view.

The results from our maze-task experiment surprisingly do not meet the prediction we made in that regional usage preferences reflect prototypicality effects: the less prototypical PP contexts (recent past) do not show regional divergence whereas resultatives and persistent situations, which are more prototypical, do show regional divergence. The adverbial phrase *in recent years* in baseline sentence (3) a. for the recent past may have something to do with this as it is clearly associated with the PP. The adverbial phrase *in some instances* in baseline sentence (3) b., by contrast, is somewhat less clearly associated with the resultative. Finally, for the persistent situation type, baseline sentence (3) d. combined the clear temporal adverbial *since* with a preterite in the initial clause of the sentence, making it a less ideal choice than a context without a verb.

Interestingly, none of the other social variables (AGE, GENDER, MOBILITY) is selected by the model shown in Figure 4. On the one hand, this confirms the results of previous, corpus-based research that have shown the choice between the preterite and the present perfect to be relatively stable across time. Had there been recent or ongoing change, we would expect to see an apparent-time effect to emerge from the AGE factor in the model. With ongoing change, GENDER and MOBILITY are likely factors to emerge as additional predictors (with women often at the vanguard of ongoing change and mobile participants more likely to change than people who are non-mobile, see e.g. Labov 2001, Britain 2016).For the mixed modelling of the priming condition, we only included those instances where participants had actually completed the prime. Interestingly, under the priming condition, the regional variety does not appear to have any influence on the constructional choice, as the glmertree in Figure 5 shows (with Somers2 returning a very good model fit at $C = 0.840$). In other words, structural priming appears to have a swamping effect on regional preferences in the choice between SP and PP. Again, none of the sociolinguistic variables were selected by the model.

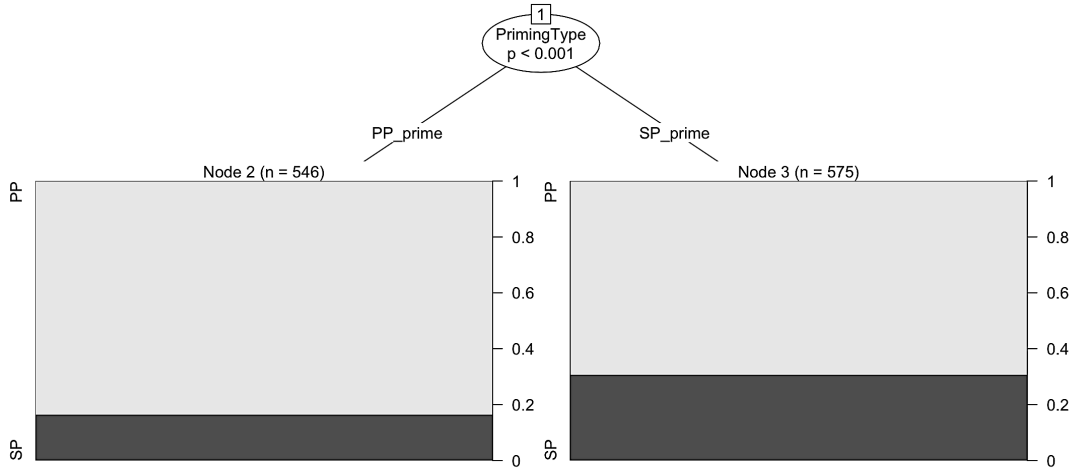

Figure 5: glmertree for usage and sociolinguistic predictor variables (priming condition)

4.2 Metalinguistic knowledge

In the following, we model the metalinguistic awareness based on the ratings that assigned the patterns in examples (6) and (7) to one of the four varieties, excluding 859 'unclear' ratings (out of a total of 4,785). Seeing that the temporal adverbial might have an impact on participants' metalinguistic awareness, we model the two contexts – with *yet* and *ever* – separately.[10]

Figure 6 shows the results of the glmertree that was fit for the choice between *did you eat yet* vs. *have you eaten yet*. Metalinguistic awareness of the participants splits preferences across the Atlantic (node 1). Interestingly, participants from England, Scotland and Canada differ from those in the US in how they perceive differences across varieties for this construction: For EngE and ScotE, US participants attribute a somewhat higher proportion of SPs than participants from the other countries (node 2) whereas for CanE and USEng, they attribute a lower proportion of SPs (see nodes 10 and 13). Overall, participants in the study attribute a clearer preference for the PP to EngE and ScotE than they do to CanE and USEng (compare nodes 5-8 and 11-15). Participants from England, Scotland and Canada also attribute a slightly higher use of the SP to ScotE than to EngE. At a lower level of the tree, we see an interaction effect with GENDER for EngE which, according to female participants from Canada, England and Scotland has the lowest proportion of SPs (node 5).

[10] For a glmertree that models awareness across the two contexts, see the supplementary materials at the same link where the data and analysis files can be found in the Data availablility statement). We also attempted to include 'context' as a predictor variable in a model across both contexts; this approach proved problematic as it resulted in too many Eigenvalues.

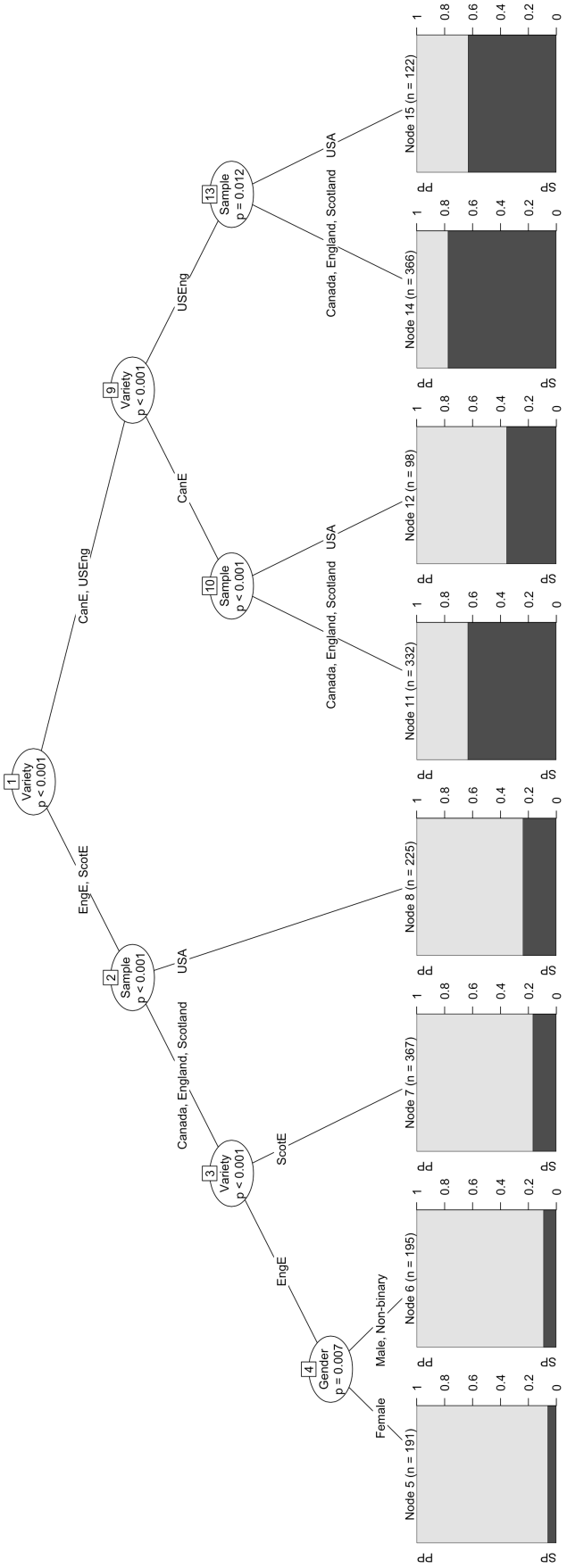

Figure 6: glmertree for metalinguistic awareness of PP:SP choice with *yet* (Somers2 at C = 0.859)

Turning to the results of the glmertree for the choice between *ever saw* vs. *ever seen* (Figure 7), we see a similar overall trend: a split across the Atlantic into British and American varieties (node 1). At the next level, though, participants from both Canada and the US perceive the preferences differently from those in England and Scotland (nodes 2 and 9). Overall, ScotE is perceived to use more SPs than EngE, albeit at somewhat higher levels by participants in the US and Canada than in England and Scotland (compare nodes 4 and 5 with nodes 7 and 8). Turning to the perception of differences between CanE and USEng, we see that participants from England and Scotland, overall, attribute a higher rate of SP to the North American varieties than informants from the US and Canada (compare nodes 13-16 with nodes 19-23). This tendency is more marked for English than for Scottish informants (compare nodes 13 and 15 to node 16). Participants from Canada perceive CanE to use fewer SPs with *ever* than USEng (compare nodes 19 and 22) whereas informants from the US attribute similar proportions of PP and SP to both North American varieties (see nodes 20 and 23). Viewed differently, Canadians can be said to be construing their usage preference as different from US usage even though participants from both countries are in fact very similar with respect to the elicited usage data. At a lower level of the tree, we again see an interaction effect with GENDER, this time for the participants from England, where female participants attribute a higher usage of SPs to CanE than male participants (see nodes 13 and 14).

To sum up, participants' metalinguistic awareness of regional preferences in the choice of PP and SP reveals a consistent transatlantic divide across the two contexts, with British varieties perceived as favouring the PP and North American varieties perceived as preferring the SP. Within these broad groupings, finer distinctions emerge depending on the participants' own regional background: for example, fewer US participants perceive SP usage to be typical of North American varieties than British and Scottish participants do, while ScotE is consistently attributed slightly more SP usage than EngE across both contexts.

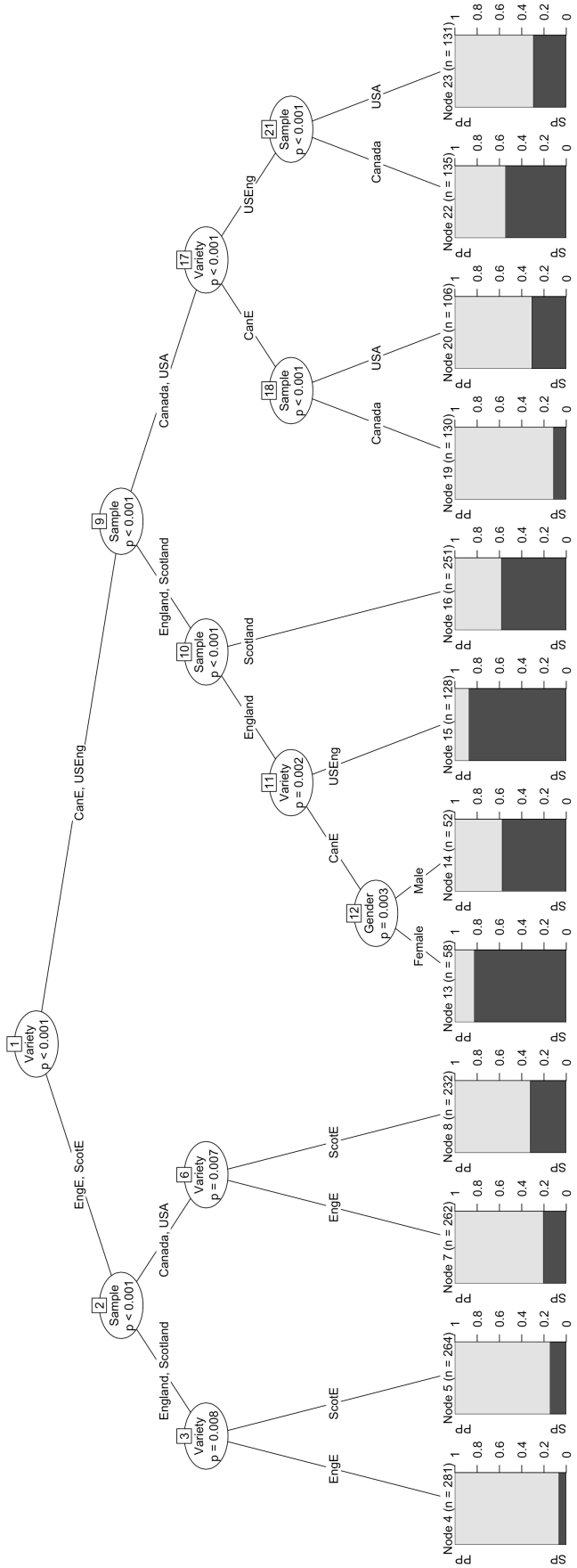

Figure 7: glmertree for metalinguistic awareness of PP:SP choice with *ever* (Somers2 at C = 0.863)

Let us now compare the regional elicited usage differences that emerge from the baseline condition with participants' perception of regional variation: Here, the most noticeable discrepancy is the trans-Atlantic difference in preferences that emerges from participants' perception, on the one hand, and the split between EngE and the three remaining varieties evidenced in the maze-task data, on the other hand. Rather than reflecting actual regional distributions of the variable choice available to speakers, the glmertree models of participants' perceptions appear to reflect the different construals of regional preferences that participants perceive. The latter might have to be attributed to the shibboleth of a – predominately US American – colloquial preterite in perfect contexts (see Vanneck 1958). In the metalinguistic awareness survey, this shibboleth may have been attributed to USEng from the outside more than from inside the US. This brings us to asymmetries in perception.

For our second RQ, we predicted that participants from Scotland and Canada would know more about usage in their respective neighbouring countries than participants from England and the US. In other words, we expected metalinguistic knowledge to be asymmetrical. The results from our survey do provide evidence of asymmetrical awareness for the choice in the resultive context with *yet*, but not exactly in the direction that we predicted. While we assumed that participants from Scottland would be more aware of English preferences and those from Canada would show a heightened awareness of choices preferred in the US than those in the neighbouring countries, we found that it is actually US Americans who tend to differ in their perception from those in Canada, England and Scotland.

4.3 Norm orientation

We focus here on norm orientation in informal spoken and formal written contexts (as the two extreme endpoints on the formality index) as well as more global norm orientation with respect to attractiveness and correctness (Figures 8-11). With respect to norm orientation in informal spoken contexts, all participant groups prefer their own variety by far and thus have a clear endo-normative orientation in informal speech. English participants rate USAmE as the next likely variety they would use in such a situation whereas for all other participant groups, EngE ranks second (see Figure 8, with perctages for exo-normative orientation added).

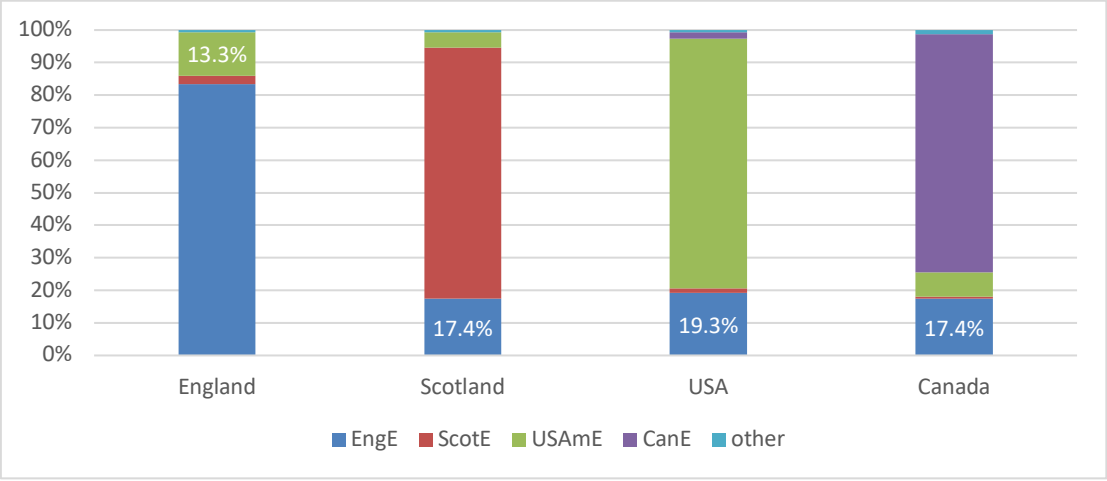

Figure 8: Norm orientation in spoken informal context

The ratings for written formal contexts are quite different in that participants in Scotland rate EngE higher than their own variety at 64.4% vs. 20.1%; the suitability of EngE for written formal use by Canadians also increases to 31.5% compared to the suitability in informal speech (17.4%), but they still rate CanE higher than an external norm at 50.7% in written formal usage; USAmE is rated at a comparatively low level with only 11% of informants naming it as the variety they would choose in this context. Participants in England and the US show a clear endo-normative orientation with only 6% and 10% rating the other country's variety as more suitable in written formal contexts (see Figure 9, with explicitly discussed percentages highlighted in the graph).

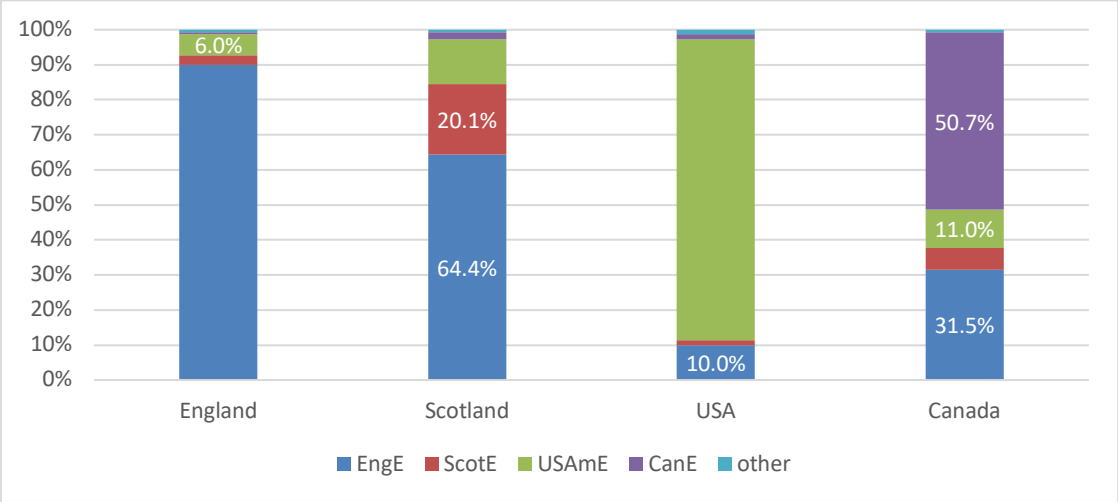

Figure 9: Norm orientation in written formal context

With respect to pleasantness (see Figure 10), the majority of participants in England and Scotland rate their own variety as the most pleasant one, despite a sizeable 36.9% of informants in Scotland rating the variety in England as the most pleasant one. Canadians are the only participant group that rate their own variety lower than the exo-normative, EngE model. The ratings for pleasantness by informants from America show that EngE is still considered the most pleasant variety, even more so in the US (56 %) than in Canada (54.4%).

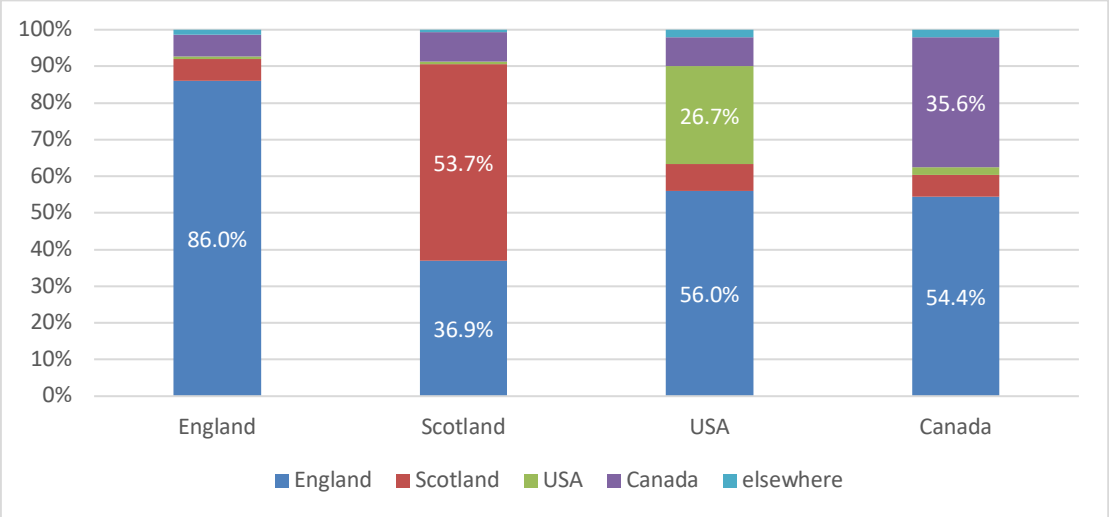

Figure 10: Most pleasant variety

Turning to the results for the ratings on the correctness dimension (*Where is the best English spoken?*) in Figure 11, we see that EngE receives the highest ratings across all participant groups, with Canadians showing stronger exo-normative orientation than people in Scotland. Interestingly, on the correctness dimension participants in the US rate their own variety higher than on the pleasantness dimension at just over 40% (compared to only 26.7%).

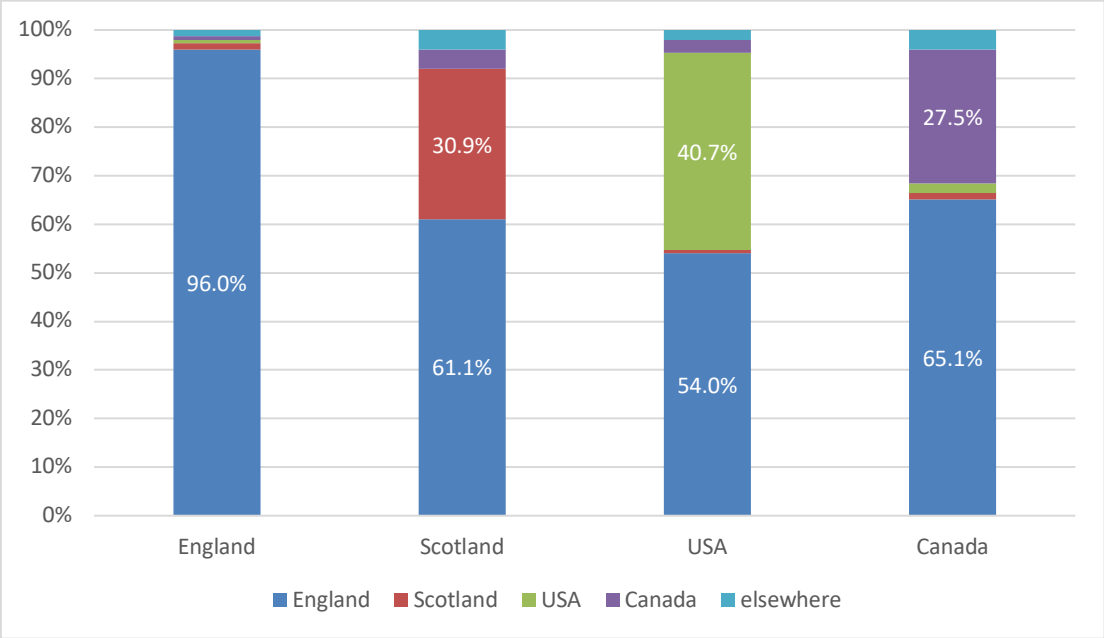

Figure 11: Best variety (aka 'correctness')

Norm orientation in the four countries thus presents us with a picture that goes beyond a simple binary distinction into exo- and endo-normative orientation. With respect to participants' own usage in Scotland and Canada, we see a shift from their own variety to a somewhat more exo-normative orientation between the two formality poles (informal speech vs. formal writing). This contrasts with participants from the US and England, who shift towards their own variety rather than away from it in more formal contexts. Somewhat surprisingly, participanats from the US rate EngE more highly than their own variety in the more global norm orientation (pleasantness and correctness, Figures 10 and 11, respectively). There is thus some lingering exo-normative orientation in North America (cf. Schneider 2007). Overall, our findings stress that a fine-grained, context- and dimension-sensitive approach towards questions of norm-orientation is called for as a background for studies on regional variation in usage.

4.4 Bringing it together: Usage and norm orientation

In a next step, we modelled usage in the baseline and priming conditions, adding norm orientation as a further predictor variable (alongside the social variables that had been included in the original models). The factor norm orientation distinguished between endo-normative, exo-normative and 'other' orientation, with exo-normative further divided into the two epicentres (see the separate levels 'exoENG' and 'exoUS' in Table 2).

We are not showing the results of the models here because they are identical to the ones shown in Figures 3 and 4 above. This means that norm orientation did not change the models that were returned for the baseline and priming

condition without norm orientation. In other words, norm orientation does not affect the regional effects for the baseline choices between the SP and the PP among the participants in this study (see Figure 3). Likewise, it does not counteract the swamping effect that the priming has on usage, where regional differences disappear (see Figure 4).

We predicted that participants' awareness of regional difference might result in their resisting to converge on a pattern that is not typical of their variety under priming conditions. This does not find any support from our data as all participants converge, regardless of their norm orientation. The same holds for our prediction on divergence between usage and norm orientation.

5 Towards integrating norm orientation, metalinguistic and constructional knowledge

The aim of this paper was to gauge the potential effect that norm orientation might have on the use of a variable grammatical pattern in WEs, which in turn would provide a strong argument in favour of including metalinguistic knowledge into what speakers know about the constructions under investigation here. In order to achieve this aim, we conducted an experiment to elicit usage data for a baseline and a priming condition. The post-experiment survey provided us with the relevant information on participants' norm orientation and metalinguistic knowledge.

With respect to the baseline condition in the experiment, we observe a split between EngE and the other varieties. These usage preferences are subject to priming effects with regional differences disappearing or being swamped under the priming condition. Contrary to our prediction that knowledge about the regional distribution of the two constructions would make participants less likely to converge on an exo-normative variant (see RQ 2), priming did not affect our participants differently.

While structural priming affects usage, norm orientation does not: Including norm orientation in the model for the baseline condition did not reveal any interaction effects with the choice between the preterite and the perfect, regardless of whether participants held strong beliefs about their own variety or not. In other words, while our experimental data appear to lend support to regional variation in the choice between PP and SP, norm orientation does not have a noticeable effect on usage. Moreover, participants construe the regional distribution of usage preferences differently from what our usage data would support: Metalinguistic awareness suggests a cross-Atlantic split whereas usage patterns revealed a split between EngE, on the one hand, and ScotE, CanE and USEng, on the other hand.

What are the implications of these findings for a usage-based approach to grammatical variation and, more specifically, for the question whether or how sociolinguistic information should be included in the representation of constructions in the construct-i-con? In order to answer this question, we need to return to the tripartite distinction between indicator, marker and stereotype and consider them as reflecting different degrees of awareness or consciousness. Since stereotype carries with it evaluative connotations we might want to make use of terminology introduced by the developmental psychologist Gopnik (2009), who observes that as we mature, we increasingly focus our attention in a spotlight-like manner from a previous state of mind where consciousness is likened to a

lantern with a more diffuse, but at the same time broader, focus.[11] Adopting the conceptual metaphors of lantern and spotlight to sociolinguistics allows us to move away from the established sociolinguistic terminology and the close association between individual features and the function they serve for specific groups of speakers as either indicator, marker or stereotype. This more fluid view of social meaning, we argue, fits in better with the move in third-wave sociolinguistics towards a more dynamic view that stresses human agency in employing the features for social meaning-making. The results of our experiment suggest that for the same group of people, the same feature can serve as both a lantern and a spotlight, depending on the perspective. The results of the maze task experiment (baseline) reveal the areal distribution and reflect the usage-based norm for the PP and SP, including more construction-internal factors. Modelling the usage-based norm thus corresponds to speakers' lantern-like focus. The results of the meta-linguistic awareness survey suggest that what speakers' beliefs about regional differences in the choice between the PP and SP can function like a spotlight.

The usage data thus provides support on semantic differences between subtypes of the present perfect modelled in previous accounts of the construction (Michaelis 1993, 1998). Regional variation only plays a role for persistent situation and resultative PPs, with ScotE, CanE and USEng participants even displaying different preferences for the two sub-types of PP. At the same time, these usage patterns were not exactly reflected in participants' perception of regional differences. Nevertheless, Canadians could be said to be using the SP:PP choice as a spotlight. In a sense, they are doing distancing work in more often assigning the pattern that they and US American actually use with a similar frequency as something that is typical of speakers on the other side of the border. Usage-based norms potentially also include interaction with register-specific variation that is difficult (if not impossible) to elicit in experimentation but typically emerges from corpus data (see Morin, Coats & Dunn 2026). The results from the maze task under priming conditions, finally, could be argued to be turning the lantern- and spotlight-like foci off and putting users on auto-focus as they align with the previously presented usage.

The apparent mismatch between usage and perception/beliefs highlights the complexity of the processes and the dynamic nature of social meaning (and meaning-making) in interaction. It is thus not possible to suggest a simple solution that would result in adding 'regional preference' to the constructional space for the PP and the SP that Michaelis (1998) proposed. Importantly for theories of WEs, participants' norm orientation did not impact their usage at all, whereas structural priming did. While our results reveal that usage, norm orientation and metalinguistic knowledge do not necessarily align with each other, this mismatch between usage and perception/awareness – what speakers think they know about variation – is in line with previous research in cognitive sociolinguistics on sociophonetics (e.g. Niedzielski 1999). Thus, on the basis of the experimental evidence and the survey results, simply including information on the regional distribution of the PP and SP across different WEs into the constructional templates runs the risk of conceptualising the social meaning dimension as too static and neglecting the more dynamic nature of the meaning-reflecting and meaning-making processes at work. Importantly, the findings of our study

[11] These differences in focussing attention represent tendencies: as adults we are also able to broaden our conscious attention if need be.

caution against a view that would move from the modelling of corpus data (usage-based norm) to an interpretation of the probabilistic patterns that can be observed in terms of the social significance that these distributions have for speakers per se, as suggested in the feature-focussed contribution to this volume (Morin, Coats & Dunn 2026). In other words, observing region- and/or register-specific distributions does not automatically allow us to infer what these distributions 'mean' to speakers of the varieties nor how they might employ them in social meaning-making. The methodological approach taken in this study thus complements that taken in Morin, Coats and Dunn (2026) in that it allows us to tap into the different levels of relevance that grammatical patterns have as diagnostic of a regional variety (what Morin, Coats and Dunn are modelling) as opposed to the more conscious – and potentially even contradicting ways – in which speakers may choose to bring them to bear on their beliefs about variation and thus, in a next step, constructing a regional identity for themselves. Interestingly, our study also shows that participants are able to access a relatively abstract construction in the domain of tense and aspect marking for the sociolinguistic purposes of identity construction. The Canadians in our study are doing so on an exemplar-based level in the post-experiment survey on metalinguistic awareness.

The present study has only started to explore the relationship between usage, norm orientation and priming in the choice between PP and SP. In future studies, baseline contexts could include different types of adverbial, such as *ever* (currently only used in the PP prime) to tap into the effect that these have on usage across the different varieties. They also merit fine-tuning with respect to their typicality for the respective subtypes of PP (see our comments on baseline sentence (3) d. in section 4.1). For the priming contexts, rather than opting for a binary (PP vs. SP prime), a more fine-grained approach could take into account different types of potential PP contexts. Overall, however, we are fairly confident that priming is likely to neutralise any regional variability that might be observed in a more fine-grained set of baseline conditions and that norm orientation will remain distinct from usage.

Acknowledgements

We are grateful for the two anonymous reviewers and the editors of this special issue for the constructive and useful feedback on earlier versions of this paper. Annika Heuser, Linda Steiner and Sofia Carpentieri all helped with the creation of distractors for the experiment, and Giancarlo Bonilla and Helene Hundt with the data-preprocessing.

Conflict of interest statement

The authors declare none.

Data availability statement

The data and analysis scripts for the present paper are available at https://osf.io/3cy7w/.

Ethics statement

All participants provided their written informed consent to participate in this study.

References

Bao, Chenyao, Xiaowen Zhang, Yunhua Qu & Zhiwei Feng. 2018. American English perfect construction across registers. *Journal of Quantitative Linguistics* 25(4). 314–341.

Bencini, Giulia M. L. & Adele E. Goldberg. 2000. The contribution of Argument Structure Constructions to sentence meaning. *Journal of Memory and Language* 43(4). 640–651.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Bock, Kathryn. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18. 355–387.

Bowie, Jill, Sean Wallis & Bas Aarts. 2013. The perfect in spoken British English. In Bas Aarts, Joan Close, Geoffrey Leech & Sean Wallis (eds.), *The Verb Phrase in English*, 318–350. Cambridge: Cambridge University Press.

Boyce, Veronica & Roger P. Levy. 2023. A-maze of natural stories: Comprehension and surprisal in the Maze task. *Glossa Psycholinguistics* 2(1): X. 1–34.

Britain, David. 2016. Sedentarism, nomadism and the sociolinguistics of dialect. In Nikolas Coupland (ed.), *Sociolinguistics: Theoretical Debates*, 217–241. Cambridge: Cambridge University Press.

Clark, Lynn. 2018. Priming as a motivating factor in sociophonetic variation and change. *Topics in Cognitive Science* 10. 729–744.

Crystal, David. [7]2002. *Rediscover Grammar*. London: Longman.

Dallas, Bethany & Marianne Hundt. Submitted. Priming perfects and preterites: A bottom-up, corpus-based approach at modelling aspectual variation in English.

Davydova, Julia. 2011. *The Present Perfect in Non-Native Englishes*. Berlin: De Gruyter.

Deshors, Sandra C. & Gaëtanelle Gilquin. 2018. Modeling World Englishes in the 21st century: New reflections on model-making. In Sandra C. Deshors (ed.), *Modeling World Englishes: Assessing the Interplay of Emancipation and Globalization of ESL Varieties*, 281–294. Amsterdam: Benjamins.

Diessel, Holger. 2023. *The Constructicon: Taxonomies and Networks*. Cambridge: Cambridge University Press.

Drager, Katie K. 2015. *Linguistic Variation, Identity Construction and Cognition*. Berlin: Language Science Press.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 435–376.

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41. 87–100.

Eckert, Penelope. 2018. *Meaning and Linguistic Variation: The Third Wave in Sociolinguistics*. Cambridge: Cambridge University Press.

Eckert, Penelope & John R. Rickford (eds.), 2001. *Style and Sociolinguistics Variation*. Cambridge: Cambridge University Press.

Elsness, Johan. 1997. *The Perfect and the Preterite in Comtemporary and Earlier English*. Berlin: de Gruyter.

Elsness, Johan. 2009. The present perfect and the preterite. In Günter Rohdenburg & Julia Schlüter (eds.), *One Language, Two Grammars? Differences between British and American English*, 228–245. Cambridge: Cambridge University Press.

Elsness, Johan. 2014. The present perfect and the preterite in Late Modern and Contemporary English. A longitudinal look. In Kristin Davids, Caroline Gentens, Lobke Ghesquière & Lieven Vandelanotte (eds.), *Corpus Interrogation and Grammatical Patterns*, 81–104. Amsterdam: Benjamins.

Fokkema, Marjolein & Achim Zeileis. 2019. Package "glmertree". https://cran.r-

project.org/web/packages/glmertree/glmertree.pdf (accessed 19 June 2025).
Fokkema, Marjolein, Niels Smits, Achim Zeileis, Torsten Hothorn & Henk Kelderman. 2018. Detecting treatment-subgroup interactions in clustered data with generalized linear mixed effects model trees. *Behavior Research Methods* 50(5). 2016–2034.
Fokkema, Marjolein, Julian Edbrooke-Childs & Miranda Wolpert. 2020. Generalized linear mixedmodel (GLMM) trees: A flexible decision-tree method for multilevel and longitudinal data. *Psychotherapy Research* 31(3). 329–341.
Forster, Kenneth I., Christine Guerrera & Lisa Elliot. 2009. The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods* 41(1). 163–171.
Franco, Karlien & Sali A. Tagliamonte. 2022. The most stable it's ever been: the preterit/present perfect alternation in spoken Ontario English. *English Language & Linguistics* 26(4). 779–806.
Fuchs, Robert. 2016. The frequency of the present perfect in varieties of English around the world. In Valentin Werner, Valentin, Elena Seoane & Cristina Suarez-Gómez (eds.), *Reassessing the Present Perfect,* 223–258. Berlin: De Gruyter.
Gathercole, Virginia. C. 1986. The acquisition of the present perfect: Explaining differences in the speech of Scottish and American children. *Journal of Child Language* 13(3). 537–560.
Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.
Goldberg, Adele E. 2019. *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions.* Princeton, NJ: Princeton University Press.
Gopnik, Alison. 2009. *The Philosophical Baby: What Children's Minds Tell Us about Truth, Love, and the Meaning of Life*. New York: Farrar Strauss and Giroux.
Hänsel, Eva & Philipp Meer. 2023. Comparing attitudes toward Caribbean, British, and American accents in Trinidad and Tobago, the United Kingdom, and the United States. *World Englishes* 42(1). 130–142.
Heuser, Annika. 2022. Transformer-Maze. MA thesis, MIT. https://dspace.mit.edu/ handle/1721.1/147233?show = full (last accessed 29 September 2025).
Hundt, Marianne & Nicholas Smith. 2009. The present perfect in British and American English: Has there been any change, recently? *ICAME Journal* 33. 45–63.
Johnson, Matt A. & Adele E. Goldberg. 2013. Evidence for automatic accessing of constructional meaning: Jabberwocky sentences prime associated verbs. *Language and Cognitive Processes* 28(10). 1439–1452.
Klein, Wolfgang. 1994. *Time in Language*. London: Routledge.
Kootstra, Gerrit Jan, & Pieter Muysken. 2019. Structural priming, levels of awareness, and agency in contact-induced language change. *Languages* 4(3). 65.
Labov, William. 1972. *Sociolinguistic Patterns*. University of Pennsylvania Press.
Labov, William. 2001. *Principles of Linguistic Change, Volume 2: Social Factors*. Oxford: Blackwell.
Michaelis, Laura A. 1993. *Toward a Grammar of Aspect: The Case of the English Perfect Construction.* Berkeley: University of California Doctoral Dissertation.
Michaelis, Laura A. 1998. *Aspectual Grammar and Past Time Reference.* London and New York: Routledge.
Miller, Jim. 2000. The perfect in spoken and written English. *Transactions of the Philological Society* 98(2). 323–352.
Miller, Jim. 2004. Problems for typology: Perfects and resultatives in spoken and non-standard English and Russian. In Bernd Kortmann (ed.), *Dialectology meets Typology: Dialect Grammar from a Cross-Linguistic Perspective,* 305–334. Berlin: de Gruyter.
Morin, Cameron, Steven Coats & Jonathan Dunn. 2026. How register and region shape the language network: Evidence from Computational Construction Grammar. *Constructions* 18 (aop).
Niedzielski, Nancy A. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18. 62–85.
Pickering, Martin J., & Victoria S. Ferreira. 2008. Structural priming: a critical review.

Psychological Bulletin 134(3). 427–59.
Preston, Dennis. 1996. Whaddayaknow?: The modes of folk linguistic awareness. *Language Awareness* 5(1). 40–75.
Preston, Dennis. 2016. Whaddayaknow now? In Anna M. Babel (ed.), *Awareness and Control in Sociolinguistic Research*, 177–199. Cambridge: Cambridge University Press.
Sandow, Rhys J. 2022. Attention, identity and linguistic capital: inverted style-shifting in Anglo-Cornish dialect lexis. *English Language and Linguistics* 26(4). 677–695.
Schmid, Hans-Jörg. 2020. *Dynamics of a Linguistic System: Usage, Conventionalization, and Entrenchment*. Oxford: Oxford University Press.
Schneider, Edgar W. 2007. *Postcolonial English. Varieties Around the World*. Cambridge: Cambridge University Press.
Schneider, Edgar W. 2025. *World Englishes as Components of a Complex Dynamic System*. Cambridge: Cambridge University Press.
Seoane, Elena & Cristina Suarez Gomez. 2013. The expression of the perfect in East and South-East Asian Englishes. *English World-Wide* 34(1). 1–25.
Sharma, Devyani. 2025. Testing sociolinguistic theory and methods in world Englishes. *World Englishes* 44(1). 26–42.
Silvennoinen, Olli O. 2025a. Going forward as an adverb. *English Language and Linguistics* 29(2). 336–361.
Squires, Lauren. 2016. Processing grammatical differences: Perceiving versus noticing. In Anna M. Babel (ed.), *Awareness and Control in Sociolinguistic Research*, 80–103. Cambridge: Cambridge University Press.
Szmrecsanyi, Benedikt. 2005. Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory* 1(1). 113–150.
Trousdale, Graeme & Elizabeth C. Traugott. 2026. *Constructions, Networks and Linguistic Change*. Cambridge: Cambridge University Press.
Vanneck, Gerard. 1958. The colloquial preterite in modern American English. *Word* 14. 237–242.
Werner, Valentin. 2014. *The Present Perfect in World Englishes: Charting Unity and Diversity*. Bamberg: University of Bamberg Press.
Yao, Xinyue. 2024. *The Present Perfect and the Preterite in Late Modern and Contemporary English*. Amsterdam: Benjamins.
Yao, Xinyue & Peter Collins. 2013. Functional variation in the English present perfect: a cross-varietal study. In Gisle Anderson and Kristin Bech (eds.), *English Corpus Linguistics: Variation in Time, Space and Genre*, 91–111. Amsterdam: Rodopi.

Appendix

Table 1A: Participant distribution by gender

	Female	Male	Non-binary	Total
Canada	74	74	1	149
England	74	76	0	150
Scotland	72	77	0	150
USA	75	75	0	150
Total	295	302	1	598

The sample includes one non-binary participant whose gender identity emerged from our demographic questionnaire rather than via the Prolific filter.

Table 2A: summary of the GLMM in Figure 4

Model formula:
CHOICE ~ 1 | PPsubtype + Gender + Age + Mobility + Sample + Speak_INF + Speak_FOR + Write_FOR + Write_INF + Best_E + Pleasant_E

Fitted party:
```
[1] root
|   [2] PPsubtype in past: n = 407
|       (Intercept)
|         0.7083225
|   [3] PPsubtype in persistentS, recentP, resultative
|   |   [4] PPsubtype in recentP: n = 513
|   |       (Intercept)
|   |        -1.294354
|   |   [5] PPsubtype in persistentS, resultative
|   |   |   [6] Sample in England: n = 233
|   |   |       (Intercept)
|   |   |        -0.8219117
|   |   |   [7] Sample in Canada, Scotland, USA
|   |   |   |   [8] PPsubtype in persistentS: n = 384
|   |   |   |       (Intercept)
|   |   |   |        -0.4439419
|   |   |   |   [9] PPsubtype in resultative: n = 385
|   |   |   |       (Intercept)
|   |   |   |        -0.04527551
```

Number of inner nodes: 4
Number of terminal nodes: 5
Number of parameters per node: 1
Objective function (negative log-likelihood): 1082.016