A social turn for Construction Grammar: double modals on British Twitter

CAMERON MORIN
Ecole Normale Supérieure de Lyon

GUILLAUME DESAGULIER
Université de Paris Vincennes Saint-Denis

and

JACK GRIEVE
University of Birmingham
(Received 1 December 2022; revised 7 September 2023)

Construction Grammar is an emerging theory of language, but the analysis of sociolinguistic variation is still relatively underdeveloped in the framework. In this article, we consider the representation of social meaning in Construction Grammar through a corpus-based analysis of double modals in British English on social media. We describe the use of double modals in a large corpus of geolocated Twitter posts, including presenting an inventory of observed double modals and maps showing the regional distribution of each of these forms. We find that double modals show a general northern pattern and are concentrated in the Scottish Borders. We also find various rare double modals that occur more widely across the UK. To account for these results, we propose a Construction Grammar account of double modals. We argue that defining double modals as grammatical constructions requires that aspects of their social meaning be delimited, especially register and region. Furthermore, we argue that double modals may be enregistered as dialect constructions, distinguished from standard constructions of British English. We conclude by considering the importance of incorporating social meaning into Construction Grammar, underlining the value of a Cognitive Sociolinguistic approach to grammatical theory.

Keywords: Construction Grammar, sociolinguistics, double modals, British English, Cognitive Linguistics

1   Introduction

There has recently been a 'social turn' in Cognitive Linguistics (CL). Prominent members of the field have emphasised the importance of accepting what is known as the 'Sociosemiotic Commitment' (Geeraerts 2016)[1] to address the lack of engagement in CL with social aspects of language knowledge and use (see also Dąbrowska 2016). A

---

[1] Geeraerts (2016: 537) defines the Sociosemiotic Commitment as follows: 'To complement the Cognitive Commitment, we define a *commitment to make one's account of human language accord with the status of*

number of related approaches in CL, such as Cognitive Sociolinguistics (Claes 2018; Kristiansen *et al.* 2021), the Entrenchment-and-Conventionalization model (Schmid 2020) and Word Grammar (Hudson 2021), are grounded in the understanding that a theory of language and grammar must integrate accounts of both individual cognitive entrenchment and community variation and convention, keeping with the assumptions of inherent variability in language (Labov 1969), embodied cognition (Croft & Cruse 2004) and the sociocognitive roots of language in the human species – the 'ultra-social animal' (Croft 2009; Tomasello 2014).

A subset of research questions in this area focuses on modelling socially conditioned grammatical variation. For example, Word Grammar analyses dialect syntax as a point of intersection between linguistic and social conceptual structure (Hudson 2007). Alternatively, the Entrenchment-and-Conventionalization model views sociolinguistic variation as the product of variable conventions whose source is the rich mutual understanding of multidimensional usage events in 'co-semiosis',[2] including elements of situational and sociocultural setting associated with utterance types (Schmid 2022). In addition, over the past decade, the Cognitive Sociolinguistics literature has offered a wide range of case studies analysing socially salient constructions in a number of languages, most especially in Dutch (Kristiansen *et al.* 2021).

In the prominent CL framework of Construction Grammar (CxG), however, interest in sociolinguistic variation has been more limited (Morin *et al.* 2020). In earlier work, a few researchers expressed awareness of something akin to the Sociosemiotic Commitment, including Goldberg (2013: 16), who wrote that in CxG 'the role of social cognition [is] viewed as essential to accounts of learning and meaning'. Similarly, Hollmann (2013) argued that a cognitive sociolinguistic approach could bring many benefits to the study of grammatical constructions in CxG, while Östman & Trousdale (2013) introduced the modelling of regional variation in CxG, focusing on its potential to explain dialect-specific cognition and social meaning as part of the knowledge of different 'subnetworks' of speakers whose usage-based conventions differ. More recently, Hoffmann (2022) cites two main notions of social meaning relevant for CxG: static meaning and dynamic indexicality. Both are inspired by the three waves of variationist sociolinguistics (Eckert 2012). Static meaning constraints on the knowledge and use of grammatical constructions – for instance, in terms of local and supralocal social categories including age, sex and region (Östman & Trousdale 2013) – are associated with the first and second waves of variationist theory. Dynamic indexicality, i.e. the inference and association of social contextual features with constructions based on acts of identity in interaction (Hall-Lew *et al.* 2021), is associated with the third wave (see also Hoffmann 2015).

---

language as a social semiotic, i.e., as an intersubjective, historically and socially variable tool, and to base that account on a methodology that likewise transcends the individual' (original emphasis).

[2] Schmid (2020: 30) defines co-semiosis as 'the activity of negotiating and establishing mutual beliefs of the mutual understanding of an utterance in a given context'.

Although there has been some interest in incorporating ideas from the study of sociolinguistic variation into CxG, linguists in this tradition have only begun to engage with these ideas. One likely factor limiting development is the methodological difficulties associated with the study of dialect constructions. Two of the main empirical sources of data in CxG are corpora and experiments (Gries 2013) and both present challenges for the study of grammatical variation. Dialect-specific and informal language use has been notoriously difficult to capture through observational corpus-based methods, especially involving rare and localised syntactic forms, while experiments are often inappropriate for the study of variation, because the social variables of interest cannot be manipulated directly (Grieve 2021).

However, recent developments in observational linguistic methodology, especially corpus-based computational sociolinguistics (Dunn 2018; Schmid *et al.* 2021; Grieve *et al.* 2023), have begun to lessen these empirical obstacles, allowing us to study a range of dialect constructions, including the rarest and most localised ones, through a corpus-based approach that samples informal written data at scale from online sources, especially social media. By drawing on a computational sociolinguistic methodology, in this article, we aim to encourage the development of CxG into a cognitive sociolinguistic theory. Specifically, we analyse a particularly rare, unusual and socially salient grammatical variant in dialects of British English: the Double Modal (DM) construction.

DMs have been attested in limited regions of the British Isles, most commonly in the Scottish Borders (Brown 1991; Bour 2014; Morin 2021a, 2021b), and to a lesser extent in the Central Belt of Scotland including Glasgow and in areas of Dumfries and Galloway in the southwest (Smith *et al.* 2019). They are attested even less frequently in the northeastern English counties of Tyneside and Northumberland (Beal 2004; Beal & Corrigan 2000), as well as in Ulster Scots-speaking counties of northeastern Ireland (Montgomery & Nagle 1993). Because they are so infrequent, restricted to specific dialect regions, and potentially archaic in these same regions, studies of these features have relied on limited amounts of data collected through elicitation by a small number of sociolinguistic fieldwork studies. DMs have long been considered impossible to analyse using observational empirical approaches (Butters 1973), although a recent study by Coats (2023) reports the first corpus-based description of DMs in British and Irish speech, based on automated speech transcripts from YouTube. Despite their scarcity, there is evidence that DMs are strongly associated with a social identity linked to the regions they are used in, especially the Scottish Borders (Brown 1991), and that they are generally perceived as 'non-standard' grammatical features, thus conveying salient social meaning when they are used (Bour 2014).

At first sight, DMs appear to have a very simple grammatical structure that should be amenable to CxG analysis, combining two distinct word forms of the same class. Nevertheless, DMs are very rare and unusual grammatical features in the English language, with a limited regional distribution, generally being perceived as unacceptable by most speakers. Consequently, DMs are likely to convey fine-grained and distinctive social meaning (Irvine 2001), making them especially well suited for the analysis of social meaning in CxG for two reasons. First, the social indexicality of

these constructions demands that we are clear about the integration of social meaning and sociolinguistic variation into CxG. Second, the empirical analysis of DMs can extend the descriptive scope of CxG, enabling us to study a range of socially meaningful constructions in future research. The analysis of DMs from a CxG perspective thus has the potential to increase both the explanatory adequacy and predictive power of CxG as a theory of language.

In this article, we therefore analyse the social meaning of DMs in a large corpus of social media, showing that British DMs as used on social media are dialect constructions. We present an observational account of British DMs on the social media platform Twitter, comparing the results with past research on these features, and we then present a discussion of social meaning in CxG in light of these results. Specifically, we pursue the following research objectives:

1. Describing the use of DMs in contemporary British English based on an unprecedentedly large corpus of social media, including their inventory, frequency, regional distribution and modal meaning;
2. Proposing a CxG model of DMs in British English on Twitter;
3. Considering, based on this model and the observed quantitative and qualitative regional and social patterns of DMs in our corpus, why and how social meaning should be incorporated into CxG.

Our article is structured as follows: in section 2, we review past research on the inventories and regional distribution of British DMs. In sections 3 and 4, we describe our methods and results in terms of the inventory, frequency and regional distribution of DMs in a large corpus of geolocated social media posts collected from Twitter across the UK in 2014. In section 5, we analyse DMs as a network of constructions through a close examination of their combination patterns and a distributional study of their modal meaning. In section 6, we argue that examining their social meaning is required to fully account for them as socially specific constructions based on non-standardness, informality and regional variation, which may entrench and conventionalise them in dialect-specific grammar networks. We conclude that future research should directly engage with social meaning in CxG and use new empirical approaches to studying language variation so as to transform CxG into a Cognitive Sociolinguistic theory of language, broadening the range of phenomena it can explain.

2   Double modal inventories and regional distribution

In previous research, the most commonly attested DMs in Scots include *will can*, *might can* and *used to could* (1) (Bour 2014; Smith *et al.* 2019; Morin 2023). Other attested forms include *must can*, *would might*, *should can*, *may should* and *must could*.

(1)   (a)   He**'ll can** help us the morn. (Scotland, Miller 2004: 34)
      (b)   She **might can** get away early (Scotland, Miller 2004: 34)
      (c)   He **used to could** do it when he was younger (Scottish Borders, Brown 1991: 109)

In Northumberland and Tyneside, *might could* and *might can* are less frequently attested, whereas *would could* seems specific to these counties and rarer in Scots (Beal 2004). In Ulster Scots, typical forms include *will can*, *might could*, *used to could*, *used to would* and *should could* (Robinson 2018).

(2)   (a)   He **wouldn't could**'ve worked, even if you had asked him (Tyneside, Beal 2004: 128)
       (b)   We don't have that, but you **might could** find that across the street (Antrim, Montgomery & Nagle 1993: 103)

Comparatively, DMs are more widespread and diverse in dialects of American English, especially in the southeast (Huang 2011), including the combinations *might can*, *might could* and *used to could*, but not typically *will can* and *would could*. They are also attested in several English-based creoles such as Gullah, Jamaican and Guyanese (Kortmann *et al.* 2020). Otherwise, they are unattested in most varieties of English, including Standard English (Huddleston & Pullum *et al.* 2002), and unacceptable to a majority of British speakers, as suggested in the quote below.

> what they do if you go up north there are certain places in Scotland that I've read about that use double modal auxiliary verbs […] so for example they'll say a sentence I *would could* do that which to us sounds ridiculous […] I *should might* be able to which just sounds really really wrong to us (Spoken BNC 2014, dialect: Received Pronunciation)

There is much uncertainty concerning the status of DMs as dialect features: are they still used in rural northern dialects? Are they disappearing? Have they disappeared? Two decades ago, Beal (2004) proposed that DMs are recessive in northern English dialects, based on their near absence in the Newcastle Electronic Corpus of Tyneside English, while Trousdale (2000) adopted an agnostic position – acknowledging their attestation, but calling for caution due to a lack of naturalistic data. Corrigan (2011) suggested that DMs had acquired 'relic status' in the British Isles overall, and that they are passive in Ulster Scots. As for (Borders) Scots, where DMs are reportedly most common, claims about these features vary. While Smith *et al.* (2019) suggest that typical DMs such as *might can* may be disappearing based on lower acceptability rates among younger generations outside the Borders, Bour's fieldwork studies (Bour 2014) suggest that they are still used among younger speakers, although this was a decade ago.

Most accounts of the combinatorial constraints on British DMs are based on limited and dated data. Brown (1991: 106) proposes a grid of possible modal combinations (table 1). Notably, first modals in DMs have been found to most commonly be *might*, *should*, *will*, *would*, *must*, *used to* and *have to*, while second modals have been found to most commonly be *can*, *could*, *must*, *be to* and *have to*[3] for a total of 26 possible combinations. The modals *shall* and *may* and the semi-modal *ought to* are not included

---

[3] A modal followed by the semi-modal *have to*, however, is an acceptable construction in Standard English, and arguably does not constitute a 'double modal': as opposed to modal verbs, which only have finite forms and therefore cannot iterate in a clause (Huddleston & Pullum *et al.* 2002), a non-finite semi-modal is an acceptable VP complement of a modal.

Table 1. *The range of modal combinations in Borders Scots (Hawick) according to Brown ([1991](#)): X = attested, ? = uncertain*

	Second modal								
First modal	*can*	*could*	*might*	*should*	*will*	*would*	*must*	*be to*	*have to*
can									
could							X	X	
might	X	X						X	X
should	X	X						X	X
will	X	?						X	X
would	?	X						X	X
must	X	X						X	X
used to		X						X	
be to									
have to	X	X							

in this inventory because they are considered rare in Scots and Scottish English. Brown ([1991](#)) captures the general consensus that *might*, *will* and *used to* are always first and *can* and *could* are always second in the typical combinations *might can*, *might could*, *will can*, *would could* and *used to could* across the relevant British dialects featuring DMs (Beal [2004](#); Corrigan [2011](#); Bour [2014](#)).

More recently, extensive fieldwork coverage of DMs in Scots is presented by the *Scots Syntax Atlas* (SCOSYA, Smith *et al*. [2019](#)), which collected data on the geographical distribution and mean acceptability judgment rates of seven typical DMs: *used to could*, *will can*, *might can*, *might would*, *must can* and *must could*. Maps for each DM are shown in [figure 1](#). Based on these maps, DMs seem most acceptable in the Scottish Borders, as expected from previous research, but also in the Central Belt and the Urban Scots-speaking areas of Glasgow and Edinburgh, as well as Dumfries and Galloway.

Finally, Coats ([2023](#)) recently used a 112-million-word corpus of YouTube transcripts tagged at the county level in the UK and Ireland, finding 187 DM tokens spanning 44 types, which is a larger inventory than considered in past research. Coats proposes that they are both distributed in the northern dialect region identified in past research, but also to a large extent throughout the UK and Ireland, with no clear regional pattern.

3   Data

The data collected for this study come from a 2-billion-word corpus of posts from British Twitter, containing around 180 million geolocated tweets from 1.9 million users produced between January and December 2014 (see Grieve *et al*. [2019](#)). The metadata include the longitude and latitude of the user's GPS-enabled smartphone. The corpus was also stratified into 124 subcorpora based on UK postcodes in order to map and analyse the regional distribution of each form. Rather than taking geolocation as an indicator of
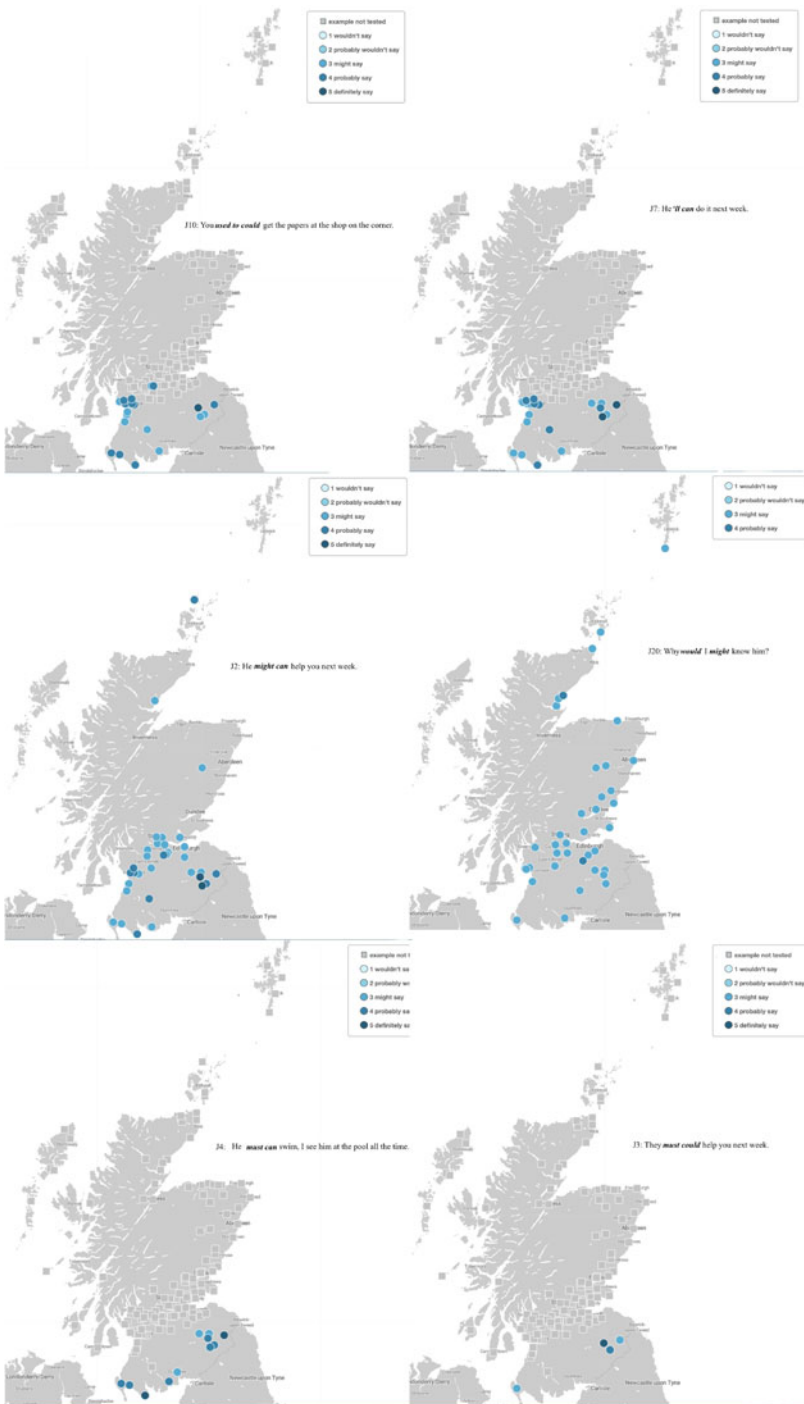
Figure 1. Mean acceptability ≥3 of 7 DMs in Scotland from the *Scots Syntax Atlas* (Smith *et al.* 2019)

geographic origin, we take it as an indicator of the geographic location of language use, which is crucial as a measure of regional variation, both for a dialectological study and for a usage-based, CxG approach to DMs. This corpus has previously been used in studies on lexical (Grieve *et al.* 2019) and phonological (Nini *et al.* 2020) variation in dialects of British English; these studies have shown that phonological maps based on this corpus align well with traditional dialect surveys.

DM token candidates were collected by automatically extracting all combinations involving the nine primary modals of English (*can, could, may, might, shall, should, will, would, must*) and three 'semi-modals' (*used to, ought to, have to* in first position). Duplicates such as *might might* were discarded because they were mostly used emphatically. To focus on the central phenomena of interest, other semi-modals (*have to* in second position, *going to/gonna*) and adverb or pronoun-split combinations (*might as well could, will you can*) were also ignored during data collection. This method resulted in 7,766 collected tokens of potential DMs in the corpus. However, a substantial number were removed by thoroughly hand-cleaning the data to only keep 'true' DMs, or at least unequivocally valid combinations of modals in a single tensed clause. Obvious types of problematic combinations or 'false DMs' included words that are isomorphic to modals but which are not modals or even verbs, clear cases of autocorrects and typing errors given the context of the tweet, and adjacent modals across clause boundaries that aren't marked by punctuation in spontaneous online communication. The final cleaned dataset that is used in this study consists of 962 examples of DMs, which were produced by a total of 935 unique Twitter accounts, indicating that this dataset is by no means dominated by the style of a small number of Twitter users.

4 Results

4.1 Forms and frequencies of double modals

Our dataset includes 60 distinct DM types, which represents exactly half of all possible combinations (120). The tweets in (4)–(7) feature four of the most common DMs found in our corpus. Table 2 displays the raw frequencies of all tokens per type listed in decreasing order.

(4) doable though; italy weren't great; i think we **will can** get 6 points and get out the group
(5) nice! i**'ll shall** look forward to all of the above 😊
(6) we **would could** be back if the gateshead public supported us !!!
(7) ok will do! just so frustrated with this woman!sorry to of contacted you just thought you **would might** of been able to help

The forms *will can* and *would could* are two typical British DMs that appear in the literature, respectively first and third in the overall frequencies. More strikingly, *will shall*, a previously unattested combination in Scots, is the second most frequent DM in the distribution. Overall, the top ten DMs in the corpus have *will* or *would* as the first

Table 2. *All DMs in the Twitter dataset*

DM	Raw freq.	DM		DM		DM	
will can	156	will must	16	should will	5	may used to	2
will shall	129	would may	12	can should	4	might may	2
would could	84	would will	11	could might	4	might must	2
would might	65	could may	10	should can	4	shall can	2
will might	59	might could	10	should might	4	shall will	2
will should	41	should would	10	would must	4	used to could	2
would should	41	may will	9	could can	3	might would	2
will may	31	might will	9	may would	3	can would	1
will could	31	could will	7	might should	3	have to should	1
will would	31	may can	7	must will	3	have to will	1
would can	23	must can	7	must would	3	will used to	1
would can	23	could used to	6	should must	3	may might	1
can will	21	should could	6	can may	2	may must	1
might can	17	could should	5	can might	2	may should	1
can could	16	should may	5	may shall	2	might shall	1

modal. The first DM to have another modal is *can will*, which is only the twelfth most frequent combination.

It is especially notable that only a few of these common DMs, namely *will can*, *would could* and *would might*, correspond to the typical, dialect-specific forms identified in section 2. In contrast, other typical DMs are much lower in frequency. The DM *might can*, for instance, has a relatively higher frequency than most DM types, but is still remarkably uncommon compared to the top twelve DMs. This infrequency is clearer in the case of *might could* and *must can*. The typical DMs *used to could*, *may should* and *should can* are very infrequent with a couple of tokens at most. Finally, *must could* is totally absent from the dataset. Another important finding is the unexpected diversity of DM types, which would likely be even bigger if a larger corpus were analysed. We have 60 unique modal combinations, over twice as many as proposed by Brown (1991) and 26 more than found in Coats (2023). A large number of these types – over half of the distribution – have extremely low frequencies, as seen in table 2, including various forms with a frequency of 1.

The status of these forms, of course, needs to be questioned before being invariably accepted. In particular, it is reasonable to question their basic validity: for example, these combinations may have been caused by typing errors or the autofill function of smartphone keyboards predicting the wrong word. However, there are several reasons why the number of such instances appears to be small. Most importantly, as discussed in section 3, we manually checked each token with care to make sure their usage appeared purposefully and meaningfully in context in order to analyse their modal meaning, as discussed below. During this process, we removed clear errors and made sure that the examples retained were directly interpretable in context.

The long tail of low frequency DMs appreciable in table 2, of the type that is familiar in frequency distributions across linguistics, is also evidence of the validity of these results. For example, there is no clear break or inflection point before which valid forms occur regularly, and after which invalid forms occur sporadically. More generally, the diversity of DMs observed in this study is broadly supported by Coats (2023), who also identifies a long tail of very infrequent, yet authentic DMs, based on a manual check of YouTube video excerpts with detailed recordings of oral speech. It is also important to note that even those forms produced in error can still be perceived as genuine combinations by other readers of the tweet: within CxG, Hoffmann (2020) has even argued that errors are a prime source of linguistic innovation, which means that even accidental DMs can turn out to be innovative grammatical constructs due to co-semiotic pressures. Therefore, despite their considerable variability and their low frequency in language use, the infrequent DMs in our dataset, many of which were previously unattested, do appear to be used in a functionally valid way to communicate specific meanings, as can be seen in (8), which presents three sample tweets with very infrequent DMs rarely (a) or never (b, c) attested in past research.

(8) (a) we **may can** do somethin aboutt it hahah ;p
     (b) if tell yur granda ye luv him then you**'d must** be together or something; idk??x
     (c) maybe we **can should** lend hull some of our buses hahaha come on hull

4.2 Regional patterns of double modals in the UK

Overall, the inventory identified above suggests that DMs make up an intricate system, with a distribution featuring a relatively short head and an especially long tail. Furthermore, many traditional DMs from past research in dialectology appear to be especially uncommon or even absent. As a way of completing our overview of this system, we can map tweets containing DMs to see if they show regional patterns in use on Twitter, thanks to the geolocation information included in the metadata of our dataset. To do this, we normalised the frequencies of each DM in each of the postcode region subcorpora, dividing them by the total number of words in each subcorpus and multiplying them by 1 billion, as the relative frequencies are generally low and often smaller than 1 when normalised by 1 million. Normalisation allows us to counterbalance the substantial differences in the amount of data available per region, usually following county population and social media usage: for instance, urban regions including large cities such as London provide a large quantity of tweets compared to rural regions such as the Scottish Highlands. In addition, we mapped the precise location of each DM based on the longitude and latitude in the metadata. In figure 2, we show an overall map with the combined relative frequencies of all 60 DMs, along with the individual maps for the first five most common DMs, *will can*, *will shall*, *would could*, *will might* and *would might*. The occasional misalignment between the shades and dots is explained by the substantial variation in the number of tweets as well as the general rarity of the feature, which is why it is useful to visualise both types of information.
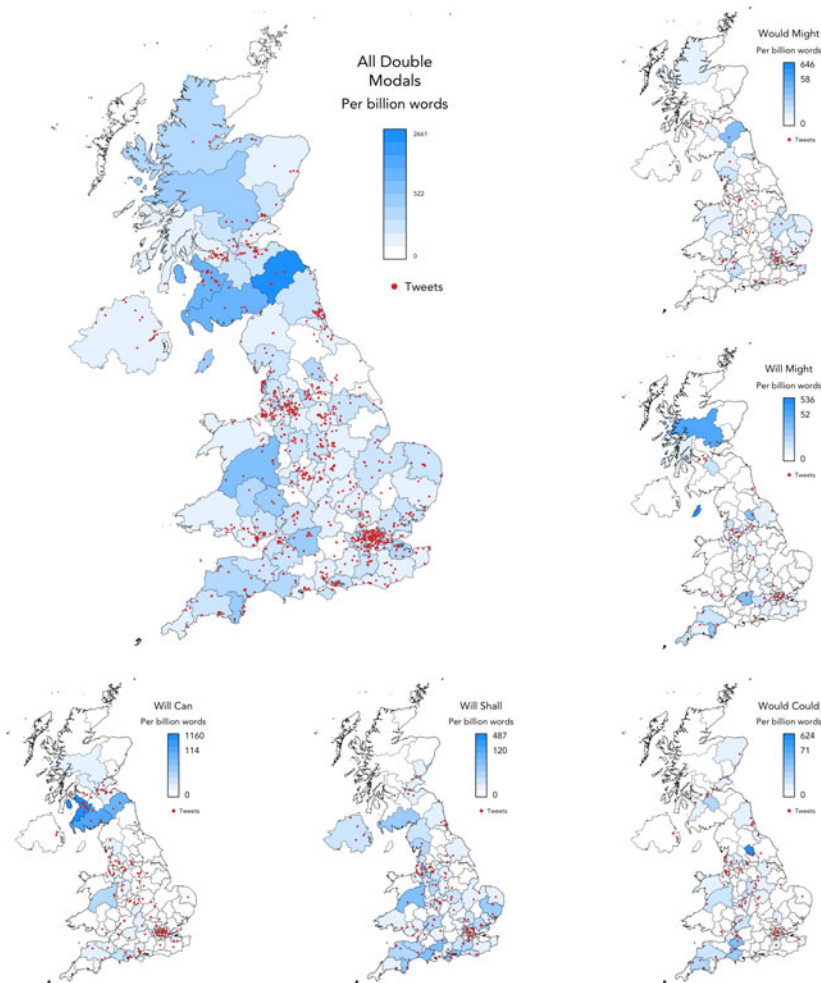
Figure 2. Frequency maps for common DMs on British Twitter (2014)

In the overall map, we find a prominent northern regional distribution, especially the Scottish Borders, as predicted by the literature, and to a much lesser extent in Northumberland, Tyneside and Northern Ireland. We also see smaller regional patterns in the Scottish Highlands and the urban area of the Scottish Central Belt, and some unexpected regions such as London (3,759 tokens per billion words), Manchester and the southwest of England more generally, but not Birmingham or Leeds. There is also a notable use of DMs in Wales, especially on the border.

In the individual DM maps, the establishment of regional patterns becomes more difficult, at least in part due to low frequency. The most common DM, *will can*, shows a clear distribution in the Scottish Lowlands and the Borders, although it is used to a lesser

extent in regions of England and Wales, including in cities, as can be seen from the clusters of data points. The second most common DM, *will shall*, appears to be most common in England and Wales, especially the south of England, although it also appears to some extent in the Scottish Lowlands and Northern Ireland. For the three remaining maps of *would could*, *will might* and *would might*, patterns are less clear in part due to the much smaller number of tokens, but the forms seem to be distributed across the UK, without regional specificities, except for *would might*, which appears to be slightly more common in the Scottish Borders. Importantly, then, based on these maps it appears that DMs in our dataset exhibit both regional patterns and a sporadic use of forms across the UK, corroborating the tendencies put forward by Coats (2023).

Figure 3 features maps only showing the tokens for nine traditional DMs in the Scottish Lowlands, northeast England and northeast Ireland (Beal 2004; Robinson 2018; Smith *et al.* 2019). Most of these DMs have very low frequencies and thus no clear regional distributions. The DMs *will can* and *would might* are common in the Scottish Lowlands and the Borders. The DMs *might can*, *used to could*, *must can* and *should can* occur at least once in either the Borders or Northumberland. Otherwise, these combinations do not show any clear regional specificity. Importantly, they appear to be almost completely absent from Northern Ireland. Alongside these DMs, the numerous other low-frequency combinations that were attested for the first time in our dataset, including those in (8), also do not show clear regional patterns. Broadly speaking, they seem consistent with the overall distribution in the Scottish Lowlands, as well as the North, the southwest of England and Wales, with sporadic uses across the UK.

In sum, DMs occur on British Twitter in a wider range of types, frequencies and regional distributions than assumed previously. However, several of the Scottish DMs appear where expected, and it seems rather unlikely that they are disappearing forms in these regions, since modern speakers appear to be using them in new registers, i.e. social media. The low frequencies of traditional DMs and their occasional occurrence outside the North can be interpreted in several ways. Although some of these forms may be especially rural, and restricted to older speakers, it is also likely that our corpus, despite being very large, is just too small and to some extent underrepresents their use. However, despite a distinct northern pattern, DMs do appear to be much less used in Northumberland, Tyneside and Ulster than in Scotland, with sparser relative frequencies and data points. This supports the claim of their relic, recessive, or even passive status in these varieties (Beal 2004; Corrigan 2011). In addition, and importantly, our results join those of Coats (2023) in questioning the assumption that British DMs are restricted to northern varieties, since we find various forms in smaller patterns across southwest England, Wales and even large urban areas in the southeast. We now turn to a step-by-step analysis of DMs as grammatical constructions.
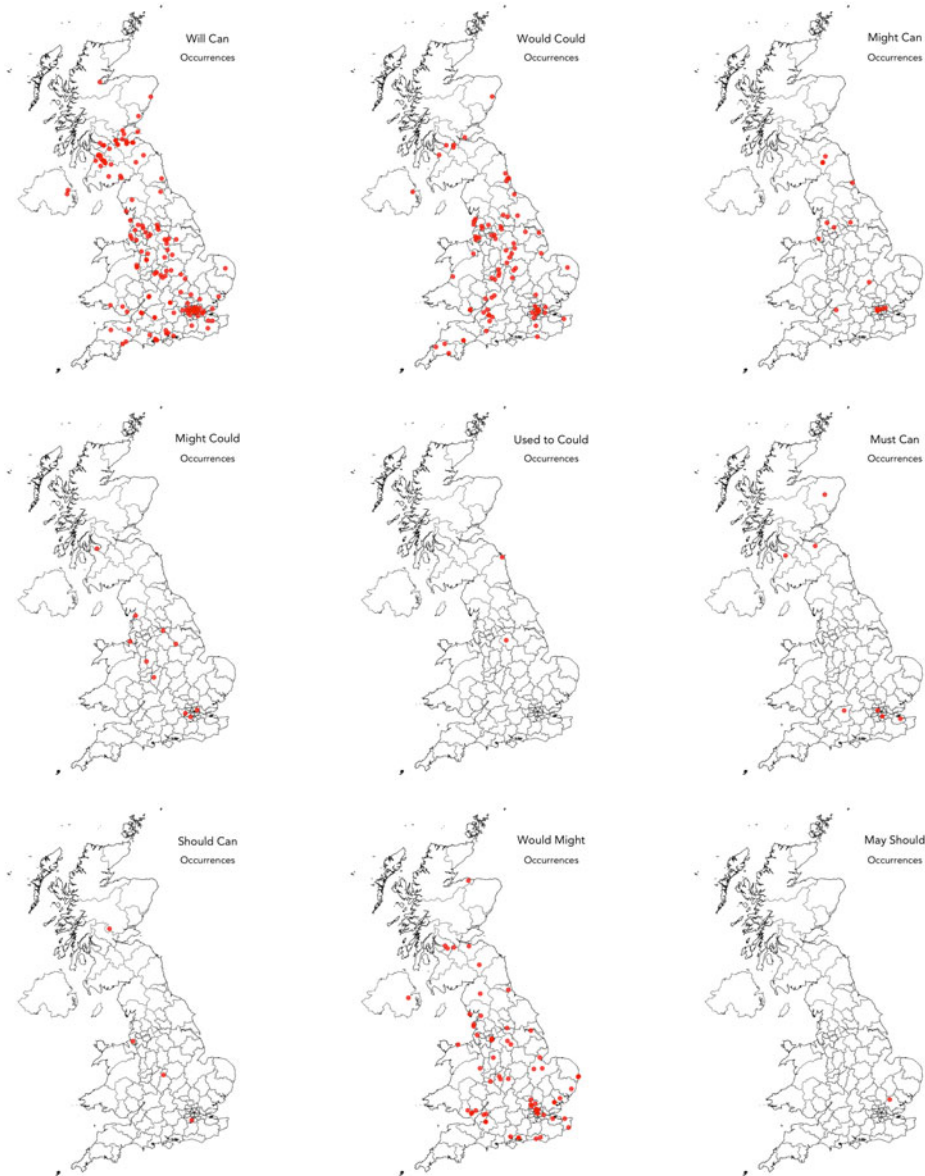
Figure 3. Frequency maps for DMs identified as dialect-specific in the literature on British Twitter (2014)

5 The construction network of double modals

5.1 The network structure of double modal forms

Following our inventory overview where DMs were taken as units, we also find important patterns when taking into account the internal structure of DMs. The network graph in
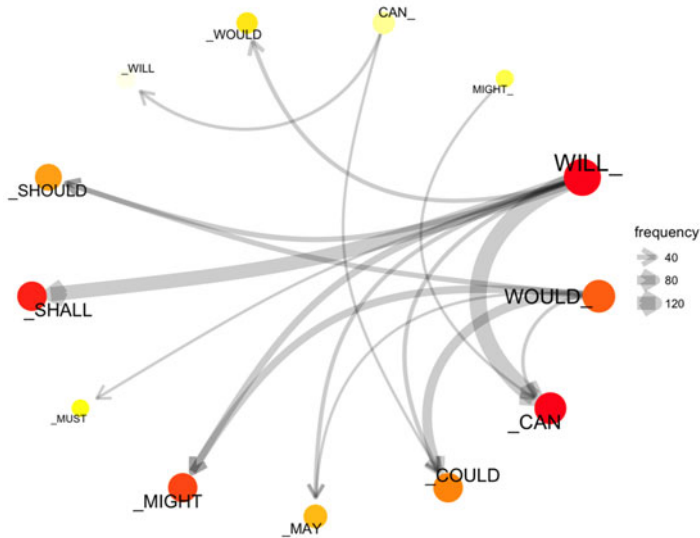
Figure 4. Eigenvector network graph of DMs modal position

figure 4 shows the co-occurrence patterns of modals, where the least frequent pairs (<15) have been excluded, resulting in the modals *ought to*, *used to* and *have to* being dropped.

The thickness of the edges represents the co-occurrence frequencies of the nine modals in first and second position. The nodes denote combinatorial schemas, coloured according to eigenvector centrality following a yellow to red gradient scheme. Shades of red indicate a higher eigenvector score, meaning that the node is connected to many nodes with high scores. We thus see that *will_X*, *would_X*, *X_can*, *X_shall*, *X_might* and *X_could* are the most productive DM schemas, including *will can*, *will shall* and *would could*, while also being productive for lower-level combinations such as *will must*, *would may* or *can could*. We see both clear positional preferences for individual modals (such as *will* and *would* as first modals, and *can*, *shall* and *might* as second modals) and some positional flexibility. The distribution of DMs thus appears to display a network structure.

This network structure is further evinced in figure 5, where correspondence analysis (CA) is used to summarise and explore attraction patterns between modals in combination as well as their relative flexibility in more detail. The basic idea behind CA is to plot variables across dimensions and see if some variables share similar profiles and appear to cluster together. First, it tests whether the association between the modal in first position (mod_1) and the modal in second position (mod_2) is statistically significant.[4] Second, CA plots the observations of mod_1 and mod_2 and

---

[4] More precisely, given a table of co-occurrence data, CA relies on the $\chi2$ test to determine whether rows and columns are independent. It tests the significance of the overall deviation of the table from the independence model. The test computes the contribution of each cell to $\chi2$ and sums up all contributions to obtain the $\chi2$ statistic.
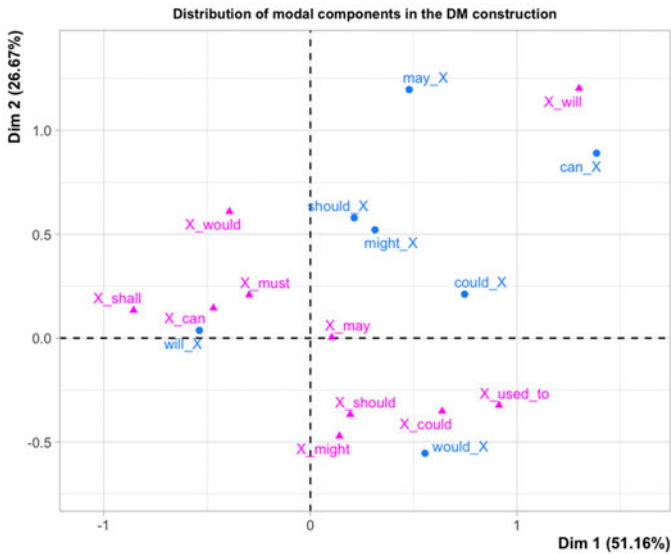
Figure 5. DM correspondence analysis graph across the first two dimensions

we can then see if they share similar distributional profiles. In figure 5, each type of mod_1 is in blue, and each type of mod_2 is in magenta.

The plot shows the first two dimensions, which account for 51% and 27% of the variance, 78% overall. This plot therefore visualises the most important trends in the joint distribution of mod_1 and mod_2 types. Judging from their extreme coordinates in the plot, three patterns of attraction stand out. *Will_X* is the only mod_1 type in the left part of the plot. It displays a strong association with *X_shall/can/must/would*. Another distinctive association is observed between *may/can_X* and *X_will* in the upper right corner of the plot. Last, *would_X* displays a preference for *X_might/should/could/used to*, as indicated by the cluster in the lower-right quadrant. The mod_1 types *should/might/could_X* and the mod_2 type *X_may* appear between these patterns, a sign of their greater combinatorial flexibility.

The CxG paradigm posits the existence of a constructicon (Goldberg 2013; Hoffmann & Trousdale 2013), a taxonomic network of knowledge containing more templatic (*schematic*) and more specified (*substantive*) form–meaning pairings known as constructions that are governed by hierarchical links. Such links include vertical default inheritance of formal and functional properties. Our data show evidence that DMs are grouped in such a structured inventory, with notable distributional skews. Based on this information, the representation principles of CxG as a theoretical model appear to provide a good fit for the network structure of our data, giving us further evidence that DMs may be grammatical constructions. The emergence of this network, however, is a surprising finding in light of what we know of the conventional constructicon of English (Kim & Michaelis 2020; Hoffmann 2022), which tends to pre-empt modal
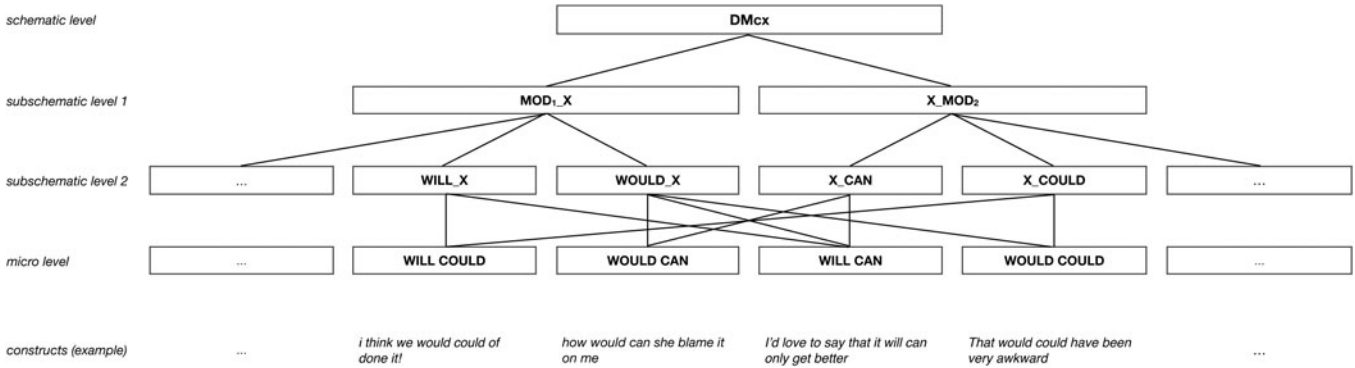
Figure 6. Network sketch of the DM construction based on the network graph and correspondence analysis of the Twitter dataset

stacking constructions, as will be discussed in section 6.2. Judging from their morphosyntax alone, the DMs subschemas display idiosyncratic connections within a consistent network structure, as revealed in both the eigenvector network graph and the CA plot. These graphs show that certain modals in certain positions are far more frequent than others and are connected to sets of modals of varying sizes. For instance, *will_X* and *X_can* are particularly productive templates, and *will can* is a very common instantiation of these templates.

In figure 6, we propose a sketch of the grammatical network of DM forms, based on a cline from the most substantive instantiations, i.e. the constructs themselves in tweets viewed as co-semiotic usage events (Schmid 2022), up to semi-schematic and fully schematic levels, licensing not only these two typical DMs, but other combinations with *will* and *would* first and *can* and *could* second, as well as other modal combinations.

There is one important caveat with this analysis: the inclusion of two levels of fully schematic macro-constructions is not obvious if we adopt a wholly bottom-up approach, given the constraint-like formal tendencies of the DM network as found on Twitter. We included these levels mainly to account for the long tail of low-frequency DMs that do not appear to follow specific combinatorics. We discuss these low-frequency DMs in more detail below, but for now we argue that their production, which is probabilistically dispreferred by the core subschemas of the network though not categorically pre-empted by them, justifies the postulation of DM macro-constructions and their conventional plausibility.

Although a CxG model seems explanatorily adequate so far, in its current state our network features only the formal side of DM constructions. An overview of their basic modal meanings is also necessary for a more complete account of grammatical constructions as *form–meaning pairings* (Hoffmann 2022).

5.2 The modal meaning of DMs

To provide a provisional meaning-based classification of DMs, we use the traditional categories defined by Palmer (1990) and Coates (1983) for English, especially Epistemic and Root, with Root further subdivided into Deontic (participant-external) and Dynamic (participant-internal). We also use a tweaked distinction between modal Possibility and Necessity based on Biber *et al.* (1999), subsuming Prediction under Epistemic Necessity and Volition under Root/dynamic Necessity.

First, we established a quantitative estimate of the meaning of the ten most frequent DMs in our corpus, by counting proportions of manually annotated modal meaning orderings of DMs in random samples of fifty tokens when possible for these top ten DM types. This estimate is shown in figure 7, and qualitative individual samples for three of the five first DM types are shown in (9). We carefully coded each token by considering the immediate context of the tweet containing a DM, relying on features that generally enable annotators with English as a native language to recognise the type of modality at play and identify the most likely intended meaning. These include the subject and main verb type of the sentence in which the DM is used in order to
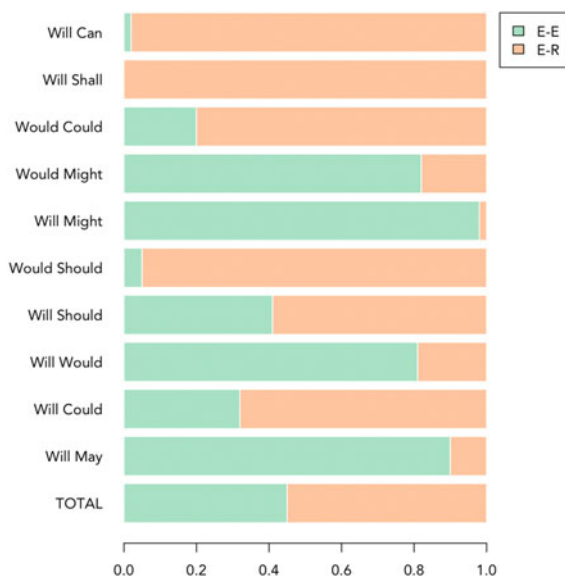
Figure 7. Proportions of modal meaning ordering (Epistemic–Epistemic or Epistemic–Root) of DMs in random samples for the top ten DMs (ranked on the Y axis) in the corpus (n = 50, all tokens for DMs with a frequency < 50)

establish the scope of the modal over the subject participant inside the proposition or the entire proposition itself, interjections, punctuation marks and emojis or emoticons as forms of gesture that help speakers convey attitudes and judgments about events and propositions online (McCulloch 2019).

(9)   (a)   few things i need to sort. bills 😔 then i**'ll can** make a decision xxxxx (Epistemic Root/
              dynamic, Necessity–Possibility)
       (b)   oh my god yes and they can bow down to me and i **will shall** be named the king of all the
              penguins (Epistemic–Root/deontic, Necessity–Possibility)
       (c)   thought you**'d might** like to see it seems as though he was named after you! (Epistemic–
              Epistemic, Necessity–Possibility)

Overall, a general ordering pattern of Epistemic > Root/deontic > Root/dynamic appears to account for modal meaning, reflecting previous accounts of the cross-linguistic ordering of modal combinations (Zullo *et al.* 2020; Werkmann Horvat 2021): two typical cases are Epistemic > Root(deontic/dynamic) and double epistemics (Epistemic–Epistemic). The first configuration (E–E) is predominant, accounting for six DMs: *will can*, *will shall*, *would could*, *would should*, *will should* and *will could*. The second configuration (E–R) is also very common, accounting for the majority of tokens of the DMs *would might*, *will might*, *will would* and *will may*. Notably, none of the tokens in these high frequency DMs are double roots (R–R) and none follow the inverse order of Root > Epistemic. Importantly, few of the individual DM types feature

Table 3. *Proportions of semantic ordering of DMs with less than five tokens, from* can should *to* might shall *(n = 66)*

DM	Epistemic First		Root First	
	E–E	E–R	R–R	R–E
can should (...) might shall	0.3	0.21	0.2	0.29

a balance between E–E and E–R, and they are often skewed towards one or the other ordering.

The general skew of each of these common forms suggests that DMs are generally specialised in the expression of one of two sequences of modal meaning, which can be described in cognitive semantic terms (Talmy 1988):

(10) (a) **Epistemic – Root**: the propositional reasoning-based (lack of) exertion of force on the physical/psychosocial (lack of) exertion of force on the agonist:[5] *will can*, *will shall*, *would could*, *would should*, *will should*, *will could*.

   (b) **Epistemic – Epistemic**: the propositional reasoning-based (lack of) exertion of force on *another* propositional (lack of) exertion of force on the Agonist: *would might*, *will might*, *will would*, *will may*.

We find another important meaning tendency in the co-occurrence patterns of modals based on what we know of the meaning patterns of these individual modals in English. Firstly, *will* and *would*, the only two modals in first position, are typically understood as expressing modal necessity as opposed to possibility. Alternatively, *can*, *could*, *may* and *might*, which account for a majority of modals in second position, are possibility modals (Palmer 1990). We thus see that necessity is the primary modality across positions, and possibility is only second, despite its involvement in a number of combinations. Most notably, DMs appear to primarily involve the sequential addition of necessity modality and possibility modality, especially the addition of the necessity modals *will*/*would* and a possibility modal.

Reflecting the formal distribution of DMs, however, the sequences of Necessity + Possibility and Epistemic > Root appear to best describe the core nodes of the network, breaking down as we move down the long tail of the distribution. This is illustrated in table 3 and in (11), which shows surprising proportions of double roots and especially the inverse order of Root > Epistemic, and overall similar rates for all possible orderings. These unusual meanings seem to correlate with unusual forms; these unusual constructs in the periphery of the network are interpreted below.

---

[5] Within force dynamics and cognitive construals of interactions between entities, Talmy (1988: 53) defines the agonist as the 'force-exerting entity ... singled out for focal attention': it is characterised by whether it 'is able to manifest its force tendency or, on the contrary, is overcome'.

(11) (a) sexy af. she **could might** get it ..

        (Root/dynamic–Epistemic, Possibility–Possibility)

    (b) hmm that's true! 🙍 hopefully we **can may** be able to get some tomorrow 😆 🙏

        (Root/dynamic–Epistemic, Possibility–Possibility)

    (c) shit how something **shall can** ruin your whole night 😫

        (Root/deontic – Epistemic, Necessity–Possibility)

So far in our demonstration, DMs appear to make up a network of grammatical constructions that combine two modals. Their primary function is to express a propositional, reasoning-based exertion on the physical or psychosocial absence of force on the agonist. The propositional basis of force is associated with epistemic modality, while the physical and social basis of force is associated with root modality. The presence of exertion is associated with necessity modality, while the absence of exertion is associated with possibility modality. Although this general ordering is the most common, other meanings are possible, especially the double epistemic sequence, associated with specific DMs in the network.

However, in the next section we argue that our description in its current state is insufficient, and that another type of meaning other than basic modal meaning is needed to achieve a more explanatorily adequate analysis of DM constructions: social meaning. To do this, we first discuss DMs in terms of necessary and sufficient criteria to assess their constructionhood.

6 The social meaning of double modals

6.1 *Issues of constructionhood*

The traditional approach in CxG to justify the existence a construction is to check whether a number of 'criteria of constructionhood' apply (Ungerer 2023): in earlier frameworks, high frequency and non-compositionality of form or meaning were the two main diagnostics (Goldberg 1995). Both types of criteria appear to be problematic when we apply them to the case of DMs: crucially, however, we believe that the explanatory gaps between these criteria and the data we found of DM constructions can be filled by including the factor of social meaning.

Firstly, DMs do not appear to fully meet the high frequency criterion. That is not to say that we do not consider frequency pressures to play a role in the description of DMs in our corpus: indeed, we found a substantial and complex dataset of DMs distributed in a network structure, which suggests that frequency must constitute an influential process in the DM system to some extent, for instance at the meso-level where the combinatorial schemas have acquired a higher degree of prototypicality with respect to the set of lower-order instantiations. Furthermore, in keeping with the usage-based commitment of CxG (Bybee 2013), it is also likely that a few of the most frequent DMs, such as *will can* and *would could*, are typical exemplars which, through repetition and entrenchment alongside other combinations with *will* and *would* as first modal and

can and *could* as second modal, have become influential nodes of the DM construction. Despite these patterns, however, the overall frequencies in our large corpus of natural language are noticeably low, reflecting the fact that DMs are rare syntactic constructions with low frequencies across dialects and registers of English, including Twitter. This challenges the sufficiency of frequency as the sole basis for their constructionhood.

Secondly, to justify the status of construction, one criterion that is often proposed is non-compositionality: a construction C is 'a form–meaning pair $<F_i, S_i>$ such that some aspect of F_i or some aspect of S_i is not strictly predictable from C's component parts or from other previously established constructions' (Goldberg 1995: 4). Based solely on the analysis of the modal meaning of DMs above, it would appear that these constructions do not meet this criterion: they consist of two combined modals, which appear to add meaning in a compositional manner. The only alternative explanation of DM constructions likely to satisfy the non-compositionality criterion would be that it is in fact their *social meaning*, discussed in more detail in section 6.2, which acts as a non-compositional constraint on their knowledge and use: such social meaning would complement (low) frequency pressures, and perhaps in fact arise as a result of this low frequency characterising a stigmatized, indexical and highly socially salient, yet maintained, dialect feature. If this hypothesis were to be confirmed in future research, this would have important consequences for CxG theorising, giving us a case of a construction that relies on a previously underappreciated combination of usage-based and social pressures, namely social non-compositionality correlated with low frequency.

Keeping this hypothesis for the time being, however, it is also the case that the non-compositional criterion is not considered a necessary and sufficient condition anymore since Goldberg (2006), and patterns that are both general and predictable deserve a place in the constructicon as long as some distributional and functional constraints apply. This may more adequately explain the case of DMs, as evidenced above. Through repetition, even patterns that are originally general and decomposable into constituents end up being treated as units. This is known as *chunking*, a domain-general cognitive process allowing fully compositional constructions to be stored and accessed directly (Bybee 2013). Given the statistical evidence presented above, the most frequent DMs are likely to be treated as chunks. We thus believe that the process of chunking, alongside low frequency and non-compositional social meaning, is an important explanatory factor of the formation of DM constructions deriving some of its main probabilistic constraints, including the modal sequences of Necessity > Possibility, Epistemic > Root, and Epistemic > Epistemic. These sequences and social meaning should thus be included in the meaning component of the constructions in the provisional network above, accounting for the central DM system.

What remains to be explained in our dataset is the substantial number of low-frequency DMs that do not appear to conform to these formal and modal meaning ordering tendencies, occurring unconstrainedly across the UK. We propose the following hypothesis: first, in regions of the UK where the DM system is prominent, these infrequent DMs may be spontaneous innovations by speakers experimenting with the DM construction. Their atypical form and meaning result in relative pre-emption by the

central system, although this does not discourage speakers from engaging in this process consistently as part of their linguistic creativity. Second, in regions of the UK where the DM system is not prominent, they may also be spontaneous innovations that are even less likely to thrive because of statistical pre-emption factors, most importantly the unacceptability of any kind of modal stacking, as suggested in the quote of an RP speaker in section 1 and further discussed below. In both cases, innovated DMs are likely to attract non-standard social meaning as a result of their formal markedness with respect to standard conventions, which we discuss in more detail below.

6.2   The distinctive force of social meaning

To sum up our discussion so far, although we postulated the DM construction based on formal patterns, modal meaning patterns, and the usage-based pressures of frequency and chunking in the previous sections, we believe these are insufficient to adequately describe DMs, and that social meaning is an essential explanatory factor, especially in light of the low frequency of these constructions. In addition, the integration of this factor enables us to explain the maintenance of DM use against the strong pre-emptive force of standard modal constructions, which involves a finite-verb constraint on their use.

Indeed, a crucial generalisation characterising the conventional network of standard English grammatical constructions is that modal constructions inherit a strong constraint from auxiliary constructions (Kim & Michaelis 2020: 186–212), itself inherited from verbal constructions (Hoffmann 2022: 206–11), which restricts the category of finiteness (i.e. tense) to one verb maximum per clausal construction. This highly general constraint of English grammar interacts with the lack of non-finite (i.e. untensed) forms in the lexical construction paradigm of the English modals. As a result, modal stacking inside a clause is strongly probabilistically disfavoured, compared to an alternate construction which would convey a very similar complex modal meaning and otherwise satisfy the one-tensed-verb constraint: Brown (1991) cites two prominent alternatives, which are semi-modal constructions (e.g. *will be able to*) or adverbial constructions (*maybe* instead of *might* in *might could*).

If DMs are constructions, there must be an equally strong distinctive aspect of their meaning explaining their entrenchment and conventionalisation against this constraint. By limiting our view of the meaning component of DM constructions to their modal meaning, we are unable to explain why they are not blocked by standard modal alternatives expressing overlapping modal meanings (Brown 1991), in the same way that a double negative (*I don't have no homework*) is prescriptively dispreferred, though not blocked, by a standard negative (*I don't have any homework*). This type of meaning distinction, expected from the co-existence of DMs and standard modal constructions in a general grammar network, would be captured by modern versions of the CxG principle of No Synonymy (Goldberg 2006; Leclercq & Morin 2023), where social meaning is featured as one of the drivers of non-synonymous constructions.

A defining factor likely to explain the constructionhood of DMs is thus the following: the specific association of modal periphrastic or adverbial constructions with a standard

network of grammatical constructions on the one hand, and the specific association of formally marked DMs with a non-standard network of grammatical constructions on the other, creates a relationship between the two likely to be conveyed in terms of social meaning, based on the usage-based pressure of social *salience* (Schmid 2012). This social meaning is a function of positions of varieties and grammatical networks of English in a hierarchy based on prescriptive systems of values and norms, and the results of dynamic acts of identity (Eckert 2012). The non-standard network is stigmatized towards informal language registers, hence the occurrence of DMs in the most informal registers of British Twitter shown in the sample tweets above, which include many other features of non-standard online language, such as lack of final punctuation, acronyms, swearing and emojis (McCulloch 2019).

If the most distinctive aspect of DMs is their social meaning, this has important consequences for CxG as a theory of language, which generally seems to underestimate the role of social meaning in networks of grammatical constructions. Social meaning deserves its own subcomponent in the pragmatic meaning of constructions (Hoffmann 2022), which crucially includes elements of situational and sociocultural setting, two important sources of variable sociolinguistic conventions (Schmid 2022). In the following paragraphs, we further discuss the potential richness of social meaning, based on the regional variation in DM use, as observed in this study.

British DMs on Twitter are rare grammatical constructions that appear to be distinguished by the social meaning of non-standard and informal register, enabling them to resist pre-emption by standard modal constructions. In addition, we propose
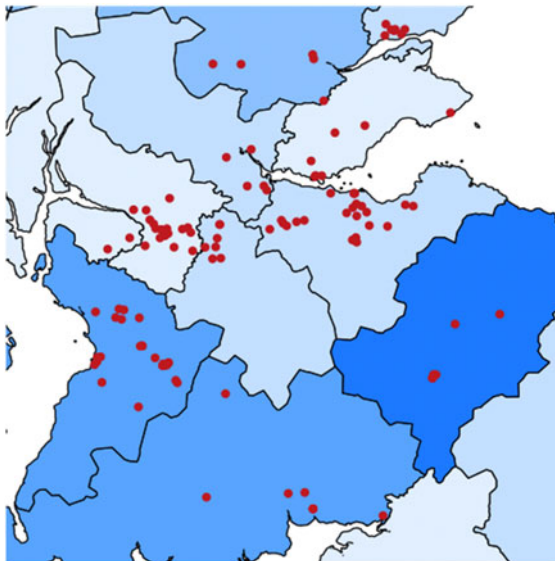

Figure 8. Relative frequencies by postcode region and individual tokens of all DMs in the Scottish Lowlands

that their social meaning includes a regional dialect meaning, inferred from regional variation in their use. Here we focus on the northern dialect region and the Scottish Lowlands in particular, zooming in on the general map of all DMs of figure 2 in figure 8 below.

By analysing a single, streamlined corpus of geolocated natural language, we see how it is now possible to substantiate these claims in finer-grained detail than before. Firstly, we are able to establish the correlation of DM usage with regional variation: the most prominent regional pattern is the northern dialect pattern previously assumed in the literature, especially in the Borders and the Scottish Central Belt. This correlation is taken to have given rise to a specific social meaning of these forms compared to other regions, in addition to the concepts 'non-standard' and 'informal'. Thus it is very likely that DMs are also most typically associated with the concept of 'Scots' and especially 'Borders Scots', to a lesser extent 'Urban Scots', and that this concept contributes to distinguishing them from alternatives associated with the concepts of 'Standard Scottish English' or more generally 'Standard British English' (see also Östman & Trousdale 2013).

Secondly, we can also take individual natural, geolocated tokens of DMs in tweets viewed as co-semiotic usage events (Schmid 2020) and look at their regional and linguistic context to see whether DM constructions are found alongside other constructs known to be prominent in grammar networks of Scots. Below are a few qualitative examples of tweets including the specific town and postcode region in which they were produced, based on their latitude and longitude.

(12) (a) aye **will can** just use ginger bottles; what a beaut?? (Arbroath, Dundee)

  (b) dan wallace still a total twat then; **might can** get his teeth fixed with the £20k he's totalled in price money. sorry jak if you read this. (Hawick, Galashiels/Selkirk)

  (c) **might can** no dae it get in there (Hawick, Galashiels/Selkirk)

  (d) will say anything they think will worry voters. if they said "indy means more tory govts" we**'d might** vote no! (Bearsden, Glasgow)

  (e) young keegans an mclaughlin? nice young men them.. we**'ll can** get a wee sesh sorted soon dear (Grangemouth, Falkirk)

  (f) will there be anywhere in hawick that **will can** repair a laptop screen? (Hawick, Galashiels/Selkirk)

  (g) aww naw a don't think i**'ll can** deal with your paranoia the day (Cumnock, East Ayrshire)

  (h) wit the fuck has pregnancy done to me?! greeting because i**'ll can** feed the ducks when i go on holiday! (Cumnock, East Ayrshire)

  (j) they're shite at soccer so they**'ll may** as well look shite playing it. glitter on a jobby springs to mind. (Dumfries, Dumfries and Galloway)

Upon close examination of these individual tweets, DMs seem natural in pragmatic/discursive contexts favouring the use of constructions from a dialect grammar, co-occuring with a number of well-known dialect features of Scots (Miller 2004; Smith *et al.* 2019): for example, the phonological variants of the morphological constructions *dae* for *do*, *no* for *not*, *a* for *I*, *wit* for *what*, *naw* for *no*; dialect-specific

lexical items such as *wee* (small), *aye* (yes), *the day* (today), *greeting* (crying), *jobby* (turd); swear words (*twat*, *fuck*, *shite*), and socially specific cultural references and topics (Daniel Wallace is a Scottish Olympic swimmer; 2014 saw the debates over the Scottish Independence Referendum; Hawick is a typical location for DM use; football is the most popular sport in working-class Scotland as in the rest of the UK). This is important qualitative evidence that DMs are syntactic constructions with particularly salient dialect indices used in sociolinguistic acts of identity, occupying a specific place in the modal expressions of a dialect grammar distinguished from the standard grammar, and including the social concepts 'non-standard', 'informal' and 'region' in the form of constraints on their knowledge and use.

7 Conclusion

In this article, we unpacked some of the consequences of the social turn of Cognitive Linguistics for Construction Grammar. To do this, we analysed a particularly rare, unusual and socially salient grammatical construction in English, which we argued could not be well defined without taking the Sociosemiotic Commitment: the double modal construction in British English used on social media. We took advantage of the computational sociolinguistic approach, analysing language variation in a large corpus of informal natural language produced online, as is increasingly used in cognitive sociolinguistics. This study demonstrates the necessity of modelling social meaning and dialect variation in CxG for more general issues in this theoretical framework, for instance a more refined view of what defines a grammatical construction. It has also highlighted the importance of building a finer-grained model of semantic and pragmatic meaning in CxG, notably through the concept of language register and regional variation, which can become indexical dimensions of constructional meaning, and act as constraints on their knowledge and use. The consequences of this case study of DMs for the framework of Construction Grammar are the following, from the most general to the most specific:

- The study of sociolinguistic variation in this CL theory of language is not just opportunistic or a luxury, but a necessity for its capacity to explain all of language, including its social aspects (Dąbrowska 2016), and following the Sociosemiotic Commitment (Geeraerts 2016).
- This necessity is illustrated by the case of dialect constructions (e.g. DMs), which are difficult to adequately explain without a working concept of social meaning.
- The integration of social meaning in a CxG model raises new questions on how to define constructions, and what criteria to use for establishing them, especially in terms of frequency, compositionality of meaning and chunking. Specifically, we have a case where a network of low-frequency construction appears to emerge and be maintained thanks to social meaning constraints.
- New methodological paradigms such as computational sociolinguistics enable the study of dialect constructions, including the rarest and most unusual ones (e.g. DMs).

- The social meaning of constructions is primarily represented as static constraints on knowledge and use, but is also the result of dynamic acts of identity in interaction (Eckert 2012) and rich mutual understanding of situational and sociocultural setting in co-semiosis (Schmid 2022).
- For dialect constructions, social meaning needs to be represented as a subcomponent of the extended PRAGMATICS component of the meaning of constructions (Hoffmann 2022).

We hope this study will prompt further research on construction-based sociolinguistics and dialectology by asking more precise questions on the respective roles of these meaning features in constructions and grammar networks, and testing hypotheses observationally on a wider range of constructions using online natural language data. The pursuit of these questions can only contribute to the growth of CxG as a theory of language.

Authors' addresses:

Ecole Normale Supérieure de Lyon
15 parvis René Descartes
69007 Lyon
France
cameron.morin@ens-lyon.fr

Université de Paris Vincennes Saint-Denis
2 rue de la Liberté
93526 Saint-Denis
France
gdesagulier@univ-paris8.fr

University of Birmingham
Frankland Building
Edgbaston
Birmingham B15 2TT
United Kingdom
j.grieve@bham.ac.uk

References

Beal, Joan. 2004. English dialects in the North of England: morphology and syntax. In Kortmann & Schneider (eds.), 114–41.

Beal, Joan & Karen Corrigan. 2000. Comparing the present with the past to predict the future of Tyneside English. *Newcastle and Durham Working Papers in Linguistics* 6, 13–30.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.

Bour, Anthony. 2014. Description of multiple modality in contemporary Scotland: double and triple modals in the Scottish Borders. MS, University of Freiburg.

Brown, Keith. 1991. Double modals in Hawick Scots. In Peter Trudgill & J. K. Chambers (eds.), *Dialects of English: Studies in grammatical variation*, 74–103. Harlow: Longman.

Butters, Ronald. 1973. Acceptability judgments for double modals in Southern dialects. In Charles James Bailey & Roger W. Shuy (eds.), *New ways of analyzing variation in English*, 276–96. Washington, DC: Georgetown University Press.

Bybee, Joan L. 2013. Usage-based theory and exemplar representations of constructions. In Hoffmann & Trousdale (eds.), 49–69.

Claes, Jeroen. 2018. Probabilistic Grammar: The view from Cognitive Sociolinguistics. *Glossa* 2(1), 1–32

Coates, Jennifer. 1983. *The semantics of modal auxiliaries*. London and Canberra: Croom Helm.

Coats, Steven. 2023. Double modals in contemporary British and Irish speech. *English Language and Linguistics* 27(4), 693–718.

Corrigan, Karen. 2011. Grammatical variation in Irish English. *English Today* 27(2), 39–46.

Croft, William. 2009. Towards a Social Cognitive Linguistics. In Vyvyan Evans & Stéphanie Pourcel (eds.), *New directions in Cognitive Linguistics*, 395–420. Amsterdam: John Benjamins.

Croft, William & David Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Dąbrowska, Ewa. 2016. Cognitive Linguistics' seven deadly sins. *Cognitive Linguistics* 27(4), 479–91.

Dunn, Jonathan. 2018. Finding variants for construction-based dialectometry: A corpus-based approach to regional CxGs. *Cognitive Linguistics* 29(2), 275–311.

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

Geeraerts, Dirk. 2016. The sociosemiotic commitment. *Cognitive Linguistics* 27(4), 527–42.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago and London: University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, Adele. 2013. Constructionist approaches. In Hoffmann & Trousdale (eds.), 15–31.

Gries, Stefan Th. 2013. Data in Construction Grammar. In Hoffmann & Trousdale (eds.), 93–110.

Grieve, Jack. 2021. Observation, experimentation, and replication in linguistics. *Linguistics* 59(3), 1343–56.

Grieve, Jack, Dirk Hovy, David Jurgens, Tyler S. Kendall, Dong Nguyen, James N. Stanford & Meghan Sumner (eds.). 2023. *Computational Sociolinguistics*. Lausanne: Frontiers Media SA.

Grieve, Jack, Chris Montgomery, Andrea Nini, Akira Murakami & Diansheng Guo. 2019. Mapping lexical dialect variation in British English using Twitter. *Frontiers in Artifical Intelligence* 2, 1–18.

Hall-Lew, Lauren, Emma Moore & Robert Podesva. 2021. Social meaning and linguistic variation: Theoretical foundations. In Lauren Hall-Lew, Emma Moore & Robert Podesva (eds.), *Social meaning and linguistic variation: Theorizing the third wave*, 2–23. Cambridge: Cambridge University Press.

Hoffmann, Thomas. 2015. Cognitive sociolinguistic aspects of football chants: The role of social and physical context in Usage-based Construction Grammar. *Zeitschrift für Anglistik und Amerikanistik* 63(3), 273–94.

Hoffmann, Thomas. 2020. Speakers are creative, within limits – a response to Peter Uhrig. *Cognitive Semiotics* 13(1), 1–7.

Hoffmann, Thomas. 2022. *Construction Grammar: The structure of English*. Cambridge: Cambridge University Press.

Hoffmann, Thomas & Graeme Trousdale (eds.). 2013. *The Oxford handbook of Construction Grammar*. Oxford: Oxford University Press.

Hollmann, Willem. 2013. Constructions in cognitive sociolinguistics. In Hoffmann & Trousdale (eds.), 491–510.

Huang, Nick. 2011. Multiple modals. *Yale Grammatical Diversity Project: English in North America*. https://ygdp.yale.edu/phenomena/multiple-modals (accessed 31 August 2023).

Huddleston, Rodney & Geoffrey Pullum *et al*. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Hudson, Richard. 2007. English dialect syntax in Word Grammar. *English Language and Linguistics* 11(2), 383–405.

Hudson, Richard. 2021. Word Grammar. In John Taylor & Xu Wen (eds.), *The Routledge handbook of Cognitive Linguistics*, 111–26. New York and London: Routledge.

Irvine, Judith. 2001. 'Style' as distinctiveness: The culture and ideology of linguistic differentiation. In Penelope Eckert & John R. Rickford (eds.), *Style and sociolinguistic variation*, 21–43. Cambridge: Cambridge University Press,

Kim, Jong-Bok & Laura Michaelis. 2020. *Syntactic constructions in English*. Cambridge: Cambridge University Press.

Kortmann, Bernd, Kerstin Lukenheimer & Katharina Ehret (eds.). 2020. *The electronic world atlas of varieties of English*. Zenodo. https://ewave-atlas.org/ (accessed 31 August 2023).

Kortmann, Bernd & Edgar W. Schneider (eds.). 2004. *A handbook of varieties of English*, vol. 2: *Morphology and syntax*. Berlin: Mouton de Gruyter.

Kristiansen, Gitte, Mariana Montes & Zhang Weiwei. 2021. Cognitive Sociolinguistics in the 21st century. In Gitte Kristiansen, Karlien Franco, Stefano De Pascale, Laura Rossel & Weiwei Zhang (eds.), *Cognitive Sociolinguistics revisited*, 3–20. Berlin and Boston: De Gruyter.

Labov, William. 1969. Contraction, deletion, and the inherent variability of the English copula. *Language* 4(4), 715–62.

Leclercq, Benoît & Cameron Morin. 2023. No equivalence: A new principle of no synonymy. *Constructions* 15, 1–16.

McCulloch, Gretchen. 2019. *Because Internet: Understanding the new rules of language*. New York: Riverhead Books.

Miller, Jim. 2004. Scottish English: morphology and syntax. In Kortmann & Schneider (eds.), 47–72.

Montgomery, Michael & Stephen Nagle. 1993. Double modals in Scotland and the United States: Trans-atlantic inheritance or independent development? *Folia Linguistica Historica* 14(1), 91–107.

Morin, Cameron. 2021a. Construire une triangulation méthodologique en variation syntaxique restreinte. *Echanges Linguistiques en Sorbonne* 7, 48–71.

Morin, Cameron. 2021b. Double modals in Scots: A Speaker's Choice hypothesis. In Laura Baranzini & Louis de Saussure (eds.), *Aspects of tenses, modality, and evidentiality*, 165–82. Leiden: Brill.

Morin, Cameron. 2023. Social meaning in Construction Grammar: Double modals in dialects of English. MS, Université Paris-Cité.

Morin, Cameron, Guillaume Desagulier & Jack Grieve. 2020. Dialect syntax in Construction Grammar: Theoretical benefits of a constructionist approach to double modals in English. *Belgian Journal of Linguistics* 34, 252–62.

Nini, Andrea, George Bailey, Diansheng Guo & Jack Grieve. 2020. The graphical representation of phonological dialect features of the North of England on social media. In Patrick Honeybone &

Warren Maguire (eds.), *Dialect writing and the North of England*, 266–96. Edinburgh: Edinburgh University Press.

Östman, Jan-Öla & Graeme Trousdale. 2013. Dialects, discourse, and Construction Grammar. In Hoffmann & Trousdale (eds.), 476–90.

Palmer, Frank. 1990. *Modality and the English modals*, 2nd edn. London and New York: Longman.

Robinson, Philip. 2018. *Ulster-Scots: A grammar of the traditional written and spoken language*, 3rd edn. Belfast: Ullans Press.

Schmid, Hans-Jörg. 2012. Entrenchment, salience, and basic levels. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of Cognitive Linguistics*, 117–38. Oxford: Oxford University Press

Schmid, Hans-Jörg. 2020. *The dynamics of the linguistic system*. Cambridge: Cambridge University Press.

Schmid, Hans-Jörg. 2022. Social and cognitive foundations of linguistic variation: Towards a unified complex-adaptive framework. *Abralin ao Vivo*, www.youtube.com/watch?v=AMVkj97egDw (accessed 31 August 2023).

Schmid, Hans-Jörg, Quirin Würschinger, Sebastian Fischer & Helmut Küchenhoff. 2021. *That's cool*. Computational sociolinguistic methods for investigating individual lexico-grammatical variation. *Frontiers in Artificial Intelligence* 3. https://doi.org/10.3389/frai.2020.547531 (accessed 31 August 2023).

Smith, Jennifer, David Adger, Brian Aitken, Caroline Heycock, E. Jamieson & Gary Thoms. 2019. *The Scots syntax atlas*. University of Glasgow. https://scotssyntaxatlas.ac.uk (accessed 31 August 2023).

Talmy, Leonard. 1988. Force dynamics in language and cognition. *Cognitive Science* 12(1), 49–100.

Tomasello, Michael. 2014. The ultra-social animal. *European Journal of Social Psychology* 44, 187–94.

Trousdale, Graeme. 2000. Variation and (socio)linguistic theory: A case study of Tyneside English. MS, University of Edinburgh.

Ungerer, Tobias. 2023. A gradient notion of constructionhood. *Constructions* 15(1), 1–20.

Werkmann Horvat, Ana. 2021. *Layers of modality: On double modal constructions by the example of Croatian*. Berlin and Boston: Mouton de Gruyter.

Zullo, David, Simone Pfenninger & Daniel Schreier. A pan-Atlantic 'multiple modal belt'? *American Speech* 96(1), 1–46.