Language Variation and Change (2025), **37**, 293–318
doi:10.1017/S0954394525100628

CAMBRIDGE UNIVERSITY PRESS

Lexical variation and change: Integrating lexis into variationist sociolinguistics

Rhys Sandow[1], Jack Grieve[2], Rose Stamp[2] and Adam Schembri[2]

[1]Faculty of Media, Arts and Humanities, University of Sussex, Brighton, UK and [2]Department of Linguistics and Communication, University of Birmingham, Birmingham, UK
Corresponding author: Rhys Sandow; Email: rhys.sandow@sussex.ac.uk

Abstract
In this paper, we advocate for the analysis of lexical variation being central to variationist sociolinguistics. We demonstrate that lexical variation is systematic and argue that this systematicity must be accounted for by a comprehensive variationist theory that explains the general causes and mechanisms of language variation and change. We present three empirical studies, which focus on lexical variation in Anglo-Cornish, British Sign Language, and online American English. These studies differ greatly in terms of their methods and results, but each reaffirms that lexis can be studied rigorously and informatively within the variationist paradigm, extending our understanding of language variation and change.

Keywords: lexical variation; words; signs; variationist sociolinguistics; sociolinguistic theory; language change

Introduction

Variationist sociolinguistics has long resisted incorporating observations about lexis into theories of language variation and change. This is, to a large extent, predicated on a belief exemplified by, among many others, William Labov, one the founders of our field, in his 2012 book (Labov, 2012:17):

> There is of course a great deal of fluctuation in vocabulary. We all find it interesting to learn that what is called *soda* in one place is called *pop* in another place, and *coke* in another. Yet the change of one word does not tell us much about change in another, and the long list of words that differ from one place to another does not form a coherent pattern or give us much insight into the machinery of speaking and listening.

From this perspective, lexical variation is seen as defying systematic description and thus falling outside the scope of sociolinguistic theory. Although far from a consensus, this assumption appears to be tacitly accepted across much of the field and has been

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

echoed by many other prominent sociolinguists, including Adams (2014:164), who wrote that "words do not behave systematically," and Trudgill (1986:25), who wrote that lexical variation is "(mostly) non-systematic." Sociolinguists do not doubt that the use of individual lexical items is socially patterned (although, see Nagy, 2011); they doubt whether these patterns are systematic, recurring across lexical items, at least in ways that are relevant to our understanding of the basic mechanisms of language variation and change more generally. Each lexical item is assumed to show its own pattern, reflecting the unique history of that word or sign, as opposed to reflecting general sources of sociolinguistic variation.

While the study of lexis has always been central to dialectology, the relative lack of engagement with lexical variation in variationist sociolinguistics[1] can be clearly seen in the publication patterns of this journal, *Language Variation and Change*—the journal of record in our field. For example, between 2012 and 2023, 177 research articles were published. Of these articles, only 10 (6%) were on lexis, compared to 98 (55%) on phonetics and phonology, and 58 (33%) on morphology and syntax (see the Appendix).[2] Of the 10 articles on lexical variation, there were four articles on quotatives (D'Arcy, 2012; Denis et al., 2019; Labov, 2018; Louro et al., 2023), as well as articles on loan word development (Poplack & Dion, 2012), the grammaticalization of variants of FINISH in Auslan (Johnston et al., 2015), change over time in the function of *key* in English (De Smet, 2016), variation in adverbs of evidentiality in English (Tagliamonte & Smith, 2021), variation in lexis and other levels of linguistic analysis in Japanese dialects (Takano, 2021), and change from Old to Middle English in the use of third-person adult male referents (Stratton, 2023).

Underlying the general assumption that lexical variation is asystematic, there are three primary challenges, in our view, that have limited the study of lexis within the variationist paradigm.

The first challenge is methodological. Although collecting data through sociolinguistic interviews (Labov, 1972a, 1984) has been highly productive for the study of variation at other levels of linguistic analysis, ***sociolinguistic interviews do not generally allow for the systematic analysis of variation in vocabulary***. While acknowledging the high levels of interest in lexical variation and the centrality of the social function of lexis in communication, Labov himself acknowledged this methodological challenge (Labov, 1973:340) (see also Weiner & Labov, 1983):

> The reason for this neglect is certainly not a lack of interest, since linguists like any other speakers of a language cannot help focusing their attention on the word, which is the most central element in the social system of communication. It is the difficulty of the problem, and its inaccessibility to the most popular methods of inquiry, which is responsible for this neglect.

Compared to most speech sounds and many grammatical constructions, the vast majority of lexical items, particularly those in open-class parts-of-speech (i.e., content words), are far too uncommon to be studied in collections of sociolinguistic interviews or even traditional multi-million word corpora of natural language. Typical sociolinguistic datasets simply do not allow for most lexical items to be observed with sufficient frequency to allow for lexical variation to be analyzed in the quantitative

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

detail demanded in variationist sociolinguistics, precluding analysis of lexical variation at a systemic level. This issue can be overcome, in part, by analyzing relatively high-frequency words, such as intensifiers (e.g., Ito & Tagliamonte, 2003) or by conducting/analyzing a vast number of interviews (e.g., Tagliamonte & Jankowski, 2025). However, the latter is not feasible for most researchers who are investigating variation and change across the lexicon. Indeed, when Labov has engaged with lexical variation, he has done so not using the Labovian sociolinguistic interview (e.g., Labov, 1972a, 1984) but by engaging with alternative sources of data, such as questionnaires/direct questions (Labov, 1972b; Labov et al., 2006)[3] and advertisements (Labov, 2003).

The second challenge is also methodological. ***The analysis of lexical variation is often incompatible with standard approaches to the measurement of sociolinguistic variation*** (see Durkin, 2012; Labov, 1978; Lavandera, 1978). It can be especially difficult to reconcile lexical variation with the notion of the "sociolinguistic variable" as "alternative ways of 'saying the same thing'" (Labov, 1969:738) because all lexical items are ultimately meaningful, unlike phonetic and phonological variables. Every lexical item tends to encompass a range of senses and functions, each denoting a distinct set of possible meanings.[4] In other words, there are no true synonyms (Cruse, 1986; Leclercq & Morin, 2023). Furthermore, the full set of possible lexical items that can be used to express a given concept in a particular context can be enormous and thus difficult to enumerate (Kretzschmar, 2009), while many words do not appear to be in alternation with any other words (Grieve, 2023). For these reasons, when analyzing variation and change in lexis, especially compared to phonetics and phonology, it is generally much more difficult to meet what is known as the "principle of accountability" (Labov, 1972a), which insists that *all variants* of a sociolinguistic variable be considered.[5]

The third challenge is theoretical: ***observations of lexical variation are difficult to integrate into sociolinguistic theory.*** Lexical variation is often closely related to variation in the referential *meanings* words express, which is quite different from variation in phonetics and phonology, for example, where the alternation between variants can be divorced from referential meaning. Lexical variation is therefore seen as being distinct from structural variation—related to language content rather than language form—and consequently asystematic and therefore on the periphery, or even outside the scope, of variationist theory, as Labov claimed in the quote at the start of this paper (2012:17). There is, however, a lack of empirical evidence to support this theoretical assumption, given the inherent methodological challenges associated with the analysis of lexical variation, as discussed above, which make the systematic observation and measurement of lexical variation so difficult within the traditional variationist paradigm. In this case, it seems that methodological challenges are driving the development of theoretical assumptions and not the other way around.

In fact, there is increasing empirical evidence for both the systematicity of lexical variation and its relevance to our general understanding of language variation and change. While there has always been some research on lexical variation in sociolinguistics (e.g., Labov, 1972b; Osser & Endler, 1970), over the past two decades there has been a dramatic increase in the amount of research on lexical variation—not only in adjacent

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

fields, but on the periphery of variationist sociolinguistics, too. In general, we believe research on lexical variation has primarily come from five main sources, all of which were made possible, to an extent, by moving away from the sociolinguistic interview as the main source of data, often taking advantage of new methodological opportunities for the collection of language data, especially by compiling increasingly large corpora of social media and by delivering surveys and other tasks via crowdsourcing platforms to large numbers of informants online.

First, research in ***historical sociolinguistics***, which is also generally conducted within the variationist paradigm, has considered change over time in word alternations, primarily focusing on function word alternations (e.g., Nevalainen & Raumolin-Brunberg, 2016; Romaine, 1982), although several studies in this tradition have focused on the use of high frequency content words, as well (e.g., D'Arcy, 2017; Stratton, 2022). As noted above, two exceptions that have been published in this journal are De Smet's (2016) analysis of change in the use of the word *key* from a noun into adjective in English between the 1970s and 1990s, and most recently Stratton's (2023) analysis of change over time in the use of the nouns *man* and *wer* to refer to human males in English over 600 years. Notably, research in historical sociolinguistics has generally been based on corpora of historical written documents, although, because these corpora are still relatively small, analyses must generally focus on high frequency lexical items.

Second, over the past decade, there has been a rapid increase in research in ***computational sociolinguistics*** in both sociolinguistics (Grieve et al., 2023) and natural language processing ([NLP] Nguyen et al., 2016). These studies are characterized by the analysis of very large datasets collected online, including corpora of social media (Eisenstein et al., 2014) and crowd-sourced surveys (Britain et al., 2020). Lexis has received more attention than any other level of linguistic analysis in this research, in part because big data generally facilitates the systematic observation of variation in a far larger sample of words than is otherwise possible. In most cases, this research has focused on the analysis of individual words, as opposed to lexical alterations. Notably, in NLP, much of this research has been pursued with the goal of increasing fairness in artificial intelligence systems, a topic that is now becoming especially important with the rise of large language models.

Third, research on the ***sociolinguistics of sign languages*** has naturally placed great emphasis on lexical variation, as this level of analysis is especially prominent in sign languages (Lucas et al., 2001; Ma, 2020; Mudd et al., 2020). For example, Johnston et al. (2015), which was published in this journal, found evidence for grammaticalization of signs used to mean FINISH in Auslan, based on an analysis of variation in the grammatical function, the semantics, and the formational features of this sign in a corpus of elicited conversational data. Notably, given challenges in creating large-scale datasets of sign languages (Fenlon & Hochgesang, 2022), rather than collecting data through relatively informal sociolinguistic interviews, sociolinguists who study lexis in sign languages may follow more formal protocols, often involving the direct elicitation of forms of interest in relatively controlled conditions suitable for video recording.

Fourth, research has been conducted at the intersection of variationist sociolinguistics and cognitive linguistics. Research in ***cognitive sociolinguistics*** (Geeraerts et al., 2010; Kristiansen & Dirven, 2008; Kristiansen et al., 2022) has

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

often focused on the systematic, structural analysis of lexical variation (e.g., Geeraerts et al., 2024; Robinson, 2010). For example, Geeraerts (1997) outlined the role of prototypicality in processes of diachronic lexical change. Research in cognitive sociolinguistics also tends to draw on increasingly large corpora of natural language harvested online for the analysis of lexical variation (e.g., Geeraerts et al., 2024; Gries, 2013).

Finally, a growing body of lexical research has been conducted broadly within ***variationist sociolinguistics***. Much of this research has investigated variation on the edge of grammar and lexis based on the analysis of high frequency words, generally based on relatively large collections of sociolinguistic interviews collected over time. In many cases, these studies have focused on function word alternations (e.g., Cheshire & Fox, 2009; Tagliamonte & Baayen, 2012), as function words are both sufficiently frequent to be observed in relatively small datasets and are generally understood to involve the expression of grammatical information, as opposed to content word alternations. Researchers in this tradition have also used similar approaches to study a small number of relatively highly frequent content word alternations in corpora, including quotatives (e.g., Buchstaller et al., 2010; D'Arcy, 2012; Denis et al., 2019), adjectives of strangeness (Tagliamonte & Brooke, 2014), words for an evening meal (Jankowski & Tagliamonte, 2019), third-person male referents (Franco & Tagliamonte, 2021), and large clusters of lexical items (Poplack et al., 1988). Other researchers have utilized non-vernacular oriented methods such as questionnaires (e.g., Labov, 1972b; Osser & Endler, 1970) or repurposed dialectological survey data (e.g., Burkette, 2012, 2017; Johnson, 1993; Montemagni & Wieling, 2016) to highlight the social distribution of lexical items according to dimensions such as age, socioeconomic class, and gender.

Developing from within this tradition, third-wave sociolinguists have also often taken lexical variation into consideration (e.g., Bucholtz, 2009; Denis, 2021; Kiesling, 2004), in part due to the emphasis on understanding the expression of identity through language variation. Speakers undeniably express social meaning through word choice, which is often highly socially salient. Third-wave sociolinguistics has also challenged the principle of accountability, with Campbell-Kibler (2011) arguing that it is the individual variant, as opposed to the set of variants that constitutes a sociolinguistic variable, that carries social meaning. Notably, research in this tradition is increasingly based on new approaches to fieldwork, including moving online, where the expression of social meaning is of central importance. For example, among many other features, Ilbury (2020) showed how certain lexical features originating from African American Language (e.g., *bae, yass*) are used by gay British men on Twitter to index specific social identities.

Building on recent research on lexical variation, the goal of this paper is therefore to affirm the general importance of this research to our understanding of language variation and change by challenging the notion that lexical variation is unsystematic and therefore less of a priority for variationist sociolinguistics. To support this claim, we consider three empirical studies that show how lexical variation is richly patterned in a wide range of ways. These studies exemplify three of the traditions we have identified above—third-wave sociolinguistics, sign language sociolinguistics, and computational sociolinguistics. In each case, we discuss the primary results of these studies and how they can inform sociolinguistic theory in unique and important ways.

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

We also stress how these studies have each adopted distinctive techniques for data collection and data analysis to overcome the methodological challenges that have previously limited research in this area. We then conclude this paper by claiming that lexical variation is not only systematically patterned, but that these patterns allow us to observe language variation and change from a unique and important perspective that must ultimately be accounted for by any comprehensive variationist theory, even if they challenge long held assumptions about the machinery of language variation and change.

Study 1: Anglo-Cornish dialect lexis

Our first study focuses on lexical variation in Anglo-Cornish, the variety of English spoken in Cornwall in the far southwest of Great Britain (Sandow, 2021, 2022, 2023, 2024a). In Sandow (2024a), the four lexical variables—TOURIST, LUNCH BOX, WOMAN, and WALK—explored in this study were analyzed in the aggregate, highlighting systematic sociolinguistic patterns, such as associations of Anglo-Cornish variants with strong local identity and an inverted style pattern. The case-study that we present here focuses primarily on the first of these variables, TOURIST, as well as related processes of semantic change.

Methodologically, this study is based on a novel combination of tasks to elicit lexical data from 80 speakers. Spot-the-difference tasks (see Fig. 1) were employed to elicit a relatively casual speech style, while naming-tasks (see Fig. 2) elicited a relatively careful speech style. In the former, participants focused on task-completion, where task-completion was conditional on identifying and lexicalizing differences. In the latter, participants were informed that their word-usage was being observed, shifting their focus to their lexical choices. Accounting for stylistic variation provides insight into the perceptions of these forms, as well as broader insight into language ideology (Sandow, 2022). This combinatory method of eliciting lexical items at varying levels of attention means that we can observe lexical variation in the context of the attention-to-speech (Labov, 1972a) model which has seldom been applied to words. In addition, to elicit information on semantic variation and change, Robinson's (2010) "who/what is [polysemous word]" elicitation tasks were employed. By asking for a referent, rather than a meaning, these tasks are less direct than asking for meanings explicitly. Here, semantic variants are defined as the same way of saying different things (see Robinson, 2010). These tasks ensure that it is clear where particular variants could occur, thus circumscribing the variable context and satisfying Labov's (1972a) principle of accountability.

In this first study, the analysis of lexical variation provides insights into social attitudes at a more granular level than is generally possible through the analysis of other levels of linguistic analysis, making it possible to track the history of sociocultural changes through lexical variation. This duality of sociocultural and lexical change is evidenced in Cornwall by shifting attitudes toward tourism and the words that are used for tourists. As we see, accounting for changes pertaining to tourism and tourists in Cornwall alongside changes in the semantic domain of tourism in Cornwall allows us to better understand the actuation of these changes. This work builds on existing research that highlights the role of social, cultural, and attitudinal influences on the

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press



Figure 1. An example of the spot-the-difference elicitation tasks (from Sandow, 2021).



Figure 2. An example of a naming-task (from Sandow, 2021).

socially mediated incrementation of lexical change (e.g., Labov, 1972b; Tagliamonte & Brooke, 2014; Tagliamonte & Jankowski, 2025).

Tourism in Cornwall largely began at the turn of the 19th to 20th centuries, but there was a notable quantitative upturn in visitor numbers in the post-war years and then again in the 1960s (see Meethan, 2002:24). This nascent tourist industry, which now accounts for a greater proportion of Cornish economic output than any comparable United Kingdom region (Office for National Statistics, 2016), has had wide-reaching effects, including on the Cornish economy and language. Since the mid/late-mid 20th century, negative attitudes to tourism, particularly the ways in which it was perceived to be diluting traditional Cornish identity, have been evident (see discussion in Perry, 1999).

In Cornwall, the Old English word *emmet* 'ant' has been retained but, with the rise in tourism, since at least the 1970s, it has also been used to mean 'tourist.' This polysemous state of affairs was induced by pejorative parallels being drawn between arthropods and the tourists who were increasingly visiting Cornwall in the post-war years and who were perceived to be red (due to sunburn on [overwhelmingly] white skin), to move in large groups, and to be seen only in fair weather. The semantic shift from 'ant' to 'tourist' can be seen in apparent-time, with the 'ant' sense being retained most by

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

older middle-class speakers with a strong sense of local identity (for further detail, see Sandow, 2023). The shift away from this 'ant' sense, toward 'tourist,' is predicated upon a less than favorable attitude toward tourists, which can also be observed in the frequent collocates of *emmet* 'tourist,' such as *fucking* and *bloody*.

While many, particularly younger, Cornish people are keen to distance themselves from such discourses, which "other" tourists and non-Cornish people more broadly, it is perhaps surprising that there is no quantitative difference in the overall rates of use of *emmet* as a variant of TOURIST across apparent time. However, there are differences in the social factors that best account for variation among older and younger speakers. Within the category of older speakers, strength of local identity was the best predictor of the use of *emmet*, while socioeconomic class was a stronger predictor for younger speakers, with *emmet* being favored by those with a strong local identity in the older group and working-class speakers in the younger group.

This generational change must be understood within the context of broader changes in Cornish identity that are underway (see Sandow, 2024b). Older Cornish people primarily orient toward a type of Cornish identity that has been described as the "Industrial Celt" (Sandow, 2024b:201), characterized by an orientation toward Cornwall's industrial and Celtic histories and a conceptualization of the Cornish people as Celtic and, consequently, non-English. By contrast, many, particularly younger, Cornish people orient to a "Lifestyle Cornwall" (Sandow, 2024b:201) version of Cornish identity which is predicated on an orientation to Cornish aesthetics, including its landscapes and seascapes, and conceptualizing Cornwall as a county of England, albeit a distinct one. This latter identity is also characterized by a mock-othering of the English while accepting, even embracing, tourism in Cornwall.

The specific indexicalities for *emmet* are clearly expressed in the following quote from a young man from Cornwall, highlighting the broader awareness of social meaning at the level of lexical variation found in this community:

> *Emmet* is a word that is only likely to be used by Cornish people […] it is a pejorative word for tourists […] the sort of person who is likely to use this word is someone with an inherent dislike of outsiders and who feels threatened by influxes of tourism […] I think that if there was a class thing to it, it would probably be used more by working-class Cornish people. Gender, I don't think it matters nor does age, I think it's more of an identity thing, a national identity […] I associate this word with a slightly insular quality based on a dislike of others, in this case, tourists. It is reflective of a traditional attitude. I think people use it to advertise their Cornishness as a sort of Cornish badge of honour to dislike tourism. I don't find it problematic but I know lots of people who do.

The majority of participants agreed on some core social meanings of *emmet*: it indexes Cornishness, frustration, localness, and working-classness. However, there was also considerable variation in the identities and stances associated with *emmet*. Industrial Celts perceived *emmet* to index solidarity, yet this same usage was perceived by those who orient toward a Lifestyle Cornwall identity as being bigoted, hostile, insular, and anti-tourist. While those who orient to Lifestyle Cornwall acknowledged such social meanings, many do use this word, often commenting that when they do so the butt

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

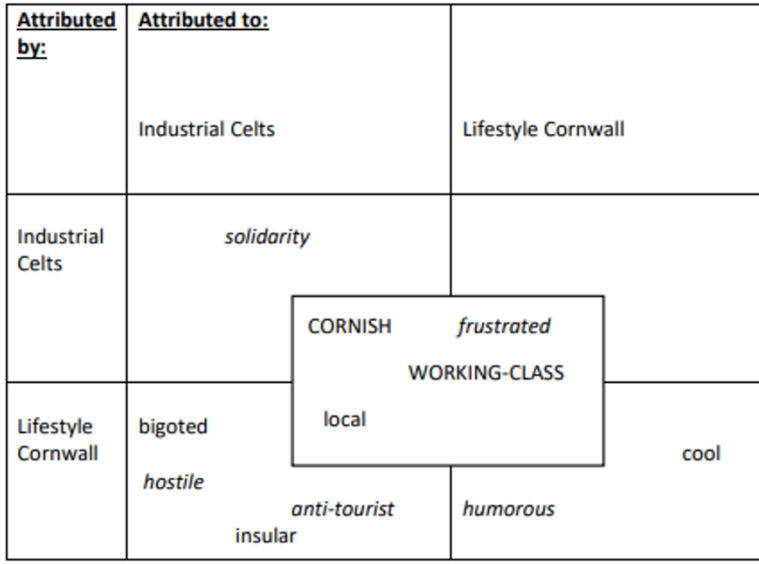


Figure 3. Indexical field for *emmet* 'tourist.' Capital letters denote social groups. Regular text denotes persona types. Italics denote stances (from Sandow, 2021).[6]

of the joke are not tourists but other Cornish people who they perceive to be bigoted. Those who orient toward Lifestyle Cornwall are therefore likely to use *emmet* to index humorousness and a cool identity with their (Cornish and non-Cornish) peers. Thus, while many Industrial Celts use *emmet* to increase social distance between themselves and non-Cornish people, many who orient toward the Lifestyle Cornwall identity use the same word but with the inverse pragmatic function: to decrease the social distance between themselves and non-Cornish folk.

Overall, meta-linguistic commentaries from speakers therefore show that *emmet* has a rich and nuanced indexical field in Cornwall, as presented in Figure 3. Crucially, the indexical field highlights the ways in which the referential meaning of lexical items can bleed into the social meaning. This contrasts with the meaning of, for example, phonological variants, in which the relationship between the phonological form and the social meaning is generally assumed to be entirely abstract (although, see Eckert, 2017). The use of *emmet* to style a Cornish identity is evident from an inverted style pattern (cf. Labov, 1972a): it is more frequent in careful (naming-task) as opposed to casual (spot-the-different task) speech. This highlights the way in which, while *emmet* is not typical of "vernacular" speech in Cornwall, it is utilized to do social identity work. This stylistic pattern, which challenges classic sociolinguistic theories of style, is shared with the three other variables analyzed in this study, highlighting a systematic nature to this cluster of lexical items (see Sandow, 2022).

Ultimately, the sociocultural changes pertaining to tourists are also seen in the semantic field of tourism in Cornwall. Tourism in Cornwall can be reduced to three key stages: (1) an increase in tourism in the post-World War II years, (2) the subsequent development of negative affective stances toward tourism, and (3) more recently, a shift

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

toward an acceptance of tourism. These stages each have a lexical analogue: (1) the semantic shift from *emmet* 'ant' to *emmet* 'tourist,' (2) the use of *emmet* to "other" tourists, and (3) the pragmatic shift in the function of *emmet* to mock the "other." The lexical-social analogues provide explanation for why this change happened in this place, providing an explanation for the actuation of the semantic shift from 'ant' to 'tourist,' as well as the shift in its social meaning.

The combination of the socially-mediated incrementation and the degree of transparency and non-arbitrariness of the actuation of the changes presented in this case-study highlights the valuable insights lexis can facilitate in understanding sociocultural and attitudinal changes and vice versa. Indeed, it highlights the way in which lexical change provides a window into society. The findings also speak to the ways in which semantic change can drive changes in patterns of lexical alternation. These insights extend our understanding of processes of sociolinguistic variation and change which cannot be observed so clearly when focusing on other levels of linguistic analysis.

Study 2: Evidence of dialect leveling in British Sign Language

Our second study focuses on lexical variation in British Sign Language (BSL). Variation may occur in the sublexical structure of signs: signs consist of a combination of distinctive handshapes, locations on the signer's body or in the space around the signer, and movements (e.g., Johnston & Schembri, 2007). For example, studies of American Sign Language, Auslan in Australia, and New Zealand Sign Language (NZSL) have shown that the location of signs specified in citation form for a forehead location can vary, produced on locations lower on the head or body (Lucas et al., 2001; Schembri et al., 2009). These patterns are partly driven by assimilation effects in connected signing: if other signs are produced at lower locations before or after the target sign, the target's location also tends to be lower. This variation, however, also correlates with social factors, including gender, age, and region. For example, in Auslan, lowering was seen more often in younger women from urban centers (Schembri et al., 2009). Variation may also occur in the grammatical organization of sign languages. For example, this includes variable presence of subject noun phrases. Like many spoken languages, most clauses in Auslan and NZSL narratives exhibit subject drop (McKee et al., 2011), but this is conditioned by several linguistic factors, such as whether the referent of the noun phrase is the same as in the previous discourse, or whether there is evidence of English influence on other aspects of the clause structure.

Just like in spoken languages, variation and change in signed languages can thus be identified at all levels of analysis. Although there is interest in variation across all linguistic levels in sign languages, lexis is especially important, as our case-study illustrates. Before introducing this study, however, there are two important differences between lexical variation in spoken and signed languages that should be highlighted. The first relates to the origins of lexical variation in sign languages in the Global North, where schools for deaf children have been found to be the origin of regional variation and the key site for the transmission of sign language variants (Eichmann & Rosenstock, 2014; Quinn, 2010; Stamp et al., 2014). Because the majority of deaf children are born to hearing families (Mitchell & Karchmer, 2004), schools for the deaf,

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

which were often residential in the 19th and early 20th centuries, served as "speech communities" where deaf children learned and passed on sign languages, often from older to younger children outside of the classroom, as the language of instruction was often the majority spoken language of the region. It is also possible that minimal interaction between children at different schools allowed sign language variants to remain distinct from one another. When deaf children left school, the use of these school variants therefore became associated with the surrounding region (Sutton-Spence & Woll, 1999). In fact, Stamp et al. (2014) described the use of regional variants in BSL as "school-lects" (for more discussion, see Quinn, 2010), reflecting the fact that schools may be a better predictor of regional variation than other social factors.

The second important difference is that regional variation has been found to occur primarily at a lexical level. Studies show that lexical variation in sign languages is systematically conditioned by a range of social factors, including age (Lucas et al., 2001), gender (Mudd et al., 2020), region (Ma, 2020), and ethnicity (McCaskill et al., 2011). In contrast, studies on regional variation in phonology are fewer in number (Bayley et al., 2000; Lucas et al., 2001), reflecting the fact that in some cases no consistent patterns of regional sociophonetic variation have been identified (Hamilton & Hochgesang, 2017). An implication of these studies is that lexical variation in BSL may index regional and national identities in a similar way as accent in British English. In fact, BSL signers often refer to this lexical variation as "accent" (Rowley & Cormier, 2021:928), which may be partly motivated by a recognition of its role indexing identity. We feel, however, that the primacy of lexical, compared to phonetic, variation in marking regional identities in signing communities is not as well known in the wider sociolinguistics literature as it ought to be.

Against this background, our second study relates to work on lexical variation and change in BSL (Stamp et al., 2014). This quantitative investigation into BSL regional lexical variation drew on the BSL Corpus dataset (Schembri et al., 2013). As part of the BSL Corpus, 249 deaf signers of BSL were filmed in eight sites across the United Kingdom: five cities in England (London, Bristol, Manchester, Newcastle, and Birmingham), as well as cities in Wales (Cardiff), Scotland (Glasgow), and Northern Ireland (Belfast). These sites were identified based both on documented lexical variation in BSL (Brien, 1992) and discussions with deaf community consultants and fieldworkers as part of the BSL Corpus Project.

The data used in this study were collected by eliciting from all participants signs for 102 concepts, using slides with an illustration and a prompt in the form of an English word equivalent (see Fig. 4). Participants were asked to produce the sign that they preferred, or felt that they most often used, for this concept. Comparison with participants' usage in conversation revealed that 78% of lexical variants used in the lexical elicitation task were also found in spontaneous dyadic conversational data, suggesting that the variants produced through elicitation were an accurate representation of their actual lexical preferences.

Variation was analyzed in a subset of 41 lexical items, including sign variants for colors ('brown,' 'green,' 'grey,' 'purple,' 'yellow'), names of countries ('America,' 'Britain,' 'China,' 'France,' 'Germany,' 'India,' 'Ireland,' 'Italy'), number signs from one to 20, and the United Kingdom place-names for the sites included in the study ('Belfast,' 'Birmingham,' 'Bristol,' 'Cardiff,' 'Glasgow,' 'London,' 'Manchester,'

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

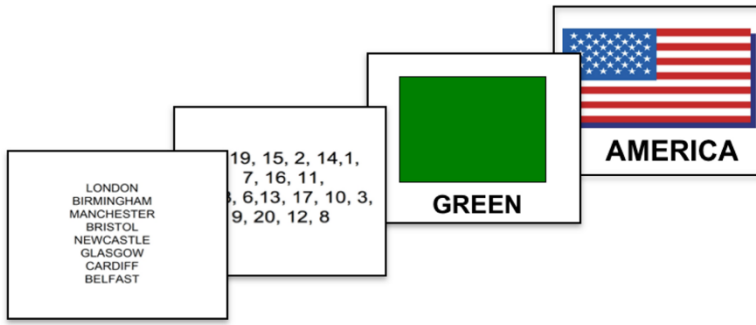


Figure 4. Examples of the elicitation stimuli used as part of the BSL corpus (from Stamp et al., 2014).

'Newcastle'). Responses were grouped into signs that were "traditional" for the signer's region and "non-traditional." Quantitative analysis found that age, school location, and language background were significant predictors of lexical variation: older signers, those who attended a local school, and those with deaf parents produced more traditional regional signs. Furthermore, results suggested that dialect leveling was taking place, with a decrease in the use of traditional, regional variants by younger signers and movement toward the use of supraregional variants (particularly from southern England), although this was clearer with particular sets of signs, like those for numbers, while others, like color signs, appeared to be changing more slowly. A comprehension task further corroborated the direction of change, showing that younger signers were more familiar with supraregional variants or those variants associated with the London area.

The United Kingdom place-name data were also analyzed, suggesting that place-name signs may work to index local, in-group, versus non-local, out-group, identities. Based on an analysis of 1,992 tokens, the study found the use of place-name variants differed significantly depending on whether the signer was a resident in that location, with the exception of signs for 'Glasgow,' 'London,' and 'Manchester.' Residents of Newcastle, Belfast, Bristol, Birmingham, and Cardiff were found to strongly favor the use of a local variant (or "endonym") that differed from signs from outside the community. The endonym was often a reduced fingerspelled form, such as 'M' and 'C' for 'Manchester,' while the exonym (the name used by individuals from outside that city) was often a calque—that is, a literal translation of the equivalent English word. For example, one exonym variant for 'Manchester' consists of a compound of individual signs MAN and CHEST. This result confirmed ethnographic observations that residents of a city use a different sign variant for their city's name than non-residents.

Overall, the findings from this study show that variation in BSL lexis is a systematic phenomenon: place-names, as well as regional color and number signs, clearly work to index local identities. Systematic change is also taking place in the BSL deaf community with a decline in lexical variation and a shift toward lexical variants with a wider currency (sometimes associated with southern England). This might reflect societal changes, including the closure of centralized schools for deaf children, the possible increased use of southern English variants in the media, and the increase in mobility

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

and contact between signers from different regions. In the case of signs for the names of countries, this change reflects international language contact, with BSL signers increasingly shifting from traditional BSL signs to borrowing signs from other sign languages (e.g., the American Sign Language sign for 'America' has largely replaced traditional BSL signs). Furthermore, these findings support the claim that there is no consistent evidence for patterns of regional sociophonetic variation in sign languages (Hamilton & Hochgesang, 2017) and that lexical variation is perceived by signers as equivalent to sociophonetic variation in spoken languages in terms of marking regional identities (Rowley & Cormier, 2021), implying that cross-modal differences may exist in this domain. Therefore, these findings are crucial for informing sociolinguistic theory in general, which should be derived from evidence in both spoken and signed languages.

Crucially, findings like these can only be obtained by examining variation at the lexical level in many sign languages. Because there have been no studies as yet demonstrating consistent patterns of regional variation at the phonological or morphosyntactic levels, by overlooking lexical studies, we inadvertently push aside a significant proportion of sociolinguistic research on sign languages. In doing so, the field runs the risk of missing out on crucial cross-modal generalizations. Studies, as aforementioned, point to the fact that there is only clear evidence of regional identities realized through lexis in sign languages; this lack of differentiation at other levels might represent a cross-modal difference, or it may relate to the comparatively short histories of many sign languages (i.e., there has not yet been time for regional sociophonetic patterns to emerge). Answering this question is an important extension of variationist sociolinguistics into the areas of cross-modal typology: is the lack of regional sociophonetic distinctions in BSL and other sign languages a modality difference or a language-age issue? This represents a new challenge for variationist theory, and one that we can only begin to address by comparing lexical variation across signing communities of different ages, work which has only just begun (Lutzenberger et al., 2023). Thus, studies of lexical variation may play a key role in understanding language emergence and evolution (Lutzenberger et al., 2021). This is a field of study which only the relatively recent emergence of many sign languages makes possible, since emerging sign languages may represent some of the "youngest" languages on the planet. For this reason, it is crucial to continue investigations cross-modally to understand to what extent differences are modality dependent or not.

Study 3: Diffusion of lexical innovation in American English online

Our third study focuses on tracking the diffusion of lexical innovation based on a large corpus of American social media. Despite being central to the variationist endeavor (Weinreich et al., 1968), the diffusion of linguistic innovations at scale, especially above the levels of phonetics and phonology, has largely resisted analysis using more conventional approaches to sociolinguistic inquiry. In general, analyzing the spread of innovations in natural discourse requires access to large amounts of language data because innovations are necessarily rare, at least initially, generally occurring far less than once per million words, requiring datasets consisting of hundreds of millions of words to observe sufficient instances of these words for meaningful generalizations to be made. Collections of sociolinguistic interviews do not typically allow for sufficient

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

data to be observed, especially over large numbers of time points or informants. Even traditional multi-million-word corpora of natural language are often far too small to reliably observe lexical innovations. Although surveys allow for rare forms to be elicited directly, they must be known beforehand, making diffusion difficult to study, especially in its early stages. The focus of variationist research on phonetic and phonological variation, which is generally more amenable to traditional approaches to data collection, as opposed to grammar and lexis, has also limited the study of diffusion, given the limited number of speech sounds and relatively slow nature of sound change. The spread of grammatical innovation also appears to often be relatively slow. As a consequence, we have not been able to describe in detail how the usage of large numbers of innovations has changed over time, precluding generalizing. Alternatively, lexical innovation can be a much more active phenomenon, as new words and signs are constantly being introduced, as catalogued by lexicographers and as analyzed by linguists interested in identifying common word formation processes (Bauer, 1983; Miller, 2014).

Although it has traditionally been difficult to build datasets suitably large for studying lexical innovation[7] at scale, this situation has recently changed with the growing availability of very large and highly stratified corpora, especially from social media, leading to the development of computational sociolinguistics (Grieve et al., 2023; Nguyen et al., 2016). Notably, while these new corpora have thus far only allowed for linguistic variation to be analyzed in a small number of very specific online varieties of language, they have greatly facilitated the analysis of the diffusion of lexical innovation within these varieties at scale, in real time, and across huge numbers of language users engaged in casual interaction, with a level of temporal and geographical resolution that has not previously been possible (e.g., Eisenstein et al., 2014; Stewart & Eisenstein, 2018; Würschinger, 2021). In this section, we therefore consider a series of studies that analyzed lexical innovation in American English based on the analysis of billions of words of geolocated social media data collected from across the contiguous United States from 2013 to 2016 (Grieve, 2018; Grieve et al., 2017, 2018), which clearly illustrate how working with very large corpora of natural language allows for the study of lexical innovation to be pursued at a scale that has not previously been possible.

In the first study (Grieve et al., 2017), a general method for identifying "emerging words" in a large time-stamped corpus was introduced. The application of this method was illustrated through an analysis of a multi-billion-word corpus of American tweets collected between 2013 and 2014, identifying word forms that were very rare at the start of this period but whose relative frequency increased rapidly over time on a day-by-day basis, excluding proper nouns and words found in standard dictionaries. Through this process, 29 unique emerging words were identified, many of which can be considered to be "slang," and many of which are now, a decade later, fairly well established (e.g., *baeless, boolin, famo, (on) fleek, faved, gainz, rekt*).

By analyzing the properties of these words, generalizations about the systematic nature of lexical emergence in this variety were also uncovered. For example, although s-shaped curves are commonly found in historical sociolinguistics over much longer time-spans (e.g., Nevalainen & Raumolin-Brunberg, 2016; Stratton, 2023), in this case, it was found that the relative frequencies of individual emerging words in this variety tend to show an s-shaped curve over real time, even though the corpus spans little more than a year. It was also found that shifts in meanings of some of these words could be

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

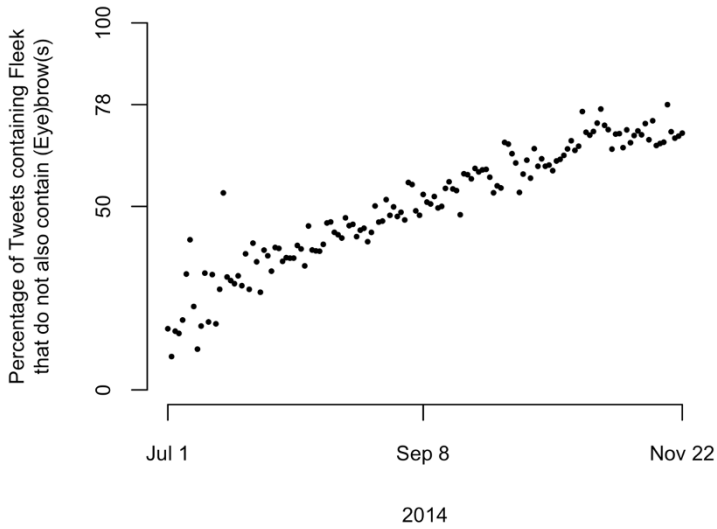


Figure 5. Change in the meaning of *(on) fleek* (from Grieve et al., 2017:120).

observed over real time, illustrating how new words can quickly undergo semantic change during the early stages of diffusion, perhaps helping to propel their spread. For example, it was found that although the term *on fleek* was originally used only to refer to carefully manicured eyebrows, its meaning quickly generalized over just a few months, as it started to be used to refer to various other well-presented things (see Fig. 5).

This study was also able to quantify the rate at which different word formation processes were used in this corpus, showing, for example, that compounding was most common, whereas true coinages were uncommon. Crucially, none of these insights are accessible through conventional approaches to data collection in sociolinguistics, illustrating the theoretical value of analyzing lexical innovations over very short periods of time, as made possible through large-scale corpus-based research. For example, as the case of *on fleek* demonstrates, understanding the semantics of a word can be directly relevant to understanding the mechanisms through which language changes.

In the second study (Grieve et al., 2018), an even larger set of 54 emerging words was extracted from the same corpus and the diffusion of these lexical items was tracked over time and across the approximately 3,000 counties in the contiguous United States, taking advantage of the precise geolocation of all tweets in the corpus. In addition to the lexical items included in the first study, new words included *balayage, litty, senpai, thottin, traphouse,* and *waifu*. Crucially, because many emerging words were analyzed, it was once again possible to draw generalizations about the nature of the geographical diffusion of lexical innovation in this variety. It was found that new words tended to spread out from urban centers, identifying five main regional hubs of lexical innovation, as illustrated in Figure 6, centered within five distinct American cultural regions. New words were also found to spread within these cultural regions before spreading to the rest of the United States, illustrating the importance of cultural patterns

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

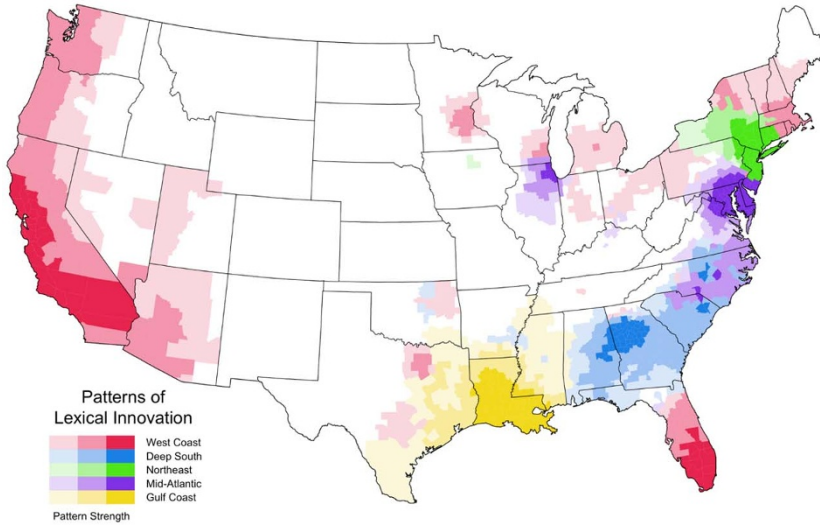


Figure 6. Hubs of lexical innovation in American English (US Twitter, 2013-2014) (from Grieve et al., 2018:307).

for constraining the geographical diffusion of lexical innovation, as opposed to simply geographical distance and population density, as had traditionally been theorized to predict patterns of diffusion (Labov, 2003; Trudgill, 1974). Finally, the importance of African American Language as a source of lexical innovation in modern American English was affirmed, with three of the five hubs notably being centered on different parts of the South with especially large Black populations.

Finally, in the third study (Grieve, 2018), these 54 emerging words were reanalyzed in a parallel multi-billion-word corpus of American Twitter data from 2016 to see how these words fared over time. For each word, the logged factor by which the relative frequency changed between 2014 and 2016 was calculated, and a linear model was used to predict this change based on four independent variables—word length, part-of-speech, word formation process, and whether or not the word expressed a new referential meaning (as opposed to being a synonym for an existing word). Overall, it was found that marking new meanings was the strongest predictor of lexical perseverance. Shorter words and words that were created using standard word formation processes were also found to be more likely to be retained over time. These results provide systematic evidence that new words that express new referential meaning are more likely to succeed, whereas words that are primarily used to express social meaning (i.e., that are synonyms for existing words) are less likely to succeed, perhaps because they tend to lose social meaning over time as they are adopted by a greater proportion of the population, demanding that more new words be introduced to fill this gap, creating a social meaning treadmill.

Overall, these studies clearly show the general theoretical value of analyzing lexis in very large corpora sampled densely over time and space to understand the diffusion of linguistic innovations. Crucially, many of the types of generalizations arrived

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

at in these studies can only be obtained when conducting aggregated analysis of many linguistic innovations in real time, which is much easier for lexis compared to sound change and grammatical change due to the pace and diversity of lexical innovation. By taking a lexical perspective, it was therefore possible to track the emergence of linguistic innovations in far greater detail than in the past, challenging the well-established wave and gravity models, which have been based primarily on theoretical assumptions and limited empirical data. Most notably, these studies have shown that the spread of new words is not only constrained by geography and population density, as predicted by the wave and gravity models, but by their cultural associations and communicative functions.

Conclusion

In this paper, we have argued that lexical variation should occupy a more central place in the variationist paradigm, despite having long been relegated to the fringes of the field. Given the recent surge of lexical studies, both in adjacent fields and within variationist sociolinguistics, we believe the need to incorporate lexis methodologically and epistemologically into our field is clearer and more important than ever. To support this claim, we have presented three sociolinguistic studies of lexical variation that reaffirm the value of sociolexical research, as is increasingly evident from the growing body of literature (e.g., see Sandow & Braber, in press). While our studies represent only three of the five modern research traditions that we have identified as pursuing research on lexical variation and change, our hope is that this paper will help to draw attention within the variationist community not only to variationist (specifically, third-wave) sociolinguistics, sign language sociolinguistics, and computational sociolinguistics, but to research being conducted in all five (additionally, cognitive sociolinguistics and historical linguistics) of these traditions and beyond. These case studies illustrate the broader body of research that highlights lexis as a key component in understanding the interface between language, culture, and society.

Notably, our three studies offer very different solutions to the two key methodological challenges that have limited lexical research in variationist sociolinguistics, as we outlined in the introduction to this paper. Furthermore, these studies have shown how, in various ways, the analysis of lexis can result in the identification of systematic patterns of lexical variation and change that can both extend and challenge theories of variationist sociolinguistics, which have generally been based on other levels of linguistic analysis. We believe these results, as well as the growing body of research on lexical variation and change in sociolinguistics and adjacent fields, therefore clearly show not only that the analysis of lexical variation is possible within the variationist paradigm, but that lexical variation is systematic and must therefore be accounted for by variationist theories if we are to have a complete understanding of the nature of language variation and change.

First, our studies illustrate different ways lexis can be systematically observed despite the inherent methodological challenges of studying linguistic forms that are generally incredibly infrequent, rarely occurring in natural discourse. This makes most forms of lexical variation difficult to observe via sociolinguistic interviews or other traditional approaches to data collection. Our first study shows how this issue can be overcome

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

through carefully designed task-oriented approaches to eliciting lexical variation. By developing bespoke elicitation procedures such as spot-the-difference tasks, we were able to identify complex sociolinguistic patterns in lexical variation, with clear links to the expression of sociolinguistic identity. Similarly, our second study shows how lexical variation in BSL can be observed through the use of a picture naming activity. Although any elicitation task in sign language is hindered by the fact that signers are videotaped, minimizing the spontaneity of the task and increasing the severity of the Observer's Paradox, task-oriented approaches have been developed that help to distract signers' attention from the presence of cameras (Stamp, 2016). However, it is notable that sign language researchers have problematized assumptions around the Observer's Paradox (see Schembri et al., 2013). Finally, our third study illustrates a complementary approach to analyzing lexis, overcoming the inherent challenge of studying very low frequency forms by working with extremely large corpora of natural language harvested online. Taking this approach created a unique opportunity to observe lexical variation and change in detail in natural language use while directly avoiding the Observer's Paradox.

Second, our three studies exemplify different ways that lexical variation can be meaningfully and rigorously measured. Because lexical items express referential meaning, measuring lexical variation within the standard sociolinguistic paradigm is challenging. While two word types are never perfectly synonymous across all tokens, either at the level of individuals or communities, actual tokens can be functionally interchangeable in context. As we exemplify in our first two case-studies, lexical variation can therefore be isolated through carefully constructed task-based elicitation methods, which allow us to track precisely where a variant could occur, thereby broadly satisfying the principle of accountability. Alternatively, in our third study, we eschewed the principle of accountability, as is common in corpus linguistics and many other fields, analyzing variation instead based on the relative frequencies of *individual* lexical types. This approach is necessary to obtain a complete picture of lexical variation and change, and by extension language variation and change more generally, because many words, including lexical innovations, are not meaningfully in alternation with any other existing forms.

Finally, our three studies clearly show that lexical variation not only is characterized by systematic social patterns, but that the analysis of lexical variation provides a unique perspective that extends and challenges existing theories of language variation and change. In our opinion, it is therefore not only possible to incorporate explanations for lexical variation into variationist theories but *necessary* if our goal is to develop robust and general theories that account for language variation and change across levels of linguistic analysis. In our first study, we see not only how lexical variation plays a critical role in the expression of sociolinguistic identity, but how the analysis of lexical variation provides a uniquely detailed perspective for understanding the complexity of the actuation of change and how closely it can be linked to the cultural context in which change occurs. In our second study, we see how lexical variation is an integral part of the variationist analysis of sign languages, where fundamental processes of language variation and change, such as the role of lexis in marking regional identity and the decrease in lexical variation from dialect leveling, appear to primarily occur at this level of linguistic analysis. To exclude lexical variation from sociolinguistic theory therefore

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

amounts to the de facto exclusion of key aspects of sign language research from sociolinguistic theory. Lastly, in our third study, we see how the aggregated analysis of lexis in very large corpora of natural language can allow for the spread of lexical innovations to be tracked in far greater resolution—given both the speed and volume of lexical change—than is possible at other levels of linguistic analysis. Most notably, this study shows how this process is driven, in part, by the meanings that these words express, as well as the cultural contexts in which these meanings are expressed. Furthermore, although each word has its own meaning and history, this study shows that generalizations about lexical innovation can still be discovered by analyzing many innovations in the aggregate, demonstrating that, in fact, change in one word can tell us much about change in another.

Most importantly, in all three studies, we clearly see that lexical variation and change is not only amenable to rigorous variationist analysis, but that lexical variation is systematic, with many lexical forms varying together in a structured way, driven by external factors that shape how language is used in society, despite widespread assumptions in our field. In this regard, each study provides novel insight into the mechanisms and processes of sociolinguistic variation and change. In study 1, we see that the actuation of lexical change was driven by important cultural changes in Cornish society, with changes in tourism and attitudes to tourism influencing the sociolinguistic patterns of lexicalization of the concept TOURIST. Similarly, in study 2, we see how regional variation in education systems has resulted in systematic patterns of lexical variation in sign language. Finally, in study 3, we see that when regional variation in large numbers of lexical innovations is analyzed in the aggregate, it is possible to identify a small number of common patterns of lexical innovation in American English that are driven by well-established patterns of cultural variation in the United States. In all cases, these patterns of lexical variation are systematic, reflecting the importance of general extra-linguistic sources of language change. It is therefore clear that the analysis of lexis is critical to fully understanding language variation and change in all its complexity. Further integrating a lexical perspective in variationist sociolinguistics, through more widespread adoption of lexis-oriented methods, can complement existing knowledge at other levels of language structure and facilitate a more holistic perspective in the pursuit of understanding the processes and mechanisms of language variation and change.

Finally, the analysis of lexical variation makes it clear that variation in linguistic structure is generally linked not only to the expression of *social meaning* but to the expression of *referential meaning*. We see that this link between form and meaning, which has largely been excluded from variationist sociolinguistics for methodological and theoretical reasons, is actually an important driver of language variation and change. When we analyze lexical variation, this is a fact that we must acknowledge. Our claim is, therefore, not simply that lexical variation should be integrated into variationist sociolinguistics, but that the analysis of lexical variation *must* be integrated into variationist theory because it provides the basis for investigating aspects of the actuation and diffusion of language change that are not otherwise generally accessible when analyzing phonological or grammatical variation. Our three studies offer clear evidence of this claim: to obtain a complete picture of how people express social meaning in the real world, how sign languages emerge and change over time, and how linguistic innovations diffuse across a population, it is not only possible but *necessary*

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

to study lexis. If variationist sociolinguistics does not at least attempt to account for variation at all levels of linguistic analysis, we cannot claim to be pursuing a complete socio*linguistic* theory.

Acknowledgements. The authors would like to thank the reviewers of this article for their considered feedback, as well as the editors of *Language Variation and Change*. Any remaining errors are the responsibility of the named authors. We would also like to acknowledge the receipt of funding that made this work possible. Rhys Sandow was supported by the Arts and Humanities Research Council [AH/L503861/1]; Adam Schembri was supported by European Research Council Advanced Grant [ERC-ADG 885220 SignMorph]; and Jack Grieve was supported by the AHRC, ESRC, and Jisc (grant reference number 3154) as part of Round 3 of the Digging into Data Challenge.

Data availability statement. The data that support the three studies presented in this article can be made available upon request to the authors.

Notes

1. Here and throughout, we do not wish to make a sharp distinction between dialectology and variationist sociolinguistics, as these boundaries are often blurred, as exemplified by study 3.
2. We use the example of this journal to exemplify broader trends in the field, not to suggest anything in particular regarding this journal or its editors. As this is the only purely variationist journal, it is the only one through which such trends in variationist objects of study can be attested.
3. For example, "what's the general term you use for a carbonated beverage in your area" (Labov et al., 2006:289).
4. Many sociolinguists have identified socially variable patterns with more flexible requirements of referential equivalence, for example, see Sankoff and Pierrette (1981) on "weak complementarity" and Thibault (1991) on semantic overlaps (for more recent applications, see Stratton, 2022; Brook, in press, respectively).
5. Considerable research on grammatical variation also studies language variation and change without strictly conforming to the definition of the sociolinguistic variable, both by relaxing the condition of equivalency (e.g., Tagliamonte & D'Arcy, 2009; Weiner & Labov, 1983) and by ignoring the principle altogether, measuring the relative frequency of individual grammatical forms measured relative to the total number of words in the corpus, without reference to other variants (e.g., Biber, 1988; Szmrecsanti, 2013).
6. This visualization distinguishes between the social meanings that are attributed to and attributed by those who embody the Industrial Celt and Lifestyle Cornwall identities. For example, those who align with Lifestyle Cornwall may interpret the use of *emmet* 'tourist' to be "bigoted" when used by someone they perceive to be an Industrial Celt, yet "humorous" when used by a speaker who aligns with Lifestyle Cornwall. The square in the middle contains indexicalities that are broadly shared across the community.
7. We use the term "lexical innovation" to refer to words that (i) occur very infrequently in the corpus at the start of the period being represented, (ii) increase in frequency rapidly over time, and (iii) did not occur in standard American dictionaries at the time (although many have since been added).

References

Adams, Michael. (2014). Vocabulary analysis in sociolinguistic research. In J. Holmes & K. Hazen (eds.), *Research methods in sociolinguistics*. Oxford: Blackwell. 163-193.
Bauer, Laurie. (1983). *English word-formation*. Cambridge: Cambridge University Press.
Bayley, Robert, Lucas, Ceil, & Rose, Mary. (2000). Variation in American Sign Language: The case of DEAF. *Journal of Sociolinguistics* 4(1):81-107. https://doi.org/10.1111/1467-9481.00104
Biber, Douglas. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
Brien, David. (1992). *Dictionary of British Sign Language/English*. London: Faber and Faber.
Britain, David, Blaxter, Tamsin, & Leemann, Adrian. (2020). East Anglian English in the English Dialects app. *English Today* 36(3):14-30. https://doi.org/10.1017/S0266078420000206
Brook, Marisa. (in press). The origins of pretend like: A syntactic-semantic puzzle in American English and beyond. To appear in *American Speech*.
Bucholtz, Mary. (2009). From stance to style: Gender, interaction, and indexicality in Mexican immigrant youth slang. In A. Jaffe (ed.), *Stance: Sociolinguistic perspectives*. Oxford: Oxford University Press. 146-170.

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

Buchstaller, Isabelle, Rickford, John R., Traugott, Elizabeth C., Wasow, Thomas, & Zwicky, Arnold. (2010). The sociolinguistics of a short-lived innovation: Tracing the development of quotative all across spoken and internet newsgroup data. *Language Variation and Change* 22(2):191-219. https://doi.org/10.1017/S0954394510000098

Burkette, Alison. (2012). The lion, the witch, and the *armoire*: Lexical variation in case furniture items. *American Speech* 84(3):315-339. https://doi.org/10.1215/00031283-2009-024

Burkette, Alison. (2017). To start a fire: The material effect on vocabulary. *American Speech* 92(3):321-339. https://doi.org/10.1215/00031283-4312075

Campbell-Kibler, Kathryn. (2011). The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change* 22(3):423-441. https://doi.org/10.1017/S0954394510000177

Cheshire, Jenny & Fox, Sue. (2009). *Was/were* variation: A perspective from London. *Language Variation and Change* 21(1):1-38. https://doi.org/10.1017/S0954394509000015

Cruse, Alan. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.

D'Arcy, Alexandra. (2012). The diachrony of quotation: Evidence from New Zealand English. *Language Variation and Change* 24(3):343-369. https://doi.org/10.1017/S0954394512000166

D'Arcy, Alexandra. (2017). *Discourse-pragmatic variation in context: 800 years of LIKE*. Amsterdam: Benjamins.

De Smet, Hendrik. (2016). How gradual change progresses: The interaction between convention and innovation. *Language Variation and Change* 28(1):83-102. https://doi.org/10.1017/S0954394515000186

Denis, Derek. (2021). Raptors vs. bucktees: The Somali influence on Toronto slang. *Journal of Multilingual and Multicultural Development* 42(6):565-578. https://doi.org/10.1080/01434632.2021.1895181

Denis, Derek, Gardner, Matt H., Brook, Marissa, & Tagliamonte, Sali A. (2019). Peaks and arrowheads in vernacular reorganization. *Language Variation and Change* 31(1):43-67. https://doi.org/10.1017/S095439451900005X

Durkin, Philip. (2012). Variation in the lexicon: The 'Cinderella' of sociolinguistics? *English Today* 28(4):3-9. https://doi.org/10.1017/S0266078412000375

Eckert, Penelope. (2017). Comment: The most perfect of signs: Iconicity in variation. *Linguistics* 55(5):1197-1207. https://doi.org/10.1515/ling-2017-0025

Eichmann, Hanna & Rosenstock, Rachel. (2014). Regional variation in German Sign Language: The role of schools (re-)visited. *Sign Language Studies* 14(2):175-202. https://doi.org/10.1353/sls.2014.0001

Eisenstein, Jacob, O'Connor, Brendan, Smith, Noah A., & Xing, Eric P. (2014). Diffusion of lexical change in social media. *PLOS ONE* 9:e113114.

Fenlon, J. & Hochgesang, Julie A. (eds.) (2022). *Signed language corpora*. Washington, DC: Gallaudet University Press.

Franco, Karlien & Tagliamonte, Sali A. (2021). Interesting fellow or tough old bird? Third-person male referents in Ontario. *American Speech* 96(2):192-216. https://doi.org/10.1215/00031283-8661833

Geeraerts, Dirk. (1997). *Diachronic prototype semantics*. Oxford: Clarendon Press.

Geeraerts, Dirk, Kristiansen, Gitte, & Peirsman, Yves (eds.) (2010). *Advances in cognitive sociolinguistics*. Berlin: Mouton.

Geeraerts, Dirk, Speelman, Dirk, Heylen, Kris, Montes, Marianna, De Pascale, Stefano, Franco, Karlien, & Lang, Michael. (2024). *Lexical variation and change: A distributional semantic approach*. Oxford: Oxford University Press.

Gries, Stefan Th. (2013). Sources of variability relevant to the cognitive sociolinguist, and corpus- as well as psycholinguistic methods and notions to handle them. *Journal of Pragmatics* 52:5-16. https://doi.org/10.1016/j.pragma.2012.12.011

Grieve, Jack. (2018). Natural selection in the Modern English lexicon. In *Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)*, Torun, Poland.

Grieve, Jack. (2023). Register variation explains stylometric authorship analysis. *Corpus Linguistics and Linguistic Theory* 19(1):47-77. https://doi.org/10.1515/cllt-2022-0040

Grieve, Jack, Hovy, Dirk, Kendall, Tyler, Nguyen, Dong, Stanford, James N., & Summer, Meghan. (2023). Computational sociolinguistics. *Frontiers in Artificial Intelligence*. https://doi.org/10.3389/978-2-8325-1760-4

Grieve, Jack, Nini, Andrea, & Gou, Diansheng. (2017). Analysing lexical emergence in Modern English online. *English Language & Linguistics* 21(1):99-127. https://doi.org/10.1017/S1360674316000113

Grieve, Jack, Nini, Andrea, & Gou, Diansheng. (2018). Mapping lexical variation on American social media. *Journal of English Linguistics* 46(4):293-319. https://doi.org/10.1177/0075424218793191

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

Hamilton, Heather & Hochgesang, Julie A. (2017). Is the Philadelphia American Sign Language (ASL) variety a real thing? Perceived lexical frequency and comprehension of ASL signs by Philadelphian and non-Philadelphian ASL users. Poster presented at 9th International Corpus Linguistics Conference, Birmingham, UK, July 24-28.

Ilbury, Christian. (2020). "Sassy Queens": Stylistic orthographic variation in Twitter and the enregisterment of AAVE. *Journal of Sociolinguistics* 24(2):245-264. https://doi.org/10.1111/josl.12366

Ito, Rika & Tagliamonte, Sali. (2003). *Well* weird, *right* dodgy, *very* strange, *really* cool: Layering and recycling in English intensifiers. *Language in Society* 32(2):257-279. https://doi.org/10.1017/S0047404503322055

Jankowski, Bridget L. & Tagliamonte, Sali A. (2019). *Supper* or *dinner*? Sociolinguistic variation in the meals of the day. *English World-Wide* 40(2):169-200. https://doi.org/10.1075/eww.00027.jan

Johnson, Ellen. (1993). The relationship between lexical variation and lexical change. *Language Variation and Change* 5(3):285-303. https://doi.org/10.1017/S0954394500001514

Johnston, Trevor, Cresdee, Donovan, Schembri, Adam, & Woll, Bencie. (2015). FINISH variation and grammaticalization in a signed language: How far down this well-trodden pathway is Auslan (Australian Sign Language)? *Language Variation and Change* 27(1):117-155. https://doi.org/10.1017/S0954394514000209

Johnston, Trevor & Schembri, Adam. (2007). *Australian Sign Language (Auslan): An introduction to sign language linguistics*. Cambridge: Cambridge University Press.

Kiesling, Scott. (2004). Dude. *American Speech* 79(3):281-305. https://doi.org/10.1215/00031283-79-3-281

Kretzschmar, William A. (2009). *The linguistics of speech*. Cambridge: Cambridge University Press.

Kristiansen, Gitt & Dirven, René (eds.) (2008). *Cognitive sociolinguistics: Language variation, cultural models, social systems*. Berlin: Mouton.

Kristiansen, Gitte, Franco, Karlien, De Pascale, Stefano, Rosseel, Laura, & Zhang, Weiwei (eds.) (2022). *Cognitive sociolinguistics revisited*. Berlin: Mouton.

Labov, William. (1969). Contraction, deletion and inherent variability in the English copula. *Language* 45(4):715-762. https://doi.org/10.2307/412333

Labov, William. (1972a). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. (1972b). The recent history of some dialect markers on the island of Martha's Vineyard, Massachusetts. In L. M. Davis (ed.), *Studies in linguistics in honor of Raven I. McDavid Jr*. Alabama: University of Alabama Press. 81-121.

Labov, William. (1973). The boundaries of words and their meanings. In C.-J. N. Bailey & R. W. Shuy (eds.), *New ways of analysing variation in English*. Washington: Georgetown University Press. 340-373.

Labov, William. (1978). Where does the sociolinguistic variable stop? A response to Beatrice Lavandera. *Working Papers in Sociolinguistics* 44:7-22.

Labov, William. (1984). Field methods of the project on linguistic change and variation. In J. Baugh & J. Sherzer (eds.), *Language in use: Readings in sociolinguistics*. Englewood Cliffs, NJ: Prentice Hall. 28-53.

Labov, William. (2003). Pursuing the cascade model. In D. Britain & J. Cheshire (eds.), *Social dialectology: In honour of Peter Trudgill*. Amsterdam: Benjamins. 9-22.

Labov, William. (2012). *Dialect diversity in America*. London: University of Virginia Press.

Labov, William. (2018). The role of the avant garde in linguistic diffusion. *Language Variation and Change* 30(1):1-21. https://doi.org/10.1017/S0954394518000042

Labov, William, Ash, Sharon, & Boberg, Charles. (2006). *The atlas of North American English: Phonetics, phonology and sound change*. Berlin: Mouton.

Lavandera, Beatriz R. (1978). Where does the sociolinguistic variable stop? *Language in Society* 7(2):171-182. https://doi.org/10.1017/S0047404500005510

Leclercq, Benoit & Morin, Cameron. (2023). No equivalence: A new principle of no synonymy. *Constructions* 15(1). https://doi.org/10.24338/cons-535

Louro, Celeste R., Collard, Glenys D., Clews, Madeleine, & Gardner, Matt H. (2023). Quotation in earlier and contemporary Australian Aboriginal English. *Language Variation and Change* 35(2):129-152. https://doi.org/10.1017/S0954394523000169

Lucas, Ceil, Bayley, Robert, & Valli, Clayton. (2001). *Sociolinguistic variation in American Sign Language*. Washington, DC: Gallaudet University Press.

Lutzenberger, Hannah, de Vos, Connie, Crasborn, Onno, & Fikkert, Paula. (2021). Formal variation in the Kata Kolok lexicon. *Glossa* 6(1):119. https://doi.org/10.16995/glossa.5880

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

Lutzenberger, Hannah, Mudd, Katie, Stamp, Rose, & Schembri, Adam C. (2023). The social structure of signing communities and lexical variation: A cross-linguistic comparison of three unrelated sign languages. *Glossa* 8(1). https://doi.org/10.16995/glossa.10229

Ma, Yunyi. (2020). *A study of lexical variation, comprehension and language attitudes in Deaf users of Chinese Sign Language (CSL) from Beijing and Shanghai*. Doctoral dissertation, UCL (University College London, London).

McCaskill, Carolyn, Lucas, Ceil, Bayley, Robert, & Hill, Joseph. (2011). *The hidden treasure of Black ASL: Its history and structure*. Washington, DC: Gallaudet University Press.

McKee, Rachel, Schembri, Adam, McKee, David, & Johnston, Trevor. (2011). Variable "subject" presence in Australian Sign Language and New Zealand Sign Language. *Language Variation and Change* 23(3):375-398. https://doi.org/10.1017/S0954394511000123

Meethan, Kate. (2002). Selling the difference: Tourism marketing in Devon and Cornwall, South-west England. In R. Voase (ed.), *Tourism in Western Europe: A collection of case histories*. Oxon: CABI Publishing. 23-42.

Miller, Gary. (2014). *English lexicogenesis*. Oxford, UK: Oxford University Press.

Mitchell, Ross E. & Karchmer, Michael A. (2004). When parents are Deaf versus hard of hearing: Patterns of sign use and school placement of Deaf and hard-of-hearing children. *The Journal of Deaf Studies and Deaf Education* 9(2):133-152. https://doi.org/10.1093/deafed/enh017

Montemagni, Simonetta & Wieling, Martijn. (2016). Tracking linguistic features underlying lexical variation patterns: A case study on Tuscan dialects. In M.-H. Côté, R. Knooihuizen, & J. Nerbonne (eds.), *The future of dialects*. Berlin: Language Science Press. 117-135.

Mudd, Katie, Lutzenberger, Hannah, de Vos, Connie, Fikkert, Paula, Crasborn, Onno, & de Boer, Bart. (2020). The effect of sociolinguistic factors on variation in the Kata Kolok lexicon. *Asia-Pacific Language Variation* 6(1):53-88. https://doi.org/10.1075/aplv.19009.mud

Nagy, Naomi. (2011). Lexical change and language contact: Faetar in Italy and Canada. *Journal of Sociolinguistics* 15(3):366-382. https://doi.org/10.1111/j.1467-9841.2011.00488.x

Nevalainen, Terttu & Raumolin-Brunberg, Helena. (2016). *Historical sociolinguistics: Language change in Tudor and Stuart England* (2nd ed.). London: Routledge.

Nguyen, Dong, Doğruöz, A. Seza, Rosé, Carolyn P., & De Jong, Franciska. (2016). Computational sociolinguistics: A survey. *Computational Linguistics* 42(3):537-593. https://doi.org/10.1162/COLI_a_00258

Office for National Statistics. (2016). *The regional value of tourism in the UK: 2013*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/leisureandtourism/articles/theregionalvalueoftourismintheuk/2013. Accessed September 12, 2023.

Osser, Harry & Endler, Norman S. (1970). Lexical choice and social class. *Language and Speech* 13(4):254-261. https://doi.org/10.1177/002383097001300404

Perry, Ronald. (1999). The changing face of Celtic tourism in Cornwall, 1875-1975. *Cornish Studies* 7:94-106.

Poplack, Shana & Dion, Nathalie. (2012). Myths and facts about loanword development. *Language Variation and Change* 24(3):279-315. https://doi.org/10.1017/S095439451200018X

Poplack, Shana, Sankoff, David, & Miller, Christopher. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26(1):47-104. https://doi.org/10.1515/ling.1988.26.1.47

Quinn, Gary. (2010). Schoolization: An account of the origins of regional variation in British Sign Language. *Sign Language Studies* 10(4):476-501. https://doi.org/10.1353/sls.0.0056

Robinson, Justyna A. (2010). *Semantic variation and change in present-day English*. Doctoral dissertation, University of Sheffield.

Romaine, Suzanne. (1982). *Socio-historical linguistics*. Cambridge: Cambridge University Press.

Rowley, Katherine & Cormier, Kearsy. (2021). Accent or not? Language attitudes towards regional variation in British Sign Language. *Applied Linguistics Review* 14(4):919-943. https://doi.org/10.1515/applirev-2020-0144

Sandow, Rhys J. (2021). *Anglo-Cornish dialect lexis: Variation, change, and social meaning*. Doctoral dissertation, University of Sussex, Brighton, UK.

Sandow, Rhys J. (2022). Attention, identity and linguistic capital: Inverted style-shifting in Anglo-Cornish dialect lexis. *English Language & Linguistics* 26(4):677-695. https://doi.org/10.1017/S136067432100040X

Sandow, Rhys J. (2023). An emmet's tale: The duality of social and lexical change. *Language & Communication* 93:15-29. https://doi.org/10.1016/j.langcom.2023.08.002

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

Sandow, Rhys J. (2024a). The role of identity in the usage and recognition of Anglo-Cornish dialect lexis. *Journal of English Linguistics* 51(1):38-64. https://doi.org/10.1177/00754242231220343

Sandow, Rhys J. (2024b). Time, space, and identities of place: Chronotopic identities in Cornwall. *North American Journal of Celtic Studies* 8(2):201-225. https://doi.org/10.1353/cel.00006

Sandow, Rhys J. & Braber, Natalie (eds.) (in press). *Sociolinguistic approaches to lexical variation in English*. Oxford: Routledge.

Sankoff, David & Pierrette, Thibault. (1981). Weak complementarity: Tense and aspect in Montreal French. In B. B. Johns & D. R. Strong (eds.), *Syntactic change*. Ann Arbor: University of Michigan Press. 205-216

Schembri, Adam, Fenlon, Jordan, Rentelis, Ramas, Reynolds, Sally, & Cormier, Kearsy. (2013). Building the British Sign Language corpus. *Language Documentation and Conservation* 7:136-154.

Schembri, Adam, McKee, David, McKee, Rachel, Pivac, Sara, Johnston, Trevor, & Goswell, Della. (2009). Phonological variation and change in Australian and New Zealand Sign Languages: The location variable. *Language Variation and Change* 21(2):193-231. https://doi.org/10.1017/S0954394509990081

Stamp, Rose. (2016). Do signers understand regional varieties of a sign language? A lexical recognition experiment. *Journal of Deaf Studies and Deaf Education* 21(1):83-93. https://doi.org/10.1093/deafed/env044

Stamp, Rose, Schembri, Adam, Fenlon, Jordan, Rentelis, Ramas, Woll, Bencie, & Cormier, Kearsy. (2014). Lexical variation and change in British Sign Language. *PLOS ONE* 9(4):e94053. https://doi.org/10.1371/journal.pone.0094053

Stewart, Ian & Eisenstein, Jacob. (2018). Making "fetch" happen. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*. 4360-4370. Available at: https://aclanthology.org/D18-1467/

Stratton, James M. (2022). Tapping into German adjective variation: A variationist sociolinguistic approach. *Journal of Germanic Linguistics* 34(1):63-102. https://doi.org/10.1017/S1470542721000088

Stratton, James M. (2023). Where did *wer* go? Lexical variation and change in third-person male adult noun referents in Old and Middle English. *Language Variation and Change* 35(2):199-221. https://doi.org/10.1017/S0954394523000108

Sutton-Spence, Rachel & Woll, Bencie. (1999). *The linguistics of British Sign Language: An introduction*. Cambridge: Cambridge University Press.

Szmrecsanti, Benedikt. (2013). *Grammatical variation in British English dialects: A study in corpus-based dialectometry*. Cambridge: Cambridge University Press.

Tagliamonte, Sali A. & Baayen, R. Harald. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24(2):135-178. https://doi.org/10.1017/S0954394512000129

Tagliamonte, Sali. A. & Brooke, Julian. (2014). A weird (language) tale: Variation and change in the adjectives of strangeness. *American Speech* 89(1):4-41. https://doi.org/10.1215/00031283-2726386

Tagliamonte, Sali A. & D'Arcy, Alexanda. (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language* 85(1):58-108. https://doi.org/10.1353/lan.0.0084

Tagliamonte, Sali A. & Jankowski, Bridget L. (2025). Lexical variation of woods and bush in Ontario English. *World Englishes* 44(1-2):285-299. https://doi.org/10.1111/weng.12704

Tagliamonte, Sali A. & Smith, Jennifer. (2021). Obviously undergoing change: Adverbs of evidentiality across time and space. *Language Variation and Change* 33(1):81-105. https://doi.org/10.1017/S0954394520000216

Takano, Shoji. (2021). Lifespan "changes from above" in the standardization of Japanese regional dialects: Levels of grammar, lexical properties and community characteristics. *Language Variation and Change* 33(3):297-329. https://doi.org/10.1017/S0954394521000211

Thibault, Pierrette. (1991). Semantic overlaps of French modal expressions. *Language Variation and Change* 3(2):191-222. https://doi.org/10.1017/S0954394500000521

Trudgill, Peter. (1974). Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in Society* 3(2):215-246. https://doi.org/10.1017/S0047404500004358

Trudgill, Peter. (1986). *Dialects in contact*. Oxford: Blackwell.

Weiner, Judith E. & Labov, William. (1983). Constraints on the agentless passive. *Journal of Linguistics* 19(1):29-58. https://doi.org/10.1017/S0022226700007441

Weinreich, Uriel, Labov, William, & Herzog, Marvin. (1968). *Empirical foundations for a theory of language change*. Austin: University of Texas Press.

Würschinger, Quirin. (2021). Social networks of lexical innovation. Investigating the social dynamics of diffusion of neologisms on Twitter. *Frontiers in Artificial Intelligence* 4:648583. https://doi.org/10.3389/frai.2021.648583

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

Appendix

Research Articles by Topics in *Language Variation and Change* (2012-2023)

Volume	Issue	Year	Phonology/ phonetics	Morphology/ syntax	Pragmatics/ discourse	Lexis	Other/ unclear	Total
24	1	2012	4	1	0	0	0	5
24	2	2012	2	2	0	0	1	5
24	3	2012	1	2	0	2	0	5
25	1	2013	1	3	0	0	1	5
25	2	2013	2	2	1	0	0	5
25	3	2013	3	1	1	0	0	5
26	1	2014	3	2	0	0	0	5
26	2	2014	4	1	0	0	0	5
26	3	2014	1	3	0	0	1	5
27	1	2015	3	1	0	1	0	5
27	2	2015	2	2	0	0	0	4
27	3	2015	2	2	0	0	1	5
28	1	2016	3	2	0	0	0	5
28	2	2016	3	1	0	0	1	5
28	3	2016	2	2	0	1	0	5
29	1	2017	4	0	0	0	1	5
29	2	2017	2	2	1	0	0	5
29	3	2017	4	1	0	0	0	5
30	1	2018	2	2	0	1	0	5
30	2	2018	4	1	0	0	0	5
30	3	2018	4	1	0	0	0	5
31	1	2019	3	1	0	1	0	5
31	2	2019	3	2	0	0	0	5
31	3	2019	3	2	0	0	0	5
32	1	2020	4	0	0	0	1	5
32	2	2020	3	2	0	0	0	5
32	3	2020	2	2	0	0	0	4
33	1	2021	2	1	0	1	1	5
33	2	2021	4	1	0	0	0	5
33	3	2021	2	2	0	1	0	5
34	1	2022	2	3	0	0	0	5
34	2	2022	4	0	0	0	0	4
34	3	2022	4	1	0	0	0	5

(*Continued*)

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press

(*Continued.*)

Volume	Issue	Year	Phonology/ phonetics	Morphology/ syntax	Pragmatics/ discourse	Lexis	Other/ unclear	Total
35	1	2023	2	3	0	0	0	5
35	2	2023	2	1	0	2	0	5
35	3	2023	2	3	0	0	0	5
		Total	**98**	**58**	**3**	**10**	**8**	**177**
		Percent	**55.37**	**32.77**	**1.69**	**5.65**	**4.52**	–

Cite this article: Sandow, Rhys, Jack Grieve, Rose Stamp & Adam Schembri. (2025). Lexical variation and change: Integrating lexis into variationist sociolinguistics. *Language Variation and Change* **37**:293–318. https://doi.org/10.1017/S0954394525100628

https://doi.org/10.1017/S0954394525100628 Published online by Cambridge University Press