Please cite as: MacKenzie, Laurel, and Mary Robinson. 2025. Spelling out grammatical variation. In *English Sociosyntax: Theory, Evidence, Approaches*, ed. Daniel Duncan and Mary Robinson, 59–95. Berlin: De Gruyter Mouton.

Publisher's formatted version available on request.

Laurel MacKenzie and Mary Robinson

3. Spelling out grammatical variation

Laurel MacKenzie, New York University

Mary Robinson, Newcastle University

Abstract

This chapter revisits a long tradition of literature that proposes that grammatical variables are unlikely to be used for social functions, and submits these proposals to a rigorous test. We carry out a meta-study of two sociolinguistics journals, searching for evidence of grammatical variables carrying social significance. We follow previous work to objectively identify and subcategorize grammatical variables into two types. We find robust evidence that one of the two types can be socially meaningful in a variety of ways. Evidence for the other type is much harder to find. However, this absence of evidence is largely due to a dearth of targeted research on the question. This empirical gap opens up a new research program.

1 Introduction

It is abundantly clear that certain elements of language can be socially significant to their users. A long tradition of sociolinguistic work has explored various facets of this phenomenon, including the social attributes that linguistic elements can index (Eckert 2008), the development of social evaluation as a language change progresses (Labov 2001), and the possible independence of the variants of a variable where their social meanings are concerned (Campbell-Kibler 2011). However, despite decades of work on this topic, one question has not yet been satisfactorily answered: Are there linguistic constraints on which elements carry social significance? In other words, are all elements of language fair game for social functions? And if not — if some elements are less likely to be socially meaningful than others — then why?

In this chapter, we explore this question as it pertains to variation in morphology and syntax. After revisiting a long tradition of literature that proposes that these so-called "grammatical" variables are unlikely to be used for social functions (e.g. Weiner and Labov 1983; Cheshire 1999; Meyerhoff and Walker 2013; a.o.), we submit these proposals to a rigorous test. We carry out a meta-study of two sociolinguistics journals, searching for evidence of grammatical variables carrying social significance. This chapter thus provides the first attempt to systematically confirm or refute this long-held assumption by testing it on a large body of sociolinguistic research.

Our meta-study follows previous work (Mansfield, Leslie-O'Neill, and Li 2023) in distinguishing two types of grammatical variables. We find robust evidence that one of the two types can be socially meaningful in a variety of ways. Evidence for the other type is much harder to find. However, this absence of evidence is largely due to a dearth of targeted research on the question. This empirical gap opens up a new research program.

The chapter is structured as follows. Section 2 reviews sociolinguistic literature on grammatical variation. It introduces what we call the *Grammatical Invisibility Principle*: the proposal, put forward in decades' worth of sociolinguistic research, that grammatical variables are invisible to social factors in a way that phonetic and lexical variables are not. Section 3 breaks down the two components of this principle, identifying what counts as a grammatical variable, subdividing grammatical variables into two types, and identifying what counts as evidence of social significance. In that section, we provide two testable hypotheses about social significance, one for each type of grammatical variable. In Section 4, we test these hypotheses through a meta-study of two sociolinguistics journals, *Language Variation and Change (LVC)* and *Journal of Sociolinguistics (JSlx)*, compiling evidence of social significance for each type of

variable and looking for patterns. Sections 5 and 6 summarize, conclude, and set up a program for future work.

2 The existing literature on the social significance of grammatical variables

A long tradition of sociolinguistic work interrogates whether sociolinguistic variation manifests in morphology and syntax in the same way that it has been shown to in phonetics and lexis. Much of this work is summarized by Levon and Buchstaller (2015); we review and critique their summary here.

Before evaluating their summary, it is important to note a terminological inconsistency that arises throughout this body of literature. It is clear in all cases that researchers are talking about variables that incorporate units larger than a single phoneme, and that frequently implicate the structural relations between words. That is, the variables in question are more than what are typically called "lexical" variables, i.e. variation in content words, such as the choice between *soda, coke,* or *pop* for a carbonated beverage. However, the term used to capture the family of variables under discussion differs from researcher to researcher, and includes *syntactic, morphosyntactic, morpholexical,* and *grammatical*. Though there are independent diagnostics from linguistic theory that can be used to assess whether a particular variable has its locus in the morphology or in the syntax (depending, of course, on one's particular theory of morphology and syntax), many researchers in sociolinguistics have not invoked these (as pointed out by Romaine [1984] and Cheshire [1987]). In fact, it is not uncommon to find sociolinguists describing as "syntactic" variables that have been argued in the syntax literature to be morphological in nature (such as English relative pronouns and English negative concord, two variables to which we return later in the chapter).

For this part of the chapter, we avoid this terminological slippage by adopting the term *grammatical variable* as a catch-all term for variables that are not obviously phonetic or lexical. We recognize the drawbacks of this term: it constitutes an oversimplification that lumps together variables that differ structurally, and it could be seen as vague, since variables at all levels of language, not simply the morphological and syntactic ones, are part of speakers' "grammars". Notwithstanding these issues, we choose it due to its longstanding use as a cover term for variables of this type (going back at least to Wolfram [1969: 52]) and to its common appearance in introductory-level presentations of these types of variables (e.g. Tagliamonte 2012: 206, 238; Wolfram and Schilling 2016: 79). Our intention at this point in the chapter is to diverge from work in sociolinguistics that has described variables with technical terms like *syntactic* absent an accompanying syntactic analysis. Instead, in this section, we provide an atheoretical overview of what sociolinguists have said about this class of non-phonetic, non-lexical variables. In Section 3.1, we make our terminology more precise to allow objective hypothesis testing.[1]

One major question running through several decades' worth of sociolinguistic literature is why sociolinguistic research has seemed to focus more on phonetic than on grammatical variables. Answers to this question have cited the methodological and conceptual difficulty of studying grammatical variables (Labov 1978; Lavandera 1978; Hudson 1980; Romaine 1984; Hudson 1996; Cheshire, Kerswill, and Williams 2005) and differences in the actuation of phonetic versus syntactic change (Cheshire, Kerswill, and Williams 2005: 136). But researchers have also proposed what we call the Grammatical Invisibility Principle: that grammatical variables are underrepresented in the sociolinguistic literature because they are somehow less

[1] The variables that sociolinguists call "phonetic" (or "phonological") are not necessarily a homogeneous group, either. For instance, Tamminga, MacKenzie, and Embick (2016) note the structurally ambiguous locus of apparently phoneme-level variables such as coronal stop deletion (*swep'* for *swept*) and *ing–in'* variation (*runnin'* for *running*). We follow previous literature in referring to "phonetic" variables in opposition to "grammatical" and "lexical" ones, while recognizing that this, too, is an oversimplification.

prone to social evaluation or social stratification than phonetic and lexical variables are (Lavandera 1978; Levon and Buchstaller 2015: 322).[2]

The Grammatical Invisibility Principle may be connected to variable frequency: some researchers hypothesize that the infrequency of grammatical variables in conversation makes them, in Levon and Buchstaller's (2015: 320) words, "less accessible for social and identity marking functions" (e.g. Hudson 1980; Rickford et al. 1995; Winford 1996; Cheshire 1999; Cheshire, Kerswill, and Williams 2005). As a result, these variables may display internal (linguistic) conditioning,[3] but not external (social) conditioning, i.e., social or stylistic stratification. It has also been proposed that grammatical variables are "abstract" in a way that phonetic and lexical variables are not (Weiner and Labov 1983; Labov 1993, 2001: 29; Meyerhoff 2001; Meyerhoff and Walker 2013; Levon and Buchstaller 2015), and that something about this abstractness prohibits language users from attaching social significance to the varying item or structure. This prohibition in turn keeps the variable from socially stratifying in the community.[4]

The Grammatical Invisibility Principle is sometimes stated as a categorical observation and sometimes as a gradient tendency. Labov (1993: 5) asserts categorically that "members of

---

[2] Other researchers (Meyerhoff 2001; Levon and Buchstaller 2015) have called this proposal the *Interface Principle*, after Labov (1993). However, Labov's (1993) Interface Principle asserts that not only grammatical, but also certain phonological variables will be socially invisible. It is a broad assertion, and its incorporation of phonological variables is often neglected when other sociolinguists invoke it. For instance, in Meyerhoff's (2001: 78) phrasing, the Interface Principle pertains to "structural (syntactic) variables": that is, Meyerhoff equates Labov's "structural" with "syntactic" when she rephrases the principle. Because this narrow interpretation of the Interface Principle is common in the sociolinguistic literature, it is this interpretation that we set out to test: the part that asserts that specifically grammatical variables are socially invisible. For precision, we avoid using the term *Interface Principle* and instead use the term *Grammatical Invisibility Principle*, which clarifies that we are testing the grammatical variation part of the principle. For an in-depth treatment of the part of the principle that pertains to phonological relationships, see Eckert and Labov (2017).

[3] While we set aside the internal conditioning of sociolinguistic variables in this chapter, see Childs (This Volume) for discussion of such effects.

[4] Why language users are purportedly unable to socially evaluate abstract structures has not been given a satisfying explanation by the researchers who have proposed it. It may have something to do with how the linguistic system interfaces with social cognition, but the details of that interface are not well understood (Campbell-Kibler 2016).

the speech community evaluate" some forms, "but not" others. Levon and Buchstaller (2015: 319), by contrast, offer the generalization that "phonetic and/or lexical features show greater degrees of social stratification than systemic features at a "deeper" level of linguistic structure, such as grammatical relationships". "Greater degrees" phrases the difference in a gradient way, though leaves the precise specification unclear: is the proposal that phonetic and/or lexical variables will show greater differences in usage between groups at opposite ends of the class spectrum than, say, grammatical variables?[5] Or that more phonetic and lexical variables will show any social stratification at all than grammatical ones? There are many ways in which these types of variables could differ from one another, and the lack of clarity on this point has made the proposals difficult to test.

As Levon and Buchstaller (2015) point out, despite researchers' intuitions that grammatical variables are underrepresented in the sociolinguistic literature, it is certainly not the case that these variables are entirely unrepresented. Levon and Buchstaller (2015) list several grammatical variables in English, French, and other languages that have been shown to show social and/or stylistic stratification (see also Cheshire, Kerswill, and Williams 2005). These counterexamples cast doubt that grammatical variables are any different sociolinguistically from lexical or phonetic ones. At the same time, as Levon and Buchstaller (2015) also note, these social correlates manifest in language production (e.g. social class groups using variants at different rates), but little research has been done to look for social effects in the *perception* of grammatical variables. Perhaps the Grammatical Invisibility Principle is a fact about language perception, then, and not about language production — though the necessary existence of a link

[5] If so, that would contradict a much-cited pattern, first identified by Wolfram (1969), that grammatical variables tend to show sharp social stratification and that phonetic variables tend to show gradient social stratification (see also Cheshire, Kerswill, and Williams 2005).

between perception and production at the speaker level (see e.g. D'Onofrio 2016) complicates this possibility.[6]

An additional complication to the Grammatical Invisibility Principle, also emphasized by Levon and Buchstaller (2015), is that sociolinguists often lack independent criteria to identify the architectural locus of a variable when the item that varies is larger than a single phoneme (see Cheshire 1987; MacKenzie 2012, 2013 for additional discussion along these lines). Perhaps the Grammatical Invisibility Principle holds of some subclasses of grammatical variables but not others. After all, as indicated earlier, "grammatical" is often used as a catch-all descriptor in sociolinguistic research when the precise locus of a variable cannot be identified. Given this, there may in fact be generalizations that can be drawn about which types of variables are less likely to be socially significant; however, without a consistent and objective way of classifying grammatical variables, such a generalization will go unnoticed.

A final, related concern raised by Levon and Buchstaller (2015) is that the discussion of these issues has at times become circular, with the presence or lack of social stratification of a variable used to diagnose its grammatical locus. So, for instance, a variable that previous researchers had described as morphosyntactic will be reanalyzed as a lexical variable when it is

[6] Also complicating this possibility is that Levon and Buchstaller (2015) themselves find clear evidence of social effects in the perception of a grammatical variable. However, these effects are restricted compared to the perception effects they find for a phonetic variable. Specifically, evaluation of the phonetic variable is effectively widespread across their speaker population, while evaluation of the grammatical variable is restricted to listeners from a particular region and those with a particular cognitive profile. They interpret this asymmetry as "tentative evidence for a weak formulation of the [Grammatical Invisibility] Principle that states that while listeners do [socially] evaluate abstract structural features, they do so in a comparatively more complex fashion than for phonetic ones" (Levon and Buchstaller 2015: 340).

However, there is an important confound between the two variables tested in Levon and Buchstaller's (2015) study: the phonetic variable is widespread throughout Britain (where subjects in their study were from), while the grammatical variable is "a relic form that remains fairly localized" to Northern varieties of British English (Levon and Buchstaller 2015: 326). Therefore, another interpretation of the region-specific perception they find is that listeners are less likely to have social associations with a variant that is not prevalent in their community. Moreover, as Levon and Buchstaller acknowledge (2015: fn. 11), their stimuli were recorded by a speaker with a Standard Southern British accent; her using a localized Northern dialect feature may have confused listeners, making the results of the study difficult to interpret. We thus do not see Levon and Buchstaller's (2015) findings as decisively supporting the assertion that listeners socially perceive grammatical and phonetic variables differently.

found to show social conditioning (Meyerhoff and Walker 2013: 409). This is another consequence of sociolinguists' lack of independent structural diagnostics of the locus of a variable.

The idea that some elements of language might behave differently from others with respect to social factors is an intuitively appealing one. Sociolinguists continue to draw generalizations about the differing social availability of different types of variables (e.g. Eckert 2019b). However, proposals that "abstract", "structural", or "syntactic" variables are unlikely or less likely to show social significance than other types of variables lack theoretical grounding. Researchers who put them forward tend to provide little discussion of the formal linguistic analysis that they draw on to support their designation of these variables as "syntactic" versus morphological, or "abstract" despite their evidence on the surface. It is also unclear whether different researchers are using these terms to mean the same thing; this in turn makes it difficult to support or refute others' claims. For this reason, we believe that it is important to start over with the Grammatical Invisibility Principle and to approach it more meticulously. There may be some truth to it, but in order to determine this, it needs to be made precise and testable.

We maintain that testing for potential sociolinguistic differences between different levels of grammar is an important enterprise that has implications for the nature of language change and the nature of the mental representation of language. On the nature of language change in the phonetic domain, Eckert (2019a: 1) takes a strong position on the role of social factors, saying "while I cannot say that sound change never progresses without taking on social meaning, I have never seen a contemporary example of one that did". Does the same hold true for the set of grammatical variables that are involved in diachronic change? We know from a large body of historical syntax work that syntax changes over time (e.g. Kroch 1989). From Eckert's statement,

we reach the conclusion that either syntactic changes are fundamentally different from phonetic ones, in that they proceed without taking on social significance, or social factors are an essential component of any change, and grammatical variables are just as likely to be subject to them as phonetic variables are. The first conclusion would provide support for the Grammatical Invisibility Principle; the second would refute it.

If it turns out that grammatical variation truly is sociolinguistically special compared to phonetic variation, many new questions arise. When grammatical variables are involved in diachronic change, how do they progress through a speech community without being embedded in a system of social meaning? Is there something cognitively different about grammatical variables as compared to phonetic ones that leads the two to be socially differentiated? More generally, to the extent that there are sociolinguistic differences between different elements of language, this suggests something important about how the linguistic system interfaces with systems of social cognition (Campbell-Kibler 2016). This in turn has implications for models of how language is mentally stored and dynamically produced (e.g. Tamminga, MacKenzie, and Embick 2016). We cannot answer these questions here, but we see our exploration of the social significance of grammatical variation as the beginning of an exciting conversation about these issues.

We know of only one study that has attempted to test questions related to the Grammatical Invisibility Principle in a rigorous way: Mansfield, Leslie-O'Neill, and Li (2023), which takes a different approach than ours. Rather than looking for grammatical variables and their social correlates in the sociolinguistic literature, Mansfield, Leslie-O'Neill, and Li identify grammatical variables in reference grammars of 42 typologically diverse languages. For each grammatical variable, they classify it as one of three types — variation in *form*, variation in

order, or variable *omission*. They then ask whether grammarians report it to differentiate regional dialects, and, if so, whether those regional dialects are in a relationship of low, medium, or high social distance from each other. They find, first of all, that "grammatical markers often differentiate dialects" (Mansfield, Leslie-O'Neill, and Li 2023: 263), but second of all, that the three types of grammatical variables do not differentiate dialects equally. Specifically, in their data, "dialect differentiation by form variables applies equally under close or distant social contact […] By contrast, order and omission variables are rarely dialectal in situations of close contact, but order variables become more likely to differentiate dialects as they become distant from one another" (Mansfield, Leslie-O'Neill, and Li 2023: 264). They conclude from this that "form variables are driven to a greater extent by social signaling, compared to order and omission variables", because "social signaling is only relevant to the extent that groups are in social contact" (Mansfield, Leslie-O'Neill, and Li 2023: 264). That is, non-form variables that differentiate regional dialects must have arisen for reasons other than social marking, because language users would have little social motivation to differentiate themselves from groups that they are not in close contact with. These results would seem to refute the Grammatical Invisibility Principle for variation in form, but support it for variation in order.

An advantage of Mansfield, Leslie-O'Neill, and Li's (2023) approach is its crosslinguistic diversity. They criticize the sociolinguistic literature for its "heavy focus on a small number of politically dominant, cosmopolitan languages" (Mansfield, Leslie-O'Neill, and Li 2023: 244), a point to which we return in Section 5. Drawbacks include its narrow focus on regional differentiation as the only social factor under consideration, and a lack of clarity regarding the evidence given to support this regional differentiation. We see our study as complementing

theirs: the social significance that we identify in synchronic language behavior may be what underlies the diachronic patterns of dialect differentiation they hypothesize.

3 Refining the Grammatical Invisibility Principle

In this section, we make previous researchers' framing of the Grammatical Invisibility Principle more precise by providing two additions: an objective classification scheme for so-called "grammatical variation" and a widely-cast net of what it means for a variable element of language to be socially significant.

3.1 Identifying and classifying grammatical variables

As we saw in Section 2, there are several proposals in the literature that so-called "grammatical" variables are somehow sociolinguistically different than phonetic and lexical variables. But what actually are grammatical variables linguistically?

Many introductory-level presentations of grammatical variables subdivide them into different types. For instance, Tagliamonte (2012: 206–207) differentiates variation in morphology — such as the variable realization of inflectional and derivational suffixes — from variation in syntax — which she says "typically involve[s] phenomena that distinguish languages from each other, particularly the ordering relationships between verbs and their complements, but also the behavior of clitics, agreement patterns, etc.". Wolfram and Schilling (2016) similarly differentiate morphological and syntactic variation, with morphological variation comprising variation in inflectional markers and pronouns, and syntactic variation comprising variation in verbal argument structure, agreement, and word order.

Already, however, we can find inconsistencies in these categorization schemes. Tagliamonte (2012: 207) refers to the variable realization of third person *-s* in English (as in *he come ~ he comes* and *they come ~ they comes*) as both "a tried and true morphological feature" and a "morpho-syntactic variable", though its involvement in the system of subject-verb agreement would seem to classify it as "syntactic" under her subcategorization scheme. English negative concord is another grammatical variable that sociolinguists have struggled to localize. A review of the literature finds negative concord described as "lexical" (Meyerhoff and Walker 2013: 409, following Labov 1993), "grammatical" (Tagliamonte 2012: 26), "structural" (Labov 2001: 29), "morpho-lexical" (Romaine 1984: 420), "morphosyntactic" (Smith 2007: 28), "syntactic" (Moore 2021: 73), and a "syntactic agreement relation" (Wolfram and Schilling 2016: 84).

This lack of terminological precision was identified as a problem as early as Cheshire (1987), but the situation has hardly improved in the decades since. However, one thing that has improved is the cross-pollination between sociolinguists and formal linguists. More sociolinguists are now using the tools of formal analysis to better understand the variables they study (see, for instance, the papers in Lignos, MacKenzie, and Tamminga [2018]). For this reason, we are now in a position to disagree with Cheshire (1987: 263) when she says that "there are not necessarily any clear grounds for distinguishing between phonological and morphophonemic variables on the one hand and between morphophonemic, morphosyntactic, and "purer" syntactic variables on the other". On the contrary, certain theories of morphology and syntax provide very clear grounds for ascertaining a variable's level of grammar.

Undertaking rigorous formal analysis of so-called "grammatical" variables is essential if we want to identify generalizations about the social significance of different types of variables,

and it will help us finally standardize the terminology across the literature. It has other benefits, too. Formal analysis of sociolinguistic variation has revealed that many variables' surface patterns conceal variation at multiple levels of grammar (Tamminga, MacKenzie, and Embick 2016). Most pertinently to our discussion here, this is the case for some forms of negative concord, which are best analyzed as comprising both variable morphology and variable syntax (Robinson and Thoms 2021). We can now start to ask whether variables with multiple loci of variation show comparable patterns of conditioning across their multiple loci, opening up new lines of research (see e.g. Tamminga 2016; MacKenzie 2020). Finally, a willingness to implement formal frameworks in sociolinguistic research opens the door for increased collaboration between sociolinguists and formal linguists, à la Cornips and Corrigan (2005) and Adger and Smith (2005). In sum, if sociolinguists were to more consistently undertake formal grammatical analyses of the variables they study, they would not only facilitate the drawing of generalizations about how different subclasses of variables behave, but they would also produce more well-rounded analyses by connecting different strands of literature that are not often brought together.

All this being said, at this stage in our project, we refrain from applying a particular formal treatment to the variables in our meta-study, though we return in Section 6 to a brief discussion of how we might do that at a later point. Our reasons for avoiding formal analysis include a desire not to alienate readers who might work in a different formal theoretical framework, and a lack of detail in many sociolinguistic treatments of grammatical variables that would allow us to carry out this type of analysis convincingly. But we see this kind of collaboration between sociolinguists and formal linguists as essential for future work.

For the time being, we follow Mansfield, Leslie-O'Neill, and Li (2023: 234) in defining grammatical variables as "grammatical meanings or functions that can be expressed in more than one way". We also adopt their subdivision of grammatical variables, with one change. Mansfield, Leslie-O'Neill, and Li (2023) subdivide grammatical variables into three categories:

> (a) Form**:** Variants have the same structure, but are distinguished by the form of a grammatical marker (either affix, clitic, or function word).
>
> (b) Order**:** Variants use the same lexical and grammatical elements, but are distinguished by linear ordering.
>
> (c) Omission**:** Variants are identical except that a grammatical marker is present in one but absent in the other. (Mansfield, Leslie-O'Neill, and Li 2023: 246–247)

We group together form and omission variables, seeing omission as simply a subtype of form: that is, omission is the alternation of a form with zero, as opposed to its alternation with another overt form. This decision is strengthened by Mansfield, Leslie-O'Neill, and Li's (2023) lack of a significant difference in social distance effects between form and omission variables, suggesting that the social patterning of the two is not meaningfully different in their data. This provides a binary categorization scheme for sociolinguists' grammatical variables: do they reflect the variable realization of grammatical markers (henceforth *realization*) or the variable order of linguistic elements (henceforth *order*)?[7]

Under the approach taken here, many variables that sociolinguists have previously named "grammatical" are reclassified as variation in realization. For instance, variable agreement, such as the British English Northern Subject Rule studied by Levon and Buchstaller (2015), or the variable choice of *was* or *were* (Adger and Smith 2005), constitutes variation in realization. So do English auxiliary contraction (Labov 1969; MacKenzie 2012, 2013), variation in English past

[7] We discuss cases that do not obviously fall into either category below.

tense formation (e.g. the variation between *dive* and *dove*; see Embick [2008]), and some forms of English negative concord (Robinson and Thoms 2021). By contrast, variables such as the English verb-particle alternation (Kroch and Small 1978), dative alternation (Tagliamonte 2014), and quantifier float (McCloskey 2000) represent variation in order.[8]

Applying this categorization scheme casts the literature summarized in Section 2 in a new light. All of the counterexamples of socially-significant grammatical variables that Levon and Buchstaller (2015) cite, including their own example of the Northern Subject Rule, are reclassified as variation in realization. If this holds up across a wider range of examples, the Grammatical Invisibility Principle would need to be refined to refer to variation in order only, not variation in realization. This, of course, would align with Mansfield, Leslie-O'Neill, and Li's (2023) findings. It additionally already has a precedent in the sociolinguistic literature. Romaine (1984) breaks down grammatical variables into two types. The first type, which she calls "morpho-syntactic or morpho-lexical", comprises those variables that reflect "the [variable] presence or absence of some linguistic item"; these are subject to social conditioning (Romaine 1984: 420). But the second type, which she calls ""pure" syntactic," and may involve word order variation (Romaine 1984: fn. 7), appears, in her estimation, to show no social or stylistic conditioning. This suggests that the difference between variation in realization and variation in order may indeed be sociolinguistically real, though we need a more comprehensive survey of the literature to say for sure.

For this chapter, we restrict our testing of the Grammatical Invisibility Principle to just these two types of variation only: variation in realization and variation in order. We exclude cases that do not fit neatly into either category, such as variation in English quotative verbs (*say,*

[8] Technically, the English dative alternation (e.g. *I gave her the book ~ I gave the book to her*) involves both variation in order and variation in realization, because the preposition *to* alternates with zero. We accordingly treat it as representing both categories and count it toward each.

go, be like, etc. [Tagliamonte 2012: 247]). From our perspective, this constitutes variation in choice of lexical item, even if in some cases the chosen lexical item brings along with it a syntactic difference. Its inclusion as a "grammatical" variable is arguable; indeed, quotatives are sometimes treated by sociolinguists as "discourse-pragmatic variables" instead (e.g. Tagliamonte 2012: 247). We hope future collaborative work between sociolinguists and syntacticians will more closely investigate the syntactic structure and sociolinguistic conditioning of these variables.

3.2 Social significance

In the previous subsection, we recategorized "grammatical variables" into two types. Here, we take a closer look at the social part of the Grammatical Invisibility Principle. As in Section 3.1, our intentions in this subsection are to problematize concepts that have been presented with minimal elaboration and to provide a framework to test prior claims of linguistic differences in social significance.

Evidence for the Grammatical Invisibility Principle is often taken from patterns of language production. For instance, Weiner and Labov (1983: 56) conclude that "social factors operate primarily on surface patterns rather than abstract syntactic alternatives" after finding that variation in speakers' use of the agentless passive versus the generalized active shows minimal social differentiation. By contrast, Labov's (1993) formulation of the principle is as a fact about sociolinguistic perception. For instance, in Levon and Buchstaller's (2015: 319) paraphrase, the Grammatical Invisibility Principle "posits that [grammatical] variation does not elicit the same kinds of perceptual reactions as phonetic variables". Grammatical variables are thought to be somehow too "abstract" for language users to attach social significance to them; the lack of social stratification in production is taken to be a consequence of that (Meyerhoff and Walker

2013). Also occasionally cited as evidence for the Grammatical Invisibility Principle is the lack of explicit metalinguistic commentary about grammatical variables: among the evidence that Labov (2001: 28–29) cites for "the isolation of abstract linguistic structures from social evaluation and differentiation" is the absence of overt commentary about these variables during metalinguistic discussions.

The three areas of evidence cited above — production patterns, perceptual reactions, overt metalinguistic commentary — correspond to Campbell-Kibler's (2016: 128) identification of three fundamental types of sociolinguistic abilities. Language users can "create sociolinguistic performances", that is, *produce* variants in a socially meaningful way. They can also extract social attributes of the speaker, situation, etc., from others' language productions, that is, *perceive* social meanings in others' use of linguistic forms (whether these meanings were intended by the other language user or not). Finally, language users can engage in metapragmatic or *metalinguistic behaviors*, including explicitly expressed attitudes, media representation of language, and stylized performance (Coupland and Jaworski 2004). From this, it follows that there are three different types of evidence for the social significance of a linguistic form: patterns of production, perceptual judgments, and metalinguistic behaviors. We suggest that the most comprehensive way to test for the Grammatical Invisibility Principle is to see whether our two types of grammatical variables are sensitive to social factors across these three domains of sociolinguistic abilities.

These three types of sociolinguistic abilities are naturally interconnected in actual language use. This is evident from the literature on the Grammatical Invisibility Principle, which has at times taken the lack of social patterns in production as a consequence of something being different about these variables in perception. If listeners cannot perceive the social meanings of

linguistic forms produced by others, this line of reasoning goes, those listeners cannot subsequently reproduce those meanings when they go to speak themselves. Patterns in production are also linked to metalinguistic commentary; one influential treatment of the link is Labov's (2001: 196) pathway from indicator to marker to stereotype, by which a variable's social stratification becomes more extreme as it moves from "zero degrees of social awareness" to the eventual "subject of overt comment". This being said, we will see in Section 4 that these three types of sociolinguistic abilities are often studied separately by sociolinguists. Few studies explicitly ask whether a given variable is sensitive to social factors in all three domains. For that reason, we keep the three domains of sociolinguistic ability separate, while recognizing that they are necessarily interconnected. It is our hope that the meta-study we report on in Section 4 prompts further research on the grammatical variables we identify that have not yet been considered from all three angles.

Additionally, we recognize that using the umbrella terms production, perception, and metalinguistic behaviors is an oversimplification. This is most evident in the area of production. Social patterns in production can take many forms and be understood in many ways. Classic sociolinguistic work, in what Eckert (2012: 88) calls the "first wave" tradition, sought to identify macrosocial patterns of variation in production, that is, population-level correlations between sociodemographic factors and linguistic forms. The variables that are held up as counterexamples to the Grammatical Invisibility Principle are frequently those that show macro-social patterns of this type (see Levon and Buchstaller 2015: 323 for a review). Research in the "third wave" tradition, however, puts the focus instead on the social meaning of variation, and argues that "variables connect only indirectly to the macrosocial" (Eckert 2016: 69). That is, a linguistic form may have social significance without necessarily showing macro-level

correlational production patterns. It may show other production patterns instead: for instance, correlations with persona, personality traits, stance, or affect (Eckert 2016, 2019b).

It is crucial to be aware of this range of meanings when testing for the Grammatical Invisibility Principle. Indeed, a weakness of past attempts to test or refute the Principle, argues Moore (2021), is a failure to recognize the varied types of social meanings a linguistic form can take on. Much of the work on the Grammatical Invisibility Principle has restricted the focus of study to the patterning of macro-level social factors like class, or formality as assessed on a vector of attention to speech. When a grammatical variable shows neither of these, researchers are often hasty to conclude that this reflects a deeper fact about the social significance of grammatical variation. Conversely, when a grammatical variable does show these types of patterns, researchers conclude that there is no truth to the principle. However, both of these social factors, Moore (2021: 56) points out, are reflective of codification and overt prescription processes. In her words, "it is possible that perceived differences between the social meanings of [phonetic] variation and [grammatical] variation [are] a consequence of standardization processes" having been applied to the particular grammatical variables that have been studied, rather than being due to "any inherent property" of grammatical variables themselves. There may still be social differences between different types of variables, but we need to adduce evidence from more variables than those that have been the subject of overt standardization processes, and we need to look for evidence of social significance in domains other than production patterns by social class and attention to speech. Our meta-study (Section 4) provides a first step in this direction.

In sum, it is our view that research on the Grammatical Invisibility Principle will benefit from additional nuance applied to the "social" side of the principle just as it will from additional

nuance applied to the "grammatical" side. Explicitly breaking down types of evidence for social significance into production patterns, perception patterns, and metalinguistic behaviors, as we do, is a start. Recognizing the range of social meanings that variables can take on — beyond formal or informal, prestigious or stigmatized — will add another important dimension on which we can look for generalizations.

Now that we have identified the different types of grammatical variables that there are, and the different types of social patterns they may show, we can put those two pieces together and ask whether different types of grammatical variables are more or less likely to show differences in social significance.

3.3 Putting it all together

The aim of this section has been to provide a framework in which to test the Grammatical Invisibility Principle, the proposal put forth by a number of researchers that grammatical variables are less likely to be socially significant than phonetic or lexical variables. In Section 3.1, we introduced a subcategorization scheme to reclassify "grammatical variables" as either variation in realization or variation in order. In Section 3.2, we identified three types of evidence of social significance: production of social patterns, perception of social meanings, and metalinguistic behaviors. Both of these categorization schemes bring necessary rigor to the task of identifying what counts as a grammatical variable, on the one hand, and what types of social significance such variables may show, on the other. They allow us to look for new, more complex generalizations about the kinds of social significance different types of variables may carry.

Putting these pieces together, we can now restate the Grammatical Invisibility Principle as a testable hypothesis. In fact, because we have argued that grammatical variation is actually a

cover term for two distinct types of variables, we need to expand it into two testable hypotheses, one addressing variation in realization and the other addressing variation in order. We also need to address the issue, raised in Section 2, that some researchers have framed their version of the principle as a categorical observation and others as a gradient tendency. Because a categorical statement is easier to disprove than a gradient tendency, which requires identifying what it means for something to be "more" or "less likely" than something else, we formulate our rephrasings as categorical hypotheses; however, we also discuss quantitative patterns when we present our results, in Section 4. Our rephrasings are in (1) and (2). We formulate them as affirmative, rather than negative, statements to avoid the complications that arise when trying to prove the non-existence of something.

(1) Social Realization Hypothesis: Variation in the form of a grammatical marker will bear social significance.

(2) Social Order Hypothesis: Variation in the linear ordering of linguistic elements will bear social significance.

Both hypotheses should be assessed, ideally, through research targeting social patterns in production, social evaluation in perception, and metalinguistic behaviors. Evidence from all three domains that supports hypotheses (1) and (2) will cast doubt on the proposal that there is any difference between variables at different levels of grammar with respect to how likely they are to be socially significant.

What if evidence supporting hypotheses (1) and/or (2) proves elusive? Absence of evidence is of course not evidence of absence. But the body of research demonstrating that variables at other levels of grammar — namely, phonetic and lexical ones — can take on social significance is sizeable, with evidence from the domains of production, perception, and

metalinguistic behavior easy to find (for instance, in the domain of perception alone: for phonetic variables, Fridland, Bartlett, and Kreuz 2004; Hay, Warren, and Drager 2006; Fridland 2008; Johnstone and Kiesling 2008; Walker 2007; Walker et al. 2014; for lexical variables, Dailey-O'Cain 2000; Violin-Wigent 2007; Hesson and Shellgren 2015; Beltrama and Staum Casasanto 2017; Maddeaux and Dinkin 2017). If a comparably-sized body of literature has looked for the social significance of realization and/or order variables but failed to find it, this raises the possibility that those variables may indeed be different in some way.

4 Testing the hypothesis with a meta-study

In the previous section, we established that a research program designed to test the Grammatical Invisibility Principle needs to look for evidence of the social significance of both realization and order variation. In this section, we look for that evidence. We review a selection of the sociolinguistic literature to assess whether there is evidence for social asymmetries between variables of these two subtypes.

Though meta-analyses are not common in sociolinguistics, they can be an excellent vehicle for testing sociolinguistic theories (Stausland Johnsen 2015). On the most basic level, our meta-study aims to assess whether either subtype of grammatical variable shows evidence of social significance, and, if so, whether that evidence is attested in all three domains of sociolinguistic knowledge. This creates a wider testing ground for the Grammatical Invisibility Principle than we have seen in previous work, which has almost exclusively focused on realization variables and their production patterns.

The meta-study can also speak not only to whether variables of our two different subtypes show social significance, but also how often. That is, we can calculate the rates at

which realization and order variables are found to show social significance in the meta-study corpus. This constitutes a first step toward addressing the "gradient" version of the Grammatical Invisibility Principle: that grammatical variables are "less likely" to be socially significant than variables of other types. Of course, thoroughly testing this gradient version of the principle requires comparing the rates at which different subtypes of grammatical variables are found to be socially significant to the rates at which (different subtypes of) phonetic and lexical variables are. Such a comparison is beyond the scope of our chapter, but we take the first step by computing the numbers for our two subtypes of variables.

Careful attention to the specific social details of the variables turned up by our meta-study will also allow us to identify commonalities among them. Specifically, we address proposals by Cheshire (1987, 1999, 2005) and Moore (2020, 2021) that those grammatical variables that are studied from a social perspective tend to be those for which one variant has been standardized, where the social significance takes the form of associations with overt prestige/stigma. This, in turn, can speak to Eckert's (2018: 190) contention that grammatical variables tend to have "quite fixed social meanings associated with external facts like class and particularly education". Thus, in addition to answering questions of whether and how often grammatical variables are socially significant, we contribute data on how that social significance manifests. Our meta-study provides a macro-level view to complement the careful micro-level work on these questions that has been done by researchers such as Cheshire and Moore.

Finally, beyond simply confirming or refuting others' claims concerning the social significance of grammatical variables, our meta-study has the additional goal of bringing to light variables that have been understudied in the literature. As others have observed, past work on the Grammatical Invisibility Principle has tended to draw repeatedly on the same set of variables,

often from English (Cheshire, Kerswill, and Williams 2005; Moore 2021). The meta-study uncovers several realization and order variables that have not been thoroughly studied from any social angle. Ascertaining the potential social significance of these lesser-known grammatical variables is a fruitful direction for future work.

In sum, our meta-study does two things: it provides large scale empirical data on a longstanding question in sociolinguistics, and it identifies understudied variables to be pursued in future research.

4.1 Methodology

Our meta-study analyzes two journals, *Language Variation and Change* (*LVC*) and *Journal of Sociolinguistics* (*JSlx*), from their first year of publication to the end of 2023 (comprising 495 articles for *LVC* and 581 articles for *JSlx*).[9] These journals are somewhat complementary in their focus: *LVC* does not require discussion of social factors while *JSlx* does.

While this testing ground naturally constitutes only a small fraction of the sociolinguistic literature, we maintain that it is nonetheless an improvement over the data that previous researchers have used to support or refute claims of the social invisibility of grammatical variables. Previous work on the Grammatical Invisibility Principle has tended to draw evidence from variables that have been well studied or are otherwise familiar to the researchers. In contrast, combining all issues of two long-lived journals gives us a representative sample of the sociolinguistic literature, and eliminates the possibility that our choice of examples will be influenced by confirmation bias. Still, we recognize that the publication process itself is not free of bias, with Western and Anglophone societies overrepresented in sociolinguistic research (Adli

[9] For *JSlx*, we counted articles published under the headings of Original Articles (called "Articles" in later issues) and Research Notes, but excluded articles published under the headings Editorial, Review Articles/Essay(s), Book Reviews, Dialogue, Introduction, Commentaries, Focus Issue, and Debate (these last four used for special issues).

and Guy 2022), not to mention the "file drawer problem" by which null results are unlikely to be published (Rosenthal 1979). We intend this to be merely the first step in a much larger research program. Testing the hypotheses in (1) and (2) will require dedicated follow-up work on particular variables and, ideally, an expansion of such work to understudied languages and speech communities. We outline some steps for future research in Sections 5 and 6.

To conduct the meta-study, we divided up the published studies between ourselves (though we consulted with each other when uncertainties arose). For every article, the responsible author examined it, identified the linguistic variable (or variables) under study, and determined whether the variable(s) fell into either (or both) of the two subcategories of grammatical variation identified in Section 3 (i.e., realization and order). Every variable that involved realization and/or order was logged in a database.[10] To keep the purview of the meta-study narrowed to only those variables of the type that sociolinguists have reliably called "grammatical", variables of the following types were not logged: phonetic/phonological (i.e. those that implicate the pronunciation of units smaller than a grammatical marker),[11] word choice (i.e. involving the phonological realization of non-grammatical markers, and including discourse-pragmatic variables [see Section 3.1]), discourse/conversation structure (i.e. the analysis of how interactions are structured), and code-switching (i.e. the variable choice of language by multilinguals). We counted only presentations of new research results and did not count reviews or references to work that had been published elsewhere. We also followed researchers' leads when identifying variables, a notoriously thorny point in the study of

[10] We have made the database publicly available at https://github.com/laurelmackenzie/grammatical-variation-metastudy.

[11] (ING) and (TD) were also excluded on these grounds. Tamminga (2016) makes a convincing case for these variables being multifactorial: surface variation in (ING) and (TD) stems from the application of a phonological variable layered over a morphological one. That makes these variables surface-ambiguous between having their locus in the phonology and having their locus in the realization of a grammatical marker. For this reason, we exclude these two variables from the meta-study.

grammatical variation (Rickford et al. 1991; MacKenzie 2012). To give an example, Britain (2002) studies past tense *BE* leveling to *was* in affirmative contexts separately from past tense *BE* leveling to *were* in negative contexts. We accordingly entered these in the database as two separate variables even though this contrasts with the approach of Schilling-Estes (2002), who treated *BE* leveling to *was* in affirmative contexts and to *were* in negative contexts as a single variable. We trusted the judgment of a paper's author(s) that the patterning of the variation in their data supported the analysis they chose.

Following the discussion in Section 3.2, we also coded each linguistic variable for social significance. This followed a tripartite scheme: no social significance investigated (i.e. social factors were not studied through novel analysis for the variable at all); social significance investigated but not found (i.e. researchers explicitly reported a null social result); social significance found. These latter two categories allowed for a further tripartite breakdown, for whether the social significance was studied via production patterns (i.e. covariation with social factors, such as class stratification), perception patterns (for instance, as ascertained by the matched-guise technique [Lambert et al. 1960]), and/or metalinguistic behaviors (e.g. explicitly expressed attitudes toward a form, use of a form in a stylized performance).

Finally, we coded each linguistic variable for two remaining factors to inform generalizations about which variables sociolinguists tend to select for research. The first factor is the language variety the variable was studied in, as reported by the author(s). The language variety was subsequently grouped into one of five categories: English, French, Spanish, Portuguese (if the authors used one of these languages when naming the variety they studied), or a residual "other" category. The second factor is whether the author(s) noted a variant of the

variable as being overtly prestigious or overtly stigmatized.[12] Contrasting with our coding protocol for social significance described in the previous paragraph, this judgment of overt prestige/stigma did take into account references to previous research. To exemplify: imagine a paper that analyzed only the linguistic constraints on some variable, but cited previous research showing that the variable bore overt prestige and had been found to show social stratification in production. This variable would be logged in our database and coded as "no social significance investigated" — because the social stratification in production did not come from novel analysis presented by the paper — but also coded as being relevant to the overt prestige/stigma axis. This coding procedure allows us to answer the question of whether studies of variation have tended to prioritize variables that are salient on this axis, even if those studies have not necessarily been designed to investigate social significance. It also allows us to ask, for the variables that qualify as grammatical under our approach, how often have researchers looked for their social significance, and how often have they found it?

4.2 Social significance of variation in realization

Our meta-study found 386 variables between the two journals that we classified as variation in realization (including variables classified as involving both realization and order simultaneously).[13] This comprises 247 from *LVC* and 139 from *JSlx*. (3) gives examples of realization variables.

(3) Examples of realization variables

[12] A variable qualified for this coding either if one variant was described by the author(s) as being overtly prestigious (e.g. prescribed by grammarians) or if one variant was described by the author(s) as being overtly stigmatized (e.g. prescribed against by grammarians and/or associated with explicitly negative metalinguistic commentary by users).
[13] This is the number of studied variables, not the number of articles examined in the meta-study. If one article looked at two different realization variables, or one variable in two distinct varieties, each analysis of a variable in a single variety counted toward this total.

a. Variation in whether or not a postposition is used to express ablative nominal relationships in Xining Chinese (Dede 1999)

b. Variable leveling in Late Middle and Early Modern English of the second-person subject pronoun *ye* to the object form *you* (Raumolin-Brunberg 2005)

c. Variation in expression of possession in child Latino English, e.g. *her~his* variation and *s~Ø* variation for the possessive morpheme on noun phrases (Wolford 2006)

Not every researcher examined social factors in their analysis; as mentioned earlier, *LVC* does not require consideration of social factors for publication, and several *LVC* studies reported on variables for which social factors were unrecorded or unavailable.[14] Narrowing down the data to only those realization variables for which the researcher looked for social patterning of any kind left us with 290 variables (75%).

Table 1 presents results that speak to our Social Realization Hypothesis (*Variation in the form of a grammatical marker will bear social significance*). We separate the three different domains of social significance (production, perception, metalinguistic behaviors). Where a variable was examined from more than one angle, the variable is counted once for each, meaning that the sum of the first row of numbers is greater than 290.

[14] A few *JSlx* studies were also counted as explicitly examining no social factors; these were typically studies of variation in a community that was sociolinguistically "interesting" (e.g. isolated, engaged in language contact) but without making comparisons between socially-demarcated subgroups.

Table 1. Number of realization variables in *LVC* and *JSlx* publications for which social factors were studied, and found to have an effect, separated by whether the social factors were assessed in the domain of production, perception, or metalinguistic behaviors.

	Domain in which social factors were examined		
	Production	Perception	Metalinguistic behaviors
Number of realization variables for which social factors were examined	261	12	69
Number of realization variables for which social factors were found to affect the variation	230	11	66
%	88%	92%	96%

It is clear from Table 1 that, for realization variables, researchers who look for social effects in any domain tend overwhelmingly to find them.[15] Some examples of realization variables with social significance are given in (4–6).

(4) Social significance of realization variation: Production

[15] A reviewer asks whether the variables for which social significance was examined but not found have anything in common. We examined these "null result" studies and couldn't see any obvious patterns. In the production domain, some of the null results may simply come from studies not being high-powered enough for the social factors to achieve significance, since much sociolinguistic work is done on small samples. In the metalinguistic domain, null results reflect authors noting that speakers in their sample seemed not to notice or comment on a particular variable, or that the variable under study was not mentioned by prescriptive grammars; the difficulty of proving the nonexistence of something means that metalinguistic commentary could still be out there for these variables.

a. Variation in whether or not a postposition is used to express ablative nominal relationships in Xining Chinese covaries with a speaker's age, urban vs. rural location, and education level (Dede 1999).

b. Variable leveling in Late Middle and Early Modern English of the second-person subject pronoun covaries with a speaker's region, social rank, gender, and addressee (family vs. nonfamily [Raumolin-Brunberg 2005]).

c. Variation in expression of possession in child Latino English covaries with a child's gender, grade in school, and language origin (Mexican vs. Puerto Rican, potentially attributable to contact among Puerto Ricans with speakers of African American English [Wolford 2006]).

(5) Social significance of realization variation: Perception

a. The formal second person pronoun in German is perceived as older, socially distant, and associated with certain regions and domains (Clyne et al. 2006).

b. Leveled *there's* (for *there are*) and *don't* (for *doesn't*) in American English are perceived as more likely to have been uttered by a lower class speaker than a middle class one (Squires 2013).

c. Listeners from Northern England, and those from Southern England with a particular cognitive profile, associate nonstandard subject-verb agreement with non-professionalism in a matched-guise study (Levon and Buchstaller 2015).

(6) Social significance of realization variation: Metalinguistic behaviors

a. Speakers of Japanese report "vivid recollections" of learning to use honorific forms and observe that their use is changing (Wetzel 1994: 134).

b. A teacher in Teesside, England was observed during fieldwork "express[ing] strong views against the use of [mi] for the possessive singular, evidently believing it to be an "incorrect" grammatical form"[16] (Snell 2010: 639).

c. Lack of plural concord in Brazilian Portuguese was the subject of a televised 2011 debate concerning whether sentences lacking concord should be presented in a textbook (Scherre and Naro 2014: 332).

It is clear from the above that variation in realization can show social significance in production, perception, and metalinguistic behaviors. The Social Realization Hypothesis holds; the Grammatical Invisibility Principle cannot be said to hold of variation in realization, because there is a preponderance of evidence that realization variables can be socially significant.

In the production domain, this is something we knew already. As we mentioned previously (Section 3.1), Levon and Buchstaller (2015) have already observed several apparent counterexamples against what we call the Grammatical Invisibility Principle, all of which constitute variation in realization under our analysis. However, these were all in the domain of production. The addition of data from perception and metalinguistic behaviors further refutes the theory that realization variables are immune to social significance.

Table 2 provides data on the number of socially significant realization variables for which the social significance pertains to overt prestige/stigma.

[16] I.e., usage of object *me* as opposed to standard possessive *my*.

Table 2. Number of realization variables in *LVC* and *JSlx* publications for which social factors were found to have an effect, and which are implicated on an overt prestige/stigma axis, separated by whether the social factors were assessed in the domain of production, perception, or metalinguistic behaviors.

	Production	Perception	Metalinguistic behaviors
Number of socially significant variables	230	11	66
Number of socially significant variables on overt prestige/stigma axis	80	7	36
%	35%	64%	55%

The figures in Table 2 lend some support to the suggestion that the grammatical variables that have been studied from a social perspective have tended to be above the level of conscious awareness and/or have one variant that is codified, with at least a third of all studies reporting an overt prestige/stigma dimension to the variation. It should also be recognized that authors do not always mention when a variable bears overt prestige/stigma, so these numbers likely underreport the reality. For instance, 74 realization variables in the database come from some variety of African American English. The stigma against this variety is well known (e.g. Rickford and King 2016), and yet only 19 of those 74 variables are accompanied by an explicit mention of their social stigma. We thus cautiously confirm suggestions by Cheshire (1987, 1999, 2005), Moore (2020, 2021), and Eckert (2018) that socially-meaningful grammatical variables tend to fall on an axis of overt prestige/stigma, at least in the realization domain.[17]

[17] A reviewer comments that it seems unusual for a variable to receive metalinguistic commentary but no overt prestige/stigma associations. Some variables that fit this description are *T/V* pronouns (e.g. Clyne et al. 2006) and honorific forms (Wetzel 1994), where there is considerable overt awareness of the variation yet no form is necessarily stigmatized *per se*: *T* and *V* pronouns, as well as honorific and humble forms, are all acceptable given the right social context. We also counted media representations of language and stylized performances as metalinguistic behaviors; we find evidence in our database of speakers engaging in these behaviors even with variables that are not on an overt prestige/stigma axis. One example is leveled *were* in Outer Banks English, which is "assigned a special sociolinguistic role that sets it apart from the classic dichotomy in which a leveled nonstandard English form is

At the same time, this should not be taken to mean that realization variables cannot take on more nuanced social meanings, too. In (7) we sample a few instances of realization variables that communicate local or discourse-dependent social meanings instead of or in addition to bearing overt stigma due to their nonstandardness.

(7) Social significance of realization variation beyond overt prestige/stigma

a. Leveling of negated past *be* to *weren't* in Ocracoke English is not only "relatively free of stigma […] but it also has been assigned at least a covert positive value related to islander identity" (Schilling-Estes and Wolfram 1994: 297).

b. Appalachian English *a*-prefixing and nonstandard past tense are used by younger speakers for story-telling, perhaps tapping into a "community voice" forged by elders (Burkette 2007).

c. Southern Californian Salvadoran Spanish speakers use second person pronouns to show regional solidarity and shared identity (Raymond 2012).

d. Contracted *was* in Appalachian English provides a non-stigmatized way of "contest[ing] the social push toward a fully standardized system" as use of leveled *was* declines (Hazen 2014: 77).

The variables surveyed in (7a–d) have taken on the kinds of locally-salient meanings commonly demonstrated for phonological variables (see Eckert 2019b: 760–762 for a recent review). In so doing, they differ from the frequently-cited grammatical variables that Eckert (2019b: 758–760) surveys, which, by virtue of one variant's codification, index institutional and educational

socially stigmatized" (hence not on the overt prestige/stigma axis) and used in written renditions of the dialect (hence a target of metalinguistic behaviors [Schilling-Estes and Wolfram 1994: 279]).

orientations, and "qualities and activities associated with this orientation" (Eckert 2019b: 759). The findings surveyed in (7a–d) demonstrate that other social meanings are open to realization variables as well.

4.3 Social significance of variation in order

Our meta-study found 72 variables between the two journals that we classified as variation in order (including variables classified as involving both realization and order simultaneously). This comprises 70 from *LVC* and 2 from *JSlx*. Some examples of order variables are given in (8).

(8) Examples of order variables

a. Shift from subject-verb-object to subject-object-verb order in Korlai Creole Portuguese (Clements 1990)

b. Variable subject-verb order in spoken Arabic (Owens, Dodsworth, and Rockwood 2009)

c. Variation in the placement of adverbs in English (Waters 2013)

d. Variable particle placement in English (Röthlisberger and Tagliamonte 2020)

Narrowing down the data further to only those order variables for which social patterning was investigated left us with 31 studied variables (43%). Compare this to the percentage of realization variables for which social patterning was investigated, which was 75%. Sociolinguists and variationists who study order variables are apparently less likely to look for the effect of social factors than are sociolinguists and variationists who study realization variables. This is partly a consequence of the fact that many of the order variation studies published in *LVC* are

done on historical data, where social factors are often unavailable. Additionally, our numbers show that research on order variables is almost never published in *JSlx*, which publishes only research that has a strong social component. This may suggest that order variables are less likely to show social patterns than realization ones, but it may also suggest a true research gap to be filled (see Section 5 for further discussion).

Table 3 presents results that speak to our Social Order Hypothesis (*Variation in the linear ordering of linguistic elements will bear social significance*). As we did in Table 1, Table 3 separates the three different types of sociolinguistic evidence. Also as before, several variables are counted twice due to having been studied from multiple angles, meaning the sum of the first row of numbers is greater than 31, the total number of order variables studied from a social perspective across the two journals.

Table 3. Number of order variables in *LVC* and *JSlx* publications for which social factors were studied, and found to have an effect, separated by whether the social factors were assessed in the domain of production, perception, or metalinguistic behaviors.

	Domain in which social factors were examined		
	Production	Perception	Metalinguistic behaviors
Number of order variables for which social factors were examined	31	0	2
Number of order variables for which social factors were found to affect the variation	29	0	2
%	94%	N/A	100%

It is clear from Table 3 that, like realization variables, order variables can show social effects. Though numbers are small, the domain of production, at least, shows a 94% rate of social

conditioning that does not differ significantly from the 88% rate found for realization variables in Table 1 ($p = 0.55$, Fisher's exact test).

However, research on the social significance of order variation in the other two domains of sociolinguistic knowledge is very difficult to come by. We found no examples that could possibly fit in our perception category: no matched guise studies of order variation were reported in either journal, something we see as a pressing necessity for future work (see Sections 5 and 6). Examples of socially-conditioned order variables in the other two domains are given in (9–10).

(9) Social significance of order variation: Production

a. The shift from subject-verb-object to subject-object-verb order in Korlai Creole Portuguese shows generational change, being led by young people under influence from Marathi, a locally prestigious language (Clements 1990).

b. Variation in the order and case of English co-ordinate object noun phrases (e.g. *me and Jim* vs. *Jim and I*) covaries with level of education, and is age-graded (Angermeyer and Singler 2003).

c. The English dative alternation shows a register effect, with the prepositional dative favored in formal written registers and the double object construction favored in more casual registers (D'Arcy and Tagliamonte 2015).

d. Variable particle placement shows effects of age, community, and occupation level (Röthlisberger and Tagliamonte 2020).

(10) Social significance of order variation: Metalinguistic behaviors

> In a diachronic corpus of Brazilian Portuguese plays, playwrights "come to represent oral speech more accurately over time"; this, in turn, coincides with an increased use in the plays of certain *wh*-interrogative structures, suggesting "the semiconscious adaptation of innovative variants that the writers experience in their everyday life" (Rosemeyer 2019: 176–179).

There is clear evidence that order variables can be socially conditioned in production: order variation can show register, age, education, and occupation effects. However, there are few documented examples in our corpus of other social behaviors where order variables are concerned. The matched guise and explicit attitudes studies that are so common for phonetic and lexical variation are vanishingly rare in the domain of order variation. This absence comes despite the fact that our meta-study does turn up a number of synchronic studies in which an order variable is found to show social patterning in production — which could in principle allow social perceptions and metalinguistic awareness to be tested in follow-up work.

Table 4 provides data on the number of socially significant order variables for which the social significance pertains to overt prestige/stigma.

Table 4. Number of order variables in *LVC* and *JSlx* publications for which social factors were found to have an effect, and which are implicated on an overt prestige/stigma axis, separated by whether the social factors were assessed in the domain of production, perception, or metalinguistic behaviors.

	Production	Perception	Metalinguistic behaviors
Number of socially significant variables	29	0	2
Number of socially significant variables on overt prestige/stigma axis	4	N/A	1
%	14%	N/A	50%

The majority of order variables are not discussed from the perspective of overt prestige/stigma, though we again urge caution in interpreting these numbers, given that a lack of discussion shouldn't necessarily be taken as lack of an effect. However, unlike what was the case for realization variation, we do not find any examples of order variation taking on locally-salient or discourse-dependent social meanings, either. For most of the order variables in our sample, researchers report broad sociodemographic patterns, but leave details of social meaning unexplored. This does not mean that there is no nuance to the social significance, simply that it was not the focus of the research.

5 General discussion

The meta-study presented in Section 4 documented that social effects in production, perception, and metalinguistic behaviors for grammatical variables, when searched for, have been found in both realization and order. We confirm our hypotheses (1) and (2), casting doubt on the Grammatical Invisibility Principle in its broadest categorical formulation, as an assertion that grammatical variation will not bear social significance.

To evaluate a more nuanced, gradient version of the Grammatical Invisibility Principle, we compared rates of social significance between our two types of grammatical variation,

realization and order. In the domain of production, at least, where most studies have been carried out, these rates do not meaningfully differ, and in fact are very high (≥88% for each type of variable).

A concern that can be raised here is the "file drawer problem" (Rosenthal 1979): namely, the possibility that there have been many studies of realization and/or order variation that found no social significance and accordingly went unpublished. If this is the case, it means the ≥88% rates cited above are inflated. We think this is unlikely, for two reasons. First, demonstration of social significance is not a prerequisite for publication in *LVC*. *LVC* thus should not be subject to publication bias on this front, so the rate of social significance found in the papers published therein should, we believe, be reflective of real-world rates. In fact, when we narrow down the data to realization variables studied from a production angle (which characterizes the majority of studies in our sample), and calculate the proportion of *LVC* papers that found social significance out of those that looked for it, we find a rate of 92% (143 out of 156), significantly *higher* than the analogous rate of 83% for *JSlx* (87 out of 105; $X^2 = 4.653$, $p = 0.031$). In other words, even in the journal where demonstration of social significance is not required for publication, social significance is still uncovered at a very high rate. Second, the fact that the proportion of *JSlx* papers that found social significance is less than 100% means that *JSlx* must be publishing studies where social significance is looked for but not found — that is, papers containing a null social result — and indeed it is (e.g. Cukor-Avila and Bailey 2001;[18] Sharma 2005; Nagy 2011). This does not eliminate the possibility that other null studies have gone unpublished, but it provides reassurance that publication of null results does happen. Taken together, these two lines of reasoning lead us to a cautious conclusion that our results do not inflate the incidence of social

[18] Cukor-Avila and Bailey (2001) report a null result for interviewer race, the focus of their study, but they do find differences in the use of the studied variables across different interviewers, which could show that the variation is socially significant in terms of the interlocutor.

conditioning, at least for the most well-represented segment of our sample, studies of realization variation from a production angle.

As the preceding paragraph indicates, the nature of the evidence in favor of our hypotheses is strongly stacked in the production domain. Few published studies have examined either type of variable in perception, or probed the metalinguistic behaviors associated with order variation. This dearth of studies seems to have led some researchers to conclude that these variables do not show social patterning in these domains (see Section 2), but this is absence of evidence being taken as evidence of absence. In reality, the types of studies that have the potential to provide evidence on these points have largely not been published in the two most notable journals in sociolinguistics. As before, this could be due to a preponderance of null results in these domains preventing publication, but, again, we believe this is unlikely. More plausibly, the research has simply not been done. We return to this point in Section 6.

A notable characteristic of the socially significant variables in our study is that, for a number of them, their social patterning reflects their presence on an overt prestige/stigma dimension. This is not to say that this is the only dimension on which grammatical variables can carry social associations (see [7]), but we confirm others' intuitions that it is a prominent one (Cheshire 1987, 1999, 2005; Eckert 2018; Moore 2020, 2021). Additionally, a large majority of the variables studied come from varieties of English or other major European languages (see also Adli and Guy 2022). Specifically, out of the 427 grammatical variables in our sample, 235 of them (55%) come from some variety of English, and an additional 87 (20%) from some variety of French, Spanish, or Portuguese, leaving only 105 from other languages.

All this is to say that the prototypical study of a grammatical variable published in *LVC* or *JSlx* is a realization variable from some variety of English, studied from a production angle.

This qualifies the findings of our meta-study: we refute the Grammatical Invisibility Principle, but in a narrow way. It still remains to be seen what types of social meanings can be found for order variables, and whether social meanings are picked up on by listeners in perception for either type of variable. We see these as exciting and essential ways forward for future research.

6 Conclusion and future directions

This chapter has been concerned with the general question, of interest to sociolinguists for decades, of which elements of language can bear social significance. We carried out a meta-study of two sociolinguistics journals in order to evaluate what we have called the Grammatical Invisibility Principle: the proposal put forward in much sociolinguistic work that grammatical variables are socially special compared to phonetic and lexical ones. First, we broke down the category of "grammatical variables" into two categories: variation in realization and variation in order. We also identified three sources of evidence of the social significance of a linguistic form, following Campbell-Kibler (2016): social patterning in production, social evaluation in perception, and metalinguistic behaviors.

Our meta-study found a number of counterexamples to the Grammatical Invisibility Principle in both realization and order. Both of these types of variables can be socially stratified in production and be the target of metalinguistic behaviors. Moreover, realization variables can elicit social judgments in perception and be used to convey locally salient meanings beyond the domain of overt prestige/stigma. But whether the same holds for order variables is undetermined. We have no evidence that it does; at the same time, the body of research on order variation from a social perspective is small, and the possibility has not been directly tested. In general, our meta-study leads us to the conclusion that it is the paucity of variationist research on

grammatical variation — particularly in the domain of order — that has led researchers to propose the Grammatical Invisibility Principle, not a preponderance of evidence that these variables are exempt from social significance.

There are two major ways forward that we see as essential in order to continue the line of work that we have started here. The first involves filling the research gaps we have uncovered. We see a pressing need for matched-guise studies of order variation; for other types of studies that explicitly test listeners' perception of a speaker's social attributes (e.g. Squires 2013) in a way that targets order variation; and for careful analysis of order variation in production that attempts to discern whether it can carry social meanings beyond overt prestige/stigma. This research will need to address a wide range of languages; in addition to English (Waters 2013; Grafmiller and Szmrecsanyi 2018; Röthlisberger and Tagliamonte 2020), our meta-study turned up order variation in Brazilian Portuguese (Tarallo and Kato 1993), Arabic (Owens, Dodsworth, and Rockwood 2009; Owens, Dodsworth, and Kohn 2013), Korlai Creole Portuguese (Clements 1990), Danish (Ladegaard 2000), and Hungarian (Kontra 2001). Given the lack of linguistic diversity in grammatical variation research (see Section 5), and in variationist sociolinguistics more generally (Adli and Guy 2022), expanding the range of languages under consideration will be welcome.

The second is increased collaboration between sociolinguists and syntacticians (see Childs [This Volume] for further discussion of where these perspectives may overlap). This chapter has until now taken an atheoretical approach to classifying grammatical variation. But certain formal theories of grammar localize our two types of variation — realization and order — in well-defined ways. For instance, in Distributed Morphology (Halle and Marantz 1993), most cases of realization variation would be analyzed as variation in Spell Out, the grammatical

module where functional morphemes receive phonological content. By contrast, order variation would be analyzed as variation in the application of some movement operation in the Narrow Syntax, the grammatical module where syntactic objects are assembled into a hierarchical tree structure (see Robinson 2022: 18-23). If this architectural difference between realization (Spell Out) phenomena and order (Narrow Syntax) phenomena does indeed correspond to an identifiable sociolinguistic difference — as previous sociolinguists have hypothesized (e.g. Romaine 1984) — this is an important finding that other models of grammatical architecture may have difficulty explaining. Sociolinguistic evidence can thus bear on formal linguistic questions.

Attempting to map our atheoretical classification scheme (realization vs. order) onto a formal model of grammar will introduce new complications, however. For instance, variables that appear a priori to be cases of realization may in fact have their grammatical locus in the Narrow Syntax, the home of order variables. Take as an example variable relative pronouns in English, such as the variation between *who, that,* and zero, as in *The boy who/that/Ø I saw* (e.g. D'Arcy and Tagliamonte 2015). This variable has been called "a clear case of syntactic variation" in the sociolinguistic literature (Guy and Bayley 1995: 148), but it is not obvious from the surface forms that there are structural differences between the alternants. In fact, syntacticians continue to debate three popular analyses of relative clauses, which differ in whether the variation is localized to Spell Out or to the Narrow Syntax (see Bhatt 2015 for a detailed overview). Thus, our atheoretical categorization scheme may mask true structural differences between variables. For this reason, we encourage sociolinguists interested in grammatical variation to engage with the syntactic literature and seek out collaborations with syntacticians. The benefits of collaboration go both ways.

Acknowledgements

We are very grateful to all those who have read and commented on our earlier drafts of this paper. We would particularly like to single out Dan Duncan, the two anonymous reviewers for this volume, and members of the NYU Sociolinguistics Lab and the Newcastle Variation Research Group.

References

Adger, David & Jennifer Smith. 2005. Variation and the Minimalist Program. In Leonie Cornips & Karen P. Corrigan (eds.), *Syntax and Variation: Reconciling the Biological and the Social*, 149–178. Amsterdam: John Benjamins.

Adli, Aria & Gregory R. Guy. 2022. Globalising the study of language variation and change: A manifesto on cross-cultural sociolinguistics. *Language and Linguistics Compass*. Article e12452. https://doi.org/10.1111/lnc3.12452.

Angermeyer, Philipp S. & John Victor Singler. 2003. The case for politeness: Pronoun variation in co-ordinate NPs in object position in English. *Language Variation and Change* 15(3). 171–209. https://doi.org/10.1017/S0954394503152027

Beltrama, Andrea & Laura Staum Casasanto. 2017. *Totally* tall sounds *totally* younger: Intensification at the socio-semantics interface. *Journal of Sociolinguistics* 21(2). 154–182. https://doi.org/10.1111/josl.12230

Bhatt, Rajesh. 2015. Relative clauses and correlatives. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax - Theory and Analysis: An International Handbook,* vol. 1, 708–749. Berlin: De Gruyter Mouton.

Britain, David. 2002. Diffusion, levelling, simplification and reallocation in past tense BE in the English Fens. *Journal of Sociolinguistics* 6(1). 16–43. https://doi.org/10.1111/1467-9481.00175

Burkette, Allison. 2007. Constructing identity: Grammatical variables and the creation of a community voice. *Journal of Sociolinguistics* 11(2). 286–296. https://doi.org/10.1111/j.1467-9841.2007.00321.x

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

Campbell-Kibler, Kathryn. 2016. Towards a cognitively realistic model of meaningful sociolinguistic variation. In Anna Babel (ed.), *Awareness and Control in Sociolinguistic Research*, 123–151. Cambridge: Cambridge University Press.

Cheshire, Jenny. 1987. Syntactic variation, the linguistic variable, and sociolinguistic theory. *Linguistics* 25(2). 257–282. https://doi.org/10.1515/ling.1987.25.2.257

Cheshire, Jenny. 1999. Taming the vernacular: Some repercussions for the study of syntactic variation and spoken grammar. *Cuadernos de Filologia Inglesa* 8. 59–80. https://revistas.um.es/cfi/article/view/65681

Cheshire, Jenny. 2005. Syntactic variation and spoken language. In Leonie Cornips & Karen P. Corrigan (eds.), *Syntax and Variation: Reconciling the Biological and the Social,* 81–106. Amsterdam: John Benjamins.

Cheshire, Jenny, Paul Kerswill & Ann Williams. 2005. Phonology, grammar, and discourse in dialect convergence. In Peter Auer, Frans Hinskens & Paul Kerswill (eds.), *Dialect Change: Convergence and Divergence in European Languages*, 135–167. Cambridge: Cambridge University Press.

Childs, Claire. This Volume. Sociosyntax and the weight of the linguistic within sociolinguistics.

Clements, J. Clancy. 1990. Deletion as an indicator of SVO → SOV shift. *Language Variation and Change* 2(2). 103–133. https://doi.org/10.1017/S0954394500000302

Clyne, Michael, Heinz-Leo Kretzenbacher, Catrin Norrby & Doris Schüpbach. 2006. Perceptions of variation and change in German and Swedish address. *Journal of Sociolinguistics* 10(3). 287–319. https://doi.org/10.1111/j.1360-6441.2006.00329.x

Cornips, Leonie & Karen P. Corrigan. 2005. Convergence and divergence in grammar. In Peter Auer, Frans Hinskens & Paul Kerswill (eds.), *Dialect Change: Convergence and Divergence in European Languages*, 96–134. Cambridge: Cambridge University Press.

Coupland, Nikolas & Adam Jaworski. 2004. Sociolinguistic perspectives on metalanguage: Reflexivity, evaluation and ideology. In Adam Jaworski, Nikolas Coupland & Dariusz Galasinski (eds.), *Metalanguage: Social and Ideological Perspectives*, 15–52. Berlin: De Gruyter Mouton.

Cukor-Avila, Patricia & Guy Bailey. 2001. The effects of the race of the interviewer on sociolinguistic fieldwork. *Journal of Sociolinguistics* 5(2). 252–270. https://doi.org/10.1111/1467-9481.00150

Dailey-O'Cain, Jennifer. 2000. The sociolinguistic distribution of and attitudes toward focuser *like* and quotative *like*. *Journal of Sociolinguistics* 4(1). 60–80. https://doi.org/10.1111/1467-9481.00103

D'Arcy, Alexandra & Sali A. Tagliamonte. 2015. Not always variable: Probing the vernacular grammar. *Language Variation and Change* 27(3). 255–285. https://doi.org/10.1017/S0954394515000101

Dede, Keith. 1999. An ablative postposition in the Xining dialect. *Language Variation and Change* 11(1). 1–17. https://doi.org/10.1017/S0954394599111013

D'Onofrio, Annette. 2016. *Social meaning in linguistic perception*. Palo Alto, CA: Stanford University dissertation.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453–476. https://doi.org/10.1111/j.1467-9841.2008.00374.x

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41. 87–100. https://doi.org/10.1146/annurev-anthro-092611-145828

Eckert, Penelope. 2016. Variation, meaning and social change. In Nikolas Coupland (ed.), *Sociolinguistics: Theoretical Debates*, 68–85. Cambridge: Cambridge University Press.

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

Eckert, Penelope. 2019a. The individual in the semiotic landscape. *Glossa: A Journal of General Linguistics* 4(1). Article 14. https://doi.org/10.5334/gjgl.640

Eckert, Penelope. 2019b. The limits of meaning: Social indexicality, variation, and the cline of interiority. *Language* 95(4). 751–776. https://doi.org/10.1353/lan.2019.0072

Eckert, Penelope & William Labov. 2017. Phonetics, phonology and social meaning. *Journal of Sociolinguistics* 21(4). 467–496. https://doi.org/10.1111/josl.12244

Embick, David. 2008. Variation and morphosyntactic theory: Competition fractionated. *Language and Linguistics Compass* 2(1). 59–78. https://doi.org/10.1111/j.1749-818X.2007.00038.x

Fasold, Ralph W. 2013. Variation and syntactic theory. In Robert Bayley, Richard Cameron & Ceil Lucas (eds.), *The Oxford Handbook of Sociolinguistics*, 185–202. Oxford: Oxford University Press.

Fridland, Valerie. 2008. Regional differences in perceiving vowel tokens on Southerness, education, and pleasantness ratings. *Language Variation and Change* 20(1). 67–83. https://doi.org/10.1017/S0954394508000069

Fridland, Valerie, Kathryn Bartlett & Roger Kreuz. 2004. Do you hear what I hear? Experimental measurement of the perceptual salience of acoustically manipulated vowel variants by Southern speakers in Memphis, TN. *Language Variation and Change* 16(1). 1–16. https://doi.org/10.1017/S0954394504161012

Grafmiller, Jason & Benedikt Szmrecsanyi. 2018. Mapping out particle placement in Englishes around the world: A study in comparative sociolinguistic analysis. *Language Variation and Change* 30(3). 385–412. https://doi.org/10.1017/S0954394518000170

Guy, Gregory R. & Robert Bayley. 1995. On the choice of relative pronouns in English. *American Speech* 70(2). 148–162. https://doi.org/10.2307/455813

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale & Samuel Jay Keyser (eds.), *The View from Building 20,* 111–176. Cambridge, MA: The MIT Press.

Hay, Jennifer, Paul Warren & Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34(4). 458–484. https://doi.org/10.1016/j.wocn.2005.10.001

Hazen, Kirk. 2014. A new role for an ancient variable in Appalachia: Paradigm leveling and standardization in West Virginia. *Language Variation and Change* 26(1). 77–102. https://doi.org/10.1017/S0954394513000215

Hesson, Ashley & Madeline Shellgren. 2015. Discourse marker *like* in real time: Characterizing the time-course of sociolinguistic impression formation. *American Speech* 90(2). 154–186. https://doi.org/10.1215/00031283-3130313

Hudson, Richard A. 1980. *Sociolinguistics*. Cambridge: Cambridge University Press.

Hudson, Richard A. 1996 [1980]. *Sociolinguistics*, 2nd edn. Cambridge: Cambridge University Press.

Johnstone, Barbara & Scott F. Kiesling. 2008. Indexicality and experience: Exploring the meanings of /aw/-monophthongization in Pittsburgh. *Journal of Sociolinguistics* 12(1). 5–33. https://doi.org/10.1111/j.1467-9841.2008.00351.x

Kontra, Miklós. 2001. Hungarian verbal puzzles and the intensity of language contact. *Journal of Sociolinguistics* 5(2). 163–179. https://doi.org/10.1111/1467-9481.00146

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3). 199–244. https://doi.org/10.1017/S0954394500000168

Kroch, Anthony & Cathy Small. 1978. Grammatical ideology and its effect on speech. In David Sankoff (ed.), *Linguistic Variation: Models and Methods*, 45–55. New York: Academic Press.

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

Labov, William. 1978. Where does the linguistic variable stop? A response to Beatriz Lavandera. In Richard Bauman & Joel Sherzer (eds.), *Working Papers in Sociolinguistics 44,* 6–22. Austin, TX: Southwest Educational Development Laboratory.

Labov, William. 1993. The unobservability of structure and its linguistic consequences. Paper presented at New Ways of Analyzing Variation in English 22, University of Ottawa.

Labov, William. 2001. *Principles of Linguistic Change, Volume 2: Social Factors*. Oxford: Blackwell.

Ladegaard, Hans J. 2000. Language attitudes and sociolinguistic behaviour: Exploring attitude-behaviour relations in language. *Journal of Sociolinguistics* 4(2). 214–233. https://doi.org/10.1111/1467-9481.00112

Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner & Samuel Fillenbaum. 1960. Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology* 60(1). 44–51. https://psycnet.apa.org/doi/10.1037/h0044430

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

Levon, Erez & Isabelle Buchstaller. 2015. Perception, cognition, and linguistic structure: The effect of linguistic modularity and cognitive style on sociolinguistic processing. *Language Variation and Change* 27(3). 319–348. https://doi.org/10.1017/S0954394515000149

Lignos, Constantine, Laurel MacKenzie & Meredith Tamminga (eds.). 2018. *The Locus of Linguistic Variation*. Amsterdam: John Benjamins.

MacKenzie, Laurel. 2012. *Locating variation above the phonology*. Philadelphia: University of Pennsylvania dissertation.

MacKenzie, Laurel. 2013. Variation in English auxiliary realization: A new take on contraction. *Language Variation and Change* 25(1). 17–41. https://doi.org/10.1017/S0954394512000257

MacKenzie, Laurel. 2020. Comparing constraints on contraction using Bayesian regression modeling. *Frontiers in Artificial Intelligence: Language and Computation* 3. Article 58. https://doi.org/10.3389/frai.2020.00058

Maddeaux, Ruth & Aaron Dinkin. 2017. Is *like* like *like*? Evaluating the same variant across multiple variables. *Linguistics Vanguard* 3(1). Article 20150032. https://doi.org/10.1515/lingvan-2015-0032

Mansfield, John, Henry Leslie-O'Neill & Haoyi Li. 2023. Dialect differences and linguistic divergence: A crosslinguistic survey of grammatical variation. *Language Dynamics and Change* 13(2). 232–276. https://doi.org/10.1163/22105832-bja10026

McCloskey, James. 2000. Quantifier float and *wh*-movement in an Irish English. *Linguistic Inquiry* 31(1). 57–84. https://doi.org/10.1162/002438900554299

Meyerhoff, Miriam. 2001. Dynamics of differentiation: On social psychology and cases of language variation. In Nikolas Coupland, Srikant Sarangi & Christopher N. Candlin (eds.), *Sociolinguistics and Social Theory*, 61–87. London: Routledge.

Meyerhoff, Miriam & James A. Walker. 2013. An existential problem: The sociolinguistic monitor and variation in existential constructions on Bequia (St. Vincent and the Grenadines). *Language in Society* 42(4). 407–428. https://doi.org/10.1017/S0047404513000456

Moore, Emma. 2020. The role of syntax in the study of sociolinguistic meaning: Evidence from an analysis of right dislocation. In Karen V. Beaman, Isabelle Buchstaller, Susan Fox & James A. Walker (eds.), *Advancing Socio-grammatical Variation and Change: In Honour of Jenny Cheshire*, 73–90. London: Routledge.

Moore, Emma. 2021. The social meaning of syntax. In Lauren Hall-Lew, Emma Moore & Robert J. Podesva (eds.), *Social Meaning and Linguistic Variation: Theorizing the Third Wave*, 54–79. Cambridge: Cambridge University Press.

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

Owens, Jonathan, Robin Dodsworth & Mary Kohn. 2013. Subject expression and discourse embeddedness in Emirati Arabic. *Language Variation and Change* 25(3). 255–285. https://doi.org/10.1017/S0954394513000173

Owens, Jonathan, Robin Dodsworth & Trent Rockwood. 2009. Subject-verb order in spoken Arabic: Morpholexical and event-based factors. *Language Variation and Change* 21(1). 39–67. https://doi.org/10.1017/S0954394509000027

Raumolin-Brunberg, Helena. 2005. The diffusion of subject YOU: A case study in historical sociolinguistics. *Language Variation and Change* 17(1). 55–73. https://doi.org/10.1017/S0954394505050039

Ravid, Dorit. 1995. Neutralization of gender distinctions in Modern Hebrew numerals. *Language Variation and Change* 7(1). 79–100. https://doi.org/10.1017/S0954394500000909

Raymond, Chase Wesley. 2012. Reallocation of pronouns through contact: In-the-moment identity construction amongst Southern California Salvadorans. *Journal of Sociolinguistics* 16(5). 669–690. https://doi.org/10.1111/josl.12001

Rickford, John R., Arnetha Ball, Renée Blake, Raina Jackson & Nomi Martin. 1991. Rappin on the copula coffin: Theoretical and methodological issues in the analysis of copula variation in African-American Vernacular English. *Language Variation and Change* 3(1). 103–132. https://doi.org/10.1017/S0954394500000466

Rickford, John R. & Sharese King. 2016. Language and linguistics on trial: Hearing Rachel Jeantel (and other vernacular speakers) in the courtroom and beyond. *Language* 92(4). 948–988. https://doi.org/10.1353/lan.2016.0078

Rickford, John R., Thomas A. Wasow, Norma Mendoza-Denton & Juli Espinoza. 1995. Syntactic variation and change in progress: Loss of the verbal coda in topic-restricting *as far as* constructions. *Language* 71(1). 102–131. https://doi.org/10.2307/415964

Robinson, Mary. 2022. *Negative concord as a window into social perception of morphological and syntactic variables*. New York: New York University dissertation.

Robinson, Mary & Gary Thoms. 2021. On the syntax of English variable negative concord. In Alexandros Kalomoiros & Lefteris Paparounas (eds.), *University of Pennsylvania Working Papers in Linguistics 27.1: Proceedings of the 44th Annual Penn Linguistics Conference*, 195–204. Philadelphia: Penn Graduate Linguistics Society.

Romaine, Suzanne. 1984. On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory. *Folia Linguistica* 18(3–4). 409–438. https://doi.org/10.1515/flin.1984.18.3-4.409

Rosemeyer, Malte. 2019. Actual and apparent change in Brazilian Portuguese *wh*-interrogatives. *Language Variation and Change* 31(2). 165–191. https://doi.org/10.1017/S0954394519000097

Rosenthal, Robert. 1979. The file drawer problem and tolerance for null results. *Psychological Bulletin* 86(3). 638–641. https://psycnet.apa.org/doi/10.1037/0033-2909.86.3.638

Röthlisberger, Melanie & Sali A. Tagliamonte. 2020. The social embedding of a syntactic alternation: Variable particle placement in Ontario English. *Language Variation and Change* 32(3). 317–348. https://doi.org/10.1017/S0954394520000174

Scherre, Maria Marta Pereira & Anthony Julius Naro. 2014. Sociolinguistic correlates of negative evaluation: Variable concord in Rio de Janeiro. *Language Variation and Change* 26(3). 331–357. https://doi.org/10.1017/S0954394514000143

Schilling‐Estes, Natalie. 2002. On the nature of isolated and post‐isolated dialects: Innovation, variation and differentiation. *Journal of Sociolinguistics* 6(1). 64–85. https://doi.org/10.1111/1467-9481.00177

Schilling-Estes, Natalie & Walt Wolfram. 1994. Convergent explanation and alternative regularization patterns: *Were/weren't* leveling in a vernacular English variety. *Language Variation and Change* 6(3). 273–302. https://doi.org/10.1017/S0954394500001691

Sharma, Devyani. 2005. Dialect stabilization and speaker awareness in non-native varieties of English. *Journal of Sociolinguistics* 9(2). 194–224. https://doi.org/10.1111/j.1360-6441.2005.00290.x

Smith, Jennifer. 2007. Techniques of analysis (ii) morphosyntactic variation. In Carmen Llamas, Louise Mullany & Peter Stockwell (eds.), *The Routledge Companion to Sociolinguistics*, 28–40. London: Routledge.

Snell, Julia. 2010. From sociolinguistic variation to socially strategic stylisation. *Journal of Sociolinguistics* 14(5). 630–656. https://doi.org/10.1111/j.1467-9841.2010.00457.x

Squires, Lauren. 2013. It don't go both ways: Limited bidirectionality in sociolinguistic perception. *Journal of Sociolinguistics* 17(2). 200–237. https://doi.org/10.1111/josl.12025

Stausland Johnsen, Sverre. 2015. Dialect change in South-East Norway and the role of attitude in diffusion. *Journal of Sociolinguistics* 19(5). 612–642. https://doi.org/10.1111/josl.12156

Tagliamonte, Sali A. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Blackwell.

Tagliamonte, Sali A. 2014. A comparative sociolinguistic analysis of the dative alternation. In Rena Torres Cacoullos, Nathalie Dion & André Lapierre (eds.), *Linguistic Variation: Confronting Fact and Theory*, 297–318. London: Routledge.

Tamminga, Meredith. 2016. Persistence in phonological and morphological variation. *Language Variation and Change* 28(3). 335–356. https://doi.org/10.1017/S0954394516000119

Tamminga, Meredith, Laurel MacKenzie & David Embick. 2016. The dynamics of variation in individuals. *Linguistic Variation* 16(2). 300–336. https://doi.org/10.1075/bct.97.06tam

Tarallo, Fernando & Mary A. Kato. 1993. Filling syntactic boundaries in spoken Brazilian Portuguese. *Language Variation and Change* 5(1). 91–112. https://doi.org/10.1017/S0954394500001423

Violin-Wigent, Anne. 2007. Regional vocabulary leveling: The example of south-eastern French. *Journal of Sociolinguistics* 11(3). 408–422. https://doi.org/10.1111/j.1467-9841.2007.00330.x

Walker, Abby. 2007. The effect of phonetic detail on perceived speaker age and social class. In Jürgen Trouvain & William J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken*, 1453–1456.

Walker, Abby, Christina García, Yomi Cortés & Kathryn Campbell-Kibler. 2014. Comparing social meanings across listener and speaker groups: The indexical field

of Spanish /s/. *Language Variation and Change* 26(2). 169–189. https://doi.org/10.1017/S0954394514000088

Waters, Cathleen. 2013. Transatlantic variation in English adverb placement. *Language Variation and Change* 25(2). 179–200. https://doi.org/10.1017/S0954394513000082

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

Wetzel, Patricia J. 1994. Contemporary Japanese attitudes toward honorifics (*keigo*). *Language Variation and Change* 6(2). 113–147. https://doi.org/10.1017/S0954394500001617

Winford, Donald. 1996. The problem of syntactic variation. In Jennifer Arnold, Renée Blake, Brad Davidson, Scott Schwenter & Julie Solomon (eds.), *Sociolinguistic Variation: Data, Theory, and Analysis: Selected Papers from NWAV 23 at Stanford,* 177–192. Stanford, CA: CSLI Publications.

Wolford, Tonya E. 2006. Variation in the expression of possession by Latino children. *Language Variation and Change* 18(1). 1–13. https://doi.org/10.1017/S0954394506060054

Wolfram, Walt & Natalie Schilling. 2016 [1998]. *American English: Dialects and Variation*, 3rd edn. Hoboken, NJ: John Wiley & Sons, third edition.

Wolfram, Walt. 1969. *A Sociolinguistic Description of Detroit Negro Speech*. Washington, DC: Center for Applied Linguistics.

Zilles, Ana M. S. 2005. The development of a new pronoun: The linguistic and social embedding of *a gente* in Brazilian Portuguese. *Language Variation and Change* 17(1). 19–53. https://doi.org/10.1017/S0954394505050027