CrossMark

Language Variation and Change, **27** (2015), 255–285.
© Cambridge University Press, 2015 0954-3945/15
doi:10.1017/S0954394515000101

Not always variable: Probing the vernacular grammar

ALEXANDRA D'ARCY
University of Victoria

SALI A. TAGLIAMONTE
University of Toronto

ABSTRACT

Written and spoken language are known to differ substantially (Biber, 1988; 1995; Biber, Johansson, Leech, Conrad, & Finegan, 1999). Standard written language is highly uniform and governed by prescription, whereas the vernacular is most revealing of structured heterogeneity (Weinreich, Labov, & Herzog, 1968). We focus on four English morphosyntactic variables that problematize assumptions about the nature of variation in the vernacular: the genitive, the comparative, the dative, and relative pronouns. Each is characterized in casual speech by functional divides that reflect discrete configurations of variant use. After detailing the patterning of these variables in speech, we explore a characteristic arguably shared by each: its historical pathway into the language, where analogy and prestige were powerful motivations for variant choice. We suggest that this combination of systemic and social factors contributed to the nature of these variables in the vernacular grammar. Furthermore, we advocate for greater scrutiny of written and spoken data and the outcomes of change from above and below within each register. The type of innovation and its trajectory may affect the nature of the emergent variable grammar.

In variationist sociolinguistics, inherent structured variation within a grammatical sector is considered the sine qua non of the vernacular (Weinreich et al., 1968). The foundational assumption is that language variation is governed by probabilistic constraints and that systematic choices among variants are

We gratefully acknowledge the support of the Social Sciences and Humanities Research Council of Canada for grants to D'Arcy and Tagliamonte, the National Science Foundation for a grant that partially supported Tagliamonte's work on genitives and datives, and the Work Study Programs at the Universities of Victoria and Toronto. The extensive data extraction and coding was completed by a skilled team of research assistants: Hannah Coulter, Bridget Jankowski, Lidia Jarmasz, Alesia Malec, Shannon Mooney, Michael Ritter, and Dylan Usher. An earlier version of this work was presented at New Ways of Analyzing Variation (NWAV41), University of Indiana Bloomington, 2012; we thank Josef Fruehwald, Greg Guy, Bill Kretzschmar, Gillian Sankoff, Carola Trips, James Walker, and others for their responses and questions. We also appreciate our anonymous reviewers and especially the wise commentary of Anthony Warner, who helped us think more critically about the issues we discuss in this paper.

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

constrained by social and linguistic factors (Labov, 1963 et seq.). A further assumption is that these factors operate most robustly, and most systematically, in vernacular speech, where corrective norms exert less influence (Labov, 1984). Circumscribing the variable context is vital, as it defines the loci of variation and categoricity in the grammar. This procedure, coupled with rigorous linguistic interpretation and robust statistical modeling, has substantially advanced understanding of the linguistic system and how it changes. The challenge we face in this paper is the following. When linguistic variables have curtailed variability, how can they be reconciled with standard conceptions of variation so as to contribute to the building knowledge base of the field? Most linguistic variables in the literature exhibit the expected structured variation. In this paper, we chronicle a collection of variables from our own research that exhibit an apparently exceptional characteristic—not variation, but partitioning. We then explore what this type of variation contributes to the ontology of language variation and language change and suggest that a combination of inherited historical patterns and functional requirements may explain the unexpected profile of these linguistic variables.

Central to our discussion are four grammatical phenomena that are putatively ideal candidates for variationist study: the genitive (1), the comparative (2), the dative (3), and relative pronouns (4).

(1) a. He was the manager of the *Toyland of the T. Eaton Company*, the *T. Eaton Company's Toyland*, and so they were fairly well off. (N/Q/f/72)[1]
b. He was the *movie star's son* … the *son of the movie star*, he died. (R/O/m/11)

(2) a. A cottage can be a lot *quirkier* … it is obvious that cottages are *more quirky*. (N/j/m/47)
b. I would say it was much *more strict* when I began teaching. … They've introduced much *stricter* curriculum ideas, and you know, that kind of thing. (N/P/m/73)

(3) a. I emailed *them a letter* … I wrote *a letter to them*. (I/8/m/32)
b. It's like, 'I've known you for so long. Give *me some pizza*.' And I finally gave up 'cause he got annoying so I gave *it to him*. (R/6/f/12)

(4) a. Neighbours who are everything from a woman *who* runs a exercise place to a guy *that* works in a nursery. (R/j/m/47)
b. The biggest change *that* my grandmother's always teasing me about is we're not as fluid with our speaking. She says a lot of talking Ø she notices is putting into like sweeping generalizations. (N/k/f/25)

Each system is reported to have systematic variation in the written language (both historical and contemporary), with a complex set of constraint effects operating on variant choice (e.g., Ball, 1996; Hinrichs & Szmrecsanyi, 2007; Hinrichs, Szmrecsanyi, & Bohmann, 2014; Mondorf, 2003; Szmrecsanyi & Hinrichs, 2008; Wolk, Bresnan, Rosenbach, & Szmrecsanyi, 2013). In the spoken language, a different reality obtains (cf. D'Arcy, 2014; D'Arcy & Tagliamonte,

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

2010; Jankowski & Tagliamonte, 2014; Tagliamonte, 2014). Variation is apparent in the aggregate, but analysis reveals at least one attribute that divides the system. When such key elements are taken into account, variation virtually disappears. Thus, it seems that contrary to current variationist epistemology, not all change gives rise to the familiar structured heterogeneity. We suggest that when an extant linguistic system is modified from above (i.e., outside the community), a possible outcome is the entrenchment of separate subsystems in the vernacular while the written language retains structured variation on account of functional complexity and/or inherited historical patterns.

DICHOTOMIES OF LANGUAGE CHANGE

Change may arise through internal systemic shifts, but it can also be initiated or buttressed by external circumstances, such as war, cultural revolution, resettlement, migration, geological phenomena, and so on. For the most part, these linguistic influences have been treated as discrete and antithetic. For example, in historical linguistics, language change is viewed as a contrast between *endogenous* and *exogenous*—that is, internally versus externally triggered. In variationist sociolinguistics, language change is framed in terms of *above* and *below*. Changes from below arise from "a gradual shift in the behavior of successive generations, well below the level of conscious awareness" (Labov, 1966:206). Changes from above result from the operation of social forces and consist of importations from external languages or dialects (Labov, 2010:185). These pathways are distinguished from one another along a number of parameters, though our research is beginning to show that the line between the two is not always clear-cut.

Change from below (or *from within*; Labov, 2006:203) is the "normal type of internal change," originating within the linguistic system (Labov, 2010:§51.2; examples are also found in Beddor, 2009; 2012; Garret & Johnson, 2013; Kirby, 2013; Ohala, 1993; Stevens & Harrington, 2014). Outside the phonology, such changes are generally initiated through the mechanisms of reanalysis and extension (e.g., Harris & Campbell, 1995). Because actuation is internal, change from below is *evolutive*. This means that it is explicable with reference to the community-based linguistic system (e.g., Andersen, 1973), emerging from the inherited structure of grammatical systems (i.e., via adult to child transmission; cf. Labov, 2007). This kind of change is systematic and initially unconscious. That is, despite early descriptions (Labov, 1966:128) *above* and *below* are only partially defined along the continuum of consciousness. The key attribute of change from below is the point of actuation (system internal) and the trajectory of development (evolutive). Although the etiology of endogenous change is discernable with reference to language-internal factors alone, the diffusion of change in social context must be motivated by speaker-based factors (Armstrong & Mackenzie, 2013:124). Change originates in speakers, not languages (Milroy & Milroy, 1985:347), necessitating a distinction between the conditions that give

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

rise to a change and those that have to do with its diffusion (see also Joseph & Janda, 2003, among others; Stevens & Harrington, 2014:2). As an innovation moves from speaker to speaker or from group to group, its diffusion is exogenous, propelled by influences outside the community grammar, what Andersen (1973) referred to as *adaptive* change (cf. Milroy, 2003:219).

Change from above (or *from without*; Labov, 2006:203) refers to the importation of elements from other systems (Labov, 2010:§51.2). It is sporadic, conscious, and "usually recognize[d] … by the fact that it involves high-prestige features" (Labov, 2010:§9.1). Indeed, the motivations for change from above derive from overt social pressures in society, of the type imposed by teachers and the conscious responses of certain sectors of the population (e.g., the middle class). The effect is socially observable in public performances and reactions. Crucially, change from above is fragmentary. It involves isolated features from outside the local speech community, and so may affect only parts of the linguistic system rather than sweeping through the language as a whole. This renders the outcome less predictable and less systematic than change from below. Consequently, it is not particularly informative with respect to the "systematic forces that mold the history of dialect divergence" (Labov, 2010:§9.1). Nonetheless, it remains an important phenomenon for explaining the synchronic patterns of language in use.

The outcomes of these different pathways of change are less well documented, prompting Sankoff (2013:261) to ask: "Does change from above have a different profile from change from below?" The four variables we examine are arguably the result of change from above, whether direct or indirect, suggesting that this type of change may well have a distinct profile. The constraints we report on here, which operate to partition these systems, are well documented. Indeed, the extensive quantitative results reported in the literature describe their operation in detail; however, few analyses fully explore the extent and nature of the variable contexts as we do here. Therefore, we cannot make any definitive claim as to how widespread or generalizable the partitioning may be. Instead, we document the ways the vernacular systems of our data operate. We suggest that the traditional dichotomization of language change as from above or from below is explained not only through the manner of actuation, with reflexes in transmission or diffusion and distinct pathways through social structure, but may also affect the living language in fundamentally distinct ways. In this respect, evidence regarding the nature of linguistic change may be retrievable from footprints left in the variable grammar.

DATA AND VARIABLES

The four linguistic variants that are central to our discussion developed as historical changes in English: the *of* genitive, comparative *more*, the *to* dative, and the *wh*–relative pronouns. Our goal is to document and synthesize across their contemporary patterning. The data come from contemporary Canada, where we have access to some of the largest representative corpora of vernacular spoken

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

English. All four variables have previously been subjected to full-scale quantitative study in these materials (D'Arcy, 2014; D'Arcy & Tagliamonte, 2010; Jankowski & Tagliamonte, 2014; Tagliamonte, 2014).

Data

Our analyses are (mostly) based on the Toronto English Archive (Tagliamonte, 2003–2006). For one variable, the dative, the analysis includes the Ontario Dialects Archive (Tagliamonte, 2007–2010; 2010–2013).[2] The Toronto English Archive database comprises more than 350 hours of recorded spontaneous speech, representing approximately 1.8 million words. The speakers—all of whom were born and raised in metropolitan Toronto—range between the ages of 9 and 92 (years of birth: 1912–1994) and are stratified by gender, occupation, and education. The Ontario Dialects Archive is similarly constituted and comprises speakers born and raised in small cities and towns across Ontario. All of the data were collected following sociolinguistic interviewing techniques (Labov, 1970, 1971), and individuals are represented by anywhere from one to three hours of informal conversational data. The discourse is varied but is uniformly lively and interactive and includes numerous narrative sequences. The materials thus capture the contemporary vernacular norms of these localities. Although Canadian English is a relatively standard variety, many nonstandard variables are present in the data, including nonagreement, multiple negation, verb forms, double comparatives, and zero inflection, as in (5a) to (5g). Variation is not restricted to any particular social group in the sample but reflects the broader, community grammar.

(5) a. So the rest of the world *don't* know what they're doing. (O/d/m/83)
b. We *didn't* get *no* damage. (T/G/f/76)
c. You get a lot of them *come* back to us. (T/T/m/49)
d. They had *selled* their things. (N/S/f/34)
e. She *come* out and she *says*, 'In the name of God, what have you got now?' (N/b/m/85)
f. They're speaking a *little more fancier*. (N/E/m/36)
g. It's amazing how quick–Ø the people do panic. (N/v/f/55)

Variables

The variables are well known in the variationist literature as well as in historical linguistics and corpus linguistics. For example, relativization—the domain of the *wh–* relatives—has been described as "one of the most researched topics in the history of the English language" (Nevalainen & Raumolin-Brunberg, 2002:109). A common characteristic of studies of the genitive, the comparative, the dative, and relativization is that the research has drawn (most of) its evidence from standard language and from written genres. The findings have uniformly reported robust variation and multidimensional patterns of variation (e.g., Gries,

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

2002; Guy & Bayley, 1995; Hilpert, 2008; Hinrichs & Szmrecsanyi, 2007; Hinrichs et al., 2014; Kytö & Romaine, 1997, 2000, 2006; Leech & Culpeper, 1997; Ljung, 1997; Mondorf, 2003, 2009; Nevalainen & Raumolin-Brunberg, 2002; Rissanen, 1984; Romaine, 1980; Rosenbach, 2003, 2005, 2006; Schlüter, 2001; Tagliamonte, Smith, & Lawrence, 2005; Tottie, 1995; Wolk et al., 2013). Our prior work on these variables has also uncovered considerable conditioning, both external and internal (D'Arcy, 2014; D'Arcy & Tagliamonte, 2010; Jankowski & Tagliamonte, 2014). However, in this paper, we abstract away from the social aspects of variation to probe the linguistic patterns of variant choice.[3]

The genitive. The ancient genitive form is the synthetic Germanic *–s*. The analytic Latinate *of* construction developed late in Old English. During Middle English, a semantic division emerged: *–s* became associated with human and other animate possessors (*the turtle's face*), and *of* specialized for inanimate possessors (*spokes of the bike*) (Altenberg, 1982:146–147; Rosenbach, 2002:180, 190–192). This correlation continues in contemporary English (e.g., Biber et al., 1999:302–303; Ljung, 1997:25; Rosenbach, 2005:614; 2008:152) and has critical ramifications for variation within the sector. It also has ramifications for variationist methods when applied to data such as ours. Specifically, vernacular discourse largely revolves around people: family, friends, colleagues, etc. As a consequence, the initial sample of 100 speakers resulted in just 304 nonhuman antecedents in 953 tokens (Tagliamonte & Jarmasz, 2008). To mitigate the natural imbalance in the data, only inanimate genitives were extracted from a further 90 speakers (Jankowski & Tagliamonte, 2014:313).

Our analysis focuses on those contexts where either variant could have been used or where such alternation would not change the meaning (see Hinrichs & Szmrecsanyi, 2007:445; Jankowski & Tagliamonte, 2014:310; Rosenbach, 2002:29). Idioms and set phrases (*for God's sake, Murphy's Law*), descriptives (*men's room*), and partitives (*some of the people*) were excluded, as either these do not vary or their propositional content is affected by the change in form. Also excluded were human proper names (*Marisa's dog*) and units of temporal measure (*years of my life, height of summer*), as these exhibited exceptional rates of use: 99% *–s* and 95% *of,* respectively (Jankowski & Tagliamonte, 2014:311).

This approach ultimately yielded 1,368 genitive tokens for analysis, distributed as in Table 1 (190 speakers, ages 17–92). Consistent with the literature, the variant proportions suggest that the genitive alternation is robust in these materials.

A number of factors constrain variation. *Relative end weight* dictates that shorter constituents will precede those that are "heavier" (Hinrichs & Szmrecsanyi, 2007:453–455; Szmrecsanyi & Hinrichs, 2008:292–293, 299–300). This predicts that when a possessor consists of fewer words than the possessum, the possessor will take the *–s* genitive (*Americans' grand plans*). Conversely, the *of* genitive is more likely when the possessum consists of fewer words than the possessor (*the feelings of certain professors*; nb. determiners do not contribute

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

TABLE 1. *Overall distribution of genitives*

Comparative Variant	%	*n*
Synthetic *–s*	46.7	639
Analytic *of*	53.3	729
Total		1368

Source: Data from Jankowski & Tagliamonte (2014:314, Table 2).

to possessum length). *Final phonological segment* operates such that a final sibilant in the possessor disfavors the synthetic genitive (*nuns, research*) (Hinrichs & Szmrecsanyi, 2007:452). Semantics are also relevant, functioning via a *possessor relation* constraint. This effect is based in typology and differentiates between prototypical and nonprototypical relationships (Grafmiller, 2014; Rosenbach, 2002; Wolk et al., 2013). Prototypical possessors include body parts, kinship terms, whole/part relationships of concrete inanimate things, and permanent/legal ownership. Because the "possessor and possessum form a very tight, if not inseparable, relation" that "should be encoded by more implicit linguistic means" (Rosenbach, 2002:123), these are predicted to take synthetic *–s* (Wolk et al., 2013:398). Nonprototypical possessors, which comprise all other relations, are expected to favor the analytic genitive, *of*. *Persistence*, a processing constraint, predicts that if an *–s* genitive has been used within the previous 100 words, then the likelihood increases that the next genitive will be *–s*, whereas if the previous genitive was analytic, it is less likely that the next instance will be synthetic (Hinrichs & Szmrecsanyi, 2007:455; Szmrecsanyi, 2006:87–101). Finally, the determining predictor in the genitive system is *animacy*: the *–s* variant favors humans and other animate beings, as well as collective nouns, whereas concrete or abstract inanimate possessors largely retain the analytical *of* variant (Hinrichs & Szmrecsanyi, 2007; Mair, 2006a; Rosenbach, 2002, 2006).

The comparative. The default comparative mechanism in Old English was synthetic, *–ra/–ost* (the predecessors of modern *–er/–est*). The analytic mode, *more/most*, diffused slowly until approximately the end of the fourteenth century, then began to increase markedly in historical texts (Mitchell, 1985). However, the analytic forms did not continue to spread; their use peaked during Late Middle English (Pound, 1901:19), after which synthetic comparison experienced a resurgence (Kytö & Romaine, 2000:172). In the modern language, the majority of comparison is synthetic, and a linguistic "division of labor" obtains (Mondorf, 2009:127). Synthetic comparison is specialized for short (and frequent) adjectives, while analytic comparison favors longer (and less frequent) adjectives. As will be shown, this has synchronic consequences, particularly for the vernacular grammar.

Analyses of English adjective comparison distinguish between suppletive forms (*good, better, best*), which block analytic forms (**more good*), and nonsuppletive forms, where alternation between synthetic and analytic comparison is licensed

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

TABLE 2. *Overall distribution of comparatives*

Comparative Variant	%	*n*
Synthetic *–er/–est*	90.4	906
Analytic *more/most*	08.0	80
Double (synthetic + analytic)	01.6	16
Total		1002

Source: Data from D'Arcy (2014:234, Table 3).

(*happier, more happy*) (D'Arcy, 2014; Hilpert, 2008; Kytö, 1996a; Kytö & Romaine, 1997, 2000, 2006; Leech & Culpeper, 1997; Mondorf, 2003, 2009). Unlike typical linguistic variables, the variants may also co-occur (e.g., *more sillier*). Among nonsuppletive adjectives, however, certain forms are exceptional in that comparison is restricted to one mode or the other. These include adjectives whose citation form consists of more than three syllables (e.g., *more adaptable, more intellectual*) and adjectives ending in *–ic(al)* and *–ous*. Both are conventionally analytic (e.g., *more magical*) (e.g., Bauer, 1994; Hilpert, 2008; but see Mondorf, 2009:11, 37–38).

The Toronto materials include a total of 1002 comparative contexts for analysis, distributed as in Table 2 (127 speakers, ages 20–92 years).

The vast majority of comparatives in these materials are synthetic—just over 90%. Variation is not trivial, however, in that it is constrained by a robust set of contextual effects.

Phonologically, the choice between comparative variants is affected by the *final segment* of the citation form. Consonant clusters favor analytic comparison, as do final /li/ and /r/, whereas final /i/ and /l/ favor the synthetic mode (Hilpert, 2008; Lindqvist, 1998, 2000; Mondorf, 2003, 2009). *Stress* is also implicated; forms with final stress typically favor analytic comparison (Hilpert, 2008; Leech & Culpeper, 1997; Mondorf, 2003, 2009). The operation of *morphological complexity* is visible in the fact that complex adjectives favor the analytic comparative while monomorphemic ones favor the synthetic form (Mondorf, 2003, 2009). Syntactic constraints include *syntactic position*—attributives favor analytic comparison (Leech & Culpeper, 1997), *than complementation*—a following *than* favors the analytic form (Hilpert, 2008; Leech & Culpeper, 1997; Lindqvist, 1998), and *premodification*—a preceding degree modifier favors analytic comparison (Hilpert, 2008; Leech & Culpeper, 1997; Lindqvist, 1998). *Frequency* is also a key determinant because more frequent adjectives tend to be compared synthetically (Hilpert, 2008; Quirk, Greenbaum, Leech, & Svartvik, 1985). A final constraint is the *number of syllables* in the citation form of the adjective. Most of the variation occurs in disyllabic adjectives (Bauer, 1994; Hilpert, 2008; Kytö & Romaine, 1997; Leech & Culpeper, 1997; Mondorf, 2003, 2009). We will demonstrate the crucial repercussions this effect has for variation in the comparative sector.

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

TABLE 3. *Overall distribution of datives*

Dative Variant	%	*n*
Double object	80.3	1687
Prepositional	19.7	414
Total		2101

Source: Based on Tagliamonte (2014).

The dative. The dative has been variable between the double object (*give me some*) and the prepositional dative (*give some to me*) for centuries. The double object construction is the historic one. The prepositional dative was not widely available in Old English (Mitchell, 1985; Traugott, 1982) and appeared only sporadically in Early Middle English, limited to *give, grant, show, teach,* and *tell*. By Late Middle English, it was the majority form (Allen, 2006:214; Fischer & wan der Wurf, 2006:108; McFadden, 2002:1113), though in the Modern period, the prepositional dative has ebbed overall (Wolk et al., 2013:393).

The data were searched by seeking out verbs from current lists of verbs that subcategorize for the dative (Bresnan, Cueni, Nikitina, & Baayen, 2007; Wolk et al., 2013) as well as by examining every verb in the interview materials.[4] Any constructions that could alternate between the two variants without a change in meaning were included. Constructions in which the dative construction is likely a benefactive (e.g., *found me a job* vs. *found a job for me*) were excluded due to the fact that this is an independent construction. A number of other types (noun phrase shift, reverse double objects, cases of elided recipients and themes, etc.) outlined in Wolk et al. (2013:390–392) were also excluded because they do not alternate without a change in meaning.

This procedure resulted in 2101 tokens, distributed as in Table 3 (200 speakers, ages 17–92 years). As with comparatives, the variation is skewed toward the historical variant, in this case the double object.

A number of constraints are consistently reported to operate on the dative alternation, including *animacy, definiteness, type of subject* (e.g., noun phrase vs. pronoun), *construction length,* and *complexity*. According to the literature, these trends are so strong and consistent that their influence on the dative alternation has been referred to as "quantitative harmonic alignment" (Bresnan & Ford, 2010:181). Harmonic alignment is the existence of "a statistical pattern in which all else being equal, animate, definite, pronominal, discourse accessible and shorter arguments tend to precede inanimate, indefinite, non-pronominal, less discourse accessible or longer arguments in both dative constructions" (ibid.:183). Table 4 shows each dative construction. For the prepositional dative, the first element is a theme. In this case, animate, definite, pronominal, shorter, discourse accessible items come first (shaded). For the double object construction, the first element is a recipient. Here, too, animate, definite, pronominal, shorter items come first. The main point of contrast is the order of

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

TABLE 4. *Contrastive order in dative constructions*

Prepositional Dative			Double Object Construction	
Theme	To	Recipient	Recipient	Theme
Animate		Inanimate	Animate	Inanimate
Definite		Indefinite	Definite	Indefinite
Pronominal	To	Nonpronominal	Pronominal	Nonpronominal
Shorter		Longer	Shorter	Longer

the grammatical function, that is, theme versus recipient. This overarching pattern has been found in dative constructions across varieties of English, including American, Australian, New Zealand, and British varieties, as well as in written and spoken discourse (Bresnan et al., 2007; Bresnan & Hay, 2008; Collins, 1995; Grimm & Bresnan, 2009; Theijssen, ten Bosch, Boves, Cranen, & van Halteren, 2013; Thompson, 1995).

The relative pronouns. The English relative pronoun system has been variable since Old English, when at least three relativization strategies operated: the inflected demonstratives *se* and *seo*, invariant *þe*, and zero (i.e., deletion). The *wh–* forms entered the paradigm during Middle English, when the interrogative pronouns (*hwa, hwilc*, etc.) were introduced to the relative domain (Romaine, 1980:222). The last to develop a relativizing function was *who* (Romaine, 1980:223); its earliest attestation occurs in a Paston letter from 1426 (Rydén, 1983:127).

The development of the relative markers in subject position (*a woman who runs a exercise place*) has been studied extensively by Ball (1996). In this context, *which* was already well established by the 16th century and could be used for reference to both nonhuman and human antecedents. As Ball (1996) documented, however, the situation changed markedly during the 17th century, when the frequency of *who* gradually came to mark human antecedents, and *that* endured with nonpersonal antecedents. According to Ball (1996:247), this shift "laid the foundation for the modern dominance of *who* for personal subject restrictive relatives." Indeed, the restriction of *who* to human antecedents is a foundational and longitudinal characteristic of this variant (Romaine, 1982a:61; Rydén, 1983:128; Traugott, 1972:155–156).

In this analysis, we focus on restrictive relative clauses only, the locus of variation between *wh–* forms, *that*, and zero. This is the context that has received the bulk of attention in variationist sociolinguistics (e.g., Ball, 1996; D'Arcy & Tagliamonte, 2010; Guy & Bayley, 1995; Levey, 2006; Nevalainen & Raumolin-Brunberg, 2002; Tagliamonte et al., 2005; Tottie, 1995, 1997; Tottie & Rey, 1997).[5] As we will demonstrate, the longitudinal linguistic variability among relative pronoun forms and the systemic functional split between *who* and *that* endures in contemporary speech communities.

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

TABLE 5. *Overall distribution of relative markers*

Relative Marker	%	*n*
That	55.6	1790
Ø	26.7	859
Who	17.2	554
Whose	00.3	11
Which	00.1	4
Whom	00.1	2
Total		3220

Source: Data from D'Arcy & Tagliamonte (2010:389, Table 2).

To isolate the restrictive relative clauses and make our results comparable to earlier research, we followed the methods outlined in Tagliamonte et al. (2005:84–86), which are based on those of Tottie and Harvie (2000). According to this procedure, only those relative clauses that "serve to identify their antecedent" (Denison, 1998:278) were included in the analysis. Accordingly, a broad range of nonrestrictive clause types were excluded, most of which were identified by specific syntactic characteristics (e.g., antecedent = whole sentence or proper noun). Also excluded were adverbial relative clauses, as these allow *where* and *when*, which are functionally distinct from, and pattern to the exclusion of, other *wh–* relative markers (Tottie & Harvie, 2000:206). Finally, we excluded tokens of WHIZ deletion, in which the relative marker and the verb *be* are deleted, as well as incomplete utterances and clauses with resumptive pronouns. This procedure provided 3220 relative clauses for analysis, with an overall distribution as shown in Table 5 (82 speakers, ages 10–92 years).

The majority of restrictive relative clauses in these materials appear with *that*, nearly 56%. The next most frequent form, accounting for nearly 27% of the data, is the zero form. Notably, the *wh–* paradigm is almost categorically represented by *who*; occurrences of *whose, which*, and *whom* are negligible. A number of factors constrain the variation between *that, who,* and *Ø*.

Sentence structure is a critical determinant for the use of a zero relative. For example, the null form is highly favored in existential contexts and in possessive clauses (Ball, 1996; Quirk et al., 1985). On the other hand, *length of the relative clause* is implicated in the use of overt relative markers (*that, who*), which are favored in longer clauses (Cofer, 1975). This is presumably a processing constraint; the relative marker clarifies structure and aids comprehension of what follows. Because length does not map directly to complexity (Chomsky, 1975), Tagliamonte et al. (2005) also tested for *complexity*, distinguishing between simple relative clauses (verb + nonclausal argument) and complex ones (clausal argument). Length and complexity are not orthogonal (simple clauses may be either short or long but complex clauses are long); an interaction group is required to model these effects (see Tagliamonte et al., 2005:97–98). *Adjacency* has also been implicated, again as a processing constraint: overt *who* and *that* are

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

favored over zero when material intervenes between the matrix and subordinate clause (Cofer, 1975; Guy & Bayley, 1995). Finally, *type of antecedent* is relevant, though the effect does not appear to be universal. It has been suggested that indefinites favor *who* and *that* (Huddleston, 1971), whereas definites and pronominal antecedents favor zero (Guy & Bayley, 1995), and yet Tottie and Harvie (2000) reported that indefinites favor zero. Kautzsch (2002) reported a differential effect across dialects, and a lack of parallelism is likewise reported by Tagliamonte et al. (2005).

Perhaps the most challenging of the underlying constraints in the study of English relative pronouns is the interaction of the *syntactic status* of the gap or trace in the restrictive relative clause (subject, nonsubject), and the *animacy type* of antecedent (human, nonhuman, etc.). These two factors have strongly constrained the choice of relative marker since at least Early Modern English. Indeed, for every combination of antecedent type and grammatical role, there are very different frequencies of *that*, zero, and *wh–*. As a consequence, Ball (1996:233) suggested that the interaction of these factors creates "different populations" of variant choice across contexts, a partitioning of the relative sector that operates in the same way in formal writing and vernacular speech (Aarts, 1993; Ball, 1996; Guy & Bayley, 1995; Kautzsch, 2002; Quirk, 1957; Tagliamonte et al., 2005). In fact, much of the writing on the relative system in Early Modern and Modern English appears to take for granted that *who* is used for human (or personal) antecedents, *which* for nonhuman (or nonpersonal) antecedents, and *that* for either (Denison, 1998:278; Quirk, 1957:97–98; Swan, 1995:473). Although *which* essentially fails to surface in the Toronto materials; or in other vernacular materials for that matter (see Tagliamonte et al., 2005:87), the foundational association of *who* with human antecedents effectively blocks its use with nonhuman ones. This, as we will demonstrate, has crucial repercussions for variation in the restrictive relativizer sector.

CATEGORICITY WITHIN VARIATION

Based on the extant literature, these four variables are ideal for variationist inquiry because, as part of structured sets, they capture dynamic form/function asymmetry. They are also relatively frequent, which enables distributional and probabilistic modeling. Crucially, studies report not only active variation but also strong and multiple internal constraints operating on the choice mechanism. On genitives, Wolk et al. (2013:392) stated, "the genitive alternation exhibits robustly fluctuating variant proportions," which in an earlier draft they had summarized as "variability galore." On comparatives, Mondorf (2009:201) noted, "The true extent of variability in this area appears to have been underestimated in the past … comparative alternation is a showcase of grammatical variation." On the dative, Bresnan et al. (2007:75) suggested that "dative verbs have more syntactic flexibility than we thought, occurring more freely in alternative constructions." And on relative pronouns, Ball (1996:243) wrote, "We find a wide range of

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

variation both in the relative paradigm and in relative frequencies within a regional paradigm."

Such observations allude to the type of variation that yields valuable insights into grammar, linguistic change, mechanisms of change, their implications for social meaning, and so on. However, when we apply the techniques and methods of variationist sociolinguistics to vernacular spoken language, a different reality obtains. Variationist sociolinguistics advocates exhaustive scrutiny of data. Cross-tabulation exposes the intersection of factors, and inspection of individual items uncovers lexical effects, constructions, and formulae that may reflect the influence of frequency and/or semantic perseverance. When the grammatical variables under study are subjected to these procedures, it becomes clear that a nontrivial proportion of their characteristic aggregate variation is the artifact of collapsing distinct linguistic sectors (i.e., mixing apples and oranges). Specifically, for each variable there is at least one grammatical factor that splits the system, partitioning the variable context in vernacular use. For a variationist, such a result is at odds with both observations and expectations of robust and structured variability. We now turn to highlighting the linguistic partitions that underlie these variables.

The genitive

Consider the genitive. Animacy is crucially implicated in the distribution of genitive variants, splitting the system (Hinrichs & Szmrecsanyi, 2007; Mair, 2006a; Rosenbach, 2002, 2006). The key distinction is whether the head noun is human or not. Figure 1 shows that 89% of human possessors take the synthetic genitive and 92% of nonhuman possessors take the analytic genitive; there is extremely limited functional overlap between the genitive variants. This result, though dramatic, marks a well-established, historical distinction that has operated for centuries. However, the total specialization that is captured here has not been reported in studies of written texts and seems to exist only in vernacular speech, where human possessors—for all intents and purposes—categorically trigger the *–s* genitive (Hinrichs & Szmrecsanyi, 2007).

In addition, a second constraint is involved—possessor relation (Rosenbach, 2002; Wolk et al., 2013). Its effect renders the partitioning of the sector by animacy even more dramatic. As discussed, prototypical possessors (body parts, kinship terms, etc.) favor the *–s* genitive, and this holds in Toronto. Overall, 78% of prototypical possessors occur with *–s* ($n = 390$). When animacy is factored in, as in Figure 2, this variability largely disappears. Prototypical human possessors are realized with the synthetic form 96% of the time (*his friend's brother*), whereas prototypical nonhuman possessors are realized with the analytic genitive 95% of the time (*joints of the body*). In short, these two contexts pattern in complementarity and are virtually categorical. This is because animacy and possessor relation are not wholly orthogonal predictors. In Figure 2, prototypical possessors are mainly human ($n = 312$) and nonprototypical ones are mainly nonhuman ($n = 78$); for the full discussion, see

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

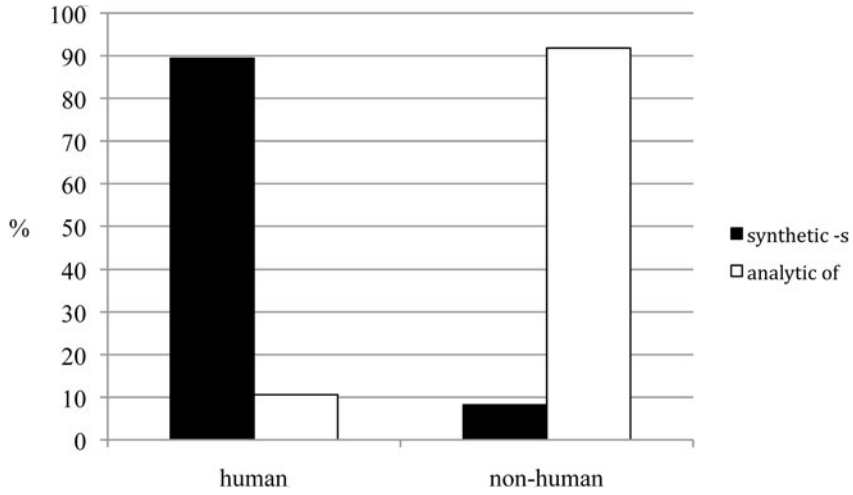


FIGURE 1. Distribution of genitive variants by animacy of the possessor, $n = 1{,}368$ (Jankowski & Tagliamonte, 2014:316, Figure 5).

Jankowski and Tagliamonte (2014:316–317). The result is a robust partition within the data.

On the other hand, the nonprototypical domain—though also heavily skewed—provides a margin of variability. Thus, although rigorous accountable methods were applied in the extraction and coding of the data, resulting in 1368 tokens, the analysis of variation is limited to nonprototypical possessors (social

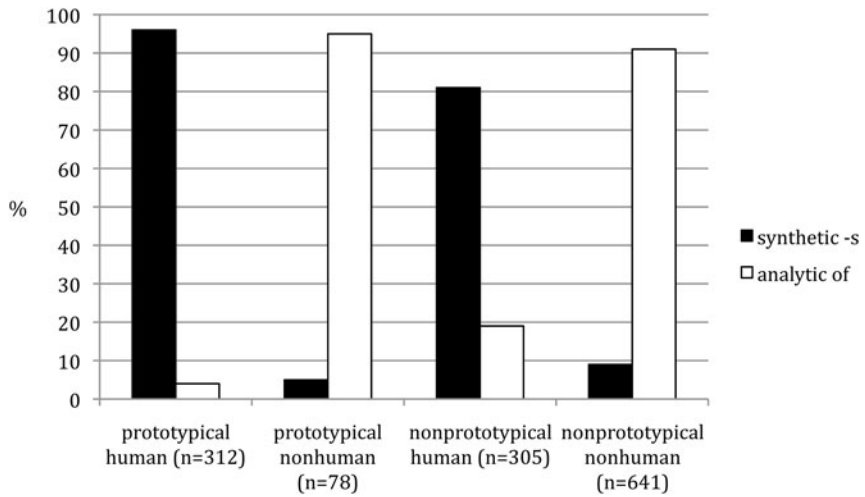


FIGURE 2. Distribution of genitive variants by animacy and possessor relation, $n = 1{,}336$ (based on Jankowski & Tagliamonte, 2014:317, Figures 6 and 7).

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

relations, states, and abstract possession), just 946 tokens. Even in this category, however, are forms that display categorical behavior; inanimates (e.g., *the wall of the apartment building, the edges of the problems*) take the *of* genitive at a rate of 99% ($n = 262$). This reduces the envelope of variation further, to just 684 tokens. In other words, half of the original dataset (48.8%) is not appropriate for probabilistic modeling. It is heavily partitioned, segmented, and effectively invariant.

The comparative

In the case of the comparative, Table 2 indicates that variation is weighted toward the historical synthetic variant (90%, $n = 1002$). This result is consistent with prior research, but it also masks underlying, categorical partitions within the data. As illustrated in Figure 3, the number of syllables in the citation form of the adjective divides the variants into distinct sectors.[6] Monosyllabic adjectives, which constitute the bulk of the data, take synthetic *–er* in 97% of cases ($n = 860$). Trisyllabic adjectives are compared categorically via analytic *more* ($n = 14$). The apparently sole locus of productive variation is disyllabic adjectives ($n = 128$).

This system presents an additional nuance. There are 1745 nonsuppletive comparative contexts in these materials, yet as discussed, only 1002 of them constitute possible sites of variation. These forms have the *potential* to alternate, but as Figure 3 illustrates, in these vernacular materials variation is restricted to disyllables.[7] This constitutes a mere 7% of the original extraction set, and just 13% of the subset. At best, comparative alternation can be considered a niche variable, as the window for variation (in the variationist sense) is quite narrow and is marginalized within the sector. This appearance of "variability" is further

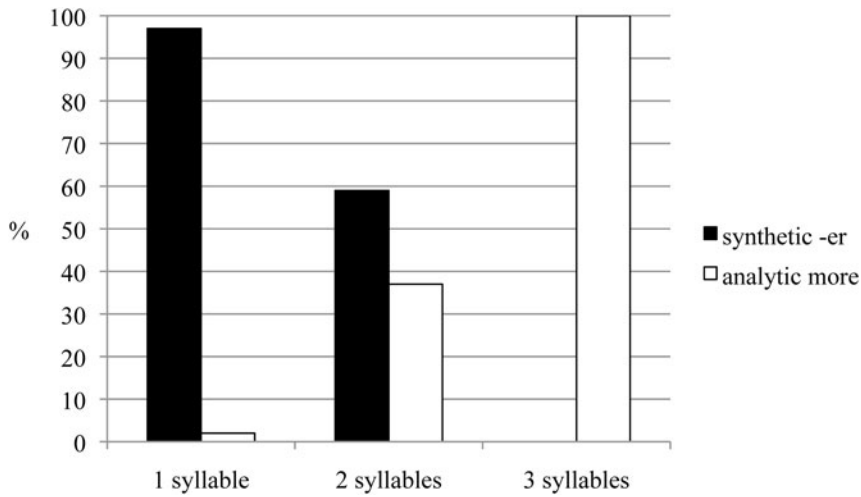


FIGURE 3. Distribution of comparative variants by syllable structure, $n = 1{,}002$ (D'Arcy, 2014:234, Table 3).

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

reduced when lexical items are considered. Of the 54 disyllabic adjective types that are attested in the data, only 8 actually alternate (*dirty* [$n = 2$], *friendly* [$n = 7$], *gentle* [$n = 3$], *healthy* [$n = 2$], *quiet* [$n = 6$], *quirky* [$n = 2$], *simple* [$n = 3$], and *sloppy* [$n = 2$]). From a variationist perspective, this is not the patterning expected of a bona fide linguistic variable. It is also notable that for this variable, super tokens—variant forms from the same speaker in the same stretch of discourse and the key evidentiary hallmarks of structured heterogeneity (Tagliamonte, 2006, 2012)—are exceptionally rare (see D'Arcy, 2014).

The dative

For the dative, the effect of pronoun versus noun phrase as recipient and theme is generally the strongest effect on variation (e.g., Bresnan & Ford, 2010:15, Table 5), exceeding all other predictors. Figure 4, however, exposes a striking distributional pattern. Variation is limited to constructions with noun phrases for both theme and recipient. Elsewhere, as summarized by Tagliamonte (2014), prepositional datives virtually never occur in constructions with a nominal theme and a pronominal recipient (*give her a coffee*), and yet they occur (nearly) categorically with pronominal themes (regardless of recipient type) (*give it to her; give the coffee to her*).

The redefined locus of variation is therefore theme noun + recipient noun constructions. Even here, however, structured heterogeneity does not fully obtain. One construction is virtually unattested in the Toronto materials (and in vernacular British data also; see Tagliamonte [2014])—animate theme with animate recipient (*give the dog to the vet*), while another changes the meaning to a locative and so is excluded altogether—inanimate recipient with animate

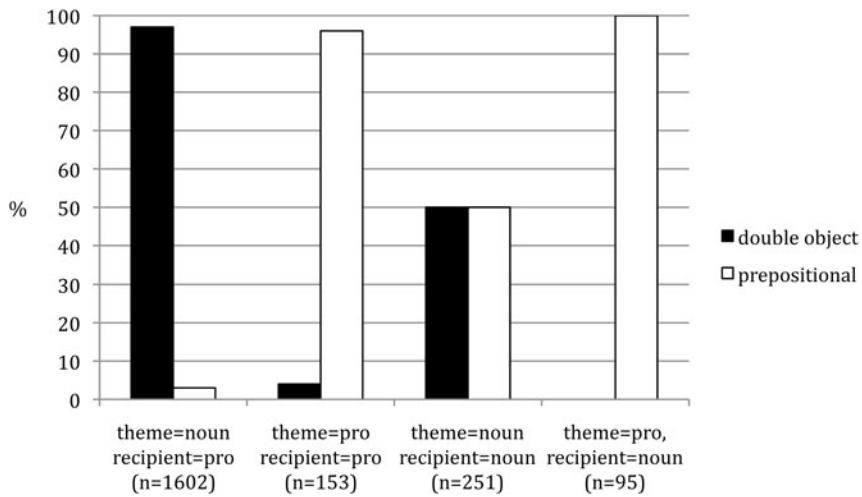


FIGURE 4. Distribution of dative variants by (pro)nominal status and theta role, $n = 2{,}101$ (based on Tagliamonte, 2014, Table 2).

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

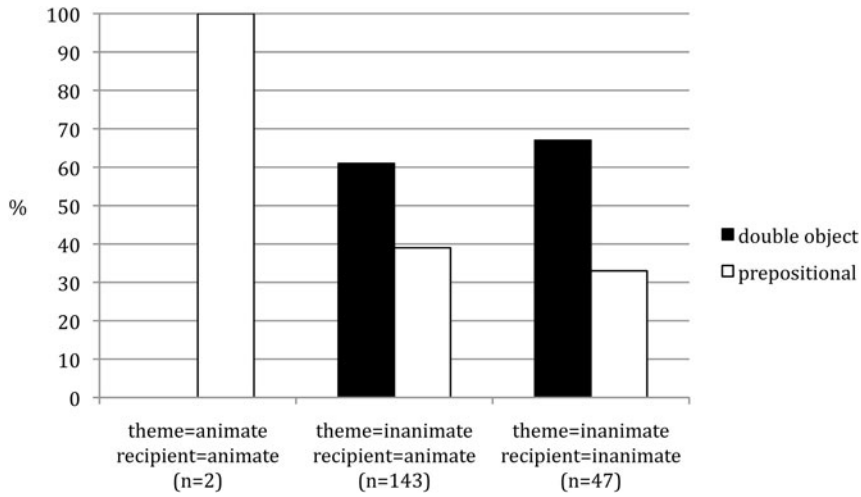


FIGURE 5. Distribution of dative variants by animacy of theme and recipient, $n = 192$ (based on Tagliamonte, 2014, Table 3).

theme (*sent a psychiatrist to the house*; see Wolk et al. [2013:390]). The consequence is that variation is circumscribed to a very small corner of the dative sector: inanimate themes, regardless of recipient type. This constitutes just 9% of the original extraction set ($n = 190$); the remainder is simply not variable.

The relative pronouns

The final linguistic variable that we consider is relativization. The overall results in Table 5 suggest a healthy alternation between *that*, a zero relative, and *who* in restrictive clauses. The remaining forms (*whose, which, whom*) together account for a mere 0.5% of the total dataset; we set them aside from this point. This overall picture seemingly indicates a ternary choice mechanism, and yet in actuality, the variation is effectively binary (D'Arcy & Tagliamonte, 2010). As is well established in the literature, this is because the relative pronoun forms have specialized according to syntactic function, subject (1a) and nonsubject (i.e., direct object, indirect object, object of preposition; [1b]). As illustrated in Figure 6, in subject position, only *who* and *that* vary productively; zero is rare. In nonsubject position, only *that* and zero are robust; *who* is rare. In other words, there is a sharp division across contexts, and only *that* is actively operational across the restrictive domain. The key role of syntactic function is a bona fide, longitudinal constraint for relative pronouns. That it should operate in these speech materials is not surprising, but the extent to which it divides the restrictive relative sector in these speech materials is striking. In more formal styles and written genres, the number of competing forms is greater.[8]

Underlying this main effect is a factor that again partitions the restrictive relative system: animacy of the antecedent. This predictor strictly concerns subject

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

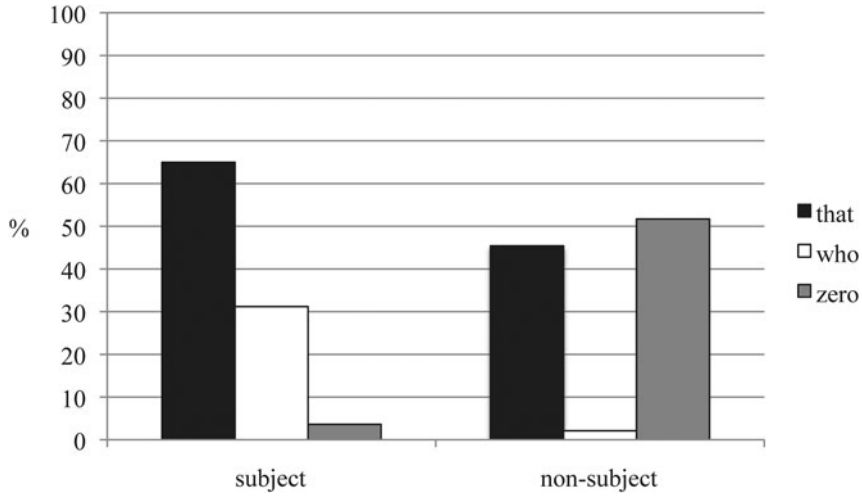


FIGURE 6. Distribution of restrictive relative pronouns according to syntactic function, n = 3,202 (based on D'Arcy & Tagliamonte, 2010:391, Table 3).

relatives, the domain of *who*, which is specified for animate reference (human and otherwise; cf. Figure 6). Figure 7 illustrates the defining role of animacy in this syntactic context.

Only in the animate contexts is variation robust. As with syntactic function, this observation is not new or unique; the restriction of *who* to animate antecedents is inherent to its development as a relative pronoun. However, this has a decisive

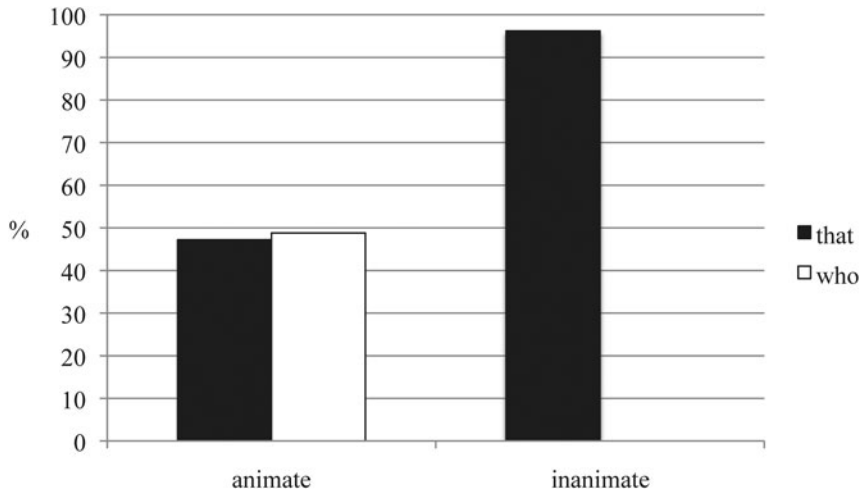


FIGURE 7. Distribution of SUBJECT restrictive relative pronouns according to animacy of the antecedent, n = 1,610 (based on D'Arcy & Tagliamonte, 2010:392, Table 4).

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

consequence for inanimate antecedents. In this category, over 96% of tokens are introduced by *that* ($n = 606$). In the variationist tradition, we exclude such near categorical contexts from our models as standard practice (Guy, 1988:131), and for good reason. In this case, animacy would dominate the model, not because it is a true (i.e., probabilistic) predictor but because it is the ultimate predictor, variation being restricted to a specific context (animate subject reference). Our main point is that there is a fundamental, systemic divide operating on the variable, one that produces categoricity within the vernacular grammar.

ATYPICAL SOCIOLINGUISTIC VARIABLES

Style represents a long and diverse continuum of contextual, interactional, and topical influences, but the vernacular is the baseline. It represents the model first acquired, before the standardizing influence of prestige and/or education becomes an overlay on usage, eroding the regularity of the constraints on linguistic variation (Labov, 1984:29). All four variables discussed here exhibit healthy overall form/function asymmetry (cf. Tables 1–3, 5), and for each the literature reports a complex and multidimensional variable grammar. However, the operation of key linguistic constraints lays bare dramatic contrasts and disproportionate distributions in vernacular data on a scale that is generally unfamiliar in variationist research. The genitive is split by animacy of the possessor; the comparative by the syllabic structure of the adjective; the dative by the interaction of (pro)nominal status and theta role; and the relative system by the syntactic status of the clause. Individual forms are then shunted to a restricted corner of the system, where they specialize for some specific function. Genitive *–s* is used for human possessors and *of* for inanimates; comparative *–er* is used for disyllables and is lexically specialized with few actively variable forms; the prepositional dative is restricted to nominal recipients and the double object construction to pronominal recipients; and relative *who* is used for human antecedents in subject relatives. In each case, the locus of variation is highly restricted, undermining the impression from overall frequencies of variable occurrence.

It is important to emphasize that observations regarding the restrictions on these variables are not new. Each constraint is well attested in the literature and none has emerged or evolved. Indeed, reexamination of the literature shows hints of partitioning for these variables—for instance, "division of labor" (Mondorf, 2009:127), "different populations" (Ball, 1996:233)—yet exhaustive delimitation of the contexts of variability or lack thereof is not fully explored. What is new is the realization that something outside the variationist norm obtains with these variables and that their patterns are either particular to or exaggerated in vernacular language. Despite occasional recognition of the potential for partitioning, the literature regularly reports the operation of variation and constraints in written materials.[9] However, as we have demonstrated, in the vernacular, variation is reduced. This contrast therefore challenges blanket

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

assumptions about structural heterogeneity in the vernacular. The question is: Are these linguistic variables exceptional or is there a principled explanation for their behavior in casual spoken language?

We suggest that the exceptional distributional qualities are not arbitrary and that they have the potential to lead to an expanded understanding of pathways and outcomes of linguistic change. As we will outline, there is reason to believe that analytic *of*, analytic *more*, the prepositional dative, and the *wh–* relative pronouns are united by a similar history. Each variant evolved from historical change that accelerated during Middle English through language contact and ultimately, through underlying grammatical mechanisms—in particular, analogy.

Analogy in language change

French and Latin influence were heightened during the Middle and Late Middle English periods. They were clerical and liturgical languages, respectively, with robust literary histories and traditions, whereas written traditions in English had faltered in the century after the Norman Conquest. Later in Middle English, a literary standard gradually emerged among the educated literate strata. These individuals were capable of shifting between French, Latin, and English easily. In this kind of social milieu, it is particularly important to consider the impact of multilingualism on developmental pathways of linguistic change.

The genesis of the analytic genitive comes late in the Old English period (Fischer, 1992:226; Mustanoja, 1960:74). However, it remained infrequent for centuries; vigorous competition did not develop until Middle Middle English, when the historical synthetic genitive lost ground dramatically to the innovative analytic form (Rosenbach, 2002; Rosenbach & Vezzosi, 2000; Thomas, 1931). Indeed, by the 14th century, the *of* form accounted for the vast majority of genitives in learned writing (Mustanoja, 1960:75). The synthetic genitive began to increase again during the 16th century; since then, variation between *–s* and *of* is said to be robust.[10] Notably, the rise of the analytic form aligns with the apex of French prestige in the history of English; due to its structural equivalency with the *de* genitive of Romance languages, its spread is believed to have been helped by contact with French (Fischer, 1992; Hinrichs & Szmrecsanyi, 2007; Mustanoja, 1960). Consistent with its Latinate roots, the *of* genitive has had formal connotations since at least the 17th century. Altenberg (1982:268–269) described it as being "generally restricted to situations demanding some degree of formality and politeness or when the language is raised to the level of rhetoric."

The history of the analytic comparative is strikingly similar to that of the analytic genitive. The default comparative mechanism in Old English was synthetic, *–ra*/*–ost*. The analytic mode, *more*/*most*, diffused slowly until approximately the end of the 14th century, when its use began to accelerate in historical texts (Mitchell, 1985), peaking during Late Middle English (Pound, 1901:19). A somewhat contentious issue concerns its origins. Historically, the analytic forms were argued to have been introduced during the 13th century (Kytö, 1996a;

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

Kytö & Romaine, 1997, 2000, 2006; Mitchell, 1985; Mustanoja, 1960; also Pound, 1901) as a change from above, introduced via analogy with either French (Morris, 1872) or Latin (Mitchell, 1985; Mustanoja, 1960; Pound, 1901; Strang, 1970). Mitchell (1985:84–85) reported a few examples of analytic comparison from Old English, but because these come from texts translated from Latin, they tend to be treated parenthetically. More recently, González-Díaz (2006) provided evidence that the analytic forms emerged in the second half of the ninth century, well before the Middle English Period. Under this analysis, they are a native development (see, too, Curme, 1931; Faiss, 1977; Mondorf, 2009). González-Díaz (2006:730–733) ruled out a connection with Latin, but she cannot preclude one with French (see also Mondorf, 2009:121). In Old English, use of analytic comparison was incipient. It was only during Middle English that its frequency escalated. Thus, while *more* and *most* may have evolved from below, the resulting forms were analogous to the (prestigious) French construction (on grammatical change in contact situations, see also Heine & Kuteva, 2005; Johanson, 2002). This may well have contributed to their spread, suggesting that notions of prestige were involved at a critical time in their development and diffusion. As with the analytic genitive, timing is particularly critical: analytic *more* was a development internal to the vernacular grammar, but it remained infrequent, low-level, and incipient until the Middle Middle English period (e.g., González-Díaz, 2006; Kytö & Romaine, 1997).

The history of the dative construction is less well documented, but it appears to follow a similar trajectory (Wolk et al., 2013:384–385). The prepositional construction was rare in Old English (Mitchell, 1985; Traugott, 1992) and a construction resembling the double object was variable (McFadden, 2002). According to De Cuypere (2010), the alternation was already subject to constraints of animacy and pronominality. Again it is the Middle English period when the new form accelerates (Fischer & wan der Wurf, 2006; Visser, 1963–1973), leading to (apparently) robust variation.

It is widely assumed that the loss of case distinctions during Middle English triggered the rise of the prepositional dative (e.g., Fischer & wan der Wurf, 2006; McFadden, 2002), yet Düringer (1923:30) argued that its sole function cannot have been to replace dative inflections (see also Allen, 2006; Rantavaara, 1962). Gerwin (2014:412) argued that the "real boost" was provided by contact with French and the borrowing of loan verbs with the preposition *à* (e.g., *obey, avail, command* [Visser, 1963–1973:624]). These loans increased the rate of the prepositional dative, but also buttressed, through analogy, the "already burgeoning" *to* dative with native verbs (see also Allen, 2006:215). As summarized by Gerwin (2014:143), in Late Middle English, the prepositional construction was "en vogue," a trend that was particularly evident in French-based texts (Trips & Stein, forthcoming). In short, there is mounting evidence from historical linguistics that the rising use of the prepositional dative in Middle English was influenced by French. In contemporary use, the prepositional dative is favored in more formal written registers, whereas the double-object construction is favored in registers approaching the oral end of the spectrum

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

(Bresnan et al., 2007; Wolk et al., 2013). This division may well reflect historical influences (cf. the analytic genitive).

In the case of relativization, the *wh–* forms were introduced when the interrogative pronouns (e.g., *hwa, hwilc*) came to be used as relatives, initially as generalizing forms in Old English but then expanding to refer to a clearly definable antecedent during Middle English (Romaine, 1980:223; also Mustanoja, 1960:191). The consensus in the historical literature is that their relativizer function gained prestige under Latin influence and then strengthened in Middle English under French influence (Mustanoja, 1960:199–200; Nevalainen & Raumolin-Brunberg, 2002; Romaine, 1982a:213). Romaine (1982a:234) suggested that "even though *wh–* is internal to the system (i.e., from the ME interrogative), its use as a relativizer can be thought of as a borrowing from a Romance acrolectal adstratum in Middle English."

This interpretation is supported by their linguistic pathway: The *wh–* relatives entered the language through the most complex and least frequently relativized syntactic positions in the case hierarchy (Romaine, 1980:233). Following from this, they were also initially confined to formal contexts (Jones, 1972:140). Meier (1967:280) argued that early epistolary uses came from models provided in Latin letter-writing manuals and notarial form books (see also Rydén [1983] on the critical role of French and Latin models for English constructions). Stylistic diffusion is sustained by Ball (1996:248), who reported "a clear lag" in the use of *wh–* relative pronouns between literary texts and spoken usage. This distinction is reflected in modern usage by the fact that their "infiltration" into the written language "can be seen as completed," whereas they have yet to affect the spoken language (Romaine, 1982b:212).

Romaine (1980:224) observed that the entry of *who* "had the effect of making English structurally more similar to Latin and French, which for reasons of prestige would have been desirable." The footprints of diachronic prestige are also evident in the longitudinal stability of social constraints on their use (D'Arcy & Tagliamonte, 2010). Timing, however, is critical. The *wh–* relatives remained incipient until roughly the Middle Middle English period (e.g., Ball, 1996; Romaine, 1980, 1982b). As we have outlined, this was a time when French had few native speakers in England yet "was quite commonly spoken and enjoyed the status of a prestige language" (Ingham, 2010:1), and Latin played a key superstratal role in Medieval England as well (e.g., Hickey, 2010; Mufwene, 2013; Munske, 1982).

Sociohistorical context

In the multilingual setting that characterized the educated elite of the Middle English period, English forms that were structurally analogous to their Romance counterparts would have carried social cachet. Indeed, numerous authors have argued, independently, that this is why the innovations in question diffused so rapidly during Middle English. They were used by the literate classes, groups who themselves held prestige in the broader sociopolitical context. That the

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

majority of Middle English speakers were illiterate may in part explain the anomalous patterning of these variables in the vernacular. The innovations were not widely available in the speech community. The fact that the major drivers were "top down" is evident in patterns of linguistic diffusion; each innovation spread through formal, literary genres, initially restricted in some cases to complex syntactic constructions. In short, each innovation appears to have been a formal, learned variant. The *wh–* relatives are the paradigm case, but there is evidence that the analytic genitive, comparative, and dative followed similar pathways. These forms remained embryonic until the social milieu changed, fostering their use in written channels as buttressed by the "Romance acrolectal adstratrum" (Romaine, 1982a:213). In contrast, the historical variant was the vernacular variant, and it populated matter-of-fact text types reflecting spoken or colloquial genres and registers.[11] These changes, which may have been evolutive in origin, were arguably adaptive in their actuation. In other words, a language external mechanism (social prestige) worked in tandem with a language internal mechanism (structural analogy) to propel the new forms into the written language (cf. Johanson, 2002).

In Middle English, formalized ideologies of standard language were incipient. By the end of the 16th century, however, normative ideas about English linguistic practice were entrenched (Kytö & Romaine, 2006:189), and complaints regarding (in)correct usage had replaced those concerning the relative status of French and English (Milroy & Milroy, 1985:26). Notably, these complaints habitually focused on the inadequacies of English relative to Latin, Greek, and "other more fashionable non-classical languages such as French and Italian" (Milroy & Milroy, 1985:27). By the middle of the 18th century, a written standard had emerged that was almost as invariable as that of prescribed Present Day English (Strang, 1970:107). Thus, the codification of the language resulted in a set of conventions for written language. However, it likely did not reflect how language was used by the majority of speakers at the time, especially in their everyday speech. Written language has always had special status as the bastion of "the standard"—this is where ideologies of *pure, correct, elegant, precise, informational, logical*, and the like are enshrined. Indeed, the function of written language (and the writing system) is to enforce or maintain standardization (Milroy & Milroy, 1985:51), a function that remains unchanged despite diachronic shifts in the underlying internal structure of written registers across the Modern period (Biber & Finegan, 1997).

CONCLUSION: CHANGE AND THE SPEECH COMMUNITY

Our understanding of variation in the genitive, comparative, dative, and relatives has been largely, though not exclusively, based on written data. Where speech has been considered, variation is either marginalized or favored in registers other than informal, spoken, conversational interactions (Börjars, Denison, Krajewski, & Scott, 2013; Guy & Bayley, 1995; Hilpert, 2008; Kytö, 1996b; Kytö &

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

Romaine, 1997; Mair, 2006a; Quirk, 1957; Szmrecsanyi & Hinrichs, 2008; Tottie, 1995; though see Hinrichs et al., 2014; Jankowski, 2012; Jankowski & Tagliamonte, 2014; Tagliamonte et al., 2005). We suggest that the predominance of written language data in the available information on these variables is not without consequence. In essence, the nature of the variation in written versus spoken data may be qualitatively different because, for these variables, variation in the vernacular is neither inherent nor heterogeneous.

Marked differences between written and spoken language are well attested and expected (Biber, 1988, 1995; Biber et al., 1999). Written English, particularly standard written English, is thought to be highly uniform and governed by prescription. In contrast, speech has long been recognized as the locus of synchronic variation (Labov, 1984; Milroy, 1992). Now consider the variables we have discussed here. The extant literature tells us that variation in these systems is robust and structured in written language. Our research on the same variables in spoken language exhibits something intrinsically different. This sets up a radical juxtaposition. The vernacular is widely held to be heterogeneous, yet these variables expose it as organized and surprisingly controlled. The written language is claimed to be rule-governed and regular, yet these variables depict it as dynamic and richly variable.

Change from below is considered the default type of linguistic change. It is "systematic," arising from the vernacular (Labov, 1994:78). Change from above is imposed from outside the community grammar. It is typically introduced by the dominant social groups, constituting borrowings from communities (broadly defined) that have higher prestige in the view of the innovators (Labov, 1994:78). Careful and formal language is affected first, reflecting a "superposed dialect learned after the vernacular is acquired" (ibid.). Both the historical trajectories and the contemporary patterning of genitives, comparatives, datives, and relativization suggest that the innovative forms were not diffused via the vernacular. A distinctive footprint remains; structured heterogeneity is evident in the written language, yet casual spoken language is marked by partitioned subsystems and minimal variation.

Labov (1994:78) remarked that in change from above, the new features are frequently "inconsistent with the vernacular system." Citing Fries and Pike (1949), he continued: "This inconsistency may prevent the change from above from being integrated with the rest of the system, and the borrowed element with its associated changes may form a separate subsystem for many generations." Elaborating on Fries and Pike (1949), Andersen (1973), and Labov (1994), the features we have considered here, where the systems are not aggregate but coexistent and reflect the separate origins of the variants, suggest that change from above—including those innovations simply interpreted as such—may never fully nativize. Instead, a functional divide may become entrenched in the vernacular and endure for centuries. In other words, it may be the case that change from below, incremented via transmission (Labov, 2007), results in what we have come to understand as structured, orderly heterogeneity. In contrast, when diffusion is the primary conduit of change, the outcome may be linguistic

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

variables without fully variable patterns in the vernacular, at least not within the accessible time frame of our corpora. In other words, a possible outcome of adaptive change is heterogeneity in the written language while the vernacular tolerates the overlay only in narrow sectors of the system.

Variation, after all, is not absent. Examples (1)–(4) come from the same speaker in the same stretch of discourse, the proverbial "super tokens" of linguistic variability and the paramount embodiments of form/function asymmetry (e.g., Tagliamonte, 2006). However, when we peer beneath grammatical variation for these features, the points of overlap are exceptional. We do not want to argue that separate subsystems must obtain in every instance of change from above (it appears, for example, that the vernacular can successfully squelch a prestige innovation, as in the case of *which* in restrictive relative clauses [Hinrichs, Szmrecsanyi & Bohmann, forthcoming]). Rather, we suggest that one of the possible outcomes is a marked speech/writing divide, with the written language bearing the hallmarks of structured heterogeneity and the spoken language exhibiting subsystemic partitioning. Perhaps one variant is considered more precise, or more elaborated, or more elegant. Perhaps one is considered more informational or more "correct." Perhaps one is used overwhelmingly by a particular (or prestigious) social group. These factors conspire for one variant to be taken up in the language "from above" and permit its use in written registers and/or formal styles for aesthetic impact. Such variables develop linguistic patterns in the written language, complete with probabilistic trends and complex constraints. The vernacular, however, can remain impenetrable, making the distribution and patterning of variants appear rigid.

Variationist sociolinguistics has spent half a century studying the structured heterogeneity of vernacular systems. We offer two new perspectives. First, we suggest that what *can* be variable grammatically need not be variable in the same way in every register of the language. The inherent structured heterogeneity of the vernacular has been a foundational assumption of the field, but these four features demonstrate that speech is not always the key locus of this kind of variation. Second, the vernacular can be remarkably resistant to fully incorporating imposed variants. It can be hospitable to prestige forms, but it also has a keen internal fortitude that keeps these variants straightforwardly separated. Such a view allows for the possibility of sociolinguistic partitioning—variation in the exponents of form, but (near) categoricity in the manifestation of function. If this scenario is accurate, new research questions arise full force. Crucially, consistent comparisons across written and spoken registers matched for date and place are needed. This will make it possible to substantiate the range of possibilities across variable systems and their social reflexes. Future research will then be able to evaluate whether we have simply tapped a stage in the trajectory of development of these variables or whether the dialectic between variation in the written and spoken language that is inferred by our work is a defining characteristic of linguistic heterogeneity in the social world.[12]

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

NOTES

1. Examples are from the Toronto English Archive (Tagliamonte, 2003–2006); the parenthetical information indicates subcorpus followed by speaker code, sex, and age at time of recording.

2. For the dative, where numerous communities are combined, we found no significant differences among them.

3. Other details can be obtained from the relevant papers. The analyses we report have been configured consistently with previous research.

4. This method ensured that newer coinages (e.g., *fax, email, zap*) were included.

5. Nonrestrictive relative clauses are virtually always introduced by a *wh–* form (e.g., Dekeyser, 1986; Lass, 1987; Poutsma, 1926; Visser, 1963–1973), and they are semantically distinct from restrictive ones, introducing supplementary information as opposed to propositional content (e.g., Huddleston & Pullum, 2002; Quirk et al., 1985).

6. Double comparatives account for less than 2% of the data. They have been omitted from the figure; cf. Table 2.

7. The lack of variability in these materials is surprising, and it is important to note that what we discuss here are bona fide patterns of use in vernacular data; they are not driven, for example, by variety or corpus size. As set out systematically in D'Arcy (2014), the primary explanatory factor is genre—unreflective vernacular speech versus other more reflective genres, both spoken and written.

8. A less clear-cut division of the subject domain obtains in some regional varieties (see e.g., Tagliamonte et al., 2005).

9. See D'Arcy (2014) for the effect of register on variation in the comparative.

10. Researchers report that use of the *–s* genitive has continued to increase in frequency throughout the 20th century, particularly in press registers (Hinrichs & Szmrecsanyi, 2007; Jucker, 1993; Leech, Hundt, Mair & Smith, 2009; Mair, 2006b; Rosenbach, 2002; Szmrecsanyi & Hinrichs, 2008).

11. Romaine (1982a:221), for example, noted that the introduction of the *wh–* restrictive relative pronouns resulted in "a 'squish' of two strategies which are opposed in stylistic meaning."

12. As one of our reviewers suggests, historical dialectology would be a profitable venue for further investigation (e.g., Wright [1898–1905], *The English Dialect Grammar*).

REFERENCES

Aarts, Flor. (1993). Who, whom, that and Ø in two corpora of spoken English. *English Today* 9:19–21.

Allen, Cynthia L. (2006). Case syncretism and word order change. In A. V. Kemenade & B. Los (eds.), *The handbook of the history of English*. Oxford: Blackwell. 201–223.

Altenberg, Bengst. (1982). *The genitive vs. the of-construction. A study of syntactic variation in 17th Century English*. Malmö: CWK Gleerup.

Andersen, Henning. (1973). Abductive and deductive change. *Language* 49:765–793.

Armstrong, Nigel, & Mackenzie, Ian E. (2013). *Standardization, ideology and linguistics*. Basingstoke: Palgrave-Macmillan.

Ball, Catherine. (1996). A diachronic study of relative markers in spoken and written English. *Language Variation and Change* 8:227–258.

Bauer, Laurie. (1994). *Watching English change: An introduction to the study of linguistic change in standard Englishes in the twentieth century*. London: Longman.

Beddor, Patrice S. (2009). A coarticulatory path to sound change. *Language* 85:407–428.

_____. (2012). Perception, grammars and sound change. In M.-J. Solé & D. Recasens (eds.), *The intitiation of sound change: Perception, production and social factors*. Amsterdam: Benjamins.

Biber, Douglas. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

_____. (1995). *Dimensions of regiser variation: A cross-linguistic comparison*. Cambridge: Cambrige University Press.

Biber, Douglas, & Finegan, Edward. (1997). Diachronis relations among speech-based and written registers in English. In M. Rissanen, T. Nevalainen, & L. Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*. Helsinki: Société néophilologique. 253–275.

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

Börjars, Kersti, Denison, David, Krajewski, Grzegorz, & Scott, Alan. (2013). Expression of possession in English: The significance of the right edge. In K. Börjars, D. Denison, & A. Scott

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

(eds.), *Morphosyntactic categories and the expression of possessio*. Amsterdam: John Benjamins. 123–148.
Bresnan, Joan, Cueni, Anna, Nikitina, Tatiana, & Baayen, R. Harald. (2007). Predicting the dative alternation. In G. Boume, I. Krämer, & J. Zwarts (eds.), *Cognitive foundations of interpretation*. Amsterdam: Royal Netherlands Academy of Arts and Sciences. 69–94.
Bresnan, Joan, & Ford, Marilyn. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86:168–213.
Bresnan, Joan, & Hay, Jennifer. (2008). Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118:245–259.
Chomsky, Noam. (1975). Questions of form and interpretation. *Linguistic Analysis* 1:75–109.
Cofer, Timothy. (1975). Performance constraints on relative pronoun deletion. *Linguistics* 157:13–32.
Collins, Peter. (1995). The indirect object construction in English: An informational approach. *Linguistics* 33:35–49.
Curme, George O. (1931). *A grammar of the English Language, III, Syntax*. Boston: D.C. Heath.
D'Arcy, Alexandra. (2014). Functional partitioning and possible limits on variability: A view of adjective comparison from the vernacular. *Journal of English Linguistics* 42:218–244.
D'Arcy, Alexandra, & Tagliamonte, Sali A. (2010). Prestige, accommodation and the legacy of relative *who*. *Language in Society* 39:389–410.
De Cuypere, Ludovic. (2010). A discourse based account of the Oled English double object alternation. *Sprachwissenschaft* 35:337–368.
Dekeyser, Xavier. (1986). English contact clauses revisited: A diachronic approach. *Folia Linguistica Historica* 7:107–120.
Denison, David. (1998). Syntax. In S. Romaine (ed.), *The Cambridge history of the English language, 1776–present day*. Cambridge: Cambridge University Press. 92–329.
Düringer, Hermann. (1923). Die Analyse im Formenbau des englischen Nomens. In W. Horn (ed.), *Beiträge zur Erforschung der Sprache und Kultur Englands und Nordamerikas*. Giessen: Im Verlag des Englischen Seminars der Universitet Giessen. 1–32.
Faiss, Klaus. (1977). *Aspekte der Englischen Sprachgeschichte*. Tübingen: Narr.
Fischer, Olga. (1992). Syntax. In N. Blake (ed.), *The Cambridge history of the English language*. Vol. 2: *1066–1476*. Cambridge: Cambridge University Press. 207–408.
Fischer, Olga, & wan der Wurf, Wim (2006). Syntax. In R. Hogg & D. Denison (eds.), *A history of the English language*. Cambridge: Cambridge University Press. 109–198.
Fries, Charles Carpenter, & Pike, K. (1949). Coexistent phonemic systems. *Language* 25:29–50.
Garret, Andrew, & Johnson, Keith. (2013). Phonetic bias in soudn change. In A. Yu (ed.), *Origins of sound change: Approaches to phonologization*. Oxford: Oxford University Press. 51–97.
Gerwin, Johanna. (2014). *Ditransitives in British English dialects*. Berlin: Walter de Gruyter.
González-Díaz, Victorina. (2006). The origin of English periphrastic comparatives. *English Studies* 87:707–739.
Grafmiller, Jason. (2014). Variation in English genitives across modality and genre. *English Language and Linguistics* 18:471–496.
Gries, Stefan Th. (2002). Evidence in linguistics: Three approaches to genitives in English. In R. M. Brend, W. J. Sullivan, & A. R. Lommel (eds.), *LACUS Forum XXVIII: What constitutes evidence in linguistics*. Fullerton: LACUS. 17–31.
Grimm, Scott, & Bresnan, Joan. (2009). Spatiotemporal variation in the dative alternation: A study of four corpora of British and American English. Paper presented at the Third International Conference Grammar and Corpora, Mannheim, September 22–24.
Guy, Gregory R. (1988). Advanced VARBRUL analysis. In K. Ferrara, B. Brown, K. Walters, & J. Baugh (eds.), *Linguistic change and contact*. Austin: Department of Linguistics, University of Texas at Austin. 124–136.
Guy, Gregory R., & Bayley, Robert. (1995). On the choice of relative pronouns in English. *American Speech* 70:148–162.
Harris, Alice C., & Campbell, Lyle. (1995). *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
Heine, Bernd, & Kuteva, Tania. (2005). *Language contact and grammatical change*. Cambridge: Cambridge University Press.
Hickey, Raymond. (2010). Language contact: Reconsideration and reassessment. In R. Hickey (ed.), *The handbook of language contact*. Malden: Wiley-Blackwell. 1–28.
Hilpert, Martin. (2008). The comparative: Language structure and language use. *English Language and Linguistics* 12:395–417.

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

Hinrichs, Lars, & Szmrecsanyi, Benedikt. (2007). Recent changes in the function and frequency of Standard English genitive constructions: A multivariate analysis of tagged corpora. *English Language and Linguistics* 11:437–474.
Hinrichs, Lars, Szmrecsanyi, Benedikt, & Bohmann, Axel. (2014). Which-hunting and the standard English relative clause. *International Society for the Linguistics of English 3 [ISLE]*. Zurich, August 24–27.
_____. (forthcoming). Which-hunting and the standard English relative clause. *Language* 91.
Huddleston, Rodney (1971). *The sentence in written English*. Cambridge: Cambridge University Press.
Huddleston, Rodney, & Pullum, Geoffrey K. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
Ingham, Richard (ed.). (2010). *The Anglo-Norman language and its contexts*. York: York Medieval Press/Boydell Press.
Jankowski, Bridget. (2012). *Cross-register language variation and change in Canadian English*. PhD dissertation. University of Toronto.
Jankowski, Bridget, & Tagliamonte, Sali A. (2014). On the genitive's trail: Data and method from a sociolinguistic perspective. *English Language and Linguistics* 18:305–329.
Johanson, Lars. (2002). Contact-induced change in a code-copying framework. In M. C. Jones & E. Esch (eds.), *Language change: The interplay of internal, external and extra-linguistic factors*. Berlin: Mouton de Gruyter.
Jones, Charles. (1972). *An introduction to Middle English*. London: Holt, Rinehard & Winston Inc.
Joseph, Brian D., & Janda, Brian D. (2003). On language, change, and language change—or, of history, linguistics, and historical linguistics. In B. D. Joseph & B. D. Janda (eds.), *The handbook of historical linguistics*. Maldon: Blackwell Publishers. 3–180.
Jucker, Andreas H. (1993). *The genitive versus the of-construction in newspaper language*. Heidelberg: Carl Winter.
Kautzsch, Alexander. (2002). *The historical evolution of earlier African American English: An empirical comparison of early sources*. Berlin: Mouton de Gruyter.
Kirby, James. (2013). The role of probabilistic enhancement in phonologization. In A. C. L. Yu (ed.), *Origin of sound change: Approaches to phonologization*. Oxford: Oxford University Press.
Kytö, Merja. (1996a). "The best and most excelletest way": The rivalling forms of adjective comparison in Late Middle and Early Modern English. In J. Svartvik (ed.), *Words: Proceedings of an international symposium*. Stockholm: Kungl. Vitterhets Historie och Antikvitets Akademien. 123–144.
_____. (1996b). *Manual to the diachronic part of the Helsinki Corpus of English texts: Coding conventions and lists of source texts*. Helsinki: Department of English, University of Helsinki.
Kytö, Merja, & Romaine, Suzanne. (1997). Competing forms of adjective comprison in Modern English: What could be more quicker and easier and more effective? In T. Nevalainen & L. Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rssanen*. Helsinki: Société Néophilologique. 329–352.
_____. (2000). Adjective comparison and standardisation processes in American and British English from 1620 to the present. In L. Wright (ed.), *The development of Standard English, 1300–1800: Theories, descriptions, conflicts*. Cambridge: Cambridge University Press. 171–194.
_____. (2006). Ajdective comparison in nineteenth-century English. In M. Kytö, M. Rydén, & E. Smitterbeg (eds.), *Nineteenth-century English: Stability and change*. Cambridge: Cambridge University Press. 194–214.
Labov, William. (1963). The social motivation of a sound change. *Word* 19:273–309.
_____. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
_____. (1970). The study of language in its social context. *Studium Generale* 23:30–87.
_____. (1971). Some principles of linguistic methodology. *Language in Society* 1:97–120.
_____. (1984). Field methods of the project on linguistic change and variation. In J. Baugh & J. Sherzer (eds.), *Language in use: Readings in sociolinguistics*. Englewood Cliffs: Prentice-Hall. 28–54.
_____. (1994). *Principles of linguistic change.* Vol. 1: *Internal factors*. Cambridge: Blackwell Publishers.
_____. (2006). *The social stratification of English in New York City*. 2nd ed. Cambridge: Cambridge University Press.
_____. (2007). Transmission and diffusion. *Language* 83:344–387.

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

_____. (2010). *Principles of linguistic change*. Vol. 3: *Cognitive and cultural factors*. Malden: Wiley Blackwell.
Lass, Roger. (1987). *The shape of English: Structure and history*. London: Dent.
Leech, Geoffrey, & Culpeper, Jonathan. (1997). The comparison of adjectives in recent British English. In M. Kytö, M. Rydén, & E. Smitterberg (eds.), *Nineteenth-century English: Stability and change*. Cambridge: Cambridge University Press. 194–214.
Leech, Geoffrey, Hundt, Marianne, Mair, Christian, & Smith, Nicholas. (2009). *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.
Levey, Stephen. (2006). Visiting London relatives. *English World-Wide* 27:45–70.
Lindqvist, Hans. (1998). The comparison of English disyllabic adjectives in /-y/ and /-ly/ in Present-day British and American English. In H. Lindquist, S. Klintborg, M. Levin, & M. Estling (eds.), *The major varieties of English: Papers from MAVEN 97*. Växjö: Acta Wexionensia. 205–212.
_____. (2000). 'Livelier /or/ more lively?' Syntactic and contextual factors influencing the comparison of disyllabic adjectives. In J. M. Kirk (ed.), *Corpora galore: Analyses and techniques in describing English*. Amsterdam: Rodopi. 125–132.
Ljung, Magnus. (1997). The *s*-genitive and the *of*-construction in different types of English texts. In U. Fries, V. Müller, & P. Schneider (eds.), *From AElfric to the New York Times: Studies in English corpus linguistics*. Amsterdam: Rodopi. 21–32.
Mair, Christian. (2006a). *Inflected genitives are spreading in present-day English but not necessarily to inanimate nouns*. Heidelberg: Universitätsverlag Winter.
_____. (2006b). *Twentieth-century English*. Cambridge: Cambridge University Press.
McFadden, Thomas. (2002). The rise of the *to*-dative in Middle English. In D. W. Lighfoot (ed.), *Syntactic effects of morphological change*. Oxford: Oxford University Press. 107–123.
Meier, Hand Heinrich. (1967). The lag of relative who in the nominative. *Neophilologus* 51:277–286.
Milroy, James. (1992). *Linguistic variation and change*. Oxford: Blackwell Publishers.
_____. (2003). On the role of the speaker in language change. In R. Hickey (ed.), *Motives for language change*. Cambridge: Cambridge University Press. 143–157.
Milroy, James, & Milroy, Lesley. (1985). *Authority in language: Investigating language prescription and standardisation*. London: Routledge and Kegan Paul.
Mitchell, Bruce. (1985). *Old English syntax*. Vols. 1, 2. Oxford: Clarendon Press.
Mondorf, Britta. (2003). Support for more-support. In G. Rohdenburg & B. Mondorf (eds.), *Determinants of grammatical variation in English*. Berlin: Mouton de Gruyter. 252–304.
_____. (2009). *More support for more-support: The role of processing constraints on the choice between synthetic and anlytic comparative forms*. Amsterdam: John Benjamins.
Morris, S., & Richard, L. L. D. (1872). *Historical outlines of English accidence*. London: Macmillan and Co.
Mufwene, Salikoko. (2013). Driving forces in English language contact linguistics. In D. Schreier & M. Hundt (eds.), *English as a contact language*. Cambridge: Cambridge University Press. 204–221.
Munske, Horst Haider. (1982). Die Rolle des Lateins als Superstratum im Deutschen und in andersen germanischen Sprachen. In U. Sture (ed.), *Die Leistung der Strategforschung in der Kreolistik: Typologische Aspekte der Sprachkontakte. Akten des 5. Symposiums über Sprachkontakt in Europa, Mannheim, 1982*. Tübingen: Niemeyer.
Mustanoja, Tauno F. (1960). *A Middle English syntax*. Helsinki: Société Néophilologique.
Nevalainen, Terttu, & Raumolin-Brunberg, Helena. (2002). The rise of relative *who* in early Modern English. In P. Poussa (ed.), *Relativisation on the North Sea littoral*. Munich: Lincom Europa. 109–121.
Ohala, John J. (1993). The phonetics of sound change. In C. Jones (ed.), *Historical linguistics: Problems and perspectives*. London: Longman. 237–278.
Pound, Louise. (1901). *The comparison of adjectives in the 15th and 16th centuries*. Heidelberg: Carl Winter.
Poutsma, Hendrik. (1926). *A grammar of Late Modern English*. Groningen: P. Noordhoff.
Quirk, Randolph. (1957). Relative clauses in educated spoken English. *English Studies* 38:97–109.
Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffry, & Svartvik, Jan. (1985). *A comprehensive grammar of the English language*. New York: Longman.
Rantavaara, Irma. (1962). On the development of the periphrastic dative in Late Middle English prose. *Neuephilologische Mitteilungen* 63:175–203
Rissanen, Matti. (1984). The choice of relative pronouns in 17th century American English. In J. Fisiak (ed.), *Historical syntax*. Berlin: Mouton de Gruyter. 417–435.
Romaine, Suzanne. (1980). The relative clause marker in Scots English: Diffusion, complexity and style as dimensions of syntactic change. *Language in Society* 9:221–247.

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

_____. (1982a). *Socio-historical linguistics: Its status and methodology*. Cambridge: Cambridge University Press.
_____. (ed.). (1982b). *Sociolinguistic variation in speech communities*. London: Edward Arnold.
Rosenbach, Anette. (2002). *Genitive variation in English: Conceptual factors in synchronic and diachronic studies*. Berlin: Mouton de Gruyter.
_____. (2003). *Aspects of iconicity and economy in the choice between the s-genitive and the of-genitive in English*. Berlin: Mouton de Gruyter.
_____. (2005). Animacy versus weight as determinants of grammatical variation in English. *Language* 81:613–644.
_____. (2006). Descriptive genitives in English: A case study on constructional gradience. *English Language and Linguistics* 10:77–118.
_____. (2008). Animacy and grammatical variation: Findings from the English genitive variation. *Lingua* 118:151–171.
Rosenbach, Anette, & Vezzosi, Letizia. (2000). *Genitive constructions in Early Modern English: New evidence from a corpus analysis*. Amsterdam: Benjamins.
Rydén, Mats. (1983). The emergence of *who* as a relativizer. *Studia Linguistica* 37:126–134.
Sankoff, Gillian. (2013). Longitudinal studies. In R. Bayley (ed.), *The Oxford handbook of sociolinguistics*. Oxford: Oxford University Press. 261–279.
Schlüter, Julia. (2001). Why worser is better: The double comparative in 16th- to 17th-century English. *Language Variation and Change* 13:193–208.
Stevens, Mary, & Harrington, Jonathan. (2014). The individual and the actuation of sound change. *Loquens* 1. doi:10.3989/loquens.2014.003.
Strang, Barbara M. H. (1970). *A history of English*. London: Methuen & Co.
Swan, Michael. (1995). *Practical English usage*. Oxford: Oxford University Press.
Szmrecsanyi, Benedikt. (2006). *Morphosyntactic persistence in spoken English: A corpus study at the intersection of variationist soicolinguistics, psycholinguistics, and discourse analysis*. Berlin: Mouton de Gruyter.
Szmrecsanyi, Benedikt, & Hinrichs, Lars. (2008). Probabilistic determinants of genitive variation in spoken and written English. In T. Nevalainen, I. Taavitsainen, P. Pahta, & M. Korhonen (eds.), *The dynamics of linguistic variation: Corpus evidence on English past and present*. Amsterdam: John Benjamins. 291–309.
Tagliamonte, Sali A. (2003–2006). Linguistic changes in Canada entering the 21st century. Research Grant. Social Sciences and Humanities Research Council of Canada (SSHRC). #410-2003-0005.
_____. (2006). *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
_____. (2007–2010). Directions of change in Canadian English. Research grant. Social Sciences and Humanities Research Council of Canada (SSHRC). #410-070-048.
_____. (2010–2013). Transmission and diffusion in Canadian English. Research grant. Social Sciences and Humanities Research Council of Canada. (SSHRCC). #410-101-129.
_____. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Malden: Wiley-Blackwell.
_____. (2014). A comparative sociolinguistic analysis of the dative alternation. In R. Torres-Cacoullos, N. Dion, & A. Lapierre (eds.), *Linguistic variation: Confronting fact and theory*. London: Routledge. 297–318.
Tagliamonte, Sali A., & Jarmasz, Lidia-Gabriela. (2008). Variation and change in the English genitive: A sociolinguistic perspective. *Linguistic Society of America*. Chicago, January 4.
Tagliamonte, Sali A., Smith, Jennifer, & Lawrence, Helen. (2005). No taming the vernacular! Insights from the relatives in northern Britain. *Language Variation and Change* 17:75–112.
Theijssen, Daphne, ten Bosch, Louis, Boves, Lou, Cranen, Bert, & van Halteren, Hans. (2013). Choosing alternatives: Using Bayesian Networks and memory-based learning to study the dative. *Corpus Linguistics and Linguistic Theory* 9:227–262.
Thomas, Russel. (1931). *Syntacticial processes involved in the development of the adnominal periphrastic genitive in the English language*. PhD dissertation. University of Michigan. Available at: http://deepblue.lib.umich.edu/handle/2027.42/78568. Accessed January 31, 2014.
Thompson, Sandra A. (1995). The iconicity of "dative shift" in English: Considerations from information flow in discourse. In M. E. Landsberg (ed.), *Syntactic iconicity and linguistic freezes*. Berlin: Mouton de Gruyter. 155–175.
Tottie, Gunnel. (1995). The man ø I love: An analysis of factors favouring zero relatives in written British and American English. In G. Melchers & B. Warren (eds.), *Studies in Anglistics*. Stockholm: Almqvist and Wiksell. 201–215.

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press

_____. (1997). Relatively speaking: Relative marker usage in the British National Corpus. In T. Nevalainen & L. Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*. Helsinki: Société Néophilologique. 465–481.
Tottie, Gunnel, & Harvie, Dawn. (2000). It's all relative: Relativization strategies in early African American English. In S. Poplack (ed.), *The English history of African American English*. Oxford: Blackwell Publishers. 198–230.
Tottie, Gunnel, & Rey, Michel. (1997). Relativization strategies in Earlier African American Vernacular English. *Language Variation and Change* 9:219–247.
Traugott, Elizabeth Closs. (1972). *A history of English syntax: A transformational approach to the history of English sentence structures*. New York: Holt, Rinehart & Winston.
_____. (1982). From propositional to textual to expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In W. P. Lehmann & Y. Malkiel (eds.), *Perspectives in historical linguistics*. Amsterdam: John Benjamins. 245–271.
_____. (1992). Syntax. In R. M. Hogg (ed.), *The Cambridge history of the English language.* Vol. 1: *The beginning to 1066*. Cambridge: Cambridge University Press. 168–289.
Trips, Carola, & Stein, Achim. (forthcoming). Contact-induced changes in the argument structure of Middle English verbs on the model of Old French. In E. Grossman, I. Serzants, & A. Witzlack-Makarevich (eds.), Special issue on valency and transitivity in contact, *Journal of Language Contact*.
Visser, Fredericus T. (1963–1973). *An historical syntax of the English language*. Leiden: E. J. Brill.
Weinreich, Uriel, Labov, William, & Herzog, Marvin. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (eds.), *Directions for historical linguistics*. Austin: University of Texas Press. 95–188.
Wolk, Christoph, Bresnan, Joan, Rosenbach, Anette, & Szmrecsanyi, Benedikt. (2013). Dative and genitive variability in Late Modern English: Exploring cross-constructional variation and change. *Diachronica* 30:382–419.
Wright, Joseph. (1898–1905). *The English dialect grammar*. Oxford: Clarendon Press.

https://doi.org/10.1017/S0954394515000101 Published online by Cambridge University Press