LEGACIES OF SYNTACTIC CHANGE IN A CONSERVATIVE DIALECT, YORK, ENGLAND*

SALI A. TAGLIAMONTE
UNIVERSITY OF TORONTO

LAURA RUPP
VRIJE UNIVERSITEIT AMSTERDAM

"Perhaps this knowledge will also lead us to appreciate the dying forms when we hear them — precious artefacts of linguistic history left in the linguistic archaeology of rural speech."

(Jankowski & Tagliamonte 2022)

ABSTRACT Several linguistic traditions have yielded important insights into syntactic change: these include historical linguists (e.g. Meillet 1967), historical dialectologists (e.g. Fisiak 1988), theoretically-informed dialect syntax (e.g. Henry 1995), and variationists (e.g. Labov 1969). We advocate an approach that draws strategically from the principles and techniques of these practices in order to refine the method for probing syntactic change, to employ vernacular speech as syntactic data, and to understand syntactic change in terms of structure as well as social and discourse context. We demonstrate how different perspectives provide essential and complementary contributions to understanding linguistic change. We use a case study of a linguistic feature that has been undergoing syntactic change through obsolescence in the variety of English spoken in York, England: the non-standard use of a zero form with singular count nouns (e.g. *They used to follow Ø river*) which we refer to as a 'zero definite article'. The path from the emergence of a syntactic feature towards its demise is typically a protracted

development. Historical (corpus) linguistics can trace the first attestations of a feature and its earlier meanings, historical dialectology its geographical distribution, and theoretically-informed research on dialect syntax can circumscribe its syntactic structure. We highlight the additional benefit of a variationist sociolinguistics approach, which focusses on community-based samples of spoken vernacular language data and quantitative methods. For example, in this case study we can document the last vestiges of the zero definite article in a conservative dialect and capture grammatical changes in the process of loss by comparing older to younger generations of speakers.

1 Introduction

Several distinct linguistic traditions have made important contributions to the study of syntactic change, each one developing its own assumptions, methods and techniques for researching and uncovering principles of syntactic change. Among these traditions are: historical (corpus) linguistics, historical dialectology, theoretically-informed dialect syntax, and language variation and change. We consider these traditions in light of a syntactic dialect feature that has been undergoing syntactic change through obsolescence in the dialect spoken in the city of York, England (YrkE) (Rupp & Tagliamonte 2019). We demonstrate that: (1) fading syntactic dialect features may undergo specific and systematic trajectories of loss and therefore, like innovations, provide a window on the nature of syntactic change in progress; and (2) the principles and techniques of different lines of evidence, historical and synchronic, text and speech, formal registers and vernacular, gender and generation all play a key part in exposing the evolution of an obsolescent feature from its emergence towards its loss. However, no single approach is definitive; different linguistic subdisciplines complement one another and not considering the insights provided by any one of them may lead to an analysis of syntactic change that is incomplete.

A well-known stereotype of York(shire) English is a phonologically reduced article, e.g. *during t'war* (Jones 2002, Tagliamonte & Roeder 2009). This paper focusses on a less-studied variant, a zero form, which occurs with singular count nouns, as in (1-3), from the York English Corpus (YEC; Tagliamonte 1996-1998). We refer to this feature as a 'zero definite article'.

(1)   That was during Ø war, forty-four, I think it was.   (bhamilton, 92)[1]

---

1 Codes in parentheses indicate the speaker's pseudonym, first initial, last name and age at the time of interview.

(2) That's near Ø river side you know and when Ø river come up it used
to flood up. (gwalton, 87)

(3) ... she'll pick me up from Ø bus stop or, you know, er go and pick Ø
kids up then from school and such like you know. (mlarkin, 70)

The zero article has been declining in (Standard) English and more generally
across languages that have (in)definite articles (Greenberg 1978, Harris 1980,
Keenan 2011). In the YEC (collected largely in 1997) speakers still used the
zero article in a definite function that has been lost from Standard English
(StdE). In this way, the zero definite article in YrkE provides evidence for
earlier patterns of change and incremental recession. While rates of the zero
variant were low in 1997, the proportion of use was stable from elderly to
adolescent speakers allowing us to spotlight characteristics of the declining
use of zero represented in this place and time.

Different linguistic subdisciplines generate key evidence for the analysis.
Historical corpus linguistics reports first attestations of the zero article, and
documents its use in English. Historical dialectology reports its geographi-
cal spread. Theoretically-informed dialect syntax circumscribes its syntactic
structure and distinguishes it from similar surface variants. Sociolinguistics
documents its occurrence by social factors and variationist methods assess the
statistical significance of social and linguistic patterns. What can be gained
by taking all the evidence into account?

The structure of the paper is as follows. Section 2 details the contributions
that distinct traditions have made to the study of syntactic change. Section
3 highlights the importance of consulting sociolinguistic corpora in study-
ing syntactic change, with a focus on fast disappearing features. Section 4
presents an analysis of the YrkE zero definite article and demonstrates how
incorporating insight from each tradition, and triangulating it effectively in a
single analysis, offers the best explanation of trajectories towards loss. Sec-
tion 5 provides a discussion and interpretation of the results for the study of
syntactic change.

2 DISCIPLINARY PERSPECTIVES

In this section, we review the theoretical practices of several sub-disciplines
of linguistics as they relate to syntactic change.

2.1 Dialectology, historical dialectology and dialect atlases

Dialectology and dialect atlases are an important source of linguistic data. Dialectology is founded in the idea that all languages change over time but change does not diffuse evenly. Dialectology focuses on the characteristics of language change that are influenced by distance, physical barriers and cultural, economic or social divisions (Boberg, Nerbonne & Watt 2017: i). Important for our discussion is the fact that much of the variation observed across localities is "the synchronic manifestation of diachronic processes, or changes in progress: newer forms, before being uniformly adopted, compete for dominance with older forms" (Boberg et al. 2017: 2). The distribution of older to newer forms often becomes a regional contrast, with urban centres advancing in language change in progress while rural places lag behind (e.g. Bailey, Wikle, Tillery & Sand 1993) inevitably leading to a core contrast between rural and urban dialects (Chambers & Trudgill 1980). Thus, differences from one dialect to the next provide important insight into linguistic change (Campbell 1998).

The long tradition of dialectology has left historical snapshots of language in dialect atlases and grammars. These materials offer the analyst a vast reservoir of real-time linguistic information. Dialect atlases comprise maps that record regional distribution of (mostly) lexical items and pronunciations; however, they can also offer insight into morphological alternations, syntactic constructions and expressions. In recent years, the original hard copies of many of these atlases have been digitized and made available in searchable databases with easy access for research purposes, for example the FRED corpus.[2]

This type of data presents many difficulties for the study of syntax. Importantly, the material is (mostly) elicited from what has been referred to as "NORMs", non-mobile, older, rural male-presenting individuals, and so restrict the sample to a particular sector of society. The data represents a single style of speech that is relatively formal or careful, which is well known to restrict a person's vernacular and thus privileges standard uses. Further, the data in traditional dialect atlases, at least up to the middle of the 20th century, is at least one step removed from what the person actually said, since the data were written down from memory, not mechanically recorded. An additional problem is that the rapport between interviewer/fieldworker and informant influences the nature of the data, so there is no consistency across individual fieldworkers or projects. Perhaps most thorny for syntactic inquiry is that most atlases do not record the structural details of forms or constructions.

---

2 Fred Corpus of English Dialects: https://fred.ub.uni-freiburg.de. Accessed 1-15-22 6:58 am.

2.2   Language Variation and Change

Language variation and change is a disciplinary subfield of linguistics that developed from dialectology, anthropology and statistics. From dialectology came the foundational concept of variation; from anthropology the fact that variation is influenced by social and cultural groups; and statistics provided the quantitative methodology. Studies of language variation and change incorporate nuances of variation within speech communities based on the socio-economic hierarchy such as wealth, power, education, and the ethnic and social identity of speakers (Boberg et al. 2017). These factors tend to dictate how language is perceived. The variety spoken in an urban centre with economic wealth and cultural dominance is perceived as prestigious while that spoken in a rural area where there is less wealth, culture and often a less educated population and working-class occupations (e.g. farming, fishing, mining) is perceived as inferior. A language variation and change approach is ideal for identifying features of language that are below the radar of the prestige language and that often go unnoticed. The object of inquiry is the vernacular, "real language in use" (Milroy 1992: 66). The basic unit of analysis is the "linguistic variable", the alternation or layering of forms, in basic terms "two or more ways of saying the same thing" (Labov 1972: 8). Variationist methods offer insight into syntactic structure in many ways, particularly the discovery of constructions not present in mainstream varieties. The standard language may have one form; a peripheral dialect, many and related dialects, a full range of slightly different variants. In addition, the focus on patterns of linguistic or social factors adds deeper insight. For example, the standard relative pronoun for relative clauses with human/animate head NPs is *who*; but in vernacular uses *that* is more common and peripheral dialects often preserve non-standard variants such as *which*, *at*, *what* and others (Tagliamonte, Smith & Lawrence 2005). Sociolinguistic corpora tend to be stratified by broad social factors enabling usage frequencies and patterns to be revealed based on the social characteristics of individuals. When this information is quantitatively examined together with linguistic contexts of use, i.e. one variant may be more favorable in one context over another and even more favorable in that context among a certain social sector (e.g. male vs. female; educated vs. less educated), this information adds insight to the progression of change, how it originated and the process of innovation or obsolescence.

The difficulty faced by studies of language variation and change is that corpora of naturally occurring vernacular speech are typically small making it difficult, if not impossible, to study syntactic phenomena which tend to be rare. Another complication is that vernacular speech from rural communities is difficult to obtain. Fieldwork, data documentation, and corpus build-

ing are all fraught with complications (e.g. Tagliamonte 2007). This makes most extant dialect corpora haphazard by nature, unevenly stratified by social characteristics, location and other factors, which compromises systematic comparison.

2.3   Syntactic Variation and Change

The study of syntactic variation and change arose from scholars of syntax and historical linguistics who were interested in studying linguistic change. The study of Old English (OE) was particularly motivating as its grammar contrasts dramatically with contemporary English. In particular, the relatively free word order of OE alongside robust case inflection and complex person/number agreement was refashioned over the centuries, leading to variation in the positioning of arguments. What sets syntactic variation and change studies apart from historical linguistics more generally is its reliance on historical texts, poetry, court and merchant records, etc. Additional insight that comes out of this line of study is how syntactic units map onto metrical units in poetry or are influenced by register (e.g. Pintzuk & Kroch 1989). More recent research in syntactic variation and change has been developing the idea that there are important influences on syntactic change from grammatical weight, complexity and information structure (e.g. Traugott 2007, Taylor & Pintzuk 2011). Importantly, some researchers introduced quantitative methods to the study of historical documents, beginning with an early analysis of *do* support (Ellegård 1953), which was taken up by Kroch (1989a) and reanalyzed using statistical modeling.

Syntactic variation and change research discovered the famous constant rate effect (Kroch 1989b), a finding that refuted the commonly held notion that when a linguistic form appears frequently in a context, then that is the context in which the form first appeared and in which it advances most rapidly. Instead, syntactic change proceeds in an S-curve with the same rate across all contexts of use (Kroch 1989b *et seq*), a function of the fact that, when there is a single underlying change in the grammar, constructions change together as a manifestation of that change. Where differences in frequency across varying contexts of use exist, this is explained by independent factors such as style or information structure (e.g. Taylor & Pintzuk 2011).

Studies of syntactic variation and change based on corpora present many of the same challenges for analysis as synchronic sociolinguistic corpora. The data is often limited in various ways, sparse, syntactic features are often scant and sometimes non-existent, and the analyst can never be sure if the absence of a feature is due to lack of grammaticality or contextual, register or other chance factors. Moreover, some periods of history or regions are better rep-

resented than others. Across the board, conversational interaction is virtually absent in historical texts, so the focus of attention is based in comparisons across registers and text types. Nevertheless, the study of syntactic variation demonstrated that historical processes taken in light of theoretical issues can astutely inform new explanations of language change.

2.4   Historical syntax

Historical corpus linguistics, also called historical syntax, is founded in the study of grammatical change in historical data. Interrogating historical material presents significant challenges for syntactic inquiry since, as Kroch (1989b: 199) points out, "grammatical analysis depends on negative evidence, the knowledge that certain sentence types are unacceptable", but this information is not available in historical material. Instead, two assumptions guide the inquiry: 1) that general principles of language change will apply; and 2) that absence of a grammatical construction in a substantial data source can be interpreted as ungrammaticality in that source. Guided by these assumptions, the benefit of historical inquiry is that the time course of language change can be observed and studied, offering the analyst a view of the process as it is happening. In so doing, the principles of organization involved in how languages change over time, and whether and how these developments can be explained by general principles of language and/or historical or cultural influences, can be obtained. The language data of historical syntax is typically based on materials collected from religious and epistolary texts from across the centuries of a language. Indeed, corpus-building has been one of the major contributions of historical syntax.

   Another contribution of historical syntax has been to use statistical tools for analyzing corpus data. In this tradition, syntactic theory provides the framework and models of linguistic change but the evidence comes directly from the data.

2.5   Theoretically-informed dialect syntax

Theoretically-informed dialect syntax studies language-internal principles that govern the forms and structures of dialect grammars. This line of inquiry is a relatively recent branch of linguistics because formal syntax initially abstracted away from variation and change in adult language, ((in)famously) assuming an "ideal speaker-listener, in a completely homogeneous speech community" (Chomsky 1965: 3). In this area of research the mechanisms that were held responsible for syntactic variability across (standard) languages can be extended to account for variation across dialects of a language. Amongst

these are the Principles and Parameters approach which assumes that syntactic variation is caused by particular settings of binary parameters (e.g. Henry 1995 on Belfast English). Another approach is that the source of variation lies in different values for sets of features (e.g. Adger 2006 on Buckie English). Formal insights into syntactic structure are fruitful for identifying a syntactic variable and the grammatical contexts which condition its use. Once established, structural relationships to other variables and any changes in grammatical conditioning can shed light on the development of a feature.

Following pioneering work by Cornips & Corrigan (2005), many collaborative projects between syntacticians and variationists have arisen (Adger & Smith 2005, Rupp & Tagliamonte 2019, Rupp & Britain 2019, and others). The literature has also been augmented by theoretically-informed studies of dialects of many languages, e.g. the Scandinavian Dialect Syntax project (Vangsnes & Johannessen 2019). Consistent with other disciplines, syntacticians studying dialects are also challenged by the scarcity of many grammatical contexts in natural speech data, for example, questions (because interviewees typically respond to, but do not ask questions). For this reason, natural speech data have frequently been complemented by judgment data for a set of test sentences (e.g. Cornips & Jongenburger 2001). Another complication is that what appear to be grammatical features may involve discourse-pragmatic nuances (e.g. Cheshire 1982), which cannot be captured by test sentences and require close scrutiny of the discourse in which the feature occurs.

2.6 Syntactic change: the case of obsolescing dialect features

There has been a building body of research on disappearing dialect features (e.g. Wolfram 1995, Wolfram & Schilling-Estes 1995): e.g. dialectal realizations of the diphthongs /ay/ and /aw/ (Schilling-Estes & Wolfram 1999); subject-verb concord, copula absence, *was* regularization (Hazen 2002); verbal -s and *old* preterite forms (e.g. Jankowski & Tagliamonte 2017). It might be expected that moribund dialect features would be found to simply 'dissipate'; that is, die (Schilling-Estes & Wolfram 1999: 487); however, waning features may have a different fate. Instead of fading away, they may 'concentrate' (Schilling-Estes & Wolfram 1999: 488), for example, by retreating to a particular social group: middle-aged and younger speakers cling to the dialectal realizations of the diphthongs /ay/ and /aw/ compared to the older speakers (p. 494, 506). Concentration may also happen though endurance of a feature in a particular linguistic context, e.g. verbal –s in linguistic environments where it was historically first favoured, i.e. relative clauses (Jankowski & Tagliamonte 2017: 525), as in 4 and 5.

(4) And the farmers that *was* thrashing just let them run.   (astarz, 72)

(5) Mangoes which *is* like a turnip.   (bthickson, 70)

Schilling-Estes & Wolfram 1999: 517 conclude that communities that have preserved obsolescing dialect features may show linguistic distinctiveness. They also conclude that while the decline of features may accelerate in the face of loss or external social pressure, these changes will exhibit known patterns of socio-linguistic conditioning and processes of change.

In the remainder of this paper, we explore how the linguistic sub-disciplines outlined in this section complement one another. We pose the following question: to what extent are extant analyses of syntactic change complete if they fail to take multiple analytic perspectives into account?

3 SITUATING THE ANALYSIS

In historical linguistics, the study of peripheral dialects is considered one of the most informative means to shed light on the origins and development of languages (Anttila 1989; Hock 1986: 442). Peripheral dialects by their nature are geographically removed and/or isolated by social and/or political circumstance and thus tend not to be affected by the changes in comparable cohorts in mainstream communities. Data from such dialects has traditionally been the province of dialectology (Chambers & Trudgill 1980), in which the practice has been heavily descriptive, with a focus on lexis. In contrast, historical and comparative linguistics (Anttila 1989) have typically resorted to historical written sources and formal theories for their interpretation, while focusing on syntax. Recent research suggests that synchronic dialect data can contribute gainful evidence for many types of linguistic inquiry, including the study of language structure and meaning, language contact and language endangerment, in addition to the more common studies relating to linguistic change over time and space. Moreover, researchers have shown that contemporary dialect phenomena provide important insights into the links between diachronic and synchronic linguistics. More recently, the value of sociolinguistic data to language typology has been explored (e.g. Chambers 2003). All these studies highlight the important contribution that sociolinguistic corpora from peripheral dialects can make to ongoing developments in syntax, language variation and change, dialectology and corpus linguistics.

Conservative dialects tend not to participate in ongoing linguistic change at the same rate as others. Meillet (1967) observes that "very often it is sufficient to arrange facts geographically to understand their history." Although synchronic dialects cannot exactly mirror varieties of English as they were

spoken in earlier days, many peripheral communities even in the early 21st century retain features of 17th and 18th-century nonstandard English vernaculars, providing at least a partial 'snapshot' of earlier stages in the history of the language.

This paper mines the linguistic assets of one such dataset, a corpus collected in York, England (in large part) in 1997 (Tagliamonte 1998). The goal of this project was to obtain a representative sample of the vernacular speech of the city and to document its linguistic features. At the time of data collection, York was a distinct community for at least two reasons: 1) it had been protected from economic upheaval, population growth and rebuilding found in other English cities; and 2) 19th century migrations to the city were predominantly from local dialect areas (Armstrong & Mackenzie 2013: 145), making the variety of English in York conservative in character.

What makes the YEC important for the study of language change and syntactic change in particular is that it is substantial (132 hours and 1 million words) and comprises conversational speech (see Tagliamonte 1998). The socio-historical nature and informality of the language materials combine to provide the language analyst with the greatest chances of tapping into an earlier stage in the history of a language in a particular place (see Hay & Foulkes 2016 for the importance of reverie in eliciting traditional variants). In the following section, we focus on one of these features, the zero definite article.

4 A case study in York English

This section presents an analysis of the zero definite article in YrkE, synthesizing historical, dialectological, structural and synchronic evidence focussing on its grammatical development.

4.1 The zero definite article

Speakers of YrkE have a non-standard use of the zero article with definite singular count nouns, as in (1-2) above and (6-7).

(6)   I crashed Ø car twice.   (rmitchell, 20)

(7)   … she said: "They painted Ø door"   (hphillips, 72)

StdE has bare nouns, but these are virtually restricted to (8) proper names, (9) non-count (mass and abstract) nouns, (10) generic or kind-denoting plural nouns, and (11) a circumscribed set of nouns naming places such as *church*, *school*, and *prison* (see e.g. Soja 1994).

(8)   *John* is nice.

(9)   *Green tea* is healthy.

(10)   *Firefighters* are brave.

(11)   They got married in *church*.

4.2   Historical perspective

The study of Old English (OE) by historical linguists can shed light on the development of the English article system. OE did not have a separate definite or indefinite article, just as the ancestors of many languages that currently have articles (Greenberg 1978). Instead, OE had an elaborate demonstrative system. The distal demonstrative *se* ('that'; used here as a shorthand for the range of gender, case, and number forms of the distal demonstrative) could be deployed as a marker of definiteness, and it ultimately gave rise to the definite article. However, OE *se* did not function in the same way as contemporary *the*; for example, it had a less wide usage in definite contexts (Mitchell 1985: 133-135), allowing for zero forms elsewhere. Some historical linguists postulate that designated definite article use developed later in the (Early) Middle English (ME) period (e.g. Millar 2000) while others (e.g. Allen 2019: 131) assume the occurrence of a "definite determiner" in OE already. Importantly, in the development toward the present-day definiteness marking system, there has been "a constant loss of ground on the part of the zero-form" (Christophersen 1939: 84). In a somewhat later development, the indefinite article derived from the numeral *one* (Givón 1981).

The replacement of the zero article happened incrementally in the history of the English language according to the referential properties of nouns in discourse. First, the definite article initially marked demonstrative-type situational and discourse-anaphoric reference (e.g. Pass me *that/the stool*, please; Lyons 1999: 164), thereafter extending to other definite contexts. The indefinite article initially marked specific indefinite nouns, before spreading "to a broader (and progressively less referentially 'strong') non-specific range of noun phrases" (Keenan 2011: 2). Second, in some syntactic structures and types of noun, the zero article endured a long time, as detailed in Jespersen (1954). For example, in the set of nouns with unique denotations, some acquired the definite article as early as OE (e.g. *sun*, *moon*), some in the Middle English period (e.g. *world*, *earth*) while a few have not to this day (e.g. *God*) (see also Christophersen 1939: 77, 87, 182). In contemporary StdE, the (in)definite articles have infiltrated the old zero contexts to such a degree

that its usage has largely been reduced to noun phrases that are referentially the least strong; namely, non-count nouns and generic/kind-denoting nouns (9-10), and the strongest; namely, proper names (8), which Jespersen (1954: 417-418) has described as involving "familiarity so complete that no article is needed." Following Greenberg (1978), Harris (1980) and Keenan's (2011) work on the grammatical development of articles in a typologically diverse range of languages, (in)definite articles may eventually oust zero.

In the context of the diachronic and cross-dialectal decline of the zero article in English, it seems likely that the more extensive use of zero in YrkE as compared to StdE indicates conservative usage, showing an earlier stage in its path towards loss. However, in principle it could also be an innovation. Evidence can be sought in the geographical distribution of zero definite articles in dialect atlases.

4.3 Dialectological perspective

Dialectology studies have commonly reported that use of the zero definite article is restricted to the far south east peninsula of Yorkshire, Holderness (e.g. the Survey of English Dialects (SED) of Orton & Halliday 1962). The absence of reports in York may be due to the fact that surveys such as the SED are based on data from only one informant per location. As Beal (2010: 49) argues: "[i]t is possible that these uses of the definite article were formerly more widespread throughout the North, or, indeed, that they exist elsewhere in the North but have not been captured by other dialect surveys." For example, the zero form has been reported in Bolton, Greater Manchester (Shorrocks 1999: 23-31) and Lancaster (Hollmann & Siewierska 2011: 42).

Therefore, it appears that the zero definite article is largely reported in dialects in the North of England. Since northern English varieties are known to be "remarkably conservative" in comparison to other UK varieties (see Ihalainen 1994: 261, 271, fn. 4), this evidence makes it likely that the occurrence of the zero article in YrkE is not a contemporary development but a remnant from the once wider use of zero in the history of English.

4.4 Theoretically-informed perspective

While evidence from historical linguistics and dialectology suggests that the zero article in YrkE is a holdover from OE, in principle it may also constitute a maximally lenited form of the phonologically reduced definite article also observed in YrkE. Note that a reduced article has been documented in a systemic set of phonologically restricted forms, namely [t, θ, ʔ] (e.g. Jones 2002). In the framework of theoretically-informed syntax, the YrkE zero article and the

reduced article potentially have different structures: (1) a "truncated" NP-structure from which a definite article and hence a DP-layer is absent; and (2) a full DP with a phonologically lenited D-head, respectively. It seems possible that the first kind of structure occurred in OE because at the time "the definite article is needed only when ... definite and indefinite readings are not disambiguated by other means" (Crisma referred to by Allen 2019: 132); for example, by the uniqueness of the referent. Accordingly, we assume that the DP layer may be absent when the referent of a noun is identifiable without the formal marking of a definite article.[3] We also assume that in the history of English, when the use of the definite (and indefinite) article became progressively compulsory across definite (and indefinite) nouns, those nouns projected a DP-layer, leaving mass- and kind-denoting nouns as well as proper nouns as NPs in contemporary mainstream English. However, tokens of reduced articles are likely to be full DPs, in which the presence of D is not dependent on information in the discourse, but D may be phonologically lenited to a greater or lesser extent.

The possibility of a truncated DP has previously been considered by, amongst others, Oosterhof & Rawoens (2017) and Radford (forthcoming) for the case of articleless nouns in newspaper headlines – to our knowledge, the only type of zero article with definite singular nouns that has been considered in the theoretical syntax literature on English. For Dutch headlines, Oosterhof & Rawoens (2017: 221) assume that where proper names are used without an overt article, "the category D ... is not projected at all", referring to Zwart (2009) for discussion of a "dynamic" account of syntactic structure. In research on null subjects in written registers like diaries and instructions, Haegeman (1997) shows that assuming truncation of a structural layer can be tested empirically: it should go together with the absence of linguistic material or processes associated with that layer. For example, truncation of a CP-layer from IP should go together with the impossibility of preposing arguments (cf. *More problems don't need __; Thrasher 1977 cited in Haegeman 1997: 249). Accordingly, Radford (forthcoming) has argued that a truncation analysis cannot be extended to account for article-less nouns in the register of newspaper headlines. For example, he points out that if NPs are treated as having no functional superstructure in headlines, it is unclear how they could contain prenominal modifiers; e.g. *Free kits <are> to be available for every adult in England* (p. 293). On the same grounds, a truncated DP-analysis of the YrkE zero article might seem untenable from the following usage: e.g. *it was blazing hot summer-time and **whole** road was like this out here* (rfielding,

---

3 This might be a pragmatic version of Longobardi's (1994: 653) syntactic Last Resort principle that overt articles are used "if no synonymous raising derivation is available".

87).[4] However, contemporary developments in theoretically-informed syntax postulate that various grammatical categories occur in separate phrases, such as an Adjective Phrase, where premodifiers may be in a lower projection than DP and so survive truncation of DP.[5]

Thus, theoretically-informed syntax allows for the possibility that the zero article is a separate grammatical phenomenon and not simply a reduced variant. Supportive to this interpretation is that Tagliamonte & Roeder (2009) demonstrated through statistical modelling that the zero definite article is conditioned differently by grammatical context than reduced variants. One key result was that the zero article is favoured with definite nouns with historically attested patterns of zero, e.g. geographic places, as in (12), adding still more weight to the possibility that it is a relic form.

(12) There was only us three left in **Ø street**   (eburritt, 82)

Discourse-pragmatic factors have played a role in the historical trajectory of English article use; therefore, further evidence for the nature of the YrkE zero form may come from variationist study that looks into its discourse-pragmatic behaviour.

4.5 *Synchronic perspective*

Rupp & Tagliamonte (2019) examined the grammar of the zero definite article further and focussed on specific nouns of two types: 1) nine nouns covering a range of types known to have varied with respect to the zero article historically; and 2) nine nouns that had no such record, excluding tokens that can appear with the zero article in StdE (e.g. fixed expressions as in *They are just playing cat and mouse* (kdilks, 26)).[6] This procedure provided a total of 1335 tokens, all of which would take the definite article in StdE but which could occur with a full form, reduced form, or zero form in YrkE as in (13-15).

(13) and it's about two minutes walk down Ø road.   (pgregory, 23)

(14) "Blown this one" and I got Ø job!   (rjones, 50)

(15) down onto Fulford Road where Ø police station is.   (ddavis, 19)

---

4 Reg Fielding's age is extrapolated to what his age would have been in 1997 when the YEC was collected. He was born 1905 and interviewed by Bob LePage in 1986 at age seventy. For further details, see Tagliamonte & Roeder (2009: 463n1).
5 We would like to thank an anonymous reviewer for pointing this out to us.
6 See Rupp & Tagliamonte (2019) for details of method, analysis and results.

Quantitative analysis considered the zero form out of all other realizations of the article and the distribution of variants by social (sex, education, occupation, age group) and linguistic (type of noun, grammatical category, and modification of the noun) factors.

Rupp & Tagliamonte (2019) went on to examine discourse-pragmatic influences, highlighted in syntactic variation and change studies generally (see section 2.3) and for variable article usage in formal syntax (e.g. Weir 2009) and variationist work (Sharma 2005). Specifically, Rupp & Tagliamonte (2019) coded for Prince's (1981) main categories of 'information status', in particular: discourse-old, hearer-old information and discourse-new, hearer-old information, as in nouns with discourse-anaphoric and unique reference, as in (16-17), respectively:

(16) *The station* got badly damaged. And *Ø station* caught fire.

(ajackson, 66)

(17) I many a time go and sit outside *Ø Minster*.   (eburrit, 82)

The results showed that the zero form made up 17 percent (N=225) of the data. Every age group had the zero article and, in each group, males were the more frequent users. In females, the level of zero article usage declined from 23 percent to between 7 and 9 percent from the oldest generation (75+) to the other age groups, suggesting "ongoing obsolescence of an older form" (Tagliamonte & Roeder 2009: 18). In contrast, the youngest generation of males used the form at similar rates to the very oldest generation (35 vs. 37 percent), with a drop in the middle-age cohort (of 11 percent). A U-shaped pattern of this type suggests a social pattern of age-grading, whereby middle-aged speakers suppress non-standard forms. Another significant finding for social factors was that individuals with less education favoured the zero article, and for the linguistic factors, that 'information status' exerted the strongest effect, with discourse-new, hearer-old information favouring the zero article whereas discourse-old, hearer-old information disfavoured.

In the history of English (see section 4.2), the zero article was first lost from definite nouns with discourse-old, hearer-old (situational and discourse-anaphoric) reference. In a subsequent stage it was lost with other definite reference nouns, and in StdE it survives (at least in the prescriptive canon) in proper names (and in abstract and kind-denoting nouns). The finding that speakers favour the zero definite article with discourse-new, hearer-old reference suggests that the YrkE zero definite article is a synchronic reflex of an earlier use from the second, intermediate stage in its diachronic trajectory, in which referents that are known between the speaker and the hearer were

not overtly marked. Given this pathway, the results of distributional patterns reflect the development of the zero article linguistically, which is echoed in apparent time. The combined social and historical evidence leads to the interpretation of the zero definite article in YrkE as a relic from earlier English and illuminates its course of change in English more generally. The analysis concurs with Schilling-Estes & Wolfram (1999) that obsolescing dialect features show normal patterns of variation and change and further adds 'retention' of an obsolescing syntactic feature to the inventory outlined in section 2.6.[7]

5 Discussion

We have now summarized the procedures we have taken to document, analyze and explain a disappearing syntactic feature. We examined the zero definite article in YrkE, using evidence from historical, dialectological, structural and sociolinguistic insights. Each tradition contributes important information; however, no single method offers the explanatory power of a composite approach. From historical linguistics, we determined that the zero form is the oldest of the articles and historically had a wider range of use before *the* and *a* took over. However, to the best of our knowledge, historical linguistics has not distinguished a stage in the history of English in which the zero variant occurs with discourse-new, hearer-old nouns. From dialectology, we established that the form is attested in conservative northern English dialect areas, an indication of relic status. That it was not reported in York is undoubtedly due to the fact that the dialectological method had a restricted scope of survey techniques and coverage. From syntax, we substantiated that the zero variant can be associated with a distinct grammatical structure. However, a syntactic approach did not generate conclusive evidence that it is not a phonologically maximally lenited form. From variationist sociolinguistics we performed quantitative analysis of social and linguistic factors and used the cultural setting to offer an explanation: 1) We assessed the variable patterns of usage and discovered that the zero variant shows unique syntactic and discourse-pragmatic conditioning. The YrkE speakers favour the zero definite article with nouns that are discourse-new, hearer-old. We postulate that this contrast reflects a previous stage in the deployment of the zero article that YrkE speakers have maintained while the form was lost from StdE; 2) From the socially stratified data, we demonstrated that the zero variant is a low frequency feature that is declining in apparent time among female

---

7 The zero definite article is attested in two L2 English varieties and dialects of Dutch and French. It seems the form behaves somewhat differently there and/or has innovative uses (e.g. Sharma 2005). This shows that there is more to learn about zero definite articles and that they may still have a future.

speakers in the community, used mostly by male speakers with less education, but showing near identical frequency of usage between the older and the younger generation, demonstrating maintenance of usage. Even the middle-aged groups who rarely use the zero article due to age-grading show the same discourse-governed use of the form; 3) Finally, the endurance of the zero definite article in YrkE can be interpreted as a facet of local allegiance to place and identity (see Tagliamonte & Roeder 2009, Tagliamonte 2017).

To conclude, finding remnant syntactic features is a feat in itself and as linguists we should take pride in the precious documentation of such features because many will become moribund in the future. While we have focussed on the zero definite article in this case study, there are many more *artefacts* of linguistic history in the fading dialects of contemporary rural speech (e.g. double demonstratives (Rupp & Tagliamonte 2022), zero indefinite articles (Bigelow & Tagliamonte to appear) and many others). In the case of the zero definite article, bringing together interpretations of the structure, patterning, frequency and geographic and social embedding of the zero definite article from multiple disciplines offers the best possible means of coming to a fulsome explanation; in the case at hand, of the grammatical nuances of the English zero definite article and its development over the past centuries in a context that offers a glimpse of the past.

References

Adger, David. 2006. Combinatorial variability. *Journal of Linguistics* 42. 503–530.

Adger, David & Jennifer Smith. 2005. Variation and the minimalist program. In Leonie Cornips & Karen Corrigan (eds.), *Syntax and variation: reconciling the biological and the social*, 149–178. Amsterdam: John Benjamins.

Allen, Cynthia L. 2019. The definite article in Old English: evidence from Ælfric's Grammar. In Nuria Yáñez-Bouza, Willem B. Hollmann, Emma Moore & Linda van Bergen (eds.), *Categories, constructions and change in English syntax*, 130–146. Cambridge: Cambridge University Press.

Anttila, Raimo. 1989. *Historical and comparative linguistics*. Amsterdam: John Benjamins.

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

Bailey, Guy, Tom Wikle, Jan Tillery & Lori Sand. 1993. Some patterns of linguistic diffusion. *Language Variation and Change* 5. 359–390.

Beal, Joan. 2010. *An introduction to regional Englishes: dialect variation in England*. Edinburgh: Edinburgh University Press.

Bigelow, Lauren & Sali A. Tagliamonte. To appear. Where have all the articles

gone? Bare nominals in Marmora and Lake, Ontario. *American Speech* to appear.

Boberg, Charles, John Nerbonne & Dominic Watt. 2017. *The handbook of dialectology*. Hoboken, NJ: John Wiley & Sons.

Campbell, Lyle. 1998. *Historical linguistics*. Edinburgh: Edinburgh University Press.

Chambers, J. K. 2003. Dynamic typology and vernacular universals. In Bernd Kortmann (ed.), *Dialectology meets typology: dialect grammar from a cross-linguistic perspective*, 127–145. Berlin: Mouton de Gruyter.

Chambers, J. K. & Peter Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.

Cheshire, Jenny. 1982. *Variation in an English dialect: a sociolinguistic study*. Cambridge: Cambridge University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Christophersen, Paul. 1939. *The articles: a study of their theory and use in English*. Copenhagen: Einar Munksgaard.

Cornips, Leonie & Karen P. Corrigan. 2005. *Syntax and variation: reconciling the biological and the social.* Amsterdam: John Benjamins.

Cornips, Leonie & Willy Jongenburger. 2001. Het design en de methodologie van het SAND-project. *Nederlandse Taalkunde* 3. 215–232.

Ellegård, Alvar. 1953. *The auxiliary* do*: the establishment and regulation of its use in English*. Stockholm: Almqvist and Wikwell.

Fisiak, Jacek. 1988. *Historical dialectology: regional and social*. Berlin: Mouton de Gruyter.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals of human language, vol. 3: word structure*, 47–82. Stanford: Stanford University Press.

Haegeman, Liliane. 1997. Register variation, truncation, and subject omission in English and French. *English Language and Linguistics* 1. 233–270.

Harris, Martin B. 1980. The marking of definiteness: a diachronic perspective. In Elizabeth C. Traugott, Rebecca Labrum & Susan C. Shepherd (eds.), *Papers from the fourth international conference on historical linguistics*, 75–86. Amsterdam: John Benjamins.

Hay, Jennifer & Paul Foulkes. 2016. The evolution of medial /t/ over real and remembered time. *Language* 92. 298–330.

Hazen, Kirk. 2002. Identity and language variation in a rural community. *Language* 78. 240–257.

Henry, Alison. 1995. *Belfast English and Standard English: dialect variation and*

parameter setting. Oxford: Oxford University Press.

Hock, Hans Henrich. 1986. *Principles of historical linguistics*. Berlin: Mouton de Gruyter.

Hollmann, Willem B. & Anna Siewierska. 2011. The status of frequency, schemas and identity in Cognitive Sociolinguistics: a case study on definite article reduction. *Cognitive Linguistics* 22. 25–54.

Ihalainen, Ossi. 1994. The dialects of England since 1776. In R. Burchfield (ed.), *The Cambridge history of the English language, vol. 5: English in Britain and overseas: origin and development*, 197–270. Cambridge: Cambridge University Press.

Jankowski, Bridget L. & Sali A. Tagliamonte. 2017. A lost Canadian dialect: the Ottawa valley 1975-2013. In Tanja Säily, Arja Nurmi, Minna Palander-Collin & Anita Auer (eds.), *Exploring future paths for historical sociolinguistics*, 239–274. Amsterdam: John Benjamins.

Jankowski, Bridget L. & Sali A. Tagliamonte. 2022. 'He come out and give me a beer but never seen the bear': vernacular preterites in Ontario dialects. *English World-Wide* 43. 267–296.

Jespersen, Otto. 1954. *A modern English grammar on historical principles, part VII: syntax*. London: George Allen and Unwin Ltd.

Jones, Mark J. 2002. The origin of definite article reduction in northern English dialects: evidence from dialect allomorphy. *English Language and Linguistics* 6. 325–345.

Keenan, Caitlin. 2011. Greenberg revisited: diachronic development of article systems and the structure of DP. Paper presented at the 13th International Diachronic Generative Syntax conference, University of Pennsylvania.

Kroch, Anthony. 1989a. Function and grammar in the history of English: periphrastic *do*. In Ralph Fasold & Deborah Schiffrin (eds.), *Language change and variation*, 133–172. Washington, DC: Georgetown University Press.

Kroch, Anthony. 1989b. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1. 199–244.

Labov, William. 1969. Contraction, deletion, and inherent variability of the english copula. *Language* 45. 715–762.

Labov, William. 1972. *Sociolinguistic patterns*. Pennsylvania: University of Pennsylvania Press.

Longobardi, Giuseppe. 1994. Reference and proper names: a theory of N-movement in syntax and Logical Form. *Linguistic Inquiry* 25. 609–665.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Meillet, Antoine. 1967. *The comparative method in historical linguistics*. Paris: Librairie Honoré Champion.

Millar, Robert McColl. 2000. *System collapse, system rebirth: the demonstrative pronouns of English 900–1350 and the birth of the definite article*. Bern: Peter Lang.

Milroy, James. 1992. *Linguistic variation and change*. Oxford: Blackwell.

Mitchell, Bruce. 1985. *English syntax, vol. 1: concord, the parts of speech, and the sentence*. Oxford: Clarendon Press.

Oosterhof, Albert & Gudrun Rawoens. 2017. Register variation and distributional patterns in article omission in Dutch headlines. *Linguistic Variation* 17. 205–228.

Orton, Harold & Wilfrid J. Halliday. 1962. *Survey of English Dialects (B): The basic material, vols. 1.1 and 1.2: the six Northern counties and the Isle of Man*. Leeds: E. J. Arnold.

Pintzuk, Susan & Anthony Kroch. 1989. The rightward movement of complements and adjuncts in the Old English of Beowulf. *Language Variation and Change* 1. 115–143.

Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In Peter Cole (ed.), *Radical pragmatics*, 223–255. New York: Academic Press.

Radford, Andrew. Forthcoming. *Sentence structure: an intermediate course in syntax*. Cambridge: Cambridge University Press.

Rupp, Laura & David Britain. 2019. *Perspectives on a variable morpheme: let's talk about -s*. London: Palgrave.

Rupp, Laura & Sali A. Tagliamonte. 2019. "They used to follow Ø river": the zero article in York English. *Journal of English Linguistics* 47. 279–300.

Rupp, Laura & Sali A. Tagliamonte. 2022. Loss and emergence: (double) demonstratives and indefinite *this* in Ontario dialects. *Journal of Pragmatics* 187. 186–198.

Schilling-Estes, Nathalie & Walt Wolfram. 1999. Alternative models of dialect death: dissipation vs. concentration. *Language* 75. 486–521.

Sharma, Devyani. 2005. Language transfer and discourse universals in Indian English article use. *Studies in Second Language Acquisition* 27. 535–566.

Shorrocks, Graham. 1999. *A grammar of the dialect of the Bolton area, part 2: morphology and syntax.* Bamberger Beiträge zur Englischen Sprachwissenschaft. Frankfurt am Main: Peter Lang.

Soja, Nancy N. 1994. Evidence for a distinct kind of noun. *Cognition* 51. 267–284.

Tagliamonte, Sali A. 1996-1998. Roots of Identity: variation and grammaticization in Contemporary British English. Economic and Social Sciences Research Council (ESRC) of Great Britain. Reference #R000221842.

Tagliamonte, Sali A. 1998. *Was/were* variation across the generations: view from the city of York. *Language Variation and Change* 10. 153–191.

Tagliamonte, Sali A. 2007. Representing real language: consistency, trade-offs and thinking ahead! In Joan Beal, Karen Corrigan & Hermann Moisl (eds.), *Using unconventional digital language corpora, vol. 1: synchronic corpora*, 205–240. Basingstoke: Palgrave.

Tagliamonte, Sali A. 2017. Changing places: tracking innovation and obsolescence across generations. In Chris Montgomery & Emma Moore (eds.), *Language and a sense of place: studies in language and region*, 15–37. Cambridge: Cambridge University Press.

Tagliamonte, Sali A. & Rebecca V. Roeder. 2009. Variation in the English definite article: socio-historical linguistics in t'speech community. *Journal of Sociolinguistics* 13. 435–471.

Tagliamonte, Sali A., Jennifer Smith & Helen Lawrence. 2005. No taming the vernacular! Insights from the relatives in northern Britain. *Language Variation and Change* 17. 75–112.

Taylor, Ann & Susan Pintzuk. 2011. The interaction of syntactic change and information status effects in the change from OV to VO in English. *Catalan Journal of Linguistics* 10. 71–94.

Traugott, Elizabeth C. 2007. Old English left-dislocations: their structure and information status. *Folia Linguistica* 41. 405–441.

Vangsnes, Øystein & Janne Bondi Johannessen. 2019. The Nordic research infrastructure for syntactic variation: possibilities, limitations and achievements. *Glossa: a journal of general linguistics* 4. 1–23.

Weir, Andrew. 2009. *Article drop in English headlines*. University College London MA thesis.

Wolfram, Walt. 1995. On dialect obsolescence: description, theory, and action. Paper presented at the ADS Annual Meeting, Special Session on Endangered Dialects. Chicago.

Wolfram, Walt & Nathalie Schilling-Estes. 1995. Moribund dialects and the endangerment canon: the case of the Ocracoke Brogue. *Language* 71. 696–721.

Zwart, C. Jan-Wouter. 2009. Uncharted territory? Towards a non-cartographic account of Germanic syntax. In Artemis Alexiadou, Jorge Hankamer, Thomas McFadden, Justin Nuger & Florian Schäfer (eds.), *Advances in comparative Germanic syntax*, 59–83. Amsterdam: John Benjamins.

Sali A. Tagliamonte
University of Toronto
Sid Smith Hall, Room 4075, 4th Floor
100 St George Street
TORONTO, Ontario, M5S 3G3
sali.tagliamonte@utoronto.ca
individual.utoronto.ca/tagliamonte/

Laura Rupp
Vrije Universiteit Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
l.m.rupp@vu.nl
research.vu.nl/en/persons/lm-rupp