“They used to follow Ø river”: The Zero Article in York English

Laura Rupp¹ and Sali A. Tagliamonte²

Abstract
Speakers of York English (UK) use a zero article with definite singular nouns (e.g., “They used to follow Ø river”), which is impossible in Standard English. We probe the possibility that this form is a remnant from Old English, when there were no articles as they are currently found in Modern English, rather than a more contemporary development. We trace the diachronic trajectory of the zero article in historical-descriptive grammars and test social and linguistic constraints on its use in York English in a logistic regression analysis. The results show that information structure is a significant predictor of the zero article across all generations of the community and that the zero article is used in the same way as it was used as far back as Old English. However, it exhibits heightened usage among the older and younger generations, exhibiting a U-shaped curve. We suggest that this pattern demonstrates longitudinal maintenance of a conservative feature, which is suppressed in middle-age as the result of social pressures. In this way, this case study adds insight into the fate of dialect features in contemporary speech communities. It also highlights the importance of combining insights from different strands in linguistics for understanding the evolution of syntactic variants like the zero article.

Keywords
zero article, York English, Old and Middle English, conservative dialect features, information structure, age-grading

¹Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
²University of Toronto, Toronto, ON, Canada

Corresponding Author:
Laura Rupp, Vrije Universiteit Amsterdam, De Boelelaan 1105, Amsterdam, 1081 HV, the Netherlands.
Email: l.m.rupp@vu.nl
I. Introduction: The Issue

Traditional dialect features in conservative varieties of English have often figured prominently in explanations of language variation and change. The variety of English spoken in the city of York in northern England is a case in point. It has three non-standard determiners: a zero article (Christophersen 1939), as in (1) and (2), a reduced determiner (Wright 1905:259), as in (3) and (4), and a complex demonstrative (e.g., Bernstein 1997), as in (5) and (6). The examples come from the York English Corpus (YEC) (Tagliamonte 1996-1998). The reduced determiner is a determiner form without a vowel and is known to occur in a range of different variants (e.g., Jones 1999).

(1) I many a time go and sit outside Ø Minster. (eburrit, 82)
(2) I could see it from Ø landing window. (nhudson, 17)
(3) And we’d been in t’ pub for two clock. (rjones, 50)
(4) Does [ʔ] teacher play it on the guitar? (maspel, 24)
(5) What is that there red book do you know? (ajackson, 66)
(6) You - know what these here police are. (rfielding, 87)

Previous studies of determiners in York English (YrkE) have mostly considered reduced determiners and complex demonstratives (e.g., Jones 2002; Rupp & Page-Verhoeff 2005; Rupp & Tagliamonte 2019). Regarding the zero article, its use in Standard English (StdE) is known to be more or less restricted to: (7) proper names, (8) non-count nouns (mass and abstract nouns), and (9) generic or kind-denoting plural nouns. Another use that has received attention in the literature concerns a circumscribed set of nouns naming social and geographical places such as church, school, and prison (see, e.g., Soja 1994 and Stvan 2007 for an inventory and discussion). These are frequently deployed as the complement of a prepositional phrase (PP) (Baldwin et al. 2006), as in (10). Otherwise, as Dayal (2011:1089) has stated, “English typically does not allow bare singular arguments,” namely a singular subject or object without an article, “setting aside exceptions like man is mortal etc.” This restriction is illustrated in (11) with an example from Stvan (2007:172).

(7) John is brave.
(8) Green tea is healthy.
(9) Firefighters are brave.
(10) They got married in church.
(11)* I put the book on desk.

However, the use of the zero article in YrkE is particularly noteworthy because speakers deploy it with definite singular nouns (like Minster in 1 and window in 2), the context we concentrate on here. There are two current hypotheses that could predict the occurrence of the zero article in contexts such as (1) and (2) in YrkE. First, Rupp’s (2007) analysis of reduced determiners in YrkE suggests that the zero article may be
derivative of the reduced determiner, as in (3) and (4). An idea that has long had much currency in the literature is that reduced determiners have arisen from the Middle English definite article *þe*, through definite article reduction. However, Rupp (2007) has argued that reduced determiners are best analyzed as reduction of the early Middle English distal demonstrative *þæt* followed by lenition, resulting in the range of variants that exists today: [t > θ > ?]. Under this analysis, it seems natural that the zero article constitutes the last stage in a lenition continuum (an analysis previously hinted at by Jones 2002:326). The perspective that the YrkE zero article is an essentially novel form may tie in with Krug and Lucas (2018:262), who have postulated that speakers of Maltese English have innovated a zero article under “internal pressures towards greater economy.” In spoken data collected for ICE-Malta, amongst other sources, they observed that Maltese English speakers used a number of nouns (e.g., *government, cabinet, budget, and beginning/majority of*) without a definite article specifically where it “may be omitted when the uniqueness or identifiability of a referent is salient in context” (Krug & Lucas 2018:261).

In contrast to an approach that derives the zero article from the reduced determiner, Tagliamonte and Roeder (2009) have proposed that the YrkE zero article is a relic from Old English, a period in which the language did not yet have separate articles (e.g., Mitchell 1985; see section 3 for discussion). Tagliamonte and Roeder’s (2009) interpretation is a plausible perspective because northern English is known to have preserved historical features such as verbal –s and the Northern Subject Rule (whereby speakers may attach the morpheme –s to verbs in the context of a plural NP-subject; Klemola 2000), the *for to* complementizer, and the deictics *yon* and *thon*, among other features (Tagliamonte, Smith & Lawrence 2005:82). Still, Hollmann and Siewierska (2011:32) have critically argued that proposals such as Tagliamonte and Roeder’s (2009) “would be more convincing if they were supported by data showing that the variation obtaining in the present day somehow still bears some reflection of the historical facts.”

Thus, more robust evidence of the current social and grammatical function of the zero article in YrkE is needed in order to further the investigation of its origins there and complement existing knowledge concerning the use of this variant of the definite article. We will therefore revisit the zero article in YrkE and probe its nature more systematically and in more depth than previous studies have done. Our specific research questions are: what is the social function of the zero article in YrkE, and what is its grammatical function in this variety? With this information we can also ask: what can the use of the zero article in YrkE tell us about its evolution from earlier stages of English, both diachronically and synchronically? Finally, how has the development of the zero article in YrkE differed from StdE? We will address these questions from the joint perspectives of historical linguistics, language variation and change, descriptive grammar, and discourse-pragmatics.

The structure of the remainder of our paper is as follows. In section 2 we outline the findings of previous dialectological and variationist research on the zero article. In section 3 we take account of the historical development of articles in English, focusing on the zero form. Our data and analytic method are described in section 4. Section 5
presents the results. In section 6 we will provide both a quantitative and qualitative analysis of the data and discuss our findings and conclusion.

2. Previous Dialectological and Variationist Research on the Zero Article

Dialectology studies have commonly reported that use of the zero definite article is restricted to the far southeast peninsula of Yorkshire, Holderness (the Survey of English Dialects [SED] of Orton & Dieth 1962-1971; Barry 1972:175), with the Wolds forming a boundary (Jones 1952:86). It may simply be that the zero form was not covered in York in surveys like the SED because they collected data from 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.” Shorrocks (1999:23-31) has documented the use of the zero form in particular contexts in Bolton, Greater Manchester. Hollmann and Siewierska (2011:42) have also reported the zero form for Lancaster on the basis of recordings from two speakers in the North West Sound Archive that were collected in the 1970s–1990s.

Tagliamonte and Roeder (2009) is the first large-scale variationist study of the definite article that includes both the reduced and the zero variants in the envelope of variation. From the YEC, they extracted 10,000 tokens of definite singular nouns that can take the definite article in StdE amongst fifty male and female speakers in four age groups: ≤30, 31–50, 51–74, and 75+ years of age. They compared the behavior of reduced and zero forms and found that they were very different in a number of respects. Tagliamonte and Roeder (2009:457) therefore concluded that the zero article is not related to the process of reduction that gave rise to the YrKE reduced determiner but is of an independent nature. One differentiating effect that Tagliamonte and Roeder (2009) observed was strong conditioning by type of following noun. Specifically, the zero form was significantly more frequent with definite nouns which historical grammars have shown to have had a long history of appearing without a definite article in English. It is this resemblance to uses of zero in the history of English that led Tagliamonte and Roeder (2009) to suggest that the zero article in YrkE goes back to an historical precedent in the region. However, Chesterman (1991:44) has remarked on these historical attestations that “there are so many ‘exceptional’ uses of ‘no article’ [. . .] that some significant generalizations seem to be badly missed.” In the next section, we will delve into the historical attestations of the zero article, seeking to observe a pattern that can help us shed light on the evolution of the zero article.

3. The History of the Zero Article in English

In order to understand the present system of definite articles in YrkE, it is useful to go back to the beginning of their developmental trajectory, in Old English (OE). Like many languages (Greenberg 1978), English did not have a separate definite or indefinite article at the earliest stages of its history. Instead, English had a complex
demonstrative system. According to Mitchell (1985:133), the OE distal demonstrative *se* (‘that’; used here as a shorthand for the range of gender, case, and number forms of the distal demonstrative) could be used as a marker of definiteness, namely, in the function of singling out a referent as distinct and identifiable from others of the same kind. However, Mitchell and other scholars (e.g., Traugott 1992:172) have remarked that *se* nonetheless behaved very differently from contemporary *the/that*: “there is no regular correspondence between OE and MnE [Modern English] usage. *Se* is found where we would not use ‘the/that’ [. . .] It can be absent where we would use ‘the/that’ [. . .]” (Mitchell 1985:134-135), for example, with identifiable referents. Epstein (2011:117) concludes that “[a]t this early stage in the history of English, it appears that the preference is for an NP to have no determiner (in effect, this is the default case).” He provides quantitative evidence from the OE epic poem Beowulf and demonstrates that *se* had a range of other uses, beyond identifiability, which determined its occurrence, such as the relative importance of the referent of a noun phrase (Epstein 2011:118; see Epstein [2011] and Mitchell [1985:127-136] for illustration from OE sources).

It seems fairly safe to say, therefore, that the zero article dominated in OE. Since then, however, the article system has shown “a constant loss of ground on the part of the zero-form” (Christophersen 1939:84), as it gave way to the definite and indefinite articles. The diachronic trajectory of the zero article is outlined in Table 1, which we have extrapolated from the literature.

First, in Early Middle English (ME), a specialized definite article developed from the nominative masculine and feminine distal demonstrative forms (*se, seó > þe > the*); this happened first in the north of England (McColl Millar 2000). The definite article initially overlapped with demonstrative usage (as in “Pass me the/that stool, please”; Lyons 1999:164) but it subsequently acquired distinct uses and evolved into a more general, full-fledged definite article. One example of this is the associative-anaphoric use, whereby the referent is evoked by association. For example, in (12), “exhaust fumes” can be definite because they are associated with the previously mentioned event of driving a car.

(12) The man drove past our house. *The/*These exhaust fumes were terrible. (from Chesterman 1991:52)

Later in the ME period, the indefinite article emerged from a weakened form of the numeral *an* ‘one’ (Fischer 1992:218). The indefinite article encroached on another function of the zero article, that of introducing a discrete (but non-identifiable) entity to the discourse (Epstein 1995:163).

In contemporary StdE, the use of the zero article has reduced to generic/kind-denoting and non-count nouns, as well as proper names. Jespersen (1954:417-418) comments that while *the* is used in situations of “nearly complete familiarity,” zero continues to be used in situations of “complete unfamiliarity” and “familiarity so complete that no article is needed.” Chesterman (1991:84) concludes: “[s]uch observations suggest that null is somehow ‘even more definite’ than *the.’” Zero may also be deployed for construing particular meanings that are consistent with its current function, for example, to personify nouns (essentially turning them into proper names; Hewson 1972:130), as in (13).
Meanwhile, as Table 1 shows, there are a number of historical attestations of zero in contexts where it endured relatively long, varied over time, or currently occurs with a more restricted scope. These attestations have been discussed in depth by historical-descriptive grammars (e.g., Christophersen 1939; Zandvoort 1945; Jespersen 1954; Quirk et al. 1985; Chesterman 1991). We can outline only a few cases here and refer the reader to these grammars for more examples and discussion. One use of zero that persisted into the Early Modern English period pertained to locative prepositional phrases such as at gate, while in another function zero was deployed to “deindividuate” nouns, that is, to render them “abstract, non-referential or non-specific” (Mathieu 2009:131). The examples in (14) and (15) are from Jespersen (1954:460) and Rissanen (1999:192), respectively.

(14) Bruce Dudley had just come down river.
(15) Nay sweete Hodge say truth, and do not me begile.

Uses of zero that are documented to have varied over time include nouns that have a sense close to a proper name, like geographical places and man-made structures
and unique denotations. Regarding uniques, Christophersen (1939:87) states that they appeared with the definite article relatively late. For example, while sun and moon have been combined with the definite article since OE, this practice was not established with world and earth until the ME period (Christophersen 1939:182), while certain uniques never acquired a definite article, such as God (Christophersen 1939:77). Zandvoort (1945:118) has drawn attention to an extension of unique denotation that occurs “when class-nouns like lord, king, tower, house, river, etc. are understood to refer to a person or thing unique in its own sphere. Thus, ‘the river’ in London is the ‘Thames.’”

Finally, with superlatives/comparatives like first and best, which circumscribe the referent of the noun (Quirk et al. 1985:270), the zero article was still variably used in ME, as in (16), whereas this use of zero is now restricted to particular contexts, as in (17).

(16) Others were singly engaged in fiercest battle with large groups. (from Mustanoja 1960:258)
(17) The banns are going up next Sunday. (from Zandvoort 1945:121)

A particular context of use of zero currently also applies to nouns that denote seasons. Following Jespersen (1954:540), a season has a zero article when it is referred to generically as a recurring period in the year, in “statements of their mere existence,” as in (18). By contrast, “when interest centers on a particular specimen, which is perhaps further characterized” (Christophersen 1939:183), the definite article is used. This use is shown in (19).

(18) In winter the Neva is a broad silent thoroughfare . . . (Merriman S 83)
(19) The first time I was in America was in the winter of 1923. (from Zandvoort 1945:120)

In the following sections we will show how taking account of the historical uses of the zero definite article helps decide on a historical precedent for its use in YrKE.

4. Method

The YEC comprises people born and raised in the city of York (Tagliamonte 1996-1998), representing approximately 1.2 million words from individuals between the ages of 17 and 92 and stratified by age, sex, education, and job type. These materials can be taken to represent the linguistic system of the locally-born population at the time of data collection in 1997.

Tagliamonte and Roeder (2009:447) observed that rates of the zero article were relatively low in the YEC when viewed as a proportion of all definite nouns found in the corpus. The zero form amounted to 3 percent, compared to 81 percent full and 16 percent reduced forms of the definite article. From this analysis we know that the full form of the definite article, the, patterns in parallel to phonologically reduced forms;
therefore, in our study we focus on the use of zero out of all other realizations of the article. Further, due to the extreme rarity of the zero forms as a proportion of all definite nouns, we chose an alternative sampling procedure. We selected for analysis singular nouns that were among the most frequently occurring nouns in the YEC. Hollmann and Siewierska (2011) noted that frequent definite noun phrases promote zero based on assumptions from Construction Grammar that postulate a connection between token frequency and processes of language change. However, when they tested this possibility in the North West Sound Archive, reduced and zero forms of the definite article (which they treated as one category) did not significantly correlate with frequent nouns. They concluded that “sheer frequency in a dialect or idiolect [does not] automatically lead to reduction” (Hollmann & Siewierska 2011:44). Nonetheless, extracting all tokens of a set of frequently occurring nouns offered us the best possibility of finding both variants.

A word list of the corpus enabled us to select the appropriate set of nouns to include in the analysis. In order to build a maximally comprehensive picture of the use of the zero article in YrkE, we selected the most frequent nouns that fulfilled three criteria. First, we sampled two types of noun: a) nouns that are reported to have occurred with the zero article in the history of English; and b) nouns that have no such record and are ungrammatical in zero contexts in StdE. Second, we stratified for lexical variety within these two groups of nouns. For example, upon having selected the noun river, we did not select another noun referring to a body of water. Third, we aimed to address a sampling concern raised by Hollmann and Siewierska (2011:40-41n9). They remarked that frequency counts from interview corpora will be potentially skewed because “interviews are conversations about speakers’ past, and thus not very accurate representations of their ordinary conversation. Specifically, words from certain domains are likely to be overrepresented.” To demonstrate, they compared figures for the words war (900 times per million words) and school (2000 times per million words) in the North West Sound Archive to figures for these words in the British National Corpus, where numbers turned out to be much lower (170 and 430, respectively). Therefore, we have sought to balance our data by selecting nouns that did not have gendered or generational associations (e.g., road and job). Applying these four criteria, we ended up with nine nouns which covered a range of types known to have varied with respect to the zero form historically, and nine nouns to control for these, for a total of 1728 tokens from eighteen different lexical items.

We excluded tokens of apparently singular nouns that can appear with the zero article in contemporary StdE, such as name+descriptor constructions (20), fixed phrases (21), parallel structures (22), antithesis (23), enumeration (24), and prosopopeia (25; see Jespersen 1954:413-416). In the YEC data, these occurred with the zero article 89 percent of the time ($N = 393$). Since we focused on the zero definite article, we also excluded zero occurrences of the indefinite article, as in (26).

(20) Yeah. Um, born off Heslington Road. (tlois, 35)
(21) [. . .] and then things just started to go down hill. (ocavell, 40)
(22) We were going across from bus to bus. (blowe, 62)
They are just playing *cat-and-mouse*. (kdilks, 26)

Had marks down *door and window*, weren’t it? (jrobinson, 27)

“Look what I’ve got here Mum.” Oh *bloody cat*. (emichaels, 63)

She was on *placement* and she got *job* on her placement. (rallen, 26)

These sampling procedures provided a total of 1335 tokens for statistical analysis, all of which could take the definite article in StdE but which could occur with a full form, reduced form, or zero form in the YEC.

Each type of noun was coded for social and linguistic factors that we extrapolated from the existing literature as possible determinants and competing influences on zero and non-zero contexts (e.g., the historical descriptive grammars in section 3). The social factors that we examined were determined based on the nature and distribution of individuals in the sample: sex, education (up to sixteen years of age, the [current] legislated age for compulsory education, or beyond sixteen years of age), occupation (student, white- or blue-collar job), and age, adopting four age groups (≤30, 31-50, 51-74, 75+) from Tagliamonte and Roeder (2009). Of the eighty-seven individuals in the YEC sample, the number of tokens of zero per speaker ranged from 1 to 50, with seventy-five individuals using the zero form at least once. Low numbers of tokens and low token counts per speaker motivated binning the age groups following Tagliamonte and Roeder (2009). The selected nouns (N = 1335) were distributed across age and sex as follows, respectively, with counts indicated in brackets: ≤30 (201), 31-50 (210), 51-74 (653), 75+ (271), and Male (625), Female (710). 7

We investigated the following linguistic factors: TYPE OF NOUN, GRAMMATICAL CATEGORY OF THE NOUN, and MODIFICATION OF THE NOUN. Given the role of articles in the referential properties of nouns in the discourse, we also coded for the factor INFORMATION STATUS. To our knowledge, this factor has not separately been explored in relation to the zero article before in the systematic way that we have approached it here, except for Sharma’s (2005) study of article use in L2 Indian English. The linguistic factors are described below.

The first factor is TYPE OF NOUN. As outlined above, nine of the 1335 frequent nouns that we selected for analysis have historically exhibited variation in zero definite article usage. We therefore termed these nouns “H(istorical) nouns.” We compared these nouns to nine (what we will call) “C(ommon) nouns.” These are forms that have not been associated with the zero article historically and that the literature flags as ungrammatical in StdE (e.g., StdE *I phoned office*; Carlson 2003). The historical nouns under investigation were as follows (with counts indicated in brackets): *Barbican* (23), *bridge* (22), *hill* (21), *Minster* (49), *museum* (22), *river* (74), *road* (192), *station* (45), and *winter* (43). Recall from section 3 that historical nouns include names of buildings; for this reason, we included both *Minster* (York’s landmark cathedral) and *Barbican* (an indoor entertainment venue located in York). Also note that zero article tokens with the noun *winter* included referential uses, as in (27), which are currently non-standard. Other examples are given in (28) and (29).
That was in nineteen-first year, Ø nineteen-seventy-nine and forty winter. (rfielding, 81)

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

There’s so much to do in Ø Railway Museum. (mpeters, 70)

The common nouns under investigation were as follows (with counts again indicated in brackets): bus (56), car (106), cat (19), corner (83), door (118), job (73), (news)paper (38), war (321), and window (30). Examples from the YEC are given in (30) and (31).

[. . .] she said “They painted Ø door” (hphilips, 72)

I crashed Ø car twice. (rmitchell, 20)

The examples in (32)–(37) involve grammatical category, illustrating use of the zero article in the context of the three different grammatical functions of nouns that we examined in this factor group: subject (32, 33), object of the verb (34, 35), and complement of a preposition (36, 37).

And Ø station got fire. (ajackson, 66)

 [. . .] well when Ø war come out they pulled me. (bhamilton, 91)

“Blown this one” and I got Ø job! (rjones, 50)

They used to follow Ø river I think did bombers. (mmichaels, 67)

[. . .] he always goes on Ø bus first. (mtoovey, 40)

And there was this bloke, he was asleep in Ø corner. (pgregory, 23)

Tagliamonte and Roeder (2009:457) previously found that in the YEC, the reduced article was favored following a preposition, while the zero variant was favored following a noun, adjective, or conjunction. Our categorization differs somewhat from theirs in that the contexts of noun and adjective were included in the group modification of the noun, which is discussed below, rather than as part of grammatical category.

In terms of modification of the noun, we distinguished bare noun (no modification), as in (38), and modification by an adjective (39), another noun (40), a comparative/superlative expression (e.g., first, last, same; see section 3), as in (41), and a relative clause (42).

He was in Ø paper. (gwalton, 87)

We had to get someone to get in Ø front window. (mlarkin, 70)

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

I thought it wasn’t Ø same cat. (mmichaels, 67)

[. . .] that’s the job that I hate every day. (cbiggs, 33)

It turned out that there were too few tokens of the factors comparative/superlative and relative clause for statistical analysis and we therefore removed them from this part of the analysis.
Finally, we examined the effect of information status, since articles are key indicators of referential information in the discourse (e.g., Prince 1981). Accordingly, we looked at all selected nouns in their discourse context. We based the coding for information status on Prince’s (1981:237) taxonomy of given-new information. Table 2 displays the coding scheme that we deployed accompanied by exemplification adapted from Prince.

Table 2. Coding Scheme for the Factor Group Information Status Based on Prince (1981)

| Factor                                      | Exemplification                                      |
|---------------------------------------------|-------------------------------------------------------|
| Discourse-new, hearer-new information       | I bought a beautiful dress.                          |
| (= brand new)                               |                                                       |
| Discourse-new, hearer-old information       | I will meet you at the river.                        |
| - shared speaker-hearer knowledge           | What can we do to save the earth?                   |
| - unique                                    |                                                       |
| Discourse-old, hearer-old information       | Did you see the dog just stepping into something?    |
| - situationally-evoked entities             |                                                       |
| - discourse-evoked entities                 | Susie went to visit her grandmother and the sweet    |
|                                           | lady was making Peking Duck.                         |
| - inferable entities (cf. Prince 1992:309)  | I went to the post office but the clerk couldn’t find|
|                                             | the parcel.                                          |
| Discourse-old, hearer-new information       | Discourse-old, hearer-new information is unattested. |

Within this scheme, we followed Hollmann and Siewierska (2011) in operationalizing “discourse-evoked” entities by means of Givón’s (1983:13) notion of “referential distance,” and “analysed anything occurring within a 20-clause window of its previous mention as [discourse-old], and anything occurring further away as [discourse-new]” (Hollmann & Siewierska 2011:39). However, given the nature of spoken language, we counted propositions (“a semantic unit composed of a predicate plus its arguments for which a truth value can be determined”; Seoane 2012:149) rather than clauses. We gauged that the interview data of the YEC were suitable to identify “situationally-invoked” referents, because the interviewer and the informant met face-to-face, and to code “discourse-new, hearer-old” information, because interviews concern a situation in which people frequently allude to something that is known to both interlocutors (Seoane 2012:153), for example, the river Ouse, the main river in York, being implicated in the noun river. Since the number of tokens of apparent “discourse-new, hearer-new” information was very low ($N = 15$), and nothing excluded familiarity, we subsumed them under the category discourse-new, hearer-old. Thus, we restricted our coding to the binary opposition: discourse-new, hearer-old, as in (43) and (44), where there is either unique or shared speaker-hearer knowledge, versus discourse-old, hearer-old, where the knowledge is situationally-invoked, as in (45), discourse-invoked, as in (46), or inferable from context, as in (47).
5. Results

5.1. Distributional Analysis

The zero article makes up 17 percent \( (N = 225) \) of the data. This frequency is comparatively high in relation to the figure of 3 percent from Tagliamonte and Roeder (2009:447). The difference is the result of a different way of circumscribing the variable context. As detailed in section 4, while Tagliamonte and Roeder’s (2009) analysis included all definite nouns, ours restricts the analysis to a specific set of carefully justified singular nouns. Figure 1 shows the distribution of the zero forms according to age and sex.

![Figure 1. Distribution of the Zero Article by Age and Sex of the Speaker in Apparent Time](image_url)

Figure 1 shows that every age group has the zero article and that in every age group, males are more frequent users. The level of zero article usage among the female speakers (the dashed line) declines dramatically in apparent time from the...
oldest generation (75+) to the other age groups (from 23 percent to between 7 percent and 9 percent). According to the apparent time construct, incremental decreasing use of linguistic features from oldest to youngest individuals in a community indicates linguistic change in progress (e.g., Bailey 2002). Tagliamonte and Roeder (2009:451) note that a trajectory of this kind, whereby one variant gradually recedes from oldest to youngest individuals, is suggestive of “ongoing obsolescence of an older form.” In this case, thecline is consistent with the idea that the zero article in the context of singular nouns is a historical remnant in YrkE. In contrast, the level of zero article usage among the male speakers (the solid line) shows a U-shaped curve. Although frequencies in the male speakers drop noticeably from the oldest cohort to the 51- to 74-year-olds (just as with the females), use of the zero article rises again to the extent that the youngest generation uses the form at similar rates (35 percent) to the oldest generation (37 percent).

Tagliamonte and Roeder (2009) found a similar U-shaped pattern for zero amongst the male speakers (albeit with rates that were much lower). Before we attempt to interpret the pattern for the zero definite article among the men, however, let us seek out further evidence by scrutinizing the patterns of variability by internal linguistic constraints.

5.2. Mixed Effects Modeling

We now turn to statistical modeling by carrying out a mixed effects logistic regression analysis using the lme4 package in R (R Development Core Team 2018), using the “bobyqa” optimizer. This enables us to determine the statistical significance of the social and linguistic factor groups when they are all considered simultaneously, as well as their patterning and relative strength. Table 3 displays the results.

Three social factors are statistically significant: occupati on, age, and sex. A fourth, education, only just meets the threshold for significance, undoubtedly due to the overwhelming effect of job type, which sets blue-collar workers apart from students and white-collar workers. This finding is consistent with other variables studied in the YEC (e.g., Tagliamonte [2001] on preterit come). That older, blue-collar males favor a non-standard form is a common finding in sociolinguistics (e.g., Labov 2001). However, that the zero article has significant social constraints in this dataset runs counter to the finding by Tagliamonte and Roeder (2009:454) and therefore requires closer scrutiny. We will return to the social constraints in the discussion in section 6, after we have considered the impact of the linguistic constraints.

The model in Table 3 shows that information status exerts the strongest effect on the zero article, while grammatical category of the noun and modification of the noun have lesser but also significant effects on this form. Type of noun is not significant. Zooming in on the highest ranked constraint, information status, zero articles are favored when the referent of the noun constitutes information that is discourse-new, hearer-old, whereas discourse-old, hearer-old information disfavors (see the examples 43-47 in section 4 for illustration).
Table 3. Mixed Effects Model of the Contribution of Linguistic Factors to the Probability of the Zero Form in the YEC

| Random effects: | Variance | Standard Deviation | Z Value | Individuals | Total N |
|----------------|----------|-------------------|---------|-------------|---------|
| individual (Intercept) | 1.218 | 1.104 | 1.166 | 87 | 1288 |

| Fixed effects: | Estimate | Std. Error | Pr (>|z|) | Percent | Ns/Cell |
|----------------|----------|------------|----------|---------|---------|
| (Intercept) | 0.7770 | 0.6664 | 0.24366 | | |
| Occupation | | | | | |
| Blue-collar | | | | | |
| White-collar/student | 1.5216 | 0.3881 | 3.920 | 8.84e-05 | *** | 8% | 804 |
| Information status | | | | | |
| Discourse-new; hearer-old | | | | | |
| Discourse-old; hearer-old | 1.1032 | 0.2132 | 5.175 | 2.28e-07 | *** | 13% | 1007 |
| Age | | | | | |
| 30 | −0.1566 | 0.5538 | −0.283 | 0.77730 | 24% | 231 |
| 31-50 | | | | | |
| 51-74 | 1.3074 | 0.4845 | 2.698 | 0.00697 | ** | 9% | 271 |
| 75+ | 0.2578 | 0.5209 | 0.495 | 0.62070 | 28% | 652 |
| Sex | | | | | |
| Female | | | | | |
| Male | −0.7643 | 0.3565 | −2.144 | 0.03204 | * | 22% | 719 |
| Grammatical category of the noun | | | | | |
| Object | | | | | |
| Preposition | 0.4956 | 0.2584 | 1.918 | 0.05511 | . | 15% | 966 |
| Subject | −0.3405 | 0.3408 | −0.999 | 0.31777 | 25% | 144 |
| Modification of the noun | | | | | |
| Adjective | | | | | |
| Bare | −0.7298 | 0.4056 | −1.799 | 0.07197 | . | 16% | 1149 |
| Noun | −0.7634 | 0.5134 | −1.487 | 0.13702 | 21% | 102 |
| Education | | | | | |
| Less educated | | | | | |
| More educated | 0.9302 | 0.5234 | 1.777 | 0.07553 | . | 9% | 265 |
| Type of noun | | | | | |
| Common | | | | | |
| Historical | −0.3080 | 0.2019 | −1.526 | 0.12710 | 21% | 516 |

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1.

In order to gain an understanding of the U-shaped pattern among male speakers (see Figure 1), we subsequently delved more deeply into the patterning of INFORMATION STATUS by examining how the strongest factor group patterned by social factors. Figure 2 shows a conditional inference tree, a procedure that permits visualization of complex patterns in linguistic data (Strobl, Malley & Tutz 2009).
Figure 2. Conditional Inference Tree Visualization of Social and Linguistic Factors
Note: Not displayed here is that blue-collar (B), young (Y), and old (X) individuals, too, favored zero in contexts of “disc-new” (discourse-new) over “disc-old” (discourse-old) information.
The two most noticeable results in Figure 2 are, first, that rates of zero are generally higher when the referent of the noun is discourse-new, hearer-old (“disc-new”), regardless of whether the speakers are white-collar/students (W, S) or blue-collar (B). Second, the effect of age shows that middle-aged individuals (M, 31-50) have the highest frequency of zero forms among the white-collar/student group (under Node 4). Of further note is the fact that individuals with less education (N), whether they are older (O, X, 51-74, 75+) or younger (Y, ≤30) (under Node 8), use the zero variant more. This complex interplay of linguistic and social factors suggests: 1) enduring discourse-governed use of the zero form; and 2) resurgence of the zero form among younger, less educated speakers. We will address these findings in the discussion in section 6.

6. Discussion

Our examination of the social and linguistic factors contributing to the use of the zero article in YrkE revealed a complex system. The main determinants of use are the individual’s type of job and information status: the zero article occurs favorably when the referent of the noun is discourse-new, hearer-old. We believe that this result can be understood in the historical context of how the zero article evolved in English. As outlined in section 3, the zero form is the oldest of the three articles in English, but the form dropped from most contexts when the and a gradually spread and took over many of its former uses. Currently, the zero article is deployed in StdE in the specific case of proper names (among other contexts). While this use may be described as a contemporary exception to the use of definite the (by virtue of proper names being inherently definite already, and the therefore unnecessary), history suggests that it represents a historical use that has been retained. Indeed, we would like to suggest that the significant rates of zero that we observed in the context of discourse-new, hearer-old information in YrkE can be understood on a par with proper names, namely, as a continuation of a once current use of the zero article with singular nouns. Note that discourse-new, hearer-old nouns have a sense close to proper names: their referent can be identified because it is unique (in its own sphere) or known by the speaker and the hearer (e.g., in York, river in example 35 will be taken to refer to the river Ouse). In this relation, the occurrence of the zero form with (what we have termed) Historical nouns seems to us to derive from these Historical nouns being particularly amenable to a familiar/unique reading (e.g., geographical places; see section 3) and, therefore, apt to have been retained with zero in the history of English. However, YrkE speakers do not only demonstrate this use but are even more conservative by deploying the zero article with Common nouns in this meaning.

In the light of the perspective that we think is unfolding, we also seem in a position to explain why the zero article is used to a lesser extent for reference to hearer-old, discourse-old information (e.g., 46). This is a still earlier use of the zero article. As we outlined in section 3, nouns whose referent can be retrieved from a previous mention in the discourse (anaphoric deixis) were the first to abandon zero and to assume the definite article. We thus find that our research provides sufficient grounds to subscribe
to Tagliamonte and Roeder’s (2009) assessment that the use of the zero definite article in YrkE is a synchronic reflex of an old zero article that has been carried over from earlier periods in the history of the English language.

Social factors were also significant: males and those in blue-collar jobs favored the zero variant, but it was eschewed by middle-aged speakers. A conditional inference tree revealed further patterns: the middle-aged group rarely uses the zero article, but the linguistic patterns of use are parallel across generations. Taking the results from all the different analyses together, we can make the following interpretation for the male speakers: those in the middle-aged groups, particularly between ages 51 and 74, have suppressed the zero definite article while retaining its internal mechanisms. Such a pattern is indicative of age grading. Age grading happens when people of different ages use language differently simply because they are at a particular stage in life (e.g., Wardaugh 2002:194) and there is no evidence for linguistic change in progress (Labov 1994:84). First, in the results we have presented here, the older and younger generation have near identical frequency of usage demonstrating maintenance of usage. Second, individuals across age groups share the same internal constraints, suggesting that the community does not change. Third, because the zero definite article is a non-standard form, it seems likely that it would be sensitive to stigmatization and external pressure leading to infrequent usage among individuals in the work force, a factor attested as a strong predictor of conservative linguistic behavior and of age grading in particular (Holmes 1992:186). While we cannot be absolutely certain of this interpretation without real time evidence, we suggest that not only does this explanation fit the facts, it also demonstrates how grammatical options may become intertwined with locally situated indexical meanings (e.g., Eckert 2000) under conditions of longitudinal stability.

We conclude that the story of the zero definite article in York offers an example of longitudinal endurance of a linguistic feature and a stable pattern of its internal linguistic patterning. The maintenance of conservative linguistic features by men is a longstanding finding for stable variables enshrined in Principle 2 of Labov’s principles of linguistic change (Labov 2001:266), namely that men will use more of the old linguistic variants than women. More broadly, these results add to the building picture of the many ways dialect features may pattern in contemporary English. Sometimes dialect features are recycled by individuals in different age cohorts (e.g., Dubois & Horvath 1999), sometimes they extend to new generations (e.g., Schilling-Estes & Wolfram 1994), sometimes they occur among more restricted sub-groups (e.g., Jankowski & Tagliamonte 2017), and sometimes they endure at low, but regularly conditioned, levels, as in the case at hand where an age-old mechanism remains and the form is structured by social prestige and local meaning.

To summarize, we have extended previous work on the zero article by examining social and linguistic factors in the local vernacular spoken in York, England, at the end of the twentieth century. We have added new insights into the grammatical function of the zero article in this variety by exposing its underlying character and structured heterogeneity in the speech community. Moreover, we have highlighted the usefulness of unifying contributions from historical linguistics, descriptive grammar, language
variation and change, and discourse-pragmatics for advancing the understanding of how the grammatical function of the zero article has evolved synchronically and diachronically. Historical linguistics provides the foundation for understanding the nature of the zero article at the beginning of its developmental trajectory in OE, and historical descriptions have further traced its grammatical development across the centuries. Language variation and change captures the linguistic mechanisms and social embedding of this feature in the community grammar. Finally, following Cheshire (2005), we would like to underline the importance of analyzing syntactic forms in their discourse context. Discourse-pragmatics has supplied an additional and key interpretive piece to our puzzle by showing that the use of the zero article in YrkE corresponds to its historical usage. Together these complementary strands of linguistics enable us to explain the zero definite article in contemporary YrkE as a retention of an old grammatical feature rather than an innovation.

YrkE demonstrates that bare singular nouns remain a viable option in a variety of contemporary English, contrary to what is commonly thought. In the context of the grammaticalization trajectory of the definite article the (Greenberg 1978), one might think that usage of the zero article will progressively decline and fade into obsolescence (cf. Hopper & Traugott’s 2003 “Hypothesis of Unidirectionality”); however, the zero article has actually persisted in YrkE in certain uses longer than in StdE. Rather than giving way as a transitional phenomenon, expressions may remain part of the evolving system. Chesterman (1991:86) even envisages that “we might expect certain uses of the to give way in time to null,” while Christophersen (1939:150) has identified a novel use of zero with bare plurals.

To what extent the deployment of the zero article with definite singular nouns is productive and whether it will continue to evolve in YrkE remains to be seen. Further detailed empirical investigation is needed to unravel the full story of this form in contemporary English and to compare the way in which it operates to other null uses of the definite article across varieties and languages, like French, the L2-variety of Indian English, and the Noord-Groningse dialect of Dutch, where Epstein (1995), Sharma (2005), and Oosterhof (2008:99) have identified other primary contexts of use: deindividualization, discourse-old, hearer-old information, and expected definite reference, respectively. However, we would agree with Epstein (1995:163) that “the very meaning of the zero article [. . .] allows it to serve useful discourse functions,” making this area of grammar a rewarding possibility for integrating across sub-fields of linguistics in the pursuit of explaining linguistic variation and change.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The second author gratefully acknowledges support from the Economic and Social Science Research Council of the United Kingdom for Research Grants 1997-2001.
Notes

1. The quote in the title refers to bombers during World War II who used the River Ouse for directions.

2. We assume the presence of a zero article, rather than the absence of the definite article, because (as we will see) bare singular count nouns can function as arguments in York English, and it is determiners that allow nouns to function as arguments (Dufresne, Tremblay & Déchaine 2018:24-25, and references therein). We note that the properties of the York English zero article can fruitfully feed back into generative syntactic theories about the structure of bare noun phrases, particularly the DP-hypothesis of Abney (1987), but we will not directly engage in this discussion here (see, e.g., Ghomeshi, Paul & Wiltshko 2009).

3. This is the speaker’s pseudonym as first initial and last name followed by age. 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 and Roeder (2009:463n1).

4. While we will address a number of other occurrences of the zero article in section 3, it is impossible to mention all the cases in which nouns can be used without a determiner. See Quirk et al. (1985:265-297) and Chesterman (1991), amongst others, for detailed overviews. These include headlines like “India Budget May Hasten Show-down” (Jespersen 1954:416).

5. Other languages more readily admit zero articles in the context of singular count nouns, e.g., Norwegian (Borthen 2003). However, the behavior of the zero articles in these languages differs from that of the zero article that we have observed in the YEC (see Borthen 2003:108).

6. In other languages, like French, the role of zero has even further decreased as the definite article is used with non-count nouns and generic/kind-denoting plural nouns (e.g., “J’aime le fromage” ‘Generally, I like (*the) cheese’) (Epstein 1995 and references therein). According to Greenberg’s (1978) analysis of definite article development, a definite article will eventually be attached to nouns indiscriminately and grammaticalize into a general marker of nominality.

7. Instances of the noun war were much higher in the 51-74 group than in others; however, this was also true of the nouns road, door, paper. The 51-74 group had comparable token rates for nouns like winter and station to the other age groups.

8. Tagliamonte and Roeder (2009:458-459) in effect coded “information structure” as “discourse parallelism,” which they described as “the influence of the form of the definite article on an identical noun in the immediately preceding context.” We conceived of this factor group as “reference to old information,” the referent of a noun being identifiable from a previous mention.

References

Abney, Stephen P. 1987. The English noun phrase in its sentential aspect. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Bailey, Guy. 2002. Real time and apparent time. In Jack K. Chambers, Peter Trudgill & Nathalie Schilling-Estes (eds.), The handbook of language variation and change, 321-332. Oxford: Blackwell.
Baldwin, Timothy, John Beavers, Leonoor van Beek, Francis Boond, Dan Flickinger & Ivan A. Sag. 2006. In search of a systematic treatment of determinerless PPs. In Patrick Saint-Dizier (ed.), *Computational linguistics dimensions of syntax and semantics of prepositions*, 163-180. Dordrecht: Springer.

Barry, Michael V. 1972. The morphemic distribution of the definite article in contemporary regional English. In Martyn F. Wakelin (ed.), *Patterns in the folkspeech of the British Isles*, 164-181. London: Athlone.

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

Bernstein, Judy B. 1997. Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102(2-3). 87-113.

Borthen, Kaja. 2003. Norwegian bare singulars. Trondheim: Norwegian University of Science and Technology dissertation.

Carlson, Greg. 2003. No lack of determination. In Lisa Cheng & Rint Sybesma (eds.), *The second Glot international state-of-the-article book*, 147-178. Berlin: Mouton de Gruyter.

Cheshire, Jenny. 2005. Syntactic variation and beyond: Gender and social class variation in the use of discourse-new markers. *Journal of Sociolinguistics* 9(4). 479-508.

Chesterman, Andrew. 1991. *On definiteness: A study with special reference to English and Finnish*. Cambridge: Cambridge University Press.

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

Dayal, Venecta. 2011. Bare noun phrases. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook on natural language meaning*, vol. 2, 1088-1108. Berlin: Mouton de Gruyter.

Dubois, Sylvie & Horvath Barbara. 1999. When the music changes, you change too: Gender and language change in Cajun English. *Language Variation and Change* 11(3). 287-313.

Dufresne, Monique, Mireille Tremblay & Rose-Marie Déchaine. 2018. The variable use of determiners in Old French and the argument DP hypothesis. *Linguistic Variation* 18(1). 23-48.

Eckert, Penelope. 2000. *Language variation as social practice*. Oxford: Blackwell.

Epstein, Richard. 1995. The later stages in the development of the definite article: Evidence from French. In Henning Andersen (ed.), *Historical linguistics 1993*, 159-175. Amsterdam: John Benjamins.

Epstein, Richard. 2011. The distal demonstrative as discourse marker in *Beowulf*. *English Language and Linguistics* 15(1). 113-135.

Fischer, Olga. 1992. Syntax. In Norman Blake (ed.), *The Cambridge history of the English language*, vol. 2, 1066-1476, 207-408. Cambridge: Cambridge University Press.

Ghomeshi, Jila, Paul Ileana & Martina Wiltschko (eds.). 2009. *Determiners: Universals and variation*. Amsterdam: John Benjamins.

Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In Talmy Givón (ed.), *Topic continuity in discourse: A quantitative cross-language study*, 1-42. Amsterdam: John Benjamins.

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.

Hewson, John. 1972. *Article and noun in English*. The Hague: Mouton.

Hollmann, William B. & Anna Siewierska. 2011. The status of frequency, schemas and identity in cognitive sociolinguistics: A case study on definite article reduction. *Cognitive Linguistics* 22(1). 25-54.

Holmes, Janet. 1992. *An introduction to sociolinguistics*. London: Longman.
Hopper, Paul J. & Elizabeth C. Traugott. 2003. Grammaticalization. 2nd edn. Cambridge: Cambridge University Press.

Jankowski, Bridget & Sali A. Tagliamonte. 2017. A lost Canadian dialect: The Ottawa Valley 1975-2013. In Tanja Säily, Arja Nurmi, Minna Palander-Collin & Anita Auer (eds.), Historical linguistics: A festschrift for Terttu Nevalainen on the occasion of her 65th birthday, 239-274. Amsterdam: John Benjamins.

Jespersen, Otto. 1954. A modern English grammar on historical principles, part VII, Syntax. London: George Allen & Unwin.

Jones, Mark J. 1999. The phonology of definite article reduction. In Clive Upton & Katie Wales (eds.), Dialectal variation in English: Proceedings of the Harold Orton Centenary Conference 1998. Leeds Studies in English 30. 103-121.

Jones, Mark J. 2002. The origin of definite article reduction in northern English dialects: Evidence from dialect allomorphy. English Language and Linguistics 6(2). 325-345.

Jones, William E. 1952. The definite article in living Yorkshire dialect. Leeds Studies in English 7-8. 81-91.

Klemola, Juhani. 2000. The origins of the Northern Subject Rule: A case of early contact? In H. Tristam (ed.), Celtic Englishes II, 329-346. Heidelberg: Winter.

Krug, Manfred & Lucas Christopher. 2018. Definite article (omission) in British, Maltese, and other Englishes. Language Typology and Universals 71(2). 261-305.

Labov, William. 2001. Principles of linguistic change, vol. 2, Social factors. Oxford: Blackwell.

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

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

Mustanoja, Tauno F. 1960. A Middle English syntax. Helsinki: Société Néophilologique.

Oosterhof, Albert. 2008. Deletie van het bepaald lidwoord in het Noord-Gronings. Taal & Tongval 21. 73-108.

Orton, Harold & Eugen Dieth (eds.). 1962-1971. Survey of English dialects: The basic material. 4 vols. Leeds: E. J. Arnold.

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

Prince, Ellen F. 1992. The ZPG letter: Subject, definiteness, and information-status. In Sandra A. Thompson & William C. Mann (eds.), Discourse description: Diverse analyses of a fund-raising text, 295-325. Amsterdam: John Benjamins.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartik. 1985. A comprehensive grammar of the English language. London: Longman.

R Development Core Team. 2018. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. www.R-project.org.

Rissanen, Matti. 1999. Syntax. In Roger Lass (ed.), The Cambridge history of the English language, vol. 3, 1467–1776, 187-331. Cambridge: Cambridge University Press.

Rupp, Laura. 2007. The (socio-)linguistic cycle of Definite Article Reduction. Folia Linguistica Historica 28(1-2). 215-249.
Rupp, Laura & Hanne Page-Verhoeff. 2005. Pragmatic and historical aspects of Definite Article Reduction in northern English dialects. *English World-Wide* 26(3). 325-346.

Rupp, Laura & Sali A. Tagliamonte. 2019. *This here town*: Evidence for the development of the English determiner system from a vernacular demonstrative construction in York English. *English Language and Linguistics* 23(1). 81-103.

Schilling-Estes, Natalie & Walt Wolfram. 1994. Convergent explanation and alternative regularization patterns: *Were/weren’t* levelling in a vernacular English variety. *Language Variation and Change* 6(3). 273-302.

Seoane, Elena. 2012. Givenness and word order: A study of long passives from Early Modern English to Present-Day English. In Anneli Meurman-Solin, Maria José López-Couso & Bettelou Los (eds.), *Information structure and syntactic change in the history of English*, 139-163. Oxford: Oxford University Press.

Sharma, Devyani. 2005. Language transfer and discourse universals in Indian English article use. *Studies in Second Language Acquisition* 27(4). 535-566.

Shorrocks, Graham. 1999. *A grammar of the dialect of the Bolton area*, part 2, *Morphology and syntax*. Frankfurt am Main: Peter Lang.

Soja, Nancy Narva. 1994. Evidence for a distinct kind of noun. *Cognition* 51(3). 267-284.

Strobl, Carolin, James D. Malley & Gerhard Tutz. 2009. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* 14(4). 323-348.

Stvan, Laurel Smith. 2007. The functional range of bare singular count nouns in English. In Elisabet Stark, Elisabeth Leiss & Werner Abraham (eds.), *Nominal determination: Typology, genesis, and historical emergence of DP*, 171-187. Amsterdam: John Benjamins.

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. 2001. *Come/came* variation in English dialects. *American Speech* 76(1). 42-61.

Tagliamonte, Sali A. & Rebecca V. Roeder. 2009. Variation in the English definite article: Socio-historical linguistics in t’speech community. *Journal of Sociolinguistics* 13(4). 435-471.

Traugott, Elizabeth Closs. 1992. Syntax. In Richard M. Hogg (ed.), *The Cambridge history of the English language*, vol. 1, *The beginnings to 1066*, 186-189. Cambridge: Cambridge University Press.

Wardaugh, Ronald. 2002. *An introduction to sociolinguistics*. Oxford: Blackwell.

Wright, Joseph. 1905. *The English dialect grammar*. Oxford: Henry Frowde.

Zandvoort, Reinard W. 1945. *A handbook of English grammar*. Groningen: Wolters.

**Author Biographies**

**Laura Rupp** is Associate Professor of English Linguistics at the Vrije Universiteit Amsterdam, the Netherlands. Her research focuses on historical and contemporary functions of grammatical dialect features. She has a special interest in the different ways in which dialect features may be maintained over time.

**Sali A. Tagliamonte** is Professor of Linguistics at the University of Toronto, Canada. She is the author of six books, including *Making Waves* (Wiley-Blackwell, 2015). She has published on Canadian and British English dialects and teen language. Her research focuses on morphosyntactic and discourse-pragmatic features, using cross-community comparisons in apparent time to explore linguistic change.