

Anna KARABAN, PhD (Philol.), Assoc. Prof.

ORCID ID: 0000-0002-1101-4227

e-mail: a.karaban@knu.ua

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

ORTHOGRAPHIC VARIATION IN MID-SEVENTEENTH-CENTURY ENGLISH: A CORPUS STUDY OF LONDON, OXFORD, CAMBRIDGE, AND YORK ALMANACS

Background. Early Modern English is pivotal for understanding the transition from regionalized writing to a supraregional standard, yet systematic dialect cartography for the period remains sparse. Previous research has centred on literary and legal genres; popular almanacs, though printed annually in every major town, have been ignored. Because they fix place of printing on the title-page, repeat identical calendar formulae, and circulated in large numbers, almanacs are ideal but under-used witnesses to mid-seventeenth-century orthographic practice. This study evaluates whether geographically specific spellings in four imprints can trace dialect boundaries and the diffusion of metropolitan norms.

Methods. All four texts were digitised through OCR and manually corrected. Five high-frequency variables were extracted: mute-e deletion, medial schwa deletion, adjectival -ie/-y reduction, geminate-l reduction, and the third-person-singular present suffix variation. Tokens were counted only in structurally parallel calendar headings and prognostications, then collapsed into variable-level totals. Qualitative inspection of copy-sources and genre conventions contextualised quantitative patterns.

Results. Three variables: mute-e, schwa loss, and -ie/-y display a clear south-to-north diffusion: London is categorical for the innovative spellings, Oxford approximates London but retains a thin conservative layer, Cambridge is mixed, and York remains almost entirely traditional. Geminate-l reduction proves not dialectal but an index of orthographic modernization: only provincial presses preserve archaic -all forms, while London is fully re-Latinised. The third-person-singular present suffix behaves conversely: despite its northern origins, innovative -s is rare because printers across the country retain biblical -eth in learned contexts; York's heavy use of -eth reflects stylistic conservatism, not local speech.

Conclusions. Almanacs capture both conservative diction and emerging metropolitan norms, offering a fine-grained lens on orthographic change during a decisive standardizing decade. The south-to-north gradient confirms London's leadership in spelling reform, while variable-specific behaviour highlights the mediating role of genre and register. Scaling the procedure to the larger pool of provincial almanacs promises a data-driven dialect atlas that will refine our understanding of Early Modern English regional boundaries and the mechanics of print-based language change.

Keywords: Early Modern English; almanacs; dialectology; historical linguistics; regional variation; isoglosses; orthography; corpus study

Background

The Early Modern English period (c. 1500–1700) was marked by considerable regional linguistic variation, much of which was captured in print before standardization norms took hold. Despite the prominence of Early Modern English in historical linguistics, systematic mapping of dialect variation during this period remains uneven. Traditional research has focused primarily on literary or legal texts, with relatively little attention paid to more formulaic printed genres such as almanacs, despite their wide circulation and stable structure. Almanacs, due to their widespread geographical distribution, consistent format, and regular publication, provide a promising resource for examining regional dialect features, specifically orthographic variation. Thus, this study aims to examine orthographic differences recorded in almanacs published in distinct geographical locations (London, Oxford, Cambridge, York) to identify dialect boundaries and assess dialectal change from the mid-17th to early 18th centuries. Specifically, the objectives of this study are to: (1) extract and document orthographic variants systematically; (2) map these variants geographically; and (3) assess the state of dialect diffusion across England. This study contributes to the reconstruction of EModE dialect boundaries by using orthographic variation in almanacs as linguistic evidence. Its aim is to test whether geographically-distributed spelling variants can be used to approximate isoglosses and trace dialectal change over time. Drawing on digital methods and geospatial analysis, the project situates almanac spelling within wider efforts to digitally reconstruct historical English dialects.

Literature overview and theoretical framework

Although dialectological research on Middle English has advanced significantly due to resources like LALME and LAEME, the Early Modern English (EModE) period remains comparatively understudied,



especially with regard to regional variation. As Meurman-Solin (2017) notes, this gap stems from both the scarcity of regionally anchored, digitized data and the influence of prescriptive norms that obscured local linguistic features in writing. With the spread of printing and a growing supraregional standard, vernacular spelling became increasingly uniform, reducing the textual visibility of dialectal traits. However, she emphasizes that local variation persisted, and calls for corpus-driven, spatially sensitive research to uncover these subtler patterns by treating geographic space as a core variable in historical dialectology.

The period's scholarship establishes that regional dialect profiles well-known in Middle English largely carried into Early Modern English, though they were gradually mitigated by the expansion of a London-based standard. Over the 1500–1700 period, many such features were either spread beyond their original region into the standard or retreated under pressure from standard forms, but they left clear traces in the historical record.

Horobin and Smith (2002) observe that, as English relinquished official functions to Latin and Anglo-Norman French, scribes vernacularized spelling, inflections, vocabulary, and even syntax to reflect local speech, yielding what they term an era of maximal written variation. Between c. 1350 and 1500, however, the combined pressures of humanist translation projects, Reformation-era printing, and expanding state administration fostered four emergent written types with characteristic spellings that presaged later print conventions. The sixteenth-century introduction of the printing press then regularized orthography and morphology, nonetheless, regional and social differences remained visible in prescriptive grammars, spelling manuals, and orthographic-reform proposals even as printers advanced a unified standard. Fisiak (1995) highlights how the post-1500 surge in print – and the corresponding drive toward a standardized orthography – erased much direct manuscript evidence of regional speech, compelling researchers to mine more conservative proxies such as place-names.

Nevalainen (2006), in her analysis of Early Modern dialectal diversity, draws on three seminal sources: Alexander Gil's *Logonomia Anglica* to chart phonological and select morphological isoglosses across six named varieties; George Puttenham's *Arte of English Poesie* to illustrate how prestige norms were prescribed for literary production, thereby defining a sociolectal isogloss between cultured and rustic speech; John Hart's *Methode* to demonstrate how mid-sixteenth-century orthographic reform proposals explicitly drew on contemporary pronunciation, demonstrating the influence of vernacular speech forms on orthographic standardization. Alexander Gil's *Logonomia Anglica* (1619/1621, as translated by Dixon 1951) identifies six principal Early Modern English dialects – Common, Northern, Southern, Eastern, Western, and Poetic – each defined by characteristic phonological and morphological features. The Poetic variety permits selective regional forms for literary effect, all set against a Common (prestige) dialect that serves as a socially universal norm despite occasional ambiguity.

One of the clearest early-modern isoglosses was the north–south divide inherited from Middle English. Early Modern English also notes many lexical differences between regions as Northern English continued to employ words of Norse origin or unique local terms that set it apart from southern usage. In 1674 the naturalist John Ray published a landmark *Collection of English Words Not Generally Used* (Ray, 1674/2015), explicitly dividing his glossary into words "proper to the Northern" and "to the Southern Counties" as well as documenting dozens of regional terms, affirming a broad lexical split between north and south in the 17th century.

Ruano-García (2020) argues that while standardization narratives have traditionally focused on the emergence of a consensus dialect, especially centered on London and the southeast, literary and lexicographical texts reveal ongoing awareness of regional forms, particularly Northern English. Through enregisterment Northern dialect traits were circulated and indexed as socially marked, often linked to characterological types such as the "honest northerner". In an earlier study Ruano García (2012) provides evidence of how specific northern phonological and orthographic forms were enregistered in Early Modern English through consistent representations of northern variants (e.g. reflexes of ME /a:/, /o:/, /u:/, and /k/) that appear with striking frequency across genres, albeit with different distributional patterns. These forms, often rendered semiphonetically are associated with characterological types and serve as indexical markers of region and identity. Although these literary texts were not linguistically realistic, their repeated use of northern forms contributed to the circulation of a recognizable linguistic repertoire.

Nevalainen and Raumolin-Brunberg (2003) emphasize two watershed social upheavals – the sixteenth-century Reformation and the mid-seventeenth-century Civil War and Revolution – as catalysts for intensified linguistic diffusion across Early Modern England. Protestant reforms not only dismantled monastic schooling (thereby reshaping literacy and promoting lay reading) but also redistributed vast tracts of land to a more mobile, landowning gentry, while the Civil War famously displaced armies and civilians alike,

breaking down the insularity of parochial speech communities and creating unprecedented contact among speakers from London, East Anglia, the North, and beyond.

By about 1650, English printed texts had largely adopted uniform spellings and usages (Rissanen, Kytö, & Palander-Collin, 1996), meaning that, unlike Middle English manuscripts, most Early Modern texts do not straightforwardly reflect an author's local speech.

Orthographic variation in Early Modern English often serves as a direct proxy for underlying pronunciation, since scribes and printers routinely spelled words to mirror their own speech. By the late seventeenth century, even printed forms signal phonetic targets that Chamberlain explicitly documents in his *Tractatus*, anchoring orthography-to-phonology correspondences at a known date and locale (Salmon, 1996). Thus, charting orthographic isoglosses allows us to infer the spread – and lag – of dialect-specific features across time and space.

McEnery and Baker (2019) offer a turnkey model for our isogloss mapping by showing how a large-scale, diachronic corpus can be mined to reveal fine-grained shifts in word meaning and social evaluation. This real-time collocation framework can be directly repurposed to chart the spread of orthographic and phonological variants as isoglosses across Early Modern dialects.

Contemporary corpus-driven studies – most notably Taavitsainen & Schneider's (2019) long-range analysis of scholastic medical prose – demonstrate how bottom-up, token-based methods can surface previously unseen stylistic patterns by letting the data itself define the features of interest. We intend to adopt the same real-time to our radically different strand of evidence: the orthographic tokens of almanacs, by triangulating these micro-data with known print-market and mobility histories, for this pilot study to both validate the token-mapping workflow and pave the way for a larger, dialect atlas of almanac orthography.

Methods

To investigate regional orthographic variation in mid-17th-century English almanacs, we implemented a small-N comparative design, drawing on five provenance-clear imprints from Early English Books Online (EEBO). To compile the dataset for this study, almanacs were systematically selected from the EEBO database restricted the publication date range narrowly to the period between 1500 and 1700 to maintain chronological comparability in print conventions and linguistic usage across texts and filtered through Universal Short Title Catalogue (USTC) subject classification, specifically targeting the category 'Calendars, almanacs and prognostications' to ensure the identification of relevant texts within a uniform genre, minimizing potential variability in orthographic practices arising from differences in style or content. All selected almanacs fall within the same genre, primarily containing calendrical information, astrological data, and practical advice. This genre consistency minimizes stylistic differences, ensuring orthographic variations captured primarily reflect regional dialect distinctions rather than genre-driven lexical or stylistic choices.

Additional considerations included selecting almanacs of similar length and physical format matching as closely as possible regarding their panel organization, and typography, minimizing potential print-related variation that could confound orthographic analysis and thereby maintaining consistency in the type of textual data extracted for dialect mapping. After initial selection based on the above criteria, each almanac was carefully verified manually to confirm readability and completeness of digital facsimiles and to ensure accurate representation in subsequent OCR processes and only those sheets with legible calendar tables and daily prognostications were retained; fragmentary or heavily damaged signatures were excluded to ensure consistent spatial coverage.

Four almanacs were ultimately chosen based on their explicit provenance: one each from York, Oxford, Cambridge, and London. These cities were strategically selected to represent distinct geographical regions of England (Northern, Midlands, East Anglia, and Southeastern dialect areas, respectively), providing a balanced spatial distribution for dialectal analysis. Four primary publication cities were targeted to achieve a balanced regional representation: London (Southeast), as a baseline standard representing the printing capital and linguistic norm of the period. Oxford and Cambridge, representing the English Midlands, to capture dialect features from the central region and estimate the strength of the London dialect in this area in this specific time frame. York, chosen explicitly to represent the northern dialectal region as one of the stronger attested south-north contrast in Early Modern English period.

By anchoring our study precisely to as narrow of frame across selected geography, we settled on the 1655–1666, which ensures that every text lies squarely within the same socio-political milieu – post-Commonwealth but pre-Restoration stabilization – so that orthographic differences reflect genuine regional preference rather than diachronic shifts in prescriptive norms. This twelve-year span corresponds to the apex of almanac production, when provincial presses in Oxford, York, Cambridge, and London churned out annual prognostications under remarkably similar technical conditions, yielding a high density of

geographically indexed spellings. By confining our analysis to timeframe, we thus bracket out both the upheavals of the 1640s and the later eighteenth-century moves toward standardized orthography, while providing a precisely bounded slice of mid-seventeenth-century English in which to trace nascent isoglosses with maximal clarity and scholarly comparability.

It is noteworthy that the EEBO database did not yield almanacs published in (south)western English cities within the specified date range. Contemporary linguistic observers recognized these areas as dialectally distinct. The absence of (south)western city almanacs during this period thus presents a limitation in our current dataset. However, an intriguing avenue for future research is available, as EEBO holds an almanac from Bristol published starting in 1700, offering potential insights into dialectal shifts and continuity in that distinct regional area.

Optical Character Recognition (OCR) software and subsequent manual checking were used to digitize and correct the texts, facilitating accurate extraction of linguistic data. Orthographic variants for targeted tokens were identified and catalogued, recording detailed contextual information (including location within text and panel heading). This linguistic data was compiled into a structured format to enable systematic geographical comparison.

To capture orthographic variation likely to reflect underlying regional differences rather than mere idiosyncrasies, from each text we extracted all graphically distinct forms of a handful of high-frequency variables. From an initial list of ca. 30 candidate words (compiled from secondary literature on Early Modern orthography and our pilot readings), we shortlisted lexemes that met all of the following criteria:

- closed-class status, e.g. calendrical terms whose grammatical function and frequency are uniform across texts;
- multiple attestable spellings with each lexeme exhibiting at least two graphically distinct variants in the imprints;
- prosodic neutrality to ensure absence of stress-driven alternations;
- universal presence through occurrence in every exemplar's headings or entries, yielding comparable token counts across sites.

For each site and variable, we designated the dominant spelling as the variant with the highest raw count. No ties arose for our variables.

Results

For quantitative comparison the raw spellings extracted from each almanac were first sorted into analytically coherent variables. A variable is defined here as a set of orthographic realisations that differ in a single historically meaningful feature (e.g. presence vs. absence of final mute *-e*, retention vs. loss of *-eth*, etc). Each lemma was then assigned to one structural variable, defined not by the surface spelling itself but by the linguistic feature at issue. Tokens were tallied lemma-by-lemma and then collapsed into these variables, so that each represents the aggregate frequency of one variant within a given press-site. This procedure yields directly comparable figures and allows the discussion to focus on interpretable contrasts (innovation vs. conservatism, regional diffusion, etc.) rather than on isolated spellings.

Final-e deletion

Every form containing or lacking a final silent *-e* (after long vowel or diphthong) was counted once per occurrence in the calendar tables, headings and paratext of each document; totals were then grouped by lemma. Only clearly parallel items (year/*e*, four/*e*, fair/*e*, air/*e*, etc.) were tabulated, so that site-to-site comparisons remained valid.

Two clear patterns emerge:

1. A south-to-north gradient in the adoption of the innovative "e-less" forms.

London (1658) is already virtually categorical for *year*, *four*, *fair*, *air*, *wind* etc. Cambridge and especially York still favour the conservative *-e* spellings, while Oxford occupies an intermediate position, oscillating between old and new forms.

2. Token-by-token chronology that matches external historiography.

Final-*e* deletion is known to accelerate in metropolitan printing during the 1630s–50s (Horobin 2013; Salmon 1988). The London *Stationers' Company* imprint conforms to this timeline, whereas the provincial presses of York (Stephen Bulkeley) and Cambridge (John Field) continue Elizabethan spellings well into the 1660s. The data thus replicate, on a micro-corpus scale, the classic southern innovation to northern diffusion pathway repeatedly documented for Early Modern English orthography.

Taken together, the mute-*e* variable proves to be a sensitive, quantifiable proxy for orthographic modernity. Its geographical distribution in the four almanacs traces the leading role of London typography and the slower uptake in the Midlands–East Anglia corridor and the far North, providing an independent linguistic argument for the commonly assumed imprint-driven standardization centred on the capital.

Similarly, during the sixteenth century London compositors began simplifying the digraph in monosyllables, dropping the final *e* and writing *dry*, *try*, *cry* while leaving the vowel length to be inferred from context, which is the case for all almanacs save for Cambridge which alternates *dry/drie*, which on the surface looks like loss of the final *-e*, *dry* lost the final *-e* because the whole digraph was collapsing, whereas *year* or *four* lost a truly silent *-e*.

Media schwa deletion

A second, much rarer, variable concerns loss of unstressed medial *-e-* (historical /ə/). The innovative form *month* is categorical in the London imprint (4 tokens) and equally entrenched in the Oxford exemplar (3). Cambridge shows mixed usage (3 *moneth*, 3 *month*), while York remains fully conservative with *moneth*. The geographical cline is thus unmistakable: schwa-loss radiates outwards from the south-eastern metropolitan press, reaches Oxford by the mid-1650s, begins to penetrate Cambridge, and is still absent in the northern imprint. Orthographically, deletion of the medial vowel simultaneously indexes (i) an underlying phonological coalescence and (ii) the wider mid-seventeenth-century programme of pruning superfluous *-e*'s already noted by spelling reformers. A single additional item corroborates the pattern: *darkned* occurs only in the London almanac, reflecting the same schwa-syncope in the weak participial suffix *-en* that contemporary orthoepists describe as characteristic of London speech. The form never diffuses northward in our material and is later reversed even in the capital, where printers restore the etymological vowel under rising standardisation pressures. By contrast, the reduced spelling *month* stabilises and is retained nationwide.

Adjectival -ie/-y suffix

Throughout Middle and Early Modern English, the orthographic sequence *-ie/-y(e)* served as the default grapheme for unstressed /i:/ in adjectives and adverbs (e.g. *windie*, *rainie*, *frostie*). London compositors, steered by Mulcaster's advice that final *-y* was sufficient where the accent is light, progressively regularised the ending to bare *-y* after c. 1620, whereas presses on the northern and eastern periphery continued to print the earlier *-ie* for at least another half-century. The *-ie* to *-y* change behaves exactly like the mute *-e* loss examined earlier: an innovation diffusing outwards from London, first penetrating Oxford-stationers' copy, reaching Cambridge more slowly, and scarcely touching a small York press by mid-century. The patchy survival of *windie*, *frostie* etc. in East-Anglia therefore offers another useful dialectometric lever for dating and localising Early Modern printed matter. Because the alternation is orthographic, we cannot isolate direct phonetic correlates, yet contemporary orthoepists (Gil, Butler, etc.) note a concomitant qualitative change of the happyY vowel from tense [i:] to shortened [i], which plausibly motivated the graphic truncation.

Geminate-l Reduction

Geminate-l reduction is the term adopted here for the optional spelling contrast between Middle-English *-all* and its single-consonant reflex *-al* in words derived from the Latin suffix *-ālis* (e.g. *principall/principal*, *royall/royal*, *severall/several*, *festivall/festival*) and in the calendrical month-name *Aprill/April*.

The token set was restricted to those seven lemmas because (i) they are the only items that fluctuate inside almanacs under study, and (ii) each occurs at least once in more than one imprint, permitting cross-text comparison without genre bias.

In late Middle English scribal practice final *-ll* was analogical: French loans in *-al* were regularly respelt *-all*, and the doubling also served the phonotactic purpose of signalling a short or tense vowel in preceding position. The habit became conventional in high-frequency items and was still common in sixteenth-century printed calendars (*Aprill*, *Royall Exchange*). As Salmon (2000) states 1580–1660 was the key period when London compositors systematically cleared away redundant consonant doublings and spelling reformers such as Richard Mulcaster argued in *The Elementarie* that a second *l* was superfluous where no syllable division followed and should be removed in favour of the classical single-*l* form. In our data set the variable therefore operates not as a pronunciation or regional marker – the syllable is invariably /-əl/ and London, the linguistic centre, is already fully simplified – but as a micro-indicator of orthographic modernization. Its uneven distribution shows that, by the 1650s, geminate-*l* spellings survived only in provincial presses and documents the uneven uptake of spelling reform in the middle of the seventeenth century. If the doubled *-ll* still marked vowel quantity in the 1650s, we would expect parallel forms like *total*/*total* or *partial*/*partial* but those never occur. Instead: York prints *principall*, *royall* but *April*, *total*, *presbyterial*; Oxford prints *severall*, *principall*, *Aprill* but *festival*, Cambridge prints *severall*, *festivall* but *total*, *partial*, *usual*; London, the typographic pacesetter, prints only the single-*l* forms.

Present tense singular suffix

The present-tense third-person-singular suffix (*-s* vs *-(e)th*) was extracted from every finite verbal token in the four almanacs. Only forms that are unequivocally verbal and concord with a singular subject were

counted (e.g. *beginneth, riseth, passeth, beareth* alongside *begins, passes*). Auxiliaries (*hath, doth etc.*) were retained because they are the last stronghold of -th in Early Modern English.

Thus, out of 61 verb-final inflections, 36 still carry the conservative -th, but the innovative Northern -s already appears in calendrical clichés such as *begins* (London 4, Oxford 2, York 1). No Cambridge token adopts -s, confirming that press as the most conservative in the set.

Historically, -th (<OE -þ) was the southwestern standard until the late Middle Ages; -s, originally a northern dialect marker, diffused southwards along the London trade roads, gradually replacing -th except in the frozen auxiliaries. By the period covered here (1655–1666) the process is advanced but incomplete: London leads with the highest proportion of -s, provincial presses lag, and auxiliaries remain in -th across the board.

Although -s originated as a northern present-tense ending, by the mid-seventeenth century its diffusion did not work in a simple north and south divide. While c. 1650 speakers everywhere used -s in speech, -th still carried a learned aura in print as The King James Bible had already made -eth the prestige printed norm, so provincial compilers echoed that register to lend gravity. The York tally (18 -th, 1 -s) reflects orthographic conservatism and textual register, not local pronunciation. The same forces keep *hath* frozen in every imprint – including London – while more mundane verbs (*begins, passes*) begin to appear with -s only where printers feel a freer style is acceptable.

Discussion and conclusions

The examination of four mid-seventeenth-century almanacs confirms that this ostensibly formulaic genre is a sensitive register for tracing orthographic change. Across all variables analysed – mute-e deletion, medial-schwa loss, reduction of the adjectival -ie ending, geminate-l reduction, and the present-tense third-person-singular suffix – London emerges as the principal locus of innovation, with Oxford following at a respectful distance, Cambridge showing mixed practice, and York retaining the most conservative spellings. This south-to-north gradient aligns with established accounts of metropolitan norms radiating outward through the English print network and demonstrates that high-frequency calendrical prose can still reveal regional differentiation even after the initial wave of standardisation. At the same time, the data caution against equating spelling directly with pronunciation. The retention of -eth in York and, to a lesser extent, in Oxford reflects the prestige aura attached to biblical diction rather than a spoken northern residue; conversely, London printers restrict innovative -s to mundane verbs while freezing auxiliaries in -th, underscoring the role of register in modulating apparent dialect features. Likewise, the persistence of geminate -ll in select provincial lemmas is best understood as slow orthographic modernization: once typographic economy and Latinate etymology took hold in the capital, provincial presses adopted the single-l forms only gradually, retaining older standing type where costs or local expectation favored it.

Methodologically, the study shows that careful token selection and cross-text alignment can turn even a small corpus into a credible dialectometric probe. By comparing identical calendrical headings and prognostications, it is possible to filter out topical variation and focus on structural contrasts that map directly onto broader spatial patterns. The workflow – OCR conversion, regex harvesting, manual validation, and spatial plotting – scales readily to larger samples and promises finer-grained dialect atlases when applied to the hundreds of surviving almanacs in Early English Books Online. The absence of southwestern imprints in the present window remains a limitation, yet it also highlights future directions: integrating later Bristol almanacs will clarify the west-east axis, while a substantially enlarged dataset will allow multivariate modelling of press networks, compositor habits, and readership effects.

REFERENCES

Primary sources

Burton, W. (1655). An almanack for the yeare 1655 being the third since leap yeare : Wherein is contained the severall lunations, quarterly ingresse of the sun, and several tables of necessary use, with moderne transactions, a discription of highwayes, and principall faires in england and wales / by VV. burton. Oxford: Retrieved from <https://www.proquest.com/books/almanack-yeare-1655-being-third-since-leap/docview/2248552506/se-2>

Eagle 1666 a new almanack and prognostication for the yeare of our lord god 1666 : Being the second after bissextile of leap-yeare : Wherein is contained the feasts both moveable and fixed, the signes, [sun] rising and setting, [sun] place, the moon rising, full : Sea in most of the principall parts in england, with the change, full, and quarters of the moon, and weather, as it will most incline to, from the quarter to quarter : As also the eclipse, and quarters of the year. &c. fitted to the meridian of the north, whose latitude is 53d 40m and longitude 20d. 12m. ad will serve for the whole kingdome of england (1666). York: Retrieved from <https://www.proquest.com/books/eagle-1666-new-almanack-prognostication-yeare-our/docview/2240915461/se-2>

F., M. (1663). An almanack for the yeare of our lord god 1663 being the third after bissextile or leap-yeare. by M. F. philomath. Cambridge: Retrieved from <https://www.proquest.com/books/almanack-yeare-our-lord-god-1663-being-third/docview/2240902817/se-2>

Jinner, S., fl.1658–1664. (1658). An almanack or prognostication for the year of our lord 1658 being the second after bissextile or leap year : Calculated for the meridian of london, and may indifferently serve for england, scotland, and ireland / by sarah jinner ... London: Retrieved from <https://www.proquest.com/books/almanack-prognostication-year-our-lord-1658-being/docview/2240936054/se-2>

Secondary sources

Bergs, A., & Brinton, L. (Eds.). (2017). *Early Modern English* (Vol. 4). De Gruyter Mouton. Chapter 8, "Dialects", by A. Meurman-Solin, 128–149. <https://doi.org/10.1515/9783110525069-008>

Dixon, D. (1951). *Alexander Gil's "Logonomia Anglica" edition of 1621 translated with an introduction and critical and explanatory notes* (Doctoral dissertation, University of Michigan). ProQuest Dissertations & Theses Global.

Fisiak, J. (1995). Standardization, printing, and the evidence for local dialects: The case of Early Modern English kirk. In W. Winter (Ed.), *On languages and language: The presidential addresses of the 1991 meeting of the Societas Linguistica Europaea* (145–166). De Gruyter Mouton. <https://doi.org/10.1515/9783110881318.145>

Horobin, S., & Smith, J. (2002). *An introduction to Middle English*. Blackwell.

McEnery, T., & Baker, H. (2019). Language surrounding poverty in Early Modern England: A corpus-based investigation of how people living in the seventeenth century perceived the criminalised poor. In *Historical sociolinguistics: Language change in Tudor and Stuart England* (226–251). Brill.

Nevalainen, T. (2006). *An introduction to Early Modern English*. Edinburgh University Press.

Nevalainen, T., & Raumolin-Brunberg, H. (2003). *Historical sociolinguistics: Language change in Tudor and Stuart England*. Longman.

Ray, J. (1674/2015). *A collection of English words not generally used: With their significations and original, in two alphabetical catalogues* (P. Schintu Martínez, Trans.; J. Ruano-García, Rev.). Salamanca Corpus, Universidad de Salamanca. <https://salamancacorpus.usal.es>

Rissanen, M., Kytö, M., & Palander-Collin, M. (Eds.). (1996). *The Helsinki Corpus of English Texts: Diachronic part: Early Modern English section*. VARIENG – Research Unit for the Study of Variation, Contacts and Change in English. <https://varieng.helsinki.fi/CoRD/corpora/HelsinkiCorpus/earlymodern2.html>

Ruano-García, J. (2012). On the enregisterment of the northern dialect in Early Modern English: An evaluation across literary text types. In *At a time of crisis: English and American studies in Spain* (60–70). Departament de Filologia Anglesa i Germanica.

Ruano-García, J. (2020). Shaping the other in the standardization of English: The case of the "northern" dialect. *International Journal of English Studies*, 20(2), 185–205.

Salmon, V. (Ed.). (1996). *Language and society in Early Modern England: Selected essays 1982–1994*. John Benjamins.

Salmon, V. (2000). Orthography and punctuation. In R. Lass (Ed.), *The Cambridge history of the English language* (Vol. 3, 13–55). Cambridge University Press.

Taavitsainen, I. A. J., & Schneider, G. (2019). Scholastic argumentation in early English medical writing and its afterlife: New corpus evidence. In C. Suhr, T. Nevalainen, & I. Taavitsainen (Eds.), *From data to evidence in English language research* (191–221). Brill. https://doi.org/10.1163/9789004390652_010

Отримано редакцією журналу / Received: 12.05.25

Прорецензовано / Revised: 04.06.25

Схвалено до друку / Accepted: 11.06.25

Опубліковано / Published: 07.05.26

Анна КАРАБАН, канд. філол. наук, доц.

ORCID ID: 0000-0002-1101-4227

e-mail: a.karaban@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

ОРФОГРАФІЧНА ВАРІАТИВНІСТЬ В АНГЛІЙСЬКІЙ МОВІ СЕРЕДИНИ XVII СТОЛІТТЯ: КОРПУСНЕ ДОСЛІДЖЕННЯ ЛОНДОНСЬКОГО, ОКСФОРДСЬКОГО, КЕМБРИДЖСЬКОГО ТА ЙОРКСЬКОГО АЛЬМАНАХІВ

Вступ. Ранньомовноанглійська мова є ключовою для розуміння переходу від регіоналізованого письма до над-регіонального стандарту, проте систематична діалектна картографія цього періоду залишається мало-дослідженою. Попередні дослідження зосереджувалися на літературних та юридичних жанрах; популярні альманахи, хоча й друкувалися щорічно в кожному великому місті, залишалися поза увагою. Оскільки вони фіксують місце друку на титульній сторінці, повторюють ідентичні календарні вирази та поширюються великими накладками, альманахи є недостатньо використовуваними джерелами орфографічних практик

середини XVII ст. Це дослідження оцінює, чи можуть географічно специфічні написання в чотирьох стародруках простежити діалектні межі та поширення столичних норм.

Методи. Усі чотири тексти були оцифровані за допомогою оптичного розпізнавання тексту (OCR) і відкориговані вручну. Було відібрано п'ять високочастотних змінних: вилучення німого -e, вилучення медіального шва, редукція -ie/-u у межах прикметників, редукція подвоєного -l та варіація суфікса теперішнього часу третьої особи однини. Токени підраховувалися лише в структурно паралельних календарних заголовках і прогнозах, а потім зводилися до загальних показників на рівні змінних. Якісний аналіз копій джерел та жанрових конвенцій контекстуалізував кількісні закономірності.

Результати. Три змінні: німе -e, вилучення шва та суфікс -ie/-u демонструють чітку дифузію з півдня на північ: Лондон є абсолютним провідником інноваційного правопису, Оксфорд наближається до Лондона, але зберігає тонкий консервативний прошарок, Кембридж займає змішану позицію, а Йорк залишається майже повністю традиційним. Зредуковане l виявляється не діалектним, а показником орфографічної модернізації: лише провінційні друкарні зберігають усталені форми, тоді як Лондон повністю релативізований. Суфікс теперішнього часу третьої особи однини поводить протилежним чином: незважаючи на північне походження, інноваційний -s зустрічається рідко, оскільки друкарі по всій країні зберігають біблійний -eth у книжкових текстах; активне використання -eth у Йорку відображає стилістичний консерватизм, а не місцеве мовлення.

Висновки. Альманахи фіксують як консервативну дикцію, так і нові столичні норми, фіксуючи тонкі грані орфографічних змін у вирішальне десятиліття стандартизації. Градієнт з півдня на північ підтверджує лідерство Лондона в реформі правопису, тоді як варіативна поведінка підкреслює посередницьку роль жанру та регістру. Масштабування процедури на більшу кількість провінційних альманахів може дати змогу створити діалектний атлас на основі даних, який поглибить наше розуміння регіональних кордонів ранньомовноанглійських діалектів та механізму мовних змін на основі друкованих видань.

Ключові слова: ранньомовноанглійська мова; альманахи; діалектологія; історична лінгвістика; регіональна варіативність; ізоглоси; орфографія; корпусне дослідження

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analysis or interpretation of the data; in the writing of the manuscript; or in the decision to publish the results.