

as taken with a violent fit of trembling. five minutes, ten minutes, a quarter of an hour  
e party, which was not dispelled for full five minutes. After it had passed away, they wer  
, which would bring in, if obtained, full five -and-sixpence vISSN 2299-5900 g Crat  
hilling. Come back with him in less than five minutes and I'll give you half-a-crown!" Th  
It shake his arm off. He was at home in five minutes. Nothing could be heartier. His nie  
autions. That, the proof would go back five years, and would show the prisoner already  
w many times? Two or three times. Not five or six? Perhaps. Of what profession? Gent

# TOKEN

## A Journal of English Linguistics

s nat'rally onwholesome; and we've had five dirty and onwholesome children, as is all d  
thout the least consideration, to bestow five pounds on the brickmaker. "I made ten po  
persuaded out of the sacrifice of these five pounds by being convinced that it would d  
it. "Let me see!" he would say. "I saved five pounds out of the brickmaker's affair, so if  
er as he laid it on the table, "more than five and forty years ago. He was then the most  
figure for a lengthened period. A rise of five has since taken place, and a further rise of  
s since taken place, and a further rise of five is guaranteed at the expiration of a term n  
es, &c., &c., &c. It is somewhere about five or six o'clock in the afternoon, and a balm  
work, and that if you gave him out, say, five and forty folio on the Wednesday night, yo  
bed, whose path in life has lain through five and forty years, lies there with no more tra  
ortense, been in my Lady's service since five years and always kept at the distance, and t  
oofs of the plan. She's going to put out five thousand new circulars, and she knows you  
ttle furniture was a mite of a boy, some five or six years old, nursing and hushing a hear



### Volume 5/2016

office, down to Zoodle, shall set right in five hundred years--though born expressly to d  
you are to move on. I have told you so five hundred times." "But where?" cries the bo  
ays Jo, with dirty tears, "fur I had to pay five bob, down in Tom-all-Alone's, afore they'd  
d then a young man he thieved another five while I was asleep and another boy he thie  
of more money. Now the consequence's five mile off, she better for all parties." Wish th

---

Token: A Journal of English Linguistics

Volume 5

---

JAN KOCHANOWSKI UNIVERSITY OF KIELCE

# Token: A Journal of English Linguistics

Volume 5

Edited by

John G. Newman

Marina Dossena

Sylwester Łodej



Jan Kochanowski University Press

Kielce 2016

#### **EDITORS**

John G. Newman (University of Texas Rio Grande Valley)

Marina Dossena (University of Bergamo)

Sylwester Łodej (Jan Kochanowski University of Kielce)

#### **LANGUAGE EDITOR**

Sheila Dooley (University of Texas Rio Grande Valley)

#### **ADVISORY BOARD**

John Anderson (Emeritus Professor, University of Edinburgh, Scotland)

Michael Bilynsky (Lviv National University, Ukraine)

Fran Colman (Emeritus Reader, University of Edinburgh, Scotland)

Hans-Juergen Diller (University of Bochum, Germany)

Małgorzata Fabiszak (Adam Mickiewicz University, Poland)

Jacek Fisiak (Emeritus Professor, Adam Mickiewicz University, Poland)

Elena Gomez-Parra (University of Cordoba, Spain)

Robert Kiełtyka (University of Rzeszów, Poland)

Grzegorz A. Kleparski (University of Rzeszów, Poland)

Zoltán Kövecses (Eötvös Loránd University, Hungary)

Marcin Krygier (Adam Mickiewicz University, Poland)

Lilo Moessner (University of Aachen, Germany)

Rafał Molencki (University of Silesia, Poland)

Heinrich Ramisch (University of Bamberg, Germany)

Piotr Ruskiewicz (Pedagogical University of Cracow, Poland)

Aleksander Szwedek (Emeritus Professor, Adam Mickiewicz University, Poland)

Irma Taavitsainen (University of Helsinki, Finland)

Akinobu Tani (Hyogo University of Teacher Education, Japan)

Jerzy Welna (University of Warsaw, Poland)

**Cover design** Jakub Patryk Łodej

**Formatting** Marzena Buksińska

© Jan Kochanowski University of Kielce Press 2016

#### **Editorial Correspondence**

Uniwersytet Jana Kochanowskiego w Kielcach

Instytut Filologii Obcych

ul. Świętokrzyska 21D

25-406 Kielce

Poland

[www.ujk.edu.pl/token](http://www.ujk.edu.pl/token)

[token@ujk.edu.pl](mailto:token@ujk.edu.pl)

*Token* is the original medium of publication for all articles that the journal prints.

*Token* is not an electronic publication.

This publication has been supported financially by the Institute of Foreign Languages, Jan Kochanowski University.

# **The language of medicine in the *Philosophical Transactions*: Observations on style**

Elisabetta Lonati

*Università degli Studi di Milano*

## **ABSTRACT**

Medicine is one of the most fully represented disciplines in the *Philosophical Transactions* (*PT*), particularly in the materials dating from before the establishment of medical societies and of specialized journals. Medical events were recorded and described, data were collected, interpreted, and discussed. The need for a faster communication among professionals brought about new written forms. Shorter texts were adopted to exchange up-to-date information.

This investigation focuses on a selected number of texts to verify the origin and the nature of any rhetorical and stylistic changes which may have occurred in the *PT* during the eighteenth century. The selection of extracts constitutes the basis for a detailed discussion of rhetorical and stylistic issues.

Between 1702 and 1801 medical writing in the *PT* undergoes major changes: these are gradual shifts along a continuum highlighting an essentially author-centered approach at the outset of the century and an object-centered perspective at the end of the period.

**Keywords:** eighteenth century, *Philosophical Transactions*, medicine, research article, style.

## **1. Introduction**

The *Philosophical Transactions* of the Royal Society of London were established in 1665 as a “private venture” (Gross et al. 2000: 372) by its secretary Henry Oldenburg. He was not a man of science but he strongly believed in the advancement of learning through widespread, systematic and periodical

communication (Bazerman 1988: 129). He thought that the Society needed a place in which members could report observations on natural events, case studies, experiments, research results and anything that could be of interest to spreading new ideas and new discoveries. The elaboration of up-to-date knowledge and the possibility to communicate it in the vernacular – particularly by letter – encouraged social interaction within the Society itself and across other domestic and foreign societies, since the *Philosophical Transactions* (hereafter *PT*) were “meant to serve as a newsletter, to favour the spread of news within the Royal Society and other learned circles” (Gotti 2011: 204). Henry Oldenburg, besides inventing “the modern scientific journal” (Atkinson 1996: 334-335), helped establish a vast scientific network and “bring together a previously dispersed scientific community” (Bazerman 1988: 130). He actually was the first editor – and author as well – of that huge correspondence which constituted the bulk of the first issues of the *PT* and characterized much of seventeenth- and eighteenth-century writing across disciplines (Gotti 2014: 151-156).

The *PT* also represented the outcome of a kind of “cooperative endeavor [...] of individuals with common interests who were engaged in work that was commonly perceived to be of value” (Dear 1985: 147). This does not mean that these individuals worked together: even though the Royal Society failed in supporting shared projects and “was more of a club than a college” (Dear 1985: 147), the *PT* acted as a “collective voice” (Dear 1985: 147). They served to highlight the members’ shared interests and shared values on the new approach to knowledge. These interests and these values rejected the principles of authority and the traditional outlook on reality based on it, in favour of an innovative approach focused on direct experience and experimentation. The Royal Society represented such epistemological innovation; it embodied intellectual inquiry within a micro-community of expert and non-expert members involved in the most diverse domains.

Among the multifarious disciplines, medicine is one of the most fully represented in the *PT*, particularly in the materials dating from before the establishment of medical societies in London and of specialized journals in the 1780s. The approach to medical research had been changing over time and medical activity had been steadily increasing during the seventeenth and eighteenth centuries. This change of perspective in investigating medical events, that is the introduction of a more practical outlook on them with the aim of recording, describing, collecting data and later interpreting them to generalize medical issues, highlighted the need of new written forms. These were different from the most traditional ones, such as treatises and long dissertations. Shorter texts were useful to support a faster

communication among professionals (naturalists, physicians, anatomists, scholars, practitioners, surgeons, etc.) and non-experts alike. Moreover, the inclusion in the *PT* of “frequent references to other contemporary scientists working on the same topics” (Gray et al. 2011: 224-225) promoted and supported discussion and debate over long distances as well. This kind of writing also helped make medical innovations and case studies known to a wider reading public, thus establishing a vast informative network (Fontes Da Costa 2009). The scientific journal, both general and specialized, answered all these needs in a new communicative and dynamic context. According to Bazerman,

the invention of the scientific journal necessitated the invention of the scientific article. The experimental report, as any other literary genre, was invented in response to a literary situation and evolved through the needs, conceptions, and creativity of the many authors who took it up. [...] the genre of experimental article has origins in essay, epistolary, and journalistic writing of the seventeenth century [...] the internal dynamics of scientific communication within a journal forum reshape the initial sources to create a new communicative form, powerful enough to influence other forms of communication and the social structure of the community which uses it. (Bazerman 1988: 59, 63)

This means that what is now known as the ‘research article’ underwent major rhetorical and stylistic adaptations over time, particularly concerning discourse and linguistic features. The growth in the number of texts ran parallel to a radical change in medical prose within a more general development of scientific writing (cf. Banks 2008).

## 2. Source texts

The texts under scrutiny here are 46 epistolary and non-epistolary medical articles published in five volumes issued at intervals of about twenty-five years, between 1702 and 1801. They represent about 10% of all medical articles issued in the *PT* across the century. Articles in Latin (which are the minority) are excluded from this analysis.

The medical article is a broad categorical type including epistolary articles, variously defined in the *PT* as *letter*, *part of a letter*, *extract(s) of a letter*, *account ... in a letter*, etc., and non-epistolary articles, labelled as *account*, *instances*, *case(s)*, *relation*, *observation*, *practice*, *abridgement*, *attempt*, *remarks*,

*examination, proceeding, explication, discourse, description, dissertation, thoughts, essay, method, lecture*, etc. These denominations hardly correspond to genre differentiation (Atkinson 1999: 80-81); overlapping and merging are common<sup>1</sup>.

Volumes 23 (for the years 1702-1703, issued 1704), 47 (for the years 1751-1752, issued 1753), and 91 (for the year 1801, issued 1801) are the most relevant.

Volume 23 marks the turn of the century and also marks, under the influence of Newton, the beginning of a change in the rhetorical strategy to framing science. Newton perceived the *PT* as a “vehicle for concrete findings [...] so that they appear as concrete facts, [...] even though the events that made these facts visible [...] occurred in a private laboratory as the result of speculative ponderings and active experimental manipulations” (Bazerman 1988: 90). This means that Newton introduced new discourse practices to fit a new rhetorical situation in writing science, since he envisaged “the narrative of the scientist operating under procedures [...as] the main rhetorical resource to establish the credibility of the events and conclusions” (Bazerman 1988: 92).

Volume 47 establishes peer reviewing for the publication of articles. For previous volumes, the editorship was carried out by the secretary, whereas from 1752 a more institutional and controlled approach was introduced and proposals for vol. 47 were to be approved by a committee. In the Advertisement to the reader it is declared that:

THE Committee appointed by the *Royal Society* to direct the publication of the *Philosophical Transactions*, take this opportunity to acquaint the public, that it fully appears, as well from the council-books and journal of the Society, as from the repeated declarations, which have been made in several former *Transactions*, that the printing of them was always, from time to time, the single act of the respective Secretaries, till this present XLVII. volume. [...]

<sup>1</sup> According to Gotti (2011: 209), “As regards the *PT*, [...] this journal greatly contributed to the textual development of specialized discourse by giving rise to a number of new genres in English. The various issues of the journal consist of a variety of text types, referred to in the journal itself by several different titles: for example, ‘accounts’, ‘calculations’, ‘comparisons’, ‘descriptions’, ‘essays’, ‘experiments’, ‘explanations’, ‘extracts’, ‘inquiries’, ‘investigations’, ‘letters’, ‘observations’, ‘proposals’, ‘remarks’ and others. The same titles were often used to introduce very different types of texts, and similar text types were often referred to by different names. Indeed, at this time, textual forms were still very loosely structured, and there were not yet any clear conventions and codifications concerning them”.



But the Society being of late years greatly enlarged, and their communications more numerous, it was thought adviseable, that a Committee of their Members should be appointed to reconsider the papers read before them, and select out of them such, as they should judge most proper for publication in the future *Transactions*; which was accordingly done upon the 26 of March 1752. And the grounds of their choice are, and will continue to be, the importance or singularity of the subjects, or the advantageous manner of treating them; without pretending to answer for the certainty of the facts, or propriety of the reasonings, contained in the several papers so published, which must still rest on the credit or judgement of their respective authors.

It is likewise necessary on this occasion to remark, that it is an established rule of the Society, to which they will always adhere, never to give their opinion, as a body, upon any subject, either of nature or art, that comes before them. (*PT* 1751-1752, 47, Advertisement)

Volume 91 marks the end of the period as a whole as well as the end of the last quarter of the century, characterized by the British Medical Reform, the professionalization of medical education and activity, and the establishment of medical journals in London.

Volumes 35 (for the years 1727-1728, issued in 1729) and 67 (for the year 1777, issued 1777) highlight two middle steps, since they mark the twenty-five years' interval established for the present rhetorical and stylistic analysis. On the one hand, volume 35 marks the end of the Newtonian administration and hence the possibility to verify whether any stylistic change took place in scientific writing under his influence. On the other hand, volume 67 marks both the end of the twenty-five years after the introduction of the peer reviewing procedure in the *PT*, and the beginning of major changes in the organization and implementation of medical knowledge in the contemporary medical community.

### 3. Aims and Method

My investigation focuses on a selected number of texts to verify the origin and the nature of any stylistic changes which may have occurred in the *PT* – and in the community of writers producing texts – throughout the eighteenth century.

The focus and the approach are primarily qualitative, even though a basic quantitative survey has been used to frame the rhetorical and stylistic analyses and discuss them in a more inclusive context.

Some major studies, such as Biber (1988) and Biber – Finegan (1989) on a multidimensional (MD) approach, have helped establish the methodological perspective. Biber (1988: 21) maintains that “linguistic variation must be analyzed in terms of sets of co-occurring features”, since linguistic features considered in isolation cannot account for systematic variation. Linguistic features can be grouped into ‘dimensions of variation’, and these “dimensions comprise those features that actually co-occur [in texts], rather than a set of features that the researcher expects to co-occur” (Biber 1988: 23). Recurring linguistic features tend to be pervasive in a text – or in a set of texts – and to indicate a specific communicative function (here conveying scientific information): in this case, they become conventional *stylistic* features, rather than *functional* register features (Biber – Conrad 2009: 53-55). However, co-occurring features may belong to different dimensions, that is to say, “several overlapping co-occurrence patterns within any set of linguistic features” (Biber 1988: 23 and 101)<sup>2</sup> can make it difficult to identify a single dimension and/or a specific communicative function. The style(s) of medical writing in the *PT* may thus be characterized by different dimensions of variation.

Bazerman (1988), Atkinson (1996), Gross et al. (2000) and Gotti (2011) have analysed the research article in the *PT* as a newly developing genre. In particular, Atkinson’s works (1996 and 1999) applying Biber’s MD approach (1988) and Bazerman’s study (1988: 62) on the experimental article in the *PT* focusing on the “social dimension” of encoding scientific writing have served as background models.

The present research considers two major interdependent perspectives: a chronological one and a stylistic one. In other words, it discusses historical change in the linguistic style(s) of research articles (Biber – Conrad 2009: 162).

The chronological perspective focuses on stylistic changes in medical writing across time.

The stylistic perspective focuses on the pervasive and co-occurring linguistic features of medical articles, that is to say on their style. According to Biber – Conrad,

---

<sup>2</sup> Biber’s dimensions are six (1988: 122 and 199): 1. Involved vs. Informational; 2. Narrative vs. Non-Narrative; 3. Explicit vs. Situation-Dependent Reference; 4. Overt Expression of Persuasion; 5. Abstract vs. Non-abstract Information; 6. On-Line Informational Elaboration.

The style perspective incorporates the same kind of linguistic analysis as the register perspective: a relatively comprehensive analysis of core lexical and grammatical features [...]. However, the basis of comparison is different from the register and genre perspectives. That is, the style perspective is usually used to compare texts within a single genre/register [...]. In this case, linguistic differences are not associated functionally with the situational context, because all of these texts are produced in similar situations. (Biber – Conrad 2009: 71-72)

According to the MD approach, the analysis is based on the following dimensions of variation, each dimension being characterized by two poles along a continuum. Polarity is highlighted between brackets:

1. Involved (face-to-face interaction) vs. Informational (academic prose)
2. Narrative (fiction, account, etc.) vs. Non-narrative (academic prose, official document, etc.)
3. Abstract (academic prose) vs. non-abstract (face-to-face conversation)
4. Situation-dependent (listener's knowledge of imminent context: place, time, other adverbs, etc., ex. personal letters) vs. Elaborated/Explicit Reference (official documents, professional letters, academic prose, etc.)

This clearly implies a focus on the role of the writer as emerging from the text(s) and, especially, the relationship between the author (writer, reporter, eye witness, researcher, etc.) and the object (topic, case study, personal experience, professional experience, etc.) of the discussion, narration, account, report, description, etc. or, in other words, the relationship between an author-centered perspective and an object-centered perspective.

According to Atkinson (1992: 338 and 340), “an integrated linguistic and rhetorical description of the *written medical research article*” highlights the linguistic consequences of textual and discursive organization: this means that the relationship between linguistic choices, stylistic features and rhetorical-discursive context is a fundamental one.

This also means that the authors organize their writing according to the communicative habits and the communicative needs of their time and their audience of experts and/or non-experts alike. Indeed, the lay interest and participation in medical events and their communication is pervasive throughout the century, and medicine continues to be a frequent topic of civil conversation.

#### 4. Results: volumes contents and organization

In the century between 1702 and 1801 (from volume 23 to volume 91), the *PT* include about five hundred medical articles in English (those in Latin represent a minority, particularly in the second half of the century: just 5 out of 26 were published after 1750). About three hundred texts – included those in Latin – belong to the period between 1702 and 1750 (from vol. 23 to vol. 46), whereas between 1751-1801 (from vol. 47 to vol. 91) the number of texts strongly decreases. Moreover, in the last quarter of the century, some volumes do not deal with medical topics and when medical texts are included they are usually long articles but few in number. A significant reduction in the number of medical texts included in the *PT* is evident between 1775 and 1801: during these twenty-five years only about 30% of medical articles issued between 1751 and 1801 are published.

As regards the five volumes analysed, the number of letters included varies in each volume, even though the epistolary form is the most common one used by medical writers of the period to exchange and convey scientific information. Moreover, the length of articles varies greatly across time, from single-page articles to much longer ones: the average length in the first three volumes is about three pages, about seven in volume 67 and about 23 in volume 91. These long articles do not constitute the norm, and they are all included in volume 91. Volumes 35 (1727-1728) and 91 (1801) include 11 and 5 medical articles respectively, but none of them include letters. Volumes 23 (1702-1703), 47 (1751-1752) and 67 (1777) include 8, 18 and 4 medical articles respectively, and each of them also includes letters: 6 out of 8 in volume 23, 9 out of 18 in volume 47, and 2 out of 4 in volume 67. In the first half of the century, letters embrace medical contents and the writers use a more personal style than they do later on: in volume 67 letters do not embrace medical contents, they are short, institutional cover letters simply introducing topics. The writers are far less involved in the accounts.

Most of the articles convey practical information, without any theoretical frame: they are rich in details, particularly those related to first-hand witnessing.

In the following section, a selection of extracts drawn from the five volumes is used as the basis for a detailed discussion of stylistic features and their discursive effects.

#### 5. Discussion: Stylistic features in context

This section includes two sets of extracts: the first group exemplifies and contextualizes the linguistic and stylistic features; the second group

principally concerns the structure and the rhetorical and discursive elaboration of medical articles.

The first example is an unplanned and situation-dependent 'account', reflecting the emotion and the memory of the author-writer-witness (author-centered approach), since it records his personal and familiar experience:

(1) Author-centered rhetoric

In your last [letter] you desired an account of my Father's Cancer, which I send you, as near as I can remember, it being 20 years since he dyed, and I being then but young, [I] could not make those remarks upon it, as another might have done, and it's possible [I] might forget something material too.

It took its rise from a small bruise on the *Os Jugale*, and in process of time spread it self over the whole Cheek; [...] it ulcerated his Eye round, which I saw him take out with his own hand; and afterwards extended it self to his Ear, and through his Cheek into his Mouth, and across the upper part of his Nose, and perforated the Bone there: It likewise overrun that side of his Forehead, fouling the *Os Frontis*, which came away in pieces [...] the *Cranium*, in a few days putrified and exposed the Brain it self, and several portions of it came away [...].

This to the best of my remembrance is the summ of all. (PT 1702-1703, 23: 1069-1070)

This extract highlights a conversational-'dialogic' style and social (face-to-face) interaction within the disciplinary discourse community but also a friendly-involved relationship with the interlocutor (civil manners of conversation), opening and closing the passage: *In your last, you desired, I send you, [...] This to the best of my remembrance is the summ of all.* The second paragraph is the 'account proper' in narrative style, whose main features are represented by the use of the 3<sup>rd</sup> person singular pronoun for the father (*his/him*) and for the Cancer (*it/its/it self*), which is at the heart of the narration; dynamic verbs in the past tense (perfective aspect: *took, spread, ulcerated, extended, perforated, overrun, came away, putrified, exposed, came away*) describe the subsequent steps in the process of physical dissolution, given as a series of hallucinative images. The text is not cohesive at all: the topic *cancer* (full-noun) is used just once and then substituted by the pronouns *it, its, it self*, also used to refer to the 'brain' (*it self, it*). The attention is focused on the observer-reporter and not strictly on the object.

The second example actually includes two extracts on dissection (the 1<sup>st</sup> dated 1702-03; the 2<sup>nd</sup> dated 1752) and emphasizes the shift from an author-centered orientation to a more and more detached attitude in the observation, account and description of the medical event:

(2)

(a) Dissection – author-centered orientation (1702-1703)

The Case being not very usual, I applied my self [...] in order to Dissect him; which [...] I did the next day. Finding the Liver only something larger than ordinary, I immediately made search from the *Ventriculus* quite to the lower end of the *Intestinum rectum*: The *Ventricle* was considerably extended, a little space from the Gut *Ileon*, in the *Jejunum*, I found [...] I then proceeded to the Gut *Ileum*, I found a considerable part of it very livid, [...]. I found [...] I found another large ruption [...].

There was some other small matters to be seen, not worth while for me to mention to you. However, Sir, this being matter of fact, as a great many of his Relations can testifie, they being present, I thought fit to acquaint you with it. (PT 1702-1703, 23: 1245)

(b) Dissection – ‘concise accumulation’ (1751-1752)

That day we had him open’d. The lungs were found full of blood. Water in the *pericardium* in the usual quantity. The blood in both ventricles of the heart fluid. The *oesophagus* without any morbid appearance [...]. The *aspera arteria* full of such frothy substance as came from his mouth. The stomach fill’d with liquor, notwithstanding the small quantity he had drank [...]. No other parts were examined. (PT 1751-1752, 47: 414)

Extract (2a) is completely author-centered and written in narrative-descriptive style: the writer accounts for the several steps in the process of dissection as they emerge and are determined by his own choice, for example *I applied myself, I did the next day, I immediately made search, I found, I then proceeded, I found*, etc. It is a discovery process, strictly situation-dependent: *The case being not very usual*. The last paragraph goes back to the conversational style of civil and social interaction (*not worth while for me to mention to you, I thought fit to acquaint you with it*) and highlights the truthfulness of the report itself, a rhetorical strategy which plays a key role in eighteenth-century medical communication and ethics (*this being matter of fact*).

In (2b) the shift is clearly marked and new stylistic features, such as the 1<sup>st</sup> person plural pronoun referring to the disciplinary community and the actual team carrying on the dissection and the deletion of linking verbs (the verb to be is reduced to Ø), are introduced: *Water in the pericardium, the blood [...] fluid, the œsophagus without, the stomach fill'd*, etc. The effect is that of concise accumulation, listing and clustering of results between two dynamic poles constituting the rhetorical framework: *we had him open'd [...] No other parts were examined*.

Example 3 is an extract from a short account opening the last quarter of the century and marking a further shift towards a more abstract, object-oriented attitude in elaborating medical discourse:

(3) Abstract object-oriented attitude (1777)

The conjecture that had been formed about the complaint in the bowels proved to be [...]. Certainly these were some signs of a slight inflammation having attacked the membranes investing the contents of the thorax.

Neither can we suppose such appearances to have existed without occasioning some uneasiness: they were, perhaps, sufficient to account for that great tenderness and oppressive pain which the doctor felt from the least pressure on the sternum, or upon any part of the breast near it.

The principal seat of the disease which proved so tedious, and in the end so fatal, was, no doubt, confined to the colon only; [...]. The part first affected must have been that portion of the canal in which we observed the most mischief.

The superficial extent of the disease over so large a surface as the whole arch of the colon, [...] distinguished the part [...]. (PT 1777, 67: 610-612)

Lexico-grammatical features are particularly interesting here, especially nominalizations which transform dynamic verbs (processes) into entities, and reduce grammatical intricacy: *the conjecture that [...] proved, a slight inflammation having attacked*, etc. They are used to thematise the central content of the discussion and organize the transitivity structure.

As regards morpho-syntactic features, the main strategy is the use of complex phrases whose head nouns and adjectives are post-modified by a series of syntactic units emphasizing more elaborate and detailed conceptualizations. Post-modification is realized by full relative clauses



(the conjecture that, oppressive pain which, the disease which, the canal in which), reduced relative clauses (inflammation having attacked the membranes investing the contents, the part first affected), non-finite to-infinitive clauses (appearances to have existed, sufficient to account for), prepositional phrases (about the complaint in the bowels, signs of a slight, the contents of the, [...]) from the least pressure on the sternum, or upon any part of the breast, the principal seat of the disease, extent of the disease over so large).

Some other rhetorical devices, particularly hedging, highlight manners and tentativeness in discourse: *Certainly, neither can we suppose, perhaps, no doubt, must have been.*

In general terms, we move toward an increase in the usage of embedded items within a complex recursive structure.

The second set of examples principally concerns the organization of the scientific-medical article. Structural-discursive features may be found as isolated hints in the first half of the century, with an increase in more institutionalized articles under Newton (1703-1727), but they acquire more definite status towards the end of the period, particularly after 1777.

These features include introductory paragraphs (4a and 4b, and 6: topicalization, background information, definition, dispute), and concluding paragraphs (5a and 5b and 6: generalization and abstraction, further research):

(4)

(a) Setting the scene: introductory sections – Aneurysm 1

AN Aneurysm, without Doubt, is a Tumor arising from some Disorder in an Artery; but what that Disorder is, or whence it arises, is not so well agreed, the Accounts which are given of it, being widely different and uncertain.

The Name seems to imply, that it is a Dilatation of the Vessel; [...]. (PT 1727-1728, 35: 436)

(b) Setting the scene: introductory sections – Aneurysm 2

An Aneurysm is by all Authors defin'd to be a soft circumscrib'd Tumor, in which there is a sensible Pulsation, contemporary with the Pulsation of the Artery, to which it adheres.

As it is certain, that any Tumor of what Kind soever, lying on, or adhering to any considerable Artery, must necessarily be moved by every Pulsation of such Artery, so this Pulsation (unless understood



in such Manner as I shall hereafter explain) can no ways be admitted as the true Diagnostick, whereby to specify the Difference between this kind of Tumor and any other.

An Aneurysm is found, [...]

It is obvious that, [...].

(PT 1727-1728, 35: 440-441)

The two texts show that the topic is

- 1) emphasized, that is thematized, and that the discussion is not situation-dependent (*An Aneurysm [...] is, An Aneurysm is*);
- 2) defined, to make the discussion start, by delimiting and clarifying as regards the topic (*is a Tumor arising from, The name seems to imply, it is a Dilatation of the Vessel, is [...] defin'd to be [...] a soft*);
- 3) contextualized within the disciplinary background of a disciplinary community (*An Aneurysm, without doubt, is [...], An Aneurysm is by all Authors defin'd to be, As it is certain, It is obvious that*); and
- 4) problematized, particularly when putting forward new viewpoints and even critical and/or opposing approaches concerning previous ideas (*what [...] or whence [...] is not so well agreed; the Accounts [...] being widely different and uncertain; and Pulsation (unless understood in such manner as I shall hereafter explain) can no ways be admitted as the true Diagnostick*).

Textual cohesion is mainly enhanced by lexical repetitions (*Aneurysm, Pulsation*).

As regards concluding remarks, the two sets of examples are dated 1702-1703 and 1727-1728, respectively at the beginning and at the end of the Newtonian period. This chronological difference marks some rhetorical and stylistic differences between them:

- (5)
- (a) Generalization in concluding remarks 1 (isolated hints, 1702-1703)

The consequence of their coming off shews, that they caused the Jaundice in those two persons I have mentioned, by obstructing the Channels thro which the Bile passeth from the common receptacle into the *duodenum*. (PT 1702-1703, 23: 1582 [1282])

From hence we come to understand how the *Itch* proves to be a Distemper so very catching; [...]. (PT 1702-1703, 23: 1298)

This and other instances, make me easily concur with some Physicians in an Opinion, that in some Families the Lungs have originally a more tender Constitution than in others. (PT 1702-1703, 23: 1384 [1378])

- (b) Generalization in concluding remarks 2 (more structured conclusions, 1727-1728)

Since then it is evident that *Fistula's* in all parts of the Body are dilatable to a great width, since Nature is often able of herself to dilate the very parts in dispute, to a very extraordinary Degree; [...].

Therefore *Artificial Fistula's* in Males and the *Urethra* of Females may be dilated so as to extract any Stone without cutting the Body of the Bladder, or lacerating any of the Parts. (PT 1727-1728, 35: 322)

The 1703 set is characterized by deductive expressions marking a certain degree of abstraction, such as *The consequence of, they caused, From hence, the Itch proves to be, easily concur*, but also by an author-centered approach. Expressions such as *I have mentioned, we come to understand, make me easily concur* still emphasize an involved attitude in an effort to introduce more abstract considerations.

In contrast, the 1727 extracts show a more institutionalized and abstract orientation. Condition-consequence structure is emphasized by thematized connectors which clearly contextualize the nature and the function of the paragraph they open: *Since then, Therefore*. Expressions such as *it is evident that, Artificial Fistula's* [...] *may be dilated* mark a shift from an author-centered to object-centered perspective.

The last example is drawn from a paper published in the 1801 volume. The situation has changed completely: from personal reports at the very beginning of the century, to definitely structured Introduction-Method-Results-Discussion (Conclusions) at the end of the period considered. The two following extracts represent the introductory and the concluding sections of a long and elaborate article (about 20 pages), a research paper, as it is designated within the text itself:

- (6) Paper

[*Introduction*]

The nerves have been hitherto considered as chords that have no powers of contraction within themselves, but only serving as a medium, by means of which the influence of the brain may be communicated

to the muscles, and the impressions made upon different parts of the body conveyed to the brain.

The difficulties which attend every attempt to investigate the real state of the nerves in the living body, and the impossibility of acquiring any information upon this subject after death, may be urged in excuse for this opinion having been so universally received, since it will be found, from the following experiments and observations, to be void of foundation.

The only means by which any knowledge respecting the irritability of nerves can be procured, must be from the operations in surgery performed upon nerves, either in a healthy state, or under the influence of disease; or from experiments made upon animal bodies before they are wholly deprived of life, and instituted for that particular purpose.

My attention was directed to this subject by the following case, which explains many circumstances respecting the actions of the nerves when under the influence of disease, and gave rise to the experiments and observations contained in this Paper. [...]

[*Body*]

[*Conclusions*]

The experiments and observations which have been related, appear to illustrate an action in the nervous chords, capable of producing the symptoms which occurred in the case related in the former part of this paper, and also those met with in many other diseases, the symptoms of which have never been satisfactorily explained. [...]

To enter further into the histories of cases which afford evidence of a morbid action in the nerves, would be trespassing too far upon this learned Society, and would render the present Paper an inquiry into medical facts, which is only intended to be an investigation of the natural actions of the nervous fibres, illustrated by the phaenomena which occur while these chords are under the influence of disease. (PT 1801, 91: 1-2, 19, 22)

In the introduction four different steps may be recognized. The opening paragraph introduces the topic (*nerves*) and the background (*have been hitherto considered as*); the author and the research project are placed within the disciplinary community. The following paragraph, instead, after a long and elaborate apologetic hedging (starting with *The difficulties* and ending with *in excuse for this opinion having been so universally received*), introduces

an oppositional-argumentative approach to put forward a new research perspective, since the previous one *will be found, from the following experiments and observations, to be void of foundation*. The author and the research project are placed inside the disciplinary community with a different perspective. The third paragraph introduces the only possibility to solve the dispute: *The only means by which any knowledge [...] can be procured, must be from the operations in surgery [...]*. A different perspective seems inevitable: there is a research gap and the author is keen to try to fill it with further experimentation. This, at least, seems to be his aim. Finally, the closing paragraph focuses on the present research and the inevitability of the attempt: *My attention was directed to this subject by the following case, which [...] gave rise to the experiments and observations contained in this Paper*. In this case, the author tries to make his research accepted by his audience-readership of experts.

The concluding section of the paper is less elaborate than the introductory section, and it may be sub-divided into two steps. The first can be traced back to the introduction, and it summarizes, with the same words, the intentions expressed in the plan of the article. Now these intentions have definitely been carried out, *The experiments and observations which have been related, appear to illustrate* (in other words ‘demonstrate’) the new perspective and further changes in medical knowledge.

The closing paragraph may actually be considered as a follow-up paragraph, suggesting further research. It delimits the scope of the paper itself to lower the risk of non-acceptability: *To enter further [...] would be trespassing too far upon this learned Society, and would render the present Paper [...]*. While delimiting the investigation on the nature of the nerves, the author encourages further inquiry into *medical facts*: *[...] would render the present Paper an inquiry into medical facts, which is only intended to be an investigation of the natural actions of the nervous fibres*.

## 6. Concluding remarks

The investigation has demonstrated that between 1702 and 1801 medical writing in the *PT* undergoes major changes. In general, these changes appear to be gradual shifts along a continuum highlighting an essentially author-centered approach at the outset of the century and an object-centered perspective at the end of the period. In particular, the movement is from a personal, involved, emotional perspective to an informational one; from a descriptive-narrative style to a descriptive-argumentative one; from observations on what happens from a situation-dependent perspective in single case studies (by single physicians and/or practitioners) to elaborated

discussions. In this context, medical communication is characterized by a more general and abstract orientation emerging from more planned and structured articles whose introduction, body (description and discussion) and conclusion are clearly framed.

Observation and experimentation are at the heart of eighteenth-century medical experience, and step-by-step reports characterize medical writing throughout the period. However, change is not systematic across time since periods of increase and decrease in the use of specific features and rhetorical strategies may alternate. On the whole, these features overlap and/or coexist; what varies is their frequency of use in the texts analysed.

Epistolary writing and narrative style in medical writing are pervasive until about 1775: they were typical communicative strategies in genteel social interaction, largely adopted in scientific writing. According to Atkinson, letters attest to the “pervasive genteel social ethic and its adoption into [...] the rhetoric of science” (Atkinson 1996: 363). Ultimately, they reflect the civil and polite society which represented the readership of medical knowledge. The last quarter of the century marks a turning point, since the British Medical Reform enhanced the professionalization and diversification of medical education and practice; this definitely changed the way medical knowledge was elaborated and conveyed. Moreover, peer review, which was introduced in the *PT* in 1752, as well as the establishment of medical journals in London towards the end of the century, further stimulated the standardizing of medical research articles.

## REFERENCES

### Sources

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.* Volumes 23, 35, 47, 67 and 91.

<http://rstl.royalsocietypublishing.org/content/by/year>.

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*

1702-1703 “III. A Letter from Mr Jonathan Kay, Chyrurgeon in Newport, Shropshire, concerning a strange Cancer, of which his Father dyed”, 23, 1069-1070.

url <http://rstl.royalsocietypublishing.org/content/23/277-288/1069.full.pdf+html>

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*

- 1702-1703 "V. Part of two Letters from Mr Henry Vaughan to the Publisher, containing some uncommon Observations made upon the Dissection of some Morbid Bodies", 23, 1244-1246.

<http://rstl.royalsocietypublishing.org/content/23/277-288/1244.full.pdf+html>

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*

- 1702-1703 "IV. A Letter from Mr James Young, F.R.S. to Dr Edward Tyson, F.R.S. concerning a Ball, extracted from a Person, who had suffered by it 30 years, in which was a Plumb-stone", 23, 1279-1283.

<http://rstl.royalsocietypublishing.org/content/23/277-288/1279.full.pdf+html>

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*

- 1702-1703 "II. An Abstract of Part of a Letter from Dr Bonomo to Sigmor Redi, Containing Some Observations concerning the Worms of Humane Bodies. By Richard Mead, M.D.", 23, 1296-1299.

<http://rstl.royalsocietypublishing.org/content/23/277-288/1296.full.pdf+html>

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*

- 1702-1703 "A Letter to Dr John Wright, F. Col. Phys. Lond. to Mr Wm Cowper, F.R.S. concerning the Cure of an Apostematation of the Lungs", 23, 1372-1385.

<http://rstl.royalsocietypublishing.org/content/23/277-288/1372.full.pdf+html>

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*

- 1727-1728 "IX. Two Chirurgical Questions, stated and answered by John Douglas, Surgeon, F.R.S.", 35, 318-322.

<http://rstl.royalsocietypublishing.org/content/35/399-406/318.full.pdf+html>

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*

- 1727-1728 "IV. An Account of an Aneurysm of the Aorta, (dissected in St. Bartholomew's Hospital) by Pierce Dod, M.D. Fellow of the College of Physicians, and Physician to that Hospital", 35, 436-440.

<http://rstl.royalsocietypublishing.org/content/35/399-406/436.full.pdf+html>

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*

- 1727-1728 "Some Observations on Aneurysms in general, and in particular, on the foregoing. By F. Nicholls, M.B. Pael. of Anatomy, Oxon. & F.R.S.", 35, 440-444.

<http://rstl.royalsocietypublishing.org/content/35/399-406/440.full.pdf+html>

*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*

- 1751-1752 "LXVIII. An Account of an Hydrophoby, by Thomas Wilbraham, LL. D.F.R.S.", 47, 412-414.  
<http://rstl.royalsocietypublishing.org/content/47/412.full.pdf+html>  
*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*
- 1777 "XXXI. A short Account of Dr. Maty's Illness, and of the appearances in the dead Body, which was examined on the 3d of July, 1776, the Day after his Decease. By Dr. Hunter and Mr. Henry Watson, F.R.S.", 67, 608-613.  
<http://rstl.royalsocietypublishing.org/content/67/608.full.pdf+html>  
*Philosophical Transactions giving some Account of the Present Undertakings, Studies, and Labours, of the Ingenious, in many Considerable Parts of the World.*
- 1801 "The Croonian Lecture. On the Irritability of Nerves. By Everard Home, Esq F.R.S.", 91, 1-22.  
<http://rstl.royalsocietypublishing.org/content/91/1.1.full.pdf+html>

### Special studies

- Atkinson, D.  
 1992 "The evolution of medical research writing from 1735 to 1985: the case of the *Edinburgh Medical Journal*", *Applied Linguistics* 13, 337-374.
- Atkinson, D.  
 1996 "The *Philosophical Transactions* of the Royal Society of London, 1675-1975: a sociohistorical discourse analysis", *Language in Society* 25, 333-371.
- Atkinson, D.  
 1999 *Scientific Discourse in Sociohistorical Context. The Philosophical Transactions of the Royal Society of London, 1675-1975*. Mahwah: Lawrence Erlbaum Associates [reprinted by Routledge, 2010].
- Banks, D.  
 2008 *The Development of Scientific Writing. Linguistic Features and Historical Context*. London: Equinox.
- Bazerman, C.  
 1988 *Shaping Written Knowledge. The Genre and Activity of the Experimental Article in Science*. Madison: University of Wisconsin Press.
- Biber, D.  
 1988 *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D. – S. Conrad  
 2009 *Register, Genre, and Style*. Cambridge: Cambridge University Press.
- Biber, D. – E. Finegan  
 1989 "Drift and the evolution of English style: a history of three genres", *Language* 65, 487-517.

Dear, P.

- 1985 "Totius in Verba: rhetoric and authority in the Early Royal Society", *Isis: A Journal of the History of Science* 76, 144-161.

Fontes Da Costa, P.

- 2009 *The Singular and the Making of Knowledge at the Royal Society of London in the Eighteenth Century*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Gotti, M.

- 2011 "The development of specialized discourse in the *Philosophical Transactions*". In: I. Taavitsainen – P. Pahta (eds.) *Medical Writing in Early Modern English*. Cambridge: Cambridge University Press, 204-220.

Gotti, M.

- 2014 "Scientific interaction within Henry Oldenburg's letter network", *Journal of Early Modern Studies* 3, 151-171.

Gray B. et. al.

- 2011 "The expression of stance in early (1665-1712) publications of the *Philosophical Transactions* and other contemporary medical prose: innovations in a pioneering discourse". In: I. Taavitsainen – P. Pahta (eds.) *Medical Writing in Early Modern English*. Cambridge: Cambridge University Press, 221-247.

Gross, A.G. et al.

- 2000 "Argument and 17<sup>th</sup>-century science: a rhetorical analysis with sociological implications", *Social Studies of Science* 30, 371-396.

Taavitsainen, I. – P. Pahta (eds.)

- 2011 *Medical Writing in Early Modern English*. Cambridge: Cambridge University Press.

Address: ELISABETTA LONATI, Dipartimento di Lingue e Letterature Straniere, Università degli Studi di Milano, Piazza S. Alessandro 1, 20123 Milan, Italy.

ORCID code: [orcid.org/0000-0002-1350-6735](https://orcid.org/0000-0002-1350-6735).



# The dialogic nature of nonstandard English

Patrice Larroque

*Toulouse-Jean Jaurès University, Toulouse*

## ABSTRACT

The purpose of this text is to describe nonstandard speech events from a dialogic point of view. Dialogism, as defined by Bakhtin himself, refers to the interplay between a speaker's discourse and other uses of language which are exterior to her/him or related to former ones. Analyzing nonstandard utterances actually reveals the speaker's ambiguity which combines two distinct discursive planes – attitudinal and predicative. Starting from the assumption that some of the nonstandard constructions (i.e. negative concord, existential *it*-clauses, preverbal *done*, and double modal constructions) date back to Early Modern English, I will take the view that nonstandard speech events exist in the context of a continuum, and contribute to a better understanding of the system.

Keywords: linguistics, nonstandard English, English grammar, dialogism, variation.

## 1. Introduction

Many nonstandard constructions date back to the Early Modern English period (from about 1450), an age of transition to present-day English. According to Gachelin (1990: 221), "what is branded as nonstandard corresponds to Early Modern English usage or stands as an original innovation, which can even occasionally be already seen at work inside marginal trends of standard English, like the disaspectualization of the Present Perfect used with past reference." Thus, forms that were accepted in the past as regular usage have been deemed "deviations" or "mistakes" since then. Nonstandard English, however, has the capacity to be recognized and interpreted by a majority of English-speaking people. Standard English is arguably a minority use (Trudgill 1979, McArthur 1998), and many native

speakers do not generally question the grammaticality of nonstandard performance<sup>1</sup>. Then, it does not seem inaccurate to describe nonstandard productions in dialogic terms.

Depending on its dialogic nature both within the system and in discourse, nonstandard English may be regarded as a synchronic projection of the history of English. Language begins within the context of a situation. This paper identifies nonstandard English productions as dialogical inasmuch as they are used in a network of personal interactions, on the one hand, and can, on the other hand, be related to former language states. These linguistic features have become either archaic, or nonstandard altogether. It is the case, for instance, of negative concord or of preverbal *done* in nonstandard varieties of English.

The linguistic description of these features can be regarded as an attempt to discuss the dialogic effects of nonstandard speech events within the system of English. In other words, it aims to show that these constructions exist in a continuum which can be related to dialogism and therefore be considered as the manifestation of several underlying voices in a sentence. So before dealing with the dialogic nature of nonstandard English and analyzing sentences in which nonstandard constructions occur, I will discuss dialogism and the historical background of such nonstandard use which can be traced back to Early Modern English. Nonstandard grammar may be illustrated by examples taken from my own research corpus which includes many dialectal varieties of English spoken in the United States and in the British Isles.

## 2. Dialogism

Dialogism refers to a concept developed by the philosopher Mikhail Bakhtin to account for his work on literary theory (Bakhtin 1952/1986; Todorov 1981). It can be defined as the interplay between a speaker's discourse and other uses of language which are exterior to her/him. For Bakhtin, language is dialogical, in the sense that it is always related to former uses (diachronic or dialectic dialogism) while existing in anticipation of a response. It is, as it were, the manifestation of multiple voices which emerge at the surface

---

<sup>1</sup> Sentences are judged ungrammatical when any ordinary speaker of English would feel that something is wrong with them, that they somehow do not belong to the system and therefore cannot be accepted, e.g. *\*Peter Mary loves*, *\*Mary dined a hamburger*, *\*Peter put a car*, *\*Peter has going to Paris*, etc.

of an utterance. Dialogism, therefore, appears to be the revelation of voices which can be sensed on many different levels. This concurs with T.S. Eliot's idea that "the past should be altered by the present as much as the present is directed by the past" (in "Tradition and the Individual Talent", 1920) and points to the concept that a sentence (or a text) always exists in relation to previous expressions within a given sphere and, arguably, is embedded in socio-historical context (Linell 2009: 19).

The description of language requires that the linguist be familiar with all linguistic history. No expression of any kind has its complete meaning in itself. Its significance can only be appreciated in its relation to former uses of language. As stated in the introduction, language does not exist in a vacuum and, therefore, must be set for contrast and comparison in continual language change. Following the example of standard English, nonstandard English productions relate to language states which were in use as far back as Early Modern English and are branded as nonstandard by prescription which clings to the notion of norm, thus discarding a large amount of productions which may inform us of the reality of language usage. Historical association, however, does not suffice to supply a full picture of the relevant diachronic facts if the analysis is synchronic. The linguistic analysis seeks to provide evidence that shows the dialogic characteristics of nonstandard productions in which exchanges and recurrent discursive interplay occur.

The dialogic nature of a sentence can be seen in every domain of communication, in the sense that it reflects the complexity of another discourse, that of language. Grammatically, dialogism rests on the speaker's ambiguity which combines two levels of discourse, like, for example, negation, interrogation, comparison, concession, modalisation, etc. (cf. Brès – Mellet 2009). In nonstandard English, double constructions, which arise from Early Modern English usage (cf. Larroque 2015) and are now regarded as deviations relative to standard English, are a living illustration of that.

As a matter of fact, many aspects of Modern English can still be found in contemporary nonstandard English, although they have been condemned by prescriptive rules in an attempt to keep the language as proper, logical, and homogeneous as possible. For example, Early Modern English is rife with redundant constructions, such as negative concord, a common grammatical feature of Old English; double modal constructions which are evidence of a category change between the 16<sup>th</sup> and the 19<sup>th</sup> century; the "double perfect" (i.e. *had've* + past participle); double comparatives and superlatives which in Elizabethan English existed alongside morphological and periphrastic non-redundant structures; in the verbal system, preverbal

*done*, which is frequently used in African American English, also had its Middle English version before its importation into nonstandard speech. All these grammatical constructions stand as survivors to prescription and the continual syntactic and morphological evolution of the language. They remain in use in many nonstandard English varieties and bear witness to a former state of the English language.

### 3. Data and methodology

The data described in this paper are excerpts specifically selected for the purpose, and which exhibit some of the aforementioned grammatical phenomena. It may appear to be heterogeneous inasmuch as the samples collected come from different sources. The main criterion of choice is that the language reflects inadequate grammatical competence as regards standard English, but that the performance can be recognized and interpreted by a large majority of native speakers.

The excerpt taken from Baugh (1999) presents some African American English peculiarities that can be found in other nonstandard English varieties, and which can be analyzed from a dialogic point of view. It illustrates how the speaker's socio-historical identity (i.e. the linguistic consequence of slavery) is linked to her language. It provides an authentic representation of nonstandard speech. Other excerpts, taken separately, show traces of multiple voices. They may have a dialogic relation, not only with other similar states of language, but also within their own deep structures. Double constructions, for instance, can be analyzed according to the specific discourse role of each constituent of the pattern. The now famous example of double negatives, which, prescriptivists say, defy logic, shows that two negatives in the same sentence do not cancel each other out and yield a positive, because the basic and simple principle that there is a direct relationship between the surface syntactic structure and the logical form is false. A dialogic approach of the phenomenon may expand our understanding of negative concord in nonstandard English varieties, and thereby expand upon other double formulations or apparent redundancies.

The following analyses will be limited to four specific nonstandard phenomena, namely negative concord, existential *it*-clauses, preverbal *done*, and double modal constructions. These, I argue, constitute a fair representation of what can be described as dialogical. As mentioned, it is important to examine adequately the relevant historical insights which may shed light on the synchronic variation that will be described

thereafter. Another possible limitation stems from the corpus itself and its heterogeneity which, if extended, may not allow to fully evaluate the scope of the analyses. At this stage of the present study, I will only concentrate on the aforementioned cases that spread across dialectal boundaries.

#### 4. Analyzing sentences

Some of the most interesting examples which can be analyzed from a dialogic point of view are given in Baugh's (1999: 5) interview of an African American woman which runs as follows:

You can never forget that slavery was a bitch from the get-go. **Slaves didn't get no schoolin' and they ain't never really given us (African Americans) equal opportunities (1)**, so how we supposed to talk like white folks, and why would we want to? **It ain't no white people really care about us (2)**, 'cause if they did they wouldn't try to make you turn into a white person,, they'd take like you is. But they don't do that. All my teachers in school kept tellin' me, "if you don't speak proper, you won't get a job." That's bullshit! I know some Brothers that went to college – y'know, they did the "white things", with good grades and good English, and they still have problems on the job. **They done tol' me about this Brother who did all the work for a white boy at his job (3)**, and then they (the Whites) lied on his ass when the boss found out and he was fired, and nobody tried to help him. How can you trust motherfuckers that do shit like that, and they say we stupid 'cause we don't talk proper – I see what's goin' on, and I see what's comin' down, and **it ain't got nothin' to do with how we talk (4)**. It's all about money, power, and politics – plain and simple!

##### 4.1. Negative concord

The example of negative concord is all the more interesting since it brings together several voices. It can be accounted for by describing very precisely how attitudinal and predicative operations, marked respectively with *don't* and *ain't*, and *no schoolin'* and *never really given us equal opportunities* in example (1), combine linguistically between the speaker and the lexis, and how double negatives relate to the different former dialogic configurations in which they occur. For example, *don't* and *ain't* establish, by means of

the auxiliaries *do-* and *ai-* (= *have*), a direct relationship between the sentence and the speaker who negates a positive (*do-n't* and *ai-n't*). In *don't* and *ain't*, *do-* and *ai-* are locative markers which relate the sentence to the speaker (Larroque 1999: 125-126) and constitute a dialogic response to a preconstruct to be inferred from the context (cf. *You can never forget that slavery was a bitch...*). The sentence can indeed be reworded as:

- (1) No, slaves got no schoolin', and no, they have never really given us equal opportunities.

In this paraphrase, *no* (the adverb) indicates direct anchoring to the speaker. *Don't* and *ain't*, in (1), can therefore be construed as locatives referring to the speaker and the speech situation. As the adverb *no*, they function as boundary morphemes and as such they are invariable (cf. example 4: *it ain't got nothin'...*); *no schoolin'* and *never really given us equal opportunities* apply to the predicative (informational) level, i.e. the lexis. This may explain why a double negation construction does not yield a positive. A similar analysis can be conducted for sentence (4). Besides, the duplication of marks in a sentence has an intensifying meaning.

Negation in Old English was *ne* and was reinforced by the particle *wilt* (> *naght* > *not*). This type of negative reinforcer is a popular strategy in language and tends to become simply a negative particle, thus losing its original nominal meaning. Double negation was commonly used until the 15<sup>th</sup> century and was ruled out of standard English in the 16<sup>th</sup> century. It has survived in many nonstandard English varieties as a manifestation of the past and reminds us that language exists in constant and relational evolution. In that sense it is dialogical. There is, moreover, an incessant interplay between so-called nonstandard productions, that is the idiosyncratic way in which the verbal system has been internalized, and the grammar of the common language. This may entail a dialogic relationship between standard and nonstandard language varieties.

#### 4.2. Existential clauses

In nonstandard English, especially African American English, existential clauses are often introduced by *it* instead of *there*, as in sentence (2), reproduced below:

- (2) It ain't no white people really care about us. (in Baugh's interview)

The construction was present in standard English until the 17<sup>th</sup> century, for example in Shakespeare's "For 'tis no trusting to yond foolish lout" (*The Two Gentlemen of Verona*, IV, iv, 66). Although it is regarded as nonstandard, it has lived on until today. In the latter, 'tis can be analyzed as *there is*.

Existential *there* in English is a function word which has developed from the place adverb *there*, which is bleached of its nominal meaning in existential clauses. But, although it functions as syntactic (grammatical) subject of the clause, it retains enough of its original place meaning to provide an abstract situational reference to introduce new information (Joly – O'Kelly 1989: 145), and thereby exemplify its inherently dialogic quality: existential *there* may occur in the same clause as place adverbs *here* and *there*:

- (3) (a) There ain't nothing here, lieutenant. (taken from K. Costner's film *Dances with Wolves*, 1990)
- (b) There ain't nothing there no more. Just a lot of bones. (P. Auster, *Moon Palace*, 1989)
- (c) There's wee-er laddies than me that goes round and starts tossing stones at the laddies round there. (Lowland Scots, Hughes – Trudgill 1996: 118)

In light of existential *there* which has lost much of its original place meaning to become a function word, one might be tempted to analyze existential *it* in the same way. Most linguists would describe the pronoun as a "mere prop word" (Sweet 1891/1969: 75) devoid of meaning, just a syntactic subject, the actual notional subject being presented in the indefinite noun phrase following the copular verb. Bolinger (1977: 84), however, defines the nature of *it* as "a definite nominal with almost the greatest generality of meaning limited only in the sense that it is neuter." For example, in sentence (4), *it* can be paraphrased as *the whole situation I am referring to* (cf. the context: '*I see what's going on, and I see what's going down*')

- (4) It ain't got nothin' to do with how we talk. (in Baugh's interview)

Therefore, *it* refers to a notional content, and subsumes all the elements occurring in the paraphrase.

In sentence (2), existential *it* hinges on a problem of focalization, either on the existence, or on the identity of the new information. In (2), *it* can be reworded as *the object I am mentioning the existence of*. It can, therefore, be construed as the anaphoric substitute for the situation referred to

by the speaker *No white people really care for us* corresponds to *it*<sup>2</sup>. Thus, *it* in existential clauses appears to be dialogical, in that it establishes a link between the linguistic and the extralinguistic situation: *it* emphasizes the notional subject and at the same time identifies it as relative to the situation. As *it* is more specific in socially marked nonstandard varieties, it appears to be more adequate to translate situational immediacy.

When stating the existence of something, which can be construed as the selection of an item from the speech situation, the denoted object may be determined in relation to the speaker, or retrieved and identified, singled out amid other subjects. The dialogic nature of existential *it* rests on an assumption of similarity: it shares the same underlying representations as existential *there* while carrying representations of former popular speech, which informs us of the history of existential clauses<sup>3</sup>. In the late Middle English period, dematerialized *there* came to be used and was to become dominant in Modern English, and still is in present-day English.

### 4.3. *Preverbal done*

In order to describe the grammaticalization of preverbal *done*, in sentence (3), and its diachronic evolution, it is necessary to analyze the attitudinal (mental) operations underlying its meaning (which is still carried with the past participle form of the verb *do*), that is the perfective aspect such as it is marked by the periphrastic *have* + *V-en* construction, and the direct relationship it establishes between the speaker and the sentence. On the other hand, the recurring discursive interplay which, within the framework of dialogism, not only interlocutory, but also with reference to a former state of language, helped the morpheme to shift from an initially lexical value to a locative grammatical one (cf. *have* in standard English) and turn the past participle into an auxiliary in nonstandard utterances.

Mustanoja (1960: 605-606) argues that Middle English had a *done* which may have carried the same perfective meaning as the sentences in which *Have* + *V-en* occurs to mark the perfective aspect, but it was usually preceded by the auxiliary *have*, as in Chaucer's *Canterbury Tales* (end of the 14<sup>th</sup> century):

<sup>2</sup> Linguistically, this may be called 'anticipatory anaphora' (Huddleston – Pullum 2002: 1312).

<sup>3</sup> Black slaves may have picked up the expression from the English slavers who, in the 17<sup>th</sup> century, transported them from West Africa to the Americas.



- (5) (a) An Oratorie, riche for to see,  
In worship of Dyane of Chastitie,  
Hath Theseus doon wrought in noble wyse.  
(‘Knyghthes Tales’, st. 7)
- (b) Thise merchant han doon fraught his shippers nerve,  
And whan they han this blissful mayden say,  
Hoon to Surrye been they went ful fayn,  
And doon his nedes as they han doon you  
And liven in wede, I kan sey yow namoore.  
(‘The Tale of the Man of Law’, ll. 171-175)

These sentences show that the preverbal *done* construction is not a grammatical innovation. Ellegård (1953) and Visser (1966, 2002) point out that the usage was common in the 15<sup>th</sup> and 16<sup>th</sup> centuries, as in (c):

- (c) When the Clerkes have dooen syngyng. (Book of Prayers; 1549)

The difference is that the frequently-used auxiliary *have*, in unstressed position, has been deleted as a surface marker, not the underlying operation. *Done*, then, becomes the locator and perfective marker.

If, historically, the rise of the perfective aspect to mark the relation to a reference point can be attributed to the aoristic derivation of markers, the simple past being too objective, it follows that the preverbal *done* structure can be regarded as a synchronic response to that development by creating a new, more expressive construction to locate the event by relating it to the speech situation and the speaker. The development can be analyzed as follows:

- the past participle (*tol’* in Baugh’s interview, sentence 3) represents the temporal root (the past);
- *done* can be rephrased as *have done*, which is the intermediary stage of the evolution;
- *done* stands alone, *have* is deleted (*done* presupposes *have*) and the locative element is the time of speaking, that is the interview itself.

Thus, preverbal *done* is the result of an interplay of semantic, syntactic and grammatical (and probably phonological) change, on the one hand, and a dialogic response to a need for expressiveness, on the other. Furthermore, one may sense an added metalinguistic value, because *done tol’* (Baugh’s interview, 3) is certainly more emphatic and markedly resembles a polemic.

At the beginning, *done* was both the simple past and the past participle form of the verb *do*. Through continual use in perfective constructions in Middle and Early Modern English, the locating function has shifted from the auxiliary to the verb, as in sentence 3 in Baugh's interview in which *done* has acquired a perfective meaning. In present-day English, it appears in nonstandard utterances, especially in the southern states of the United States of America. Historical data place the latest occurrences of the marker in standard English around the 16<sup>th</sup> century and in the north of England and Scotland. In some nonstandard varieties, *done* is regarded as an archaic reintroduction (US southern states).

#### 4.4. Double modal constructions

I cannot end this examination without considering the case of double modal constructions which in nonstandard English exhibit a dialogic attitudinal ambiguity in the use of modal verbs. Examples which currently appear in some varieties of English are given below.

- (6) You may can find more information throughout the website. Here are some areas that may help. (Omega 3, "Learning for Health and Medicine", 2008)
- (7) If you're already in College, you might could well because you're looking at making sure that they, they fulfill what is a dream for you and for them. (Abbey life introduction course, recorded on January 25, 1994, spoken part of the BNC)
- (8) because, as you probably know Clyde was looking into a program which will could alleviate a lot. Yes, I know. (Conversation recorded on April 1992, spoken part of the BNC)
- (9) If a woman chooses to stay home [...] we feel that she should ought to have the opportunity. (Baptists debate, New Seminary Degree in Homemaking, Audrey Barrick, *The Christian Post*, 15 August, 2007)

Double modal constructions can be described as the result of epistemic modality in conjunction with root modality (Brown 1991: 76-77; Denison 1998; Larroque 2005: 212-213; Brandstetter 2006; Larroque 2010, 2015: 69-79). In such combinations, the two modals behave syntactically and semantically in a restricted order. In sentences (6-9), for example, the epistemic meaning (*it is possible/probable/predictable/necessary/required that...*) is carried with the first modal. In (6), *may* as first modal indicates a contingent state of affairs while *might* in sentence (7) is used in its conditional sense. *Will* can

be epistemic in the sense of prediction as in (8)<sup>4</sup>, and *should* in example (9) is used epistemically, since the proposition refers to the sphere of non-self<sup>5</sup>. The second modal carries a root meaning and applies to the subject-predicate nexus. In sentences (6-8), *can/could* refers to ability while in (9), *ought* can be interpreted as a deontic marker.

Interestingly, it is to be noted that the epistemic modal comes first in the sequence as it denotes the speaker's assessment of the propositional content. Thus, the modal applies to the whole proposition and thereby indicates direct anchoring to the speaker. That is another reason why the first modal is logically and iconically restricted to an epistemic sense and has an attitudinal function. It represents a dialogic response to the presupposed subject-predicate structure involving the root (second) modal which is close to the function and the meaning of a lexical verb.

In Old English, modal verbs shared the same syntactic and morphological properties as other verbs. In other words, Old English had a single category of verbs which included the modals (Lightfoot 1979; Denison 1993; McMahon 1994). They had an infinitive form, inflections, and a "normal" complementation including a direct object (Lightfoot 1979: 100; Denison 1993: 327; Bauer 1994: 23-24)<sup>6</sup> as in Shakespeare's lines in (10):

- (10) Let the priest in surplice white  
That defunctive music can. (*The Phoenix and the Turtle*, 14, 1601)

In Modern English, modals behaved in a way that set them apart from other verbs. English then innovated a new category of verbs which gradually lost some of their nominal meaning and developed syntactic properties as auxiliaries, especially when some of these verbs ceased to be used with a direct object (Crépin 1994: 145). It is, indeed, a case of grammaticalization accompanied with a category change which began in the 15<sup>th</sup> century, in Early Modern English.

I have mentioned that when it comes to modality, the general tendency is to distinguish two interpretations, epistemic and root modality. Epistemic modality applies to the entire subject-predicate structure. It concerns the relationship between the propositional content of the sentence and the speaker's assessment of the actualization of the subject-predicate nexus

<sup>4</sup> *Will* should not be restricted to a simple marker of futurity.

<sup>5</sup> *Should* as a single modal can be either an epistemic modal (*The weather should clear up tomorrow*), or a root modal (*You should behave yourself*), depending on whether the relationship is interpersonal or not.

<sup>6</sup> A "normal" complementation includes transitivity and intransitivity, subordination, prepositional phrases.

using what evidence he or she has of the situational context. Epistemic meanings arguably developed from root, namely deontic, modality when it was frequently used with forms marking the speaker's subjectivity (Traugott 1989). As for root modality, it applies to the subject-predicate relation and, as it were, qualifies the subject of the sentence.

Thus, due to the specific discourse role of each constituent of the construction, it is not surprising that they may foster a dialogic relation by combining speaker involvement and informational content, that is the attitudinal and the predicative levels of the sentence. When a single modal is used, it carries both the information and the speaker's judgment about the subject-predicate relation. For example, (6) above can be rephrased as:

- a) You can find some information throughout the website.
- b) You may find some information throughout the website.
- c) You may be able to find some information throughout the website.

*Can* in (a) expresses a possibility inherent in the subject (*you*): access to the information is made possible by consulting the website (= *it is possible for you to find some information throughout the website*); in (b-c), *may* encodes the speaker's assessment of the propositional content. Epistemic modality is only represented in (b-c). In (c), the suppletive *be able to*, which is not exactly a modal (it is sometimes called a semi-modal), makes the root reading of *can* more explicit.

The historical insights provide evidence of this state of affairs, and at the same time shed light on the synchronic variation that has just been described. Again, double modal constructions should not be cut off from the former dialogic configurations in which they appeared. Indeed, modal verbs have also undergone semantic changes which may be the result of metaphor or, as Traugott (1989: 50) puts it, of "conversational implicature". This involves two discursive levels and may in turn be analyzed within the framework of dialogism<sup>7</sup>.

## 5. Conclusion

I have attempted to show that nonstandard English is dialogical in the sense that it appears within the natural movement or evolution of the language.

<sup>7</sup> *Shall* (OE *sculan*), for example, is historically akin to the notion of financial debt (*icsceal* = *I owe*), hence the idea of constraint, moral debt, or of having to do something unwillingly.

Many forms that are considered deviations or mistakes in present-day English are occasions to explore grammatical phenomena that can be related to structures which were in common use in Early Modern English. Despite prescriptive rules which impose a norm (which is arbitrary) and strive “to teach us to express ourselves with propriety in that language” (Lowth 1762: xi), nonstandard usage is not inherently worse than any other variety. As such, it is worthy of analysis.

The notion of dialogism is important in grammatical studies, such as those about nonstandard grammar, because it highlights the multi-relational aspect of the language: 1) nonstandard usage appears to be directly associated with a former state of language, while prescription may disrupt the continuum, and 2) there are variable degrees of dialogic instances, where uses bring out the complexity of speech production, that is, the ambiguity which combine two discourse units: the predicative and the attitudinal levels. Double grammatical structures such as negative concord or double modal constructions provide evidence of that, with reference to Early Modern English. Furthermore, recurrent discursive interplay processes may result in the grammaticalization of lexical items (cf. *there, done*).

The present study has aimed to contribute to a better understanding of nonstandard English varieties, and thereby explicates the dialogic nature of grammatical representations, both standard and nonstandard. There is no objective boundary or straight line between the two, if not a dialogic relationship which may have social and educational consequences.

## REFERENCES

### Sources

- Auster, P.  
1989 *Moon Palace*. London: Penguin Books.
- Barrick, A.  
2007 Baptists debate, New Seminary Degree in Homemaking. *The Christian Post*, 15 August 2007.
- Baugh, J.  
1999 *Out of the Mouth of Slaves*. Austin: University of Texas Press.
- Gibson, E.C.S.  
1549 *The First and Second Prayer Books of Edward VI*. Repr.2008, Wildside Press.  
[www.wildsidepress.com](http://www.wildsidepress.com)

*British National Corpus (BNC)*

Compiled by Mark Davies (Brigham Young University).  
<http://corpus.byu.edu>

Chaucer, G.

1370-1400 *The Canterbury Tales*. Repr.1999, London: The Folio Society.

Costner, K.

1990 *Dances with Wolves*. Adaptation film based on the book *Dances with Wolves* by Michael Blake, 1988. MGM.

Hughes, A. – P. Trudgill.

1996 *English Accents and Dialects*. 3<sup>rd</sup> edition, London: Arnold.

Omega-3.

2008 "Learning for Health and Medicine"  
<http://www.omega3learningforhealthandmedecine.edu>

Shakespeare, W.

1623 *The Two Gentlemen of Verona*. Repr. 1972 by C. Leech (ed.), *The Arden Shakespeare*. London: Methuen.

**Special studies**

Bakhtin, M.

1952/1986 *Speech Genres and Other Late Essays*. Translated by Vern W. McGee. Austin: University of Texas Press.

Bauer, L.

1994 *Watching Language Change*. London: Longman.

Bolinger, D.

1977 *Meaning and Form*. London: Longman.

Brandstetter, C.

2006 "A Study in Syntactic Variation: Double Modal Constructions", communication presented at NWAV 35 (New Ways of Analyzing Variation). Columbus, OH: Ohio State University.

Brès, J. – S. Mellet

2009 "Une approche dialogique des faits grammaticaux", *Langue française* 3 (163), 3-20.

Brown, K.

1991 "Double Modals in Hawick Scots". In: P. Trudgill – J.K. Chambers (eds.) *Dialects of English. Studies in Grammatical Variation*. London: Longman, 74-103.

Crépin, A.

1994 *Deux mille ans de langue anglaise*. Paris: Nathan.

Denison, D.

1993 *English Historical Syntax: Verbal construction*. London: Longman.

1998 "Combining English Auxiliaries". In: O. Fisher – A. Rosenback – D. Stein (eds.) *Pathways of Change: Grammaticalization in English*. Amsterdam: John Benjamins, 111-147.

- Eliot, T.S.  
1920 "Tradition and the Individual Talent". In: *The Sacred Wood. Essays on Poetry and Criticism*. London: Methuen. (Electronic book, New York, Barkly.com, 2000).
- Ellegård, A.  
1953 *The Auxiliary do: The Establishment and Regulation of its Use in English*. Stockholm: Almqvist & Wiksell.
- Gachelin, J.-M.  
1990 "Aspects in non standard English". In: J.-L. Duchet (ed.) *L'auxiliaire en question*. Travaux linguistiques du CERLICO 2, 221-251.
- Huddleston, R. – G. Pullum.  
2002 *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Jolly, A. – D. O'Kelly.  
1989 *L'analyse linguistique des textes anglais*. Paris: Nathan.
- Larroque, P.  
1999 *Faits de langue en anglais non standard, Etude des marques énonciatives*. PhD thesis, Université Nancy 2.  
2005 "Variation et représentation linguistique", *Anglophonia* 18, Toulouse: Presses Universitaires du Mirail, 205-216.  
2010 "Les doubles modaux anglais, témoins d'un changement de catégorie", *Anglophonia* 28, Toulouse: Presses Universitaires du Mirail, 117-134.  
2015 *Introduction à l'anglais non standard*. Rennes: Presses Universitaires de Rennes.
- Lightfoot, D.  
1979 *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- Linell, P.  
2009 *Rethinking Language, Mind, and World dialogically*. Charlotte, NC: Information Age Publishing Inc.
- Lowth, R.  
1762/1967 *A Short Introduction to English Grammar*. Facsimile. Menston England: The Scholar Press.
- McArthur, T.  
1998 *The English Languages*. Cambridge: Cambridge University Press.
- McMahon, A.  
1994 *Understanding Language Change*. Cambridge: Cambridge University Press.
- Mustanoja, T.  
1960 *A Middle English Syntax. I. Parts of Speech*. Helsinki: Société Philologique.
- Sweet, H.  
1891/1968 *A New English Grammar, Logical and Historical*. Oxford: Clarendon.

- Todorov, T.  
1981 *Mikhaïl Bakhtine – Le Principe dialogique suivi de Ecrits du Cercle de Bakhtine*. Paris: Seuil.
- Traugott, E.  
1989 "On the Rise of Epistemic Meanings in English: An Example of Subjectification". In: "Semantic Change", *Language* 65, 31-55.
- Trudgill, P.  
1979 "Standard and Non-Standard Dialects of English in the United Kingdom", *International Journal of the Sociology of Language* 21, 9-24.
- Visser, F. Th.  
1966/2002 *An Historical Syntax of the English Language*. Leiden: Brill.

Address: PATRICE LARROQUE, Université Toulouse – Jean Jaurès, 5, Allées Antonio Machado, 31058 Toulouse, cedex 9, France.

ORCID code: [orcid.org/0000-0002-4804-7233](https://orcid.org/0000-0002-4804-7233).



## **English in multilingual education programs: some cases and perspectives**

Annalisa Bonomo

*“Kore” University of Enna*

### **ABSTRACT**

With reference to the multilingual turn and the spread of English as a lingua franca, this article discusses the pressures for multilingual education programs and the educative nation plans based on language learning as a tool for democracy. The idea of a sustainable development of new language learning strategies will be studied in light of its complexity and its pertinence to language studies. The investigation will begin with some introductory remarks about how much complexity the idea of “multilingual education” involves, reflecting on how language matters can become language issues because they concern values, usage domains, geographical variations, and many historical processes which may be irregular. The study will then focus on the “Philippines case”, still considered as a complex multilingual scenario despite the widespread presence of English in the country. Then, the challenge of English in multilingual education will be discussed in the context of CLIL programs and their effectiveness in new language policies even beyond the European borders.

Keywords: Multilingualism, English, complexity, the Philippines, CLIL.

### **1. Introductory overview**

Language learning plays an ambitious and pre-eminent role in the education of a democratic citizenship. For this reason, the global dimension of the world fosters multilingual communication which calls for a complex view of foreign language learning.

Nevertheless, the existence of a coherent system of language qualifications which aims at facilitating the attainment of “levels of proficiency required by existing standards, tests, and examinations” as the Council of

Europe stated in 2001 (*Council of Europe* 2001: 21), does not imply necessary references to the sociocultural dimension which should be an essential part of language evaluation in intercultural frameworks.

After all, it is impossible to think language education far from the awareness of the structural changes which face global society. Each act of communication translates its cultural origins and backgrounds; at the same time, communicating involves a linguistic dimension (the “textual one”), and a socio-linguistic counterpart (the “contextual one”), which makes the difference most of the time. In Kramsch’s words, this double face reflects “the fundamental polarity of linguistic discourse that describes language as both the reaction of texts and the shaping of contexts” (Kramsch 1993: 10).

Despite the paradoxical pressures for a lingua franca to be spoken across countries with the purpose of mutual intelligibility, people currently still speak a great variety of languages, which makes multilingualism a new educational goal; in a similar vein, multilingual education experiences involve a revised language knowledge which mixes what Chomsky has described as *linguistic competence*, or what people know of the language, with *performance*, what they say or write at any given moment (Jordan 2004: 6).

The idea of sustainable development of new language learning strategies which may give the individual in global society a complex status that can be studied in light of different perspectives (linguistic above all, but also sociolinguistic, philosophical, economical, and anthropological to name the most relevant today) translates linguistically *la pensée complexe* by the French philosopher Edgar Morin: “seule une pensée complexe, c’est-à-dire multidimensionnelle et ouverte, peut respecter la complexité du réel et communiquer avec elle” (Fortin 2005: 110)<sup>1</sup>.

The idea of linguistic uniformity and the spread of multilingualism, together with the unconditioned recognition of the privileged status of English, do not solve the problem of communication in all multilingual domains. Thus, the matter calls for hybridity rather than new imperialisms among languages.

As a matter of fact, English is considered a transcultural resource which facilitates mobility and fosters a wide interlingual communication. However,

<sup>1</sup> “Only a complex thought, which is a multidimensional and open one, can respect the complexity of the real and communicate with it” (my translation). On the importance of fostering new ‘strategies’ more than ‘programs’, Morin writes: “la complexité appelle la stratégie. Il n’y a que la stratégie pour s’avancer dans l’incertain et l’aléatoire” (1990: 178); “Complexity asks for strategy. Only strategy can advance in the uncertain and unplanned” (my translation).

some of the learning policies which involve the use of minority or national languages in education see English as a risk to what Cooper has described as the “corpus, status and acquisition” of language planning (Cooper 1989). As Cenoz comments:

The use of a minority language in education has an influence on status and corpus planning because the minority language is used for new and more prestigious functions. An implication of the use of the minority language for new functions is the adaptation of the corpus of the language which could include graphization, standardization and the creation of scientific and technological terminology. Moreover, it has a direct influence on acquisition planning because the number of speakers can be expanded when the language is learned as a second language, and also those who speak the minority language as their first language can acquire literacy skills (Cenoz 2009: 8-9).

Clearly, speaking of multilingual education in different countries involves challenges, above all those connected to the processes of the codification of certain patterns of minority languages and the standardization of language awareness through spelling rules, grammars and usage(s). Some restrictive separations between majority languages and minority ones may lead to relevant detrimental effects, both for teachers and students. In García's words,

in the case of minority languages that are being revitalized or that are stigmatized, there is great linguistic insecurity among the teachers, who are often reacquiring the language themselves. This linguistic insecurity may sometimes lead to the use of language that may be ‘standard’, but that is impoverishing in meaning, in metaphors, in poetry, and that is restricted in form (García 2009: 151-152).

For this reason, with respect to the research methodology to be applied to multilingual issues, many variables must be taken into account, such as syllabus contents, language proficiency of teachers, program limits, cultural backgrounds of students and social contexts in which the multilingual educative process begins. In other words, descriptivism *and* prescriptivism should be aware of the social implications of every speech act, becoming new tools of complexity according to which language matters involve values, geographic variations, and many historical processes typically perceived as irregular.

The increasing need for an intercultural communicative competence, like the one assumed by Byram in the 1990s, has moved forward in the elaboration of a new framework for learning and teaching foreign languages in multilingual contexts. Byram stated:

There can be no generalizable syllabus, neither linguistic nor cultural. A French learner of English needs a different syllabus and methods to a Greek, and different again from a Japanese, and within each of these national groups there are different needs arising from age, purpose, institution and so on. Similarly the assessment of their success as learners needs to take into account of specific learners' origins as well as the language and cultures they are learning (Byram 1997: 4).

Though such concerns may seem obvious by now, they involve contingencies related to migration flows and social stabilities in host countries, particularly as people search for mediated education there. So, schools and universities become microcosms of global society, fostering or hindering expectations and cultural conflicts of the speakers of first or second generation. For instance, in various sub-Saharan African and South-East Asian countries linguistic complexities exist beside tensions between bilingualism and monolingualism. Linguistically heterogeneous, these countries may be envisaged as multilayered repertoires whose language situation results from a highly complex setting which "is due not only to the language left behind by colonists but also the diverse tribal and linguistic settings" (Smakman – Heinrich 2015: 37) that are notably known as code-switching experiences and parallel usage of different languages especially in informal and everyday communication contexts. Yet, the linguistic interaction of the inherited colonial languages and the indigenous ones, if seen in the light of multilingual education policies, can be perceived as twofold. On the one hand, "it requires more than a knowledge of mainstream sociolinguistic approaches and the respective languages spoken in these polities" (Smakman – Heinrich 2015: 37). Take for instance the interesting linguistic identity of Indonesia, with many islands and a "shared linguistic identity, embodied in one standard language. This is the official *lingua franca* for peoples living as far as 3,000 miles apart, whose native dialects are highly different from each other and not always mutually intelligible" (Smakman – Heinrich 2015: 38). On the other hand, according to Kamwangamalu, "it has been contended that the promotion of any indigenous language for official use often elicits opposition from the elites of those languages not chosen" (Kamwangamalu 2016: 129).

Interestingly, the well-known distinction between *subtractive* and *additive bilingualism* (Lambert 1975) is reflected in the contradictory findings about the effects provoked by multilingual education programs in the 21<sup>st</sup> century. In her inspiring survey of *translanguaging* as performed by the speakers accessing different languages, García notes that “in subtractive bilingualism, the first language (L1) is taken away as the second language (L2) is added, resulting in monolingualism in a second language ( $L1 + L2 - L1 \rightarrow L2$ ). In contrast, in additive bilingualism, a second language is added without any loss of the first language ( $L1 + L2 \rightarrow L1 + L2$ )” (García 2009: 142)<sup>2</sup>.

With Lambert’s distinction in mind, we may consider the situation of Canada, whose education programs foster and monitor multilingual linguistic competencies and performances. In contrast, most language policies promoted in the U.S. (as with the Hispanic populations whose first language is gradually undermined), Australia, and Russia result in subtractive bilingualism, and monolingual outcomes.

Thus, it is worth investigating whether the multilingual turn and its effects on language learning may provoke disorientation or frustration caused by the management of language programs. In Baker’s words, the question is: are bilinguals “insiders” or “outsiders” (Baker 2000: 20) in educational contexts?

## 2. The Philippines “case”

The Philippines may be quoted as an interesting “complex multilingual case”. Philippine education administration urged for the increasing use of English, especially during the 1970s, when Marcos and his dictatorship were fond of English and supportive of the U.S. policy in Vietnam. However, the national language policy did not harmonize with concerns for the sustainability of a national identity strongly based on Philipino – spoken by 99 per cent of households as Gonzalez reported in 2007 – which is largely Tagalog-based and a widely accepted symbol of unity for the nation. According to Gonzalez,

while Philipino is the national language, the official language, the language of linguistic symbol of unity and identity, little investment has been placed in developing it as a language of scholarly work at

---

<sup>2</sup> For further reference to Lambert, see Lambert (1967 and 1975). It is also worth mentioning the definition of bilingualism provided by Fishman (1976), who distinguishes between “folk” and “elite” bilingualism, depending on the social status of particular speakers.

the universities. For basic education, a bilingual scheme of English and Philipino has been adopted, with English now more prominent than Philipino, since Philipino is used for only the Philipino Language Class and for some subjects in the curriculum, the rest being taught in English (Gonzalez 2007: 12).

Certain figures reveal developments in the Philippines well. English competence is still considered an asset of Philippine education programs and: “the last national estimate for English speakers was 64.5% of the population of 48,098,960 in 1980 (NCSO 1984)” (Gonzalez 2007: 8). The number of English speakers and their percentage in the Philippines has increased to 89,800,000, 92.58% of the population, according to more recent surveys<sup>3</sup>.

The status of multilingual education in the Philippines is complex, as it is in India: “with 33 languages used in education in India, including English, and 41 languages available for study at school (NCERT, 1999), but with an education in India, as Mohanty (2006: 279) says, that is not really bilingual” (García 2009: 150).

Looking back, we see that one policy, the *Bilingual Education Policy* (known as BEP), which was first introduced in the 1970s and then reinforced during the 1980s, strove for equal use of the national language (Philipino) and English as educational tools. Nevertheless, the Philippine system did not directly address new multilingual needs in the late 20<sup>th</sup> century; the widespread use of English was mostly the result of concerns for economic power that English competences could assure to the emigrant Filipinos. According to some nationalist groups, policies that granted English official status were hindering the national language.

Still, the motivation to learn English for utilitarian purposes such as economic gain and career advancement has remained strong. In Pefianco Martin’s words:

The preference for using the national language carries on to present times. A recent study of Go & Gustilo (2013) on the *lingua franca* of Filipino urban factory workers reveals that the workers favored Tagalog as the language of communication [...]. Tagalog, not Taglish (the term used for code-switching in Tagalog and English) was preferred

<sup>3</sup> See the map of the *Top Ten English Speaking Countries* available at [www.mapsofworld.com/world-top-ten/countries-with-most-english-language-speaker-map.html](http://www.mapsofworld.com/world-top-ten/countries-with-most-english-language-speaker-map.html) (accessed April 2016).

because the language made them feel accepted in their social groups. This finding is significant in the light of a preponderance of studies on code-switching in the Philippines, which reveal that the practice is widespread in various domains of Philippine society, including education (Pefianco Martin 2014: 77).

The prevalence of Tagalog-English code-switching in the Philippines has given rise to Taglish, a language mix that Thompson (2003: 41) describes as “Filipino street English”. It has spread rapidly through radio and popular TV programs though it has “no body of literature except in tabloids” (Thompson 2003: 41). As a hybrid vernacular, Taglish is a living language by now while English occupies a paradoxical position in the country, where it is used regularly in the media but not so in the street. Taglish seems to constitute a pragmatic compromise in the dispute between English and Tagalog.

Of the eight major broadsheet newspapers in circulation, all are published in English (Dayag 2004). Inclusions of Tagalog/Filipino wording in them are invariably marked by italics or quotation marks (Thompson 2003). However, of the sixteen major tabloids only two are in English. Twelve of them are in Tagalog/Filipino or Cebuano (if based in Cebu), and two use both Taglish and a macro-switching variety in which long stretches of English alternate with long stretches of Filipino (Smedley 2006: 37-38)<sup>4</sup>.

Mixtures of English and Tagalog vary but, as Samson (2013) points out, certain forms of code-switching are much more common in Taglish than others are. The three most frequently occurring forms are these:

*Literal, word-for-word translation into English of structures and phrases of Tagalog, resulting in a peculiar idiom.* Notable examples include expressions like *kill the light* meaning ‘turn out the light’, an idiom deriving from the action of extinguishing candles, and *I don’t know to my mother* for ‘I’m leaving it up to my mother’. These literal translations of Tagalog idioms are typically non-transparent to those who do not speak Taglish or Tagalog.

*Alternating use of phrases of English and Tagalog in single sentences.* An example of this form is *I made him sampal... tapos I left him.*, which roughly translates as ‘I discovered his infidelity so I left him’. English

<sup>4</sup> Smedley’s research is entirely available at <http://aut.researchgateway.ac.nz/handle/10292/186> (accessed March 2017).

speakers who do not know Tagalog can usually follow conversations made up of such sentences.

*Use of Tagalog words in English sentences, often with altered meanings.* Misinterpretation even among Taglish speakers of different ages, social backgrounds, etc. may result here. According to Samson (2013), such words may even lose their original meanings due to this usage.

Evidently, since the late 20<sup>th</sup> century restoration of indigenous languages as educational tools used in the earliest school years, a trilingual education system utilizing English and Filipino as well as vernaculars has developed. However, that may not be the ideal system. As García notes,

during the transitional stage, a *biomedial system* of instruction is supposed to be used. The instructor gives the gist of the lesson in the language prescribed, Filipino or English – and then explains to students in the local vernacular (Gonzalez, 1998). This policy officially moves away from the total separation of languages in instruction, although it does not go far enough in recognizing the translanguaging of the students, as they make sense of their multilingual learning environment (García 2009: 150).

As regards Philipino speakers' notions of their multilingual experience including English, we find those ideas shed light on important aspects of multilingualism in the Philippines. Blogs, diaries, text messages, and various webpages provide evidence of multilingual awareness and sociolinguistic attitudes. Even Q&A sites such as *Quora*<sup>5</sup>, quoting users' general opinions on different matters, may be used – though with caution for academic purposes – to investigate some interesting viewpoints on speakers' ideas of their multilingual experience(s). Following are some revealing excerpts from pages of the *Quora* which are dedicated to the issue of English proficiency levels in the Philippines.

**Speaker A:** *There are some native English speakers who grew up with it as their first language at home and in school. Such as myself. I don't speak Tagalog very*

<sup>5</sup> *Quora* is a famous question-and-answer website where the community of users can ask, answer and edit questions downloaded on the webpage with a voting system for each question which displays the most popular answers. See [www.quora.com/](http://www.quora.com/) (accessed April 2017).



well. It's not enough for the Philippines to be considered as a technically English-speaking country, though. As an immigrant to Australia I had to take the IELTS exam and achieve a high score to be considered as a skilled migrant.

**Speaker B:** It really depends but based on my observation, we Filipinos are highly adaptive in different language so to say, we can easily learn English at some sort. Some say Filipinos are one of the best English speakers in Southeast Asia alongside Singapore. The point is, *English isn't our first language whereas SG's primary language is English*. Another thing to pinpoint here is we speak English *regardless of grammar and pronunciation*. We happily claim that we're good speakers as long as we can utter words in accordance to our best knowledge.

**Speaker C:** I find it *very disturbing that the level of English among the younger generation is deteriorating thanks to the schools are not giving it the importance it once had*. In high school – 3<sup>rd</sup> grade – in a class with 55 students, they have *ONE HOUR of English lessons per WEEK*. Anyone can understand what comes out of that in practical and useful lessons. My high school son almost 16 years old, doesn't have any vocabulary, cannot communicate and do not have any comprehension whatsoever in English.

**Speaker D:** The level of English in the Philippines is *enough to function abroad* and enough to make us a really good *tourist trap*<sup>6</sup>.

Such disparate perceptions of English say something significant about its real role in the multilingual Philippine education system. English is a focal point in the language community's efforts at both harmonization and resistance. That community continues to search for stability as PE (Philippine English) progresses beyond nativization.

Importantly, motivated contact with the language and regular usage opportunities remain key factors in English proficiency, especially as regards accuracy in grammar and pronunciation. At the same time, what speaker C and D say about the actual status of English among the young generations speaks volumes about how much revitalization of the bilingual (or trilingual) education programs is still needed in a complex scenario where "English continues to dominate numerous domains, including intimate contexts of

<sup>6</sup> Quora excerpts are available at [www.quora.com/Whats-the-level-of-English-in-the-Philippines#!n=12](http://www.quora.com/Whats-the-level-of-English-in-the-Philippines#!n=12) (accessed April 2017).

the home, thus maintaining the status of the language as functionally native, and there is a growing awareness of PE as a language that is not deficient and may represent Filipino identity” (Pefianco Martin 2014: 79).

### 3. CLIL in multilingual education programs

Multilingual education has been considered from different viewpoints. The UNESCO position on education in a multilingual world has moved on from the resolution of the 1999 General Conferences and from the declaration of 2003 in which “multilingual education” was described as the *use of at least three languages, say, the mother tongue matched by a regional or national language and an international language in education*<sup>7</sup>.

Studies of linguistically complex education programs around the world such as those of the Philippine and Basque Country systems<sup>8</sup>, show something of the variety of policies implemented in order to integrate English as a second language of instruction. Importantly, the immersion projects developed to teach English, or to teach other subjects *in* English, represent very interesting cases in point. One of these, CLIL (*Content and Language Integrated Learning*), is the most promoted approach in European education systems. As a sort of English for Specific Purposes, CLIL aims at more than language proficiency. Developed in the 1990s, CLIL now functions as an umbrella term to indicate many language programs involved in teaching and learning even non-linguistic subjects. English is the language most commonly taught as a second or foreign language in European schools, and it is increasingly involved in the instruction of a wide range of academic and scholarly subjects. Still, although CLIL is strongly supported by the European Commission in its efforts to address the multilingual challenge, “there is no agreement about the scope of CLIL and the combination of content and language has been understood in different ways” (Cenoz – Gorter 2015b: 478).

As a matter of fact, in Cenoz – Gorter’s words:

<sup>7</sup> See [www.unesco.org/education/education\\_today/ed\\_today6.pdf](http://www.unesco.org/education/education_today/ed_today6.pdf) (accessed April 2017).

<sup>8</sup> In the Basque Country, children experience a rare language diversity from the very beginning of their school activity. Most of the children start attending school at the age of two, with Basque, the first language of some but not all pupils (others speaking Spanish as a first language), as the medium of instruction. English is introduced in the second or third year of preschool when children are three or four years old. For further references see Cenoz (2009).

CLIL can refer to teaching some subject content in the language class by having some activities or units on academic content, but it can also refer to the teaching of a school subject through the medium of English or another language. (...) CLIL has become quite popular in Continental Europe, and English is increasingly used as the language of instruction for some subjects in secondary school and higher education. There is not enough research to see the specific effect of using academic content to teach language as compared to the same amount of instruction and exposure in language classes. There is also not enough research that looks into the effects of CLIL on achievement on academic content either (Cenoz – Gorter 2015b: 478-479).

What emerges from studies of CLIL is a multi-faceted picture in which the psycholinguistic dimension is as significant as the pedagogical one. According to Pérez-Vidal – Roquet (2015: 238), language learning outcomes are measured primarily by just a few criteria:

- Quality of input
- Interaction
- Cognitive/learning abilities

Although this integrated approach to language and content learning is not limited to multilingual education policies alone, it continues to gain popularity in European language education programs where “it is a motivating force for the stakeholders, but also, and most importantly, for the learners themselves who probably see that CLIL fulfils some of the demands of their mindsets, such as new technologies, access to mobility and global communication” (Pérez-Vidal 2013: 76).

CLIL emphasizes a strong interdependence between teachers and students, and in that context the language difference becomes the main agent of a metalinguistic awareness which sees learners as varyingly efficient “users” of language from the very beginning of their language learning experience. Thus, “while traditional FL classrooms tend to treat learners as (deficient) *novices*, CLIL classrooms treat them as (efficient) *users*” (Lorenzo – Moore 2010: 24).

Increasing applications of CLIL in academic contexts, along with the utilization of second languages, particularly English, in higher education settings, can motivate and increase the knowledge of language learners, helping to actualize a multilingual scenario which may reconcile complex concerns about languages and cultural identities. This is especially true when second languages are also foreign languages. In such cases,

CLIL classrooms appear to be a clever and economic way of turning classrooms into 'streets' as it were. When there are no streets around the school in which the language could be picked up, one may try to convert school life, or parts of it, into a naturalistic environment where the toils of the foreign language classroom can be left behind (Dalton-Puffer 2007: 2).

Although the road map of CLIL experiments in language programs has been mostly restricted to European boundaries, the increasing interest in LAC (Language Across the Curriculum), EAL (English as an Additional Language) and CLIL is spreading also in countries such as "Hong Kong, mainland China, Thailand, the Philippines, Taiwan, Malaysia, Japan, and Korea" (Lin 2016: 3). The dominant role of English in the most prestigious universities in the Philippines is unchallenged and in a quite recent article Aquino explains how the five elements of CLIL (content, cognition, communication, community, and competence), may be part of General Education (GE) English courses in the University of the Philippines drawing inspiration and guidance from Rosenblatt's Transactional Theory of Reading and Writing. As she writes.

The UP Department of English and Comparative Literature (DECL) offers six General Education (GE) English courses and these can be classified into two: (1) the reading and writing courses (English 1, English 10, English 30, and Creative Writing 10) and (2) the literature courses (English 11 and English 12). Among these, English 10 (College English), English 30 (English for the Professions), English 11 (Literature and Society), and English 12 (World Literatures) more readily lend themselves to the CLIL approach. (Aquino 2016: 2).

Focusing on the reading-writing connection in GE classes, Aquino points out that "Vygotsky's and Bates's ideas of *meaning*, together with Rosenblatt's concept of a *linguistic-experiential reservoir* are useful in describing how the establishment of the reading-writing connection facilitates the blending and integration of content and language in a literature class" (Aquino 2016: 7). Being language at the heart of every negotiation of meaning (be it written or spoken), "in reading a literary text, the students learn not just the word but also the world – that is, both language and content (which is at the heart of CLIL). In writing about a literary text, the students are given an opportunity to apply what they have learned about the word and the world" (Aquino 2016: 10). However, the debate about English as a medium of instruction still

implies different pros and cons depending on the grade and the subjects learnt. As for the former and assuming the Philippines as the case in point of this paper, Sundqvist – Sylvén note that with the new policy adopted in 1974, which involved the use of both English and Filipino,

children were to be introduced to English and Filipino as L2s in grades 1 and 2, and from grade 3 onward these languages were to be used as medium of instruction. However, the results indicated that the academic achievement of these students was far from satisfactory. The downtrend in the educational achievement was attributed to a number of factors. For instance, many teachers lacked competence in the content material they were to teach; many of the teachers who were to use Filipino as the medium of instruction were not proficient enough in the language and there was a scarcity of teaching materials available (Sundqvist – Sylvén 2016: 53).

On the other hand, the effects of English as a medium of instruction of non-language subjects may be challenging and not always good. For instance, according to Costa – D'Angelo, some results may be quite negative when learning subjects such as "Mathematics, Science, Geography and History" (Costa – D'Angelo 2011: 6). For the same reason "the Philippine government has decided recently to move toward teaching in the native language beginning in 2009 (Republic of the Philippines, Department of Education) (Costa – D'Angelo 2011: 6).

#### **4. Concluding remarks**

As speakers of various languages on a global scale, we face multimodal and multilingual imperatives. Most of these imperatives have been identified by researchers studying English, the most widespread international lingua franca used today. Studies like the several mentioned here have addressed needs related to intercultural communication and the use of multiple codes. However, the multilingual turn in education programs and policies is still in its infancy. Cenoz – Gorter (2015a: 8) have claimed that it stems from a continuum along which speakers are "being and becoming multilingual" (Cenoz – Gorter 2015a: 8). Crucial to this continuum are the various multilingual contexts: "Students can 'be multilingual' because they are fluent in both the minority and the national language (Catalan/Basque

and Spanish) and at the same time ‘becoming multilingual’ because they go on learning these languages and additional languages such English” (Cenoz – Gorter 2015a: 8).

Clearly, multilingual education programs require a complex and holistic approach. Such an approach must involve the formulation of strategies for multilingualism at most or all levels of education. The Philippines “case” brings this issue to light, with its “over 120 languages, including the two official (Filipino and English) and nineteen ‘recognized’ regional languages” (Maher 2017: 2). Of course, national language education plans will vary for social, religious, political and economic reasons (take for instance the effects of the Tagalog linguistic imperialism since 1937, becoming Tagalog the linguistic base for Filipino); at the same time, the educational response will involve a wide range of pedagogical and political actors. Given sufficient attention, however, such a view of multilingual education could herald a more informed and a more equitable language world in which linguistic “otherness” fades in significance as the spread of English (beside other *linguae francae*) continues to function as an instrument of linguistic “sameness”.

## REFERENCES

### Special studies

Aquino, L.F.Y.

- 2016 CLIL in the General Education Classes in the University of the Philippines: Establishing the reading-writing connection”, *ELTWorldOnline.com, Special Issue on CLIL*, January 2016, 1-17.

Baker, C.

- 2000 *The Care and Education of Young Bilinguals*. Clevedon: Multilingual Matters.

Byram, M.

- 1997 *Teaching and Assessing Intercultural Communicative Competence*. Clevedon: Multilingual Matters.

Cenoz, J.

- 2009 *Towards Multilingual Education. Basque Educational Research from an International Perspective*. Toronto: Multilingual Matters.

Cenoz, J. – D. Gorter (eds.)

- 2015a *Multilingual Education*. Cambridge: Cambridge University Press.

Cenoz, J. – D. Gorter

- 2015b "Minority languages, State languages, and English in European education". In: W.E. Wright – S. Boun – O. García (eds.) *The Handbook of Bilingual and Multilingual Education*. Malden, MA: John Wiley & Sons, 473-483.
- Cooper, R.L.  
1989 *Language Planning and Social Change*. Cambridge: Cambridge University Press.
- Costa, F. – L. D'Angelo  
2011 "CLIL: A suit for all seasons?", *Latin American Journal of Content & Language Integrated Learning* 4(1), 1-13.
- Council of Europe  
2001 *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Dalton-Puffer, C.  
2007 *Discourse in Content and Language Integrated Learning Classrooms*. Amsterdam: Benjamins.
- Dayag, D.T.  
2004 "The English-language media in the Philippines", *World Englishes* 23(1), 33-45.
- Fishman, J.L.  
1976 *Bilingual Education: An International Sociological Perspective*. Rowley: Newbury House.
- Fortin, R.  
2005 *Comprendre la complexité. Introduction à La Méthode d'Edgar Morin*, 2 édition. Paris: Les Presses de l'Université Laval, L'Harmattan.
- García, O.  
2009 "Education, multilingualism and translanguaging in the 21<sup>st</sup> century". In: T. Skutnabb-Kangas – R. Phillipson – A.K. Mohanty – M. Panda (eds.) *Social Justice Through Multilingual Education*. Toronto: Multilingual Matters, 140-158.
- Gonzalez, A.  
2007 "Language, nation and development in the Philippines". In: L. Hock Guan – L. Suryadinata (eds.) *Language, Nation, and Development in Southeast Asia*. Singapore: Institute of Southeast Asian Studies. 7-16.
- Jordan, G.  
2004 *Theory Construction in Second Language Acquisition*. Amsterdam: Benjamins.
- Kamwangamalu, N.M.  
2016 *Language Policy and Economics: The Language Question in Africa*. London: Palgrave Macmillan.
- Kramsch, C.  
1993 *Context and Culture in Language Teaching*. Oxford: Oxford University Press.

- Lambert, W.E.  
1967 "A social psychology of bilingualism", *Journal of Social Issues* 23, 91-109.
- Lambert, W.E.  
1975 "Culture and language as factors in learning and education".  
In: A. Wolfgang (ed.) *Education of Immigrant Students*. Toronto: Institute for English Studies in Education, 55-83.
- Lin, A.M.Y.  
2016 *Language Across the Curriculum & CLIL in English as an Additional Language (EAL) Contexts*. Hong Kong: Springer.
- Lorenzo, F. – P. Moore  
2010 "On the natural emergence of language structures in CLIL. Towards a theory of European educational bilingualism". In: C. Dalton-Puffer – T. Nikula – U. Smit (eds.) *Language Use and Language Learning in CLIL Classrooms*. Amsterdam: Benjamins, 23-38.
- Maher, J.C.  
2017 *Multilingualism*. Oxford: Oxford University Press.
- Morin, E.  
1990 *Science avec conscience*. Nouvelle édition. Paris: Le Seuil-Points.
- Pefianco M.I.  
2014 "Beyond Nativization? Philippine English in Schneider's Dynamic Model". In: S. Buschfeld – T. Hoffmann – M. Huber – A. Kautzsch (eds.) *The Evolution of Englishes*. Amsterdam: Benjamins, 70-85.
- Pérez-Vidal, C.  
2013 "Perspective and lessons from the challenge of CLIL experience". In: C. Abello-Contesse – P.M. Chandler – M.D. López-Jiménez – R. Chacón-Beltrán (eds.) *Bilingual and Multilingual Education in the 21<sup>st</sup> Century*. Toronto: Multilingual Matters, 59-84.
- Pérez-Vidal, C. – H. Roquet  
2015 "CLIL in Context: Profiling Language Abilities". In: M. Juan-Garau – J. Salazar-Noguera (eds.) *Content-based Language Learning in Multilingual Educational Environments*. Berlin: Heidelberg, 237-252.
- Samson, A.R.  
2013 "Language fusion". *Business World Online*, October 28.  
<http://www.bworldonline.com/content.php?section=Opinion&title=language-fusion&id=78518>, accessed March 2017
- Smakman, D. – P. Heinrich (eds.)  
2015 *Globalising Sociolinguistics: Challenging and Expanding Theory*. London: Routledge.
- Smedley, F.  
2006 *Code-switching and Identity on the blogs: An Analysis of Taglish in Computer-Mediated Communication*. Auckland: Auckland University of Technology.  
<http://aut.researchgateway.ac.nz/handle/10292/186>, accessed March 2017



Sundqvist, P. – L.K. Sylvén

2015 *Extramural English in Teaching and Learning*. London: Palgrave Macmillan.

Thompson, R.M.

2003 *Filipino English and Taglish*. Amsterdam: Benjamins.

### Web sources

Quora.com

<https://www.quora.com/>, accessed April 2017

The Map of the Top Ten English Speaking Countries

<http://www.mapsofworld.com/world-top-ten/countries-with-most-english-language-speaker-map.html>, accessed April 2016

UNESCO

[http://www.unesco.org/education/education\\_today/ed\\_today6.pdf](http://www.unesco.org/education/education_today/ed_today6.pdf), accessed April 2017

Address: ANNALISA BONOMO, “Kore” University of Enna, Centro Linguistico Interfacoltà Kore (CLIK), Sede Centrale-Rettorato, Cittadella Universitaria, s.n., 94100, Enna, Italy.

ORCID Code: [orcid.org/0000-0002-9593-576X](https://orcid.org/0000-0002-9593-576X).

## **Genitive alternation in New Englishes: The case of Nigerian English**

Mayowa Akinlotan

*Vrije Universiteit Brussels (VUB)*

### **ABSTRACT**

This paper is concerned with influences on the forms of possessive expression in Nigerian English compared to influences reported in other Englishes such as British, American, and Canadian. The study examines and compares independently the four commonly investigated determiners of animacy, text type, prototypicality, and topicality, and it also shows the extent to which variation is attested in possessive form alternation in Nigerian English. The evidence adduced was drawn mainly from the Nigerian component of the International Corpus of English (ICE). More than 3000 data, mainly written register attestations of alternation, were analyzed. The findings suggest that animacy is the primary determiner of possessive expression form in Nigerian English. Prototypicality and syntactic weight/length, which have also been shown to exert strong influences, evidently have very little influence in this variety of English. Because multilingualism is widespread in Nigeria, these patterns likely indicate grammatical structuring partly or wholly derived from local Nigerian languages that have no such alternating system.

Keywords: genitive alternation, Nigerian English noun phrase, animacy, syntactic weight.

### **1. Introduction**

The medium of language allows us to organize and express our thoughts in different yet limited forms. However, these different structures are usually constrained by the grammar of the language in use, such that having decided what to communicate, we thus arrange them into syntactic constructions, generating output best suited to the purpose and context. In some cases, the grammar provides us with more than one choice about how to convey the same, or nearly the same message, whereas in others we have only one

choice of construction. In expressing possession, two options are ordinarily possible. The choice of one over the other has been shown to be influenced by various factors, and their influences have also been shown to differ from one variety of English to another. The common factors of animacy, topicality, prototypicality, and length have been reported to be causal in many publications, albeit in varying degrees (Rosenbach 2014).

Of all syntactic alternations in English, the genitive has been the most widely treated in current scholarship. Rosenbach (2014: 215) argues that it is “Today, [...] arguably the best researched of all syntactic alternations in English”. Still, there has been no single work demonstrating the extent to which those common factors behave independently in motivating the expression of possession in the Nigerian English noun phrase. The present study is concerned with how the factors of animacy, prototypicality, topicality, and weight varyingly influence the choices of possessive genitive construction, and how frequently they do so. Consider the alternation in (1):

- (1) a. the Federal government’s exclusive right
- b. Exclusive right of the federal government

This paper intends to show which, or which combination, of these factors is most influential in various instances in the choice between the Saxon, suffixal genitive and the Latinate, phrasal genitive as well as how those factors behave differently, as seen in their comparative frequencies of occurrence, in Nigerian English than they do in “older” varieties of English such as the British, American, and Canadian varieties.

## 2. Genitive variables influencing choices

There has been little agreement in the literature as to the relative strength of the various factors motivating genitive alternation. Thus far, animacy and syntactic weight have been analysed most often and have been found to be key factors, motivating genitive choices independently or simultaneously. Some researchers have tested the relative significance of these as well as other factors (Altenberg 1982; Jucker 1993; Leech et al. 1994; Arnold et al. 2000; Kreyer 2003; Szmrecsanyi 2013). However, common prediction is lacking because the relative independent and/or interactional strength of these factors is subject to different external and internal language variables. Predictions of the several factors are tabulated below (cf. Szmrecsanyi 2010: 2; Rosenbach 2014: 225-227).

### 3. Conceptualisation and predictions

In the context of previous research, expectations for the motivation of each factor are set out here. They are formulated in accordance with the hypothesis that grammars/constructional choices produced in Nigerian English are highly influenced by and reflective of the grammars of certain indigenous languages, languages which are in contact with this variety of English and which exhibit much less influence by these factors in regard to possessive construction choices.

#### 3.1 Animacy

Animacy is well known to be an influential parameter motivating alternation in almost all varieties of English (Alternberg 1982; Hawkins 1994; Rosenbach 2005, 2014; Hinrichs–Szmrecsanyi 2007; Szmrecsanyi 2010). More specifically, the animacy of the possessor, not that of the possessum, typically determines the choice of *s*-genitive or *of*-genitive, and this is borne out in the amount of attention that has been paid to the influence of the possessor's animacy status (cf. Gries – Stefanowitch 2004). The general conclusion is that animate possessors regularly take the *s*-genitive and inanimate possessors the *of*-genitive (Rosenbach 2014). Complicating matters, however, animacy itself is categorized variously among researchers (Kreyer 2003; Rosenbach 2005; Wolk et al. 2013). Wolk et al. (2013) propose five classes: human, collective, temporal, inanimate, and locative. In this scheme, nouns denoting animals fall into the class human. Kreyer (2003) collapses the categories of proper nouns, common nouns, and collective nouns in an effort to devise a personality scale. Common to the diverse groupings along the scale is a distinction of geographical and commercial entities as proper nouns because they behave referentially as proper nouns having one conceptually tangible referent.

The present study focuses on animacy and possessors with the expectations that not only will animate possessors decisively select the *s*-genitive, but inanimate possessors will indecisively select the *of*-genitive such that animacy will be shown to be an independent factor not decisive in predicting genitive construction choices in Nigerian English. Here animacy is categorized into three sets: human, peopled, and inanimate. The human and peopled sets are considered those of animacy, and, obviously, the inanimate set is regarded as that of inanimacy. This last set features nouns denoting animals, plants, and lifeless entities. A human sub-level accommodates proper nouns (e.g. Mary, Bandele) and supernaturals (e.g. God's, Chukwu).

The class *peopled* contains nouns referring to people-oriented entities, including, but not limited to, the names of places/countries/continents (e.g. Lagos, Amsterdam, Nigeria, Africa) as well as organizations/companies (e.g. the federal government, Shell).

### 3.2 Syntactic weight

The syntactic weight, or relative length, of possessor and possessum has been shown to determine choices in grammatical variation, and particularly, in English genitive alternation. At the most general level, it is widely accepted that language users realize short constituents before long constituents. Hawkins (1981), Rosenbach (2002, 2005), Szmrecsanyi – Hinrichs (2008), and Szmrecsanyi (2010) show that the lengths of the possessor and the possessum can influence genitive form choices. As a group, these scholars maintain that (1) the *s*-genitive is preferred when the possessor is shorter while the *of*-genitive is favored when the possessor is longer, (2) the *s*-genitive is preferred when the possessum is longer, while *of*-genitive when the possessum is shorter. Moreover, strong interrelation between animacy and weight is reported by both Hawkins (1994) and Rosenbach (2005). Hawkins (1994) argues further that weight is the primary motivator behind alternation. In this investigation, we do not expect shorter-possessor-longer-possessum-for-*s*-genitive nor longer-possessor-shorter-possessum-for-*of*-genitive. Rather, we predict short-possessor-short-possessum for both genitive forms such that length of possessor and/or possessum will be found to be important but not decisive as an independent factor determining genitive construction choice in Nigerian English.

### 3.3 Semantic relation

The conceptual distance between the possessor and possessum has been mapped and demonstrated to be a consequential factor in English genitive form choices. Relevant to this factor is Haiman's (1983) iconicity principle, by which the closer the meanings of the possessor and the possessum, the higher the probability that the *s*-genitive construction will occur. Accordingly, if the possessor and the possessum have meanings that are not significantly proximate, then the *of*-genitive is likely to occur. Mapping such semantic distance as a factor motivating English genitive alternation has typically been guided by prototype theory. Koptjevskaja – Tamm (2003) use a binary classification in which prototypical meanings are contrasted with non-prototypical ones. The designation prototypical obtains when possessor

and possessum express a commonplace relationship and may therefore be considered conceptually close (e.g. the Pastor's robe). The opposite designation is employed when the two are not in such a relationship (e.g. the Pastor's educational qualifications). Rosenbach (2002) makes a finer distinction in which prototypicality is assigned to subsets of terms for body parts, kinship members, parts of wholes, and legal ownership. In contrast, non-prototypicality is assigned to all others. Given that what is conceptually close or far in Nigerian English may or may not be so in other varieties of English, such as British English and American English, an approach to prototypicality similar to Rosenbach's is utilized here, though one including subsets which suit Nigerian semantic idiosyncrasies. It should be noted that Payne – Berlage (2011) have demonstrated that semantic relation significantly motivates genitive alternation. Here, we expect prototypicality to behave in Nigerian alternation as it has been predicted to do so by Rosenbach (2002; 2014: 229) and others (e.g. Kreyer 2003; Payne – Berlage 2011). That is, prototypical semantic relations will be expected to attract the *s*-genitive and non-prototypical ones the *of*-genitive.

### 3.4 Topicality

Knowledge of a possessor or a possessum has been shown to impact positively on the selection of the shorter *s*-genitive. If the referent of the possessor construction is known to the speaker/writer and the listener/reader, then prior or shared knowledge of the possessor is present. Focusing on topicality of possessor, Rosenbach (2002) distinguishes between a referentially given possessor and a new possessor. In the main, topicality of possessor highlights a referent which is known to the speaker/writer and the listener/reader and which, ordinarily, has recently been mentioned. Additionally, topicality of possessor involves definite expression of a referent in the real world, or one already mentioned, knowledge of whose existence is also shared. Hinrichs – Szmrecsanyi (2007:451) operationalise topicality in terms of givenness and non-givenness, and find it to have an insignificant effect on genitive alternation. However, Grafmiller (forthcoming: 18), cited by Rosenbach (2014: 228-229), finds topicality in fact to be significant as a motivator of genitive alternation.

In this investigation, following partly Rosenbach (2002) and partly Hinrichs – Szmrecsanyi's (2007), topicality and givenness are collapsed into a single factor. Since topicality correlates highly with animacy (Rosenbach 2014: 230), this factor alone is not expected to influence genitive form choices significantly.

### 3.5 Text type

The nature of the text in which we express our thoughts has been found to influence choices of genitive form. On one hand, Alternberg (1982: 284) and Dahl (1971: 172) claim that more formal texts tend to utilize the *s*-genitive and less formal ones the *of*-genitive. On the other, Biber et al. (1999: 300) show that the pressure to be brief is more permissible in certain text types than it is in others, which accounts for the comparative abundance of *s*-genitives in less formal texts. Use of this genitive helps satisfy requisites of length economy. Hinrichs – Szmrecsanyi (2007) find that texts required to disseminate large amounts of information markedly prefer the *s*-genitive. Press language, or journalese, in contrast to non-press language, has been shown to make greater use of the *s*-genitive, likely because it exhibits less formality and also because it needs to be “more compact”. This text type features a “tendency to brevity” (Rosenbach 2014: 233-234). It should be noted that most of these conclusions are based on text types of British and American varieties of English. Since, according to Hinrichs – Szmrecsanyi (2007), the tendency for text types such as journalese to prefer the *s*-genitive over the *of*-genitive seems to show the effect of lexical density, here, where the focus is on Nigerian English, we do not expect lexical density, or “tendency to brevity”, to be influential.

## 4. Method

Data analysed in this study were culled from a variety of written register texts of the Nigerian component of the International Corpus of English. However, limits were necessarily imposed on the types of data considered. Only those constructions that can exhibit alternation of genitive form without change of meaning were analyzed. Those which cannot were excluded. Examples (2a) and (3a) below illustrate what is considered zero meaning change or retention of meaning, which is key to interchangeability. Excluded were contracted forms as well as those that fall into categories (2b) and (3b) where alternation (into 2b and 3b) causes a shift or loss of the original meaning, resulting in ambiguity<sup>1</sup>.

<sup>1</sup> For a detailed description of criteria for exclusion, with examples, see Rosenbach (2002, 2014).

- (2) a. Welfare of teachers  
       b. Chief of Staff  
           Interchangeability and non-interchangeability: teachers' welfare  
           versus Staff's Chief
  
- (3) a. Bloomfield's behaviourism  
       b. Pregnancy's plans  
           Interchangeability and non-interchangeability:  
           Behaviourism of Bloomfield versus plans of pregnancy

After the exclusion criteria were applied, a sample of 3371 interchangeable genitive constructions, divisible into 1299 *s*-genitives and 2072 *of*-genitives, remained. The 3371 items were then analysed in regard to the several influential factors mentioned above. The analyses presented in sections 5.1 – 5.5 show the strength of each factor as an independent determiner of genitive form choices. The operationalisation of the determiners as well as the annotation procedures are discussed immediately below in Sections 4.1 – 4.4.

#### 4.1 Animacy

Classifying the animacy of possessors required some revision of the common animacy categorisations present in the literature (e.g. Bergen 2011, Rosenbach 2008, Zaenen et al. 2004). In the end, the four main categories of human, peopled, animal/plant, and article were employed. Human and peopled were classified as animate while animal/plant and article were classed as inanimate. This scheme differs from those of most researchers in that animals and plants are classified not as animate but rather as inanimate (cf. Kreyer 2003, Bergen 2011, Rosenbach 2003). It also differs from the common schemes in its collapse of the categories location and organisation into one labelled peopled. Human embraces names ordinarily signified by proper nouns, names that refer to distinct entities, and names for supernatural entities. Peopled takes in the names of places, countries, continents, organizations, societies, and companies. This category collapses the classes of location and organization which are kept distinct by various other researchers (e.g. Zaenen et al. 2004). By way of illustration, *world* and *Nigeria* are coded as peopled in the following constructions: *The world is at war with itself* and *Nigeria is the second most culturally diverse in the world*.



## 4.2 Syntactic weight

The syntactic weights of the possessor and the possessum concerned only adjectives, single-word adjectivals, and nouns. These lexical constituents were considered significant elements of the possessor and possessum, and they were counted in order to quantify respective syntactic weights. The annotation below details syntactic weight as length (Altenberg 1982; Rosenbach 2005). Determiners, quantifiers, and prepositions were not counted. For example, *the world's information* was counted as a 1-word possessor (*world*) and a 1-word possessum (*information*) while *Five important steps of political developments* was deemed to contain 4 words (*important*, *steps*, *political*, and *developments*), so that both possessor and possessum were counted as 2-words, respectively. Therefore, our operationalisation of syntactic weight involved the quantification of lexical constituents but not grammatical constituents.

## 4.3 Topicality

The topicality of the possessor was measured in a manner which follows Rosenbach (2003). Proper nouns signifying distinct entities such as persons (e.g. *Helen*) or countries (e.g. *Nigeria*) were annotated as topical, as were instances of the definite article. Consider (4) and (5):

- (4) the boy's eyes
- (5) her husband's needs

The use of *the* in these noun phrases suggests foreknowledge of the possessor *boy*, perhaps due to previous mention, and that knowledge is shared by the writer/speaker and the reader/listener. Although Rosenbach (2003) collapsed constructions like those of (4) and (5) into the single category of definiteness, they were viewed separately in order to obtain a fuller picture. Thus, because the use of the possessive determiner *her* in (5) implies shared knowledge of the possessor, this and similar instances were accounted as topical. However, markers of indefiniteness, such as the article *a* in the phrase *a boy's eyes*, revealed a lack of foreknowledge (in the real world) and disqualified such instances from being regarded as topical. Quantifiers such as *one* in *one's baby* were similarly disqualifying. Lastly, while Rosenbach (2003) considered items containing *a*, *some*, *many*, *any* as of the single category indefinite expression and deemed them non-topical, such items were accounted for separately

in the present study, which was meant to allow for a finer analysis of the system underlying the noun phrase configurations.

#### 4.4 Prototypicality

The notion of prototypicality may be used to measure the conceptual distance between the possessor and the possessum. Rosenbach (2002) operationalises this notion and reckons as in close relationship, and therefore prototypical, items denoting kinship members, body parts, and legal ownership. Other types of relationship between the possessor and possessum are classified as non-prototypical. In this investigation, Rosenbach's categorization of prototypicality was altered to include those types of possessive relationship that are more generic, such that the possessor-possessum relationship evident in *student's learning* would be considered as conceptually close and so prototypical. When the conceptual distances were so great that causality could not be established, the relationships were reckoned non-prototypical. Here, prototypicality was used to measure possessive relationships conceived of as either conceptually close (+) or conceptually far (–).

### 5. Results

In this section, the results of how each factor behaved independently are presented.

#### 5.1 Animacy

Table 1 shows genitive alternation by animacy.

Table 1. Genitive alternation by animacy

	s-genitive		of-genitive		Total	
	n	%	n	%	n	%
Animate possessor	1201	54	1013	46	2214	100
Inanimate possessor	98	8	1058	92	1156	100
Total	1299	39	2071	61	3370	100

As can be seen, there is some relationship between the animacy of the possessor and genitive form choice  $\{\chi^2(1) = 671, p < 0.0000\}$ . The difference

between animate possessor (54%) and inanimate possessor (8%) is significant, such that animate possessors are clearly more likely to occur in *s*-genitive construction than are inanimate possessors. However, there is the same likelihood that an animate or an inanimate possessor would occur in an *of*-genitive construction.

## 5.2 Syntactic weight of possessor

Table 2 displays length of possessor in relation to genitive construction type.

Table 2. Length of possessor in relation to genitive construction type

	s-genitive		of-genitive		Total	
	n	%	n	%	n	%
1-word	1138	44	1420	56	2558	100
2-word	133	21	508	79	641	100
3-word	20	20	114	80	134	100
4-word	8	26	23	74	31	100
5-word	0	0	2	100	2	100
6-word	0	0	4	0	4	100
Total	1299	39	2071	61	3370	100

As presented above, there is very little relationship between length of possessor and genitive form choice  $\{\chi^2(5) = 32.26, p < 0.0000\}$ . At once, a one-word possessor is shown to be more likely than a two or three-word possessor to occur in the *s*-genitive and a one-word possessor is seen to be more likely than a two or three-word possessor to occur in the *of*-genitive.

## 5.3 Weight of possessum

Table 3 shows length of possessum in relation to genitive construction choice. As displayed here, unlike with length of possessor shown in Table 4 below, there is some relationship between length of possessum and genitive form choices  $\{\chi^2(4) = 28.85, p < 0.0000\}$ . A one-word possessum is more likely than a two, three, or four-word possessum to occur in the *s*-genitive and the *of*-genitive. Also, while there is little or no likelihood of a five-word or six-word possessor occurring in the *s*-genitive (see Table 4), there is some likelihood of a five-word possessum occurring in the *s*-genitive.

Table 3. Length of possessum in relation to genitive construction choice

	s-genitive		of-genitive		Total	
	n	%	n	%	n	%
1-word	865	35	1641	65	2506	100
2-word	311	45	385	55	696	100
3-word	95	72	37	28	132	100
4-word	22	73	8	27	30	100
5-word	6	100	0	0	8	100
6-word	–		–		–	
Total	1299	39	2071	61	3370	100

## 5.4 Topicality

Table 4 exhibits topicality of possessor in relation to genitive choices.

Table 4. Topicality of possessor in relation to genitive choices

	s-genitive		of-genitive		Total	
	n	%	n	%	n	%
Topical possessor	1171	43	1575	57	2746	100
Non-topical possessor	128	21	496	79	624	100
Total	1299	39	2071	61	3370	100

As shown above, there is some relationship between topicality of possessor and genitive choice  $\{\chi^2(1) = 105.13, p < 0.0000\}$ . A topical possessor is more likely to occur as an *s*-genitive (43% versus 21%), while a non-topical possessor is more likely to occur as an *of*-genitive (79% versus 57%).

## 5.5 Prototypicality

Table 5 displays the frequency distributions of genitive choices by prototypicality.

Table 5. Relative frequency distribution of genitive choices by prototypicality

	s-genitive		of-genitive		Total	
	n	%	n	%	n	%
Prototypicality	37	23	91	18	128	100
Non-prototypicality	125	77	411	82	536	100
Total	162	24	502	76	664	100

As can be seen in Table 5, there is an inverse correlation between genitive form choice and possessor-possessum semantic relation  $\{\chi^2(1) = 0.87, p < 0.3501\}$ . When possessor and possessum have little or no semantic relationship, the choice can be either *s*-genitive (77% versus 23%) or *of*-genitive (82% versus 18%).

## 6. Conclusions

The results of the study are summarized in Table 6 below. There, the behaviour of each factor as an independent variable which motivates genitive alternation in the Nigerian English noun phrase is indicated.

Table 6. Summary of findings of factors as independent variables

Variables	's-genitive	of-genitive
Animacy of possessor	+	–
Topicality of possessor	+	+
Prototypicality	–	–
Weight of possessor	+	+
Weight of possessum	+	+

If we compare the contents of Table 6 with those of Tables 1 and 2, we can easily discern the noteworthy results. From these we may draw two main conclusions.

First, the factor of animacy behaves in Nigerian English much as it does in British English and American English. In all three varieties, an animate possessor is more likely to take *s*-genitive form and an inanimate possessor *of*-genitive form.

Second, the factors of topicality, prototypicality, and syntactic weight behave differently in Nigerian English. The choice between *s*-genitive and *of*-genitive is not in this variety of English influenced by the level of topicality. Both genitive forms occur when topicality is high and both when topicality is low. This finding for Nigerian English contrasts with Rosenbach's (2002) for other varieties of English, according to which high topicality is linked to the selection of the *s*-genitive. Similarly, genitive form choices are not determined by prototypicality in Nigerian English. Whether a possessor and possessum are semantically close is immaterial. Because this factor has been shown to be influential in British, American, and Canadian Englishes, this

is an important finding. As regards syntactic weight, the effect of length is bidirectional here rather than unidirectional, as it has been found to be in the varieties just mentioned (Szmrecsanyi 2010; Jankowski 2009). One-word possessors take *s*-genitive form as often as they take *of*-genitive form, and the same is true of one-word possessums.

The differences in the behaviours of these factors in Nigerian English suggest interference from local Nigerian languages which multilingual Nigerians also speak. Those languages exhibit no such genitive alternation. Though an *of*-genitive equivalent may in some cases be interchangeable with an adjective phrase (e.g. in the local language equivalents of *the exclusive right of the federal government* and *the federal government exclusive right*), an equivalent of the *s*-genitive is absent from the syntactically unique local Nigerian languages. In that language contact situation, including the presence of French and its *of*-genitive, transfer of patterning from languages in contact with Nigerian English seems the most reasonable explanation for these differences. Nigerian English, like the many other varieties spoken by bilinguals and multilinguals, is increasingly differentiating itself from established Englishes like British English and American English.

## REFERENCES

### Sources

*International Corpus of English (Nigeria component).*

Online at <<https://sourceforge.net/projects/ice-nigeria/>> (Accessed March 1, 2016).

### Special studies

Altenberg, B.

1982 *The Genitive v. the of-Construction. A Study of Syntactic Variation in 17<sup>th</sup> Century English*. Malmö: CWK Gleerup.

Arnold, J.E. – T. Wasow – R. Ginstrom – A. Losongco

2000 "Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering". *Language* 76.1, 28-55.

Bergen, G. van.

2011 *Who's first and what's next: Animacy and Word Order Variation in Dutch Language Production*. Doctoral dissertation: Radboud University Nijmegen

- Biber, D. et al.  
1999 *The Longman Grammar of Spoken and Written English*. London: Longman.
- Dahl, L.  
1971 "The S-genitive with Non-Personal Nouns in Modern English Journalistic Style. *Neuphilologische Mitteilungen* 72, 140-172.
- Gries, S. Th. – A. Stefanowitsch  
2004 "Extending collocation analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9.1, 97-129.
- Haiman, J.  
1983 "Iconic and Economic Motivation". *Language* 59, 781-819.
- Hawkins, J.A.  
1994 *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Hawkins, R.  
1981 "Towards an account of the possessive constructions: NP's N and the N of NP". *Journal of Linguistics* 17, 247-269.
- Hinrichs, L. – B. Szendrői  
2007 "Recent changes in the function and frequency of Standard English genitive constructions: A multivariate analysis of tagged corpora". *English Language and Linguistics* 11.3: 437-474.
- Koptjevskaja-Tamm, M.  
2003 "Possessive noun phrases in the languages of Europe". In: F. Plank (ed.) *Noun Phrase Structure in the Languages of Europe*. Berlin: Mouton de Gruyter, 621-722.
- Jankowski, B.  
2009 Grammatical and register variation and change: A multicorpora perspective on the English genitive. Paper presented at the American Association for Corpus Linguistics (AACL).
- Jucker, A.H.  
1993 "The genitive versus the of-construction in newspaper language". In: A.H. Jucker (ed.) *The Noun Phrase in English: Its Structure and Variability*. Heidelberg: Winter, 121-136.
- Kreyer, R.  
2003 "Genitive and of-construction in modern written English: Processability and human involvement". *International Journal of Corpus Linguistics* 8, 169-207.
- Leech, G. – B. Francis – X. Xunfeng Xu  
1994 "The use of computer corpora in the textual demonstrability of gradience in linguistic categories. In: C. Fuchs – B. Victorri (eds.), *Continuity in Linguistic Semantics*. Amsterdam: Benjamins, 57-76.
- Payne, J. – E. Berlage  
2011 The effect of semantic relations on genitive variation: relations hierarchy. Paper presented at ISLE 2, Boston, MA, 17-21 June 2011.

- Rosenbach, A.  
 2002 *Genitive Variation in English. Conceptual Factors in Synchronic and Diachronic Studies*. Berlin: Mouton de Gruyter.
- Rosenbach, A.  
 2003 "Aspects of iconicity and economy in the choice between the s-genitive and the of-genitive in English. In: G. Rohdenburg – B. Mondorf (eds.) *Determinants of Grammatical Variation in English*. Berlin: Mouton de Gruyter, 379-412.
- Rosenbach, A.  
 2005 "Animacy versus weight as determinants of grammatical variation in English". *Language* 81.3, 613-644.
- Rosenbach, A.  
 2014 "English genitive variation-the state of the art". *English Language and Linguistics* 18.2, 215-262.
- Szmrecsanyi, B.  
 2010 "The English genitive alternation in a cognitive sociolinguistics perspective". In: D. Geeraerts – G. Kristiansen – Y. Perisman (eds.) *Advances in Cognitive Sociolinguistics*. Berlin: Mouton de Gruyter, 141-166.
- Szmrecsanyi, B.  
 2013 "The great regression: Genitive variability in Late Modern English news texts". In: K. Borjars – D. Denison – A. Scott (eds.) *Morphosyntactic Categories and the Expression of Possession*. Amsterdam: Benjamins, 59-88.
- Szmrecsanyi, B. – L. Hinrichs.  
 2008 "Probabilistic determinants of genitive variation in spoken and written English: A multivariate comparison across time, space and genres". In: T. Nevalainen – I. Taavitsainen – P. Pahta – M. Korhonen (eds.) *The Dynamics of Linguistic Variation: Corpus Evidence on English Past and Present*. Amsterdam: Benjamins, 291-309.
- Wolk, C. – J. Bresnan – A. Rosenbach – B. Szmrecsanyi.  
 2013 "Dative and Genitive Variability in Late Modern English: Exploring Cross-constructural Variation and Change". *Diachronica* 30.3, 382-419.
- Zaenen, A. – J. Carletta – G. Garretson – J. Bresnan – A. Koontz-Garboden – T. Nikitina – M.C. O'Connor – T. Wasow  
 2004 "Animacy encoding in English: why and how". Proceedings of the Association for Computational Linguistics Workshop on Discourse Annotation, available at [http://clair.eecs.umich.edu/aan/paper.php?paper\\_id=W04-0216](http://clair.eecs.umich.edu/aan/paper.php?paper_id=W04-0216) (accessed May 2017).

Address: MAYOWA AKINLOTAN, Department of Linguistics and Literary Studies, Vrije Universiteit Brussels, Pleinlaan 2, B-1050 Brussel, Belgium.

ORCID code: [orcid.org/0000-0002-4870-1856](http://orcid.org/0000-0002-4870-1856).



# **The impact of Arabic on the English lexicon since 1801**

Julia Schultz

*University of Heidelberg*

## **ABSTRACT**

Arabic borrowings which were introduced into English during the nineteenth and twentieth centuries have as yet been relatively neglected in previous investigations. The present paper will analyse the influence of Arabic on English vocabulary since 1801. The results presented in this study rely on a comprehensive lexicographical sample of 302 nineteenth- and twentieth-century Arabic borrowings collected from the *Oxford English Dictionary* Online. The Arabic-derived words under consideration will be divided into various subject areas and spheres of life (such as the fine arts and crafts, gastronomy, politics, Islam) in order to give a rounded picture of the fields from which Arabic borrowings were adopted into English during the last two centuries.

Keywords: lexicology, language contact, Arabic influence on the English vocabulary, online dictionaries in lexicological research.

## **1. Introduction**

### **1.1 Previous studies of Arabic influence on English vocabulary**

There are very few studies of the Arabic influence on the English lexicon. Kaye's (1992) essay, for instance, concentrates on the orthographical variation of a number of widespread Arabic words and names which occur in English texts, such as the phrase *Allahu akbar* and its spelling variant *Allaho akbar*, the meaning of which can be paraphrased as 'God is great'. Kaye (1992: 32) outlines that

With the Middle East, the Arab world, and much of the rest of the Muslim world being so conspicuous in the news since the Iraqi invasion of Kuwait in August 1990, the many forms in English print for some Arabic loanwords further prompt a study of such variations.

In addition, the work entitled *The Arabic Contributions to the English Language: An Historical Dictionary* (1994) by Cannon – Kaye should be mentioned here, as it is a large collection of words and phrases taken from Arabic throughout the centuries. Numerous major English dictionaries were surveyed in order to collect both direct and indirect borrowings from Arabic, comprising, for example, the second edition of the *OED* (1989), and dictionaries of new words, such as Ayto's *Longman Register of New Words* (1989, 1990) and Tulloch's *Oxford Dictionary of New Words* (1991). The number of borrowings examined by Cannon – Kaye amounts to 2338 lexical items arranged in alphabetical order in the dictionary section. The book also encompasses three articles: the first article offers a rounded picture of the chronological distribution of the borrowings adopted from Arabic. Furthermore, the words and phrases under review are grouped into 46 semantic areas, ranging from agriculture, botany, music, literature, military to politics and sociology, so as to provide an overview of the various domains from which Arabic borrowings entered English through the ages. The focus of the second article is on grammatical assimilation, and that of the third is on the phonological assimilation of the words of Arabic provenance. This investigation gives significant insights into the contact between Arabic and English in many regards. Still, it does not include an in-depth analysis of the manifold areas from which Arabic borrowings were introduced into English during the last two centuries. The present paper will offer a more up-to-date examination of the semantic areas which have been influenced by Arabic since 1801.

The reader may notice that Cannon – Kaye consulted a new type of dictionary in order to collect Arabic borrowings. That is, a dictionary of new words similar to others like Tulloch's *Oxford Dictionary of New Words* published in 1991. There are also dictionaries concentrating on foreign lexical items which have become part of the English language, including Bliss's (1966) *Dictionary of Foreign Words and Phrases in Current English*. Durkin, the deputy chief editor of the *OED*, draws attention to the fact that the lexical items which are recorded in these dictionaries will be listed in the *OED* if they correspond to the necessary inclusion requirements<sup>1</sup>. He states that "the most important of these dictionaries have been read (or "carded") for *OED*'s files, and all of

<sup>1</sup> Durkin, p.c. in email dated 9<sup>th</sup> February 2010.

them are available for consultation by *OED* editors”<sup>2</sup>. The *OED* thus serves as the main source of the Arabic-derived words presented in this paper.

In his study *Borrowed Words: A History of Loanwords in English*, Durkin (2014: 384-385) summarizes the development of Arabic influence on the English language over the centuries and the proportion of borrowings recorded in the *OED* as follows:

The number of direct loanwords from Arabic reflected by the *OED* is relatively small, but there is an unusually large number of indirect loans via other languages. [...] It is uncertain whether there are any loanwords directly from Arabic before 1500; apparent cases, such as the star name *Aldebaran* (a. 1393), probably just reflect a gap in the historical record for the language of transmission (in this case, like most others, probably Latin or French). However, many words ultimately of Arabic origin entered English in the Middle English and Early Modern periods. [...] Small numbers of (probable) direct borrowings are found in Early Modern English and the numbers subsequently climb modestly, peaking round about 1800 [...]. In many cases, especially earlier in this period, it is difficult to be certain that a loan has come directly from Arabic rather than via another language [...]. Most of these loans relate specifically to aspects of Arabic culture (or to Islam), e.g. *hashish* (1598), *madrassa* (1616), *hammam* (1625), *kohl* (1799), *hadj* (1847), *hijab* (1885), *mujahidin* (1887). Greater currency is shown by some of the large number of later loanwords ultimately of Arabic origin that have entered English wholly or partly via French [...].

The focus of linguistic concern of the present analysis will be on those borrowings which have been adopted directly from Arabic since 1801. As will be seen, the *OED Online* represented an essential tool in identifying the various words that show an Arabic origin in their etymological description. In addition, the *Oxford Dictionary of English Etymology*, edited by Onions *et al.* in 1966, offered additional etymological information concerning the Arabic-derived words presented in this study.

## 1.2 The online version of the *OED* as a source of Arabic borrowings

The findings included in this article result from a close review of the linguistic data provided by the *Oxford English Dictionary Online* (henceforth

---

<sup>2</sup> See fn. 1.

*OED*). The *OED* is currently being subjected to its first complete “overhaul”. The digitized form of the *OED*, containing the complete text of the Second Edition from 1989 (*OED2*), the 1993 and 1997 *OED Additions Series*, and an essential number of updated and new entries which belong to the Third Edition, or *OED3*, can be searched online. The results of the revision work, including revised and new dictionary entries, are being added to the electronic *OED Online* every quarter<sup>3</sup>. The *OED Online* permits an exhaustive count and account of all the words and meanings which have their origins in Arabic<sup>4</sup>. The following search makes it possible to retrieve all the different nineteenth- and twentieth-century Arabic borrowings from the electronic *OED Online: Advanced Search: Entries containing “Arabic” in “Etymology” and “1801-” in “Date of Entry”*. The corpus data on which the present article is based was identified in the *OED Online* in the summer of 2016. At that time, the *OED* did not include any lexical item assumed to have come from Arabic in the twenty-first century. The number of nineteenth- and twentieth-century Arabic borrowings listed in the *OED Online* totals 302 lexical items. The sample of words comprises borrowings adopted from Standard Arabic, some borrowings from regional or dialectal Arabic, and several from additional national varieties of Arabic. *Riqq*, for example, ‘[a] musical instrument similar to a tambourine, used principally in Arab countries’ (*OED3*), was taken over from Egyptian Arabic into English in 1836. Furthermore, the sample contains borrowings which show a ‘mixed’ etymology in the *OED*, i.e. lexical items which were partly adopted from Arabic and partly from another language. This holds for the twentieth-century borrowing *ouguiya*, for instance, which refers to ‘[t]he principal monetary unit in Mauritania, equivalent to 5 francs of the West African Monetary Union’ (*OED3*). Details in the *OED3* make it clear that the word was derived from both the French form *ouguiya* and its Mauritanian Arabic etymon *ūgiyya*.

In this study, all of the borrowings analysed are assumed to have been borrowed directly from some dialect of Arabic regardless of their meanings and whether they had been borrowing into Arabic from some other language historically. The word *tajine*, for instance, first attested in 1898 in English in the sense of ‘[a]ny of various types of North African (orig. Moroccan) stew prepared by slowly cooking the ingredients in a shallow, earthenware cooking dish [...]’ (*OED3*), was regarded as a borrowing from (Moroccan)

<sup>3</sup> For more information concerning the revision of the *OED*, see Durkin (1999: 1-49).

<sup>4</sup> For detailed information about the production and the advantages of the electronic form of the *OED*, see Brewer (2004) and Brewer (2007: 213-257).

Arabic, despite the fact that it ultimately goes back to ancient Greek *τάγηνον* 'frying pan' (see *OED3*).

Before we consider the various subject fields of borrowings which have been taken from Arabic into English during the nineteenth and twentieth centuries, the classification of loan influences presented in this paper requires some terminological clarification.

### 1.3 Classification of loan influences

The classification undertaken in the present analysis is based on the categorization provided by Carstensen in 1968 for researching Anglicisms in German. Carstensen's model considers the most important categories of lexical borrowing and is applicable to other language contact situations including that between English and Arabic. I shall confine myself to the different categories of lexical borrowing which could be identified among the Arabic borrowings investigated in this paper. The following terms will be used: borrowing, transliteration, adaptation, and hybrid.

#### 1.3.1 Borrowing

*Borrowing* functions as the usual term for a word or a meaning which was taken over from another language. It can also refer to an affix, combining form, phrase or even speech sound taken in as well as to the process by which a word or a meaning is borrowed from another language.

#### 1.3.2 Transliteration

Most of the words borrowed from Arabic since 1801 have undergone *transliteration*. The term refers to the process by which the characters or letters of a foreign alphabet are translated into the alphabet of another language, or to the lexical item rendered in this way. The majority of transliterated Arabic borrowings presented in this article approximate or correspond to the pronunciation of their originals in the source language. An example is *niqab*, '[a] veil worn by some Muslim women, covering all of the face and having two holes for the eyes' (*OED3*), which reflects Arabic ب ا ق ن. According to the *OED2*, it was borrowed into English in 1936.

#### 1.3.3 Adaptation

*Adaptation* specifies the assimilation by which a word from a foreign language becomes an integrated unit of the language borrowing it. The term can also

relate to the naturalized lexical item itself. The borrowing *ghazeeyeh*, which denotes '[a]n Egyptian dancing girl' (*OED2*), may serve as an example. The *OED2* makes it clear that the word was adapted from *ġāziya*, which represents the Romanized form of its Arabic source term.

### 1.3.4 Hybrid

The term *hybrid* is used to designate a word which comprises both foreign and native components. The Arabic borrowing *Salafist* serves as an example. It entered English in 1974. The item was formed from the Arabic *salafī* 'Salafī', to which the English suffix *-ist* was attached.

As already mentioned, the majority of words borrowed from Arabic into English since 1801 were subjected to a transliteration process arising out of the non-Roman spelling system used by Arabic. Some Arabic borrowings show more than one orthographic form in the receiving language. Cannon (1997: 177) rightly draws attention to the fact that particularly 'recent transfers [from Arabic] have had little opportunity for a "standard" spelling to develop.' *Felafel*, designating a type of dish, may be adduced as an example. The borrowing is first recorded in the *OED2* in 1951. Two additional spelling variants exist in English which reflect vowel alternation of the word: *falafel*, the more frequent spelling, and *filafil*, which can also be used in English. All the three forms correspond to the Arabic *falāfil* (see *OED2*).

It also seems noteworthy that most of the borrowings presented in this paper show only one specific meaning, such as *tabl*, a designation of a type of drum. The word goes back to the synonymous Arabic *ṭabl*. Arabic-derived words which have developed additional senses due to semantic change are in the minority. *Hijab* might serve as an example. It was first attested as a technical term from Islam in 1885 in English, referring to '[t]he practice observed by some Muslim women of wearing concealing clothing (esp[ecially] headgear, or (in early use) living in seclusion' (*OED3*), as in:

- (1) 1885 T.P. Hughes *Dict. Islam* 174/1 *Hijab*, a term used for the seclusion of women enjoined in the Qur'an. (*OED3*)

From the linguistic evidence included in the *OED3* it emerges that the word came to specify '[a] head covering or veil worn in public by some Muslim women.' According to the *OED3*, this use, which now represents the common meaning of the borrowing, first occurred in 1980 in English:

- (2) 1980 *Associated Press Newswire* (Nexis) 15 July She said the wearing of the hijab, or veil, is a matter of choice.

The word is derived from the Arabic *hijab*, which literally translates as 'veiling'.

## 2. Subject fields influenced by Arabic during the nineteenth and twentieth centuries

As already outlined, the *OED Online* lists 302 borrowings which were taken over from Arabic in the nineteenth and twentieth centuries. The words under consideration were divided into seven major subject areas and their related domains, in order to provide insight into the various semantic fields influenced by Arabic during the last two centuries. The assignment of the Arabic-derived words and meanings relies on their classification in the *OED*.

The following list gives an overview of the numbers and percentages of borrowings in ascending order in the different subject areas enriched by Arabic during the nineteenth and twentieth centuries. For each field, a number of illustrative examples of Arabic borrowings are given:

1. **The Fine Arts and Crafts (9 borrowings, i.e. 3.0%)**
  - 1.1. Literature (1 borrowing, i.e. 0.3%), e.g. *ghazal*, n. (1801).
  - 1.2. Music (8 borrowings, i.e. 2.6%), e.g. *tabl*, n. (1831); *riqq*, n. (1836); *kissar*, n. (1864); *gimbri*, n. (1876).
2. **Language and Linguistics (13 borrowings, i.e. 4.3%),** e.g. *Sahidic*, adj. (1830); *mimaton*, n. (1873); *shadda*, n. (1896); *Lihyanic*, n. (1911).
3. **Gastronomy (24 borrowings, i.e. 7.9%)**
  - 3.1. Kitchen Utensil (1 borrowing, i.e. 0.3%), e.g. *zarf/zurf*, n. (1836).
  - 3.2. Drink and Tobacco (6 borrowings, i.e. 2.0%), e.g. *shisha*, n. (1832); *merissa*, n. (1844).
  - 3.3. Cookery (17 borrowings, i.e. 5.6%), e.g. *kibbeh*, n. (1829); *halawi*, n. (1836); *tajine*, n. (1898); *baba ganoush*, n. phr.<sup>5</sup> (1938); *felaful/falafel*, n. (1951); *tabbouleh*, n. (1955).
4. **Civilization and Politics (38 borrowings, i.e. 12.6%)**
  - 4.1. Government, Administration and Politics (18 borrowings, i.e. 6.0%), e.g. *wali*, n. (1811); *mudir*, n. (1844); *Mendoub*, n. (1923); *umma*, n. (assuming a meaning from politics in 1946); *Amal*, n. (1979).

<sup>5</sup> The grammatical terminology employed in this paper is based on Quirk et al. (2008).



- 4.2. Terrorism, War and the Military (20 borrowings, i.e. 6.6%), e.g. *redif*, n. (1836); *jihad*, n. (1869); *intifada*, n. (1985); *Hamas*, n. (1988); *al Qaeda*, n. phr. (1996).
5. **The Natural Sciences (43 borrowings, i.e. 14.2%)**
  - 5.1. Astronomy (1 borrowing, i.e. 0.3%), e.g. *Deneb*, n. (1867).
  - 5.2. Medicine (4 borrowings, i.e. 1.3%), e.g. *argel*, n. (1803); *Yunani*, adj. 1922).
  - 5.3. Zoology (6 borrowings, i.e. 2.0%), e.g. *mhorrr*, n. (1825); *ariel*, n. (1832); *raad*, n. (1858).
  - 5.4. Geography and Geology (14 borrowings, i.e. 4.6%), e.g. *wadi/wady*, n. (1839); *hammada*, n. (1853); *seif*, n. (1925).
  - 5.5. Botany (18 borrowings, i.e. 6.0%), e.g. *argan*, n. (1809); *tarfa*, n. (1858); *sim-sim*, n. (1917).
6. **Islam and Religion (79 borrowings, i.e. 26.2%)**, e.g. *Eid-al-Fitr*, n. (1823); *to halal*, v. (1855); *Eid-al-Adha*, n. (1880); *madhhab*, n. (1929); *takaful*, n. (1953); *Salafist*, adj. (1974).
7. **People and Everyday Life (79 borrowings, i.e. 26.2%)**
  - 7.1. Animal Rearing (2 borrowings, i.e. 0.7%), e.g. *saluki*, n. (1809).
  - 7.2. Tourism (3 borrowings, i.e. 1.0%), e.g. *mandarah*, n. (1836); *mudhif*, n. (1888).
  - 7.3. Monetary Units (5 borrowings, i.e. 1.7%), e.g. *fiis*, n. (1891); *halala*, n. (1961); *ouguiya*, n. (1973).
  - 7.4. Security and Police (6 borrowings, i.e. 2.0%), e.g. *razzia*<sup>6</sup>, n. (1821); *ghaffir*, n. (1831); *mutawwa*, n. (about 1977).
  - 7.5. Transport and Travelling (7 borrowings, i.e. 2.3%), e.g. *dahabeeyah/dahabiah*, n. (1846); *mahaila*, n. (1904).
  - 7.6. Communication (9 borrowings, i.e. 3.0%), e.g. *quaiss kitir*, int. (1898); *malesh*, int. (1913); *iggri/iggry*, int. (1917).
  - 7.7. Clothing (16 borrowings, i.e. 5.3%), e.g. *habara*, n. (about 1817); *'sherwal*, n. (1844); *gandoura*, n. (1851); *niqab*, n. (1936); *hijab*, n. showing a meaning from clothing in 1980).
  - 7.8. Society, Human Behaviour and Feelings (31 borrowings, i.e. 10.3%), e.g. *kef/keif/kief*, n. (1808); *ghazeeyeh*, n. (1819); *bint*, n. (1855); *faki*, n. 1872); *Saudi*, n. (1933); *Sahrawi*, n. (1976).

<sup>6</sup> According to the *Oxford Dictionary of English Etymology* (1966), *razzia* shows a mixed etymology. It was partly borrowed from French *razzia* and partly from its Algerian Arabic etymon *gāziya* (see also the OED3 for additional etymological details).



The sample of nineteenth- and twentieth-century borrowings from Arabic includes 17 lexical items which are not easily categorized. Examples are *tahalli* and *hawala*. Of these, *tahalli*, referring to '[d]ecoration', was borrowed into English in 1833 (*OED2*). It goes back to Arabic *taḥallī* 'ornamenting'. *Torba* represents a twentieth century borrowing from Arabic. The word specifies a type of cement in English. As stated in the *OED2*, it was adapted from the Arabic *turba* in the sense of 'dust, earth, soil' in 1910.

## 2.1 The fine arts and crafts

"The fine arts and crafts" constitutes the field with the smallest number of Arabic-derived items. It includes nine borrowings which have to do with literature and music. *Ghazal*, for instance, was adopted into English in 1801 as a term for '[a] species of Oriental lyric poetry, generally of an erotic nature, distinguished from other forms of Eastern verse by having a limited number of stanzas and by the recurrence of the same rhyme' (*OED2*). According to the *OED2*, the word corresponds to both Persian and Arabic *ghazal*. Another example is the borrowing *gimbri*, '[a] small Moorish guitar played by plucking the strings with a piece of dry palmetto leaf' (*OED2*), as is exemplified by the following *OED2* quotation:

- (3) 1903 *Westm. Gaz.* 18 Feb. 2/3 The slender fanatical singer, Whose fingers were skilled on the gimbri.

The borrowing can also denote 'a player of this instrument' (*OED2*). *Gimbri* is an adaptation of the Arabic *gunbrī* (*OED2*). Like most of the borrowings in this area, *gimbri* is not confined to specific contexts in the receiving language. Yet it comparatively often occurs in Arabic-speaking contexts in English, as in the following example, which is taken from a novel revealing a Moroccan locale:

- (4) 1907 F. Campbell *Shepherd of Stars* xv. 177 The Gimbri wagged his [head] from side to side. (*OED2*)

It seems noteworthy that the field of the fine arts and crafts does not encompass any word or meaning which was introduced from Arabic during the twentieth century. The afore-mentioned *gimbri* is the latest acquisition of this field documented in the *OED*: it was introduced into English in 1876.

## 2.2 Language and linguistics

Similarly, the domain of language and linguistics includes a relatively small number of adopted words and senses. An example is the nineteenth-century borrowing *mimation* and its spelling variant *mimination*, a technical term in Semitic grammar for '[t]he appending of *m* to the flexional vowels in Semitic languages, esp[ecially] Akkadian' (*OED3*). Its first attested use dates from 1873 (*OED3*):

- (5) 1873 *Eng. Cycl. Arts & Sci. Suppl.* 173 The use of mimination by the Babylonians.

The *OED3* notes that the borrowing is a hybrid which was formed from the Arabic *mīm*, the designation of the Arabic letter which reflects the Roman *m*, and the English suffix *-ation*. A further Arabic borrowing from the field of language and linguistics is *Lihyanic*, '[t]he name of an ancient Semitic language known only from north Arabian inscriptions of the 2<sup>nd</sup> and 1<sup>st</sup> centuries b.c.' (*OED2*). It is first documented in English in 1911:

- (6) 1911 *Encycl. Brit.* XXIV.626/1 A more commendable proposal is to call the inscriptions Lihyānī, since the tribe of Lihyān is sometimes mentioned in them... Other brief inscriptions ... have been discovered... Their writing is a somewhat later form of the Lihyānī, and the dialect ... seems to be very similar to Lihyānī. (*OED2*)

The word is derived from Arabic *lihyān* to which the suffix *-ic* was attached. The majority of Arabic borrowings in this domain are documented in specialized literature about language and linguistics.

## 2.3 Gastronomy

Arabic has served as the donor language of 24 borrowings in the area of gastronomy, encompassing words and meanings to do with "kitchen utensils", "drink and tobacco", and "cooking". *Zarf/zurf*, *merissa*, *shisha*, *fela fel/falafel* and *baba ganoush* serve as examples. *Zarf* and its spelling variant *zurf* relate to '[a] cup-shaped holder for a hot coffee-cup, used in the Levant, usually of metal and of ornamental design' (*OED2*), *merissa* is the name of a type of beer produced in Sudan, and *shisha* refers to a type of water-pipe or the tobacco which is smoked with this pipe. A close review of the *OED*

quotations shows that *shisha* occurs in English mainly in Arabic-speaking contexts. The following *OED3* citation from 2003, for instance, describes an Egyptian setting:

- (7) 2003 *enRoute* Mar. 26/2 At night, the cafés along the Corniche, like cafés everywhere else in Egypt at night, are full of men drinking tea, smoking *shisha* and playing backgammon.

*Shisha* was adapted from the Egyptian Arabic *šīša* ‘hookah’, ‘tobacco which can be smoked with a hookah’. The Arabic source term ultimately goes back either to Persian *šīša* ‘glass’ or to Turkish *şişe* ‘glass’ (see *OED3*). *Felafel/falafel* and *baba ganoush* specify dishes. The former has become a widespread term for a dish which includes peppers and small pea balls, usually served in a flat bread roll, and the latter is used to denote ‘[a] Middle Eastern (originally Lebanese) dish of puréed roasted aubergine, garlic, and tahini’ (*OED3*). The various Arabic-derived culinary terms under review are fairly often found in specialized literature on food and cooking, as can be seen from the following usage examples of *baba ganoush*, *tajine*, a variety of stew originating in Morocco, and *kibbeh*, a Middle Eastern dish typically prepared in Lebanon and Syria:

- (8) 1986 J. Ridgwell *Middle Eastern Cooking* 28 (*heading*) Baba Ghanoush ... Aubergine and tahini purée. (*OED3*)
- (9) 1990 *Gourmet* Sept. 165/2 The spicy, hearty meat stews cooked in clay pots called *tajines* are probably Bedouin in origin. (*OED3*)
- (10) 2003 E. Powell tr. S. Jamal *Arabian Flavours* 37 One of the most sought-after and flaunted virtues in the eyes of the family of a girl of marriageable age is her dexterity in preparing *kibbeh*. (*OED3*)

## 2.4 Civilization and politics

Arabic has also served as a source of words in the domain of “civilization and politics”, including borrowings related to “government, administration and politics”, “terrorism, and war and the military”. *Umma*, for instance, shows a political meaning some time after its first attested use in English. Since 1885 the word has been documented in English with reference to ‘[t]he Islamic community, founded by Muhammad at Medina, comprising individuals bound to one another by religious ties on a tribal model’ (*OED2*). *Umma*

assumed an additional sense in English in 1946, serving as '[t]he name of a nationalist political party founded in the Sudan in 1945' (OED2). In this use, the word is spelled with a capital initial letter, as in:

- (11) 1981 *Economist* 24 Jan. 43/2 There are parochial or communal parties which do not favour or are fearful of, absorption into larger units: these include the Christian Phalange in Lebanon, the Umma in Sudan, the Neo-Destour in Tunisia. (OED2)

*Umma* was borrowed from the Arabic word '*umma* 'nation, people, community' (see OED2). A further example is *Mendoub*, which was adopted from the Arabic *mandūb* in 1923, denoting '[t]he Sultan of Morocco's representative in the administration of the international city of Tangier, with responsibility for presiding over the legislative assembly' (OED3). Since Tangier was under the jurisdiction of Britain, Spain, France and a number of other countries from 1923 to 1956, *Mendoub*s are no longer involved in the international administration of that city. Hence, it is now confined to historical contexts in present-day English, as can be seen from an OED3 citation of 1999 which records the spelling variant *mandub*:

- (12) 1999 *Britannica Online* (Version 99.1) at Morocco, Tangier, though it had a Spanish-speaking population of 40,000, received a special international administration under a *mandub*. (OED3)

The majority of the borrowings from the field of "civilization and politics" are associated with terrorism, and war and the military. *Jihad*, a term occurring commonly in Present Day English, may be adduced as an example. It was borrowed from Arabic in the sense of '[a] religious war of Muslims against unbelievers, inculcated as a duty by the Qur'an and traditions' (OED2) in 1869. *The Oxford Dictionary of English Etymology* (1966) also documents *jihad* as a spelling variant of this borrowing. The word widened its semantic scope a few years after it entered English. It has been documented since 1880 in an extended use, referring to '[a] war or crusade for or against some doctrine, opinion, or principle' or 'war to the death' (OED2). The domain of terrorism, and war and the military also contains the latest borrowing from Arabic listed in the OED. This is *al Qaeda*, '[a] network of militant Islamic groups active internationally since the 1990s, and [...] associated with numerous attacks on military and civilian targets across the world' (OED3). The phrase frequently occurs in attributive function in English, as in *al Qaeda cell*, *al Qaeda link* and *al Qaeda terrorist* (see OED3).

## 2.5 The natural sciences

The natural sciences with their various related areas constitutes the second largest domain influenced by Arabic in the nineteenth and twentieth centuries. In all, 43 borrowings can be assigned to this field. The greatest proportion of lexical items belongs to geography, geology and botany. A number of the natural science terms under review are specialized expressions which have not made it into everyday use. Examples are *mhorr*, a borrowing from Moroccan Arabic which refers to a variety of gazelle native to the Sahara, and *sim-sim*, a synonym for 'sesame'. In addition, there is *seif*, a technical term in physical geography for '[a] sand dune having the form of a narrow ridge elongated in a direction parallel to that of the prevailing wind' (*OED2*). These types of borrowings are chiefly documented in scientific texts and encyclopedias. A perusal of the documentary evidence available in the *OED* reveals that *mhorr* and *seif* also function as adjectivals in English, as in:

- (13) 1974 *Encycl. Brit. Micropædia* IV. 443/3 Coat ... reddish brown with white rump, underparts, and head and a white spot on neck in the western races, such as the critically endangered *mhorr* gazelle. (*OED3*)
- (14) 1975 *Nature* 20 Feb. 617/2 Until the recent drought, *seif* dunes were mainly active in this region north of the 150-mm isohyet. (*OED2*)

*Seif* was adopted from the Arabic noun *saif*, which literally means 'sword'. *Wadi* is one of the few Arabic borrowings from the natural sciences with which the average speaker of English would be familiar. It corresponds to the Arabic *wādī*.

## 2.6 Religion and Islam

"Religion and Islam" and "people and everyday" life constitute the two fields where the impact of Arabic was strongest in the nineteenth and twentieth centuries. Both categories include 79 borrowings. *Eid-al-Fitr* and *takaful* constitute two representative borrowings associated with religion and Islam. Of these, *Eid-al-Fitr*, which came into English in 1823, is the designation of '[t]he Feast of Breaking the Fast, a major Islamic festival celebrating the end of Ramadan, the month of fasting, and commencing on the first day of the tenth month of the Muslim calendar' (*OED3*), e.g.

- (15) 2004 *New Yorker* 13 Dec. 52/2 The clientele that morning was mainly veiled wives in black burkas and bearded men wearing shalwar kameez,... who were shopping for Eid-al-Fitr – the sighting of the new moon that marks the end of Ramadan. (OED3)

The word is given quite a complex etymological description in the OED3: it was partly influenced by Persian *ʿīd-i fiṭr* and its Arabic etymon *ʿīd al-fiṭr*, ‘Feast of Breaking the Fast’. *Takaful* is classified as a technical term from Islam (OED3), specifying ‘[t]he principle of mutual obligation’ or, more specifically, ‘a form of mutual insurance, compliant with Sharia law, in which subscribers contribute to a fund, sharing losses and liabilities and also any profits which may accrue’ (OED3). The borrowing corresponds to the Arabic *takāful*, which literally means ‘mutual obligation’.

## 2.7 People and everyday life

Like the area of “religion and Islam”, that of “people and everyday life” takes in 79 items. These fall into eight subgroups, consisting of borrowings relating to “animal rearing” (e.g. *saluki*, the name of a breed of dog), “tourism”, for instance *mudhif*, which designates ‘a guest house’ or ‘a reception room’ (OED3) in Iraq, “monetary units” (e.g. *halala*, specifying a unit of currency which is used in Saudi Arabia), “security and police”, (e.g. *ghaffir*, ‘[a] native Egyptian policeman; a guardian, watchman’ (OED2)), and several terms associated with “transport and travelling”, such as *dahabeeyah* and its spelling variant *dahabiah*, denoting ‘[a] large sailing-boat, used by travellers on the Nile’ (OED2). In addition, it contains nine borrowings having to do with “communication”, among them a number of interjections which may render a speech or a conversation more vivid and expressive. An example is *malesh*, which entered English in 1913 in the sense of “[n]o matter!”, ‘never mind!’ (OED3). The interjection is quite often put into the mouth of characters who speak English but whose first language is Arabic in order to increase the authenticity of a scene described or an utterance. This is illustrated by an extract from the novel *Baghdad without Map*, which reveals a Middle Eastern locale:

- (16) 1991 T. Horwitz *Baghdad without Map* i. 10 ‘*Malesh*’ Ahmed said, producing a second oar. *Malesh* is an Egyptian phrase of surrender, meaning ‘never mind’ or ‘doesn’t matter’. (OED3)

*Malesh* is derived from the colloquial Arabic form *ma ‘alay-š* ‘no matter’ (see OED3). This group of OED entries also lists sixteen terms for clothing. An

example is the borrowing *habara*, a type of silk garment worn by women. It goes back to Arabic *ḥabara*. Evidently, this word, like *seif*, can function adjectivally in English, as in *habara veil* (OED2). Of the borrowings in the category “people and everyday life”, a substantial number have to do with “society, and human behaviour and feelings”. The majority of these types of words, such as *faki*, ‘[a] title given in Africa to schoolmasters’ (OED2) and *Saudi*, which represents a common term for a person from Saudi Arabia, relate to the individuals who make up a society.

### 3. Conclusion

This study constitutes an *au courant* investigation of a significant part of the foreign vocabulary adopted into English during the last two centuries: the sum of nineteenth- and twentieth-century Arabic borrowings documented in the *OED Online*. As mentioned above, even the latest Arabic-derived word, *al Qaeda*, which entered the language in 1996 according to the *OED*, is considered here.

The documentary evidence provided by the electronic *OED* has made it possible to give an overview of the manifold domains and spheres of life from which Arabic words have been adopted into the English language since 1801. As regards the proportion of borrowings in the various subject areas, the data indicate that the greatest donations have been made in the fields of “religion and Islam” and “people and everyday life”. Arabic has long served as an important donor language of lexical items, among them a substantial number of words from Islam. Cannon – VanDuinkerken (2008: 36-37) point out that

Arabic is an international language long known as a major supplier of words to Swahili, Romance languages, Persian, Turkish, Indian languages like Hindi and Urdu, and English. A tabulation of such contact data worldwide would place Arabic among the major word-suppliers to English. [...] Islam has had continuing, modern influences, with items going into Southeast Asia and the Philippines, among other distant areas. In conveying Islam to the world, Arabic surpassed Sanskrit as an international carrier of religion.

A substantial number of the fairly recent borrowings related to “religion and Islam” are specialized terms that the ordinary proficient speaker of English



would not commonly encounter. An example is *madhhab*, taken over from Arabic in 1929 as a technical term in Islam for '[e]ach of the schools of Sunni Islamic law, each of which is based on a particular system of interpretation of Islamic religious and legal texts' (*OED3*). The number of Arabic-derived words in the domain of people and everyday life, comprising eight subareas, such as animal rearing, transport and travelling, and communication and clothing might be explained, to some extent at least, by the global presence and impact of Arabia and Muslim culture. The status of Arabic as an influential language world-wide might have led to the adoption of manifold terms reflecting objects, concepts and ideas common in the Arabic world into English. Cannon – VanDuinkerken (2008: 37) draw attention to the fact that

Twentieth-century petrol wealth has given the Arabs political and strategic importance, with international events continually involving Arabia and the Muslim world. The impact of Islam has carried Arabic into the bahasas of Malaysia and Indonesia, where national contests determine the muezzin who can best call the faithful to prayer in Arabic. The war between Iraq and Iran and later Middle Eastern intercourse have had global impact.

The borrowings of the natural sciences, with their various related fields, make up the second largest category of lexical items taken over from Arabic in the nineteenth and twentieth centuries. The adoption of natural science terms has likely resulted from the advances made in these domains in Arabic-speaking countries and the consequent coinage of Arabic terms for such concepts, terms which spread with those advances to English-speaking nations.

As for the individual subgroups of the various major subject areas, the highest number of Arabic borrowings are included in the domains of "society, human behaviour and feelings" (31 borrowings), "terrorism, war and the military" (20 borrowings), "government, administration and politics" (18 borrowings), "botany" (18 borrowings) and "cooking" (17 borrowings). These findings illustrate that Arabic not only contributed to the English lexicon in the form of specialized terms restricted to scientific or technical contexts but also in the form of a variety of borrowings having to do with everyday life.

The present article serves as a prolegomenon to further study. Considering the highly political, sociological, and historical implications Arabic shows in the Western world, it would be worth delving into further research on the degree of borrowing in different English varieties, such as



British versus American English. It would also be interesting to provide a rounded picture of the various historical, political, and cultural reasons which may have led to the introduction of Arabic words into English. However, such a broad analysis would have to be carried out within a framework that is far beyond the scope of this paper.

## REFERENCES

### Sources

- Ayto, J.  
 1989 *The Longman Register of New Words*. Harlow: Longman.
- Ayto, J.  
 1990 *The Longman Register of New Words*. Vol. 2, Harlow: Longman.
- Bliss, A.  
 1966 *A Dictionary of Foreign Words and Phrases in Current English*. London and New York: Routledge.
- Onions, C.T. – R.W. Burchfield – G.W.S. Friedrichsen. (eds.)  
 1966 *The Oxford Dictionary of English Etymology*. Oxford: Clarendon.
- Oxford English Dictionary Online*  
<http://www.oed.com/> (last accessed May 2017).
- Tulloch, S.  
 1991 *The Oxford Dictionary of New Words*. Oxford: Oxford University Press.

### Special studies

- Brewer, C.  
 2004 "The 'electronification' of the *Oxford English Dictionary*", *Dictionaries* 25, 1-43.
- Brewer, C.  
 2007 *Treasure-house of the Language: the Living OED*. New Haven and London: Yale University Press.
- Cannon, G.  
 1997 "90 post-1949 Arabic loans in written English", *Word* 48, 171-194.
- Cannon, G. – A.S. Kaye (eds.)  
 1994 *The Arabic Contributions to the English Language: An Historical Dictionary*. New Haven and Wiesbaden: Harrassowitz.
- Cannon, G. – W. VanDuinkerken  
 2008 "Arabic lexical contributions to the English language", *Journal of the Dictionary Society of North America* 29, 36-61.

Carstensen, B.

- 1968 "Zur Systematik und Terminologie deutsch-englischer Lehnbeziehungen". In: H. Brekle – L. Lipka (eds.) *Wortbildung, Syntax und Morphologie. Festschrift zum 60. Geburtstag von Hans Marchand am 1. Oktober 1967*. The Hague: Mouton, 32-45.

Durkin, P.

- 1999 "Root and Branch: Revising the etymological component of the OED", *Transactions of the Philological Society* 97, 1-49.

Durkin, P.

- 2014 *Borrowed Words: A History of Loanwords in English*. Oxford: Oxford University Press.

Kaye, A.S.

- 1992 "Orthographic variation in Arabic loan words", *English Today* 8, 32-41.

Quirk, R. et al.

- 2008 [1985] *A Comprehensive Grammar of the English Language*. Harlow: Longman.

Address: JULIA SCHULTZ, University of Heidelberg, English Department, Kettengasse 12, 69117 Heidelberg, Germany.

ORCID code: [orcid.org/0000-0002-6402-7670](https://orcid.org/0000-0002-6402-7670).

# ***Maybe*: Development and Topic Marking**

Daisuke Suzuki

*Setsunan University*

## ABSTRACT

This paper discusses the development of *maybe* and related expressions (i.e., *it may be*, *mayhap*) in the history of English. I provide a quantitative analysis of their long-term histories by drawing on the *OED* and its quotations database, along with data from two different datasets, namely, the *Corpus of Late Modern English Texts*, version 3.0, and the Brown family of corpora. After extracting their instances from the datasets, this study analyzes the data to determine the position in which *maybe* is used the most within a clause. The results of the analysis indicate that both the nineteenth and twentieth centuries were crucial periods for the development of *maybe*. In addition, I further demonstrate that the target adverb, over time, began exhibiting features of a topic marker, and that the development has also interwoven with the process of subjectification.

Keywords: English modal adverbs, *OED*, corpus data, historical analysis, topic/theme, subjectification.

## **1. Introduction**

The historical corpora of English originate in the *Helsinki Corpus*, and more and more corpora have now become available for historical research. Due to the limited number of words in these corpora, however, they cannot yield sufficient instances of some target expressions; particularly, they lack low frequency items for linguistics research. In addition, few corpora (corpus series) span more than a thousand years of English usage. This study, therefore, explores an alternative way of using an even larger database for providing big data through the analysis of modal adverbs as case studies.

This study concerns the histories of the modal adverb *maybe* and related expressions, and takes a fresh look at the change in *maybe* in the history of

English. In comparison to other modal adverbs formed by the combination of two elements (morphemes), the history of *maybe* is less well documented (e.g., on *indeed*, Traugott – Dasher 2002; on *no doubt*, Simon-Vandenberg – Aijmer 2007; Davidse – De Wolf – Van Linden 2015; on *of course*, Lenker 2010). In present-day English, *maybe* is used to mark epistemic possibility, as in (1):

- (1) *Maybe* I'm right and *maybe* I'm wrong. (Swan 2005: 17)

In the history of English, modal adverbs assumed their forms during the Middle English period; in contrast, they did not have “their present-day epistemic meanings” at that time (Hanson 1987: 137). In a similar vein, according to the *Oxford Dictionary of Word Histories*, *maybe* is first evinced in the Late Middle English period, and it derives from *it may be (that)*. Additionally, based on Terasawa (1997), Table 1 indicates the first attestation and epistemic use of typical modal adverbs, including *maybe*:

Table 1. The development of the main modal adverbs (from Terasawa 1997)

Modal adverbs	First appearance in English	First epistemic usage
<i>certainly</i>	c.1300	c.1303
<i>surely</i>	?c.1300	?c.1300
<i>maybe</i>	a.1325	a.1325
<i>possibly</i>	1391	1600
<i>probably</i>	c.1535	1613

Although previous studies agree on the origin and first appearance of *maybe*, no comprehensive analysis of the semantic shift is offered. Therefore, a detailed description derived from historical data is the starting point of this study, followed by a theoretical discussion of the general linguistic change.

## 2. Methodology

The main data source of the present study is retrieved from the quotations database of the *Oxford English Dictionary* (OED), based on the CD-ROM version of the second edition (Version 4.0) (cf. Berg 1991). The use of the OED quotations database as a “corpus” is widely discussed in the literature,

including Brewer (2000), Mair (2001, 2004), Hoffmann (2004), Iyeiri (2010), and Rohdenburg (2013). The database is not a balanced or representative one, and the number of quotations it comprises fluctuates significantly throughout the different periods. However, the dataset has certain advantages that offset these drawbacks. The database contains over 2.4 million quotations, spanning more than a thousand years of English usage. This amounts to “a total of 33-35 million words” (Hoffmann 2004: 25), i.e. a sizable body of actual usage in English over a considerable period of time, which is indispensable for a linguistic analysis.

In extracting the data concerning *maybe* and related expressions from the *OED* quotations database, I proceeded as follows. First, I manually extracted all occurrences matching the target expressions, including the following spelling variants, from the search results<sup>1</sup>:

Table 2. Occurrences of the target expressions in the *OED*

Expressions	Variants	Total
<i>maybe</i>	<i>mebbe</i> (31), <i>mebby</i> (5), <i>maybees</i> (1), <i>may be</i> (53), <i>maybe</i> (542)	632
<i>mayhap</i>	<i>mayhap</i> (19), <i>mehap</i> (1), <i>mayhaps</i> (2)	22

I then identified the quotations in which the target expressions serve as sentence adverbs, because this study focuses on the epistemic function in a sentence<sup>2</sup>. This process was completed also in manual jobs, and the results are organized in Table 3.

Table 3. Instances of the target adverbials in the *OED*

Adverbial	<i>OED</i>
<i>it may be</i>	55
<i>mayhap</i>	19
<i>maybe</i>	403

<sup>1</sup> The *OED* and other etymological dictionaries suggest *may-chance*, *may-fall*, *may-fortune*, *mayhap*, and *may-tide* as the competing forms of *maybe*, which have all become obsolete; however, no instances of their forms, excluding *mayhap*, were obtained from the *OED*. With regard to the spelling variants of *maybe*, the reviewer of the journal has highlighted other forms such as *maybe*, *mabbei*, *mabby*, and *maybi(e)*, but the search of the *OED* makes no matches of them.

<sup>2</sup> For this analysis, I excluded all examples of one-word responses, such as “*Mebbe.*” or “*Maybe.*” and of the nominal use, as in (i). Also excluded from this analysis were examples that did not form a complete clause, as in (ii). I further excluded any examples in which the target expressions appeared within the phrase structure, as in (iii): (i) Without all *Maybees*, the Lord is never more gracious to his Servants. (1615 *Day Festivals* xii. 335, *OED*), (ii) *Maybe* scenario writing eventually. (1928 H. Crane *Let.* 27 Mar. (1965) 321, *OED*), (iii) Barmaids do the work, with *maybe* a barman, potman or cellarman to help. (1936 Mencken *Amer. Lang.* (ed. 4) 243, *OED*).

The other different sets of corpus databases I used to complement the above data are the *Corpus of Late Modern English Texts*, version 3.0 (CLMET3.0) and the Brown family of corpora. CLMET3.0 includes 34 million words of running text from the Late Modern English period, covering five major genres: narrative fiction, narrative non-fiction, drama, letters, and treatise. The corpus design categorizes all texts into the following three sub-periods, with each sub-period spanning 70 years: [1710-1780], [1780-1850], and [1850-1920]. The Brown family of corpora, for its part, spans one million words of running text from 1961 and 1991-2 and the two major varieties of English, American and British. They sample the following different registers:

Press (reportage); Press (editorial); Press (reviews); Religion; Skills, trades, and hobbies; Popular lore; Belles lettres, Biography and essays; Miscellaneous (government documents, foundation reports, industry reports, college catalogue, and industry house organ); Learned and scientific writings; General fiction; Mystery and detective fiction; Science fiction; Adventure and western fiction; Romance and love story; Humor

In this series, the American components in 1961 and 1992 are from the *Brown University Standard Corpus of Present-Day American English* (Brown) and the *Freiburg-Brown Corpus of American English* (Frown), respectively. The equivalents of British English in 1961 and 1991 are from the *Lancaster-Oslo/Bergen Corpus of British English* (LOB) and the *Freiburg-LOB Corpus of British English* (FLOB), respectively. Additionally, and even more importantly, these four sets were constructed under the same principles of corpus design and selection of texts, which facilitate comparability between two regions over a thirty-year period. The instances of *maybe* were obtained from CLMET3.0 and the Brown family of corpora along the same lines as in the *OED* dataset, identified in Tables 4 and 5, respectively:

Table 4. Instances of *maybe* in CLMET3.0

Adverb	CLMET3.0
<i>maybe</i>	315

Table 5. Instances of *maybe* in the Brown family of corpora

Adverb	Brown (1961)	Frown (1992)	LOB (1961)	FLOB (1991)
<i>maybe</i>	134	199	85	101

In addition to the frequencies of the target adverbials, I provided the factor interacting closely with their usage, which will be shown as significant in the following analysis. The factor is the clausal position, referring to the clause position (i.e., initial, medial, or final) in which the target adverbials occur. English permits modal adverbs to be placed initially, medially, or finally in actual use, and Quirk et al. (1985: 490-491) and Hoyer (1997: 148) present the different positions of modal adverbs as in (2a-g):

- |     |    |           |                  |   |
|-----|----|-----------|------------------|---|
| (2) | a  | <i>I</i>  | (initial)        | <i>Possibly</i> they may have been sent to London.  |
|     | b. | <i>iM</i> | (initial-medial) | They <i>possibly</i> may have been sent to London.  |
|     | c. | <i>M</i>  | (medial)         | They may <i>possibly</i> have been sent to London.  |
|     | d. | <i>mM</i> | (medial-medial)  | They may have <i>possibly</i> been sent to London.  |
|     | e. | <i>eM</i> | (end-medial)     | They may have been <i>possibly</i> sent to London.  |
|     | f. | <i>iE</i> | (initial-end)    | They may have been sent <i>possibly</i> to London.  |
|     | g. | <i>E</i>  | (end)            | They may have been sent to London <i>possibly</i> . |
- (Hoyer 1997: 148)

Despite the various approaches towards the positioning mentioned above, I followed Biber et al. (1999) in focusing on the three major categories of initial, medial, and final positions in a clause<sup>3</sup>, hand-coded this information into the above data, and determined the frequency and percentage of the target adverb in each position. This factor can be interpreted as an indicator of the speaker's or writer's perspective, and by comparing the shift of the factor across the database, I can further assess whether an individual change may indicate more general changes in grammar.

### 3. Results and discussion

#### 3.1. Historical overview

I provide an overview of the development of *maybe* and related expressions in the history of English in Table 6. Examples (3a-c) illustrate the use of the target adverbials. As shown in Table 6, *it may be* experienced a radical decline in the eighteenth century, and *maybe* in turn underwent a steady rise up to the present day; in particular, a remarkable increase can be observed from

<sup>3</sup> In other words, *I* in (2) indicates the initial position, *iM*, *M*, *mM*, and *eM* correspond to the medial position, and *iE* and *E* correspond to the final position.

the nineteenth century onwards. It is also shown that *mayhap* constitutes a very small part of the overall dataset, and it appears to be declining without being established as a functional marker of epistemicity.

Table 6. Diachronic breakdown of frequencies of the target adverbials

	-1600	1601-1700	1701-1800	1801-1900	1901-	Total
<i>it may be</i>	7	36	2	6	4	55
<i>maybe</i>	3	7	10	64	319	403
<i>mayhap</i>	0	2	4	12	1	19

- (3) a. Which, *it may be*, made the other to be the more virulently remembered. (1647 Clarendon *Hist. Reb.* II. §101, OED)  
 b. *Maybe* Mr. Chamberlain was remembering his pre-election promises. (1898 *Westm. Gaz.* 16 May 3/1, OED)  
 c. *Mayhap* she's hungry. (1840 Dickens *Barn. Rudge* Ixxii, OED)

Although *maybe* was introduced in the fourteenth century, it is evident from Table 6 that *maybe* had a limited number and was not prevalent. Its usage in Middle English can be seen in Example (4), which has been taken from the *Middle English Dictionary*:

- (4) Ther is manye of yow Faitours, and so *may be* that thow Art riht such on. (a1393 Gower *CA* (Frf 3) 1.174, MED)

This is a clear example of the onset of grammaticalization<sup>4</sup>. What must be borne in mind is that there are some parallel expressions in other languages such as French *peut-être* and Polish (*być*) *może*, and the presence of the expression in Old French, including Anglo-Norman, may have been associated with the rise of *maybe* in Middle English<sup>5</sup>. Here, I give special consideration to the periods of Late Modern English and present-day English, when the number of instances of *maybe* attests to its growing use.

The radical shift of *maybe* from the nineteenth century onwards is confirmed by the evidence from CLMET3.0. As shown in Table 7, the CLMET3.0 data also indicate that the frequency of *maybe* increased dramatically from the period [1780-1850] onwards.

<sup>4</sup> For details on the shift of *maybe* from a clause to an adverb, see López-Couso – Méndez-Naya (2016).

<sup>5</sup> I would like to thank one of the reviewers for alerting me to this point, which is clearly interesting and worth pursuing.



Table 7. Frequencies of *maybe* during the period 1710-1920 (CLMET3.0)

	1710-1780	1780-1850	1850-1920
<i>maybe</i>	2	71	242

The search results matching “maybe” in the period [1710-1780] yield many examples in which “maybe” forms the verbal phrase, as in (5a), and few examples are obtained in which *maybe* behaves as a sentence adverb, as in (5b).

- (5) a. The same thing *maybe* said of the taxes upon tea and sugar,  
(CLMET3\_0\_1\_51)  
b. And then your father, *maybe*, will stay. (CLMET3\_0\_1\_7)

In the next period [1780-1850], interestingly enough, examples can be found in which *maybe* co-occurs with the phrase *may be* and the modal verb *might*. These combinations are illustrated in Examples (6a, b), respectively. Such co-occurrence implies that the expression *maybe* is independent of, or decategorializes from, the phrase *may be*, and establishes its status as a modal adverb.

- (6) a. ... and *maybe* Aunt Barbara **may be** got to give me that much at  
(CLMET3\_0\_2\_117)  
b. ... and *maybe* in th’ dark it **might** take me for... (CLMET3\_0\_2\_173)

In the same period, further cases of *maybe* can be observed, particularly in final position, and it proliferated widely from this sub-period onwards. This usage is illustrated in (7a, b). This use of *maybe* signals the fact that *maybe* is less fixed in, and more detached from, the medial position within the clause, which is the typical modal adverb positioning. The next subsection discusses positioning in more detail.

- (7) a. To-morrow I may look on you different, *maybe*. (CLMET3\_0\_2\_173)  
b. I’ve said all this afore, *maybe*. But from that time I’ve dropped down, down down. (CLMET3\_0\_2\_173)

### 3.2. Position

In their corpus analysis, Biber et al. (1999: 872) present a tendency for stance adverbials to be positioned medially in a clause, as shown in Table 8<sup>6</sup>.

<sup>6</sup> According to Biber et al., stance adverbials have “the primary function of commenting on the content or style of a clause or a particular part of a clause” (1999: 853). They

Therefore, this general trend will be compared with the results of the present study.

Table 8. Positioning of stance adverbials across registers (from Biber et al. [1999: 872])

	Initial position (%)	Medial position (%)	Final position (%)
CONVERSATION	• • •	• • • • • • • • • •	• • • • • • •
FICTION	• • • • •	• • • • • • • • • •	• • • • •
NEWSPAPER	• • • • • • •	• • • • • • • • • •	• •
ACADEMIC	• • • • • • •	• • • • • • • • • •	•

each • represents 5%

In my analysis, Figure 1 presents a comparison with different positions of *maybe* from 1801 onwards in the *OED* dataset. The breakdown by position is illustrated in Examples (8a-f). As Figure 1 shows, the use of *maybe* in the initial position commenced with a share of over 50%, and the medial use was a strong runner-up to the initial one.

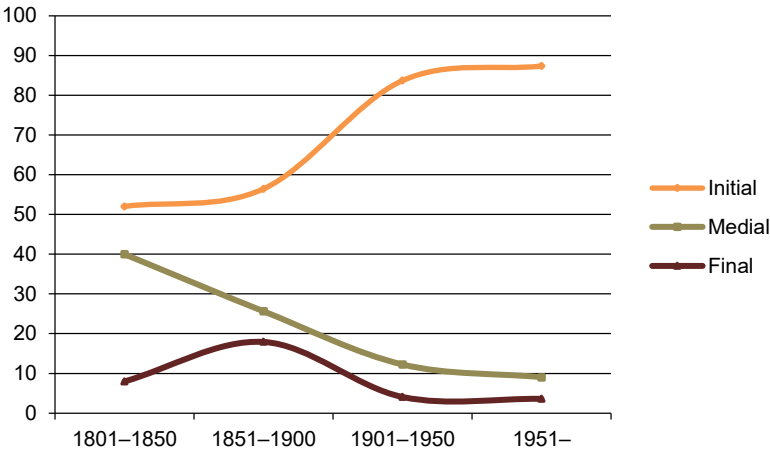


Figure 1. Positioning of *maybe* from 1801 onwards (*OED*)

further categorize stance adverbials into three sub-divisions, namely, epistemic, attitude, and style. Epistemic adverbials include *no doubt, certainly, probably, definitely, I think, in fact, really, according to, mainly, generally, in my opinion, kind of, and so to speak*; attitude adverbials include *unfortunately, to my surprise, and hopefully*; and style adverbials comprise *frankly, honestly, truthfully, and in short*.

The frequency of *maybe* in the initial position then began to rise in subsequent periods. Particularly, the dramatic rise from the twentieth century, reaching a share of between 80% and 90%, made the initial position use vastly outnumber the medial position in present-day English.

(8) **Initial**

- a. *Mebbe* I've ben hard done by all my hull life. (1898 E.N. Westcott *David Harum* (1900) xx. 197, *OED*)
- b. *Maybe* we didn't do so bad for a Dago fisherman and a sheeny storekeeper. (1977 H. Fast *Immigrants* II. 88, *OED*)

**Medial**

- c. Some people will *maybe* not crack quite so crouse by-and-by. (1824 S.E. Ferrier *Inher.* Ixvi, *OED*)
- d. I think Bullock is *maybe* ten years out of date. (1977 *Jrnl. R. Soc. Arts* CXXV. 671/2, *OED*)

**Final**

- e. [Said of a dying man] He will go out with the tide, *may be*. (1891 S.C. Scrivener *Our Fields & Cities* 10, *OED*)
- f. Jacques'll be here about then, *mebbe*. (1910 R. Brooke *Let.* 8 June (1968) 240, *OED*)

The first half of the results in Figure 1, representing the nineteenth century, parallels the results from the CLMET3.0. Figure 2 displays the shift of *maybe* from 1780 to 1920 for different positions<sup>7</sup>.

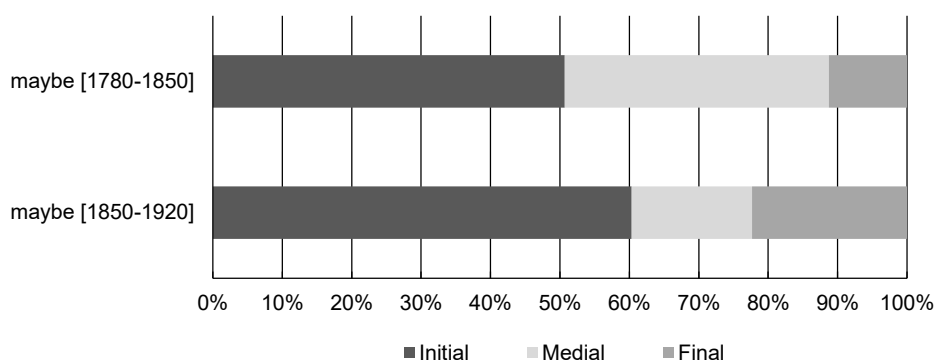


Figure 2. Positioning of *maybe* from 1780 to 1920 (CLMET3.0)

<sup>7</sup> In the first sub-period of CLMET3.0, namely, the period [1710-1780], only two examples were obtained from the data; therefore, my analysis here is restricted to the subsequent two sub-periods, [1780-1850] and [1850-1920].

As Figure 2 shows, the percentage of *maybe* occurring initially has increased from 50.7% of all occurrences in the period [1780-1850] to 60.3% in the period [1850-1920].

The latter half of the results in Figure 1, representing the twentieth century, is in line with the results from the Brown family of corpora. Figure 3 indicates that in American English, the percentages of the initial *maybe* are fairly stable, and its share is nearly 90%. Figure 4 clarifies that *maybe* in British English has experienced a steady increase in its initial position from 87.0% in LOB to 91.4% in FLOB.

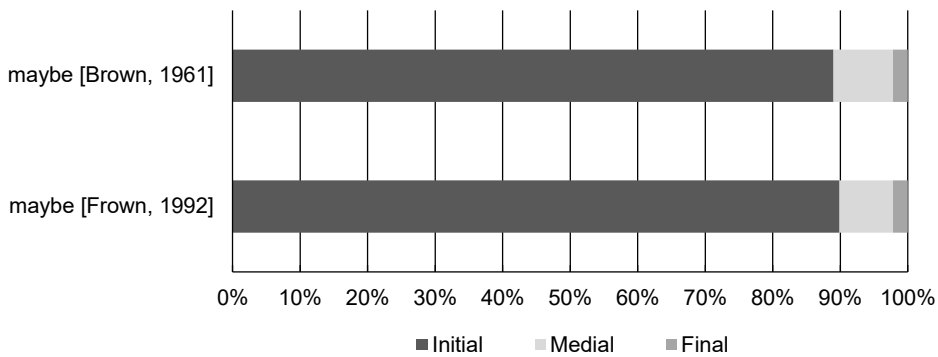


Figure 3. Positioning of *maybe* in American English in 1961 and 1992 (Brown and Frown)

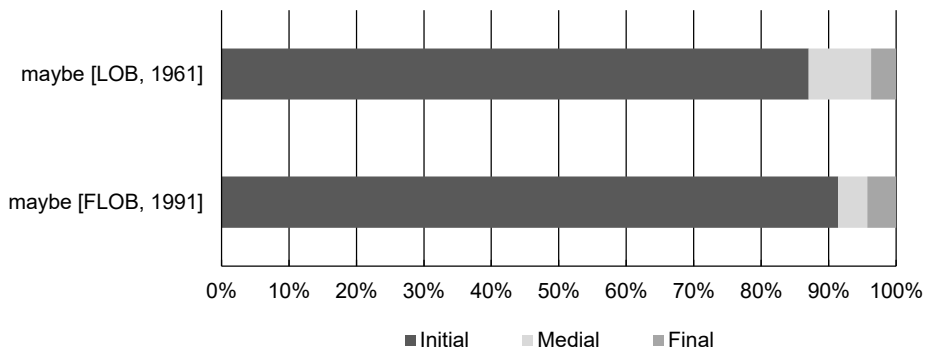


Figure 4. Positioning of *maybe* in British English in 1961 and 1991 (LOB and FLOB)

### 3.3. Topic marking and subjectification

The development of *maybe* can be attested to by a process of grammaticalization. One of the most significant features of grammaticalization, referred to as decategorialization, can apply to the present shift (cf. Hopper 1991; Hopper

– Traugott 2003). *Maybe* originated in the verbal phrase *may be*, acquired adverbial use with the epistemic function at a sentence level, and has further shown syntactically flexible mobility in a clause, developing an increased scope. In addition, the increased frequency of *maybe* in the last two centuries seems to play an important role in the process, as the path of grammaticalization is intrinsically involved in a significant increase in frequency (cf. Bybee 2003, 2006). This quantitative approach also presents a strong correlation of frequency with grammaticalization.

The close relationship between position and function with regard to English modal adverbs is now presented. Consider the following comparable examples:

- (9) a. It *may* have been Wren.
- b. *Possibly* it was Wren. (Halliday 1970: 335)

A quick look at (9a, b) reveals that *may* in (9a) and *possibly* in (9b) convey the same meaning of epistemic possibility. The difference, however, resides in their position, as *possibly* occurs initially. As argued by Halliday (1970: 335), Perkins (1983: 102-104), Høye (1997: 148-152), and Halliday – Matthiessen (2014: 105-111), a modal adverb occurring initially expresses the topic or theme (that of modality), and serves the topic marking function. The speaker or writer provides the addressee or reader with the material that will help to (re)construct the flow of the discourse. *Possibly* in (9b) fulfills the function of expressing the topic or theme of modality in addition to merely expressing modality. Thus, the present study shows that *maybe*, occupying this position much more frequently, is strongly attracted to the function of topic encoding, and has become prevalent in present-day English.

The development in the use of *maybe* as a topic marker in the contemporary stages implies that the shift of *maybe* is also accompanied by a process of general linguistic change, namely, subjectification, and that the meanings of the adverb have changed from less to more subjective in present-day English. Subjectification is the path whereby the speaker or writer recruits meanings “that encode or externalize their perspectives and attitudes as constrained by the communicative world of the speech event” (Traugott – Dasher 2002: 30). The occurrence of modal adverbs in the initial position derives from the result of the speaker’s or writer’s choice of word order, signalling the speaker’s or writer’s perspective regarding the flow of discourse. More precisely, they “serve the procedural purposes of expressing speaker’s attitude to the text under production (topicalizers, discourse markers)” (Traugott 2010: 31). In this sense, their use in this position is closely

associated with subjectivity (Traugott – Dasher 2002; Brinton 2007; Traugott 2012; Beeching – Detges 2014). The use of *maybe* thus exhibits the subjective view of the speaker or writer, and its development is characterized as subjectification in contemporary English.

Grammaticalization and subjectification as processes of linguistic change are not incompatible with each other. Grammaticalization is “a complex multilevel diachronic process leading towards grammar,” whereas subjectification is “a particular type of semantic change” (Diewald 2011: 373). Diewald further states that “[t]hough subjectification is often found as one component of grammaticalization processes, it is independent of and not restricted to it” (ibid.). Viewing the two processes in parallel, subjectification in grammaticalization is “the development of a grammatically identifiable expression of speaker belief or speaker attitude to what is said” (Traugott 1995: 32; cf. Traugott 2003). In the present study, it was shown that subjectification derived from the shift in the positioning of the modal adverb within the clause, and this development of the modal adverb can be considered an instance of grammaticalization accompanied by subjectification.

#### 4. Conclusion

Although the origin of *maybe* is well documented in the dictionaries of word history, its development from Modern English to present-day English is virtually unexplored. Despite the earlier appearance and epistemic use of *maybe* in the history of English, the later periods, namely both the nineteenth and twentieth centuries, witnessed a significant shift of *maybe*. In addition, the development of *maybe* is accompanied by grammaticalization and subjectification.

This study relied on a combination of the *OED* quotations database, CLMET3.0, and the Brown family of corpora. These datasets enabled us to describe the long-term development of modal adverbs from the Modern English period to contemporary stages. Although the *OED* dataset has drawbacks as a “corpus,” this study has shown that it can yield sufficient token counts, and the results are in line with those from other authentic corpora. The dataset can be regarded as “a useful and reliable source of data” (Rohdenburg 2013: 157), and can “provide the linguist with a wealth of useful information” (Hoffmann 2004: 26), spanning more than a thousand years of English. It is particularly helpful for offering a historical overview covering the long-term development of items that do not have a high frequency.

## REFERENCES

### Sources

*Anglo-Norman Dictionary*

1977 London: Modern Humanities Research Association.

Biber, D. et al.

1999 *Longman Grammar of Spoken and Written English*. Harlow: Pearson.

*Brown University Standard Corpus of Present-Day American English* (Brown)

1964 Compiled by W.N. Francis – H. Kučera.

*Corpus of Late Modern English Texts*, version 3.0 (CLMET3.0)

2005 Compiled by H. De Smet – H. Diller – J. Tyrkkö.

*Freiburg-Brown Corpus of American English* (Frown)

1999 Compiled by C. Mair.

*Freiburg-LOB Corpus of British English* (FLOB)

1999 Compiled by C. Mair.

*Lancaster-Oslo/Bergen Corpus of British English* (LOB)

1976 Compiled by G. Leech – S. Johansson – K. Hofland.

*Middle English Dictionary*

1954 Ann Arbor: University of Michigan Press.

*Oxford Dictionary of English Etymology*

1966 Oxford: Clarendon Press.

*Oxford Dictionary of Word Histories*

2004 Oxford: Oxford University Press.

*Oxford English Dictionary*, 2<sup>nd</sup> edn. (CD-ROM, version 4.0)

2009 Oxford: Oxford University Press.

Quirk, R. et al.

1985 *A Comprehensive Grammar of the English Language*. London: Longman.

Swan, M.

2005 *Practical English Usage*, 3<sup>rd</sup> edn. Oxford: Oxford University Press.

### Special studies

Beeching, K. – U. Detges (eds.)

2014 *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Leiden: Brill.

Berg, D.L.

1991 *A User's Guide to the Oxford English Dictionary*. Oxford: Oxford University Press.

Brewer, C.

2000 "OED Sources". In: L. Muggleston (ed.) *Lexicography and the OED: Pioneers in the Untrodden Forest*. Oxford: Oxford University Press, 40-58.

- Brinton, L.J.  
 2007 "The development of *I mean*: Implications for the study of historical pragmatics". In: S. Fitzmaurice – I. Taavitsainen (eds.) *Methods in Historical Pragmatics*. Berlin: Mouton de Gruyter, 37-80.
- Bybee, J.  
 2003 "Mechanisms of change in grammaticalization: The role of frequency". In: B.D. Joseph – R.D. Janda (eds.) *Handbook of Historical Linguistics*. Oxford: Blackwell, 602-623.  
 2006 "From usage to grammar: The mind's response to repetition", *Language* 82, 711-733.
- Davidse, K. – S. De Wolf – A. Van Linden  
 2015 "The development of the modal and discourse marker uses of (*there/it is/I have*) *no doubt*", *Journal of Historical Pragmatics* 16(1), 25-58.
- Diewald, G.  
 2011 "Pragmaticalization (defined) as grammaticalization of discourse functions", *Linguistics* 49(2), 365-390.
- Halliday, M.A.K.  
 1970 "Functional diversity in language as seen from a consideration of modality and mood in English", *Foundations of Language* 6, 322-361.
- Halliday, M.A.K. – C.M.I.M. Matthiessen  
 2014 *Halliday's Introduction to Functional Grammar*, 4<sup>th</sup> edn. Abingdon, Oxon.: Routledge.
- Hanson, K.  
 1987 "On subjectivity and the history of epistemic expressions in English". In: B. Need et al. (eds.) *Papers from the 23<sup>rd</sup> Annual Regional Meeting of the Chicago Linguistic Society*, 133-147.
- Hoffmann, S.  
 2004 "Using the OED quotations database as a corpus – A linguistic appraisal", *ICAME Journal* 28, 17-30.
- Hopper, P.J.  
 1991 "On some principles of grammaticization". In: E.C. Traugott – B. Heine (eds.) *Approaches to Grammaticalization*. Amsterdam: Benjamins, 17-35.
- Hopper, P.J. – E.C. Traugott  
 2003 *Grammaticalization*, 2<sup>nd</sup> edn. Cambridge: Cambridge University Press.
- Hoye, L.  
 1997 *Adverbs and Modality in English*. London: Longman.
- Iyeiri, Y.  
 2010 *Verbs of Implicit Negation and their Complements in the History of English*. Amsterdam: Benjamins.
- Lenker, U.  
 2010 *Argument and Rhetoric: Adverbial Connectors in the History of English*. Berlin: Mouton de Gruyter.



- López-Couso, M. – B. Méndez-Naya  
 2016 “From clause to adverb: On the history of *maybe*”. In: G. Kaltenböck – E. Keizer – A. Lohmann (eds.) *Outside the Clause*. Amsterdam: Benjamins, 157-176.
- Mair, C.  
 2001 “Early or late origin for *begin* + V-ing? Using the OED on CD-ROM to settle a dispute between Visser and Jespersen”, *Anglia* 119, 606-610.  
 2004 “Corpus linguistics and grammaticalisation theory – Statistics, frequencies, and beyond”. In: H. Lindquist – C. Mair (eds.) *Corpus Approaches to Grammaticalization in English*. Amsterdam: Benjamins, 121-150.
- Perkins, M.R.  
 1983 *Modal Expressions in English*. London: Frances Pinter.
- Rohdenburg, G.  
 2013 “Using the OED quotations database as a diachronic corpus”. In: M. Krug – J. Schlüter (eds.) *Research Methods in Language Variation and Change*. Cambridge: Cambridge University Press, 136-157.
- Simon-Vandenberghe, A.-M. – K. Aijmer  
 2007 *The Semantic Field of Modal Certainty: A Corpus-based Study of English Adverbs*. Berlin: Mouton de Gruyter.
- Terasawa, Y. (ed.)  
 1997 *The Kenkyusha Dictionary of English Etymology*. Tokyo: Kenkyusha.
- Traugott, E.C.  
 1995 “Subjectification in grammaticalization”. In: D. Stein – S. Wright (eds.) *Subjectivity and Subjectivisation: Linguistic Perspectives*. Cambridge: Cambridge University Press, 31-54.  
 2003 “Constructions in grammaticalization”. In: B. Joseph – R. Janda (eds.) *The Handbook of Historical Linguistics*. Oxford: Blackwell, 624-647.  
 2010 “(Inter)subjectivity and (inter)subjectification: A reassessment”. In: K. Davidse – L. Vandelanotte – H. Cuyckens (eds.) *Subjectification, Intersubjectification and Grammaticalization*. Berlin: Mouton De Gruyter, 29-71.  
 2012 “Intersubjectification and clause periphery”, *English Text Construction* 5(1), 7-28.
- Traugott, E.C. – R.B. Dasher  
 2002 *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Address: DAISUKE SUZUKI, Setsunan University, Faculty of Foreign Studies, 17-8 Ikedanaka-machi, Neyagawa, Osaka, 572-8508, Japan.

ORCID code: 0000-0003-4357-3119.

# **The lexical field WINE in American and British English: A corpus-based study**

Paulina Ołownia

*Jan Kochanowski University of Kielce*

## **ABSTRACT**

This study examines the development of the lexical field WINE. It is conducted on the basis of the set of lexemes denoting WINE which are extracted from the *Historical Thesaurus of English* (HTE). The lexemes have been classified according to their origin and the results are compared with the quantitative distribution of the lexeme *wine* with adjectives denoting nationality. The source for the analysis is the corpus collection compiled at Brigham Young University (COCA, COHA and BNC). My study demonstrates that there exists a correlation between the origin of a given wine and the etymology of the lexemes denoting it. Moreover, the countries enumerated as popular wine exporters to the British Isles appear to correlate with both the adjectives modifying the lexeme *wine* and the etymology of lexemes in the lexical field WINE. Furthermore, the lexemes have been studied diachronically in order to verify the correlations among the dates of occurrence in the HTE and in the corpora. My analysis shows that there are some significant differences among the dates of attestation of the lexeme under investigation. The paper encompasses also an analysis of temporal trends in adopting both native lexemes and borrowings in the lexical field WINE, as well as a quantitative comparison between the distribution of the lexemes in American English and British English; the aim is to define prototypical lexemes for the lexical field studied and their frequency in both varieties of the English language.

Keywords: lexeme, lexical field, borrowing, lexicology.

## **1. Introductory remarks**

The aim of this study is to examine the development of the lexical field WINE. It offers a diachronic and quantitative analysis and is conducted on the basis of a set of borrowings from different languages selected according

to their denotation. The source for the borrowings is the *Historical Thesaurus of English* (HTE). The analysis makes use of the *Corpus of Historical American English* (COHA), the *Corpus of Contemporary American English* (COCA), and the *British National Corpus* (BNC). The study also examines the etymology of these lexemes as outlined in the *Oxford English Dictionary* (OED). The paper provides not only a classification of the borrowings of the lexical field WINE according to their etymology, but also a discussion of cultural influences on the quantitative distribution of the selected borrowings. This empirical investigation employs lexical field theory, and it explores linguistic processes accompanying the development of lexical fields.

## 2. The theory of lexical field: A brief description

The starting point for discussing a lexical field is the notion of **lexeme**. Crystal (1997: 276) states that it is “a term used by some LINGUISTS to refer to the minimal DISTINCTIVE UNIT in the SEMANTIC SYSTEM of a LANGUAGE. [...] Lexemes are the units which are conventionally listed in dictionaries as separate entries”. It should be noted that different inflectional forms represent the same lexeme, e.g. *go*, *goes*, *went*, *gone*, and *going* belong to the lexeme *go*, but the words which are created by derivational affixation are different lexemes, e.g. *job* and *jobless* are two separate lexemes.

Lexemes are grouped in a mental lexicon into categories termed lexical fields. One of the pioneers of the notion of lexical field was the German linguist Jost Trier. Trier (in Eckardt et al. 2003: 68-69) claims that the meanings of lexemes are based on their relationship within the same semantic field and the extension of meaning of one lexeme limits the meanings of the other lexemes in the same semantic field. Most linguists use the term “lexical field” interchangeably with the term “semantic field”, for instance Lehrer (1985: 283) defines a semantic field as “a set of lexemes which cover a certain conceptual domain”. As Langacker (1987: 147) claims, “Domains are necessarily cognitive entities: mental experiences, representational spaces, concepts, or conceptual complexes”. Thus, it seems that the terms included in a lexical field, semantic field or a conceptual domain refer to the same notion. Nevertheless, Wyler (1992: 30) distinguishes the lexical field which he calls “a structure formed by lexemes” from the semantic field which is “the underlying meaning which finds expression in lexemes”. This division can be considered to refer respectively to semasiological and onomasiological

approaches to lexemes. Semasiology is a semantic approach based on the question of the meaning of a particular lexeme, while onomasiology addresses the question of which lexeme can be used to name a particular concept. Wyler's (1992: 30) definition of lexical field seems to correspond to the semasiological approach to lexemes while semantic field as defined by this author is onomasiological. It seems that in real life, as well as in the study of the development of a lexical field, onomasiology is the more common approach. The human mind works onomasiologically – we continuously coin new terms for new or existing concepts. Moreover, semasiology is not only a term antonymous to onomasiology, but semasiology appears to be the result of an onomasiological approach. It seems that we name concepts first, and only later can the question of the referent of a particular linguistic sign be asked.

An important feature of lexemes in a lexical field is that some lexemes may be prototypical, i.e. more recognizable than others as category members. Thus, "a sparrow would be a prototype of bird, whereas an ostrich (because of its atypical characteristics, notably its inability to fly) would not" (Crystal 1997: 395). Rosch – Mervis (in Geeraerts 1997: 11) similarly say that "some natural categories are analog and must be represented logically in a manner which reflects their analog structure". Prototypicality, according to the above definitions, may be established mainly by analyzing the characteristics of concepts belonging to a particular category; however, it seems that the criterion of frequency is important in order to decide whether some lexemes in a lexical field are more representative (prototypical) or more peripheral for the category. In the case of the lexical field WINE, the characteristics of its members are not crucial. The lexemes in this empirical study are going to be classified as prototypical based on the analysis of the most frequent occurrences in both American and British English.

### 3. Sources for the analysis of the lexical field WINE

The source for the analysis is a list of 138 lexical items, which was derived from the *HTE* online entry for *Types of wine* in an unmodified form in 2013 (see Appendix 1). Other tools for the analysis are the corpora compiled at Brigham Young University (BYU): *COHA*, *COCA*, and the *BNC*, and the electronic version of the *OED*.

The analyses of the selected lexemes were conducted on the basis of electronic corpora. The sources here are certain corpora which were compiled by Mark Davies at Brigham Young University. These corpora are *COHA*, *COCA*, and the *BNC*. *COHA* is a diachronic corpus of historical American English which encompasses more than 400 million words from 1810 to 2009. *COCA* is a corpus of contemporary American English which consists of more than 520 million words from 1990-2015. The *BNC* was first created by Oxford University Press and now it exists in different versions on the Web. The version used in the study is the *BNC* by Mark Davis. It is a corpus of British English which contains 100 million words from the 1980s to 1993. The corpus includes both spoken and written language, and it is arranged into subcategories. The corpora are balanced and grammatically tagged, which facilitates comprehensive searches.

#### 4. Selection and categorization of the lexemes used in the analysis of the lexical field WINE

This study constitutes a practical application of lexical field theory to an analysis of the lexical field WINE. The terms are those of the *HTE* entry mentioned above (see Appendix 1), and they are grouped according to their denotations. In order to avoid questionable groupings under the category WINE, homographs and lexemes whose cotext does not allow their denotation to be identified, have been excluded from the list; e.g., *sparkler* may mean a firework, a diamond or everything which is sparkling, or *must* may be a verb. All the data presented in this study were compiled in 2013; therefore, equivalent present-day statistics based on updated versions of the corpora may differ somewhat. Nevertheless, the conclusions drawn below remain valid within their parameters. Moreover, the periods studied seem to be sufficient to verify the correlation between some cultural factors and the lexical field WINE creation, as well as to conduct a comparative analysis of the distribution of lexemes in the field in American English and British English.

It should be noted that because the category *Types of wine* includes that of *Types of grapes*, differentiation between these two categories is not always possible due to contextual limits imposed by the corpora used in the present study. However, if a term includes *grape*, as in *labrusca grape*, and is important to the diachronic discussion, perhaps because it constitutes the

first known occurrence of the term, it is analyzed and the indication that it denotes a grape is presented below in an appropriate table or figure.

Also, in the absence of contextual clues indicating semantic contrasts, spelling and phrase variants occurring in the *HTE* are ignored in the sense that all variants of a phrase, for instance, are considered to signify a single meaning. Moreover, the entries *red*, *ruby*, and *rose* were included in the discussion only in the combination with the lexeme *wine* which is marked by *wine* enclosed in square brackets, e.g. *red* [wine]. The term *rosé* has been treated as a variant form of *rose* [wine], and because it does not occur in the data in combination with *wine*, it has been excluded. In the case of the alternative phrase *rape* (wine), only the combination *rape wine* is taken into consideration since the form *rape* may instance a different but homonymous lexeme.

Finally, the terms *ceren* and *Sauvignon* occur under different subcategories in the *HTE*. *Ceren* can be found in subcategory 08 *sweet wine* and 11 *new wine*; *Sauvignon* occurs in subcategory 19 *wine from specific grapes* 19.02 *Cabernet Sauvignon* and 19.04 *others*. For the present purposes, *ceren* has been treated as one term and *Sauvignon* has been considered to represent *Cabernet* + *Sauvignon* i.e. [Cabernet] *Sauvignon* except when it is modified by other lexemes and must be taken to signify a different type of wine, as in *Fiddlehead Sauvignon Blanc*.

## 5. Classification of the lexemes according to their origin

The terms selected from the *HTE* have also been classified according to their origin. Verification was obtained by reference to the electronic *OED*. However, because the version used was not most updated one, some lexemes were impossible to categorize due to lack of etymological information or complete absence of the lexeme from the dictionary. Eleven eligible lexemes were found in other *Oxford Dictionaries*, and nineteen were seen to have obscure etymology.

Fig. 1 presents the distribution of the lexemes according to their origin. Indications are that French has been the primary donor language as regards English borrowings in lexical field WINE. Still, native lexemes constitute twenty-four percent of those analyzed. Loans of Italian, Spanish, German, Latin and Portuguese origin constitute less than ten percent. Some borrowings of the lexical field WINE, though fewer, have come in from South African, Bulgarian, Dutch, Greek and Persian.

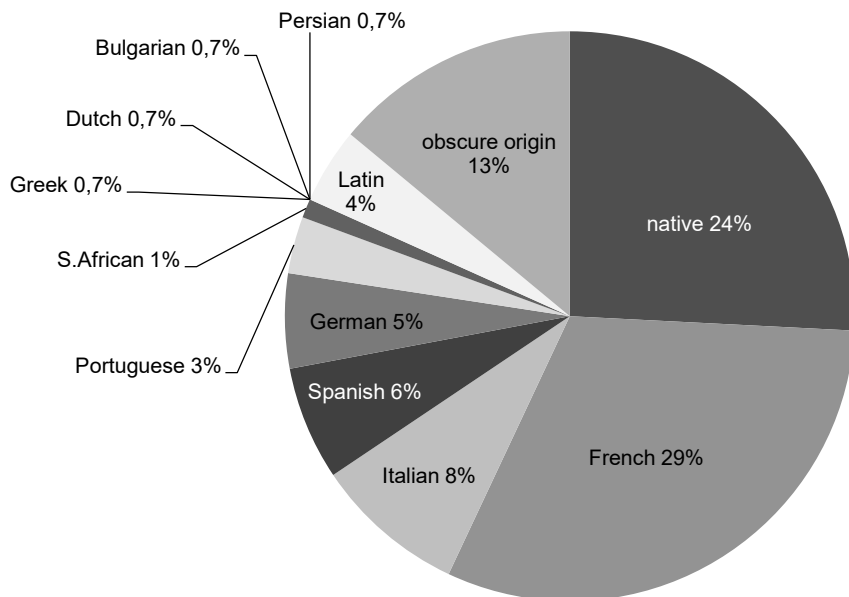


Fig. 1. Origins of lexemes in the lexical field WINE

It should be noted that one term denoting WINE, *vinho de casa* of Portuguese origin, was not present in the *HTE* under the entry *Types of wine*, but was found in the *OED*.

## 6. Quantitative analysis of the combinations of the lexeme *wine* with adjectives denoting nationality

While the previous section discussed the origins of the lexemes in the lexical field WINE, this section contains a discussion of the combinations of the lexeme *wine* with adjectives denoting nationality. The quantitative analysis is conducted with the use of *COCA* and the *BNC*. The aim here is to establish which adjectives denoting nationality modify the lexeme *wine* most often. This, together with the information concerning the origin of the lexemes in the previous section, provides the basis for drawing conclusions about cultural influence on the development of the lexical field WINE in English. A quantitative comparison of the results in *COCA* and the *BNC* serves to differentiate the combinations in American and British English.

Fig. 2 presents all the combinations of the lexeme *wine* with adjectives denoting nationality occurring in *COCA* and the *BNC*. The results are normalized per 1,000,000 words.

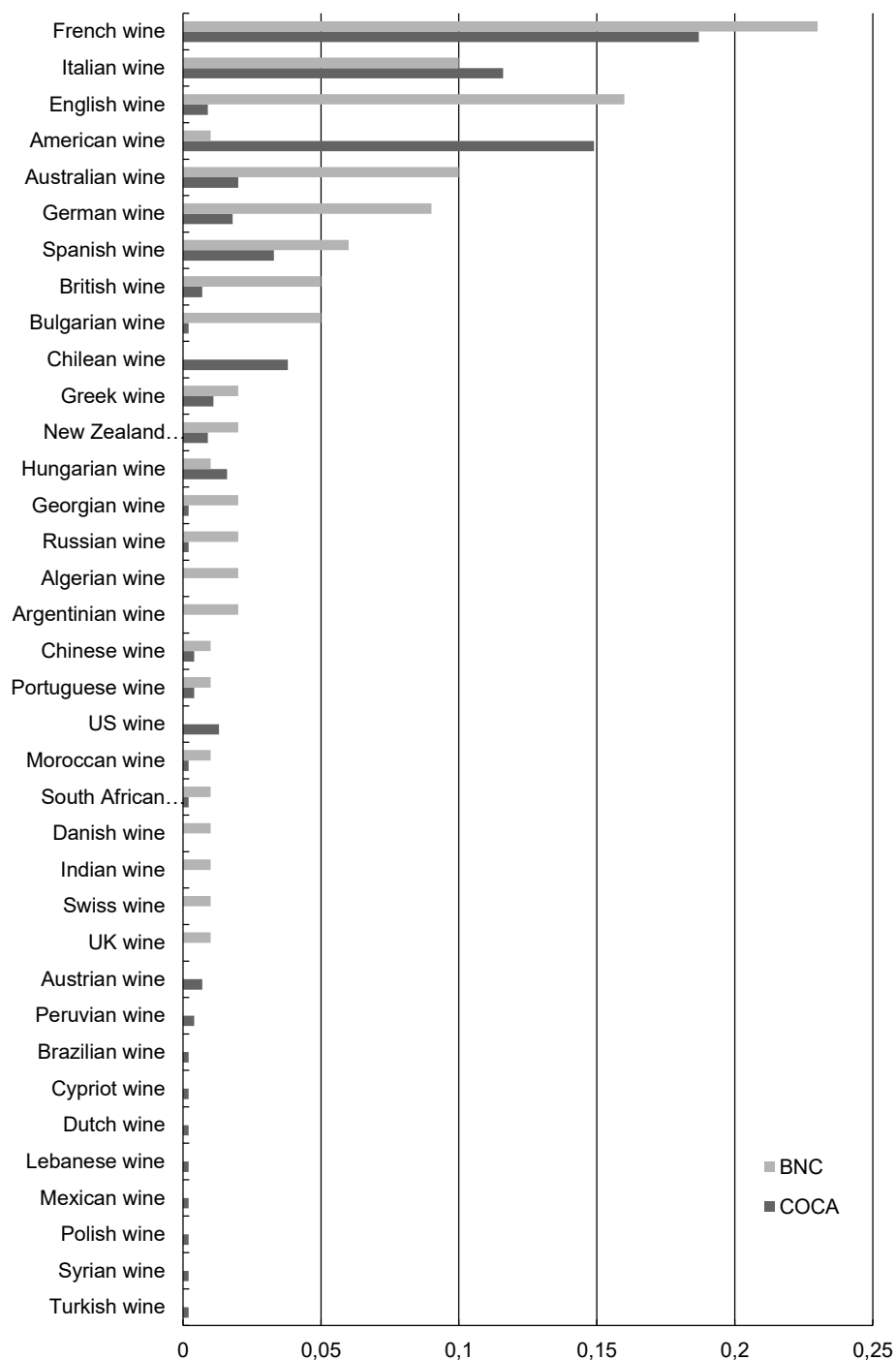


Fig. 2. Comparative analysis of the combinations of the lexeme *wine* with adjectives denoting nationality in COCA and the BNC



*French*, *American/US* and *Italian* are significantly prevailing adjectives denoting nationality in COCA. *American wine* denotes the same referent as *US wine*, thus the numbers of the combinations are considered together. Similarly, *English*, *UK*, and *British* are naturally treated as one; although the lexeme *English* does not refer to the same area as *UK* or *British*, it is considered the same in the study as it denotes wine from the British Isles. The prevailing adjectives denoting nationality in the combinations with the lexeme *wine* in the BNC are *French*, *English/British/UK*, *Australian*, *Italian*, and *German*. The most frequent adjectives modifying *wine* in the study in both COCA and the BNC are *French* and *Italian*. In COCA the second most frequent combination is *American + US wine* and in the BNC it is *English + British + UK wine*, which marks the importance of native wines in these two countries. These comparisons of the combinations of adjectives denoting nationality with the lexeme *wine* indicate that more types of such combinations exist in American English than in British English – thirty combinations are evidenced in COCA and twenty-four in the BNC. Nevertheless, if normalized results are taken into consideration, twenty-three out of thirty-six combinations (64%) prevail quantitatively in the BNC data.

It is also revealing to examine whether the trend of combining adjectives denoting nationality + *wine* correlates with the trend concerning the origin of the lexemes in the lexical field WINE. Fig. 3 presents the percentage relations between the trends – the percentage of the particular combination out of all the combinations studied and the percentage of the lexemes of particular origins out of all the lexemes studied. The data have been rounded up to whole digit percentages.

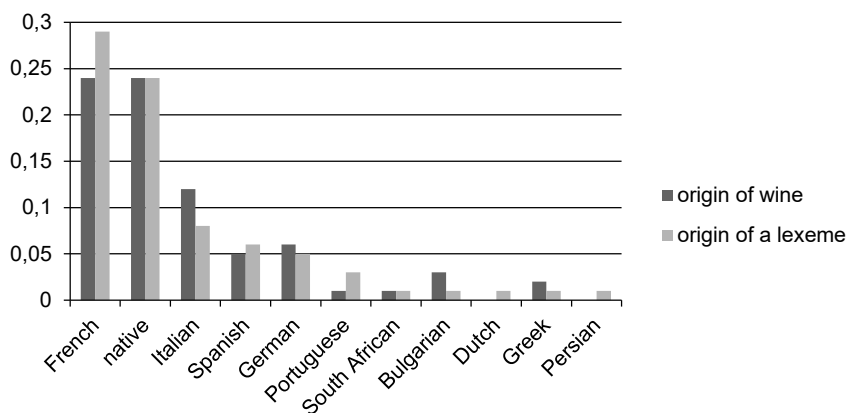


Fig. 3. Relation between *adjectives denoting nationality + wine* and adjectives denoting the origins of the lexemes in the lexical field WINE

The most popular adjective denoting nationality in combination with the lexeme *wine* is *French* (24%). It correlates with the most prevailing origin of lexemes in the lexical field WINE, which is French (29%). These data suggest that French culture has had the most significant influence on the development of the English lexical field WINE. The adjectives with the next highest percentages are *American/US* and *British/English/UK*, which also correlates with the quantity of native lexemes in the lexical field WINE (24% for both the combinations and the origins). The third highest combination is *Italian wine* (12%), which aligns roughly with the Italian lexeme origin percentage (8%). The percentage relations between the remaining combinations and origins are similar, which confirms the supposition that the countries in Fig. 3 might have had the most significant influence on the development of the lexical field WINE in the English language. Moreover, it can be expected that the process of borrowing may be influenced by various phenomena which lead to the meeting of different languages. Borrowing lexemes of the category WINE into English might have been influenced by wine trade between Great Britain and other countries. As Spahni (2000: 73) claims, “French, Italian and German wines [historically have and] still [do] make up the majority of wines sold in Britain (e.g. 57% of still wines sales in 1997), but [or even though] they have New world exporters hard on their heels (e.g. Australia, South Africa, the US and Chile in, respectively, fourth, sixth, seventh, and eighth place in 1997)”. Furthermore, Ludington (2013: 24) presents France and Spain as main wine exporters to Britain in 1675 and Estreicher (2006: 66) emphasizes the importance of trade between England and France and Portugal:

In the 1500s, England dominates the wine trade. Following the end of the Hundred Years War with France, the English merchants look for new suppliers of wine to make up for the clarets from Bordeaux. The English already do business in northern Portugal, where the light and fruity Vinho Verde is made.

There can be clearly found a correlation between the wine trading trends discussed above and the patterns of borrowing lexemes in the lexical field WINE into English. The countries enumerated above as significant wine exporters to Britain, namely France, Italy, Germany, Spain, Portugal, or South Africa are also visible in Fig. 3, which may demonstrate the cultural influence of wine trade on the expansion of the lexical field WINE in the English language.

## 7. Diachronic analysis of selected lexemes

The terms listed under the *HTE* category *Types of wine* have been selected according to the criteria mentioned above, and they have been studied diachronically. Apart from a diachronic study with the use of *COHA*, given lexemes were analyzed comparatively according to the time of occurrence in the *HTE* and the corpora (*COHA*, *COCA*, and the *BNC*) in order to verify the correlations among the dates of occurrences. This section also outlines trends in lexeme borrowing in the English lexical field WINE.

Table 1 below shows differences between the dates of occurrence of the lexemes noted in the *HTE* and those evidenced in the corpora (*COHA*, *COCA*, and the *BNC*). It contains only the terms whose dates of occurrence in the corpora differ from those given in the *HTE*. Thus, empty cells indicate that the date range of occurrence in *COHA*, *COCA*, or the *BNC* is within the range given in the *HTE*, and hyphens indicate that lexemes do not occur in a given corpus. A hyphen at the end of a date in the *HTE* column signifies that the term is current, while a semicolon between the dates suggests a gap in evidence during the period. An individual date indicates a single citation in the case of the *HTE* and one or more attestations within the same year in the case of the *BYU* corpora. If the number of occurrences in different years is more than two, the dates are marked as a range. For example, *wine of astonishment* occurs three times in *COHA* in the years 1849, 1945 and 1981, and the range is shown as 1849-1981. When two occurrences in two different years are in evidence, as in the case of the term *wine whey*, the dates are separated by a semicolon (1880; 1950). Bolded dates indicate occurrences in the corpora that are earlier than those cited in the *HTE*. A total of nine dates are bolded for this reason in Table 1.

Table 1. Dates of occurrence of the lexemes in the *HTE* and in the corpora (*COHA*, *COCA*, and the *BNC*)

Term	<i>HTE</i>	<i>COHA</i>	<i>COCA</i>	<i>BNC</i>
1	2	3	4	5
bastard wine	1436-1616	–	2003	–
charneco	1593-1631	1937	–	–
rape [wine]	1600-1733	–	2001	–
brown bastard	1603; 1609	1838-1937	–	–
wine of astonishment	1611	1849-1981	–	1985-1994
vino tinto	1673; 1901-	1843		–
wine whey	1769-1856	1880; 1950	–	–

1	2	3	4	5
lunel	1770-1841		1999	–
Traminer	1851-	<b>1848</b>		
mousseux	1861-	<b>1840</b>	–	
verdelho	1883-	<b>1848</b> (grape)		
vin rouge	1917-	<b>1908</b>		
Shiraz	1966-	<b>1860</b>		
vino nero	1968-	<b>1960</b>	–	–
vino blanco	1981-	<b>1907</b>		–
Lambrusco	1986-	<b>1950</b>		1985-94
Merlau	a1997	–	2011	–
labrusca	1988	<b>1838-1993</b> (grape)		–

The following examples show the earlier occurrence of the lexemes in COHA:

- (1) “Royal Muscadine, one vine. Red *Traminer*, one vine. White Rishng, one vine.” 1848 in *A practical treatise on the culture and treatment of the grape vine: embracing its history, with directions for its treatment, in the United States* by John Fisk Allen.
- (2) “Were you ever at Rome?” – asked the restaurateur as he finished his second bottle of *Mousseux*, and drew from the closet a larger supply of Vin de Chambertin.” 1840 in *Tales of the Grotesque and Arabesque*, volume 1 by Edgar Allan Poe.
- (3) “*Verdelho* is a small oval white grape, very good, and a great bearer.” 1848 in *A practical treatise on the culture and treatment of the grape vine: embracing its history, with directions for its treatment, in the United States* by John Fisk Allen.
- (4) “Some day she will poison the soup or the *vin rouge*;...” 1908 in *A Bundle of Letters* by Henry James.
- (5) “Just above Mr. Redruth’s head stood jars of sweet Cypress, and a carabas of *Shiraz* wine.” 1860 in *Sir Rohan’s Ghost. A Romance* by Harriet Elizabeth Prescott Spofford.
- (6) “In Italian cafes, they sit six deep around a cheap bottle of *vino nero*, dawdle away an afternoon for 30.” 1960 in *Migration of the Hairy Legs* in Time Magazine: 1960/10/25, no author.

- (7) "But the luscious fruits, the "*vino blanco*," and champagne cool our smarting palates and reconcile us to our gastronomic ventures." 1907 in *Under the Southern Cross* by Elizabeth Robins.
- (8) "While he was eating a hearty lunch accompanied by a quart of sparkling *Lambrusco*, a Communist leader burst into his room." 1950 in *Militant Mouse* in Time Magazine: 1950/10/09, no author.
- (9) "The most common, and in the opinion of many botanists the only species of American grape, vitis *Labrusca* or fox grape..." 1838 in *Kenrick's American Orchardist* in North American Review: October 1838: 423-452, no author.

In addition, nine lexemes appear in the corpora at date intervals which contrast with those indicated in the *HTE*. Falling within the range of dates given in the *HTE*, the subsequent occurrences of particular lexemes are not marked as one point of this investigation is to reveal occurrences other than those cited in the *HTE*, which clearly requires amendment.

The diachronic analysis has facilitated the discovery of certain temporal trends in the borrowing of relevant foreign lexemes which have contributed to the development of the English lexical field WINE. Fig. 4 demonstrates first occurrences of the lexemes in the lexical field WINE in time intervals of five decades. Where dates of attestation in the corpora differ from those cited in the *HTE*, the dates indicated are the earlier (see Table 1).

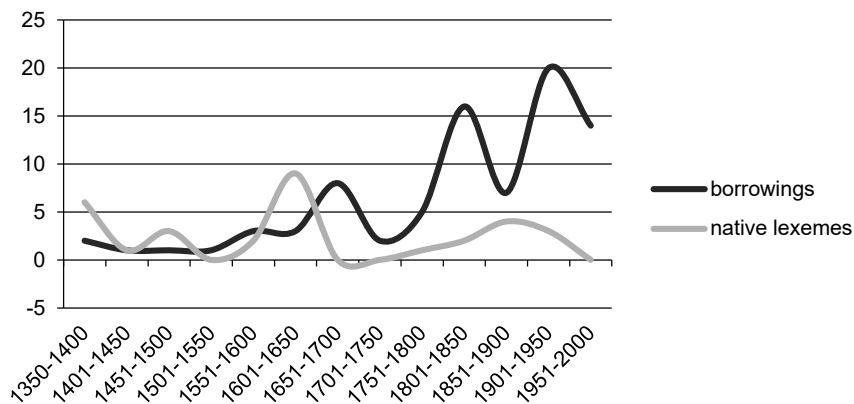


Fig. 4. The first occurrences of borrowed and native lexemes in the lexical field WINE

Fig. 4 demonstrates trends in the development of the lexical field WINE. Although relative stability characterizes borrowing patterns from the late 14<sup>th</sup> century to the early 17<sup>th</sup> century, the number of loans of this lexical domain increases somewhat sharply in the second half of the 17<sup>th</sup> century. After reaching a peak, borrowing returns approximately to previous levels in the course of the 18<sup>th</sup> century. During the next one hundred and fifty years, the number of borrowings increases eightfold (16 lexemes being borrowed in the first half of the 19<sup>th</sup> century). In the late 19<sup>th</sup> century, it decreases by more than half (7 lexemes being borrowed during that half century). The first half of the 20<sup>th</sup> century sees the greatest number of borrowings of the lexical field WINE (20 lexemes being borrowed at that time). In the late 20<sup>th</sup> century, the number of loans decreases by a third or so (14 lexemes being borrowed then).

The quantity of native lexemes in the lexical field WINE varied somewhat over those centuries. Most striking, however, is that their numbers remained comparatively small. In the early 17<sup>th</sup> century 9 lexemes were introduced. Quantitatively, that peak is followed by the late 14<sup>th</sup> century introduction of 6 lexemes and the late 19<sup>th</sup> century addition of 4 lexemes. Overall, the native terms were considerably outnumbered by the borrowed ones.

## 8. Quantitative analysis of the selected lexemes

The terms selected to represent the English lexical field WINE, listed under the *HTE* entry *Types of wine* (see Appendix 1), have been analyzed according to their numbers of instantiation in *COHA*, *COCA* and the *BNC*. Fig. 5 and Fig. 6 demonstrate the most frequent lexemes in the lexical field studied (10 or more attestations in a given corpus). As *COHA* is a diachronic corpus, quantitative analysis of attestations gathered from it has been done without statistical comparison to any other corpus, while the data from *COCA* and the *BNC* have been normalized per 1,000,000 words and compared.

Fig. 5 reveals that *red [wine]* and *white wine* are the significantly prevailing terms. They occur in *COHA* more than 500 times. Other lexemes (most of French origin) occur less than 200 times in that corpus. *Sweet wine* and *scuppernong (wine)* also occur there with some frequency, as does *Riesling* (a loanword from German).

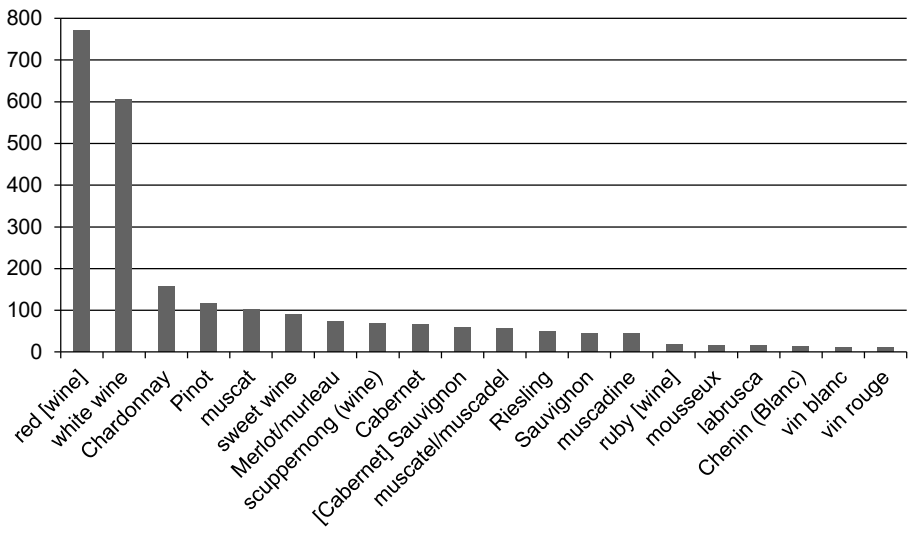


Fig. 5. Lexemes with attestation frequencies of 10 or higher in *COHA*

Fig. 6 presents comparisons of the lexemes in the lexical field WINE in *COCA* and the *BNC*, in an attempt to determine whether any relationship in lexeme usage between American English and British English exists.

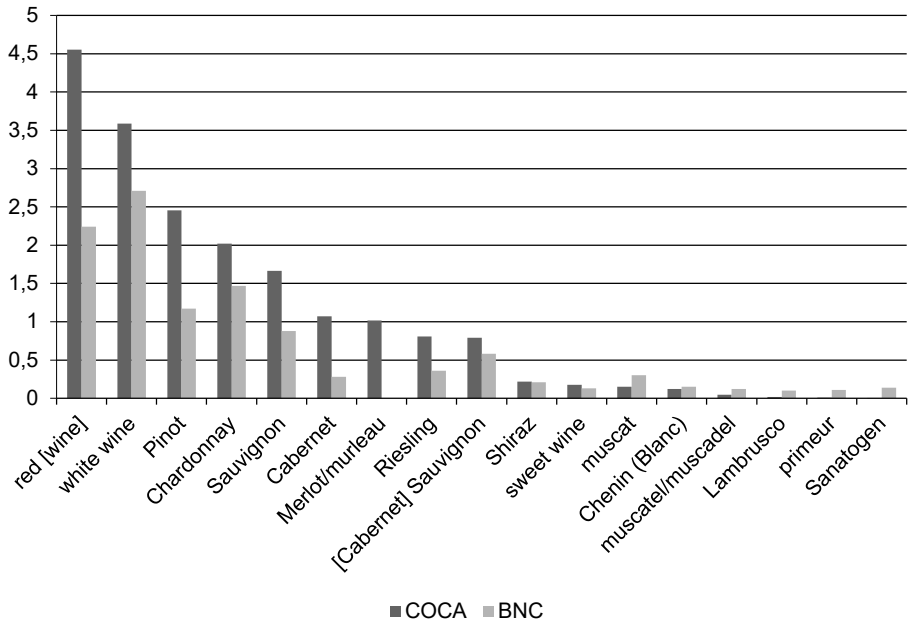


Fig. 6. Comparison of lexeme occurrences in *COCA* and the *BNC* (results normalized)

The chief terms in COCA are the same as those which predominate in COHA: *red [wine]* and *white wine*. The other lexemes, mainly of French origin, occur approximately as often in COCA as they do in COHA. The lexemes which occur in both COHA and COCA more than ten times are: *red [wine]*, *white wine*, *Pinot*, *Chardonnay*, *Sauvignon*, *Cabernet*, *Merlot/murleau*, *Riesling*, *[Cabernet] Sauvignon*, *sweet wine*, *muscat*, *Chenin (Blanc)*, *labrusca*, *muscatel/muscadel* and *scuppernong (wine)*.

As in the cases of COHA and COCA, in the BNC the most prevalent terms are *white wine* and *red [wine]* though in that corpus a pattern contrary to the one seen in the other two corpora is apparent; in the BNC, *white wine* occurs more frequently than *red [wine]*, not the opposite. The lexemes which occur more than 10 times in both COCA and the BNC are: *white wine*, *red [wine]*, *Chardonnay*, *Pinot*, *Sauvignon*, *[Cabernet] Sauvignon*, *Riesling*, *muscat*, *Cabernet*, *Shiraz*, *Chenin (Blanc)*, *sweet wine*, and *muscatel/muscadel*. *Sanatogen*, *primeur*, and *lambrusco* occur more than 10 times only in the BNC.

The above analysis suggests that the lexemes which are attested in more than 10 occasions in COHA, COCA, and the BNC might be called the prototypical WINE field lexemes in both American English and British English. Those lexemes are: *red [wine]*, *white wine*, *Pinot*, *Chardonnay*, *Sauvignon*, *Cabernet*, *Riesling*, *[Cabernet] Sauvignon*, *sweet wine*, *muscat*, *Chenin (Blanc)*, and *muscatel/muscadel*.

This comparison of data from COCA and the BNC demonstrates that more of the higher frequency lexemes of the lexical field WINE (*red [wine]*, *white wine*, *Pinot*, *Chardonnay*, *Sauvignon*, *Cabernet*, *Merlot/murleau*, *Riesling*, *[Cabernet] Sauvignon*, *Shiraz*, and *sweet wine*) occur in greater numbers in COCA than they do in the BNC, and that more of the lower frequency lexemes (*muscat*, *Chenin (Blanc)*, *muscatel/muscadel*, *lambrusco*, *primeur*, *Sanatogen*) occur in greater numbers in the BNC than they do in COCA. This seems to indicate that the prototypical lexemes of the field WINE are more popular in American English than in British English (if not also that the wines denoted are more popular in America than in the British Isles).

## 9. Conclusions

This study has investigated the history of the lexical field of WINE, and a number of etymological, cultural and other patterns are indicated. The development of lexical fields involves borrowing, which occurs when languages are in contact. The language contact which affected most the development of the English lexical field WINE resulted from wine trade



between Britain and other countries. According to Estreicher (2006: 66), Ludington (2013: 24), and Spahni (2000: 73), France has been the main importer of wine to Britain. Present findings support that claim by showing that French has been the donor language of the majority (29%) of the wine-related lexemes examined. Italy, Spain, Germany and Portugal are also enumerated as significant importers by those researchers (Spahni 2000: 73; Ludington 2013: 24; Estreicher 2006: 66). This too is confirmed in the present data representing the English lexical field WINE, wherein relatively large numbers of loans from the primary languages of those countries appear. Spahni (2000: 73) also mentions that Australia, South Africa and Chile were notable wine importers to Britain in 1997. Here again the data collected, especially those categorized as of the combination *lexeme denoting nationality + wine* and culled from COCA and the BNC, substantiate a previous assertion regarding external factors. Lastly, the patterns in the comparative quantities of the combinations *lexeme denoting nationality + wine* and the combinations featuring adjectives denoting the language origins of the borrowings correlate so as to indicate that these loans have been borrowed in response to the cultural phenomena of wine making and wine trading, which have triggered the borrowing process, as the English lexical field of WINE has developed over the centuries.

## REFERENCES

### Sources

#### (a) Dictionaries:

*The Historical Thesaurus of English*, version 2.0. 2013. Glasgow: University of Glasgow.

<http://historicalthesaurus.arts.gla.ac.uk/>, accessed March 2013

*Oxford English Dictionary Online* (OED online)

<http://www.oed.com/>, accessed March 2013

*Oxford Dictionaries*

<https://www.oxforddictionaries.com/>, accessed March 2013

#### (b) Corpora:

*Corpus of Historical American English* (COHA) Compiled by M. Davis.

<http://corpus.byu.edu/coha/>, accessed March 2013

*Corpus of Contemporary American English* (COCA) Compiled by M. Davis.

<http://corpus.byu.edu/coca/>, accessed March 2013

*British National Corpus* accessed March 2013 (BNC) Compiled by M. Davis.

<http://corpus.byu.edu/bnc/>, accessed March 2013

## Special studies

- Crystal, D.  
1997 *A Dictionary of Linguistics and Phonetics*. Oxford: Blackwell.
- Eckardt, R. – K. von Heusinger – C. Schwarze (eds.)  
2003 *Words in Time: Diachronic Semantics from Different Points of View*. Berlin: Mouton de Gruyter.
- Estreicher, S.K.  
2006 *Wine from Neolithic Times to the 21<sup>st</sup> Century*. New York: Algora.
- Geeraerts, D.  
1997 *Diachronic Prototype Semantics: A Contribution to Historical Lexicology*. Oxford: Clarendon.
- Langacker, R.W.  
1987 *Foundations of Cognitive Grammar. I, Theoretical Prerequisites*. Stanford: University of Stanford.
- Lehrer, A.  
1985 "The Influence of Semantic Fields on Semantic Change". In: J. Fisiak (ed.) *Historical Semantics – Historical Word-Formation*. Berlin: Walter de Gruyter, 283-296.
- Ludington, C.  
2013 *The Politics of Wine in Britain: A New Cultural History*. Basingstoke: Palgrave Macmillan.
- Spahni, P.  
2000 *The International Wine Trade*. Cambridge: Woodhead.
- Wyler, S.  
1992 *Colour and Language: Colour Terms in English*. Tübingen: G. Narr.

## APPENDIX

### Unmodified entry **Types of wine** from the *HTE*

Category	Term	Date of occurrence
1	2	3
01 <i>red wine</i>	red	c1386-
	ruby	1671; 1859
	vino tinto	1673; 1901-
	red fustian	a1700 ( <i>Cant dict.</i> ); 1834
	vinho tinto	1835-
	vin rouge	1917-
	vino rosso	1949-
	rouge	1957-
	tinto	1958-
	vino nero	1968-

1	2	3
<b>02</b> <i>white wine</i>	white wine	1377-
	white	c1386-
	vin blanc	1792-
	vinho branco	1835
	vino blanco	1981-
<b>03</b> <i>pink wine</i>	rose	c1460
	rosé	1897-
	pink wine	1909- ( <i>slang</i> )
	pink	1928- ( <i>colloq.</i> )
	vin rosé	1931-
	oeil-de-perdrix	1971-
<b>04</b> <i>sparkling wine</i>	vin mousseux	1789-
	mousseux	1861-
	sparkler	1868-
	méthode/methode champenoise	1928-
	spritzig	1968-
<b>05</b> <i>twice-fermented wine</i>	reboil	c1460
<b>06</b> <i>dessert-wine</i>	dessert-wine	1773-
<b>07</b> <i>liqueur-wine</i>	liqueur-wine	1872
<b>08</b> <i>sweet wine</i>	ceren	OE
	honey-tear < hunigtear	OE-a1240 ( <i>also fig.</i> )
	sweet wine	c1386-
	bastard	1399-1631; 1869 ( <i>History</i> )
	bastard wine	1436-1616
	brown bastard	1603; 1609
	dulce	1844-
	vino dolce	1902-
	vino dulce	1911-
	vin doux (naturel)	1958-
<b>09</b> <i>dry wine</i>	vino secco	1911-
<b>10</b> <i>flat/sour wine</i>	vappa	1629-1840
	vappe	1660
<b>10.01</b> <i>renewed by must</i>	stum	1664-1746

1	2	3
<b>11</b> <i>new wine</i>	ceren	OE
	must	OE
	neowe win	OE
	vin fou	1833-
	vinho verde	1958-
	primeur	1973-
<b>12</b> <i>old wine</i>	beeswing	1860-
<b>13</b> <i>strong wine</i>	high wine	c1384-
	nit	a1700 ( <i>cant</i> )
<b>14</b> <i>ceremonial wine</i>	guild wine	1597
	wine of honour	1706
	vin d'honneur	1947-
<b>15</b> <i>wine served in loving cup</i>	love-wine	1641
<b>16</b> <i>wine drawn from the wood</i>	pipe-wine	1598
<b>17</b> <i>wine made from frozen grapes</i>	Eiswein	1963-
<b>18</b> <i>medicinal wine</i>	mandrake wine	1621-1753
	viper-wine	1631-
	wine whey	1769-1856
	mandragora	1844
	tonic wine	1899-
	Sanatogen	1924-
<b>19</b> <i>wine from specific grapes</i>	muscatel/muscadel	c1400-
<b>19.01</b> <i>muscat</i>	muscadine	1541- ( <i>now History</i> )
	muscat	a1578 ( <i>Scots</i> ); a1747-
	Syracuse	1768-
	lunel	1770-1841
<b>19.02</b> <i>Cabernet Sauvignon</i>	Sauvignon	1907-
<b>19.03</b> <i>Merlot</i>	Merlot/murleau/merlot	1978-
	Merlau	a1997
<b>19.04</b> <i>others</i>	Pineau	1763-
	hanepoot	1804-
	scuppernong (wine)	1825-
	Nebbiolo	1833-
	Riesling	1833-

1	2	3
<b>19.04</b> <i>others</i>	Sauvignon	1846-
	Traminer	1851-
	Catawba	c1857-
	Sémillon	1875-
	Bual	1882-
	verdelho	1883-
	Grignolino	1894-
	mavrodaphne	1911-
	aligoté	1912-
	Pinot	1912-
	Chenin (Blanc)	1928-
	Verdicchio	1940-
	Sylvaner	1958-
	Chardonnay	1959-
	Gamza	1959-
	Cabernet	1961-
	Shiraz	1966-
	Lambrusco	1986-
	labrusca/Labrusca	1988
<b>20</b> <i>wine in natural state</i>	vino crudo	1673; 1833
<b>21</b> <i>wine used for blending</i>	vino de color	1851-
	basis wine	1905-
	vino maestro	1911-
<b>22</b> <i>wine from grape refuse</i>	raspe wine	1600
	rape (wine)	1600-1733
	beverage	1627-1721
	piquette	1688-
<b>23</b> <i>other types of wine</i>	hæcine	OE
	reamwin	OE
	roche	a1400
	tyre	1429-1587
	rospeys	c1440
	raspis	c1460-1584
	caprik(e)	c1460-1587
	rospyse	1465

1	2	3
<b>23</b> <i>other types of wine</i>	Mountrose	c1475
	garnade	c1475-c1481
	whippet	c1500 (2)-
	rampion	1519
	romanisk	1542
	Roberdavy	1542-1620
	raspis wine	1562-1662
	charneco	1593-1631
	pitch-wine	1601
	wine of astonishment	1611
	deal	1613-1635
	deal-wine	1616
	whippincrust	1616
	Sherant	1622-1623
	Sheranino	1632
	Massic	1638-
	Greco	1644;1645
	Langoon	1674-1750
	raspy wine	1703
	rasped wine	1823
	straw wine	1824-
	vin de paille	1833-
	vin jaune	1833-
	vino tierno	1911-
	rancio/Rancio	1939(2)
	boerwyn	1947- (S. Afr.)

Address: PAULINA OŁOWNIA, Instytut Filologii Obcych, Uniwersytet Jana Kochanowskiego, ul. Świętokrzyska 21D, 25-406 Kielce, Poland.

ORCID Code: 0000-0002-5467-433X.



Printed and bound by  
Jan Kochanowski University of Kielce Press.

ELISABETTA LONATI: The language of medicine in the <i>Philosophical Transactions</i> : Observations on style	5
PATRICE LARROQUE: The dialogic nature of nonstandard English	25
ANNALISA BONOMO: English in multilingual education programs: some cases and perspectives	41
MAYOWA AKINLOTAN: Genitive alternation in New Englishes: The case of Nigerian English	59
JULIA SCHULTZ: The impact of Arabic on the English lexicon since 1801	75
DAISUKE SUZUKI: <i>Maybe</i> : Development and Topic Marking	93
PAULINA OŁOWNIA: The lexical field WINE in American and British English: A corpus-based study	109

*Token: A Journal of English Linguistics* focuses on English linguistics in a broad sense, and accepts both diachronic and synchronic work, grammatical as well as lexical studies. That being said, the journal favors empirical research. Jan Kochanowski University (Kielce, Poland) publishes *Token* once annually, and all submissions are double-blind peer reviewed. The journal's website, where published articles are freely accessible, is at <http://www.ujk.edu.pl/token/>.