

Manfred Krug and Christopher Lucas*

Definite article (omission) in British, Maltese, and other Englishes

<https://doi.org/10.1515/stuf-2018-0012>

Abstract: This article investigates factors that underlie the discrepancies in article omission between Maltese English (MaltE) and British English (BrE), with reference to further ENL, ESL and EFL varieties. We investigate seasons of the year, ordinal numbers, languages, proper nouns, titles, institutions and common nouns. Our sources include text corpora, and web and questionnaire-based data. Our key proposal is that MaltE has innovated a rule that the definite article may be omitted when the uniqueness or identifiability of a referent is salient in context. Furthermore, MaltE avoids the definite article commonly when the referent is generic rather than definite. The resulting MaltE system is regulated according to fewer parameters than in BrE, but more consistently.

Keywords: Maltese English, British National Corpus (BNC), International Corpus of English (ICE), article omission, zero article

1 Introduction

As is well known, definite articles typically derive from adnominal demonstratives (e.g. Himmelmann 1997). A frequent suggestion in the literature on this development has been that obligatoriness, suitably defined, is a key criterion for deciding whether or not a demonstrative in a particular language has made the transition to definite article status (e.g. Greenberg 1978; Crisma 2011). Against this background, definite article usage in English represents an interesting case. While the article, or some other morphological marker of definiteness, is clearly obligatory in core definite referential contexts, there are also numerous contexts in which its appearance is optional – often contexts which do not allow for optionality in other European languages with definite articles. This is especially clear in the Maltese variety of English, in which article omission is extended to

*Corresponding author: Christopher Lucas, School of Languages, Cultures and Linguistics, SOAS University of London, Thornhaugh Street, Russell Square WC1H 0XG London, UK, E-mail: cl39@soas.ac.uk.

Manfred Krug, English and Historical Linguistics, University of Bamberg, 96047 Bamberg, Germany, E-mail: manfred.krug@uni-bamberg.de.

contexts and frequencies that are unheard of in norm-providing inner-circle English varieties, such as standard British or American English.

The purpose of this article is to shed light on the factors that underlie the discrepancies in article omission between Maltese English and British English, in the context of other varieties of English around the world. The key proposal is that Maltese English has innovated a rule, not shared by British English, that the article may be omitted whenever the uniqueness or identifiability of the referent of the relevant noun phrase (NP) is highly salient in context.

This contribution is structured as follows. Section 2 gives an introduction to Maltese English and the materials currently being developed to study it, especially the Malta component of the International Corpus of English (ICE-Malta). Section 3 provides an overview of article optionality in British and Maltese English, with reference also to other Englishes. Section 4 then presents a detailed investigation of article omission in referential uses of *government* and *cabinet* in the 100-million-word British National Corpus (BNC). Here it is argued that article omission with these and related items is a consequence of certain speakers choosing to treat these items as proper names. In Section 5 we show that article omission is much more frequent with *government* and *cabinet* in Maltese English, and that it also occurs in ICE-Malta with a number of other items for which it is unattested in British English. We explain this discrepancy by positing the aforementioned innovation: that in Maltese English the definite article may be omitted whenever it is not necessary to signal the uniqueness or identifiability of a particular NP. In Section 6, this analysis is lent further support by (a) the findings of a small-scale questionnaire-based investigation of Maltese English speakers' intuitions regarding the acceptability of article omission with a range of items, as well as (b) the results of a larger-scale questionnaire project comparing British and Maltese spoken and written English. Section 7 considers the question of what might have motivated speakers to make this innovation, showing that transfer from Maltese cannot be the answer, and that internal pressures towards greater economy and consistency of expression represent a more likely source. Section 8 concludes.

2 Maltese English and sources of the present study

In the sections focusing on Maltese English, we describe acrolectal English as spoken and written by university-educated Maltese who have spent their lives

mostly on the archipelago of Malta.¹ This is in line with the text collection principles laid out for the components of the International Corpus of English (ICE), whose Maltese component (ICE-Malta) is being compiled at the University of Bamberg, with considerable support from researchers at the University of Malta. Each 1-million-word ICE component consists of 400,000 words of written and 600,000 words of spoken texts, with individual text samples containing 2,000 words. The distribution of genres and text types (see Table 1 below) follows the principles laid out in Nelson (1996).

The material analyzed for the present study consists of approximately 150,000 words of spoken MaltE. This is a transitional corpus of texts which have not yet been converted into their final ICE format (missing anonymization and proofreading). The texts (overwhelmingly recorded in the early 2000s) include University of Malta Campus Radio recordings, typically interviews with academics from all fields, with a Maltese English host; TV Malta News programs; European Parliament hearings and speeches of Maltese members and candidates; classroom lessons; discussions among students, and monologues produced by Maltese tourist guides.

In addition, we use some 100,000 words of written MaltE press language. For the present study, the press report and editorial material of ICE-Malta has been expanded to 100,000 words for an extra-ICE newspaper corpus consisting of 60,000 words of press reportage and 40,000 words of editorials. For comparison, a parallel press corpus was compiled for British English (BrE), the historical norm-providing standard variety of the former colonial power, whose exonormative influence, however, is under pressure due to endonormative and globalizing tendencies in MaltE, as recent research consistently shows (cf. Bonnici 2010; Krug and Rosen 2012; Grech 2015; Krug et al. 2016; Krug and Sönning 2018). The UK corpus consists of material from *The Times*, *The Guardian* and *The Independent*. For convenience, we will refer to these two corpora from now onwards as *Press_Malta* and *Press_UK*. All texts in these corpora date from 2008 and 2009. The exact figures and newspaper sources are given in Table 2. The corpus design is modeled on that of the press section of the International Corpus of English (see Table 1). Following the principles used in the ICE-Malta compilation, in the Maltese subcorpora preference was given to texts dealing with local topics and to texts written by journalists with local-sounding last names.

¹ A more detailed discussion of the varietal range of Maltese Englishes can be found in Krug (2015).

Table 1: ICE corpus design (figures in brackets indicate the number of 2,000-word texts in each category).

Spoken (300)	Dialogues (180)	Private (100)	Conversations (90)
			Phone Calls (10)
			Public (80)
		Class Lessons (20)	
		Broadcast Discussions (20)	
	Monologues (120)	Unscripted (70)	Broadcast Interviews (10)
			Parliamentary Debates (10)
			Cross-Examinations (10)
		Scripted (50)	Business Transactions (10)
			Commentaries (20)
Written (200)	Non-printed (50)	Student Writing (20)	Unscripted Speeches (30)
			Demonstrations (10)
			Legal Presentations (10)
		Letters (30)	Broadcast News (20)
			Broadcast Talks (20)
	Printed (150)	Academic (40)	Non-Broadcast Talks (10)
			Student Essays (10)
			Exam Scripts (10)
		Popular (40)	Social Letters (15)
			Business Letters (15)
Reportage (20)	Instructional (20)	Humanities (10)	
		Social Sciences (10)	
		Natural Sciences (10)	
	Persuasive (10)	Technology (10)	
		Humanities (10)	
Creative (20)	Persuasive (10)	Social Sciences (10)	
		Natural Sciences (10)	
			Technology (10)
			Press Reports (20)
			Administrative Writing (10)
			Skills/Hobbies (10)
			Editorials (10)
			Novels (20)

As the vast majority of both the spoken and written texts used for this analysis will eventually form part of the Maltese component of the International Corpus of English, we will, for simplicity and convenience, call this transitional corpus *ICE-Malta_2017* in the remainder of this study when we refer to the total of spoken and written texts. Spoken and written subcomponents will be referred to individually as *spoken ICE-Malta_2017* and *Press_Malta*, respectively.

Table 2: Newspaper corpora (number of words and sources).

Variety	Press category	# words	Sources
Maltese English	News reports	60,728	<i>Malta Today, The Malta Independent Online, Times of Malta, Business Today</i>
	Editorials	40,254	
	<i>Total Press_Malta</i>	100,982	
British English	News reports	61,447	<i>The Guardian, Times Online, The Independent</i>
	Editorials	40,348	
	<i>Total Press_UK</i>	101,795	

For the analysis of contemporary British English, we will also use the 100m-word British National Corpus (BNCweb version, see Hoffmann et al. 2008), which contains some 90m words of written and 10m of spoken texts. The spoken words fall into two roughly equal components of more formal (so called context-governed) texts and less formal texts (spontaneous conversations). More detailed information on compilation principles, text origin and text types can be obtained from Aston and Burnard (1998).

Apart from the corpora mentioned above, two questionnaire-based sets of data will be studied. One small-scale questionnaire on definite article use and omission was specifically designed for the present study (for detail see Section 6 below). In addition, we will analyze a number of items related to article usage which figure in the Bamberg Survey of Morphosyntactic and Lexical Variation. Designed in 2008 at the outset of the ICE-Malta project and partly informed by the 76 features used by Kortmann et al.'s *Handbook of varieties of English* (2004), the Bamberg questionnaire also contains various features reported for, or thought to be of interest for, a larger project on insular and peninsular contexts of Romance–English language contact (e.g. Malta, Channel Islands, Gibraltar, Puerto Rico; see Krug and Sell 2013 for an outline; Krug et al. forthcoming for the full questionnaire). The Bamberg questionnaire elicits information on usage frequency for 138 grammatical items for an informal spoken register (an informal conversation among friends). All spoken items are recordings of a native speaker of the variety under investigation; during the questionnaire session each item's audio file is played individually (twice), followed by a brief pause for rating. 200 items are subsequently rated for frequency of usage in a moderately formal, written register (an email to a former school teacher). The 200 written items include all 138 spoken items so that stylistic differences can be investigated. We will analyze ten sentences relating to definite article use and omission in Section 6.

3 Overview of article omission in British and Maltese English, and other varieties of English around the world

In the following sections, we will see that there are two grammatical contexts in which definite article usage is variable and thus deserves detailed discussion:

- (i) contexts in which the usage of a definite marker is not justified on logical grounds because a noun phrase is not semantically definite (but generically used, for instance);
- (ii) contexts in which the definite article may be logically redundant and could thus be dispensed with for economy reasons (as is common with names).

As it is impossible within the scope of this contribution to discuss exhaustively the rules of definite article use and omission in standard Englishes (for an overview, see e.g. Quirk et al. 1985: 265–297), we will focus on areas in which use and non-use vacillate in interesting ways. It is probably no coincidence that all examples of the second type above (Sections 3.2ff.) stem from the domain of (or are related to) proper nouns, which have both obligatory and optional rules for the regulation of article use.

3.1 Seasons of the year

Let us first look at type (i) sketched above, i.e. contexts that, logically, require no explicit grammatical definiteness marking. In standard BrE and AmE, the seasons of the year occur in free variation with or without the definite article when a season is used in a generic sense, i.e. without reference to a specific year (cf. Quirk et al. 1985: 278–279):

- (1) in summer/winter/autumn (AmE fall)/spring ~ in the summer/winter/autumn/spring

BrE (as represented in the British National Corpus) uses the definite article in the clear majority of cases, even though no definite summer or winter is being referred to. In quantitative analyses of the spoken BNC, seasons with the definite article outnumber article-less instances by a margin of about 5:1 when following the preposition *in*: in absolute numbers, we find 368 instances of *in the* + season vs. 68 instances of *in* + zero article + season (with all four seasons consistently favoring article-full usage; for further detail, see Krug 2015). In our MaltE corpus, by contrast,

we find substantially higher proportions of seasons without the definite article. For example, we have 49 tokens of article-less *summer* following a preposition, and 15 tokens of *summer* with the definite article following a preposition. An example of an article-less use from spoken ICE-Malta_2017 is given in (2).

- (2) And the reason for that is uhm because if you buy a cheap house which is rated F, that means it's uh uh it has thin walls, it has single glass window panes its, its roof is very slim, it's too hot in summer and too cold in winter, you will end up spending a lot of money on your electricity and energy bills throughout the lifetime of that building.

We conclude, therefore, that for the domain of generic uses of seasons, which in standard Englishes exhibit free variation (and in BrE a preponderance of definite articles), Maltese English has developed a more consistent system: generic use (i.e. non-definiteness) tends not to be marked by the definite article.

3.2 Ordinal numbers

We now turn to the second type outlined at the beginning of Section 3. In default everyday contexts (i.e. in the absence of competing hierarchies), noun phrases containing modifying ordinal numbers (*first*, *second* etc.) are unambiguous, in fact unique and thus inherently definite. From a strictly logical perspective, therefore, an additional definite article is redundant. Nevertheless, in the major norm-providing varieties, usage with the definite article is fully grammaticalized when ordinal numbers occur in prepositional phrases to function as adverbials.

Our analysis focuses on common multi-word sequences involving ordinals in both British English and Maltese English: *first/second/third* etc. immediately followed by *year*. In what follows, we will quantitatively analyze definite-article variation in the prepositional phrases *in first year* vs. *in the first year*. In the entire British National Corpus there are a total of 413 instances of *in the first year*. Only four of these (i.e. less than 1%) occur in compound adjectives like *in the first year lecture*. These need to be excluded from the current count as compound adjectives are not the focus of our present investigation.² The overwhelming

² The remaining three are:

- (1) The work mostly took place in the first year base, which was described as ... [BNC EW7]
 (2) Students who obtain passes in the first year subjects but do not intend to complete ... [BNC HBN]
 (3) Damian's twelve. He's in the first year Wankers and Playboy's under his bed ... [BNC KPL]

majority of the remaining 409 relevant *first year* occurrences function as (parts of) time adverbials, e.g.:

- (3) It is not available in the first year of the policy and in the second year only £250 – about two hours' worth for a top London matrimonial lawyer – will be met for each party. [BNC, A3J]
- (4) We have almost as many girls as boys in the first year in the sixth form, for instance, ... [BNC, KRG]
- (5) It'll save us a hundred thousand in the first year. [BNC, KRL]
- (6) In the first year students take Housing Studies 1, Social Policy 1 and another social science course. In the second year they continue with Housing Studies 2, Social Policy 2 and a further social science course. At the end of the second year there is an eight week supervised placement with a housing association or housing co-operative. In the honours years, students concentrate more on subjects which are directly related to the professional qualification in housing. In the third year they will study Housing Policy and Practice, Housing Economics and Finance, Housing Law, Research Methods and Management Studies. ... In the final year, students complete the study of Housing Policy and... [BNC HTE]

In the 100m words of the BNC, there exist also 15 instances of *in first year*, i.e. without the definite article. Again, of these 15 examples, the following four occur in compound adjectives and therefore have to be excluded, especially since in such contexts definite articles effect a meaning change:

- (7) to cover material usually taught in first year cell biology courses, [BNC B7G 1925]
- (8) and we currently anticipate a further large increase in first year entrants. [BNC HCG 131]
- (9) She also points to the importance of performance in first year examinations as a means of identifying potential non-graduates, [BNC HX1 80]
- (10) Well, I must have been in first year juniors. [BNC KCH 3165]

Most relevant examples are to be found in more formal registers, often technical, institutional or fictional writing, e.g.:

- (11) In Information Sources in first year, the Information Science base is retained. [BNC CG9 66]
- (12) The tallest, heaviest boys were in the A group; the next size down was B and so on until F. F was the group for the real tinies, the handful of Alec Davidsons who were mostly in first year. [fiction, BNC ASN 111]

While the 409 relevant examples with definite article (*in the first year*) include university and school contexts, it is striking that article-less examples in the 100m words of the BNC of both simple adjectival ordinals and ordinal numbers in compound adjectives occur exclusively in the context of educational institutions.³ British English therefore has little variation and uses the article-less form in educational contexts (if at all). MalTE apparently omits the definite article in such contexts proportionately more often than BrE, though not categorically: out of seven tokens including *in (the) first year* in ICE-Malta_2017, for instance, five are article-less. The following extract from a Malta Campus Radio interview exemplifies this variability:

- (13) The first year you know they are all new to the system yes finding their system, you don't really want to overload them. But from second year, especially third year, and fourth year I think it is very important to...

That the same constraint operates in Maltese English as in BrE is suggested by our small-scale questionnaire investigation (see Section 6 below). On a scale from 0 (=No one) to 5 (=Everyone), the university-context example receives considerably higher usage ratings than that referring to young children:

- (14) Second year of university is usually more intense. (average rating: 4.0)
- (15) Children usually start walking in second year of life. (average rating: 2.8)

³ Sharma (2005: 558–560) finds in her data (from 12 first-generation adult immigrants in California of Indian descent) that quantifier modification of this kind has high rates of zero articles. She does not establish a correlation with institutional contexts, though.

3.3 Nominal and adjectival uses of language names

The use of the definite article with nominalized language adjectives is fairly strictly regulated in norm-providing varieties belonging to Kachru's inner circle (see e.g. Kachru 1988: 5). When the article is used (as in example (16), *the French*), the nominalized adjective refers to the people of a country in a generic sense. When referring to the language, there is no definite article (17), unless the nominalized adjective is further specified and postmodified by prepositional phrases (18) or (elliptical) relative clauses (19).

(16) The French (i.e. people) love wine.

(17) He studied French.

(18) He studied the French of Louis XIV.

(19) He studied the French (that was) used in the 17th century.

The matter becomes more complex with other types of postmodification. Consider *as* clauses, which have a certain affinity with relative clauses – *as* exists as a relative pronoun in older English and some regional dialects:

(20) He studied (?the) French as it was used in the 17th century.

(21) He studied (the) French as used in the 17th century.

Quirk et al. (1985: 287), too, mention “expressions in which the definite article is optional with the name of a language.” Their examples include:

(22) a word borrowed from (the) French/(the) Italian

(23) examples from (the) Sanskrit/(the) Hebrew

(24) He has the French.

Example (25) is considered archaic for most modern dialects by Quirk et al. (1985: 287) when meaning ‘he knows French and English’, and would today be usually understood with the curricular meaning of school or university subjects.

(25) He has French and English.

The present authors feel that language change indeed plays a role in the above examples and that the frequency and acceptability of examples (22) to (24) has further decreased since the 1980s. To us, “a word borrowed from the French” seems to border on the ungrammatical, with the only clearly acceptable examples figuring in the context of translation, e.g.:

(26) translated from the French by [Name]

When language names are used adjectivally, there is variation not only in regional dialects but also in standard Englishes. For spoken informal English, questionnaire data from the Bamberg Survey of Morphosyntactic and Lexical Variation (see Krug et al., forthcoming for a detailed description) yield rather high usage ratings for sentence (27) below not only in Malta, but also in the US and the UK:

(27) Maria has a very good knowledge of Italian language.

From 2008 to 2010, our questionnaire informants from the varieties mentioned (BrE, AmE, MaltE) rated this example on average as “Could be said by many people” (i.e. two categories below the top category “everyone”, but only one below what is usually the highest average rating actually chosen by participants taking part in the questionnaire, i.e.: ‘most people’; see Section 6 below for the full scale and numerical detail). More surprisingly, in our US data, questionnaire usage ratings did not even drop significantly from informal spoken to semi-formal written English. This is different for our BrE questionnaire data, where informants report markedly lower usage ratios for the genre ‘email to a former school teacher’. We conclude therefore that, rather than exhibiting nonstandard syntax in the case of definite article omission with language adjectives, MaltE seems to be taking part in a global trend: in 2008/9, MaltE occupies a position intermediate between AmE and BrE for informal spoken English, while siding with BrE in differentiating stylistically.

Section 6 below focuses on different types of article variation and on Maltese data from 2013. Here we present data from 2008/9 as well, so that we can consider in more detail the figures relevant to language adjectives and their short-term trends. The relevant figures and usage ratings are displayed in Table 3 below.⁴

In 2008/09, a small degree of stylistic variation (of 2.5 in the written vs. 2.8 in the spoken register) is observable in MaltE; and the situation is parallel with that of

⁴ The number of responses from Maltese informants in 2008/09 varies considerably because fewer than half of the informants who participated in the spoken part also completed the written section, which was collected separately. Due to such low rates, we soon changed the procedure and conducted both parts of the questionnaire in a single session.

Table 3: Average usage ratings by Maltese and British university students for a semi-formal written and an informal spoken register.

		Maltese English				British English	
		2013		2008/9		2010	
		Written	Spoken	Written	Spoken	Written	Spoken
(i)	Maria has a very good knowledge of Ø Italian language.	2.9 (n = 69)	3.0 (n = 74)	2.5 (n = 48)	2.8 (n = 127)	2.4 (n = 33)	2.6 (n = 33)
(ii)	Some students spent two weeks in families, enjoying Ø American way of life.	2.0 (n = 67)	2.1 (n = 72)	1.6 (n = 48)	2.1 (n = 140)	0.9 (n = 33)	1.8 (n = 33)

the former exonormative BrE standard, where usage ratings in 2010 are 2.4 vs. 2.6, respectively. However, the data suggest a narrowing stylistic gap for MaltE: from 0.3 (in 2008/09) to 0.1 in 2013. From 2008 to 2013, usage ratings increase in both MaltE registers, but, more importantly, converge on almost exactly the former spoken value (at 2.9 and 3.0 for the written and spoken genre, respectively).

If this turns out to be a robust scenario over the longer term, we are in the midst of an interesting endonormative development which is relevant for the study of world Englishes more generally: for one, while *increased* stylistic differentiation in postcolonial varieties is normally taken to be an indicator of increasing nativization and later developmental stages (Schneider 2007: 100), on the basis of our data we can at least state the theoretical possibility of a different scenario: the *loss* of stylistic differentiation can evidence linguistic nativization and endonormative stabilization as well. Secondly, increasing usage ratings that converge on (or even exceed) erstwhile spoken values can be well integrated within the commonly observed pattern of colloquialization of written norms (cf. Biber and Finegan 1992; Mair 1995).

The examples below (all from spoken ICE-Malta_2017, all from radio interviews with academics) may serve to illustrate the variation found with language adjectives in standard Englishes around the world:

- (28) Now the ideal, the idea is to introduce and or enhance the teaching of Chinese language and culture.
- (29) Malta is different from the UK and from Greece, this doesn't really take away also that even Maltese culture is becoming very quickly multicultural

- (30) yet that boy's talents and fluency in Italian language cannot be assessed
- (31) Okay so uhm when we look at Arabic language you have the script that's different

A zero article would be normal in (28) and (29), less likely in (30), and out in (31) for most speakers in formal English. As we will see in a related fashion for institutions, functions and titles, when an entity becomes name-like, the omission of the definite article becomes more likely. When language adjectives form part of a name-like entity or a phraseological unit in certain contexts (e.g. a university degree or course), article omission becomes obligatory in both norm-providing varieties and MaltE, as can be seen in the next example:

- (32) He studied (*the) French Language and Literature.

In summary, language adjectives like *Italian* are inherently definite for the ordinary language user. This renders the definite article logically redundant. Norm-providing inner-circle varieties are variable in many of the above-cited adjectival uses and omit the article obligatorily only when the adjective is nominalized (as in *his fluency in Italian*) or part of the name of a field of study (as in *studying French Language and Literature*). Many Maltese English speakers reorganize this relatively free adjectival pattern into a more systematic pattern in which inherent definiteness triggers definite article omission.

We conclude that article omission with language adjectives exists in both British and American Standard English, but is rarer in both these varieties than in Maltese English. The phenomenon of zero articles in learner Englishes is well known generally (cf. Robertson 2000; Zdorenko and Paradis 2008; Zahedi and Mehrnazmay 2011; Sun 2016; Schembri this volume), and we have observed it sporadically for language adjectives in other ESL and EFL varieties as well (e.g. Nepali English, Kurdish English, Chinese English, Polish English). It is clearly not restricted to (but equally clearly very common in) regions where the main contact language has no functional equivalent to the English definite article (as Polish or Mandarin).⁵ A similar, language-related observation was made for Indian English by Sedlatschek (2009: 207), who cites a zero article before an adjective that refers to a country (not a language, though): “the ex-captain of Ø South African Cricket

⁵ Sorani Kurdish, for instance, like Maltese, has a definiteness marker. Nevertheless, in an online article on definiteness in Sorani Kurdish and English, there are four instances of zero articles preceding “Kurdish language” and they occur also before other combinations with *language*, e.g.: “Ø Kurdish language belongs to the Indo-Iranian branch of Ø Indo-European language family” (Zahedi and Mehrnazmay 2011: 130).

Team”. Sedlatschek here also points to capitalization and the link to proper nouns, points to which we now turn and take up again in Sections 4 and 5 on *G/government* and *C/cabinet*.

3.4 Proper nouns and titles

Proper nouns and titles offer another interesting field of variation (cf. Quirk et al. 1985: 291–92). In norm-providing varieties of English, the distribution is clear-cut. When titles precede the name of the title holder, the whole phrase is treated as a proper noun. It receives no definite article and, like first names, the title is regularly capitalized in writing. Typically, only the surname follows the title, but this seems to depend on contextual familiarity. Compare:

(33) Dean [first name or function] Martin

(34) President Obama vs. Director (John) Smith

In norm-providing inner-circle varieties, the definite article is required when the title is not followed by the name of its bearer; in written representations, capitalization of the title is also common here but less consistently used than in the above cases:

(35) the P/president; the D/dean; the R/rector; the D/director; the P/prime M/minister

Exceptions to the above generalizations are appositions, e.g.

(36) the president, John Smith, has issued a statement.

In headlines, both above types occur without the definite article, e.g.:

(37) President due to sign yet another decree

In Maltese English, bare titles are not uncommon without the definite article, when the context is clear, e.g.:

(38) So when Rector went off to China last year to sign the relative agreement (spoken ICE-Malta_2017)

(39) This section is directly accountable to Dean (overheard in university context)

(40) High Commissioner is giving a speech tomorrow (overheard at a reception)

In unambiguous contexts, titles and posts become similar to proper names, which are ungrammatical with the definite article in the major norm-providing varieties of English. In a similar vein, Sharma (2005: 535) finds in her data that speakers of Indian English use zero articles more commonly when the noun phrase's referent can be identified unambiguously from the context. In Malta, uniqueness is in evidence for *rector* or *university* (see also the next section on institutions): at the time of writing, there exists only one university in the country (an American University is planned, though); and only the head of this institution (and not that of a college, for instance) is referred to as *(the) Rector*. As in Section 3.3, therefore, name-like entities in Maltese English can be treated grammatically like proper names, licensing the zero article. It is not inconceivable that common exposure to article-less bare titles in the headlines of press language functions as a contributory force in such a recategorization of name-like entities. Signs like the following (see Figure 1, picture taken at the University of Malta) may exert a similar influence:



Figure 1: Car park photos “Reserved for Director” and “Reserved for Chairman”.

3.5 Institutions

Related to proper nouns are institutions when they occur with place names or are named after persons etc. (e.g. *Oxford University*, *John Hopkins University*,

G. F. Abela Junior College, Mater Dei Hospital). In other cases, norm-providing inner-circle varieties follow the general rule that the article-less form is used when a person visits the institution or building for a default purpose it is designed for. Witness the following contrasts:

- (41) He's in Ø hospital [sc. to be cured of an illness]. (BrE; see below for AmE)
- (42) He's in the hospital [e.g. as a visitor, cleaner or electrician].
- (43) She's at (AmE in) Ø school [as an enrolled pupil].⁶
- (44) She's in the school [e.g. as a visitor, cleaner or electrician].
- (45) He's in/at church [sc. attending service].
- (46) He's in the church [e.g. as a visitor, cleaner or electrician].
- (47) She's at university [sc. as a student].
- (48) He's at the university [e.g. as a visitor, cleaner or electrician].

With regard to *university*, Quirk et al. (1985: 277) note that “the article is optional in BrE in the expressions *be at/go to (the) university*, whereas AmE requires the definite article (as also with *hospital* in AmE: ‘He’s in the hospital.’).”⁷

Specific universities, by contrast, are inherently definite (because they are generally unique) and display interesting patterns of variation in that they can:

- a) regularly be referred to as geographical names by [Place name + *University/College*] in which case *University* is typically optional;
- b) often be referred to also as [*University of* + Place name] and
- c) occasionally be referred to also as acronyms [e.g. *UCL, UCLA*].

Compare the variation typically found in inner-circle varieties and illustrated in (49) to (51) below:

- (49) I taught at the University of Cambridge. vs.
I taught at Ø Cambridge (University).

⁶ For more detail, see Quirk et al. (1985: 277; 677).

⁷ As was suggested earlier, quantitative research seems necessary in order to examine to what extent the claims made in 1985 are still valid today.

- (50) He taught at the University of California, Los Angeles. vs.
He taught at Ø UCLA.
- (51) He taught at the University of Malta. vs.
He taught at Ø UoM.

This yields the following system for norm-providing inner-circle varieties:

- (i) When followed by the prepositional phrase [*of* + place name], universities are preceded by the definite article (e.g. *the University of Bamberg*).
- (ii) When the place name precedes *University*, the compound is treated like a geographical name (e.g. a city, town, country) and occurs without the article (e.g. *Bamberg University*).
- (iii) Acronyms of both above types (i) and (ii) are treated as geographical names and consequently occur without the article (e.g. *UCLA, UoM*).

Like norm-providing inner-circle varieties, MaltE has grammaticalized the zero article for acronyms like *UCLA* or *UoM*, but the full form also often occurs without the definite article, as illustrated below:

- (52) He taught at the University of Malta. ~ He taught at Ø University of Malta.

Compared to the system outlined above, many educated MaltE speakers have therefore seemingly altered rule (i) as follows:

- (i') When *University* is followed by the prepositional phrase [*of* + place name], the entire phrase (e.g. *University of Bamberg*) may be treated like a geographical name and thus occurs without the definite article.

It is noteworthy, however, that in our data we only find relevant examples for *(the) University of Malta*. This restriction might theoretically be due to the origin of our spoken and written corpus texts, and it remains to be seen whether educated speakers of MaltE adopt (i') above at a general level. In the absence of such evidence, however, we assume that familiarity with the context of *(the) University of Malta* turns this particular phrase into a geographical name, which is therefore treated under the label of (ii) above. An apparent resetting of rules in Maltese English may thus be better described as treating the default (and only full) university of the country as a null-marked geographical or institutional name also in its full form with the preposition *of*.

3.6 Article use with *majority of* across varieties of English worldwide

Sedlatschek (2009: 204–211) presents web data for the prepositional phrase *in a/the/∅ majority of*. He finds (Sedlatschek 2009: 210) that Indian internet English in 2003 – which obviously consisted of ENL, ESL and EFL material – displayed the highest proportion of null articles in the analyzed country-specific domains (.uk, .us, .in, .za, .sg, .au); and that at 37.8% (i.e. more than a third of all examples), null articles on the Indian websites even outnumbered the remaining options in relative terms. Sedlatschek’s 2003 figures are given in the Appendix. Replicating Sedlatschek’s investigation, we collected data, via Google searches, on the distribution of the three variants (*in a/the/∅ majority of*) for twelve regional domains in January 2017. Figure 2 illustrates our findings in the form of dot-plot inspired hierarchies. For a comparison of these figures with those of Sedlatschek, see the Appendix (where relative frequencies have been added to his figures to allow for easier comparison with ours).

In each of the three hierarchies depicted in Figure 2, the top row lists regions where English is overwhelmingly spoken as a first or native language (ENL). Each middle and lower row in Figure 2 lists regions where English is predominantly spoken as a second language (ESL) and foreign language (EFL), respectively. In order to identify variety clusters, small dots on the numerical scales represent one variety (listed underneath), with dot sizes increasing if they represent two or three varieties (listed below the larger dots).

As the Appendix shows, the situation has changed over the 14 years from 2003 to 2017. Some interesting developments and patterns emerge. The country-specific domain .in still has a relatively high proportion of null articles, but the Indian (Ind in Figure 2) domain’s relative shares of the three variants have become much more similar to the remaining country-specific domains. It is quite striking just how similar the figures for the null article are: except for the American (USA) .us and the Indian .in domains (where the proportions of zero articles are 0.9% and 0.75%, respectively), all domains of countries where English is spoken as a first or second language feature zero articles in 0.2 to 0.3% of the relevant contexts. At 1.3% for .pl and 1.4% for .cn, we find higher shares of null articles for Polish (Pol) and Chinese (Chin) domains – cases in which the major contact languages have no functional equivalents of the English definite article. A third EFL variety investigated here is German (Germ) English. It features the null article less commonly in relative terms than either of the above learner varieties, which is unsurprising since German has a definite article; at 0.75%, however, the null article is still relatively frequent in German

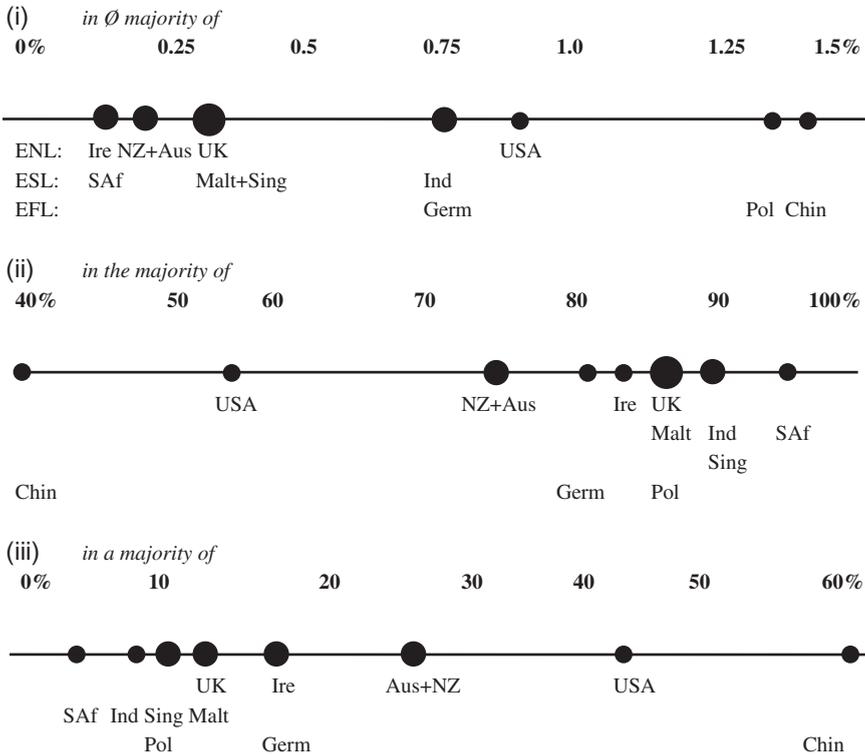


Figure 2: 2017 web data for percentages of phrases *in a/the/∅ majority of*.

internet English and on a par with Indian web English in 2017. In addition, German web English is situated consistently between BrE and American English. This result points to similar language contact intensity with both major norm-providing varieties, as well as to a fairly balanced situation as to whether British or American English should be regarded as the major target variety.⁸ In the present data, Polish web English, by contrast, is unambiguously oriented towards BrE, a pattern that is only violated for the null-article figures (and in that case probably due to the absence of a definite article in Polish, as noted above).

The two focus varieties of the present study, i.e. the country domains for Malta and the UK are almost identical as far as the three investigated variants

⁸ This scenario broadly reflects the situation in German schools, where today both BrE and AmE are usually accepted as target and reference varieties.

are concerned. In both varieties, we find 88% definite articles, 12% indefinite articles and 0.3% null articles. This suggests that current web material for the two varieties is extremely similar, and perhaps deceptively so, since we observe significant qualitative and quantitative differences not only in spoken material, but also in more carefully sampled written texts as well as questionnaire data, as we see in Sections 5 and 6.

Two other web domains are virtually identical for the phrases investigated: those for New Zealand (NZ) and Australia (Aus). Both feature the null article in 0.2%, the definite article in 74% and the indefinite article in 26% of the relevant constructions. Predictably, they figure closer to UK English than US web English. Parallels in settlement and colonial history, regional proximity and dialect contact serve together well to motivate similar linguistic (and rather more ‘British’) behavior. Just how similar the figures for New Zealand and Australia turn out to be in the present investigation, however, is still surprising.

A related finding emerges if we disregard the null article for a moment (for which quantitative differences are rather minor in terms of percentages): all other postcolonial Englishes which were investigated in the present web-based study, i.e. Indian, Singaporean and South African English (as Malta anyway), cluster with or figure closer to the web domain of their former British colonizer than to American web English.

Another noteworthy pattern can be observed for Chinese and South African (SAf) web English. Without exception, these two form the extreme poles in the three scales of Figure 3. Moreover, Chinese web English always figures close to AmE, whereas SAfE always figures closer to UK web English. It seems plausible to assume, therefore, that ChinE and SAfE are oriented towards American and British English, respectively, due to language contact and/or the adoption of one as the dominant norm-giving target variety. Intriguingly, then, both South African and Chinese web English consistently outperform, as it were, their target varieties, and exaggerate AmE and BrE tendencies, at least for definite and indefinite article use with *majority*.⁹ Even though the same pattern also holds for the zero article, we assume that, for the Chinese domain, the structural proximity to US web English here is not due to the fact that American English is the main target variety in China. Instead, a different kind of language contact is more likely the major factor

⁹ We cannot rule out ‘colonial lag’ in the case of SAfE, in which case it would adhere to older British norms. In the absence of historical BrE web material, however, we do not advance such an interpretation here.

accounting for the observed overrepresentation of the article-less form in ChinE, *viz.* the above noted absence of a definite-article equivalent in the major contact language.

While the above patterns are certainly interesting and worthy of further investigation, there remain a number of methodological caveats. For one, a comparison of absolute web-based figures shows very clearly that the material has increased dramatically over the past 14 years. The relative similarity and closing gaps between 2003 and 2017 also suggest that, in relative terms, more material from international sources (including advertising, PR agencies, companies, news agencies) can be found on country-specific domains today than in the past. Finally, a considerable (but potentially varying) proportion of the zero-article figures occur in headlines, where null articles are favored in all varieties of English.

What has apparently been overlooked in the previous pertinent literature, however, is that the variation in standard Englishes between *a majority* and *the majority* is semantically irrelevant. We are dealing with an interesting case, therefore, in which definite and indefinite article mean the same: 'greater than 50%', such that either one, from this point of view, is redundant. It is perhaps at least partly due to this semantic fact that a third variant, the zero article, becomes a further option.

4 Common nouns as names in British English

Section 3 surveyed a number of contexts that allow for variation of different kinds in the appearance of the definite article. These contexts have in common that the noun whose determination is at issue either has clear name-like properties, or is not being used referentially. There are, however, a few common nouns which, at least for some speakers, can occur without the article in standard British English, even when referential and acting as a core argument of the local predicate. The purpose of the present section is to explore in detail the use of the definite article with two such items – *government* and *cabinet* – with a view to demonstrating in the following sections that the rules underlying article omission in British and Maltese English are not just quantitatively, but qualitatively different.

4.1 Overview

In the uses of interest to us here, *government* and *cabinet* are referential count nouns. This means that, when definite, we ordinarily expect that they will take

the definite article, or some other means of determination, such as a demonstrative or possessive element. Generally speaking, this is indeed what we find when examining these items in the BNC, as in the examples in (53)–(54).

- (53) Legally, the Government is not in a position to force generators to take more coal. [K97 1625]
- (54) The Cabinet then dispersed, expecting that next day another government, presumably a Conservative–Liberal coalition, would be formed. [FU3 2797]

What sets these items apart from the vast majority of other referential count nouns, however, is that in a significant minority of occurrences they are article-less, as in the examples in (55)–(56).

- (55) Within government itself the practicalities of administering controls in peacetime became increasingly difficult. [CN9 244]
- (56) The atmosphere in cabinet, they say, is “palpably better”. [CR7 1765]

Examples (55)–(56) are representative of the majority of article-less uses of these items in the BNC, in that they occur within prepositional phrases (PPs), and are not core arguments of the local predicate, but numerous examples of truly bare (henceforth ‘free-standing’) uses of the items, as in (57) and (58), may be found too, showing that these usages cannot be subsumed under the class of article-less locational PPs (*at school* etc.) discussed in Section 3.5.

- (57) It is Government’s aim to create a class of full-time professional farmers [A6M 477]
- (58) Cabinet was asked to look at the probable growth in the economy [G1C 118]

4.2 Quantitative investigation of the BNC

We conducted a quantitative investigation of these different uses of *government* and *cabinet*, using the Simple Query Syntax to search the BNC via the *BNCweb* interface. The goal of this investigation was to obtain figures for the number of article-less uses of *government* and *cabinet* items in the corpus. For *cabinet* this

was relatively straightforward. We used a query designed to return all instances of this item where it was not immediately followed by a noun (i.e. where *cabinet* was the head of its phrase, not a modifier), and not immediately preceded by the definite article or any other element normally in complementary distribution with the definite article.¹⁰ Note that this query does not exclude tokens of *cabinet* immediately preceded by an adjective or noun. This is necessary because, as we will see below, *cabinet* sometimes occurs without an article when preceded by adjectives such as *full*. A disadvantage of not excluding tokens with immediately preceding adjectives or nouns is that it was necessary to hand-sort the 1,448 hits returned by this query and exclude the majority that were not article-less but of the form DETERMINER_MODIFIER_TOKEN. Once these are excluded, along with any tokens of *cabinet* that refer to the item of furniture rather than the instrument of government, we are left with 125 tokens of article-less *cabinet* in the whole corpus. Of these, about three quarters (92/125) occur as part of a PP, and one quarter (33/125) are free-standing. We can get an approximate sense of what proportion of the relevant total these 125 tokens represent as follows. Searching the BNC for tokens of *the* (MODIFIER) *cabinet* not followed by a noun returns 1,870 hits.¹¹ In a random 5% sample of these (93 hits), 12% (11/93) refer to the item of furniture, suggesting that *the* (MODIFIER) *cabinet* not followed by a noun occurs approximately 1,650 times (~88% of 1,870) in the BNC with the intended meaning. This suggests that roughly 7% (125/1,775) of definite uses of *cabinet* in the BNC are article-less. These figures are summarized in Table 4.

Turning to *government*, here we face the difficulty that, alongside the referential use we are interested in here, where reference is to the government of a particular nation or region, this item has a frequent abstract-noun or gerund-type use, as in *the art of government*. Like all such nouns in English, this use of *government* does not take a definite article. To make matters worse, this use of *government* and the referential use we are interested in tend to occur in very similar contexts, meaning careful consideration needs to be given to a particular instance of article-less *government* in order to ascertain which use it instantiates. On the plus side, unlike with *cabinet*, article omission with referential *government* does not seem to be compatible with any kind of premodification. This means that we were able to search for tokens of *government* that are not immediately preceded by a determiner

¹⁰ The query used was “(_AJ0|_AV*|_CJ*|_CRD|_ITJ|_N*|_PN*|_PR*|_PU*|_V*|_X*) cabinet (_A*|_CJS|_CJT|_CRD|_D*|_E*|_ITJ|_O*|_PN*|_PO*|_PRP*|_PU*|_T*|_V*|_X*)”. Details of the CLAWS-5 tagset employed in the XML Edition of the BNC can be found here: bnc2guide.htm#tagset.

¹¹ The query used for this search was “the (_N*|_AJ*)? cabinet (_A*|_CJS|_CJT|_CRD|_D*|_E*|_ITJ|_O*|_PN*|_PO*|_PRP*|_PU*|_T*|_V*|_X*)”.

Table 4: Definite uses of *cabinet* in the BNC.

	Without article			With article (approx.)	Grand total
	Free-standing	In PP	Total		
Tokens	33	92	125	1,650	1,775
Proportion of total	2%	5%	7%	93%	100%

or a noun or adjective.¹² Even so, this query returns more than 4,000 results. We therefore searched the spoken component (10% of the total corpus word count) and the written component of the corpus separately. The query returns 165 results from the spoken component of the corpus, of which 46 are tokens of article-less referential *government*. 17 of these (36%) are free-standing, the remainder are in PPs. The query returns 4,194 results from the written component of the corpus. To make the investigation more manageable, we thinned this to a random 10% sample of the total. Of these 419 results, 75 are tokens of article-less referential *government*. Extrapolation from our random 10% sample suggests that there are approximately 750 tokens in the whole written component of the corpus and approximately 800 across both components. In the 10% sample of the written component, 30 of the 75 article-less tokens were free-standing (40%), the remainder were in PPs. A search of the whole corpus for “the government”, not followed by a noun, returns approximately 21,000 results. This suggests that the definite article is omitted with definite referential uses of *government* in the BNC approximately 4% of the time. These figures are summarized in Table 5.

4.3 Explaining article omission with *government* and *cabinet*

Article omission with *government* and *cabinet* is thus a rare but solidly attested phenomenon in standard British English. In this subsection we deal with the questions of why it occurs at all and what constraints prevent it from being attested more frequently.

The first point to make here is that article omission with these items does not seem to be possible in the native-speaker idiolect of one of the present co-authors (CL). In his idiolect – and presumably the idiolects of many other native speakers of British English – these items behave just like any other referential

¹² The query used for *government* was “(_AV*_CJ*_CRD|ITJ|_PN*_PR*_PU*_V*_X*) government (_A*_CJS|CJT|CRD|_D*_E*_ITJ|_O*_PN*_PO*_PRP*_PU*_T*_V*_X*)”.

Table 5: Definite uses of *government* in the BNC (partly extrapolated).

	Without article				Total (approx.)	With article (approx.)	Grand total
	Free-standing		In PP				
	Spoken	Written	Spoken	Written			
Tokens	17	300	29	450	800	21,000	21,800
Proportion of total	< 1%	1.5%	< 1%	2%	4%	96%	100%

common noun in requiring some explicit morphological indication of their definiteness status. In order to understand which sub-group of speakers do tend to omit the article with these items, it is helpful to consider two other items (perhaps the only two) that are known to behave similarly for some speakers of British English. These are *conference*, when this refers to the annual conference of large bodies such as political parties or trade unions, and *college*, when this refers to one of the constituent colleges of traditional UK universities such as Oxford and Cambridge. Examples are given in (59) and (60).

(59) On Saturday afternoon the RAF Presentation Team [...] gave their presentation to Conference for which they received the grateful thanks and applause of those present. Conference provided a focus for Telethon 1990, when delegates and visitors donned blue ‘ears’ to raise money for this charity. [A67 327–8]

(60) Accommodation in College for Old Somervillians: We also warmly welcome graduates visiting Oxford, out of term, who would like to stay in College [J2B 203]

These are self-consciously in-group usages. The few speakers who use *conference* without an article do not do so routinely when referring to any large conference: members of the Labour Party, for example, will refer (among themselves, and perhaps also when speaking with certain non-members) to their party’s own annual conference as bare *Conference*; but they would always use an article when referring to the annual conference of the Conservative Party. Similarly with *college*. The example in (60) appears, from its wider context, to come from a letter or pamphlet addressed specifically to graduates of Somerville College, Oxford. It is only the actual members (or former members) of such a

college who will refer to it without an article: most English speakers, who have no association with such an institution, will not even be aware that this usage exists; and, as with *conference*, members of a particular college will only ever omit the article when referring to that college, not to any others.

Our claim is that article-less *government* and *cabinet* are in-group usages in the same way – they only occur, in the UK context, in the speech and writing of individuals who feel a close association with national politics: politicians, political journalists, and perhaps also a few members of the general public who take an especially keen interest in British politics. The relatively low occurrence of article-less *government* and *cabinet* in the BNC (approximately 4% and 7% of relevant attestations, respectively) is thus to be understood as the consequence of this usage being absent from the idiolects of the majority of speakers of British English.

There is of course a great deal more that could be said about the social factors that constrain this usage, but in the remainder of this section we focus on the question of the linguistic factors that make it possible at all. One potentially fruitful line of enquiry (which we will also pursue in Section 5 on article omission in Maltese English) is to ask whether article omission in these cases is licensed by the referents being very readily identifiable in context. Clearly there are many entities in the world that may accurately be described as *governments*, *cabinets*, *colleges* and *conferences*. But in the contexts in which article-less uses such as (55)–(60) occur, the referent of the noun in question is always very readily identifiable. It cannot be, however, that the ready identifiability of the referent is what licenses article omission. To see this, consider first of all what Löbner (2011) calls ‘individual’ nouns. These are nouns for which, in ordinary usage, there is only one possible referent – they necessarily refer unambiguously.¹³ Examples include *sun*, and *sky*. The (ordinary) meaning of these words is such that there is only ever one entity in a pair of interlocutors’ mutual cognitive environment, irrespective of context, that *sun* and *sky* can refer to. As such, the referents of individual nouns are maximally identifiable. And yet we never find article-less uses of such items parallel to what we have seen with *government* and the others in (55)–(60).

This point is reinforced if we look in more detail at uses of *cabinet* in corpora, specifically at the minority of tokens in which *cabinet* is premodified. In the BNC, we find that 10 of the 125 tokens of article-less *cabinet* are modified by *full*, as in (61). There are also 22 tokens of *full cabinet* in the BNC that are

¹³ Note that Löbner (2011) also classes proper names and pronouns as types of individual noun. Here we restrict our use of the term to ordinary nouns that refer unambiguously.

preceded by an article, as in (62). *Full cabinet* in these examples should be understood as referring to all ministers in the government, including more junior ones, and contrasts with *inner cabinet* and other terms that refer only to certain more senior ministers.

(61) Foreign policy tended to be the area where Churchill kept full Cabinet somewhat at arm's length. [BOH 292]

(62) It is rare for Mrs Thatcher to have more than one meeting of the full Cabinet per week. [BOH 1349]

Another frequent premodifier of *cabinet* is *shadow*, where *shadow cabinet* refers to the members of parliament who would constitute the actual cabinet if their party were in government rather than the opposition. There are no article-less uses of *shadow cabinet* in the BNC, but they can be found elsewhere, for example in Press_UK, as in (63).

(63) I don't think [Theresa Villiers] getting on to Shadow Cabinet so quickly did her any good. [Pr_UK_26]

The important thing to notice about these modifiers *full* and *shadow* is that they do not, in themselves, make the reference of the noun phrases in which they occur unambiguous: there are multiple entities in the world, in different countries that can accurately be described as a full cabinet or a shadow cabinet. As such, the phrases *full cabinet* and *shadow cabinet* pattern with unmodified *cabinet* and *government* – these are not individual nouns like *sun* and *sky*. Note, however, that there is a class of restrictive premodifiers which do have the effect of rendering the reference of a noun phrase headed by *cabinet* unambiguous. The premodifiers in question are those that restrict the reference of the noun phrase in which they occur to the (unique) cabinet of a particular nation or region, as in (64) and (65).

(64) The Dutch cabinet has agreed a plan to declare the country's North Sea territorial waters a protected environmental zone. [J39 426]

(65) Saddam Hussein appointed Col. Hussain, the PFKG Premier, as a Deputy Prime Minister in the Iraqi Cabinet. [HKX 170]

At any one time, there cannot be more than one Dutch cabinet or more than one Iraqi cabinet. These phrases thus refer unambiguously, irrespective of the

context – their referents are maximally identifiable. And, as with the individual nouns *sun* and *sky* with which they share this property, article-less uses of these phrases never seem to occur. There are 53 phrases in the BNC of the kind illustrated in (64) and (65), and none appear without an article.

In view of these findings, it does not seem helpful to pursue further the idea that article omission of the kind illustrated in (55)–(60) is a function of referents being highly salient (that is, very readily identifiable), since the article never seems to be omitted with other nouns and phrases whose referents are maximally salient. Instead, we would like to propose that, when speakers omit the article with items such as *government*, *cabinet*, *conference* and *college*, they do so because they are treating these items as names.¹⁴

An analysis of these article-less uses as names has the obvious advantage that it enables a straightforward explanation of why article dropping is possible at all here: names are a class of definite noun phrase which do not require an article in standard English. There is also a clear incentive to refer to the government and cabinet of a nation directly via a name as opposed to indirectly via a descriptive noun phrase, as these are institutions, and institutions and related groupings are generally given names in English (consider, e.g., schools, universities, corporations, rock bands, etc.). Finally, the name analysis has the added advantage of not requiring us to posit that the minority of speakers who exhibit these usages do so on the basis of different rules of definiteness marking to those followed by the majority of speakers who lack these usages. Instead, we can simply say that *government* and *cabinet* as names are not widely known – and therefore hardly used – outside of UK political circles.

5 Article omission in ICE-Malta

Turning to Maltese English, here we find what appears to be a similar situation with respect to article omission with *government* and *cabinet*, in that we find both items used both with and without an article in the transitional 2017 ICE-Malta corpus:

- (66) The unions are also urging the government to keep its word regarding the price of oil. [pr_M_056]

¹⁴ Frequent (but not consistent) capitalization of article-less uses of these items suggests that writers and editors view them as sufficiently name-like as to require similar orthographical treatment.

- (67) If these issues, along with a hundred others, are not worrying the cabinet, they are certainly worrying the business community [ed_M_028]
- (68) But government yesterday distanced itself outright from the censorship [pr_M_027]
- (69) These are precisely the sort of proposals that Cabinet should be discussing [ed_M_059]

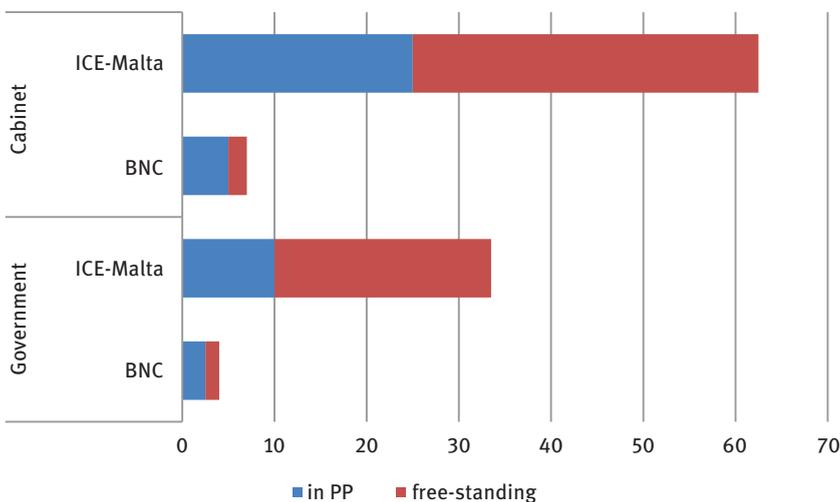
If, however, we look at the proportions of definite uses of these items with and without articles in ICE-Malta_2017, a rather different picture emerges to what we saw for British English in the previous section. There are 334 tokens of *government* as the head of a definite noun phrase in ICE-Malta_2017. Of these, 112 (34%) are without an article, and of these 112 article-less uses, 34 are in PPs and the remaining 78 are free-standing. There are eight tokens of *cabinet* as the head of a definite noun phrase, and five of these are article-less (2 in a PP, 3 free-standing). These figures are summarized in Table 6, and Figure 3 gives a visual representation of the relative proportions of definite uses of *government* and *cabinet* with and without articles in the *BNC* and ICE-Malta_2017, respectively.

We see that article omission with *government* and *cabinet* is much more prevalent in Maltese English than in British English. Moreover, free-standing uses, which are relatively rare in British English, as compared to PP uses, are the more frequently attested type of article-less use in Maltese English. Neither of these observations in themselves excludes the possibility that article-less uses with these two items are best explained the same way in Maltese English as in British English. And there is doubtless considerable influence on journalistic and political Maltese English from the same genres in British English. Nevertheless, the argument we pursue here is that the quantitative difference between the two varieties with respect to *government* and *cabinet* reflects a qualitative difference in the linguistic factors licensing article omission in each variety. Specifically, for Maltese English article omission the evidence points towards the analysis that we rejected for British English, namely that it is licensed by the referents of the relevant nouns being highly salient (very readily identifiable) in context, and not by those nouns being treated as names.

The key observation that points to a different analysis in Maltese English is that article omission occurs with a wider range of nouns than in British English.

Table 6: Definite uses of *government* and *cabinet* in ICE-Malta_2017.

		Without article			With article	Grand total
		Free-standing	In PP	Total		
<i>government</i>	Tokens	78	34	112	222	334
	Proportion of total	23.5%	10%	33.5%	66.5%	100%
<i>cabinet</i>	Tokens	3	2	5	3	8
	Proportion of total	37.5%	25%	62.5%	37.5%	100%

**Figure 3:** Article-less uses of *government* and *cabinet* in BNC and ICE-Malta_2017 as a proportion of all definite tokens (%).

In ICE-Malta_2017, in addition to *government* and *cabinet*, we found article-less uses of *budget*, *beginning*, and *majority*. None of these items occur without an article in standard British English.¹⁵

¹⁵ For example, there are more than 1,000 tokens of *budget* as the head of a definite noun phrase in the BNC. Among these, there is just one token that lacks an article:

(i) [The cost of petrol] will go up again when budget comes in. [KB1 3346]

This comes from the spoken portion of the corpus, and was uttered by a traditional northern English speaker who frequently drops definite articles as a regular feature of his dialect (or is at least transcribed in the BNC as doing so). There is therefore no solid evidence in the BNC that article omission is a possibility with *budget*. The same is true of *beginning* and *majority*.

In ICE-Malta_2017 there are nine tokens of *budget* (in the sense of the annual spending plans of a national government) as the head of a definite noun phrase, and four of these are article-less, as in the following example.

- (70) Extra funding to finance this may need to be incorporated in 2010 budget.
[ed_M_008]

There are ten tokens of *beginning of*, and two of these are article-less, as shown in (71) and (72).

- (71) [...] consumer price inflation fell less than expected in beginning of this year.
[ed_M_011]
- (72) Those recommendations were endorsed by EU finance ministers informally in Prague in beginning of April
[ed_M_011]

Finally, in the spoken portion of ICE-Malta_2017 there are nine tokens of *majority of*, and six of these are article-less, as in (73). It must be noted, however, that all six are uttered by the same speaker, meaning that this could be an idiosyncratic feature of her speech and no one else's (but see Sections 3.6 and 6 for further investigation of this point).

- (73) Water in Malta is very expensive. Majority of water we get in our houses.
[ed_M_011]

The question, then, is what do these items – *government*, *cabinet*, *budget*, *beginning*, and *majority* – have in common such that they can optionally appear without an article in the speech of at least some speakers of Maltese English? Clearly the generalization that held for the article-less nouns in British English discussed in Section 4 – that they all refer to institutions of one kind or another – does not hold here. And it makes little sense to consider whether more functional items such as *beginning* and *majority* are being treated as names. What these items have in common is that the contextual uniqueness of their referents is particularly salient. In discussions of the internal political situation in Malta, among speakers of Maltese English, identifying the referents of *government* and *cabinet* as the current government and cabinet of Malta is a trivial matter. The same is true of *budget*, particularly if the year of the budget in question is specified as it is in (70), and as it is, in fact, in all four of the article-less tokens of *budget* in ICE-Malta_2017. With *beginning* in (71)–(72), the uniqueness of the referent is inherent in the meaning of the noun, similarly to individual nouns

such as *sun* and *sky*,¹⁶ discussed in the previous section – any event necessarily has a single beginning. The same seems to be true of *majority* – the denotation of this quantifier is vague, but it typically has a definite NP as its partitive complement, and arguably divides the denotation of that complement into two parts, referring to the larger of the two, which is necessarily unique.

To summarize the arguments so far, then, we argue that the data on article omission from ICE-Malta indicates that Maltese English has innovated the option of omitting the definite article with at least some nouns, in contexts in which the uniqueness or identifiability of the referents is especially salient. This is true for each of the items with which article omission occurs in ICE-Malta_2017 (*government*, *cabinet*, *budget*, *beginning* and *majority*). This analysis also allows an explanation of the discrepancies in article omission tendencies that we have observed between British and Maltese English. *Budget*, *beginning*, and *majority* do not occur without an article in the *BNC* because the referents of these items are not suitable for being referred to via names. *Government* and *cabinet* are suitable for being treated in this way, but only the minority of speakers who feel particularly closely involved with UK politics omit the article with these items, and then only optionally, explaining why over 90% of tokens of these items as heads of definite noun phrases in the *BNC* retain the article. In ICE-Malta_2017, over 60% of tokens with *cabinet* as head of a definite noun phrase appear *without* the article, and with *government* it is approximately one third. These much higher rates of article omission in Maltese English result, we argue, from a wide range of users of Maltese English treating the article as optional with these items when their referents are very readily identifiable in context. We discuss possible motivations for this innovation in Section 7. Before that, Section 6 offers further support for this analysis of Maltese English article omission, in the form of results from an acceptability judgment questionnaire.

6 Maltese English speakers' judgments of article omission

6.1 Definite-article questionnaire at the University of Malta

To test some of the hypotheses outlined above, we conducted a small-scale questionnaire study in 2015. The participants were 11 educated speakers of

¹⁶ Note, however, that *beginning* would be classified as a 'functional' noun on Löbner's (2011) typology, since its meaning requires it to have a possessor argument, unlike with individual nouns.

Table 7: Average usage ratings of control sentences and examples related to the use of definite articles.

	Average Rating	Sentence
(i)	4.7	Do you have a car?
(ii)	4.4	Prof Muscat works at the University of Malta.
(iii)	4.3	Prof Muscat works at University of Malta.
(iv)	4.0	Second year of university is usually more intense.
(v)	3.8	Answers to all questions can be found on internet.
(vi)	3.8	These days, internet is crucial for business.
(vii)	3.7	The second year of university tends to be more intense.
(viii)	3.6	I think majority of Maltese can swim.
(ix)	3.2	I've seen the Queen on TV, but not in person.
(x)	3.0	I talked to receptionist about the broken TV in my room.
(xi)	2.9	The majority of life on Earth lives in the oceans.
(xii)	2.8	Children usually start walking in second year of life.
(xiii)	2.5	Prince William is Queen's grandson.
(xiv)	1.3	Yesterday, I am to Gozo.

assumption made earlier, that *University of Malta* is treated as a unique (geographical) name. It deserves to be studied in more detail, however, to what extent this result is owed to the fact that all our raters are University of Malta students, i.e. members of an in-group familiar with the relevant institution.

At 4.0, the next-highest rating of example (iv), *Second year of university is usually more intense*, is in fact slightly higher than that of (vii), the same sentence with the definite article. This suggests that in MaltE, given the right (educational) context, article-less usage of ordinal adjectives is at least equally acceptable as (if not preferred to) article-full usage. That educational context is critical with ordinals is shown by the fact that *Children usually start walking in second year of life* receives the second-lowest rating (2.8) of the items eliciting intuitions on article usage. Only (xiii) *Prince William is Queen's grandson* scores lower, which we attribute to the fact that *the Queen* in modern English is conventionalized with the definite article. This can also be gauged from the higher rating of example (ix), and not only from *God save the Queen*. Another likely factor contributing to this relatively low (but in absolute terms, on a scale from 0 to 5, on average exactly intermediate) rating is that, without the definite article, *Queen* shifts its meaning and refers to one of the most popular rock bands of all time, which is obviously incompatible with the intended meaning of (xiii).

It deserves mention that the two article-less examples involving *internet* – *Answers to all questions can be found on internet* and *These days, internet is crucial for business* – receive exactly the same, rather high, rating of 3.8, which

is, according to the scales presented at the beginning of Section 6.1 and in (74) below, substantially closer to “everyone” (= 5) than to “no-one” (= 0).

As was seen for ordinal numbers, the example without the definite article, (viii) *I think majority of Maltese can swim*, is rated to be more common by our informants (at 3.6) than that with the article, (xi) *The majority of life on Earth lives in the oceans* (at 2.9). We therefore find recurrent patterns in which informal conversations among friends in MaltE prefer article-less examples over parallel ones with the definite article.

Finally, rated at 3.0, the article-less example (x) *I talked to receptionist about the broken TV in my room* is judged to be roughly equally common as the two standard sentences *I've seen the Queen on TV, but not in person* and *The majority of life on Earth lives in the oceans*. We would attribute this relatively high rating of example (x) to a high degree of discourse familiarity (or givenness, cf. Sharma 2005) of *receptionist* in a given hotel context, in which case posts become similar to personal names, which in turn are ungrammatical with the definite article in the major norm-providing varieties of English.

6.2 Definite article omission in the Bamberg questionnaire project on Lexical and Morphosyntactic Variation

As part of the Bamberg questionnaire project on Lexical and Morphosyntactic Variation, 74 Maltese university students and 33 British university students rated the usage frequency of, inter alia, the ten examples related to definite article use and omission listed in Table 8. Data collection took place at the University of Cambridge in 2009/10 and at the University of Malta in 2013.

Table 8: Average usage ratings by Maltese and British university students for a semi-formal written and an informal spoken register.

		Maltese English		British English	
		Written 2013	Spoken	Written 2010	Spoken
(i)	Maria has a very good knowledge of \emptyset Italian language.	2.9 (n = 69)	3.0 (n = 74)	2.4 (n = 33)	2.6 (n = 33)
(ii)	Some students spent two weeks in families, enjoying \emptyset American way of life.	2.0 (n = 67)	2.1 (n = 72)	0.9 (n = 33)	1.8 (n = 33)
(iii)	Many people felt the threat of a nuclear war during the first half of \emptyset eighties.	2.8 (n = 68)	–	2.1 (n = 33)	–
(iv)	How fast I talk depends on \emptyset language I use.	2.2 (n = 69)	–	0.6 (n = 32)	–

In the grammar sections of the Bamberg questionnaire, we investigate many colloquial or non-standard features, which are often stigmatized among the well-educated layers of the speech communities under study. In order to avoid social desirability effects (like ‘My family/I myself/My friends would never use such a construction!’), we have questionnaire participants rate sentences not for themselves or people with whom they are closely related. Another reason for not asking raters about their own or their family’s usage is that many of them are L2-speakers of English (as in Malta) who do not regularly use English at home or with their friends. As a result, if asked whether sentences could be said in their family, informants might rate items low in terms of usage not because of their intuitions about the English they hear but for reasons beyond our control and interest. We therefore ask informants to rate whether each sentence could (a) be said in their home country in an informal conversation among friends or (b) be written in their home country in an email to a former teacher by “everyone”, “most”, “many”, “some”, “few” or “no-one”. Despite some caveats (dealt with in Krug and Sell 2013), the data can be converted for statistical processing into a quasi-interval scale as follows:

(74) everyone = 5; most = 4; many = 3; some = 2; few = 1; no-one = 0

Table 8 shows the average usage ratings of 74 Maltese and 33 British university students for a semi-formal written and an informal spoken register.

At 2.9 for written and 3.0 for spoken MaltE, *Maria has a very good knowledge of Italian language* receives the highest usage rating, and shows hardly any stylistic differentiation in MaltE. BrE has lower usage ratings and seems to differentiate stylistically, a tendency that is much more pronounced for the second item involving a language adjective: *Some students spent two weeks in families, enjoying American way of life*. Again, Maltese subjects differ only minimally (2.0 vs. 2.1) in usage ratings for informal spoken and more formal written English; British subjects, by contrast, not only have lower usage ratings for both genres but also differ significantly stylistically in having such low ratings for the written register (0.9, i.e. few) that they must be considered by many raters as ungrammatical or extremely uncommon.¹⁸

As with all items for which we have data from both Malta and Britain, (iii) *Many people felt the threat of a nuclear war during the first half of eighties* is rated higher for MaltE (2.8) than for BritE (2.2). We attribute the relatively high rating in

¹⁸ It will be evident that in order to arrive at an average rating of 0.9, many subjects must choose the option ‘This sentence could be said in my home country by *no one*’.

both varieties on the semantics of *eighties*, which has name-like properties and is thus akin to unique reference, just like the language adjectives in (i) and (ii).

How fast I talk depends on language I use, is rated considerably higher in MaltE (2.2) than for BrE, where (at 0.5) it in fact ranks lowest of all examples presented in Table 8. The only surprise is that it ranks relatively high in MaltE despite the fact that *language I use* is not inherently unique, at least not out of context. We may tentatively suggest two ways of explaining the situation in MaltE: one is that due to the typical situation of bilingualism with Maltese as the dominant, first language and English as the second language, the sentence may be more predictable in a Maltese context than in other countries. The second is that in a language continuum from acrolectal to basilectal MaltE, a judgment ‘could be written by some’ is probably more adequate than ‘could be written by few’, even for structures that are not part of the major norm-providing inner-circle varieties. It is certainly evident that Maltese raters shy away from the extreme pole ‘no-one’, and probably rightly so.

The paired structures with and without definite articles displayed in Table 9 were introduced into the written part of the questionnaire only after 2010, when article omission was identified as a particularly interesting field of investigation. We therefore have no comparative BrE data (yet). The ratings for written MaltE from 2013 alone, however, show interesting and recurrent patterns: Although we obtain relatively high usage ratings for article-less uses of 2.5 to 2.8 (i.e. between ‘Some’ and ‘Many’), the parallel structures including the definite article receive

Table 9: Average usage ratings by Maltese university students for a semi-formal written register.

	Usage rating written MaltE 2013
(v) Yesterday, I went to Ø office very early in the morning.	2.5 (n = 66)
(vi) Yesterday, I went to the office very early in the morning.	3.4 (n = 69)
(vii) I talked to Ø receptionist about the broken TV in my room.	2.5 (n = 66)
(viii) I talked to the receptionist about the broken TV in my room.	3.0 (n = 66)
(ix) Yesterday, Ø government decided on a new healthcare policy.	2.8 (n = 67)
(x) Yesterday, the government decided on a new healthcare policy.	3.4 (n = 63)

higher usage ratings for all three items (of 3.0 to 3.4, i.e. between ‘Many’ and ‘Most’), which points to a preference for article-full usage in a semi-formal written genre. That formality may play a role is also suggested by more fine-grained analyses of *government* in press editorials and press reportage (Krug et al., forthcoming), where higher ratios of zero articles are found in the more involved journalistic text type of editorials. It is also consistent with our earlier analyses and quantitative findings presented above that the noun *government* in *Yesterday, government decided on a new healthcare policy* receives the highest usage rating (of 2.8) among our examples without the definite article.

7 What motivates increased article omission in Maltese English?

As is well known, many of the features that make Maltese English distinctive are, at least in part, the result of transfer from Maltese (see Krug 2015 for some examples). It makes sense, therefore, to ask whether the relatively higher incidence of article omission in Maltese English as compared to standard British English can be explained by transfer from Maltese. This question can be straightforwardly answered in the negative. In Maltese the head of a definite noun phrase is obligatorily marked with the definite article *il-*,¹⁹ except in certain sharply circumscribed syntactic contexts, such as certain possessive constructions. None of the Maltese translation equivalents of the items discussed in Section 5 can appear without the article in ordinary definite contexts. For illustration, in the 250 million word *Korpus Malti v3.0*, a search for *fix-xitwa* ‘in.DEF-winter’ returns 1,876 hits, while *fi xitwa*, without the definite article, returns none; *fil-kabinett* ‘in.DEF cabinet’ returns 1,954 hits, while *fi kabinett* returns none; *fil-baġit* ‘in.DEF budget’ returns 4,825 hits, while *fi baġit* returns none; *fil-bidu* ‘in.DEF beginning’ returns 19,600 hits, while *fi bidu* returns none; *fil-maġġoranza* ‘in.DEF majority’ returns 2,714 hits, while *fi maġġoranza* returns none.²⁰

Transfer from Maltese can thus be ruled out as a cause of increased article omission in Maltese English. Instead, the most likely motivations of this

¹⁹ Note that the /l/ of the Maltese article fully assimilates to an initial coronal consonant in the following noun, and that Maltese orthography treats the article as a suffix on certain monosyllabic prepositions such as *fi* ‘in’.

²⁰ *Fil-gvern* ‘in.DEF government’ returns 27,629 hits. *Fi gvern*, without the article returns 192 hits, but in all of these *gvern* is interpreted as indefinite (indefiniteness is zero-marked in Maltese).

development include cognitive and grammatical forces, as well as the universal pressure towards system economy, which is widely attested in dialect and second language acquisition studies as well as language contact scenarios more generally (cf. Trudgill 1990: 82, 2010; Sharma 2005; Ellis 2008: Ch. 9). As we have seen in Sections 3 and 4, there are numerous contexts in which the article is optionally present in definite contexts in inner-circle varieties like standard British English, but the constraints on this optionality are complex and heterogeneous. Maltese English preserves the existing optionality, but extends it as a result, we argue, of simplifying this optionality to a single rule (with certain lexical exceptions such as *Queen*): if the uniqueness or identifiability of a referent is particularly salient in context, the article may be optionally omitted.

8 Conclusion

For the vast majority of our data investigated in this paper it is not contact with Maltese or other contact dialects that motivates definite article usage and omission in MaltE, but general cognitive and grammatical factors. These are probably universal and thus help to explain why some of the phenomena described here enjoy a more global spread. Maltese English (like other varieties too, we suspect) seems to use the definite article most consistently only when it is semantically or grammatically required: it frequently omits the definite article when its usage would be semantically redundant. This is not only in evidence with unambiguous posts, institutions, language adjectives and ordinal numbers, but also when the referent has been established in the discourse or has acquired name-like characteristics (as with *Government*, *Cabinet*, *University of Malta*). On the other hand, MaltE avoids the definite article commonly where the noun is not definite but generic (as with *summer*, *winter*, *autumn* and *spring*). What we seem to witness – apart from a great deal of overlap between MaltE on the one hand and BrE and AmE standard varieties on the other – is a reorganization of those aspects that are partly unregulated in norm-providing inner-circle varieties, resulting in a MaltE system that is regulated according to fewer parameters but more consistently.

References

- Aston, Guy & Lou Burnard. 1998. *The BNC handbook: Exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.
- Biber, Douglas & Edward Finegan. 1992. The linguistic evolution of five written and speech-based English genres from the 17th to the 12th centuries. In Matti Rissanen, Ossi Ihalainen & Terttu

- Nevalainen (eds.), *History of Englishes: New methods and interpretations in historical linguistics*, 688–704. Berlin: Mouton de Gruyter.
- Bonnici, Lisa. 2010. Variation in Maltese English: The interplay of the local and the global in an emerging postcolonial variety. Davis: University of California Unpublished Ph.D. Dissertation.
- Crisma, Paola. 2011. The emergence of the definite article in English: A contact-induced change? In Petra Sleeman & Harry Peridon (eds.), *The noun phrase in Romance and Germanic: Structure, variation and change*, 175–192. Amsterdam: John Benjamins.
- Ellis, Rod. 2008. *The study of second language acquisition*. Oxford: Oxford University Press.
- Grech, Sarah. 2015. Variation in English: Perception and patterns in the identification of Maltese English. University of Malta PhD thesis. University of Malta Library: Electronic Thesis & Dissertations Repository. <https://www.um.edu.mt/library/oar/bitstream/handle/123456789/5439/15PHDLIN001.pdf?sequence=1&isAllowed=y>.
- Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals of human language*, vol. iii, 47–82. Stanford: Stanford University Press.
- Himmelmann, Nikolaus. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer.
- Hoffmann, Sebastian, Stefan Evert, Nicholas Smith, David Lee & Ylva Berglund Prytz. 2008. *Corpus linguistics with BNCweb: A practical guide*. Frankfurt am Main: Peter Lang.
- Kachru, Braj. 1988. The sacred cows of English. *English Today* 16. 3–8.
- Kortmann, Bernd, Edgar W. Schneider, Kate Burridge, Rajend Mesthrie & Clive Upton (eds.). 2004. *A handbook of varieties of English. Vol. 2: Morphology and syntax*. Berlin & New York: Mouton de Gruyter.
- Krug, Manfred. 2015. Maltese English. In Jeffrey P. Williams, Edgar Schneider, Daniel Schreier & Peter Trudgill (eds.), *Further studies in the lesser-known varieties of English*, 8–50. Cambridge: Cambridge University Press.
- Krug, Manfred, Michaela Hilbert & Ray Fabri. forthcoming. Maltese English morphosyntax: Corpus-based and questionnaire-based studies. *STUF/Language Typology and Universals* (Special issue; Alexandra Vella & Ray Fabri, eds. *Towards a description of Maltese English*).
- Krug, Manfred & Anna Rosen. 2012. Standards of English in Malta and the Channel Islands. In Raymond Hickey (ed.), *Standards of English: Codified varieties around the world*, 117–138. Cambridge: Cambridge University Press.
- Krug, Manfred, Ole Schützler & Valentin Werner. 2016. Patterns of linguistic globalization: Integrating typological profiles and questionnaire data. In Olga Timofeeva, Sarah Chevalier, Anne Gardner & Alpo Honkapohja (eds.), *New approaches to English linguistics: Building bridges*, 35–66. Amsterdam: John Benjamins.
- Krug, Manfred & Katrin Sell. 2013. Designing and conducting interviews and questionnaires. In Manfred Krug & Julia Schlüter (eds.), *Research methods in language variation and change*, 69–98. Cambridge: Cambridge University Press.
- Krug, Manfred & Lukas Sönning. 2018. Language change in Maltese English: The influence of age and parental languages. In Albert Gatt & Patrizia Paggio (eds.), *The languages used on the Island of Malta*, 247–270. (Studies in diversity linguistics series). Berlin: Language Science Press.
- Löbner, Sebastian. 2011. Concept types and determination. *Journal of Semantics* 28. 279–333.

- Mair, Christian. 1995. Changing patterns of complementation, and concomitant grammaticalisation, of the verb *help* in present-day British English. In Bas Aarts & Charles F. Meyer (eds.), *The verb in contemporary English: Theory and description*, 27–35. Cambridge: Cambridge University Press.
- Nelson, Gerald. 1996. The design of the corpus. In Sidney Greenbaum (ed.), *Comparing English worldwide. The International Corpus of English*, 27–35. Oxford: Clarendon Press.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Robertson, Daniel. 2000. Variability in the use of the English article system by Chinese learners of English. *Second Language Research* 16(2). 135–172.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Sedlatschek, Andreas. 2009. *Contemporary Indian English: Variation and change*. Amsterdam: Benjamins.
- Sharma, Devyani. 2005. Language transfer and discourse universals in Indian English article use. *Studies in Second Language Acquisition* 27(4). 535–566.
- Sun, Ganzhao. 2016. The acquisition of English articles by second language learners: The sequence, differences, and difficulties. *SAGE open* (January-March). 1–8.
- Trudgill, Peter. 1990. *The dialects of England*. Oxford: Blackwell.
- Trudgill, Peter. 2010. Contact and sociolinguistic typology. In Raymond Hickey (ed.), *The handbook of language contact*, 299–319. Chichester: Wiley-Blackwell.
- Zahedi, Keivan & Roghayeh Mehrzmay. 2011. Definiteness in Sorani Kurdish and English. *Dialectologia* 7. 129–157. <http://www.raco.cat/index.php/Dialectologia/article/download/247912/332000> (accessed 22 January 2017).
- Zdorenko, Tatiana & Johanne Paradis. 2008. The acquisition of articles in child second language English: Fluctuation, transfer or both? *Second Language Research* 24(2). 227–250.

Appendix: Internet searches for *in a/the/∅ majority of* in ENL, ESL and EFL varieties in 2003 and 2017

Google searches January 2017												
Region:	United Kingdom		Australia		USA		India		Singapore		South Africa	
	.uk	%	.au	%	.us	%	.in	%	.sg	%	.za	%
<i>in the majority of</i>	17,000	0.28%	9,210	0.22%	17,300	0.93%	38,500	0.75%	2,840	0.27%	5,410	0.19%
<i>in a majority of</i>	5,380,000	87.8%	3,140,000	74.2%	1,030,000	55.3%	4,630,000	90.4%	935,000	90.0%	2,770,000	95.8%
Sum	6,131,000	100.0%	4,229,210	100.0%	1,863,300	100.0%	5,122,500	100.0%	1,038,840	100.0%	2,890,410	100.0%
Google searches February 2003 (from Sedlatschek 2009: 210)												
Region:	United Kingdom		Australia		USA		India		Singapore		South Africa	
	.uk	%	.au	%	.us	%	.in	%	.sg	%	.za	%
<i>in the majority of</i>	560	0.09%	832	0.46%	573	0.88%	779	37.82%	82	9.57%	95	0.37%
<i>in a majority of</i>	574,000	95.8%	161,000	88.1%	49,200	75.6%	653	31.7%	670	78.2%	24,200	95.0%
Sum	599,460	100.0%	182,732	100.0%	65,073	100.0%	2,060	100.0%	857	100.0%	25,465	100.0%

(continued)

(continued)

Google searches January 2017

Region:	Malta		New Zealand		Republic of Ireland		Germany		Poland		China	
	.mt	%	.nz	%	.ie	%	.de	%	.pl	%	.cn	%
<i>in majority of</i>	144	0.27%	2,510	0.19%	1,800	0.17%	20,800	0.75%	17,500	1.31%	4,980	1.40%
<i>in the majority of</i>	46,700	87.8%	999,000	74.0%	886,000	83.9%	2,270,000	81.7%	1,180,000	88.2%	141,000	39.5%
<i>in a majority of</i>	6,340	11.9%	348,000	25.8%	168,000	15.9%	488,000	17.6%	141,000	10.5%	211,000	59.1%
Sum	53,184	100.0%	1,349,510	100.0%	1,055,800	100.0%	2,778,800	100.0%	1,338,500	100.0%	356,980	100.0%