

**SYNCHRONIC VARIABILITY IN THE  
COMPLEMENTATION PROFILE OF *REMEMBER*:  
FINITE VS NON-FINITE CLAUSES  
IN INDIAN AND BRITISH ENGLISH**

**VARIACIÓN SINCRÓNICA EN EL PERFIL DE  
COMPLEMENTACIÓN DE *REMEMBER*:  
CLÁUSULAS FINITAS VERSUS CLÁUSULAS NO  
FINITAS EN LOS INGLESES INDIO Y BRITÁNICO**

**LAURA GARCÍA-CASTRO**

Universidad de Vigo  
l.garcia@uvigo.es

---

137

**Abstract**

This paper explores the prevalent simplification of morphosyntactic features occurring in Postcolonial Englishes (cf. e.g. Williams 1987) by addressing the hypothesis that finite complement clauses (CCs) should be more common in these varieties than a non-finite counterpart due to their higher explicitness. The hypothesis is tested in two varieties of English, British English as a reference variety and Indian English, exploring the complementation profile of *REMEMBER* by means of a corpus-based approach. In addition, a variable analysis aims at shedding light on the language-internal features potentially conditioning the choice between CCs in competition. The results partially confirm the hypothesis; there is a larger proportion of finite CCs in Indian English than in British English but non-finite CCs are the most common option in the two varieties. Furthermore, not only simplification, but also other factors such as substrate influence or second-language acquisition processes may be responsible for the distribution found in Indian English. As for language-internal features, while some factors condition the choice across varieties, others are variety-specific.

**Keywords:** corpus linguistics, complementation, simplification, Postcolonial Englishes.

## Resumen

Este artículo versa sobre la frecuente simplificación de rasgos morfosintácticos que se da en las lenguas inglesas postcoloniales (cf. ej. Williams 1987) y explora la hipótesis de que las cláusulas de complementación finitas fueran más comunes en estas variedades del inglés que su equivalente no finita debido a que son más explícitas. Este trabajo pone a prueba esta hipótesis en dos variedades del inglés, el inglés británico como variedad de referencia y el inglés indio, y explora el perfil de complementación del verbo *remember* mediante una metodología de corpus. Además, un análisis cualitativo de variables intenta arrojar luz sobre los factores internos del lenguaje que condicionan la elección de cláusulas de complementación que compiten entre sí. Los resultados confirman la hipótesis parcialmente; hay una mayor proporción de cláusulas finitas en inglés indio, aunque las cláusulas no finitas son la opción más común en las dos variedades. Además, se observa que no solo la simplificación, sino otros factores como la influencia del substrato o los procesos típicos de la adquisición de una segunda lengua pueden ser los responsables de la distribución hallada en el inglés indio. En cuanto a los factores internos del lenguaje, mientras algunos factores condicionan la elección en las dos variedades, otros son específicos de cada una.

**Palabras clave:** lingüística de corpus, complementación, simplificación, ingleses postcoloniales.

## 1. Introduction

In Present-day English, *to*-infinitival and gerund-participial complement clauses (CCs) are frequently attested in complementary distribution after retrospective verbs such as REMEMBER, REGRET and DENY (Quirk et al. 1985: 1193; Fanego 1996: 71), as shown in (1) and (2). These types of clauses are not freely interchangeable, due to their specific functional differentiation: while *to*-infinitival CCs encode a situation which is projected into the future and has not yet taken place in relation to the time of remembering, as in example (1), gerund-participial CCs encode a situation which has already happened and is prior to the time of remembering, as in (2) (Mair 2006: 215; Cuyckens et al. 2014: 182).

(1) *Remember to call when you get there!*

(2) *I remember her calling when she got there.*

In addition to *to*-infinitival and gerund-participial CCs, the retrospective verb REMEMBER, when it conveys the meaning ‘recall’,<sup>1</sup> can also take finite declarative clauses (Mair 2006: 216; Cuyckens et al. 2014: 182-183), as in (3). Unlike the

functional differentiation between *to*-infinitival and gerund-participial CCs (cf. (1) and (2)), declarative clauses and gerund-participial clauses seem to be freely interchangeable on the basis of their identical propositional contents, as illustrated in (3) and (4). Therefore, as has been described in the literature (Fanego 1996; Cuyckens et al. 2014), non-categorical variation exists between declarative and gerund-participial CCs after *REMEMBER*, which is part of what this paper sets out to explore.

(3) *I remember that I called her.*

(4) *I remember calling her.*

Since the complementation profile of verbs (and adjectives) is said to be one of the classic examples where Postcolonial Englishes (PCEs) exhibit differences with respect to other varieties (Schneider 2007: 86), the current study examines CC variation between finite declarative clauses and non-finite gerund-participial clauses (and some types of *to*-infinitival clauses) after the verb *REMEMBER* in two relevant World Englishes by performing a corpus-based analysis with data from British English, an L1 and Inner Circle variety of English, and Indian English, a PCE and Outer Circle variety of English (cf. B. Kachru 1985). Furthermore, the study investigates the factors that condition the choice of CC, firstly by coding the data on the basis of a number of relevant intralinguistic variables, and secondly by performing a regression analysis of such factors. The main purpose of this article then is to establish whether there are differences between British and Indian English, that is, between an L1 and an L2 variety of English, and explore the potential language-external and language-internal factors conditioning the variation between finite and non-finite CCs.

The article is organised as follows: Section 2 presents a theoretical survey of sentential verb complementation and of the complementation profile of *REMEMBER* in particular, which serves as the main focus of the study. Section 3 describes the data and the methodology followed. Section 4 shows the results of the variable analysis and the binary logistic regression analysis, which are discussed in relation to claims made in the literature regarding CC variation in general, and more specifically the variation that pertains to the envelope of variation under discussion here. Finally, Section 5 offers a brief summary of the main conclusions and their theoretical and methodological repercussions on research into complementation, particularly in Postcolonial Englishes.

## **2. Sentential Verb Complementation**

Verb complementation, and more specifically clausal verb complementation, has been a relevant area of interest in linguistics, not least within generative and

cognitive-functional frameworks. The initial studies are synchronic in nature and focus mainly on syntactic issues (e.g. the constituent structure of different complementation patterns; cf. Bresnan 1970, 1979) and semantic issues (e.g. the match between a particular verb and certain complementation options; cf. Noonan 2007). Over the last three decades, a large body of research has focused on diachronic studies on the matter, providing an account of variation and change in the complementation system of English throughout history (Cuyckens et al. 2014: 183-184; cf. Fanego 1996, 2004, 2007, 2016; Rohdenburg 2006, 2007, 2014, among others) and also in recent times (cf. e.g. Kaunisto et al. 2018; Rickman and Rudanko 2018).

Cuyckens et al. (2014), following the diachronic research tendency of recent decades, look at the envelope of variation between finite and non-finite CCs after the verbs REMEMBER, REGRET and DENY. The justification for their study is that this type of variation is non-categorical: these clauses co-exist and seem to be used depending on the speaker's preferences and the time period involved, which they explore in terms of the frequency, distribution and a number of variables that influence the choice. Among these variables, they explore a set of structural factors (e.g. TYPE SUBJECT CC and COMPLEXITY CC) based on what Rohdenburg (1996, 2006) proposes in the Complexity Principle. According to this principle, in more cognitively complex environments, the more explicit option should be preferred, which in this case are finite CCs (they have an expressed subject, code for tense and mode, among other features). Additionally, Mair (2006) points out that, in relation to this same envelope of variation, a high degree of elaboration makes the use of non-finite CCs less likely than finite CCs after REMEMBER, although the choice remains free.

The results of their binary logistic regression analysis are summarised in Table 1, in which the variables analysed are divided into those which favour non-finite complementation and those which disfavour it (Cuyckens et al. 2014: 196-197).

Favouring factors	Disfavouring factors
<ul style="list-style-type: none"> <li>- TYPE SUBJECT MAIN CLAUSE (MC)</li> <li>- COMPLEXITY CC</li> <li>- VOICE CC VERB</li> </ul>	<ul style="list-style-type: none"> <li>- MEANING CC</li> <li>- INTERVENING MATERIAL IN WORDS</li> <li>- TYPE SUBJECT CC</li> <li>- DENOTATION</li> <li>- ANIMACY CC SUBJECT</li> <li>- TEMPORAL RELATION</li> <li>- PERIOD</li> </ul>

Table 1. Factors favouring and disfavouring non-finite complementation in Cuyckens et al. (2014)

As can be seen in Table 1, there are non-structural factors that influence the choice, such as MEANING CC, DENOTATION, ANIMACY CC SUBJECT, TEMPORAL RELATION (between the clauses) and PERIOD. Furthermore, there are interesting results regarding those variables which can be used to measure the cognitive complexity of the environment. In line with Rohdenburg's (1996, 2006) Complexity Principle, grammatical manifestations of cognitive complexity such as INTERVENING MATERIAL (between the MC and the CC) and complex subjects of the CC disfavour non-finite CCs (Cuyckens et al. 2014: 199). However, complex CC predicates and passive structures, which in theory would increase cognitive complexity, favour non-finite CCs. From this, Cuyckens et al. (2014: 199) conclude that "Rohdenburg's proposed disavouring effect cannot be generalized to all structural complexity factors". This paper will examine these and other intralinguistic factors in the L1 and L2 data selected for the study (see Section 4.2).

As already noted, sentential verb complementation in English has been studied both from a synchronic and a diachronic perspective. However, the topic remains under-researched in the blossoming field of Postcolonial Englishes, that is, varieties of English that emerged in former British and American colonial contexts. This is so even though the complementation profile of nouns and verbs is considered a feature likely to be subject to innovation and change in such varieties (Schneider 2007: 86). Indian English, the PCE chosen for the present analysis, is among the most researched varieties within the field of Postcolonial Englishes. A number of studies have dealt with nominal and prepositional complementation in Indian English (cf. Nihalani et al. 1979; Leitner 1994; Shastri 1996; Olavarría de Ersson and Shaw 2003; Mukherjee and Hoffmann 2006; Schilk 2011; Schilk et al. 2013; Röthlisberger et al. 2017, among others). However, research on sentential verb complementation is scarce (cf. Steger and Schneider 2012; Bernaisch 2015; Deshors 2015; Deshors and Gries 2016; Romasanta 2017)<sup>2</sup>. The present study intends to contribute to the lexicogrammatical documentation of Indian English by describing the sentential complementation profile of the verb REMEMBER, and also to contrast Indian and British English here in order to explore whether, and if so to what extent, the factors said to influence the development of PCEs can explain a potential divergence between the two varieties. Among these factors (cf. Williams 1987; Schneider 2007: 88-90, 99-107; Brunner 2014, 2017; Suárez-Gómez 2017, among others), the most pertinent for this study are:

- (i) Innovation: It refers to "the results of internal change and linguistic creativity" (Schneider 2007: 102), which may be caused by three main types of processes: restructuring, exaptation and simplification. The most relevant for the present paper is the tendency towards simplification, which subsumes

various mechanisms and processes such as the tendency towards transparency, and which mainly affects morphosyntactic constructions (Schneider 2007: 102; cf. Wong 1983; Williams 1987; de Klerk 2003). Transparency (also known as the one-to-one principle in second-language acquisition (SLA) research; cf. Andersen 1984) is understood as “a transparent one-to-one mapping of conceptual structure and surface form” (Steger and Schneider 2012: 156; cf. Williams 1987: 179; Schneider 2013: 145; Suárez-Gómez 2017: 215, among others). This configuration minimises the acquisition effort for L2 speakers (Steger and Schneider 2012: 157), and should lead to a preference for more transparent constructions in PCEs, in that these are L2 varieties arising from individual and community SLA (Steger and Schneider 2012: 157). In this case, finite CCs are more transparent than non-finite CCs.

- (ii) Language contact: It may lead to the direct transfer of linguistic material from the substrate languages, can have an impact on the frequency of use of a certain structure, and it can also trigger a preference for particular patterns (Brunner 2014: 23; cf. Brunner 2017; Suárez-Gómez 2017). Sometimes, language contact can also trigger or accelerate other processes, such as those within innovation.
- (iii) SLA processes: They also condition the development of PCEs, since these varieties are the result of single and group SLA, and hence such processes usually intersect with innovation and language contact (Schneider 2013: 143, 148; cf. Thomason 2001). For instance, processes associated with this kind of acquisition may lead to speakers selecting options more similar to their L1, that is, a substrate language (the Shortest Path Principle; cf. Wald 1996), or sticking to structures that they know well and can use with confidence (the Teddy Bear Principle; cf. Hasselgren 1994).

I have surveyed the complementation profile of the four most spoken languages in India, that is Hindi, Telugu, Bengali and Tamil (in that order; cf. Eberhard et al. 2019). All these languages seem to have structures equivalent to declarative CCs but only Telugu has a parallel structure to English gerund-participial CCs (see Appendix 1; Annamalai and Steever 1998: 122; Krishnamurti 1998: 234-235; Y. Kachru 2006: 217; Thompson 2012: 195).

Hence, this study examines the non-categorical variation between finite and non-finite CCs following the verb REMEMBER, as shown in (5) and (6), in British and Indian English. Such variation is characterised by indeterminacy: speakers can choose freely between the two clauses at the level of usage (De Smet 2013: 27-29).

- (5) *I remember on the 1996 tour of England, I went to the gym only once (...)* (GloWbE IN)
- (6) *(...) he remembers going to Havelian and then to Golra junction near Texla to catch the Frontier Mail to Lahore.* (GloWbE IN)

With this purpose in mind my first methodological step is to explore the frequency distribution of the two CC options in British and Indian English, and then, by means of a binary logistic regression analysis, I proceed to seek which factors significantly predict (in terms of odd ratios) the choice of CC and whether these factors are the same in the two varieties. In doing so I will be testing the following hypotheses, as informed by earlier claims made in the literature:

- (a) Discrepancies between British and Indian English (if they exist) are the result of the factors that drive the linguistic evolution of PCEs, described above as (i), (ii) and (iii), and
- (b) Related to (a), in accordance with Rohdenburg's (1996, 2006) Complexity Principle (cf. Section 3.2), complex environments will favour the use of finite CCs in both British and Indian English, but the Complexity Principle will have a stronger impact on Indian English, due to factors (i) to (iii) above.

### 3. Data and Methodology

#### 3.1. Data Selection

As discussed in Sections 1 and 2, REMEMBER with the meaning 'recall' allows free variation between finite and non-finite CCs, something few verbs do. These CCs always have a retrospective temporal relationship with respect to REMEMBER (cf. Cuyckens et al. 2014). Starting with finite CCs after REMEMBER, these are declarative CCs, following the taxonomy of CCs in Huddleston and Pullum (2002: 951-953), which can be either bare declarative clauses, as in (7), or expanded declarative ones, as in (8).

- (7) *I remember Ø I asked her to buy them.* (GloWbE GB)
- (8) *I remember that I asked her to buy them.*

As for non-finite CCs after REMEMBER with the meaning 'recall', these are predominantly gerund-participial CCs with or without an explicit subject, as in (9) and (10) respectively. In addition, perfect *to*-infinitival CCs, as in (11), can also be found, although very rarely.

- (9) (...) *I remember HT doing a story on this very topic!* (GloWbE IN)  
(10) (...) *I tried to remember ever loving him.* (GloWbE GB)  
(11) *A more natural story we do not remember to have read.* (GloWbE GB)

In order to study this variation, data were extracted from the Corpus of Global Web-based English (GloWbE; cf. Davies 2013), by searching for the forms <remember, remembers, remembered, remembering> and retrieving all examples, which were then randomised. From the randomised data sets I selected a 3,000-hit random sample from each variety, that is, British and Indian English. GloWbE contains material retrieved from the Internet in 2012 and can be considered to represent English as used on the web (Loureiro-Porto 2017: 455). Additionally, it is an invaluable tool in the research of World Englishes since it contains abundant data from 20 countries (L1 varieties like British English, PCE varieties like Indian English, ESD —English as a Second Dialect— varieties such as Jamaican English, etc.) and allows for research into low-frequency phenomena such as CC variation. Other corpora containing PCEs, such as ICE (The International Corpus of English) prove too small for such an endeavour (cf. García-Castro 2018).

144

The two random samples were pruned of all spurious hits manually, which included non-valid hits (e.g. incomplete, repeated and ambiguous examples, among others) and instances in which REMEMBER takes nominal complements or clausal complements not in competition. This resulted in the following numbers. In GloWbE GB there are 138 instances (i.e. 25.5%) of <remember + finite CCs> and 404 instances (74.5%) of <remember + non-finite CCs> after REMEMBER meaning ‘recall’ while in GloWbE IN the numbers are 103 (32.5%) instances of <remember + finite CCs> and 216 instances (67.5%) of <remember + non-finite CCs>. These examples were then coded as described in Section 3.2.

### 3.2. Coding of the Data

Each relevant corpus attestation consisting of <remember + CC> was entered into an IBM SPSS (Statistical Package for the Social Sciences) Statistics 25.0 (IBM Corp. 2017) software package database. Hits were coded for a number of factors which might determine the CC choice: characteristics of the main clause (MC), characteristic of the CC, and characteristics of the combined structure of the MC and the CC, which are listed in Table 2 (See Appendix 2, where examples of each variable and value are provided). The selection of potentially significant factors was drawn from the literature, and comprises semantic, structural and other additional factors (cf. Bresnan and Hay 2008; Nam et al. 2013; Cuyckens et al. 2014; Deshors and Gries 2016; Shank et al. 2016).

## Synchronic Variability in the Complementation Profile of *Remember*

Name of variable		Values
Semantic factors	MEANING CC	<i>Event/Action State</i>
	TIME REFERENCE	<i>Independent time reference Dependent time reference</i>
Structural factors	TYPE SUBJECT MC	<i>Pronominal Non-pronominal Non-expressed</i>
	TYPE SUBJECT CC	<i>Pronominal Non-pronominal Non-expressed</i>
	COMPLEXITY CC IN NO. OF CONSTITUENTS	<i>Short (0-1) Medium (2-3) Long (4-5)</i>
	COMPLEXITY CC IN NO. OF WORDS	<i>1-7 8-14 +15</i>
	INTERVENING MATERIAL	<i>Presence Absence</i>
	VOICE CC VERB	<i>Active Passive</i>
	SUPPLEMENTATION	<i>Presence Absence</i>
	COORDINATION	<i>Presence Absence</i>
	POLARITY MC	<i>Positive Negative</i>
	POLARITY CC	<i>Positive Negative</i>
Additional factors	ANIMACY CC SUBJECT	<i>Animate Inanimate</i>
	CO-REFERENTIALITY BETWEEN THE MC AND THE CC SUBJECTS	<i>Co-referential Non-co-referential</i>
	COMPLEMENTATION TYPE	<i>Finite Non-finite</i>

145

Table 2. List of language-internal variables coded

As for semantic factors, Cuyckens et al. (2014: 196) find that MEANING CC is a determinant factor since MEANING CC = *state* disfavours non-finite complementation. They also consider TIME-REFERENCE, a factor that “designates the relation between the time referent of the complement and the meaning of the CTP [complement-taking predicate]” (Cuyckens et al. 2014: 189; cf. Noonan 2007). Structural factors are related to the Complexity Principle (Rohdenburg 1996, 2006) and, in addition to the variables TYPE SUBJECT MC, TYPE SUBJECT CC, COMPLEXITY OF CC IN NO. OF CONSTITUENTS, INTERVENING MATERIAL and VOICE CC VERB, I also included the following variables: COMPLEXITY OF CC IN NO. OF WORDS as another means to measure the overall structural complexity of the CCs, SUPPLEMENTATION<sup>3</sup>, COORDINATION and POLARITY MC and POLARITY CC to account for these features, whose presence increases the complexity of the environment (cf. Rohdenburg 1996, 2006). Finally, as for additional factors, ANIMACY CC SUBJECT and CO-REFERENTIALITY BETWEEN THE MC AND THE CC are considered since they are found to condition CC-variation in Cuyckens et al. (2014: 196), with *inanimate subjects* and *non-co-referential subjects* disavouring non-finite complementation. The variable COMPLEMENTATION TYPE was included as it classifies CCs in the envelope of variation after REMEMBER meaning ‘recall’.

146

## 4. Results

### 4.1. Distribution of Complement Clauses in British and Indian English

Table 3 presents the distribution of finite and non-finite CCs after REMEMBER across the two samples, that is, GloWbE GB (Great Britain) and GloWbE IN (India). As can be observed, the proportion of non-finite CCs in GloWbE GB (74.5%) is higher than in GloWbE IN (67.5%) and the difference is significant at  $p < 0.05$ .

	GloWbE GB		GloWbE IN	
	No.	%	No.	%
Finite CCs	138	25.5	103	32.5
Non-finite CCs	404	74.5	216	67.5
Total	542	100	319	100

Table 3. Distribution of finite and non-finite CCs after REMEMBER (‘recall’) in GloWbE ( $\chi^2 = 4.643$ ,  $df = 1$ ,  $p$ -value = 0.031)

These results are in line with the hypothesis that the proportion of finite CCs should be larger in L2 varieties of English than in L1 varieties. This hypothesis is based on claims that processes of simplification (cf. (i) in Section 2) will lead L2 speakers (here speakers of Indian English) to select finite structures proportionally more often than L1 speakers (British English speakers), due to the closer correspondence between form and meaning characteristic of finite clauses (e.g. expressed subject and tense marking; cf. Steger and Schneider 2012: 172).

Furthermore, the larger proportion of finite CCs in Indian English may be explained by substrate influence and SLA processes as well. As pointed out in Section 2, the four most widely spoken substrate languages in India (Hindi, Telugu, Bengali and Tamil; cf. Eberhard et al. 2019) have an equivalent structure to English declarative CCs. However, only Telugu has a construction similar to the English gerund-participial CC. Therefore, SLA processes such as the Teddy Bear Principle (the tendency to select structures that the speaker knows and thus feels safe using) and the Shortest Path Principle (the tendency to select the structure closer to the one existing in the L1, when several are available) may have influenced speakers to choose finite CCs, which is the option closest to their L1 and hence which they feel most comfortable using, at the expense of non-finite CCs.

Thus, our first hypothesis is confirmed, since discrepancies between British and Indian English seem to be caused by those factors that drive the linguistic evolution of PCEs, namely simplification (as part of innovation), language contact, and SLA.

### 4.2. Factors Conditioning Complement Clause Variation

The results of the manual variable analysis described in Section 3.2 show that none of the factors studied uniquely conditions one outcome or the other. Although certain values trigger one option, they also allow the other. Consider, for example, the results of the variable MEANING CC, shown in Table 4. Although the meaning *event/action* triggers non-finite CCs in both varieties, as illustrated in (12), this meaning is also expressed by means of finite CCs, as in (13).

(12) *I clearly remember my Aunt **buying** a coal mining cottage in Pontypridd (...)*  
(GloWbE GB)

(13) *(...) he remembered he **had bought** a new memory foam mattress (...)* (GloWbE GB)

		Finite CCs		Non-finite CCs		Total	
		No.	%	No.	%	No.	%
GloWbE GB	<i>State</i>	36	46.0	42	54.0	78	100
	<i>Event/action</i>	102	22.0	362	78.0	464	100
	Total	138	25.5	404	74.5	542	100
		No.	%	No.	%	No.	%
GloWbE IN	<i>State</i>	26	65.0	14	35.0	40	100
	<i>Event/action</i>	77	27.5	202	72.5	279	100
	Total	103	32.5	216	67.5	319	100

Table 4. Distribution of finite and non-finite CCs according to the variable MEANING CC<sup>4</sup>

In a similar vein, following the Complexity Principle (cf. Rohdenburg 1996, 2006), which states that complex environments should favour the use of explicit alternatives, in this case finite CCs, we could hypothesise that CCs with no or few constituents would be expressed by means of non-finite CCs, and CCs with four and five constituents would be expressed by means of finite CCs. However, as shown in Table 5, both long non-finite CCs and short finite CCs are found.

148

		Finite CCs		Non-finite CCs		Total	
		No.	%	No.	%	No.	%
GloWbE GB	<i>Short (0-1)</i>	46	23.0	152	77.0	198	100
	<i>Medium (2-3)</i>	84	27.0	229	73.0	313	100
	<i>Long (4-5)</i>	8	26.0	23	74.0	31	100
	Total	138	25.5	404	74.5	542	100
		No.	%	No.	%	No.	%
GloWbE IN	<i>Short (0-1)</i>	21	22.5	72	77.5	93	100
	<i>Medium (2-3)</i>	71	34.5	134	65.5	205	100
	<i>Long (4-5)</i>	11	52.5	10	47.5	21	100
	Total	103	32.5	216	67.5	319	100

Table 5. Distribution of finite and non-finite CCs in the variable COMPLEXITY CC MEASURED IN NO. OF CONSTITUENTS<sup>5</sup>

Tables 4 and 5 show that frequency distribution, although statistically significant, cannot be used to identify which factors are determinant in the choice and which type of CC is favoured in each case; the number of non-finite CCs is higher than that of finite CCs, and thus these are always likely to be favoured to a higher degree. In light of this, frequency distributions alone are clearly not enough to tell us which variables determine the choice. Thus, we turn to a regression model, since we are dealing with probabilistic complementation choice in a context where various factors are in play. The variables analysed, then, will be used as predictors, which in a statistical model such as the one described in the following paragraphs, are factors triggering a particular outcome which are considered to have a certain predictive value.

The next section describes the binary logistic regression analysis of the variables carried out to examine their impact on CC choice. As will be explained below, some variables had to be excluded due to issues related to their collinearity with the dependent or independent variables, and also due to problems in their analyses. Also, some variables were discarded from the model because their results were not significant. These were:

TIME REFERENCE. The analysis confirms that both declarative finite CCs and gerund-participial CCs have an independent —and the same— time reference, since the time of the action in the CC is not conditioned by the time of the action in the MC and is always anterior (cf. also Cuyckens et al. 2014).

TYPE SUBJECT CC. Findings for this variable show that CCs without an expressed subject correspond to non-finite CCs, since non-finite CCs whose subject is co-referential with that of the MC do not have an expressed subject. Therefore, *non-expressed subjects*, rather than conditioning the choice of clause, are a feature of most non-finite CCs.

#### 4.3. Binary Logistic Regression Analysis

Binary logistic regression analysis predicts the choice between two variants (in this case finite and non-finite CCs, that is, the dependent variable) based on a series of explanatory factors (i.e., each of the variables under study; Field 2009: 265). The analysis yields a value “that reflects the chances of one outcome compared with the other outcome for a given combination of values of the predictors” (Levshina 2015: 253). When we apply a logistic regression model “the algorithm tries again and again different sets of values of the model parameters and returns the combination which maximally closely models the actual outcomes” (Levshina 2015: 254). However, before running a logistic regression model, it is necessary to check that the data and the predictors used do not violate the requirements and assumptions of logistic regression.

The crucial value for the interpretation of logistic regression analysis is the odds ratio value (OR). The OR is an indicator of the change in odds that results from a unit change in the predictor and can be interpreted as follows (Field 2009: 270-271; Levhsina 2015: 260):

- If the value is higher than 1 it indicates that as the predictor increases, the odds of the outcome occurring increase (the outcome being one of the values of the dependent variable).
- If the value is lower than 1 it indicates that as the predictor increases, the odds of the outcome occurring decrease.

In this analysis, COMPLEMENTATION TYPE is the dependent variable, with the values finite CCs and non-finite CCs. The results of the regression analysis (namely *predicted odds*) indicate whether the independent variables and their values favour or disfavour non-finite CCs (*the outcome*, in statistical terms). I followed the customary steps to obtain an adequate regression model, that is, I began by establishing a model which included all the potentially relevant language-internal variables (with the exclusion of TIME REFERENCE and TYPE SUBJECT CC, as previously noted). Subsequently, the models, one for each variety of English, were refined by excluding non-suitable variables, that is, those that showed collinearity with another variable or (quasi-)complete separation from the dependent variable.

I performed two binary logistic regression analyses, one for each sample of data, as described below.

## GloWbE GB

In the analysis of the data sample for GloWbE GB, the variables INTERVENING MATERIAL and POLARITY CC were discarded because the distribution of the data did not fulfil the chi-square assumptions. Table 6 lists the significant factors out of all those included in the binary logistic regression model. The factors that favour non-finite CCs are the following: MEANING CC, where the value *event/action* increases the odds for non-finite CC by a factor of 2.477 with respect to the default *state*, and POLARITY MC, where *negative* MCs increase the odds for non-finite CC by a factor of 10.056 against the default *positive* MCs. The shadowed cells in Table 6 show the values that disfavour non-finite CCs. CO-REFERENTIALITY BETWEEN THE MC AND THE CC SUBJECTS has a slightly disfavoured effect on complementation choice; if the value is *non-co-referential*, the odds for non-finite CCs decrease by a factor of 0.949 (that is, 5.1%). As for ANIMACY CC SUBJECT, the value *inanimate subjects* disfavour non-finite CCs robustly by a factor of 0.373 (62.7%).

## Synchronic Variability in the Complementation Profile of *Remember*

	Odd ratios (OR)	95% Confidence interval	
		Lower	Upper
(Intercept)	Not significant		
MEANING CC (default state)			
<i>Event/action</i>	2.477 ***	1.266	4.848
CO-REFERENTIALITY BETWEEN THE MC AND THE CC SUBJECTS (default <i>same</i> )			
<i>Different</i>	0.949 ***	0.099	0.278
POLARITY MC (default <i>positive</i> )			
<i>Negative</i>	10.056 ***	2.715	7.246
ANIMACY CC SUBJECT (default <i>animate</i> )			
<i>Inanimate</i>	0.373 **	0.180	0.775

Table 6. Significant variables in the model and confidence interval at 95% of these variables (GloWbE GB)

151

The 95% confidence interval confirms these results (cf. Table 6). The values that correspond to *event/action* and *negative* MCs display values higher than 1, which means that as the predictor variable increases so do the odds of non-finite CCs. In contrast, the remaining values are lower than 1 (shaded cells in Table 6), which means that as the predictor variable increases the odds of non-finite CCs decrease, which also agrees with the results of the odd ratios.

These results show that MEANING CC, CO-REFERENTIALITY BETWEEN THE MC AND THE CC SUBJECTS, POLARITY MC and ANIMACY CC SUBJECT condition the choice of CC. As was the case in Cuyckens et al. (2014), in my data *non-co-referentiality* between the MC and the CC subject and *inanimate subjects* of the CC disfavour the use of non-finite CCs.

## GloWbE INDIA

For the data in GloWbE IN, the variables TYPE SUBJECT MC, INTERVENING MATERIAL, POLARITY MC and POLARITY CC are not included because they violate the chi-square assumptions. The collinearity test shows no signs of any collinearity issues, so all variables are included in the model. The significant factors, presented in Table 7, indicate that the only one that favours non-finite CCs is MEANING CC, where the

value *event/action* increases the odds for non-finite CC by a factor of 4.346 with respect to the default *state*.

In contrast, the shadowed cells in Table 7 show the values that disfavour non-finite CCs. In terms of CO-REFERENTIALITY BETWEEN THE MC AND THE CC SUBJECTS, *non-co-referential* have a robustly disfavouring effect on complementation choice since the odds for non-finite CCs decrease by a factor of 0.138 (86.2%). COMPLEXITY CC MEASURED IN NO. OF WORDS also disfavors non-finite complementation, since *medium* CCs disfavour non-finite CCs by a factor of 0.357 (64.3%).

152

	Odd ratios (OR)		95% Confidence interval	
			Lower	Upper
(Intercept)	Not significant			
MEANING CC (default <i>state</i> )				
<i>Event/action</i>	4.715	***	2.031	10.945
CO-REFERENTIALITY BETWEEN THE MC AND THE CC SUBJECTS (default <i>same</i> )				
<i>Different</i>	0.138	***	0.076	0.251
COMPLEXITY CC MEASURED IN NO. WORDS (default <i>short</i> )				
<i>Medium</i>	0.357	*	0.137	0.926
<i>Long</i>	Not significant			

Table 7. Significant variables in the model and confidence interval at 95% of these variables (GloWbE IN)

As can be seen in Table 7, the 95% confidence interval values confirm these results, in that *event/action* has values greater than 1. On the other hand, the remaining values (shadowed cells in Table 7) are lower than 1, which means that as the predictor variable increases the odds of non-finite CCs decrease, which also agrees with the results of the odd ratios.

Therefore, the factors that influence the choice of CC in GloWbE IN are MEANING CC, CO-REFERENTIALITY BETWEEN THE MC AND THE CC SUBJECTS and COMPLEXITY CC MEASURED IN NO. OF WORDS. These results are in accordance with Cuyckens et al.'s (2014) study in that *non-co-referentiality* between the MC and the CC subject disfavors non-finite CCs in both studies.

The main results from the regression analysis of the two data samples can be summarised as follows:

- Non-finite CCs are favoured by MEANING CC = *event/action* in both varieties, and by POLARITY MC = *negative* in GloWbE GB.
- Non-finite CCs are disfavoured by CO-REFERENTIALITY = *non-co-referential* in both varieties, by ANIMACY CC SUBJECT = *inanimate* in GloWbE GB and by COMPLEXITY CC IN NUMBER OF WORDS = *medium* in GloWbE IN.

In sum, there are common features that condition the choice of CC within the different sections of GloWbE, no matter whether they represent an L1 or an L2 variety. These are the values *event/action*, which favours the use of non-finite CCs, and *non-co-referential subjects*, which disfavours the use of non-finite CCs. As for differences between the L1 and L2 varieties analysed, in GloWbE GB *negative polarity* of the MC also favours non-finite CCs while *inanimate* CC subjects disfavour their use. In GloWbE IN, however, these values have no apparent effect, whereas *medium* CCs in number of words, that is, CCs between 8 and 14 words long, disfavour the use of non-finite CCs.

#### 4.4. Discussion of Results

153

The results concerning differences in the features that determine the choice of CC between sections of GloWbE are interesting in themselves, but are more relevant if discussed in relation to the Complexity Principle (see hypothesis (b) in Section 2; Rohdenburg 1996, 2006) and to claims that the evolution of PCEs is conditioned, among other factors, by simplification (cf. Williams 1987; Schneider 2007, among others; cf. Section 2). Starting with the Complexity Principle, our hypothesis was that this would be found to be at work especially in Indian English, since PCEs are claimed to prefer simpler structures (finite CCs in this case; cf. Steger and Schneider 2012), which are easier to process and produce. This is exactly what the results here show: in the case of GloWbE IN, a feature involving an increase in the complexity of the CC, particularly longer CCs in number of words, disfavour the use of (less explicit and thus more difficult to process) non-finite CCs. Therefore, the results in Indian English are in line with the Complexity Principle (Rohdenburg 1996, 2006), since more cognitively complex environments disfavour the use of the less iconic option, that is, non-finite CCs. However, in GloWbE GB, which is an L1 and then less likely to be affected by simplification, a feature that adds complexity, the *negative polarity* of the MC, favours the use of non-finite CCs, and thus goes against claims of the Complexity Principle.

Furthermore, these results also tie in with diachronic research on the complementation profile of REMEMBER. Some of the factors that are shown to determine variation between finite and non-finite CCs in Cuyckens et al.'s (2014)

study<sup>7</sup> are the same as in this study; particularly MEANING CC and CO-REFERENTIALITY OF THE MC AND CC SUBJECTS.

Finally, my results seem to confirm some of the claims made in previous research on complementation in PCEs: (i) there are common cross-varietal features shared by varieties of English which predict the choice of syntactic construction and, thus, it seems that the core probabilistic grammar determining the choice is stable across varieties (cf. Schilk et al. 2013; Bernaisch et al. 2014; Szmrecsanyi et al. 2016; Röthlisberger et al. 2017); but (ii) there are also factors not shared among varieties which may be related to SLA phenomena and language contact phenomena and which may indicate that L1 and L2 speakers do not share the same abstract knowledge of the morphosyntactic constraints associated with these constructions (cf. Deshors 2015; Röthlisberger et al. 2017).

In sum, the conclusion derived from these analyses is twofold. On the one hand, there are language-internal factors that condition the choice of CC in the two varieties of English, independently of their status (L1 or L2). On the other hand, there seem to be other language-internal factors conditioning the variation which are not shared between the L1 and L2 varieties and which are in line with claims made in the literature regarding the Complexity Principle and the alleged preference of PCEs for simpler, more transparent and more iconic structures (cf. Williams 1987; Rohdenburg 1996, 2006; Schneider 2007; Steger and Schneider 2012; Brunner 2017, among others).

154

## 5. Conclusion

In this study I have examined the variability in the sentential complementation profile of REMEMBER when it means 'recall' in British and Indian English as represented in GloWbE. I have explored the distribution of finite and non-finite CCs as happening in competition in these two varieties of English and, in order to explore the multivariate nature of complementation choice, I have applied a binary logistic regression analysis so as to shed light on what factors favour non-finite complementation and what these factors are in the case of two specific varieties of English. Finally I considered the relevance of the observed preferences for (i) the alleged factors driving the linguistic evolution of PCEs, and (ii) Rohdenburg's (1996) Complexity Principle.

The most pertinent results from this study are that: (i) although non-finite CCs are the most common type of CC in the two sections of GloWbE, the preference is statistically significant less marked in GloWbE IN, which seems in line with the tendency towards transparency that is often identified in L2 varieties of English (cf.

Williams 1987; Steger and Schneider 2012) since non-finite CCs, the less transparent option, are used less frequently than in the L1 variety. The preference for finite CCs in the L2 variety may also be the result of language contact and the subsequent influence of substrate languages, as four of the most widely spoken, Hindi, Telugu, Bengali and Tamil, can express complementation via CCs equivalent to English declarative CCs and lack a construction equivalent to the English gerund-participial CC (except for Telugu).

The conclusions derived from the variable analysis and the two binary logistic regression analyses are twofold: (i) there are some factors that condition the choice of CC across the two varieties, but (ii) there are also variety-specific factors that condition the choice in the case of each variety. In addition, these results can be connected to the Complexity Principle (Rohdenburg 1996, 2006), they agree partially with diachronic research on the verb *REMEMBER* (cf. Cuyckens et al. 2014), and tie in with previous research on complementation in PCEs.

So, the findings presented here largely corroborate earlier claims on the evolution of PCEs and the expected behaviour of morphosyntactic structures in competition, and thus support the hypotheses formulated in this article. Unlike other studies, however, the present one uses a data set large enough to study the complementation profile of *REMEMBER* and state-of-the-art methodology by applying statistical modelling to CC variation in PCEs, which hence provides a more solid empirical grounding than was the case in previous work.

## **Acknowledgements**

For their generous financial support, I am grateful to the Spanish Ministry of Economy, Industry and Competitiveness (Grants FFI2017-82162-P and BES-2015-071543) and the University of Vigo. I would also like to thank the two anonymous reviewers who provided helpful comments on earlier drafts of the manuscript.

## **Notes**

---

<sup>1</sup>. *REMEMBER* governs CCs with three different meanings: 'recall', followed by finite and non-finite CCs; 'bear in mind the fact', followed by finite CCs; and 'remember to do', followed by non-finite CCs. Therefore,

'recall' is the only meaning with which an envelope of variation is found (cf. Fanego 1996; Mair 2006; Cuyckens et al. 2014).

<sup>2</sup>. Romasanta (2017) explores the complementation profile of the verb *REGRET*,

which, like REMEMBER, allows finite and non-finite CCs in competition, across World Englishes. However, in the distribution of finite and non-finite CCs after REGRET (Romasanta 2017: 136-137) she includes all finite and non-finite CCs, not only those in competition. Therefore, her results are not suitable for comparison.

3. According to Huddleston and Pullum (2002: 1350), supplements are “elements which occupy a position in linear sequence without being integrated into the syntactic structure of the sentence”. Supplements are not accounted for in the variable that measures complexity in number of constituents (cf. Huddleston and Pullum 2002: 1353).

4. The results for both varieties are significant at  $p < 0.05$  (GloWbE GB:  $\chi^2 = 20.556$ ,  $p$ -value = 0.000; GloWbE IN:  $\chi^2 = 20.384$ ,  $p$ -value = 0.000).

5. Only the results for GloWbE IN are significant at  $p < 0.05$  (GloWbE GB:  $\chi^2 = 0.832$ ,  $p$ -value = 0.660; GloWbE IN:  $\chi^2 = 8.402$ ,  $p$ -value = 0.015).

6. \* significant at  $p < 0.05$ , \*\* significant at  $p < 0.01$ , \*\*\* significant at  $p < 0.001$ .

7. These factors are MEANING CC, INTERVENING MATERIAL, TYPE SUBJECT MC, TYPE SUBJECT CC, CO-REFERENTIALITY between the MC and THE CC, COMPLEXITY CC MEASURED IN NO. OF CONSTITUENTS, VOICE CC VERB, ANIMACY CC SUBJECT, TEMPORAL RELATION and PERIOD.

## Works Cited

156

- ANDERSEN, Roger W. 1983. “Transfer to Somewhere”. In Gass, Susan and Larry Selinker (eds.) *Language Transfer in Language Learning*. Rowley, MA: Newbury House: 177-201.
- ANNAMALI, Elay, and Sanford B. STEEVER. 1998. “Modern Tamil”. In Steever, Sanford B. (ed.): 75-99.
- BERNAISCH, Tobias. 2015. *The Lexis and Lexicogrammar of Sri Lankan English*. Amsterdam: John Benjamins. <https://doi.org/10.1075/veaw.g54>
- BERNAISCH, Tobias, Stefan Th. GRIES and Joybrato MUKHERJEE. 2014. “The Dative Alternation in South Asian Englishes: Modelling Predictors and Predicting Prototypes”. *English World-Wide* 35 (1): 7-31. <https://doi.org/10.1075/eww.35.1.02ber>
- BRESNAN, Joan. 1970. “On Complementizers: Toward a Syntactic Theory of Complement Types”. *Foundations of Language* 6 (3): 297-321.
- BRESNAN, Joan. 1979. *Theory of Complementation in English Syntax*. New York: Garland.
- BRESNAN, Joan and Jennifer HAY. 2008. “Gradient Grammar: An Effect of Animacy on the Syntax of Give in New Zealand and American English”. *Lingua* 118 (2): 245-259. <https://doi.org/10.1016/j.lingua.2007.02.007>
- BRUNNER, Thomas. 2014. “Structural Nativization, Typology and Complexity: Noun Phrase Structures in British, Kenyan and Singaporean English”. *English Language and Linguistics* 18 (1): 23-48. <https://doi.org/10.1017/S1360674313000269>
- BRUNNER, Thomas. 2017. *Simplicity and Typological Effects in the Emergence of New Englishes: The Noun Phrase in Singaporean and Kenyan English*. Berlin; Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110521801>
- CUYCKENS, Hubert, Frauke D’HOEDT and Benedikt SZMRECSANYI. 2014. “Variability in Verb Complementation in Late Modern English: Finite vs. Non-finite Patterns”. In Hundt, Marianne (ed.): 182-203. <https://doi.org/10.1017/cbo9781139507226.014>
- DAVIES, Mark. 2013. *Corpus of Global Web-Based English: 1.9 Billion Words from*

- Speakers in 20 Countries (GloWbE). <<http://corpus.byu.edu/glowbe>>. Accessed August 1, 2018.
- DE KLERK, Vivian. 2003. "Towards a Norm in South African Englishes: The Case for Xhosa English." *World Englishes* 22 (4): 463-481. <https://doi.org/10.1111/j.1467-971X.2003.00313.x>
- DE SMET, HENDRIK. 2013. *Spreading Patterns: Diffusional Change in the English System of Complementation*. Oxford: Oxford U.P. DOI:10.1093/9780199812752.001.0001
- DESHORS, Sandra C. 2015. "A Constructionist Approach to Gerundial and Infinitival Verb Complementation Patterns in Native and Hong Kong English Varieties." *English Text Construction* 8 (2): 207-235. <https://doi.org/10.1075/etc.8.2.04des>
- DESHORS, Sandra C. and Stephan Th. GRIES. 2016. "Profiling Verb Complementation Constructions Across New Englishes: A Two-step Random Forests Analysis of *ing* vs. to Complements." *International Journal of Corpus Linguistics* 21 (2): 192-218. <https://doi.org/10.1075/ijcl.21.2.03des>
- EBERHARD, David M., Gary F. SIMONS and Charles D. FENNIG. (eds.) 2019. *Ethnologue: Languages of the World*. 22 edn. <<http://www.ethnologue.com>>. Accessed April 23, 2019.
- FANEGO, Teresa. 1996. "On the Historical Development of English Retrospective Verbs." *Neuphilologische Mitteilungen* 97 (1): 71-79.
- FANEGO, Teresa. 2004. "On Reanalysis and Actualization in Syntactic Change: The Rise and Development of English Verbal Gerunds." *Diachronica* 21 (1): 5-55. <https://doi.org/10.1075/dia.21.1.03fan>
- FANEGO, Teresa. 2007. "Drift and the Development of Sentential Complements in British and American English from 1700 to the Present Day." In Pérez-Guerra, Javier, Dolores González-Álvarez, Jorge L. Bueno-Alonso and Esperanza Rama-Martínez (eds.) *Of Varying Language and Opposing Creed: New Insights into Late Modern English*. Bern: Peter Lang: 161-235.
- FANEGO, Teresa. 2016. "The Great Complement Shift Revisited: The Constructionalization of ACC-*ing* Gerundives." *Functions of Language* 23 (1): 84-119. <https://doi.org/10.1075/fol.23.1.05fan>
- FIELD, Andy P. (2000) 2009. *Discovering Statistics Using SPSS: (and Sex and Drugs and Rock 'n' Roll)*. Thousand Oaks, Los Angeles, CA: SAGE Publications. <https://doi.org/10.1348/000709906x100611>
- GARCÍA-CASTRO, Laura. 2018. "Variation in Verbal Complementation Patterns in Nativised Varieties of English: The Case of REMEMBER in Indian English" In Ferrández San Miguel, María and Claus-Peter Neumann (eds.) *Taking Stock to Look Ahead: Celebrating Forty Years of English Studies in Spain*. Zaragoza: Pressas Universitarias de la Universidad de Zaragoza: 207-213.
- HASSELGREN, Angela. 1994. "Lexical Teddy Bears and Advanced Learners: A Study into the Ways Norwegian Students Cope with English Vocabulary." *International Journal of Applied Linguistics* 4 (2): 237-258. <https://doi.org/10.1111/j.1473-4192.1994.tb00065.x>
- HUDDLESTON, Rodney and Geoffrey K. PULLUM, in collaboration with Laurie BAUER, Betty BIRNER, Ted BRISCOE, Peter COLLINS, David DENISON, David LEE, Anita MITTWOCH, Geoffrey NUNBERG, Frank PALMER, John PAYNE, Peter PETERSON, Lesley STIRLING and Gregory WARD. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge U.P. <https://doi.org/10.1017/9781316423530.001>
- HUNDT, Marianne. (ed.) 2014. *Late Modern English Syntax*. Cambridge; New York: Cambridge U.P.
- IBM Corp. Released 2017. "IBM SPSS Statistics for Windows", version 25.0. Armonk, NY: IBM Corp. <https://www.ibm.com/analytics/us/en/spss/spss-statistics-version/>. Accessed August 16, 2018.
- KACHRU, Braj B. 1985. "Standards, Codification and Sociolinguistic Realism: The English Language in the Outer Circle" In Quirk, Randolph and Henry G. Widdowson (eds.) *English in the World: Teaching and Learning the Language and Literatures*. Cambridge: Cambridge U.P. in association with The British Council: 11-30.
- KACHRU, Yamuna. 2006. *Hindi*. Amsterdam: John Benjamins. <https://doi.org/10.1075/loall.12>

- KAUNISTO, Mark, Mikko HÖGLUND and Paul RICKMAN. (eds.) 2018. *Changing Structures: Studies in Constructions and Complementation*. Amsterdam: John Benjamins.
- KRISHNAMURTI, Bhadriraju. 1998. "Telugu". In Steever, Sanford B. (ed.): 202-240.
- LEITNER, Gerhard. 1994. "Begin and Start in British, American and Indian English." *HERMES - Journal of Language and Communication in Business* 7: 99-122. <https://doi.org/10.7146/hjlc.v7i13.25077>.
- LEVSHINA, Natalia. 2015. *How to do Linguistics with R: Data Exploration and Statistical Analysis*. Amsterdam: John Benjamins. <https://doi.org/10.1075/z.195.website>
- LOUREIRO-PORTO, Lucía. 2017. "ICE vs GloWbE: Big Data and Corpus Compilation." *World Englishes* 36 (3): 448-470. <https://doi.org/10.1111/weng.12281>
- MAIR, Christian. 2006. "Nonfinite Complement Clauses in the Nineteenth Century: The Case of *remember*". In Hytö, Merja, Mats Rydén and Erik Smittberg (eds.) *Nineteenth-century English: Stability and Change*. Cambridge: Cambridge U.P.: 215-228. <https://doi.org/10.1017/CBO9780511486944.009>
- MUKHERJEE, Joybrato and Sebastian HOFFMANN. 2006. "Describing Verb-complementation Profiles of New Englishes: A Pilot Study of Indian English." *English World-Wide* 27 (2): 147-173. <https://doi.org/10.1075/eww.27.2.03muk>
- NAM, Christopher, Sach MUKHERJEE, Marco SCHILK and Joybrato MUKHERJEE. 2013. "Statistical Analysis of Varieties of English." *Journal of the Royal Statistical Society* 176 (3): 777-793. <https://doi.org/10.1111/j.1467-985X.2012.01062.x>
- NIHALANI, Paroo, Ray K. TONGUE and Priya HOSALI. 1979. *Indian and British English: A Handbook of Usage and Pronunciation*. Delhi: Oxford U.P.
- NOONAN, Michael. (1985) 2007. "Complementation". In Shopen, Timothy (ed.) *Complex Constructions*. Vol. 2 of *Language Typology and Syntactic Description*. Cambridge: Cambridge U.P.: 52-150. <https://doi.org/10.1017/CBO9780511619434.002>
- OLAVARRÍA DE ERSSON, Eugenia and Phillip SHAW. 2003. "Verb Complementation Patterns in Indian Standard English." *English World-Wide* 24 (2): 137-161. <https://doi.org/10.1075/eww.24.2.02ers>
- QUIRK, Randolph, Sidney GREENBAUM, Geoffrey LEECH and Jan SVARTVIK. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- RICKMAN, Paul and Juhani RUDANKO. 2018. *Corpus-based Studies on Non-finite Complements in Recent English*. Basingstoke: Palgrave Macmillan.
- ROHDENBURG, Günter. 1996. "Cognitive Complexity and Increased Grammatical Explicitness in English." *Cognitive Linguistics* 7 (2): 149-182. <https://doi.org/10.1515/cogl.1996.7.2.149>
- ROHDENBURG, Günter. 2006. "The Role of Functional Constraints in the Evolution of the English Complementation System." In Dalton-Puffer, Christiane, Dieter Kastovsky, Nikolaus Ritt and Herbert Schendle (eds.) *Syntax, Style and Grammatical Norms: English from 1500-2000*. Bern: Peter Lang: 143-166. <https://doi.org/10.3726/978-3-0351-0272-7>
- ROHDENBURG, Günter. 2007. "Functional Constraints in Syntactic Change: The Rise and Fall of Prepositional Construction in Early and Late Modern English." *English Studies* 88 (2): 217-233. <https://doi.org/10.1080/00138380601042824>
- ROHDENBURG, Günter. 2014. "On the Changing Status of *that*-clauses." In Hundt, Marianne (ed.): 155-181. <https://doi.org/10.1017/CBO9781139507226.013>
- ROMASANTA, Raquel P. 2017. "Contact-induced Variation in Clausal Verb Complementation: The Case of REGRET in World Englishes." *Alicante Journal of English Studies* 30: 121-147. doi:10.14198/raei.2017.30.05
- RÖTHLISBERGER, Melanie, Jason GRAFILLER and Benedikt SZMRECSANYI. 2017. "Cognitive Indigenization Effects in the English Dative Alternation." *Cognitive Linguistics* 18 (4): 673-710. <https://doi.org/10.1515/cog-2016-0051>
- SCHILK, Marco. 2011. *Structural Nativization in Indian English Lexicogrammar*. Amsterdam:

- John Benjamins. <https://doi.org/DOI: 10.1075/scl.46>
- SCHILK, Marco, Joybrato MUKHERJEE, Christopher NAM and Sach MUKHERJEE. 2013. "Complementmentation of Ditransitive Verbs in South Asian Englishes: A Multifactorial Analysis". *Corpus Linguistics and Linguistic Theory* 9 (2): 187-225. <https://doi.org/0.1515/cllt-2013-0001>
- SCHNEIDER, Edgar W. 2007. *Postcolonial English: Varieties Around the World*. Cambridge: Cambridge U.P. <https://doi.org/10.1017/CBO9780511618901>
- SCHNEIDER, Edgar W. 2013. "English as a Contact Language: The "New Englishes". In Schreier, Daniel and Marianne Hundt (eds.) *English as a Contact Language*. Cambridge: Cambridge U.P.: 131-148. <https://doi.org/10.1017/CBO9780511740060.008>
- SHANK, Christopher, Julie VAN BOGAERT and Koen PLEVOETS. 2016. "The Diachronic Development of Zero Complementmentation: A Multifactorial Analysis of *that/zero* Alternative *Think, Suppose, and Believe*". *Corpus Linguistics and Linguistic Theory* 12 (1): 31-72. <https://doi.org/10.1515/cllt-2015-0074>
- SHASTRI, S.V. 1996. "Using Computer Corpora in the Description of Language with Special Reference to Complementmentation in Indian English". In Baumgardner, Robert J. (ed.) *South Asian English: Structure, Use, and Users*. Urbana: University of Illinois: 70-87.
- STEEVER, Sanford B. (ed.) 1998. *The Dravidian Languages*. London; New York: Routledge.
- STEGEER, Maria and Edgar W. SCHNEIDER. 2012. "Complexity as a Function of Iconicity: The Case of Complement Constructions in New Englishes". In Kortmann, Bernd and Benedikt Szmrecsanyi (eds.) *Linguistic complexity: Second Language Acquisition, Indigenization, Contact*. Berlin: De Gruyter Mouton: 156-191. <https://doi.org/10.1515/9783110229226>
- SUÁREZ-GÓMEZ, Cristina. 2017. "Transparency and Language Contact in the Nativization of Relative Clauses in New Englishes". *English World-Wide* 38 (2): 211-237. <https://doi.org/0.1075/ewww.38.2.05sua>
- SZMRECSANYI, Benedikt, Jason GRAFMILLER, Benedikt HELLER and Melanie RÖTHLISBERGER. 2016. "Around the World in Three Alternations: Modeling Syntactic Variation in Varieties of English". *English World-Wide* 37 (2): 109-137. <https://doi.org/10.1075/ewww.37.2.01szm>
- THOMASON, Sarah G. 2001. *Language Contact: An Introduction*. Edinburgh: Edinburgh U.P.
- THOMPSON, Hanne-Ruth. 2012. *Bengali*. Amsterdam: John Benjamins. <https://doi.org/10.1075/loall.18>
- WALD, Benji. 1996. "Substratal Effects on the Evolution of Modals in East Los Angeles English". In Arnold, Jennifer, Renée Blake, Brad Davidson, Scott Schwenter and Julie Solomon (eds.) *Sociolinguistic Variation: Data, Theory and Analysis. Selected Papers from NNAV 23 at Stanford*. Stanford CA: Center for the Study of Language and Information: 513-530.
- WILLIAMS, Jessica. 1987. "Non-native Varieties of English: A Special Case of Language Acquisition". *English World-Wide* 8 (2): 161-199. <https://doi.org/10.1075/ewww.8.2.02wil>
- WONG, Irene F. H. 1983. *Simplified Features in the Structure of Colloquial Malaysian English*. Singapore: University Press for SEAMO Regional Language Centre.

## Appendix 1 - Complementation in substrate languages

India is linguistically complex, with 29 languages spoken by over one million speakers, and a total of 448 living languages, according to Ethnologue (Eberhard et al. 2019). I selected the most widely spoken languages: Hindi (predominant in the north of India), Bengali, Tamil (predominant in the south of India) and Telugu (Eberhard et al. 2019). In all these languages there are structures equivalent to English declarative CCs but only Telugu shows a structure similar to gerund-participial CCs.

In Hindi, finite clauses as objects are introduced by the conjunction *ki* 'that' (Y. Kachru 2006: 217), as can be seen in (1).

- (1) *us ne kaha ki voh pune me nokeri karta he*  
 he AG say.PERF.M.SG that he Pune in job do.IMPF.M.SPRES.SG  
 'He said that he has a job in Pune'. (Y. Kachru 2006: 217)

In Bengali we can find complex sentences which consist of at least one MC and one CC, both finite, which are usually, but not always, linked by a conjunction (Thompson 2012: 195). An example of a complex clause in Bengali is (2) below, where the conjunction *ye*, equivalent to English *that*, may be used (2) or not (3), as happens with expanded and bare declarative CCs in English (Thompson 2012: 195).

- (2) *ami jani ye tumi asbe*  
 I know.1st.PR.SG that you come.FUT  
 'I know that you will come'. (Thompson 2012: 195)
- (3) *ami jani tumi asbe*  
 I know.1st.PR.SG you come.FUT  
 'I know you will come'. (Thompson 2012: 195)

Tamil lacks conjunctions and complementation and coordination are marked through verb morphology. CCs can be finite and non-finite. Finite CCs can be embedded using a set of clitic particles, which combine with the host verb and form a phonological word with it. See for instance (4), where the verb *Varu-v-ān* 'come' has the clitic *ō* attached (Annamalai and Steever 1998: 122).

- (4) *rāman nālaikku Varu-v-ān=ō enakku cantēkam*  
 Raman-NOM tomorrow come-FUT-3.SG.M=or I-DAT doubt-COM  
 'I doubt whether Rama will come tomorrow'. (Annamalai and Steever 1998: 122)

In Telugu any clause can be embedded in another one as a CC if it contains a gerundive, which is formed by adding the suffix *aṭam/adam* to the verb, as in (5), where the verb *ceyyaṭam* 'do' has the *aṭam* suffix and becomes a gerund which complements the verb *mānēsaru* 'stopped' (Krishnamurti 1998: 234). The formation of the gerundive reminds of the English gerund-participial, which is formed by adding the suffix *ing* to the bare form of the verb, as in *remembering*.

- (5) *Āṭsar lēkapōwaṭam walla gumāstālu pani*  
 of the officer not be-GERUND there because the clerks the work  
*ceyyaṭam ye mānēsaru*  
 do-GER stop-PAST  
 'Because of the officer not being there, the clerks stopped doing the work'. (Krishnamurti 1998: 234)

The perfect conjunctive of the verb *ani* 'say' is used as a complementizer to introduce finite CCs with the verb *an* 'to say', as illustrated in (6), where *ani* introduces the CC headed by the verb *ceppenu* 'said' (Krishnamurti 1998: 235).

- (6) *nēnu rēpu wastān(u) ani ceppenu*  
 I tomorrow come-FUT that say-PAST  
 'I said that (I) will come tomorrow'. (Krishnamurti 1998: 235)

## Appendix 2: Variables analysed

### A. Semantic factors

#### I. MEANING CC

Value	Example
<i>State</i>	(1) <i>I remember <b>being</b> impatient with Rom that morning (...)</i> (GloWbE IN)
<i>Event/action</i>	(2) <i>I remember I <b>asked</b> her to buy them.</i> (GloWbE GB)

#### II. TIME-REFERENCE (Noonan 2007)

Value	Example
<i>ITR</i>	(3) <i>(...) Math should remember <b>that correlation does not equal causation.</b></i> (GloWbE GB)
<i>DTR</i>	(4) <i>(...) people need to remember <b>to be patient.</b></i> (GloWbE IN)

161

### B. Structural factors

#### III. TYPE SUBJECT MC

Value	Example
<i>pronominal subject</i>	(5) <i>I hadn't remembered writing a suicide note.</i> (GloWbE G)
<i>non-pronominal</i>	(6) <i><b>Jamie</b> remembers that it was he who left his bait to putrefy and cause a ghastly stink as Vin did at "Starlings".</i> (GloWbE GB)
<i>non-expressed subject</i>	(7) <i>(...) it's important <b>Ø</b> to remember to have a good sense of balance throughout your time at university; (...)</i> (GloWbE GB)

#### IV. TYPE SUBJECT CC

I use the same values for the subject of the CC as I use for that of the MC (see (III) above).

V. COMPLEXITY CC MEASURED IN NO. OF CONSTITUENTS

Value	Example
Zero	(8) (...) <i>and remember it can be renewed!</i> (GloWbE GB)
One	(9) (...) <i>I do remember doing <u>my Ranger promise</u>... #</i> (GloWbE GB)
Two	(10) <i>I remember telling <u>him to 'take me home'</u>.</i> (GloWbE GB)
Three	(11) <i>I remember that I also published my first guest post on HBB (...)</i> (GloWbE IN)
Four	(12) <i>I remembered then that I had heard Matthew there the first morning too.</i> (GloWbE GB)
Five	(13) <i>I remember first picking up this book from my brother's bookshelf when I was thirteen (...)</i> (GloWbE IN)

VI. COMPLEXITY CC MEASURED IN NO. OF WORDS

Value	Example
1-7	(14) <i>I don't remember <b>that the older series was that drab.</b></i> (GloWbE GB)
8-14	(15) <i>I remember <b>this captured my imagination as a small child.</b></i> (GloWbE GB)
+15	(16) <i>I distinctly remember <b>the Conservatives in the 2005 Election and before that saying immigration needed to be controlled.</b></i> (GloWbE GB)

162

VII. INTERVENING MATERIAL

Value	Example
Presence	(17) <i>I remember <b>very clearly</b> that in 1993 most SYs (including myself) were waiting expectantly for the Sahaj leadership to put the collective desire into action (...)</i> (GloWbE IN)
Absence	(18) (...) <i>they could remember being born (...)</i> (GloWbE GB)

VIII. VOICE CC VERB

Value	Example
Active	(19) (...) <i>I don't remember my predictions <b>going wrong.</b></i> (GloWbE IN)
Passive	(20) <i>The screams of people that she remembered <b>being echoed</b> earlier was reduced to a faint buzzing in the back of her head.</i> (GloWbE IN)

## Synchronic Variability in the Complementation Profile of *Remember*

### IX. SUPPLEMENTATION

Value	Example
<i>Presence</i>	(21) (...) <i>we need to remember that a section of the society —the armchair critics— like to say things are bad.</i> (GloWbE IN)
<i>Absence</i>	(22) <i>I don't remember having confessed.</i> (GloWbE GB)

### X. COORDINATION

Value	Example
<i>Presence</i>	(23) <i>He also remembers <u>that on that November Sunday in Calcutta, he was involved in the run-out of Allan Border, and that he dropped John Emburey on the cover-point boundary.</u></i> (GloWbE IN)
<i>Absence</i>	(24) <i>I remember watching Barfi!</i> (GloWbE GB)

### XI. POLARITY MC

Value	Example
<i>Positive</i>	(25) <i>I <b>remember</b> writing an Astro Boy story when I was about 6.</i> (GloWbE GB)
<i>Negative</i>	(26) <i>I <b>can't remember</b> having seen another national team play the way Spain do.</i> (GloWbE IN)

163

### XII. POLARITY CC

Value	Example
<i>Positive</i>	(27) <i>But I also remember <b>reading</b> a comment piece by you a few years ago (...)</i> (GloWbE GB)
<i>Negative</i>	(28) <i>I remember <b>not wanting</b> to tell anyone about it.</i> (GloWbE GB)

### C. Additional factors

### XIII. ANIMACY CC SUBJECT

Value	Example
<i>Animate</i>	(29) <i>I remember <b>my class teacher</b> coming late to the film (...)</i> (GloWbE IN)
<i>Inanimate</i>	(30) <i>I remember <b>a party</b> happening on the day we launched (...)</i> (GloWbE GB)

XIV. CO-REFERENTIALITY BETWEEN THE MC AND THE CC SUBJECTS

This variable indicates whether the subjects of the MC and of the CC are the same or not.

Value	Example
<i>Co-referential</i>	(31) <i>I remember Ø earning about 120 a day in the 1980s from the council pool (...)</i> (GloWbE GB) (32) <i>I remember once I heard this story of a lady walking on the beach after a terrible storm.</i> (GloWbE IN)
<i>Non-co-referential</i>	(33) <i>I remember <b>you</b> had earlier mentioned that you would raise around may be Rs 1400 crore of total debt.</i> (GloWbE IN)

XV. COMPLEMENTATION TYPE

Value	Example
<i>Finite</i>	(34) <i>I still remember <b>that last year I got an interview call from a very renowned company (...)</b></i> (GloWbE IN) (35) <i>But I remember <b>I stayed up all night thinking that even if my heart is into it, I can not sell my family off.</b></i> (GloWbE IN)
<i>Non-finite</i>	(36) <i>I remember <b>many representatives of countries telling me a single decision of the COP (Conference of the Parties) is an agreed outcome (...)</b></i> (GloWbE GB) (37) <i>(...) a few lines from a hymn which I remember <b>to have repeated from my earliest boyhood (...)</b></i> (GloWbE IN)

164

Received: 14 November 2018

Accepted: 2 May 2019