

***HELP* or *HELP to*: what do corpora have to say?¹**

Tony McEnery and Zhonghua Xiao
Lancaster University

[Abstract] In this paper, we will examine a range of factors that may potentially influence a language user's choice of a full or bare infinitive following *HELP*. The factors include language variety, language change, spoken/written distinction, semantic distinction, and syntactic conditions, namely, an intervening noun phrase or adverbial, the number of intervening words, *to* preceding *HELP*, the passive construction, inflections of *HELP*, and *it* as the subject. Six corpora are used in this paper, four written corpora (LOB, Brown, FLOB and Frown) and two spoken corpora (the speech section of the BNC and the Corpus of Professional Spoken American English, CPSA).²

[Key words] corpus, language variety, language change, register, syntax

1. INTRODUCTION

HELP is a frequent verb of English, with a distinctive syntax, that has generated on-going debate amongst language researchers. As such it is a verb that is often given some prominence in textbooks and grammars (e.g., Chalker 1984: 106; Eastwood 1992: 106; Murphy 1985: 110) though the treatment of the verb can be poor (e.g., Close 1988; Dixon, 1991: 199; Duffley 1992: 27-29; Quirk, Greenbaum, Leech & Svartvik 1972: 841). For example, all of these authors maintain the choice of a full or bare infinitive after *HELP* is determined by a semantic distinction between the two (cf. section 5). In this paper we will take a corpus-based approach to improve the description of the verb and to test claims made about the verb in the literature. We will also explore variation in that description between two major varieties of English, British English (BrE) and American English (AmE). In addition, we will investigate how *HELP* has varied diachronically and by register in these varieties. First, however, the claim that *HELP* is a frequent verb of English with distinctive syntactic properties must be justified.

HELP is one of the most frequent words in the English language, ranking as 245th in the word frequency list of the British National Corpus

¹ We would like to thank Susan Hunston and two readers of an earlier version of this paper for their useful comments. In this paper we use capitals *HELP* lemma to distinguish it from the word form *help*.

² The four written corpora used in this paper are well suited to the study of language change and contrasting British English and American English. Leech (2002), for example, uses these corpora to study recent grammatical changes in British English. These corpora are also used in Mair (1995, 2002) to study the changing patterns of verb complementation.

(BNC).³ When its inflected forms *helped*, *helps* and *helping* are included, there are 528.62 instances of *HELP* per million words. When we look at the most frequent verbs (lemmatized) in the BNC, *HELP* rises to 72nd in the word frequency list. Furthermore, *HELP* is the only verb that can both control either a full infinitive or a bare infinitive and occur either with or without an intervening noun phrase (NP),⁴ as in the following examples cited from the BNC:

- (1) (a) **HELP to V**
Perhaps the book *helped to prevent* things from getting even worse.
(b) **HELP NP to V**
I thought I could *help him to forget*.
(c) **HELP V**
Savings can *help finance* other Community projects.
(d) **HELP NP V**
We *helped him get* to his feet and into a chair.

In this paper, we will examine the factors that may potentially influence a language user's choice of a full infinitive or a bare infinitive as the object or object complement of *HELP*.⁵ Our work is based on the relative frequencies of *HELP* in six corpora, as shown in Figure 1. All of these corpora are used to explore the potential syntactic and semantic conditions that may be relevant to the choice of a full or bare infinitive with *HELP*.

The four written English corpora were compiled using the same sampling frame, each containing 500 segments sampled from 15 text categories, each corpus totalling one million words. LOB (*The Lancaster-Oslo-Bergen Corpus of British English*) and FLOB (*The Freiburg-LOB Corpus of British English*) represent British English (BrE) in 1961 and 1991, while Brown (*A Standard Corpus of Present-Day Edited American English*) and Frown (*The Freiburg-Brown Corpus of American English*) represent American English (AmE) in the same periods.⁶ The corpus of spoken AmE used in this paper is the Corpus of Professional

³ This paper is based on BNC version 2, which is accessible online at URL <http://escorp.unizh.ch/cgi-bin/bnc2/BNCquery.pl>

⁴ Biber, Johansson, Leech & Finegan (1999: 735) note that *dare* and *help* are the two main clause verbs that can control either a *to*-clause or a bare infinitive. Only *help*, however, can take an intervening noun phrase followed by either a full or bare infinitive (cf. also Chalker 1984: 149). Thus, while *to* in (1d) can be left out, it cannot in *Ernest...dared Archie to punch him in the stomach* (Frown).

⁵ As a reader of an earlier version of this paper points out, when an NP intervenes between *help* and an infinitive (as in patterns 1b and 1d), the intervening NP can possibly be analyzed as the object of the first clause or the subject of the second clause (biclausal analysis). In our monoclausal analysis, this NP is object of *help* while the infinitive functions as the object complement. When there is no intervening NP, the infinitive functions as the object of *help*. This paper will not include infinitives functioning as adverbials of purpose.

⁶ See corpus manuals (<http://khnt.hit.uib.no/icame/manuals/index.htm>).

Spoken American English (CPSA),⁷ including over two million words of conversations occurring between 1994 and 1998 in the context of professional activities broadly tied to academics and politics. The corpus of spoken BrE we use is a subcorpus we defined within the spoken component of the British National Corpus (BNCS), totalling around 6.43 million words. To make BNCS more representative of BrE and more comparable to CPSA, the subcorpus only includes language uttered between 1985 and 1994 by speakers whose first language is BrE.⁸

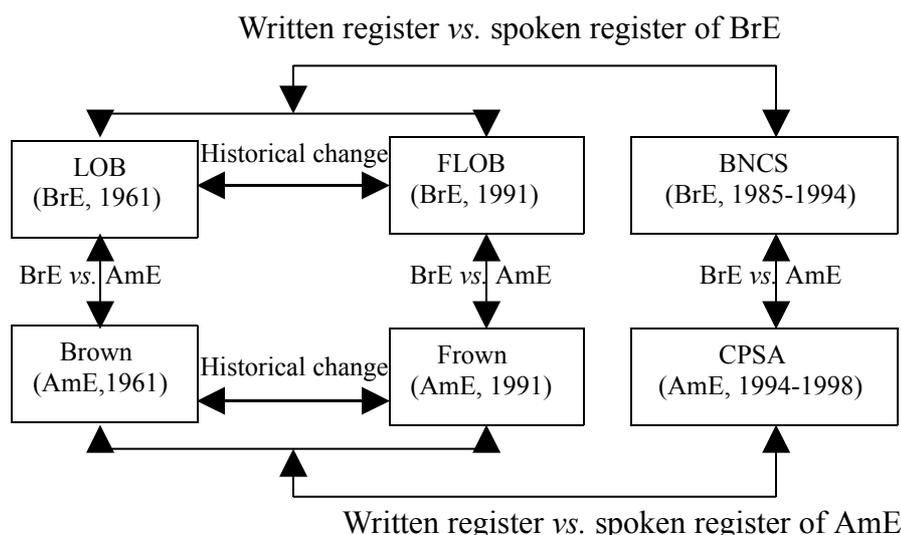


Figure 1: Corpus data

This paper is organized as follows: section 2 contrasts the BrE data and the AmE data to see whether the variety of English has an effect on the language user's choice; section 3 compares frequencies in LOB/Brown and FLOB/Frown to show the effect of language change over three decades; section 4 is concerned with factors relating to the spoken/written distinction; section 5 examines the effect of the alleged semantic distinction between a full infinitive and a bare infinitive; section 6 discusses the potential influences of syntactic conditions on the use of *HELP*; and section 7 concludes the paper.⁹

⁷ A detailed description and a sample of the corpus is available online at URL <http://www.athel.com/cpsa.html>.

⁸ Considering that a time span of less than 10 years is not likely to change the grammar of a language drastically, we assume that the slight difference in the sampling periods of the two spoken corpora will not affect our results significantly.

⁹ In addition to the factors discussed in this paper, infinitival verbs and text categories may also influence the choice of a full or bare infinitive. For example, *pay* tends to take the bare form as in *help pay*. However, the discussion of collocation and distribution needs much larger corpora to achieve reliable quantification. In the four written corpora used in this paper, the most frequent verb *make* only occurs 23 times in the positions of 1st-4th collocates on the right side of *help* as a verb. We would like to thank Professor Wolfgang Teubert for suggesting this line of inquiry.

2. LANGUAGE VARIETY

To examine the potential effect of the variety of English on *HELP*, we extracted all of the instances of *HELP*, including its inflected forms (e.g., *helps*, *helped* and *helping*), from the six corpora and classified each occurrence according to the four-fold classification in (1). The frequencies of the full and bare infinitives in the BrE and AmE corpora are shown in Figure 2. Note that the frequencies in the figure are total counts of the relevant usage of infinitives in both the data of the 1960s and the 1990s, and in both written and spoken corpora.

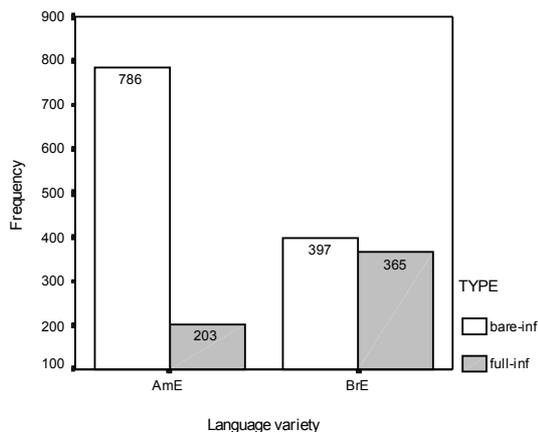


Figure 2: Contrasting BrE and AmE

As sample sizes may affect the level of statistical significance, raw frequencies must be first normalized to a common base.¹⁰ Of the six corpora used in this paper, four (Brown, Frown, LOB and FLOB) are one million words in size. Therefore, unless otherwise stated, the raw frequencies (RF) of CPSA and BNCS are normalized as frequencies per million words in order to facilitate the comparison of the six corpora. Table 1 shows both raw and normalized frequencies of infinitive variants in the AmE and BrE data.¹¹ The last two columns of the table indicate the LL (log likelihood) ratio calculated on the basis of normalized frequencies and the significance level¹².

For 1 degree of freedom (df), the critical value of significance at $p < 0.001$ is 10.83, much less than the calculated log likelihood value (LL) in Table 1. Therefore, we can confidently conclude that the difference in usage of *HELP* between BrE and AmE is statistically significant with respect to

¹⁰ A *normalized frequency* (NF) refers to a weighted frequency measure that allows for easy and reliable comparison of data sets of different sizes (cf. Ball 2002: 11).

¹¹ The counts do not include i) the instances with *it* as the provisional subject; ii) infinitives functioning as adverbials of purpose. The same applies to all of the frequencies given in this paper unless otherwise stated.

¹² Unless otherwise specified, the values for significance tests and significance levels in this paper are calculated using SPSS Release 10.1.

the choice of a full or bare infinitive. Our finding is in line with the observation of Biber *et al* (1999: 735) that ‘AmE has an especially strong preference for the pattern *verb + bare infinitives* although the bare infinitive is more common than the *to*-infinitive in both varieties.’ However, a more refined view of the differences between AmE and BrE emerges if we compare the three pairs of comparable corpora separately.

Table 1: Contrast between BrE and AmE

Variety	Full-inf		Bare-inf		LL (1 df)	Sig. level (2-sided)
	RF	NF	RF	NF		
BrE	365	43.30	397	47.04	23	<0.001
AmE	203	50.75	786	196.50		

Table 2 shows the results of such a comparison. As can be seen from the table, LOB and Brown (with an LL value of 65.265), which represent written BrE and written AmE in 1961, contrast more strikingly than FLOB and Frown (with an LL value of 24.805). For the moment we will simply note this difference, though we will return to it in section 3. The difference between the two spoken corpora (with an LL value of 18.393) is roughly similar to the FLOB/Frown difference rather than to the LOB/Brown difference. Interestingly, the spoken data is nearly contemporaneous with FLOB and Frown.

Table 2: Full infinitives and bare infinitives in BrE and AmE corpora

Corpus	Full-inf		Bare-inf		LL (1 df)	Sig. level (2-sided)
	RF	NF	RF	NF		
LOB	95	95	27	27	65.265	<0.001
Brown	58	58	125	125		
FLOB	78	78	121	121	24.805	<0.001
Frown	45	45	204	204		
BNCs	192	29.86	249	38.72	18.393	<0.001
CPSA	100	50	457	228.5		

The following example illustrates the British preference for *to*-infinitives:

- (2) You are going to *help me* make *to make* a birthday cake for Jim remember. (BNC)

The repair in this utterance is telling. The speaker first utters *You are going to help me make* but immediately changes the utterance to use the full infinitive.

By the wording ‘British preference’, we do not mean that full infinitives are more frequent in British English. Rather, the British preference for full infinitives is in relation to the domination of bare infinitives in the AmE data. As Figure 2 shows, bare infinitives account for nearly 80% in the AmE data, whereas in the BrE data they only make up about 52%. Bare infinitives are prevalent in AmE simply because this construction is of American provenance, though it has been penetrating

rapidly into BrE (cf. Lind 1983: 264; Onions 1965). Zandvoort (1966) classified this construction as an Americanism and claimed that ‘except in American English, however, *to help* usually takes an infinitive with *to*’ (cf. Lind 1983: 264). However, if we take language change into account, which we will do in section 3, we find Zandvoort’s claim does not hold any longer – *HELP* no longer necessarily takes a full infinitive in BrE; rather, the bare infinitive has also become the statistical norm in BrE (cf. also Mair 1995: 264; 2002:124).

3. LANGUAGE CHANGE

Language change over time has affected the choice of a full or bare infinitive following *HELP*. The bare infinitive after *HELP* was pronounced to be now dialectal or vulgar in the *Oxford English Dictionary* (1st ed., 1933). The *Supplement to the OED* (1989) removed this label and judged it as being ‘a common *colloq.* form’ (cf. Kjellmer 1985: 264). There is evidence that even the 1933 *OED* was not reflecting reality, however; Mair (2002: 123), based on the quotation base of the *OED*, observed a rapid increase for bare infinitives from the mid nineteenth century onwards. As such, Vallins’s (1951: 56) claim that ‘the construction is not seriously questioned now (as it might have been twenty years ago) even in normal literary writing’ is credible. Certainly, by 1991, a bare infinitive after *HELP* ‘lost the informal ring formerly associated with it’ (Mair 1995: 268).

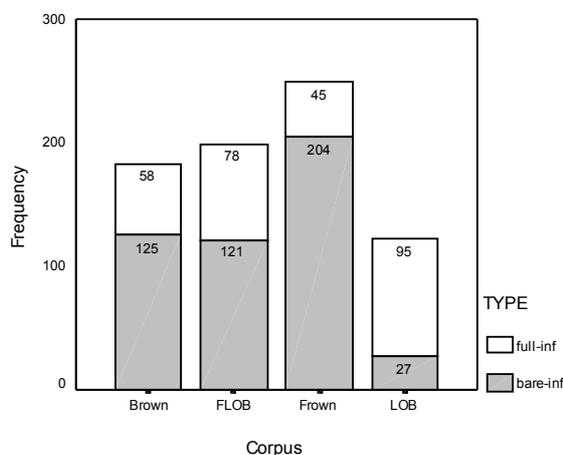


Figure 3: Frequencies in the four written corpora

Given that there is some evidence of language change related to *HELP*, this section examines recent data to demonstrate the possible effect of language change on the language user’s choice. We will only consider written English because the four written corpora used in this paper are perfect for this purpose. Figure 3 shows the relevant frequency data from the

four corpora.¹³ It can be seen from the figure that the proportion of the bare infinitives in both BrE and AmE data have increased over the period 1961-1991.

Table 3: Contrast between written English in 1961 and 1991:

Sample Period	Corpus	Full-inf	RF	Bare-inf	RF	LL (1 df)	Sig. level (2-sided)
1961	LOB	95	153	27	152	40.143	<0.001
	Brown	58		125			
1991	FLOB	78	123	121	325		
	Frown	45		204			

Table 3 shows the frequencies of the full and bare infinitives in the data for English in 1961 and 1991. As the written AmE and the written BrE data are of equal size, normalization is not needed. The calculated log likelihood value in the table is much greater than the critical value of 10.83 for significance at $p < 0.001$. Therefore, it can be argued confidently that language change over the three decades has indeed exerted influence over the language user's choice between the two infinitive variants. It is also interesting to note in the table that there is a marked increase in the total occurrence of HELP, in both the BrE and AmE data. For the moment, we will simply note this increase, though we will return to it in section 5.

Table 4: Changes in written BrE and AmE

Variety	Corpus	Full-inf	Bare-inf	% of bare-inf	LL (1 df)	Sig. level
BrE	LOB	95	27	22.13%	47.575	<0.001
	FLOB	78	121	60.80%		
AmE	Brown	58	125	68.31%	10.678	0.001
	Frown	45	204	81.93%		

As can be seen in Table 4, in the 1960s, bare infinitives account for only 22% of the BrE data, but this percentage rose to 60% in the 1990s; In the AmE data, there was also an increase, from 68% to 82%, in the proportion of bare infinitives. But the change in AmE is not as marked as that in BrE, as reflected by the lower significance level and smaller LL value for the AmE data. The difference between FLOB and LOB (LL=47.575) is significant at $p < 0.001$, whereas the significance level between Frown and Brown (LL=10.678) is 0.001. The reason for this apparent difference is that by 1961 AmE was already much more tolerant of bare infinitives than BrE (see Figure 2). Consequently a greater shift towards the use of bare infinitives in the period 1961-1991 was possible for BrE, resulting in a more marked change. It is clear that by the 1990s, the bare infinitive has become the statistical norm also in BrE. But even so, the British use full infinitives more frequently than Americans.

¹³ While it may be desirable to group the four corpora by the creation date or language variety in the figure, SPSS automatically arranged them alphabetically.

4. THE SPOKEN/WRITTEN DISTINCTION

Written language differs from spoken language in many respects, one of which is that speech is typically less formal than writing and thus more tolerant of variant forms. Earlier studies of *HELP* show that of the two variants of *HELP (NP) to do* and *HELP (NP) do*, the former is the original one and the latter a later development (cf. Kjellmer 1985: 158). As such, bare infinitives are predicted to be more common in spoken English than in written English. This prediction is generally supported by our corpus data. As can be seen in Figure 4, except in written BrE,¹⁴ bare infinitives occur more frequently in the spoken data than in the written data. In spite of the slightly larger proportion of bare infinitives in spoken English, however, we cannot conclude that the spoken/written distinction actually influences the language user's choice, as shown by the statistical test conducted below.

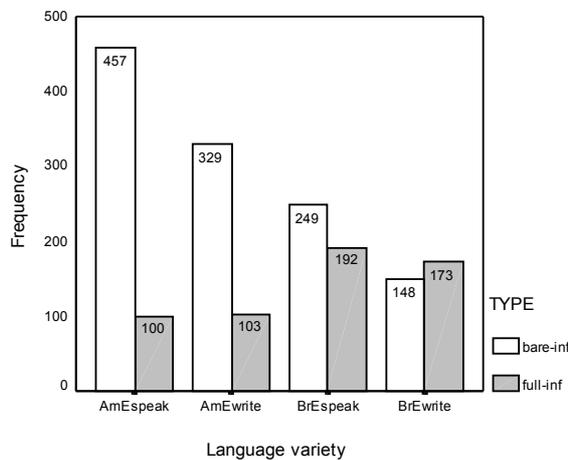


Figure 4: Full/bare infinitives following *HELP* in spoken and written English

To test the statistical significance of this difference, all of the raw frequencies were normalized to one million words, as shown in Table 5. For the difference to be statistically significant, the calculated log likelihood ratio must be greater than 3.84, the critical value for significance at $p < 0.05$ for df of 1. Table 5 shows that irrespective of whether we consider the written and spoken data in BrE and AmE separately, or ignore the language variety and take the written and spoken data in the two language variety together, the significance level is greater than 0.05 and not statistically significant. Even if we disregard the effect of language change (cf. section 3) and compare the written and spoken data of the matching period (see Table 6), we come to the same conclusion: while bare infinitives occur more frequently in spoken English, the spoken vs. written distinction does not

¹⁴ The relatively low frequency of bare infinitives in written BrE is attributable to the unusually higher frequency of full infinitives in LOB, data of the 1960s (77.87%).

significantly influence a language user's choice between the two infinitive variants.

Table 5: Contrast between spoken and written registers

Register	Full-inf		Bare-inf		LL (1 df)	Sig. level (2-sided)
	RF	NF	RF	NF		
BrEwrite	173	86.5	148	74	2.159	0.142
BrEspeak	192	29.86	249	38.72		
AmEwrite	103	51.5	329	164.5	2.711	0.100
AmEspeak	100	50	457	228.5		
Written	276	69	477	119.25	1.746	0.186
Spoken	294	34.88	706	83.75		

Table 6: Contrasting the spoken and written data of the matching period

Corpus	Full-inf		Bare-inf		LL (1 df)	Sig. level (2 sided)
	RF	NF	RF	NF		
FLOB	78	78	121	121	0.389	0.533
BNCs	192	29.86	249	38.72		
Frown	45	45	204	204	0.002	0.964
CPSA	100	50	457	228.5		
Written	123	61.5	325	162.5	0.132	0.716
Spoken	292	34.64	706	83.75		

5. SEMANTIC DISTINCTION

The debate over the semantic distinction between the two versions of infinitive has a long history (see Duffley 1992:1-14). While most researchers content themselves with stating that the omission of *to* after *HELP* is optional, a few others see a subtle semantic distinction between the two variant forms. Wood (1962: 107-108) and Lu (1996: 813), for example, argue that *to* 'can be omitted only when the helper does some of the work, or shares in the activity jointly with the person that is helped' (Wood, *ibid*). In other words, when the helper does not take part in the activity with which the help is offered, the infinitive *must* take *to*, as in (3a). Thus sentences like (3b) and (3c) are unacceptable according to Lu and Wood.

- (3) (a) This book *helped me to see* the truth. (Lu, *ibid*)
 (b) These tablets will *help you sleep*. (Wood, *ibid*)
 (c) Writing out a poem will *help you learn* it. (Wood, *ibid*)
- (4) (a) Will you *help me clear* the table? (Quirk *et al* 1972: 841)
 (b) This book will *help you to see* the truth. (Quirk *et al* 1972: 841)
- (5) (a) John *helped Mary eat* the pudding. (Dixon 1991: 199)
 (b) John *helped Mary to eat* the pudding. (Dixon 1991: 199)

Similarly to Wood, Quirk *et al* (1972: 841) argue that the choice of the infinitive variants 'is conditioned by the subject's involvement.' For example, in (4a) with a bare infinitive, 'external help is called in' whereas in

(4b) with a full infinitive, ‘assistance is outside the action proper.’ Similar views can also be found in Dixon (1991: 199), who argues that in (5a) John ate part of the pudding as Mary did, whereas in (5b) John presumably fed the pudding to an invalid Mary. Quirk *et al* (1985: 1206), though, drop the semantic distinction and claim that the only contrast being that the bare infinitive is more American.

Duffley (1992: 14, 18) uses the following minimal pairs to argue for a semantic distinction between the two infinitival variants:

- (6) (a) I saw him be impolite.
(b) I saw him to be impolite.
- (7) (a) I had nine people call.
(b) I had nine people to call.

Duffley suggests that there is a general difference in the aspectual properties of the bare and full infinitives: the bare infinitive evokes ‘a perfective view of the realization of an event’ (action-like or state-like) while the full infinitive evokes ‘an action situation referred to a point in time prior to its realization’. Thus in (7a) the bare infinitive ‘evokes the actual realization of the action of calling from beginning to end in the past time-stretch referred to by *had*’ (*ibid*: 18) whereas in (7b) *call* is supposed to follow the existence of the obligation to realize this event, denoted by *had*. On careful examination, however, we find that *saw* and *had* have different meanings, and the different readings of these minimal pairs come as a result of a lexical shift rather than the presence or absence of *to* before the infinitive. In (6a) *saw* refers to visual perception whereas in (6b) it is related to mental apprehension, or the realization of his being impolite by means of inference (cf. also Bolinger 1974: 66). Likewise, in (7a), *had* has a causative meaning while in (7b), *had* simply means ‘possess’, thus the sentence can be interpreted as *I will call nine people, and these people are my (real or fictitious) calling list*. Hunston (2002: 139) argues, on the basis of collocations, that the three main meanings of *maintain* (‘do not allow to weaken’, ‘say something strongly’ and ‘keep at a particular level’) might as well be treated as three phraseologies with their own meaning rather than as a single word with three meanings as a traditional dictionary does. We believe the same applies to *see* and *have* in (6) and (7). As long as we can approach the difference in these sentences from the semantic difference encoded in full verbs, rather than aspectual properties of the full and bare infinitives, we will not pursue this issue further here.

The semantic difference between the infinitival variants is not reported in more recent corpus-based works such as *The Longman Language Activator* (1993), *The Collins CoBUILD English Dictionary* (1995), *The Longman Dictionary of Contemporary English* (1996), and *The Longman Grammar of Written and Spoken English* (1999). The *Collins* dictionary, for example, defines *HELP* thus:

If you help someone, you make it easier for them to do something, for example by doing part of the work for them or by giving them advice or money. (*Collins*, p. 787)

The *Collins* definition does not draw a distinction between whether the helper actually shares or does not share the helping activity. One of their examples is:

- (8) My mum used to *help cook* the meals for the children. (*Collins*, p. 787)

Without more contextual information, it is not clear whether the mother did the actual cooking herself or helped the children, perhaps, by means of simply giving advice on how to cook or relieving the children from such chores as vacuuming the floor so that they could cook. But the most reasonable reading is that the mother did the cooking herself, yet the bare infinitive *cook* is used. The *Activator's* examples, as quoted in (9), are even more illustrating as they certainly seem to counter the semantic distinction:

- (9) (a) My mother's death was a very difficult time for me but my boyfriend *helped me get* over it. (*Activator*, p. 604)
(b) If I write a list, it *helps me remember* all the things I have to do in a day. (*Activator*, p. 606).

Yet assessing the claims of Lu, Wood, and Quirk *et al* on a large scale is made difficult by two factors. Firstly, most of the examples in our corpora do not cover the scenario discussed by these authors. Secondly, where an example may fit the desired scenario, it is in fact hard to make the distinction between whether or not the helper actually takes part in the helping activity. Nevertheless, the following examples provide enough evidence to undermine the semantic distinction as suggested by Lu, Wood and Quirk *et al* as being an absolute one (cf. also Lind 1983: 271):

- (10) (a) Good field techniques will not only equip linguists for better work, but also *help them overcome* negative attitudes. (Brown)
(b) Historical antecedents *help us understand* the current debate and the absence of a perfect solution to the dilemma of war coverage. (Frown)
(c) Mrs. Clinton, before she came up here today, gave a tour of the White House and the personal residence to one of the physical therapists that will be added to the White House Medical Unit team that will *help the President convalesce* when he leaves the hospital. (CPSA)
(d) And there's nothing like a poultice to *help you get* to sleep. (LOB)
(e) I *help people stop* smoking. (FLOB)
(f) Well you oh it says if you have a dose last thing at night it *helps you sleep*. (BNC)

In none of these cases, with either an animate or inanimate subject (i.e. the helper) could the helper have actively involved in helping activity, yet the bare infinitive was chosen. As such, Duffley suggests that

A better characterization of the bare infinitive structure in these uses is that it evokes ‘helping’ as direct or active involvement in the bringing into being of the action denoted by the infinitive...In contrast, *HELP* + *to* evokes help as a condition which enables the helpee to realize the event denoted by the infinitive. (Duffley 1992:28)

This characterization, however, does not add much to the argument for the semantic distinction, because there is little difference between direct/active involvement and direct/active participation discussed above. The two are practically equivalent. Duffley uses the examples in (11) to support the distinction he makes:

- (11) (a) Mrs Arthur Goldberg, wife of the Secretary of Labour, paints professionally and *helps sponsor* the Associated Artists’ Gallery in the District of Columbia. (Brown)
(b) The Bonaventure was quivering and lurching like an old spavined mare. Her stern was down and a sharp list *helped us to cut* loose the lifeboat, which dropped heavily into the water. (Brown)

According to Duffley (*ibid*: 24), it is not acceptable to use *to sponsor* in (11a) while *to cut* in (11b) cannot be replaced with the bare infinitive *cut*. However, we cannot see any contextual difference between the sentences in (11) and (12):

- (12) (a) What a thoughtful company are Ford Motors. They don’t only *help to sponsor* Sky’s TV Soccer but close down a factory and various assembly lines so that their workers will have time to watch! (BNC)
(b) Opportunity 2000 in Kingfisher has *helped us unlock* rich reserves of talent among our employees [...] (BNC)

Just as Mrs Arthur Goldberg could be actively involved in sponsoring an art gallery (11a), Ford Motors could sponsor a football match in the same way, because ‘the only way to help sponsor something is to sponsor it in part by contributing money oneself’ (Duffley 1992: 138). Yet the full infinitive was used in (12a).¹⁵ Similarly, the subjects in (11b) and (12b) are both inanimate,

¹⁵ In addition to (11a), Duffley (1992:26) provides two other examples in which ‘only the bare infinitive seems appropriate’. But we cannot see any difference between his examples and *Table 1.6 helps to provide an overall picture of the content, pattern and distribution of first degree courses* (BNC) and [...] *aid programs that are helping to provide immunization for children around the world* [...] (CPSA). Duffley appears to have come to this conclusion because the pattern is rare in AmE and is simply absent in Brown, the corpus he used in his study.

and *HELP* in both sentences means *to enable* or *to facilitate*. However, one sentence uses the full infinitive whereas the other uses the full infinitive. It would appear that neither AmE nor BrE in fact display the distinction claimed by Lu, Wood, Quirk *et al* and Duffley. Consequently, we claim that the sentences in (3) are perfectly acceptable.

Another issue that is related to the semantic distinction is the hypothesis that *HELP* preceding a bare infinitive is progressively grammaticalized as a modal idiom/catenative or ‘quasi-auxiliary’ (Mair 1995:270; 2002:124). Based on his observation that the use of *HELP* with infinitives (especially bare infinitives) has started mushrooming since the mid nineteenth century, Mair (*ibid*) argues that the meaning of *HELP* has become so general, and abstract (*contribute to/provide a favourable environment for*) that ‘it approaches those typically associated with grammatical categories’. While Mair is right that the increase of the use of *HELP* with infinitives in general and bare infinitives in particular is attributable to the extension of the meaning of *HELP*, we cannot see a link between this increase and the grammaticalization of *HELP*. In what way has grammaticalization contributed to the increase of the use of *HELP* with infinitives? Is it that only a grammatical word increases over time while a lexically full verb does not? Such issues cannot be addressed fully here, but clearly beg future investigation.

Another problematic finding of Mair (2002: 125) is based on his use of the 132 instances of the *to help* + full verb sequence in the whole BNC corpus to argue that this sequence should be analyzed as an ‘auxiliate’ rather than two separate infinitival clauses arranged in sequence. This argument seems to us to be ill founded. If the sequence is to be analysed as a modal idiom, the infinitival verb following *HELP* must be definite. For example, *going* in the ‘quasi-modal’ (Biber *et al* 1999: 484) *be going to* cannot be replaced with *coming*. Unfortunately, both in the BNC as used by Mair, and in the other corpora we use, we cannot find a clear pattern in the infinitival verbs following *to help*. Furthermore, the examples that we found were mainly in the BrE data. The normalized frequencies (per million words) are given as follows: 2 in LOB, 2 in FLOB, 1.24 in BNCS and 0.5 in CPSA. No instances were found in Brown and Frown (see Table 11 in section 6.3). It seems unusual at best, and would be quite unreasonable at worst that grammaticalization should occur in BrE alone.

As Mair (2002:122, 124) observes, bare infinitives have increased considerably in BrE, especially from the 1930s and 1940s onwards. Our discussion in section 3 also shows that, in both BrE and AmE, bare infinitives increased significantly over the period 1961-1991. If *HELP* is indeed in the process of grammaticalizing as an auxiliary, as Mair claims, there should be, by now, some clear sign of this process. Yet, apart from a considerable increase in the frequency of bare infinitives, we cannot find any evidence showing that sentences like (13a) are becoming acceptable. In

contrast, though, *need* and *dare* can be used both as a main verb and as a modal auxiliary, and thus (13b) and (13c) are quite felicitous:

- (13) (a) *Helped Mrs Arthur Goldberg sponsor the art gallery?
(b) Do we need/dare to escape? (Duffley 1992: 99)
(c) Need we/Dare we escape? (*ibid*)

Biber *et al* (1999:483-484) observe that ‘the boundary between modals and lexical verbs taking infinitive complementation is in some cases unclear.’ Examples include so-called marginal or quasi-modals like *need (to)*, *dare (to)*. It is reasonable to assume that these semi-modals derive from lexical verbs and are undergoing a process of grammaticalizing as full modals.

Mair (2002:125) provides the following examples from the quotation base of *OED* to support his grammaticalization hypothesis:

- (14) (a) Sir Kingsley Wood ... asked the House for another £1,000,000,000, to *help pay* for the next three months of war.
(b) Nor have they eliminated the unburned hydrocarbons which *help produce* the smog that blankets such a motor-ridden conurbation as Los Angeles.
(c) Negro cabbie John W. Smith, whose arrest for ‘tailgating’ a police car ... *helped spark* five days of rioting ..., was found guilty of assaulting a policeman.
(d) Part of the fun of the game comes in ‘sooping’. This is when the players sweep the ice with special brooms in front of a moving stone to *help it go* further.

According to Mair, replacing the bare infinitive *pay* with the full infinitive *to pay* in (14a) ‘would not only be stylistically clumsy because of the repetition involved; it would also produce a slight shift in perspective, from the instrument (money) to the agent who spends it.’ While we agree with the first half of his argument, we cannot accept the second half. Consider the example (15a):

- (15) (a) Money raised from tolls on roads will *help to pay* for the scheme. (BNC)
(b) The diesel also produces 90% less carbon monoxide, 60% fewer oxides of nitrogen and 90% fewer of unburnt hydrocarbons which *help to produce* acid rain. (BNC)
(c) Where the fund of damage is likely to be substantial, including future nursing costs and the like as well as loss of earnings, an accountant’s evidence can *help the court to decide* the multiplier, as well as the multiplicand, for example in the case of a one-man business [...] (BNC)

In this sentence the full infinitive *to pay* is used, yet no agent is

mentioned at all. An infinitive marker clearly does not necessarily produce a shift in perspective. One also, must, therefore doubt the reliability of Mair's proposed paraphrase test for this feature.

Mair (2002:126) argues that adding *to* before the infinitives in (14b) and (14c) 'would be slightly incongruous' because the negative effect featured in the two sentences (i.e., smog and rioting) are 'incompatible with the core semantics of *HELP*'. Nevertheless, this argument is poorly postulated, as it is not uncommon for examples featuring negative effects to take infinitives with *to*, as shown in (15b).¹⁶

Finally, Mair (*ibid*) argues that as (14d) is 'a fairly clear case of a purely causative use of *HELP*, equivalent to *make*', 'adding *to* before the infinitive is problematical.' This statement raises two problems. First, we do not see why the causative use of *HELP* should be analyzed as a modal idiom, because *make* is not a modal auxiliary. Second, we cannot find any reason why Mair should claim that the causative use of *HELP* cannot take a full infinitive, because counter examples are not uncommon, as exemplified in (15c).

On the basis of our exploration of AmE and BrE corpus data, we claim that not only is the semantic distinction between the full and bare infinitives following *HELP* not well grounded, it is also the case that the grammaticalization hypothesis is not justified.

6. SYNTACTIC CONDITIONS

A number of syntactic conditions have been suggested in the literature that may be related to the choice of a full or bare infinitive following *HELP*. In this section we will discuss the following factors:

- an intervening NP or adverbial
- the number of intervening words
- *to* preceding *help*
- the passive construction
- inflections of *HELP*
- *it* as the subject

6.1 The intervening NP or adverbial

Biber *et al* (1999: 73), Lind (1983: 269) and Kjellmer (1985: 158) observe that bare infinitives occur more frequently after *HELP* with an intervening NP than where there is no intervening NP. This observation is partially supported by our data, as shown in Table 7.

¹⁶ Here are some more examples, which are all cited from the BNC corpus: *to undermine the Weimar Republic and open the way to Hitler*, *to accelerate global warming*, *to destroy the market*, *to destroy the precious rain forest*, *to disrupt international trade*.

