Describing verb-complementational profiles of New Englishes

A pilot study of Indian English*

Joybrato Mukherjee and Sebastian Hoffmann
University of Giessen / University of Zurich

The present paper investigates the emergence of local norms in Indian English at the level of verb complementation, an area which so far has not attracted much attention in research into New Englishes. In attempting to describe the verb-complementational profile of Indian English, we offer a pilot study which combines a descriptive aim and a methodological aim. At the descriptive level, the present article focuses on ditransitive verbs and their complementation and addresses two related questions: (1) To what extent do the frequency and distribution of complementation patterns of specific ditransitive verbs (e.g. *give*) differ between Indian English and British English? (2) To what extent is the basic ditransitive pattern with two object noun phrases (e.g. in *he sent Mary his warmest wishes*) associated with different verbs in British English and Indian English? The present paper reveals that in both regards there are clear and identifiable differences in verb complementation between the two varieties. At the methodological level, this pilot study combines the use of balanced and representative subcorpora from the International Corpus of English (ICE) with the in-depth analysis of a much larger database that has been extracted from the Internet archive of the daily Indian newspaper *The Statesman*. This makes it possible to also detect examples of low-frequency constructions in Indian English, e.g. sporadic cases of ditransitive complementation of verbs such as *advise*, *gift* and *impart*.

* Some of our results were presented at the Joint Conference of the International Computer Archive of Modern and Medieval English (ICAME) and the American Association for Applied Corpus Linguistics (AAACL) in Ann Arbor, MI in May 2005. We would like to thank the audience for a fruitful and stimulating discussion. We also thank Rosemary Bock, Sarah Chevalier, Magnus Huber and Gunnel Tottie, who read earlier versions of the manuscript and provided us with very helpful feedback.

English World-Wide 27:2 (2006), 147–173.
ISSN 0172–8865 / E-ISSN 1569–9730 © John Benjamins Publishing Company

Keywords: Corpus linguistics, Indian English, verb complementation, ditransitive verb, corpus compilation, web-derived corpus

1. Introduction

Over the past decades, the emergence of New Englishes around the world has led to an ever-increasing interest in the description of differences between varieties of English at all linguistic levels. In this context, it has been frequently noted that many characteristic features of New Englishes tend to cluster around the interface between lexis and grammar (cf. e.g. Schneider 2001: 140), including, for example, collocations and idioms, particle verbs, article usage and tense usage (cf. e.g. Skandera 2003; Schneider 2004; Sand 2004). Another important field in the core area of lexicogrammar showing clear traces of regional differentiation is verb complementation. As Mair (2002) reports in a recent corpus-based study into on-going changes in British and American English, differences in verb complementation between varieties of English are usually gradual in nature and can be observed at an early stage in terms of more or less subtle shifts in frequency. Thus, it is of particular importance to analyse large and representative corpora in order to identify quantitative differences between varieties in a viable manner.

As for Indian English (IndE), which is the largest variety of English in Kachru's (1985) "Outer Circle", verb complementation so far has not attracted wide-spread interest among linguists in general and corpus linguists in particular. A notable exception is Olavarría de Ersson and Shaw's (2003) recent analysis of selected complementation patterns of specific verbs in three IndE

Table 1. Complementation of *pelt* in IndE and BrE newspapers (cf. Olavarría de Ersson and Shaw 2003: 154)

Patterns of PELT	Example	Indian newspapers	British newspapers
V NPgoal	*they pelted the man*	2	5
V NPobject	*they are pelting cans*	28	–
V NPgoal *with* NPobject	*they are pelting him with cans*	14	**92**
V NPobject *at* NPgoal	*they are pelting cans at him*	**42**	3
V NPobject *on* NPgoal	*they are pelting cans on him*	13	–
V NPobject adverbial	*they pelted stones across the road*	–	–
Others		1	–
N (total size of sample)		100	100
p		< 0.001	

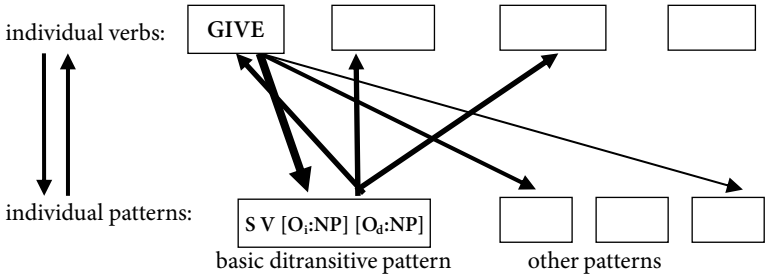

Figure 1. Two complementary aspects of a verb-complementational profile — focus on ditransitive verbs and ditransitive verb complementation

on-line newspapers. As Table 1 shows, they find, for example, that while the most frequent pattern of *pelt* in British newspapers is the *with*-pattern, this is not the case in Indian newspapers.

In the light of results like the ones given in Table 1, Olavarría de Ersson and Shaw (2003: 138) argue that "[v]erb complementation is an all-pervading structural feature of language and thus likely to be more significant in giving a variety its character than, for example, lexis". While we are wary of arguing as forcibly against the importance of lexis for the formation of individual varieties of English, we do agree with Olavarría de Ersson and Shaw (2003) on the point that verb complementation has so far been underestimated as an area of the language system in which regional differentiation figures prominently.

The present paper is largely inspired by Olavarría de Ersson and Shaw's (2003) study, both from a descriptive and a methodological perspective. At the descriptive level, our aim, too, is to shed new light on differences between varieties of English in verb complementation. We will focus on ditransitive verbs and ditransitive verb complementation and sketch out aspects of what we wish to call the "verb-complementational profile" of IndE. As shown in Figure 1, the concept of verb-complementational profile includes two complementary aspects: (a) the range (and frequencies) of the patterns of an individual verb in a variety;[1] (b) the range of verbs with which an individual pattern is associated.

We will show that IndE, which according to Schneider's (2003) dynamic model of the evolution of New Englishes is in many regards a typical example of an endonormatively stabilised variety (cf. Mukherjee 2005b), has a verb-complementational profile which diverges from British English (BrE); as our data show, this divergence is not categorical but of gradual nature.

1. Figure 1 is based on the assumption that all patterns in which a ditransitive verb like *give* is used are manifestations of a ditransitive verb complementation (cf. sec. 2).

At the methodological level, we take Olavarría de Ersson and Shaw's (2003) study as a starting point for a much more systematic utilisation of the worldwide web as a corpus-linguistic database for research into varieties of English. Specifically, we will combine traditional standard-size corpora with large amounts of data obtained from the worldwide web. We will start off by comparing the British and the Indian components of the International Corpus of English (ICE; cf. Greenbaum 1996), both of which are 1 million words in size and include 300 spoken text samples (60%) and 200 written text samples (40%).[2] Since low-frequency phenomena like new verb complementation patterns are very often not attested in the relatively small 1-million-word corpora, it makes sense to extract much larger databases from the web. In the present study, we will apply the procedure that Hoffmann (in press) has described for the extraction of CNN transcripts from the web to the compilation of a 31-million-word database made up of texts from the on-line archive of the Calcutta-based national Indian newspaper *The Statesman*. It will thus be shown how analyses of corpora and web-derived databases can fruitfully complement each other in the description of New Englishes like IndE, for which large corpora (of the size of, say, the British National Corpus) are not available.

The plan of the paper is as follows. In the following section, we will compare the complementation patterns of the ditransitive verbs *give* and *send* in IndE and BrE on the basis of ICE-India and ICE-GB (cf. sec. 2). Then, we will zoom in on the basic ditransitive pattern with two object noun phrases and discuss the range of verbs occurring in this pattern in IndE and BrE by using data from the Internet; in this context, we will elaborate on how homogeneous and well-defined web-derived databases can be generated for linguistic analysis (cf. sec. 3). The findings of our analysis will be discussed in Section 4. Finally, since our paper should best be seen as a pilot study, we will offer some concluding remarks on the descriptive and methodological issues raised in the present paper and sketch out some vital prospects for future research (cf. sec. 5).

2. Complementation patterns of ditransitive verbs: ICE-GB vs. ICE-India

In a previous large-scale study of the British component of the International Corpus of English (ICE-GB), Mukherjee (2005a) provides a detailed quantitative and qualitative analysis of the complementation of ditransitive verbs in

2. For details of the design of ICE see Nelson (1996).

present-day BrE. The analysis is based on a categorisation of all patterns with which a given ditransitive verb is attested in five basic types. These basic types are described and exemplified in (1–5) by instances of the most frequent ditransitive verb in English, i.e. *give*; all examples are taken from ICE-GB and ICE-India.

(1) a. **I** (S) GIVE [O_i:NP] [O_d:NP]
b. On Tuesday members of Parliament **gave** the government their overwhelming support <ICE-GB:S2B-030 #54>

(2) a. **II** (S) GIVE [O_d:NP] [O_i:PP_{to}]
b. I meant to **give** it to you earlier <ICE-GB:S1A-022 #176>

(3) a. **III** (S) GIVE [O_d:NP] ~~O_i~~
b. he wanted physical love and I couldn't **give** that <ICE-GB:S1A-050 #184>

(4) a. **IV** (S) GIVE ~~O_i~~ ~~O_d~~
b. The other major point he raises is in addressing the question of "why **give** in the first place?" <ICE-GB:W1A-011 #94>

(5) a. **V** (S) GIVE [O_i:NP] ~~O_d~~
b. I didn't **give** Lakshmi I had just given Sumi you know <ICE-India: S1A-098 #69–70>

Type I is the most basic type of ditransitive complementation with both objects realised as noun phrases. Type II differs from the basic type I in that the indirect object is realised as a *to*-phrase and placed after the direct object. In type III, the indirect object is omitted. At first sight, this pattern seems to be a manifestation of monotransitive verb complementation, but there are good reasons why type III should also be regarded as a realisation of ditransitivity. On the assumption that ditransitive verbs are intricately linked to what Goldberg (1995) calls the "ditransitive construction" at a cognitive level, any instance of the verb *give* invokes an event-type in which three argument roles (i.e. the "agent" = X, the "recipient" = Y, the "patient" = Z) are involved in a transfer event with the "ditransitive meaning" of 'X causes Y to receive Z'. Of course not all argument roles need to be made explicit in all contexts; not only Goldberg (1995), but many others working in different frameworks (e.g. Matthews 1981; Jackson 1990; Newman 1996; Biber *et al.* 1999) agree on the point that specific elements may be omitted because, for example, they can be recovered from the context or inferred from world knowledge. Thus, from a cognitive point of view *give* retains its ditransitive meaning in contexts like (3b). Consequently, we also

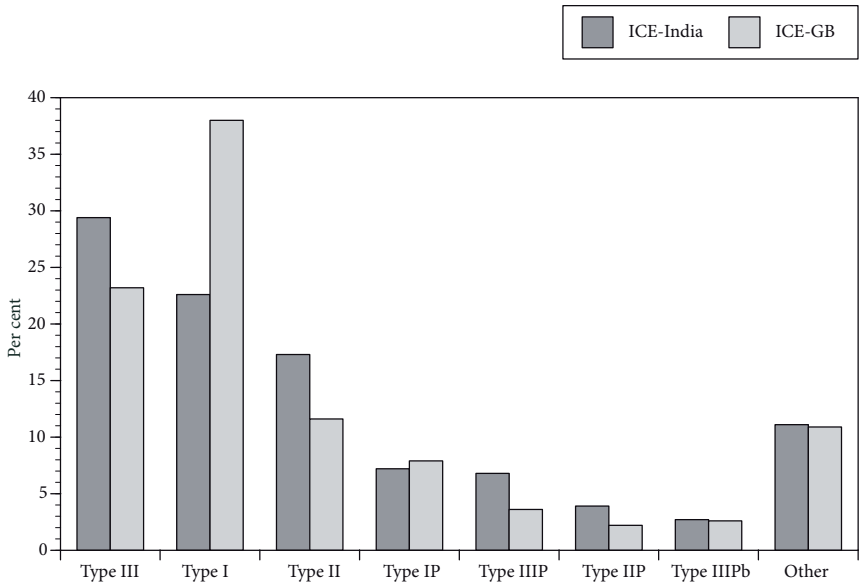

Figure 2. Complementation of *give* in ICE-India and ICE-GB — focus on the most frequent patterns.

allow for the type-IV pattern, in which both objects are deleted, and the type-V pattern, in which the direct object is not made explicit (while the indirect object is). Note that from each of the five basic patterns, it is possible to derive structurally related patterns, e.g. with specific elements in fronted positions, relative clause structures, participle constructions and passive constructions; these derivative patterns are represented by their own formulas (e.g. type IP for a passive pattern derived from the type-I pattern). A full list of all the patterns of *give* and *send* — the verbs that we will focus on in this section — in ICE-GB and ICE-India is given in Appendices 1 and 2.[3]

Figure 2 provides an overview of the frequency and distribution of the most frequent patterns of *give* in ICE-GB and ICE-India. These patterns account for more than 90% of all instances of *give* in the two corpora.

3. The syntactic analyses underlying all individual pattern formulas and the factors that make language users prefer specific patterns over others in particular contexts are described in detail in Mukherjee (2005a). Note in this context that the verb complementation patterns that we posit are much more specific than the relatively abstract traditional basic clause patterns (as, for example, described in Quirk *et al.* 1985 and Biber *et al.* 1999). However, our patterns remain more abstract than the relatively concrete lexicogrammatical patterns in Hunston and Francis's (2000) pattern grammar approach.

Before discussing the data presented in Figure 2, it should be mentioned in passing that — as Appendix 2 reveals — the total number of occurrences of the verb *give* differs considerably in the two corpora (1797 in ICE-India vs. 1064 in ICE-GB). One could thus hypothesise that speakers of IndE are much more likely to express a transfer event by way of *give* as the most prototypical ditransitive verb. However, to confirm this interpretation a detailed analysis of the complete range of possible verb choices would be required, which is beyond the scope of this paper. As far as relative frequencies are concerned, the most significant divergence between ICE-GB and ICE-India visualised in Figure 2 refers to the frequency and distribution of the type-I, type-II and type-III patterns. The markedly different frequency of the type-I pattern is a particularly interesting finding. It seems that *give* is a less prototypical ditransitive verb in IndE — in fact, it is used more frequently in the type-III pattern with only one explicit object than in the type-I pattern, which is by far the most frequent pattern in BrE. Note also that the type-II pattern, in which the patient precedes the recipient, is more frequent in IndE than in BrE.

In this context, it should be noted that there are various reasons for the assumption that the frequency of occurrence of a verb in the type-I pattern determines to a large extent the prototypicality of its status as a ditransitive verb. One of the major reasons, on which much work in construction grammar rests, is the interpretation of the type-I pattern as a syntactic realisation of a cognitive event-type which can be described, as mentioned above, along the lines of the "ditransitive meaning" of 'X causes Y to receive Z' (cf. Goldberg 1995) — this cognitive event-type is called the "ditransitive situation schema" by Beermann (2001). In this context, *give* is the most prototypical ditransitive verb — in BrE at least — because its meaning represents the very core of the ditransitive situation schema:

> In recent work on ditransitive verb semantics such as Goldberg (1995), the core meaning of ditransitive constructions is identified as *successful transfer* between a *volitional agent* and a *willing recipient*. Goldberg claims that '*give*' represents the '*conceptual archetype*' for ditransitive constructions which expresses the meaning 'CAUSE-RECEIVE'. (Beermann 2001: 5; italics in original)

In some studies, *give* is even ascribed the status of a "pathbreaking verb" (cf. Ninio 1999) which opens up the acquisition of the ditransitive construction

and the ditransitive meaning in early child language.[4] Against this background, the much lower frequency of the type-I pattern of *give* in IndE, as shown in Figure 2, is quite remarkable.

What are the reasons for these findings? It seems to us that there are at least two avenues that could be explored in future research in order to find reasons for the divergences in ditransitive verb complementation: (1) cultural motivations; (2) language-internal motivations.

Cultural motivations are difficult to pinpoint. Nevertheless, Olavarría de Ersson and Shaw (2003) speculate on the potential of a cultural difference between European and South Asian cultures as a possible *explicans* for differences in verb complementation between BrE and IndE:

> Northern European cultures could have been more influenced by subjectivism, and see the individual as being at the center of the world, while South Asian cultures might tend to view the individual as a part or a small object in a larger whole. If this were so, when offered the choice between two syntactic structures that focus either on what is provided or on the recipient to express more or less the same thing, BrE speakers would be likely to profile the recipient more frequently in their use of language than their Indian counterparts do, whereas IndE speakers will be more likely to use the structure which profiles what is provided rather than the structure where the recipient is profiled. (Olavarría de Ersson and Shaw 2003: 159)

Such a difference in cultural perspective would no doubt explain the larger proportion of the type-III pattern in ICE-India. On the other hand, one could argue that it fails, for example, to account for the larger proportion of the type-II pattern in IndE, given that in the type-II pattern the recipient is in end-focus position and could thus be viewed as being "profiled".

A more promising avenue seems to be a language-internal explanation for the differences in the complementation of *give*. One particularly striking phenomenon in IndE is the use of *give* in a light-verb construction, i.e. a construction with *give* as a semantically fairly empty ("light") verb and a lexically more specific ("heavy") noun (e.g. *give explanation*). This construction is semantically equivalent to the use of the simplex root verb (e.g. *explain*) from which the noun in the light-verb construction is derived. In IndE, light-verb constructions of this kind, which are also linked to many collocations that are unusual in BrE, can be found in particular in the type-II pattern, as exemplified in (6–9):

4. Campbell and Tomasello (2001), however, show that children may also use less prototypical verbs in their first sentences with double-object constructions.

(6) …explanatory hypothesis which should **give** explanation to [~ explain] the problem <ICE-India:W1A–015 #107>
(7) They are not **giving** development to [~ developing] their career that is my point <ICE-India:S1A–088 #151>
(8) …because he **gave** provocation to [~ provoked] the deceased <ICE-India: S2A–067 #82>
(9) …when uh he heard the accused **giving** uh threat to [~ threatening] Mashid uh that if he lodged complaint he will be killed <ICE-India: S2A–064 #68>

It should also be noted that it is not unusual for IndE speakers to use phrases with *give* where other verbs would be more common in BrE. Examples (10–13) are cases in point:[5]

(10) Our fifth standard also is **giving** problem <ICE-India:S1A–087 #87>
(11) If the students give uh **gives** any complaint uhm or if they place a letter in the suggestion box about the performance of the professors … <ICE-India:S1A–027 #208>
(12) Let them **give** viable proposals. We will consider <ICE-India:W2C–019 #96>
(13) Mr Angale while giving the evidence **give** small sentences because it is to be typed <ICE-India:S1B–061 #28>

At this stage, it seems as though usages as the ones exemplified in (6–13) could be one reason for the higher frequency of the type-II and the type-III pattern of *give* in IndE. It remains to be seen in future research whether they are also responsible for the higher overall number of occurrences of *give* in ICE-India (cf. Appendix 2). As already pointed out, it is not possible for us to offer any conclusive answers to the why-question, and the potential reasons for the findings we have just sketched out are, at best, tentative and warrant further investigation.

It has become obvious, however, that there exist clear and identifiable differences between IndE and BrE at the level of verb complementation. In the field of ditransitive verbs, this applies not only to the prototypical verb *give*, but also to other members of the class of ditransitive verbs. Figure 3 provides an overview of the frequency and distribution of the most frequent patterns of *send* in ICE-GB and ICE-India. These patterns cover around 85% of all instances of *send* in the two corpora.

5. The verbs *cause* in (10), *make* in (11) and (12) and *speak in* in (13) would be much more idiomatic choices in BrE.

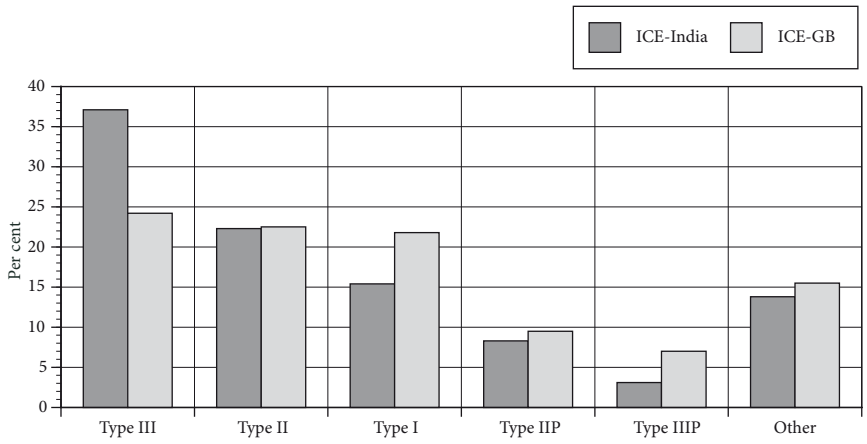

Figure 3. Complementation of *send* in ICE-India and ICE-GB — focus on the most frequent patterns

Again, there is a significant difference between ICE-GB and ICE-India in the type-I pattern, which is used less frequently in IndE than in BrE.[6] On the other hand, the type-III pattern of *send* occurs much more frequently in ICE-India than in ICE-GB. The type-II pattern, however, occurs similarly frequently in ICE-India and ICE-GB. Generally speaking, the case of *send* seems to corroborate a general tendency in IndE to use ditransitive verbs in combination with a direct object only and to use the basic ditransitive pattern — as the most prototypical representation of the ditransitive situation schema — with a significantly lower frequency.

In the light of these findings, it would be very interesting in future research to analyse a syntactically parsed corpus of IndE and to identify all instances of the patterns at hand with regard to all verbs, i.e. to map the range of attested verbs onto each individual complementation pattern. This would point to the other side of the concept of verb-complementational profile (i.e. the "bottom-up" perspective in Fig. 1). In the following section, we will zoom in on one individual pattern, the basic ditransitive pattern, and discuss the range of verbs that are attested in this pattern in present-day Standard IndE. This pattern is of particular importance because it can be seen — from a construction-grammar point of view — as an argument structure construction that is linked to the

6. The figures in Appendix 3 reveal that *send*, too, is more frequent overall in IndE. As for *give*, one could hypothesise that IndE speakers tend to use *send* instead of other, semantically related verbs — a hypothesis the verification of which would again require a detailed analysis of the range of verb choices for each individual occurrence of *send*.

cognitive event of TRANSFER (see above). According to Goldberg's (1995: 5) basic scene hypothesis, it is thus a very "simple" construction in that it encodes a basic event-type: "Simple clause constructions are associated directly with semantic structures which reflect scenes basic to human experience". Specifically, we are interested in unveiling differences between IndE and BrE with regard to the range of verbs that are used in the type-I pattern as a basic construction in English. In this context, we will make use of a large newspaper archive that we have generated from the worldwide web.

3. Focus on "new" ditransitives in Indian English: From corpora to web resources

From a syntactic perspective, it is the occurrence of a verb in the type-I pattern that defines the verb as a member of the class of ditransitive verbs. Among others, Green (1974), Goldberg (1995) and Mukherjee (2005a) discuss many examples of verbs that tend not to be used ditransitively but that are nevertheless attested in the type-I pattern, for example:

(14) **Kiss** me a kiss.
(15) **Cry** me a river.
(16) …she now **had** him a room in the basement… (Grisham 2002: 25)

In Goldberg's (1995) terminology, these usages are licensed because the verbal meaning is such that it can be "fused" with the meaning of the ditransitive construction 'X causes Y to receive Z', e.g. by way of metaphorical extension. For example, the meaning of *cry* in (15) is extended to a process of TRANSFER in which a river of tears is transferred from the agent to the recipient (by means of crying, that is). Mukherjee (2005a: 204) subsumes this process of allowing new verbs into the class of ditransitive verbs under "grammatical institutionalisation", which is defined as "a process in the course of which potential linguistic forms (here: potentially ditransitive verbs such as *provide* [in BrE]), are licensed to become possible forms (here: grammatically institutionalised ditransitive verbs)". Grammatical institutionalisation of "new" ditransitives is an on-going and creative process in all norm-producing native varieties of English, as examples (14–16) reveal.

There is general agreement that the grammatical institutionalisation of "new" ditransitives is also part of the development of local norms in institutionalised second-language varieties of English. Just as there are differences between native varieties of English with regard to the range of verbs that are

Table 2. "New" ditransitives in IndE

Verb (cf. Nihalani *et al*. 2004, Olavarría de Ersson and Shaw 2003)	**ICE-India** **Number of occurrences in the type-I pattern**	**ICE-GB** **Number of occurrences in the type-I pattern**
convey	2	type-I pattern not attested
enquire / inquire	0	
entrust	0	
furnish	1	
inform	4	
present	1	
provide	24	
put	0	
supply	0	

grammatically institutionalised as ditransitive verbs (e.g. *provide* in American and BrE), second-language varieties like IndE deviate from other native and non-native varieties in allowing a different range of verbs in the basic ditransitive pattern. The left-hand column in Table 2 gives a selection of "new" ditransitives that usually do not occur in the type-I pattern in present-day BrE but that have been described as ditransitive verbs in IndE by Nihalani *et al*. (2004) and Olavarría de Ersson and Shaw (2003).

Apart from *provide* — which is also well-established in the type-I pattern in native varieties like American English[7] — Table 2 reveals that there are only very few relevant instances or no instances at all in the 1-million-word ICE-India. Clearly, a much larger corpus is required to investigate low-frequency phenomena such as the incipient grammatical institutionalisation of "new" ditransitives. Given the lack of such a large-scale corpus of IndE, we therefore decided to follow Olavarría de Ersson and Shaw's (2003) approach and turned to the worldwide web as a source of additional data. However, in contrast to Olavarría de Ersson and Shaw (2003), we preferred not to rely on the available on-line search facilities but rather to create our own web-derived corpus of Indian newspaper English. This enabled us to search for relevant patterns in the resulting database with the help of Perl ("Practical Extraction and Report

7. According to various American native-speaker informants, also other verbs are increasingly admissible in the type-I pattern, e.g. *present*. This confirms our observation that the membership of the class of ditransitive verbs is not fixed and that the grammatical institutionalisation of new ditransitive verbs is an on-going creative process both in native and in non-native varieties of English.

Language") and its powerful regular expression engine. As we will demonstrate below, such an approach can reveal aspects of language use that would otherwise have escaped our attention.

3.1 Generating a database of *The Statesman* from the worldwide web

In recent years, scholars have become increasingly interested in the Internet as a source of empirical linguistic data (cf. e.g. Kilgarriff and Grefenstette 2003). However, although the worldwide web certainly dwarfs any of the currently available corpora in terms of its sheer size, a number of practical and methodological issues remain unresolved for the time being. Thus, fundamental concepts of corpus-linguistic analysis such as corpus representativeness and the presentation of relative frequency counts are difficult to apply when the whole of the worldwide web is considered as a single corpus. Furthermore, the available search engines are clearly not geared towards a linguistic analysis of the data and their (often non-replicable) search results are presented in a format and order that is difficult to interpret in a meaningful way.

However, it would nevertheless be rash to discount the worldwide web as a useful source of linguistic data. As Hoffmann (in press) shows on the basis of the CNN transcripts (cf. <http://www.cnn.com/transcripts>), one of the available options is to extract large and well-defined corpora from the Internet. This process involves the automated download of a selected range of pages to a local computer, to be followed by a — potentially elaborate — post-processing stage. This second stage ensures that the data are converted to a consistent format which can be searched with standard corpus tools as well as Perl scripts. While the creation of such web-derived corpora is much more time-consuming than a simple web search, the resulting database can be reliably employed to complement existing corpora.

For the purpose of the present investigation, we decided to make use of the on-line archive of the Calcutta-based national newspaper *The Statesman*, which offers access to news items dating from January 2002 onwards (cf. <http://www.thestatesman.org>). Although the complete archive can be explored on-line, the search facility suffers from the usual restrictions of data retrieval for non-linguistic purposes: The query language is relatively limited and the results are presented in a less-than-ideal format. In addition, no indication about the size of the database is given and relative frequency counts are thus impossible to calculate.

In order to make full use of the data, it was therefore necessary to create a local copy of the archive. For this purpose, all of the relevant pages containing

news items first had to be downloaded from the Internet. A simple Perl script in conjunction with the module LWP[8] offers a convenient solution for automating this type of task (cf. Hoffmann, in press). In a second step, the actual news items had to be separated from the other elements found on the page (e.g. links to other pages, advertisements, etc.). In the case of *The Statesman*, this proved to be a relatively simple undertaking because the format of the web-pages was highly consistent across the whole archive.[9] During this conversion process, the available metatextual data (date of publication, title, category of news — e.g. "home", "editorial", etc.) were also retrieved and stored separately in a relational database. Easy access to this type of data allows the researcher to explore flexible time-spans or subsections of the converted newspaper corpus. Finally, the automatic tagger EngCG was used to tag the complete *Statesman* archive with part-of speech information (cf. Voutilainen 1997). This last step greatly enhances the range of possible retrieval strategies by extending potential search algorithms to cover abstract grammatical patterns.

In its final format, the corpus consists of approximately 31 million words. Although it is of course very different in nature from ICE-India and its balanced set of text types, this homogeneous collection of news items significantly expands the available corpus data for Standard IndE. As we will demonstrate below, it is thus much more suitable than ICE-India for the description of low-frequency phenomena, as for example "new" ditransitives.

3.2 Using *The Statesman Database* for the identification of "new" ditransitives in Indian English

A syntactically parsed corpus like ICE-GB is of course ideally suited to an analysis of grammatical patterns such as ditransitive verb complementation. In the absence of this type of information, sequences of part-of-speech tags can be defined to retrieve potentially relevant constructions. In the case of ditransitive verbs with the type-I pattern, such a sequence would need to match all verbs

8. LWP is short for "Library for World Wide Web in Perl". See Burke (2002) for a comprehensive introduction to LWP and its flexible set of features.

9. The reason for this formal consistency is that individual news items are stored in a database on the server and are only retrieved whenever a web-user clicks on the relevant link. The full HTML-code for the final web-page to be displayed is then created dynamically on the basis of a boilerplate code. As a result, regardless of the original date of publication of the news item, all of the downloaded web-pages share the same structural properties and the distinction between news text and boilerplate is consistent across the whole archive.

Table 3. Verbs with type-I pattern complementation in *The Statesman Archive* (verbs listed as ditransitive verbs in Nihalani *et al*. 2004 and/or Olavarría de Ersson and Shaw 2003 in boldface)

Verb	Number of occurrences in *The Statesman Archive*
advise	10
brief	1
confer	3
convey	**0**
despatch / dispatch	1
enquire / inquire	**0**
entrust	**0**
explain	2
father	1
furnish	**0**
gift	26
impart	8
inform	**4**
intimate	1
notify	1
present	**18**
print	1
provide	**217**
put	**2**
remind	4
rob	4
submit	1
supply	**15**
threaten	1

that are immediately followed by two noun phrases. Since noun phrases can be extremely complex, it is not feasible to make a complete list of all potential realizations. While the recall of a tag-based retrieval strategy is thus unlikely to reach 100%, it can nevertheless be safely assumed that a simple noun phrase definition will be sufficient to match the large majority of cases. For our purposes, we defined a noun phrase as any noun with an optional determiner and one or several optional premodifying adjectives (which, in turn, could be premodified by adverbs). In addition, we also accounted for situations where the noun is post-modified by a prepositional phrase with the preposition *of*.

We then wrote a Perl script that successively went through the complete *Statesman Archive* and retrieved any verb that was either immediately followed

by two noun phrases or by an oblique pronoun and a noun phrase.[10] This resulted in a set of several thousand potential type-I constructions which had to be manually scanned for relevant instances.

After discarding unwanted sentences, several hundred type-I constructions with a whole range of verbs remained for closer analysis. In order to restrict our set to those instances which may indeed be typical of IndE usage, we turned to the British National Corpus (BNC World Edition; cf. Burnard 2000) as a suitable reference corpus. We discarded all verbs in *The Statesman Archive* which were also attested with a type-I pattern in the approximately 87.6 million words of the written component of the BNC.[11] Table 3 lists the remaining set of items and their number of occurrences in *The Statesman Archive*. All of the verbs discussed in Nihalani *et al.* (2004) and Olavarría de Ersson and Shaw (2003) are also included and are marked in boldface.

Not surprisingly, the most frequent verb in Table 3 is *provide* (217 instances). In addition, both *present* and *supply* were retrieved in considerable numbers (18 and 15 instances, respectively). Typical instances of the type-I pattern use of these three verbs in IndE are displayed in (17–19):

(17) With an open-air cafeteria, the park would have **provided** the city the much-needed greenery <The Statesman 2005–03–05>
(18) As a token of appreciation we **presented** each donor a travel bag and a certificate <The Statesman 2004–11–12>
(19) The matter of quota reduction aside, the FCI is not **supplying** us the foodgrain for October and November <The Statesman 2004–12–16>

For *provide*, *present* and *supply*, our data thus offer convincing empirical support for the intuition-based claims presented in Nihalani *et al.* (2004).

For the other verbs displayed in Table 2 above, however, quantitative evidence is clearly much more sparse. With only four instances, *inform* is found in equal numbers in both *The Statesman Archive* and ICE-India — albeit in very different corpus sizes. Furthermore, the verb *put* is only attested with two instances. Sentences (20) and (21) exemplify the rare pattern-I uses of these two verbs.

10. In order to reduce the number of retrieved sentences, verbs commonly known to occur with the type-I pattern were explicitly excluded from this search.

11. The only exception is *provide*, for which two instances of a type-I pattern are attested in the written component of the BNC. We decided against discarding *provide* in order to enable a full comparison of our data with all of the type-I verbs discussed in Nihalani *et al.* (2004) and Olavarría de Ersson and Shaw (2003).

(20) The employee is also required to **inform** the appointing authority the amount of monthly instalment… <The Statesman 2004–08–1>
(21) I **put** him a question as to whether he had an auspicious time for… <The Statesman 2003–08–09>

For the remaining verbs discussed in Nihalani *et al.* (2004) and Olavarría de Ersson and Shaw (2003), i.e. *convey*, *enquire* / *inquire*, *entrust* and *furnish*, *The Statesman Archive* does not supply any empirical evidence at all. However, given the low frequency of most items displayed in Table 3, the relevance of this finding should perhaps not be overrated.[12]

Interestingly, Table 3 also contains a considerable number of verbs whose potential to occur in the type-I pattern has not been discussed in previous studies of verb complementation in IndE. This seems especially true of *gift*, which was found 26 times in the type-I pattern. Considering this relatively large number of instances as well as the range of possible contexts of use displayed in *The Statesman Archive* — e.g. sports, politics, literature — it seems reasonable to claim that this verb has undergone the process of "grammatical institutionalisation" and is fully established as a potential type-I pattern verb in IndE.[13] Three typical instances are shown in (22–24):

(22) He was forced to bring down Nabi in the danger zone after **gifting** him the ball… <The Statesman 2003–12–12>
(23) Delay means serious risk of **gifting** Islamabad a talking point. <The Statesman 2002–10–26>
(24) She said she wanted to **gift** him a dream. <The Statesman 2003–02–17>

Although the verb *gift* is mentioned in the *Oxford English Dictionary* (*OED*), it is not listed as a ditransitive verb and none of the illustrative quotations displays a pattern similar to the one exemplified by sentences (22–24).[14] From a cognitive point of view, however, this use does not come as a surprise as it fully conforms to the basic meaning of the ditransitive construction ('X causes Y to receive Z').

12. For example, it could be argued that the non-occurrence of these type-I verbs in *The Statesman Archive* is simply a reflection of the relatively limited range of text types it contains.

13. There are also two instances of *gift* with a type-I pattern in ICE-India.

14. Furthermore, its use in the sense 'to bestow as a gift; to make a present of' is given as "chiefly Scottish" (*OED*, s.v. *gift*).

Apart from being an interesting finding *per se*, the discovery of *gift* as a type-I pattern verb in IndE is also testimony to the importance of working with data that are fully searchable via Perl scripts or standard corpus tools. The necessary reliance of web-based searches on purely lexical input is a serious limitation for any type of research into syntactic patterns.[15] Without having access to a tagged version of *The Statesman Archive*, we would probably never have considered looking at *gift* in the first place.

Among the other unexpected pattern-I verbs retrieved by our search algorithm, *advise* and *impart* — as shown in (25) and (26) — have the highest frequency with ten and eight instances, respectively. At least in our data, they are thus more common than some of the verbs discussed in Nihalani *et al*. (2004) and Olavarría de Ersson and Shaw (2003) and might therefore relatively safely be included among the list of potential type-I pattern verbs.

(25) I have **advised** him some technical changes like using both hands while stopping the ball . <The Statesman 2004–03–26>

(26) …teachers should study at least five times more than the students to be able to **impart** them the correct knowledge and wisdom. <The Statesman 2004–11–13>

For the remaining verbs in Table 3 (*brief*, *confer*, *despatch* / *dispatch*, *explain*, *father*, *intimate*, *notify*, *print*, *remind*, *rob*, *submit* and *threaten*), only between one and four type-I instances could be retrieved from the 31 million words of *The Statesman Archive*. Illustrative examples of some of these verbs are given in (27–30):

(27) People do strange and dangerous things when the cloak of anonymity **confers** them additional protection in a dark alley. <The Statesman 2004–02–29>

(28) When Nirbhay, who **fathered** her a son, was killed, Seema took charge of the gang. <The Statesman 2004–07–05>

(29) …and make unnecessary delay in **intimating** the student the outcome of review… <The Statesman 2004–04–30>

(30) …provided they formally **notify** us the date from which they will sign the new register. <The Statesman 2002–01–10>

Given this very low overall frequency, we would of course feel wary of making any sweeping claims about the level of integration of these type-I constructions

15. Most search engines also allow the use of wildcards in their searches. However, at least part of the query string must always be lexical.

into the grammatical system of IndE. Nonetheless, the relatively wide range of verbs that occur only sporadically in the type-I pattern in our data raises some important and interesting questions about the grammatical institutionalisation (or non-institutionalisation) of these verbs as ditransitive verbs in IndE, which we wish to briefly discuss in the following section.

4. Discussion

The data presented in the previous section make it clear that the verb-complementational profile of IndE differs from BrE — and other native varieties — not only with regard to the distribution of the patterns of a given verb, but also with regard to the range of verbs which are used in a given pattern, for example the type-I pattern. Using a large collection of web-derived data, we were able to extend the list of verbs that merit more careful attention. However, a number of questions remain and this pilot study thus calls for a more comprehensive investigation into the verb-complementational profile of IndE.

Most importantly, given the low overall frequencies of the type-I pattern verbs displayed in Table 3 above, future studies will have to determine which of these structures are indeed typical of IndE usage. As examples (14–16) demonstrated (e.g. *cry me a river*), speakers of all varieties of English — including those assigned to the Inner Circle — appear to exhibit a certain degree of flexibility in the area of verb complementation. As a result, at least some of the sentences presented in the previous section may in fact simply be a reflection of the inherent creativity of language users and should thus not be considered to form part of the grammatical system. These sentences seem to exemplify syntactic nonce-formations, as it were, and do not exemplify processes of grammatical institutionalisation.

Such a view is perhaps supported by the fact that our search in *The Statesman Archive* retrieved rare type-I pattern uses of a range of verbs that could also be attested in very low numbers in the written component of the BNC — this is the case for *avail*, *bequeath*, *deliver*, *extend*, *incur*, *issue*, *lay*, *penalize*, *prescribe* and *return*. Consider sentences (31) and (32), which illustrate this infrequent complementation pattern of the verb *return* in IndE and BrE:

(31) The accused said that the necklace did not fetch the price fixed and **returned** her a counterfeit necklace. <The Statesman 2004–06–25>

(32) I said that was fine, but could he **return** me the biography I had sent? <BNC; CAK 232>

Given the low overall frequency of the constructions under consideration, much larger — possibly also web-derived — corpora will be required to offer more conclusive evidence about the status of many of the verbs displayed in Table 3 above.

A second important observation to make is that the findings presented in our paper are fully derived from synchronic corpora. This consequently makes it difficult to evaluate our data within the larger context of language change. By using the label "new" ditransitives, we implicitly assert that the range of grammatically institutionalised type-I pattern verbs changes over time. However, without access to empirical support from diachronic corpora, such a claim must remain speculative for the time being. Diachronic evidence would also allow us to investigate the origin of the observed differences in the complementational profiles of IndE and BrE. Are they the result of post-colonial divergence or are we instead looking at the complementational profile of 19th century BrE which has been preserved by its Indian speakers in its original form? Finally, diachronic data may also permit us to ascertain more clearly whether rare type-I pattern occurrences in BrE are the transient product of creativity or whether the complementation patterns of some verbs found in outer-circle varieties are in fact entering the territory of native varieties by adding to the inventory of ditransitive verbs.

In this context, it is also interesting to see that the "new" ditransitives in IndE presented in Section 3 include verbs which had already been used in the type-I pattern in the history of native Englishes, such as *inform*, as shown in (33) and (34):

(33) I have often the satisfaction of hearing the publican, the baker, and sometimes even the parish-clerk, petitioning my housekeeper … to **inform** him the exact time by Master Humphrey's clock. (Dickens 1840: 10)

(34) I **informed** him the account I had got from John Hutson. (Rupp 1847: 349)

It seems reasonable to claim that there is a clearly rational impetus, e.g. on grounds of semantic analogy, which has always led users of English — both native and non-native — to extend the ditransitive situation schema to new verbs.[16] Some of these will become grammatically institutionalised and may as

16. More specifically, Mukherjee (2005b) argues that "new" ditransitives emerge on grounds of "nativised semantico-structural analogy", i.e. a process by means of which non-native speakers of English as a second language are licensed to introduce new forms and structures

a consequence develop into fully-fledged type-I pattern verbs. A better understanding of these processes might certainly help in an evaluation of the status of the "new" ditransitives in IndE. It should thus be more than interesting to explore in detail the diachronic development of the type-I pattern across varieties of English.

5. Concluding remarks

In the light of the results of the present pilot study of verb complementation by combining corpus data and web data, we would like to offer three concluding remarks, which point to some key issues for future research.

Firstly, we hope to have shown that verb complementation in general and ditransitive verb complementation in particular represent core areas in which different varieties of English are marked by diverging preferences and structural options. Future research into the endonormative stabilisation of New Englishes will thus have to delve much more deeply into the verb-complementational profiles of varieties of English. In this context, it will be necessary to extend the analysis of verb complementation to other than ditransitive verbs and to other than type-I patterns. It is only then that a complete picture of the entire verb-complementational profile of a given variety can be drawn.

Secondly, while the lack of parsed corpora for most varieties of English is certainly a limiting factor in research into verb complementation, tag-based pattern searches can offer a suitable alternative for detecting relevant constructions. However, since considerable manual work is required in order to discard unwanted instances, the availability of parsed corpora of other components of the International Corpus of English (ICE) would greatly facilitate exhaustive and large-scale searches for complementation patterns across all classes of verbs.

Thirdly, we would like to conclude from this pilot study that language data from the web are useful for the description of New Englishes such as IndE, for which corpora larger than ICE-India are not available. Once they have been downloaded from the Internet and converted into an appropriate format, large and homogeneous collections of authentic text accessible on the world-wide web, e.g. newspaper archives like the database of *The Statesman*, provide a goldmine of data when mining for examples of low-frequency phenomena.

into the English language because corresponding semantic and formal templates already exist in the English language system.

It is therefore desirable that larger portions of the Internet are made available to linguists in a suitable format. While carefully compiled, balanced corpora such as the BNC will always have their place, the worldwide web opens up new avenues of research that will — among other things — make it possible to offer an empirically grounded description of the distinctive features of Englishes world-wide on a much larger scale than previously envisaged.

References

Beermann, Dorothee. 2001. "Verb semantics and double object constructions". In Marc van Oostendorp and Elena Anagnostopoulou, eds. *Progress in Grammar: Articles at the 20th Anniversary of the Comparison of Grammatical Models Group in Tilburg*. <http://www.roquade.nl/meertens/progressingrammar/beermann.pdf> (8.8.2005).

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Burke, Sean M. 2002. *Perl & LWP*. Sebastopol, CA: O'Reilly.

Burnard, Lou. 2000. *Reference Guide for the British National Corpus (World Edition)*. Oxford: Oxford University Computing Services.

Campbell, Aimee L. and Michael Tomasello. 2001. "The acquisition of English dative constructions". *Applied Psycholinguistics* 22: 253–367.

Dickens, Charles. 1958 [1840]. *Master Humphrey's Clock and A Child's History of England*. London: Oxford University Press.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: The University of Chicago Press.

Green, Georgia. 1974. *Semantics and Syntactic Regularity*. Bloomington, IN: Indiana University Press.

Greenbaum, Sidney, ed. 1996. *Comparing English Worldwide: The International Corpus of English*. Oxford: Clarendon.

Grisham, John. 2002. *The Summons*. New York: Doubleday, Random House.

Hoffmann, Sebastian. In press. "From web-page to mega-corpus: The CNN transcripts". In Marianne Hundt, Nadja Nesselhauf and Carolin Biewer, eds. *Corpus Linguistics and the Web*. Amsterdam: Rodopi.

Hunston, Susan and Gill Francis. 2000. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam, Philadelphia: Benjamins.

Jackson, Howard. 1990. *Grammar and Meaning: A Semantic Approach to English Grammar*. London: Longman.

Kachru, Braj B. 1985: "Standards, codification and sociolinguistic realism: The English language in the outer circle". In Randolph Quirk and Henry Widdowson, eds. *English in the World: Teaching and Learning the Language and Literatures*. Cambridge: Cambridge University Press, 11–30.

Kilgarriff, Adam and Gregory Grefenstette. 2003. "Introduction to the special issue on the web as corpus". *Computational Linguistics* 29: 333–47.

Mair, Christian. 2002. "Three changing patterns of verb complementation in Late Modern English: A real-time study based on matching text corpora". *English Language and Linguistics* 6: 105–31.
Matthews, Peter H. 1981. *Syntax*. Cambridge: Cambridge University Press.
Mukherjee, Joybrato. 2005a. *English Ditransitive Verbs: Aspects of Theory, Description and a Usage-based Model*. Amsterdam: Rodopi.
———. 2005b. "'All mine, mine alone…': Emerging local norms in Indian English lexicogrammar". Paper presented at the University of Zurich.
Nelson, Gerald. 1996. "The design of the corpus". In Greenbaum, ed.: 27–35.
Newman, John. 1996. *Give: A Cognitive Linguistic Study*. Berlin: Mouton de Gruyter.
Nihalani, Paroo, Ray K. Tongue, Priya Hosali and Jonathan Crowther. 2004. *Indian and British English: A Handbook of Usage and Pronunciation*. 2nd ed. Delhi: Oxford University Press.
Ninio, Anat. 1999. "Pathbreaking verbs in syntactic development and the question of prototypical transitivity". *Journal of Child Language* 26: 619–54.
Olavarría de Ersson, Eugenia and Philip Shaw. 2003. "Verb complementation patterns in Indian Standard English". *English World-Wide* 24: 137–61.
The Oxford English Dictionary (OED), CD-ROM edition (1992). Oxford: Clarendon Press.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
Rupp, Israel D. 1847. *History and Topography of Northumberland, Huntingdon, Mifflin, Centre, Union, Columbia, Juniata, and Clinton Counties, Pa*. Lancaster, PA: G. Hills.
Sand, Andrea. 2004. "Shared morpho-syntactic features in contact varieties of English: Article use". *World Englishes* 23: 281–98.
Schneider, Edgar W. 2001. "Rev. D. Biber, S. Johansson, G. Leech, S. Conrad and E. Finegan, *Longman Grammar of Spoken and Written English* (Harlow: Pearson Education, 1999)". *English World-Wide* 22: 137–43.
———. 2003. "The dynamics of New Englishes: From identity construction to dialect birth". *Language* 79: 233–81.
———. 2004. "How to trace structural nativization: Particle verbs in World Englishes". *World Englishes* 23: 227–49.
Skandera, Paul. 2003. *Drawing a Map of Africa: Idiom in Kenyan English*. Tübingen: Narr.
Voutilainen, Atro. 1997. "The EngCG–2 Tagger in Outline". <http://www.ling.helsinki.fi/~avoutila/cg/doc/engcg2-outline/engcg2-outline.html> (8.8.2005).

Authors' addresses

Joybrato Mukherjee
University of Giessen
Department of English
Otto-Behaghel-Str. 10B
35394 Giessen
Germany
e-mail: Joybrato.Mukherjee@anglistik.uni-giessen.de

Sebastian Hoffmann
University of Zurich
Department of English
Plattenstr. 47
8032 Zurich
Switzerland
e-mail: sebhoff@es.unizh.ch

Appendix 1

Patterns of the ditransitive verbs *give* and *send* — an overview.
[NB: The pattern formulas are based on the following notational conventions (cf. Mukherjee 2005a: 93–101): "[…]" = obligatory element; "[…(…)]" = obligatory element with a specific form/function; "(…)" = optional element; "~~O~~$_i$"/"~~O~~$_d$" = clause element which is not part of the lexicogrammatical pattern at the level of syntactic surface structure.]

Type-I patterns

I (S) GIVE [O_i:NP] [O_d:NP]
Give me a warning next time <ICE-GB:S1A–091 #313>

I a (S) GIVE [O_d: NP] [O_i:NP]
"He's my dog. You gave him me." <ICE-GB:W2F–001 #107–108>

I b [O_d: $NP_{antecedent}$] (rel. pron.) [S] GIVE [O_i:NP]
Those batteries that you gave me lasted an hour <ICE-GB:S1A–085 #132>

I c [O_i:$NP_{antecedent}$] (rel. pron.) [S] GIVE [O_d:NP]
Uh an American lady that I gave a lecture uhm on architecture […] <ICE-GB:S2A–024 #34>

I d [O_d: $NP_{fronted}$] [S] GIVE [O_i:NP]
Somebody she gave me <ICE-GB:S1A–043 #69>

IP [S < O_i active] BE *given* [O_d:NP] (by-agent)
One was that I was being given the opportunity to uhm learn and develop uhm physical skills <ICE-GB:S1A–001 #34>

IP b [O_d:$NP_{antecedent}$] (rel. pron.) [S < O_i] BE *given* (by-agent)
And uh this the letter that the UN Secretary General has been given uh by John McCarthy <ICE-GB:S2A–008 #166>

Type-II patterns

II (S) GIVE [O_d:NP] [O_i:PP_{to}]
Well perhaps you're giving too much to other people <ICE-GB:S1A–067 #288>

II a (S) GIVE [O_d:NP] [O_i:PP_{for}]
The defendant had been ordered to attend at Leeds Crown Court to give evidence for the prosecution in a fraud trial <ICE-GB:W2B–020 #68>

II b [O_d:$NP_{antecedent}$] (rel. pron.) [S] GIVE [O_i:PP_{to}]
One of the great things about by-elections […] is of course this extra bit of power that they give to a voter <ICE-GB:S1B–029 #60>

II c (S) GIVE [O_i:PP_{to}] [O_d:NP]
It was the Queen of course who gave to Norman Schwarzkopf the knighthood that makes him now Sir Norman <ICE-GB:S2A–019 #9>

IIP [S < O_d active] BE *given* [O_i:NP_{to}] (by-agent)
more time should have been given to sanctions <ICE-GB:S2B–018 #93>

IIP b [antecedent]$_{co}$ (S < O_d)$_{co}$ (BE) *given* [O_i:NP_{to}] (by-agent)
I'll leave you with the final message that was given to the world leaders at the summit <ICE-GB:S2B–022 #148>

Type-III patterns

III (S) GIVE [O_d:NP] ~~O_i~~
he wanted physical love and I couldn't give that <ICE-GB:S1A–050 #184>

III b [O_d:$NP_{antecedent}$] (rel. pron.) [S] GIVE ~~O_i~~
The sermon he gave when his daughter was married <ICE-GB:S1A–053 #186>

IIIP [S < O_d active] BE *given* ~~O_i~~ (by-agent)
More specific implementation details are given at the end of the report <ICE-GB:W1A–005 #5>

IIIP b [antecedent]$_{co}$ (S < O_d)$_{co}$ (BE) *given* ~~O_i~~ (by-agent)
it also is of relevance when considering the evidence given by Mr Holt because there is a clear conflict <ICE-GB:S2A–068 #40>

Type-IV patterns

IV (S) GIVE ~~O_i~~ ~~O_d~~
If you give and take when there's that close bodily contact it's great <ICE-GB:S1A–003 #146>

IVP *given* ~~O_i~~ ~~O_d~~ (by-agent)
The run rate here as given by our scorer immediately and it's two point four one <ICE-India:S2A–016 #12>

Type-V pattern

V (S) GIVE [O_i:NP] ~~O_d~~
I didn't give Lakshmi I had just given Sumi you know <ICE-India:S1A–098 #69–70>

Appendix 2

Complementation of *give* in ICE-India and ICE-GB (most frequent patterns in boldface)

		ICE-India		ICE-GB	
type	**pattern**	**sum**	**%**	**sum**	**%**
I	(S) GIVE [O_i:NP] [O_d:NP]	**407**	**22.6**	**404**	**38.0**
Ia	(S) GIVE [O_d: NP] [O_i:NP]	3	0.2	1	0.1
Ib	[O_d: $NP_{antecedent}$] (rel. pron.) [S] GIVE [O_i:NP]	14	0.8	23	2.2
Ic	[O_i:$NP_{antecedent}$] (rel. pron.) [S] GIVE [O_d: NP]	1	0.1	2	0.2
Id	[O_d: $NP_{fronted}$] [S] GIVE [O_i:NP]	3	0.2	1	0.1
	miscellaneous	11	0.6	10	0.9
IP	[S < O_i active] BE *given* [O_d:NP] (by-agent)	**130**	**7.2**	**84**	**7.9**
IPb	[O_d:$NP_{antecedent}$] (rel. pron.) [S<O_i] BE *given* (by-agent)	3	0.2	12	1.1
	miscellaneous	4	0.2	0	0.0
II	(S) GIVE [O_d:NP] [O_i:PP_{to}]	**310**	**17.3**	**123**	**11.6**
IIa	(S) GIVE [O_d:NP] [O_i:PP_{for}]	5	0.3	4	0.4
IIb	[O_d:$NP_{antecedent}$] (rel. pron.) [S] GIVE [O_i:PP_{to}]	12	0.7	7	0.7
IIc	(S) GIVE [O_i:PP_{to}] [O_d:NP]	4	0.2	2	0.2
	miscellaneous	6	0.3	6	0.6
IIP	[S < O_d active] BE *given* [O_i:PP_{to}] (by-agent)	**70**	**3.9**	**23**	**2.2**
IIPb	[antecedent]$_{co}$ (S<O_d)$_{co}$ (BE) *given* [O_i:NP_{to}] (by-agent)	19	1.1	17	1.6
	miscellaneous	2	0.1	2	0.2
III	(S) GIVE [O_d:NP] ~~O_i~~	**528**	**29.4**	**247**	**23.2**
IIIb	[O_d:$NP_{antecedent}$] (rel. pron.) [S] GIVE ~~O_i~~	17	0.9	16	1.5
	miscellaneous	25	1.4	3	0.3
IIIP	[S < O_d active] BE *given* ~~O_i~~ (by-agent)	**123**	**6.8**	**38**	**3.6**
IIIPb	[antecedent]$_{co}$ (S < O_d)$_{co}$ (BE) *given* ~~O_i~~ (by-agent)	**49**	**2.7**	**28**	**2.6**
	miscellaneous	10	0.6	0	0.0
IV	(S) GIVE ~~O_i~~ ~~O_d~~	24	1.3	10	0.9
	miscellaneous	0	0.0	1	0.1
IVP	*given* ~~O_i~~ ~~O_d~~ (by-agent)	5	0.3	0	0.0
V	(S) GIVE [O_i:NP] ~~O_d~~	6	0.3	0	0.0
	miscellaneous	6	0.3	0	0.0
sum total		**1797**	**100**	**1064**	**100**
p		< 0.001			

Appendix 3

Complementation of *send* in ICE-India and ICE-GB (most frequent patterns in boldface)

		ICE-India		ICE-GB	
type	**pattern**	**sum**	**%**	**sum**	**%**
I	(S) SEND [O_i:NP] [O_d:NP]	**54**	**15.4**	**62**	**21.8**
Ib	[O_d: $NP_{antecedent}$] (rel. pron.) [S] SEND [O_i:NP]	4	1.1	6	2.1
	miscellaneous	1	0.3	1	0.4
IP	[S < O_i active] BE *sent* [O_d:NP] (by-agent)	0	0.0	7	2.5
IPb	[O_d:$NP_{antecedent}$] (rel. pron.) [S < O_i] BE *sent* (by-agent)	0	0.0	1	0.4
II	(S) SEND [O_d:NP] [O_i:PP_{to}]	**78**	**22.3**	**64**	**22.5**
IIb	[O_d:$NP_{antecedent}$] [S] (rel. pron.) SEND [O_i:PP_{to}]	0	0.0	3	1.1
IIc	(S) SEND [O_i:PP_{to}] [O_d:NP]	0	0.0	2	0.7
	miscellaneous	5	1.4	5	1.8
IIP	[S < O_d active] BE *sent* [O_i:PP_{to}] (by-agent)	**29**	**8.3**	**27**	**9.5**
IIPb	[antecedent]$_{co}$ (S < O_d)$_{co}$ (BE) *sent* [O_i:NP_{to}] (by-agent)	9	2.6	4	1.4
	miscellaneous	2	0.6	1	0.4
III	(S) SEND [O_d:NP] O_i	**130**	**37.1**	**69**	**24.2**
IIIb	[O_d:$NP_{antecedent}$] (rel. pron.) [S] SEND O_i	6	1.7	4	1.4
	miscellaneous	3	0.9	2	0.7
IIIP	[S < O_d active] BE *sent* O_i (by-agent)	**11**	**3.1**	**20**	**7.0**
IIIPb	[antecedent]$_{co}$ (S < O_d)$_{co}$ (BE) *sent* O_i (by-agent)	10	2.9	3	1.1
	miscellaneous	1	0.3	0	0.0
IV	(S) SEND O_i O_d	7	2.0	2	0.7
V	(S) SEND [O_i:NP] O_d	0	0.0	1	0.4
	miscellaneous	0	0.0	1	0.4
sum total		**350**	**100**	**285**	**100**
p		< 0.001			