

Be surprised by in the Corpus and the *OED*

Susumu Taketazu

1. Introduction

The present paper is an attempt to explore the historical shift of the prepositions of agent introduced by the *be surprised* constructions¹ used in a psychological sense.² Contrary to the explanations found in many traditional grammars, which have stated that the constructions generally take *at*, not *by*, as the agentive preposition, the *be surprised by* construction as shown in (1) occurs not sporadically but quite frequently in Present-day English (PE). In fact, it is so prolific and occurs with such an unexpectedly high frequency in books, magazines, and especially in newspapers that the real linguistic conditions seem to run counter to conventional descriptions.

- (1) a. Back in Miami, both Ainslie and Quinn were surprised by the absence of phoned-in tips following the extensive media coverage of the murders. (*Arthur Hailey, Detective*, 1997, 90)
- b. It seems hard to imagine that Hillary was truly surprised by her husband's discretion. (*Newsweek*, Aug. 31, 1998, 26c)
- c. He shut down one of the most dangerous-hitting teams in baseball, and he didn't have his breaking ball. I'm past being surprised by Hideki. (*The Hartford Courant*, May 16, 1998)

Because of the recent development of computerized corpus, a situation has emerged where an enormous amount of data, which ordinarily would take a lifetime to collect, can be obtained in an instant and with flawless accuracy. This favorable situation has induced me to undertake research to clarify the true linguistic behaviors of the *be surprised by* construction used psychologically. Thus I have utilized four computer corpora of Modern English and the *OED 2* on CD-ROM and conducted statistical investigations.

2. Previous studies

2.1 Grammars and usage books

Grammarians have discussed the *be surprised at/by* constructions in relation to passives. Svartvik (1966: 102), in his pioneering corpus-based work on voice, uses the term “quasi-agents” for the agents of “emotive passives.” Svartvik claims that the prepositions introduced by quasi-agents are not “voice-conditioned, but are selected by collocation with particular verbs: *be worried about, surprised at, interested in, annoyed with*, etc.”

Some grammarians (e.g. Close (1975), Leech and Svartvik (1975), Konishi (1980), Swan (1995)) treat the *be surprised* constructions as adjectival when they have a stative meaning and occur with *at*, and regard the constructions as passives if they are dynamic in meaning and take *by*. This is exemplified by the following sentences from Close:

- (2) a. I am very surprised at you.
- b. I was surprised by a knock at the door. (Close §11.3)

Sentences like (2b) and (3) below (presented by Swan), however, do not seem to be appropriate examples.

- (3) The burglar was surprised by the family coming home unexpectedly.
(Swan §405)

First, they have a slight physical connotation, which naturally requires *by*. Second, sentences like (2b) and (3) are observed to have occurred in the early corpora (i.e. the corpora of late Modern English), but seldom appear in the data of current corpora (i.e. in PE).³

Quirk *et al.* (1985: §3.76) refer to sentences such as *I was a bit surprised at her behaviour* as “semi-passives,” which “have both verbal and adjectival properties” and say that “it is rare to have a *by*-phrase expressing the agent.” In another section (§ 9.63), however, Quirk *et al.* present two sentences that are exactly the same except for the prepositions in (4).

- (4) a. I was surprised by their rejection of the offer.
b. I was surprised at their rejection of the offer.
(Quirk *et al.* §9.63)

They call the former a passive and the latter a participial adjective and state that the latter has adjectival characteristics (e.g. it can be preceded by *very*) and “it is with such adjectival forms that *at* [‘stimulus’] characteristically combines.”

Palmer (1988: § 5.2.2) also uses the term “semi-passives” for “some *-en* forms that appear to have corresponding actives, yet exhibit adjectival features” and “relate to some kind of emotional condition.”

As Palmer states that “many of such *-en* forms can occur not only with *by* but also with other prepositions,” it follows that the *be surprised* construction takes *by* as well as *at* as its agentive preposition.

Declerck (1991: §6.1) uses the term “true passive” to refer to the construction taking *by* and the term “pseudo-passive” for one taking other than *by*.

- (5) a. She seemed surprised at your behavior.
- b. I was surprised by what she said.
- c. They were surprised by the guard suddenly entering the room.

(Declerck §6.1)

Declerck’s pseudo-passives are those which other grammarians regard as adjectival constructions. (5b) is a type of sentence which is often seen in the current corpora, but (5c), similar to (3), may allow of a physical interpretation and hardly appear in PE.

Considering what has been discussed, there seems to be a great deal of ambivalence among grammarians in the treatment of the construction.

2.2 Dictionaries

Dictionaries also vary in their treatments. Some dictionaries, especially older ones or the ones published in Britain or America, hardly make any mention of *by*. Even the BNC-based *NODE*, does not provide any information on *by* nor present any quotation containing *by*. Some others only provide quotations with a physical meaning, which I find to have been decreasing historically.

- (6) a. They were surprised by a unit of US marines during the night.
(*COBUILD*)
- b. The robbers had just opened the safe when they were surprised
by the police. (*CIDE*)

There are dictionaries which do not seem to regard *by* as a legitimate and well-established preposition. *KDEC*'s treatment, for instance, seems to betray the editors' unsureness about the legitimacy of the usage. Under the 'verb + preposition' collocation, four examples of *be surprised at* are cited but not a single quotation of *be surprised by* is given. However, under the collocation 'verb + adverb', two quotations of *be surprised by* as shown below are cited, which could have been used for the 'verb + preposition' collocation.

- (7) a. She appeared deeply [profoundly] surprised by the revelation.
- b. I was faintly [slightly, a little] surprised by his absence.

LDCE is one of a few 'native' dictionaries which provide a reasonable treatment of the construction. *LDCE*'s treatment, however, can be confusing, depending on the edition. The 1991 edition indicates the collocation [*at*] but provides a quotation *I was surprised at/by her reaction*. In the 1995 edition it shows a more appropriate collocation [*at/by*] and gives a quotation *We were all surprised at Sue's outburst*.

Dictionaries which provide an appropriate description are very few and exceptional. *Global*, for instance, indicates a proper collocation [*at/by*] and gives an 'authentic' quotation (*She could no longer be surprised by anything he might say or do*) unlike some other dictionaries, many of which seem to adopt 'second-hand' quotations such as *I*

was surprised by a knock at the door, rarely seen in the data of the current corpora.

WEDU gives a slightly more detailed explanation: “Both *at* and *by* are possible when *surprised* means ‘struck with wonder’ or ‘taken aback.’” It also provides an authentic quotation.

- (8) . . . have been myself continually surprised . . . by the abrupt and vast changes that I have seen. (S. P. B. Mais, *The English Scene To-day*, 2nd ed., 1949)

College Lighthouse is the one which probably gives the best treatment. It provides good quotations, together with the most detailed explanation, the first of this kind, about the usage:

By is a colorless preposition used to mention a fact, whereas when *at* is used, a strong shock or an emotional response or reaction can be felt.

In the face of the profusion of examples of *be surprised by* used in a psychological sense in PE, conventional descriptions given in many grammars, dictionaries and usage books seem to be increasingly becoming inappropriate and insufficient. (Of course allowances have to be made for the limited space in dictionaries, and the inability of grammarians and lexicographers to have gained access to the enormous amount of data which have come to be available to us through computer corpora in recent years.) With computer corpora easily available, an attempt has been made to explore the *be surprised by* construction from an historical perspective.

3. The computer corpora of Modern English

Using four computer corpora of Modern English,⁴ I have examined how the constructions have been used in late Modern English. The four corpora can be grouped into the ‘early’ corpora and the ‘current’ corpora; ‘early’ in the sense that the corpora deal with relatively earlier English, mostly from the 19th to the early 20th century (which I refer to as late Modern English here), compared to the other corpora which contain the texts of current English (i.e. PE). The early corpora include translations from foreign literature, which were excluded from this research.⁵

For the sake of simplicity I will call one of the early corpora the ‘Virginia Corpus’ and the other the ‘Michigan Corpus’ as the two were created at the University of Virginia and the University of Michigan, respectively. The British National Corpus is regularly called the ‘BNC’ and I will call the CobuildDirect Corpus the ‘Cobuild Corpus’. For more detail about the corpora, see Tajima (1996: 48) for the early corpora and Saito *et al.* (1998: 10) for the current corpora.

- (i) The early corpora
 - The Modern English Collection (the Virginia Corpus)
 - Public Domain Modern English Search (the Michigan Corpus)
- (ii) The current corpora
 - The British National Corpus (the BNC)
 - The CobuildDirect Corpus (the Cobuild Corpus)

With these corpora, such collocations as *surprised at* (*by*, *with*, etc.) were searched. After the search, I screened the samples one by

one and excluded inappropriate ones. Non-psychological sentences (e.g. *surprised by a guerrilla force*) and the mechanically-searched ill-collocated ones (e.g. *I was much surprised at first*) were discarded and so were the cases where *surprised by* is used as the title of a book (*Surprised by Joy* by C. S. Lewis). The exact repetition of one person's utterance containing *surprised by* by another in a dialogue was also excluded.

Syntactically, not only full-fledged clauses but also non-finite clauses such as participle constructions (e.g. *Much surprised by the request, I took the note*) were taken into account. The constructions in which such linking verbs as *seem*, *look*, *sound* or *feel* function as main verbs (e.g. *I felt surprised at myself*) were also considered.

What troubled me most were those samples that entail not only mental but physical or physiological influence (e.g. *No potential thief likes to be surprised by a 500 watt halogen floodlight*), especially in the early corpora. To err on the cautious side, these examples are excluded. Although the exact figures will be given in the succeeding sections, these should be considered to be approximate because of the exclusion of such dubious examples just mentioned above, which may allow of different interpretations. Nevertheless, these exclusions, as they are very marginal, do not make a significant difference in the general ratio between *at* and *by*.

3.1 The early corpora

3.1.1 The Modern English Collection (the Virginia Corpus)

This is a corpus consisting of texts mostly from the 19th to the early 20th century. Some of the important authors are restricted from non-University of Virginia users (e.g. Jane Austen, Willa Cather (in

part), Daniel Defoe, Ralph Waldo Emerson, William Faulkner), but this shortcoming would not hinder one from obtaining data that may well represent the conditions of late Modern English. The corpus has been constantly updated and the version I used is that of July 15, 1999. The search yielded the following raw data: *at*: 357, *by*: 111, *with*: 17 and *about*: 1. After screening and discarding inappropriate samples, the results in Table 1 were obtained. The results show that more than one-fifth (21.1 percent) are *by*-phrases, while a little over three quarters (77.6 percent) occur with *at*.

Table 1 The Virginia Corpus

<i>at</i>	343 (77.6%)
<i>by</i>	93 (21.1%)
<i>with</i>	5 (1.1%)
<i>about</i>	1 (0.2%)
total	442 (100%)

3.1.2 Public Domain Modern English Search (the Michigan Corpus)

This corpus is similar to the Virginia Corpus in that many of the texts are mutual. The size of the Virginia Corpus is much larger but there are texts which are not contained in the Virginia Corpus. The results below were obtained after following the same procedures and screening the raw data: *at*: 141, *by*: 41, *with*: 4.

Table 2 The Michigan Corpus

<i>at</i>	137 (77.0%)
<i>by</i>	38 (21.3%)
<i>with</i>	3 (1.7%)
total	178 (100%)

Although the amount of data obtained is less than one-half of that of the Virginia Corpus, the occurrence rates of prepositions show an extreme similarity to those of the Virginia Corpus. This is presumably because a considerable number of texts overlap those of the Virginia Corpus and naturally many of the quotations are mutual. In this sense, the Michigan Corpus may be taken to be supplementary to the Virginia Corpus.

The search into the two early corpora also discloses that not a small number of great and popular writers of the late Modern English period used the *surprised by* construction, as well as the *surprised at* construction. These writers include Joseph Conrad, Charles Dickens, George Eliot, Nathaniel Hawthorne, Samuel Johnson, Sir Walter Scott, Anthony Trollope, H. G. Wells and Virginia Woolf. It was Dr. Johnson, one of the most influential men in the history of Modern English, who used the *surprised by* construction for the first time in English literature. Dickens, another great writer of late Modern English, also seems to have favored this construction.⁶ It is surprising that a host of famous and popular writers of late Modern English already used this construction, contrary to a prejudiced belief based on the explanations in traditional grammars.

It does not mean that every writer of the period used the *surprised*

by construction. Many other writers adhered only to the conventional *surprised at* construction and never used the newcomer. They are Lewis Carroll, Sir Arthur Conan Doyle, Henry James, Charles Darwin and Andrew Lang, to name only a few. The leveling of the statistics from the writers who patronized *surprised by* and those who did not leads to an occurrence rate of a little over 20 percent.

3.2 The current corpora

3.2.1 The British National Corpus (the BNC)

The BNC is a collection of over 4,000 texts of modern British English from 1975 onwards, both spoken and written, and it totals over 100 million words (Aston and Burnard 1998: Preface). The results in Table 3 were obtained after searching the same phrases and screening the raw data: *at*: 422, *by*: 422, *about*: 25, *with*: 5, *of*: 1. The same number of occurrences of *at* and *by* (i.e. 422) seem to be purely accidental.

Table 3 The BNC

<i>at</i>	412 (48.8%)
<i>by</i>	403 (47.8%)
<i>about</i>	24 (2.8%)
<i>with</i>	4 (0.5%)
<i>of</i>	1 (0.1%)
total	844 (100%)

The outcome of this search is very different from that of the early corpora. The occurrence rate of *surprised by* has increased a great deal

and shows almost an equal rate to that of *surprised at*, which has undergone a considerable decline by as much as 30 percent.

There may be two reasons for the difference between this corpus and the early corpora. First, as compared to the first two corpora, whose texts are mostly taken from books, the texts of this corpus are composed of a wide range of materials from written to spoken English and from books to newspapers. Another and more important reason is that a majority of the texts of the early corpora are from the books of the 19th and the early 20th centuries, in contrast to the more current texts of the BNC. Diachronic factors seem to have come into play in creating this gap.

3.2.2 The CobuildDirect Corpus (the Cobuild Corpus)

This is a 50-million-word sub-corpus of the 300-million-word Cobuild Word Bank (which is ever-increasing). Most of the texts are from 1983 onwards and genres include: British Books, British Magazines, British Spoken Data, British Ephemera, American Books, American Ephemera, Australian Newspapers, BBC World Service, National Public Radio, The Times, Today (British tabloid). Such a wide variety of texts, because of its freedom from bias for a particular type of English, suggests that the data well represent the actual conditions of current English. A search of the same phrases yielded the following raw data: *by*: 214, *at*: 188, *about*: 10, *with*: 8, *of*: 1. After screening, the results in Table 4 were obtained.

The corpus that I used is quantitatively half as large as the 100-million-word BNC, and the amount of data gained from the search is approximately half as large as that of the BNC (BNC 844 vs. Cobuild 403), which fits my expectations. The ratio of *at* and *by* is fairly

Table 4 The Cobuild Corpus

<i>by</i>	206 (51.4%)
<i>at</i>	176 (43.9%)
<i>about</i>	10 (2.5%)
<i>with</i>	8 (2.0%)
<i>of</i>	1 (0.2%)
total	401 (100%)

similar to that of the BNC, except for the slight increase of *by* and the decrease of *at* by about 5 percent, with the result that the rate of *by* has surpassed that of *at*.

3.3 The results

The results of the search into the four corpora are all presented together in Table 5. Only the statistics for *at* and *by* are shown, with minor prepositions omitted. The figures are rounded to the nearest whole number for the sake of clarity.

Table 5 The results of the four corpora

	early		current	
	Virginia	Michigan	BNC	Cobuild
<i>at</i>	78%	77%	49%	44%
<i>by</i>	21%	21%	48%	51%

The occurrence rate of *surprised by* was a little over 20 percent in late Modern English, while in PE the rate has risen to as high as 50

percent or close to that figure. The ratio of occurrence between *at* and *by* was 3.7 : 1 in late Modern English, whereas in PE the ratio comes to approximately 1 : 1 (BNC) or 0.9 : 1 (Cobuild). The disparity between the BNC and the Cobuild may be ascribed to the fact that the texts of the BNC contain a slightly earlier texts than those of the Cobuild. There is approximately a ten-year difference between the two corpora, which could produce some difference of occurrence rate, considering the apparent supremacy of *surprised by* in recent years. The disparity may also be accounted for by the difference of genres. The Cobuild contains a rather large amount of texts written in journalism English, which may contribute to the preference of *by* to *at* (See Taketzu 1999: 12-13).

4. The *OED 2* on CD-ROM as a corpus

It has been shown that the *be surprised by* construction has gained such currency during the Modern English period that the occurrence rate of *by* has shifted from a little over 20 percent in late Modern English to 50 percent in current English. Now we will further examine the historical shift of agentive prepositions by using the *OED 2* on CD-ROM as a corpus.

4.1 The definition of *surprise*

Let us first examine sense 5a of *surprise* used psychologically in the *OED*, which is as follows:

SURPRISE 5. a. To affect with the characteristic emotion caused by something unexpected; to excite to wonder by being unlooked

-for. *Formerly also in stronger sense (cf. SURPRISE *n.* 4 a), to astonish or alarm; also, to excite to admiration. Often *pass.*, const. *at* (**with*) or *inf.*; colloq. *to be surprised at*=to be scandalized or shocked at; also as a retort: *you'd be surprised*, the facts are not as you would think.

It tells us that the verb *surprise*, often used in passives, formerly took *with*, which is now obsolete, and then *at*, as an agentive preposition. No mention of *by* is made in the definition. There are eleven quotations in which *surprise* is used in the passive. Five of them occur with prepositions: three of them accompany *with* and two take *at*. (There is one quotation with an infinitive and in five quotations the predicate occurs independently.) There is no quotation to illustrate *surprised by*. This does not seem to reflect the true linguistic conditions of the late Modern English period because the investigations into the corpora of late Modern English yields numerous examples of *surprised by*.

4.2 The *OED 2* on CD-ROM

The size of this electronic *OED* as a linguistic corpus can be estimated to be about 35 million words in terms of the number of words in text quotations.⁷ It is a huge historical corpus in its own right and the size of the corpus is such that some meaningful results can be obtained.⁸

Using the *OED 2* on CD-ROM as a corpus, the phrases *surprised at* (*by*, *with*, etc.) were text-searched. *Surprised* itself was also searched in order to find quotations in which the verb and the preposition are separated by intervening words. Variant forms (e.g. *surprized*, *surpris'd*,

surpriz'd, surprysed) were searched as well. Quotations thus collected were then screened and inappropriate quotations such as non-psychological ones (e.g. *surprised by the army*) and inappropriate collocations (e.g. *surpriz'd at unawares*) were discarded. The same quotation used for illustrating two different entries was counted as one occurrence. The title of the book (*Surprised by Joy*) was not counted. A few ambiguous samples, which were hard to determine as to being psychological or not, were excluded. These exclusions, however, do not seem to influence the general ratio of *at* and *by*.

The search and screening yields the results below.

Table 6 The *OED 2* on CD-ROM

<i>at</i>	72 (62%)
<i>by</i>	22 (19%)
<i>with</i>	17 (15%)
<i>of</i>	4 (3%)
<i>among</i>	1 (1%)
total	116 (100%)

The century-by-century distribution of the prepositions is shown in Table 7. It begins with the 16th century when the first instances of *be surprised* constructions used psychologically with prepositions were cited. *Of* and *among* are left out because of very few occurrences. The percentages in parentheses are for the same century.

Table 7 Century-by-century distribution

	16th	17th	18th	19th	20th	total
<i>at</i>	1 (50%)	4 (57%)	23 (70%)	34 (64%)	10 (63%)	72 (65%)
<i>by</i>	0 (0%)	0 (0%)	2 (6%)	14 (27%)	6 (37%)	22 (20%)
<i>with</i>	1 (50%)	3 (43%)	8 (24%)	5 (9%)	0 (0%)	17 (15%)
total	2 (100%)	7 (100%)	33 (100%)	53 (100%)	16 (100%)	111 (100%)

With was prevalent in early Modern English but has gone into disuse in the 20th century (although the investigations of the current corpora have already shown that it is still in use in PE). *At* has been most predominant throughout the Modern English period but seems to have been decreasing since the 18th century. *By* appeared for the first time in the 18th century and has been on the increase in ratio ever since.

The occurrence rate of *surprised by* during the period from the 18th century to the 20th century, a period almost equivalent to the period covered by the Virginia Corpus, is 21.6 percent (22/102), which is surprisingly similar to the ratio of the Virginia Corpus (21.1 percent). The occurrence ratio between *at* and *by* for the same period is 3 : 1, also quite similar to the ratio of the Virginia Corpus (3.7 : 1).

The ratios between the two rival prepositions have been narrowing century by century. If we give the value of 1 to *by*, the ratio of *at* to *by* for each century would be 10 : 1 (18th), 2.4 : 1 (19th) and 1.7 : 1 (20th), which is a clear indication of the increase of *by* and the relative decrease of *at*. If this trend should continue, *by* may catch up with, or overtake, *at* as an agentive preposition. In fact, the results of the Cobuild Corpus have shown that *by* has already surpassed *at*.

5. The historical shift of *at* and *by*

5.1 The shift of occurrence ratios between *at* and *by*

The data from the *OED 2* on CD-ROM and from the current corpora can be combined to exhibit the shift of the occurrence ratios between *at* and *by* in Table 8. The *OED 2* on CD-ROM contains quotations up to 1988. The BNC contains English from 1975 on and the Cobuild from 1983 on. It is evident that *by* has been increasing at such a rate as to be overtaking *at* century by century and is currently surpassing *at* in PE.

Table 8 The shift of occurrence ratios between *at* and *by*

	<i>at</i>	<i>by</i>
18th	10	1
19th	2.4	1
20th (-1988)	1.7	1
PE		
BNC (1975-)	1	1
Cobuild (1983-)	0.9	1

5.2 The historical shift of agentive prepositions

In Present-day English, passives occur mainly with a *by*-agent. Historically, however, this is not necessarily the case and a variety of prepositions have been used. Mustanoja says (1960: 442), “The ME prepositions expressing agency are *by*, *from*, *mid*, *of*, *through*, and *with*” and Visser (1973: §§1987-2000) gives the following prepositions: *among*, *at*, *between/betwixt*, *be/by*, *for*, *fram/from*, *mid*, *of*, *on*, *to*, *pruh/through*, *with*. *By*, however, ousted other prepositions and has now become the

principal agentive preposition in Modern English passives.

With regard to the period when *by* became prevalent as an agentive preposition, Jespersen (*MEG* III: 317) says, “From about 1400 *by*, which originally meant what is now expressed by *through* or *through the means of*, became more and more usual before the converted subject.” Mustanoja (1960: 442) states, “Unambiguous cases with *by* indicating the agent of a passive verb are rare until the end of the 14th century.” Visser (1973), on the other hand, gives a different view and says, “the agentive *by*-adjunct. . . occurred fairly frequent in early Middle English” (§ 1992). Peitsara (1993) provides the most recent comprehensive research, even if it is a pilot study, of the development of *by*-agent with the use of the Helsinki Corpus.⁹ She says, “*By* has thus been clearly predominant since the fifteenth century. . .” (p. 228). Peitsara’s conclusion may be taken to be closest to the truth.

Considering the usurpation of the agentive prepositions of passives with *by* in the history of English, it can be surmised that the same encroachment of *by* has been taking place for the *be surprised* constructions, and possibly for other psychological passives.

6. Conclusion

The historical shift of agentive prepositions introduced by the *be surprised* construction in a psychological sense has been explored using four computer corpora of Modern English and the *OED 2* on CD-ROM. The investigations into the computer corpora show that the occurrence rate of *surprised by* was a little over one-fifth (*at* about three-quarters) in late Modern English, whereas in current English *by* seems to be occurring almost at an equal rate to *at*, or even occurring at a slightly

higher rate than *at*.

The outcome from the computer corpora coincided with the results of the search into the *OED 2* on CD-ROM as a linguistic corpus. Both the occurrence rate of *by* and the occurrence ratios of *at* and *by* for the late Modern English period are quite similar to the findings from the early corpora, which cover roughly the same period. Moreover, *by* has become more and more common since its first appearance in the 18th century, while *at* has been slightly decreasing.

The combined results of the computer corpora and the *OED 2* on CD-ROM show that *by* has been increasing at such a rate as to overtake *at* and is currently surpassing it. *By*, which got the better of other agentive prepositions of passives and ousted them in the history of English, may be replacing *at* for the *be surprised* constructions as well.

Notes

- 1 Not only *be* but other linking verbs like *look*, *seem*, *appear*, etc. occur with *surprised* but I use the formula '*be surprised*' to subsume *look* (*seem*, *appear*, etc.) *surprised* to avoid complexity. A simple '*surprised*' is also used when there is no need to indicate *be*.
- 2 *Be surprised* takes *by* in a physical sense such as 'to attack unawares or suddenly' as in the following example: *The American soldiers were surprised by the guerrillas and many were killed.*
- 3 My investigation into the prepositional complements of *be surprised at/by* shows that there are abundant cases in which abstract nouns are used after prepositions, especially after *by* (e. g. *He is perhaps surprised by her composure*). It also shows that gerunds after the prepositions (e.g. *The host is agreeably surprised by receiving honours and reward*) have been declining.
- 4 All the corpora I used are accessible online. The following are the corpora and their web-sites: The Modern English Collection: <http://etext.lib.virginia.edu/modeng.browse.html>; Public Domain Modern English Search: <http://www.hti>.

umich.edu/english/pd-modeng/; The British National Corpus: <http://info.ox.ac.uk/bnc>; The CobuildDirect Corpus: <http://www.cobuild.Collins.co.uk>
The data of the BNC and the Cobuild were obtained while I was staying in England in 1998.

- 5 The following authors in translation were excluded from the research: Julius Caesar, Anton Chekhov, René George Sand Doumic, Nikolai V. Gogol, Hans Gross, Victor Hugo, Gaston Leroux, Titus Livius, Plato, P. J. Proudhon, Leo Tolstoy and Jules Verne.
- 6 My research into the computer concordance of Dickens (<http://concordance.com/dickens.htm>) shows that the novelist used *surprised at* and *surprised by* 40 times each in the works contained in the concordance. It also reveals that Dickens seems to have shifted his use of preposition from *at* in his earlier works to *by* in his later works. It seems to suggest that the *be surprised by* construction was becoming popular during the 19th century.
- 7 Personal communication with the academic division of Oxford University Press. It has been suggested that the figures for the number of words contained within the text of the quotations cited in the *OED* will be approximately 60 per cent of the figures for the number of words in the *OED* database, which is estimated to be 59 million. It follows that approximately 35 million words will be the figure for the number of text quotations. Fischer (1997: 162), on the other hand, provides a figure of 25 million words. Her estimation is: "the *OED* on CD-ROM contains nearly 2.5 million illustrative quotations. If an average quotation is assumed to be ten words long, the quotation corpus of the *OED* totals about 25 million words"
- 8 For the usefulness of the *OED 2* on CD-ROM as a historical corpus, see Fischer (1997).
- 9 The Helsinki Corpus is a historical corpus created at the University of Helsinki. It contains 1.7 million words of texts of English from 1350 to 1710.

References

- Aston, Guy and Lou Burnard. 1998. *The BNC Handbook*. Edinburgh: Edinburgh University Press.
- Close, R. A. 1975. *A Reference Grammar for Students of English*. London: Longman.
- Declerck, R. 1991. *A Comprehensive Descriptive Grammar*. Tokyo: Kaitakusha.

- Fischer, Andreas. 1997. "The *Oxford English Dictionary* on CD-ROM as a historical corpus: *To wed* and *to marry* revisited." Udo Fries, Viviane Müller and Peter Schneider (eds.), *From Ælfric to the New York Times: Studies in English Corpus Linguistics*. 161-72. Amsterdam: Rodopi.
- Jespersen, Otto. 1927. *A Modern English Grammar on Historical Principles*, Vol. III. London: Allen and Unwin.
- Leech, Geoffrey and Jan Svartvik. 1975. *A Communicative Grammar of English*. London: Longman.
- Mustanoja, T. F. A. 1960. *A Middle English Syntax*. Helsinki: Société Néophilologique.
- Palmer, F. R. 1974/1988². *The English Verb*, 2nd ed. London: Longman.
- Peitsara, Kirsti. 1993. "On the development of the *by*-agent in English." Matti Rissanen, Merja Kytö and Minna Palander-Collin (eds.), *Early English in the Computer Age*. 219-233. Berlin and New York: Mouton de Guyer.
- Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Stein, Gabriele. 1979. *Studies in the Function of the Passive*. Tübingen: Gunter Narr Verlag.
- Svartvik, Jan. 1966. *On Voice in the English Verb*. The Hague: Mouton & Co.
- Swan, M. 1980/1995². *Practical English Usage*, 2nd ed. Oxford: Oxford University Press.
- Taketazu, S. 1999. "Be surprised at or by?: The choice of preposition in Present-day English." *Nagasaki Prefectural University Journal*, Vol. 33, No. 1. 1-18.
- Visser, F. Th. 1973. *An Historical Syntax of the English Language*, Part III. Leiden: E. J. Brill.
- 小西友七 (編). 1980 [英語基本動詞辞典] 東京: 研究社. (Konishi, T. ed. 1980. *A Dictionary of English Word Grammar on Verbs*.)
- 斎藤俊雄・中村純作・赤野一郎 (編). 1998. [英語コーパス言語学] 東京: 研究社. (Saito, T., J. Nakamura and I. Akano (eds.). 1998. *English Corpus Linguistics*. Tokyo: Kenkyusha.)
- 田島松二. 1996. 「'cannot help but do' 型構文はいつ頃文献に現われたか?」『The Kyushu Review』 Vol. 1. 47-53. (Tajima, Matsuji. 1996. "When did the 'cannot help but do' construction appear in literature?" *The Kyushu Review*, Vol.1. 47-53.)

Dictionaries

- 『カレッジライトハウス英和辞典』1995. 東京: 研究社. (*Kenkyusha College Lighthouse*)

- English-Japanese Dictionary*. 1995. Tokyo: Kenkyusha.) [*College Lighthouse*]
『新グローバル英和辞典』1994. 東京：三省堂. (*The New Global English-Japanese Dictionary*.
1994. Tokyo: Sanseido.) [*Global*]
『研究社英語活用辞典』1995. 東京：研究社. (*The Kenkyusha's Dictionary of English
Collocation*. 1995. Tokyo: Kenkyusha.) [*KDEC*]
Cambridge International Dictionary of English. 1995. Cambridge: Cambridge
University Press. [*CIDE*]
Collins COBUILD English Language Dictionary. 1987. London and Glasgow:
Collins. [*COBUILD*]
Longman Dictionary of Contemporary English. 1991/1995. Harlow, Sussex:
Longman. [*LDCE*]
The New Oxford Dictionary of English. 1998. Oxford: Clarendon Press. [*NODE*]
The Oxford English Dictionary, Second Edition. 1989. Oxford: Oxford University
Press. [*OED*]
The Oxford English Dictionary (Second Edition) on Compact Disc. 1993. Oxford:
Oxford University Press. [*OED 2 on CD-ROM*]
Webster's Dictionary of English Usage. 1989. Springfield, Mass.: Merriam-
Webster. [*WDEU*]