

Clause Ordering in English Complex Sentences *

Tae-Sook Park
(Seoul National University)

Park, Tae-Sook. 2002. Clause Ordering in English Complex Sentences. *SNU Working Papers in English Language and Linguistics* 1, 92-108. This paper is a corpus-based investigation of clause ordering in the English sentences which contain a subordinate clause. We analyze the sentences containing thirteen simple subordinators which can be semantically classified into time, reason, condition and concession phrases. The data is from the Lancaster-Oslo/Bergen (LOB) Corpus and the British National Corpus (BNC). Data Analysis shows that the subordinate clauses have a tendency to be positioned sentence-finally in general, with the exception of *if*-clauses. Among the clauses, reason clauses most frequently occur in the sentence-final position. Among condition clauses, the positive *if*-clause frequently precedes the main clause, whereas the negative *unless*-clause often follows it. Data also shows that spoken English reflects this tendency more clearly than written English. We suggest that the clause ordering may be analyzed in terms of the semantic principle of end-focus, on the assumption that most of the subordinate clauses have a relatively heavy meaning weight within complex sentences. (Seoul National University)

Keywords: clause ordering, subordinate clauses, subordinators, end-focus

1. Introduction

This paper is an corpus-based investigation of clause ordering in the sentences which contain a subordinate clause. One major purpose is to examine whether clause order is influenced by the relationship which the subordinate conjunction marks between clauses. The other is to see if clause order shows variation according to the style of the text, particularly according to written and spoken English. Another concern is on the overall frequency of the subordinate clauses in written and spoken corpora.

The data is from the Lancaster-Oslo/Bergen (LOB) Corpus and from some parts of the British National Corpus (BNC). LOB is 1,000,000-word

* This paper was supported by the Grant for the Reform of University Education under the BK21 Project of SNU.

corpus, representing written English. For spoken data, I draw 1,500,000-word corpus from some parts of BNC; this size is intended to be so big as to make the comparative analysis of these two corpora possible.

It will be examined if subordinate clause is positioned sentence-initially or sentence-finally. The medial position is excluded, since it is considered unimportant to our present perspective for clause order. For convenience, initial position and final position will be indicated by respective *I* and *F*. Counting the percentage of two positions, I calculate down to one place of decimals, rounding a figure off the nearest whole number at the second place.

In particular, this paper is intended as a whole picture of thirteen simple subordinators: *after, although, as, because, before, if, once, since, though, till, unless, until, when*. The subordinators analyzed are limited to those for finite clauses, excluding subordinators for nonfinite and verbless clauses.

The subordinate clauses are classified into four major semantic phrases: clauses of time, reason, condition and concession. The following three sections deal with the sentences containing these subordinate clauses. In Clauses of time we will look into subordinators marking temporal relationships: *after, as, before, once, since, till, until, when*. Clauses of reason analyzes three major subordinators marking causal relationships: *as, because, and since*. And Clauses of condition and concession deals with *if* and *unless* as conditional and with *although* and *though* as concessive.

2. Clauses of Time

Some of adverbial finite clauses of time are introduced by one of the following simple subordinators: *after, as, before, once, since, till, until, when*. Temporal relationships between clauses may be divided into three parts: time before, same time, and time after. The following tables show the distribution of those subordinators and the pattern of clause ordering in the sentences containing these temporal subordinate clauses, in LOB and BNC:

Table 1. LOB

Time relationships	Type of clause	Total tokens	Initial tokens (%)	Final tokens (%)
Time before	<i>until</i> -clause	275	18 (6.5)	257 (93.5)
	<i>till</i> -clause	35	0 (0)	35 (100)
	<i>before</i> -clause	364	66 (18.1)	298 (81.9)
Same time	<i>as</i> -clause	549	232 (42.3)	317 (57.7)
Time after	<i>after</i> -clause	136	38 (27.9)	98 (72.1)
	<i>once</i> -clause	103	34 (33.0)	69 (67.0)
	<i>since</i> -clause	54	9 (16.7)	45 (83.3)
	<i>when</i> -clause	2,009	721 (35.9)	1,288 (64.1)
	Total	3,525	1,118 (31.7)	2,407 (68.3)

Table 2. BNC

Time relationships	Type of clause	Total tokens	Initial tokens (%)	Final tokens (%)
Time before	<i>until</i> -clause	154	14 (9.1)	140 (90.9)
	<i>till</i> -clause	54	1 (1.9)	53 (98.1)
	<i>before</i> -clause	354	52 (14.7)	302 (85.3)
Same time	<i>as</i> -clause	237	97 (40.9)	140 (59.1)
Time after	<i>after</i> -clause	202	34 (13.8)	168 (83.2)
	<i>once</i> -clause	154	67 (43.5)	87 (56.5)
	<i>since</i> -clause	54	11 (20.4)	43 (79.6)
	<i>when</i> -clause	2,927	1,011 (34.5)	1,916 (65.5)
	Total	4,136	1,287 (31.1)	2,849 (68.9)

On the whole, temporal subordinate clauses are preferable in *F* in two corpora. 68.3% of total 3,525 tokens occupy *F* in LOB, and 68.9% of 4,136 tokens in BNC.

The following (1) lists the conjunctions introducing subordinate clauses in order of the proportion of occurrences in *F* in two respective corpora:

- (1) *till*(100%) > *until*(93.5%) > *since*(83.3%) > *before*(81.9%) > *after*(72.1%)
 > *once*(67.0%) > *when*(64.1%) > *as*(57.7%) (LOB)

- (2) *till*(98.1%) > *until*(90.9%) > *before*(85.3%) > *after*(83.2%) > *since*(79.6%)
> *when*(65.5%) > *as*(59.1%) > *once* (BNC)

Representing time before, *till*, *until*, and *before* are the foremost three subordinators of very high proportion in *F* in BNC. *Till* occupies *F* in 98.1%, *until* in 90.9%, and *before* in 85.3%. This is similar in LOB; *till*, *until*, and *before* occupy *F* very frequently. All tokens of *till* appear in *F*, *until* occupies *F* in 93.5%, and *before* in 81.9%.

Until, *till*, and *before* indicate that the situation in the main clause occurred before or leading up to the situation in the subordinate clause. *Until* may mark the time up to which the situation in the main clause applies. *Till* is used in the same way as *until*. *Before* also marks the time from which the situation in the main clause applies. Time order in which events happen in the real life may be reflected in clause order; the situations in subordinate clauses containing *till*, *until*, and *before* indicate that they are subsequent to the situation in main clauses. Regarding the order of the proportion in *F*, spoken language reflects this point better than written language. It may signify that spoken language is closer to our real life.

Concerning time after, the subordinators indicate a sequence in which the situation in the main clause occurs after that in the subordinate clause. *After* merely marks the sequence of the clauses. *Since* marks the beginning of the period after which the situation in the main clause applies. *When* may indicate a sequence when the two clauses are nondurative. *Once* adds the notion of proximity in time of the two situations.

The following (3) and (4) show the order of proportion in *F* only in the case of time after:

- (3) *since*-clause > *after*-clause > *once*-clause > *when*-clause (LOB)
(4) *after*-clause > *since*-clause > *when*-clause > *once*-clause (BNC)

Written and spoken English show a little different rate order. The biggest difference is that *since*-clause appears most frequently in *F* in LOB. In this data, even *since*-clause occupies *F* more frequently than *before*-clause representing time before. And *since*-clause and *after*-clause occupy *F* more frequently than *when*-clause and *once*-clause in both corpora.

- (5) This news *came* after the stock market *had closed*. (LOB)
 He *had only ridden* the horse for four months since she *was bought*
 by Sue Welch from Raul Schockemohle. (BNC)
 He *bought* a wine box in when they *were* in France. (BNC)
 It *was* the old restlessness that drove me on once I *found* myself
 alone, away from the calming influence of his presence. (LOB)

In the cases of *since* and *after*, there are many tokens which use perfect tense in one clause. Particularly, most of the *since* tokens have perfect tense in main clauses. This indicates that two situations are more distant from each other in the cases of *since* and *after* than in *when* and *once*. They are definite in the distance between two situations. They indicate a sequence in which the situation in the main clause occurs after that in the subordinate clause. If the order in real life is reflected in language according to that, *since*-clause and *after*-clause will be positioned initially in a high frequency. However, they are positioned finally in about 80 % of written and spoken tokens. Here we assume that temporal subordinate clauses tend to be positioned finally, if definite in the distance between two situations. In the case that two situations are definitely distant, the situation of subordinating clause may be largely heavier than that of the main clause in the weight of meaning. This aspect can explain that these subordinate clauses are mostly positioned finally.

Two situations may be more proximate in *when* and *once*. Proximity of situations means that the order of situations can not be definite but may be changeable. This aspect may account for the relatively low frequency in *F* in the cases of *when* and *once*. These clauses appear in *I* in a relatively high frequency, compared with other temporal subordinate clauses. Particularly, *once*, which adds proximity in time of the two situations, occupies *F* least frequently in the spoken BNC corpus. This fact also indicates that spoken language reflects the language aspect of our real life more accurately than written language.

As, as our representative example of same time, indicates the simultaneity of the situations in main and subordinate clauses, or at least an overlap in time of the two situations. Thus, *as* implies much more proximate property between two situations than *once* of proximity. This aspect is also reflected to a certain extent in our corpora. *As*-clause occupies *F* in relatively very low frequency in comparison with the rest

temporal subordinate clauses: 57.7% in LOB and 59.1% in BNC. The simultaneity or overlap of *as* implies that the order of two situations are free in real life and thus that the order of two clauses are less restricted in sentential position than other clauses showing the definite distance between two situations.

As occupies *F* in the least frequency in LOB. It applies to BNC, excluding *once*. The fact that *once* is more frequent in *I* than *as* in spoken language is not accordant with the above explanation that spoken language reflects real life aspect more clearly than written language. Leaving this matter for a further study, I suggest the possibility that *once* and *as* may be similar in the degree of proximity, *once* categorized into the subordinator representing same time.

Now, let us consider the overall frequency of eight temporal subordinate clause types in both corpora. The following (6) and (7) show the order of subordinators in frequency:

- (6) when(2,009) > as(549) > before(364) > until(275) > after(136) >
 once(103) > since(54) > till(35) - (LOB)
 (7) when(2,927) > before(354) > as(237) > after(202) > once(154)/until(154)
 > since(54)/till(54) - (BNC)

When is the most frequent temporal subordinator in both written and spoken corpora. The number of its tokens is over 2,000 and outnumbers the occurrences of the rest. *Till* is used least frequently in both. *As* and *before* are shown to be used many times next to *when*-clause. *Since* and *till* are common in having the lowest frequency as temporal subordinate clauses.

In terms of frequency, eight temporal subordinate clauses are distributed roughly similarly in the written LOB corpus and the spoken BNC corpus, except for two subordinators, *as* and *until*. *Until* appears in 275 tokens in 1,000,000-word LOB corpus, but only in 154 tokens in the 1,500,000-word corpus of BNC. This indicates that *until* is used more preferably in written English. *Till* is used in the similar frequency in written and spoken English to consider the size of our two corpora. *As* appears in 549 tokens in LOB and in 237 in BNC. This also shows that *as*-clause as a temporal function is used much more frequently in written English than in spoken English.

Next, we look into clause ordering in terms of pragmatics. Here the principle of end-focus can be considered an important pragmatic device to account for clause order in complex sentences. That principle says that people process the information in a message so as to achieve a linear presentation from low to high information value. In other words, new information tends to follow given information within sentences; new information, or the focus tends to be positioned finally, having a heavy meaning weight.

In this analysis I assume that given information includes evoked information and inferrable information. Evoked information is that referred in the previous context, and inferrable information can be inferred from the context. The new information often needs to be stated more fully than the given, having a longer, 'heavier' structure. Thus the principle of end-focus comes into operation along with an organization principle end-weight.

Now, let us see the order of the overall proportion of temporal subordinate clauses in the sentential *F*, combining LOB and BNC:

- (8) *till*(99.1%) > *until*(92.2%) > *before*(83.6%) > *since*(81.5%) > *after*(77.7%)
 > *when*(64.8%) > *once*(62.8%) > *as*(58.4%)

The more frequently a subordinate clause occupies *F*, the heavier meaning it represents. Put another way, subordinate clause types, which occupy *F* very frequently, is heavier in meaning than other clause types. For example, *till*-clause may deliver more important information than *before*-clause, and *after*-clause may be heavier than *once*-clause, etc.

- (9) You just wait *till you meet the girl who thinks you're a god*. (LOB)
 (10) She wants to be a nurse. Oh does she? . . . You can't be a nurse
till you're eighteen. (BNC)

Till-clause occupies *F* most frequently among eight clause types. This indicates that *till*-clause is the heaviest subordinate clause type in meaning. In (10), the initial main clause contains evoked information, *a nurse*, while *till*-clause represents new information.

Reverse ordering, shown in a small part of tokens, is examined in the following (12):

- (11) . . . you got ta use erm, well, they use loose tea I mean, it is different, I never used to like the loose tea when I was younger but now I sort of really taste the difference. (BNC)
- (12) On the 14th , two days before Chamberlain was supposed to start, this message was in Lytton's hands. When Lytton received the telegram, however, he was in no mood to delay. (LOB)

Here, it is supposed that the final clause, if main or subordinate, delivers the focus, or what authors or speakers intend to put a heavy meaning weight on. In the underlined sentence of (11), the main clause represents given information, containing the evoked element, *the loose tea*, while the subordinate clause represents new and heavy information. On the contrary, in (12) the subordinate clause in *I* represents given information; that Lytton received the telegram is inferred from the fact that this message was in Lytton's hands. Still, the final main clause represents new information, the focus, or the heavy meaning. This aspect is applicable across the board in our tokens.

3. Clauses of Reason

Three major reason subordinators are *because*, *since*, and *as*. They all agree in implying a reason for an occurrence or action, but have a slightly different shade of meaning, along with a stylistic difference. *Because* introduces a direct reason, while *as* and *since* imply merely circumstances attendant on the main statement.

According to English Usage, if the situation of reason clause is the most important in the sentence, *because* is used as a subordinator. And *because*-clause is mostly in *F*. If the situation of the reason clause is already known or less important than the other elements of sentence, *since* and *as* are used as subordinators. *Since*-clause and *as*-clause are often in *I*.

Now, let us see the result of our corpus analysis. The following tables show the distribution of reason subordinators and the pattern of clause ordering:

Table 3. LOB

	Type of clause	Total tokens	Initial tokens (%)	Final tokens (%)
Reason	<i>because</i> -clause	604	94 (15.6)	510 (84.4)
	<i>since</i> -clause	198	72 (36.4)	126 (63.6)
	<i>as</i> -clause	158	67 (42.4)	91 (57.6)
	Total	960	233 (24.3)	727 (75.7)

Table 4. BNC

	Type of clause	Total tokens	Initial tokens (%)	Final tokens (%)
Reason	<i>because</i> -clause	2,964	400 (13.5)	2,564 (86.5)
	<i>since</i> -clause	73	27 (37.0)	46 (63.0)
	<i>as</i> -clause	108	20 (18.5)	88 (81.5)
	Total	3145	447 (14.2)	2,698 (85.8)

The order in the proportion of subordinate clauses in *F* is *because*-clause, *since*-clause, and *as*-clause in LOB. However, the order is *because*-clause, *as*-clause, and *since*-clause in BNC. *Because*-clause occupies *F* in 84.4% of LOB tokens, and in 86.5% of BNC tokens. Combining two corpora, about 85.5% of all *because* tokens occupy *F* and 15.5% occupy *I*. This is accordant with the English usage statement. However, *since* and *as* disagree to that statement, for both *since* and *as* appear more frequently in *F* than in *I*. In the case of *since*, 63.6% are in *F* in LOB and 63.0% in BNC. *As*-clause occupies *F* in 57.6% in LOB and in as much as 81.5% in BNC. It is positioned finally much more frequently in spoken English than in written English. These show that the English usage statement is not based on the corpus analysis.

On the whole, the subordinate clause of reason tends to be in *F*. 78.6% of all LOB and BNC reason clauses are shown to occupy *F*. This means that the reason of occurrences or actions is regarded as important information in sentences. Or the reason clause represents relatively new information with a heavy meaning weight. Particularly, *because*-clauses that are positioned finally in the highest frequency are shown to have the heaviest meaning weight among the clauses of reason.

Now let us see these subordinators with the sentences. *Because*-clause, which occupies *F* most frequently, expresses the direct and specific reason of the main statement.

- (13) I knew that the woman, who turned into my sister-in-law, wouldn't understand how I felt about everything. I was desperate, because I knew I would have to stay for a holiday. (LOB)
- (14) She came and made me a ham sandwich and the maid just look at her to eat it now. Brilliant! There really was not a scrap of. That's awful because the housekeeper sorts all that out, so when she's away no one does it. (BNC)

As seen in the above underlined sentences, the main statements are all caused directly from the situation of the *because*-clause; desperation in (13) and awfulness in (14) are derived from the following subordinate statement. The reasons here are not inferred from the general situation but specified by authors or speakers. In terms of the end-focus principle, the above sentences show that *because*-clauses represent new information and that main clauses represent given information. The main statements are all inferrable from the previous sentences, and the subordinate statements have a focus and a heavy meaning weight. Thus *Because*-clause may represent the most important information and tend to be in *F*.

In reverse, 15.5% of all *because*-clause tokens occupies *I* in our corpora. What factor will make some *because*-clauses positioned initially? The end-focus principle can also account for this ordering.

- (15) Serious art, by and large, does make demands that popular art does not: sensitive discrimination, awareness of form, some familiarity with technique, and, above all, an active and concentrated attention. in so far as this is the case, serious art is not easily accessible to the untutored. Because facility with serious art requires skill and knowledge not acquired incidentally, it makes sense for the school to offer a programme of art education. (LOB)
- (16) But although the fundamental forces had already changed by 1920, it took many years for the full consequences to work themselves out. And because the process of adjustment took so long and was so slow, it was a long time before the change in the underlying situation was recognised. (LOB)

In the above sentences, the statements of *because*-clauses are inferrable from the previous sentences. In (15) we can infer that facility with serious art requires skill and knowledge not acquired incidentally, from the fact that serious art makes demands that popular art does not and is not easily accessible to the untutored. In (16) the fact that it took many years for the full consequences to work themselves out signifies that the process of adjustment took so long and was so slow. Representing lighter information than the main clause, the *because*-clause can be positioned initially. An author or a speaker may be able to use this device in order to put a focus and weight on the main statement.

Concerning *since* and *as*, they tend to express the indirect and general reason of the main statement, as seen in the following:

- (17) The Russians are obviously impressed by the standard of schools in Cheltenham. An exact replica of one of them is being shipped to Armenia, right down to blackboards and chalk, but it's taken top level agreement between Mr Gorbachev and Mrs Thatcher to make it happen. Tim Hurst reports Armenian schoolchildren are being taught in tents since earthquakes destroyed their school in Lininecarn killing 52 children and 6 teacher. (BNC)
- (18) This cage is capable of being divided into two halves, by means of a sliding hardboard partition. I can recommend this type of cage, as it is impossible for the birds to throw out any seed husks, and . . . (LOB)

The reasons of *since*-clause and *as*-clause are inferrable from the previous context or the general principle. In spite of it, they tend to occupy *F*. This may indicate that the reason is generally considered important information in sentences. In the above, the final subordinate clauses represent heavier information than the initial clauses. In (17) the initial main statement contains Armenia as evoked information and the final reason clause represents inferrable information. In (18) the cage represents evoked information, and the reason clause is inferrable.

The following examples present the initial *since*-clauses and *as*-clauses:

- (19) While not as popular as his Second Concerto (what is?), the First is a fine romantic work. And since Rachmaninov wrote it for himself

- to play*, it demands a tremendous technique from the soloist. (BNC)
- (20) Since six months ago, when I calculated that there were more than two hundred and fifty of them, I reckon that the figure is now not far off the three hundred mark. As this is one of the busiest seasons of the year for exhibitions, I am receiving daily so many invitations for private views that I have to decide which shows are not worth seeing. (LOB)

The above sentences are also explainable in terms of end-focus. Initial *since*-clauses and *as*-clauses represent less heavier information than final main clauses. In (36) the fact that this is one of the busiest seasons of the year for exhibitions is inferrable from the present situation while the main statement represents relatively new information. Here, since they represent given information, reason clauses are positioned initially. And the main statements are relatively heavy in meaning weight.

In terms of overall frequency, *because* and *since* show a big difference between two corpora. *Because*-clause appears in 510 tokens of LOB but in as many as 2,564 tokens of BNC. Considering their size, the number of tokens in BNC is about four times as big as that in LOB. On the contrary, *since*-clause appears in 126 tokens of LOB but in as few as 46 of BNC. This number in LOB is also about four times as big as that in BNC. In comparison, *because* is very preferable as a reason subordinator in spoken English, and *since* is used often in formal and literary text. On the whole, the reason subordinate clauses are very preferable in spoken English, as seen in total 2698 tokens in BNC.

4. Clauses of Condition and Concession

The two simple subordinators for conditional clauses are *if* and *unless*. The most common and most versatile of the conditional subordinators is *if*. The negative subordinator *unless* is the next most common. Clauses of concession are introduced chiefly by *although* or *though*.

The distribution of *if* and *unless* and the patterning of clause order are seen in the following tables:

Table 5. LOB

		Total tokens	Initial tokens (%)	Final tokens (%)
condition	<i>if</i> -clause	1,871	1,275 (68.1)	596 (31.9)
	<i>unless</i> -clause	153	27 (21.4)	126 (82.4)
	Total	2,024	1,302 (64.3)	722 (35.7)

Table 6. BNC

		Total tokens	Initial tokens (%)	Final tokens (%)
condition	<i>if</i> -clause	4,833	3,627 (75.0)	1,206 (25.0)
	<i>unless</i> -clause	134	37 (27.6)	97 (72.4)
	Total	4,967	3,664 (73.8)	13,03 (26.2)

If and *unless* show a considerable difference in the proportion of tokens in *F*. *If* occupies *F* in 31.9% in LOB and in 25.0% in BNC. *Unless* occupies *F* in 82.4% in LOB and in 72.4% in BNC. In conditional clauses, the position may depend on their positiveness or negativeness. *If*-clause, positive subordinate clause, is used in *I* in 71.5% of our all tokens, and *unless*-clause, negative subordinate clause, is used in *F* in 77.4% of all tokens. This implies that negative condition is heavier than positive condition in terms of meaning.

- (21) You're going to do this job for me without any more argument.
If you do it right, you'll get paid right. (LOB)
- (22) Or do you think that's enough? That's enough. Right. Well we'll do that unless they suggest somebody else Er, what to do. (BNC)

In (21), the initial *if*-clause is only a kind of condition for action of main statement, representing given information. The focus and weight is on the final main statement. In (22), however, the final *unless*-clause is the single condition set up for the main statement and has a focus and a heavy meaning weight within the sentence.

The following examples present the clauses of reverse order, which occupy a small part of our tokens.

- (23) . . . and it seems somehow absurd to accept the doctrine of precedent if we have to admit that we are not able to say what is the ratio of a particular case. The difficulty may perhaps be solved *if it is realized that there are really two problems involved in the use of cases.*
(LOB)
- (24) The rash tends to spread from the child's cheeks to his head, neck, body and limbs. It's an uncomfortable condition because of the irritation, and *unless he is checked,* the child will inevitably scratch.
(LOB)

Here, focus and weight are put in reverse order. In (23), it is not the initial main clause but the final *if*-clause that is heavy in meaning weight. The final *if*-clause implies that the condition is very important for the situation of main statement. Besides this, the initial main clause represents lighter information than the final *if*-clause. In contrast, in (24), the main clause is a focus, representing new information. Here, the initial *unless*-clause is like the final *if*-clause in that it is just a kind of small condition for the main statement action.

In terms of overall frequency, *if*-clause appears in a extremely high frequency in BNC; it appears in 1,871 tokens in LOB but in 4,833 in BNC. *Unless*-clause is used a little more often in LOB than in BNC. On the whole, conditional clauses are used in spoken English much more frequently than in written English, which is due to the high frequency of *if*-clause in spoken corpus. The number of total tokens in BNC is twice as big as that in LOB.

Concerning *although* and *though*, see the following tables:

Table 7. LOB

		Total tokens	Initial tokens (%)	Final tokens (%)
concession	<i>although</i> -clause	370	169 (45.7)	201 (54.3)
	<i>though</i> -clause	310	64 (20.6)	246 (79.4)
	Total	680	233 (34.3)	447 (65.7)

Table 8. BNC

		Total tokens	Initial tokens (%)	Final tokens (%)
concession	<i>although</i> -clause	276	109 (39.5)	167 (60.5)
	<i>though</i> -clause	118	19 (16.1)	99 (83.9)
	Total	394	128 (32.5)	266 (67.5)

Although-clauses are distributed similarly in *I* and *F* in LOB and are positioned finally a little more frequently in BNC. *Though*-clauses tend to be positioned finally in both corpora: 79.4% in LOB and 83.9% in BNC. This indicates that *though*-clause may have the heavier meaning weight than *although*-clause.

Both *although* and *though* appear more frequently in LOB than in BNC. Particularly, *though* appears much more frequently in LOB than in BNC: 246 in LOB and 99 in BNC. This is not accordant with the English usage statement that *though* is the more informal variant of *although*. According to our corpus analysis, *although* and *though* are used in written English more frequently than in spoken English. Thus the concessive clauses containing *although* and *though* are preferable in written English, and those subordinators are formal and literary. It is guessed that *even if* or *even though*, excluded in this paper, may be preferable in spoken English.

5. Conclusion

We have investigated if clause order in the complex sentences is variable according to whether the conjunction marks time, reason, condition or concession. Clause Ordering is a little variable according to the degree of importance of those relationships. It does not show a big difference between written and spoken English. Still, spoken English may reflect the characteristics of subordinate clause types a little more clearly than written English. Spoken English may be shown to be closer to our real life aspect.

In general, clauses of time and reason tend to be positioned sentence-finally. In the case of clauses of condition, their positiveness and negativeness are the factors which have an effect on clause ordering.

The positive *if*-clause is positioned initially very frequently, while the negative *unless*-clause is often positioned finally. Clauses of concession tend to be in *F* on the whole, though *although*-clause is distributed roughly similarly in *I* and *F* in LOB. Comprehensively, the subordinate clauses tend to be positioned finally, with an exception of *if*-clause.

Clause ordering has also been analyzed in terms of pragmatics. The principle of end-focus has been considered as an important pragmatic factor. Reason clauses which occupy *F* in the highest proportion (83.4%) are supposed to have the heaviest meaning weight. Temporal clauses which occupy *F* in 68.6% are regarded as relatively heavy. *Till*-clause is on the top and *as*-clause is at the bottom in the order of meaning weight. It has been suggested that the real life aspect is reflected well in clause order, particularly in temporal clauses. In conditional clauses, *if*-clause is not heavy in meaning weight, while *unless*-clause represents relatively new and important information. Concessional clauses occupy *F* in 66.6%, and also have a heavy meaning weight to some extent.

The principle of end-focus has also applied to clause ordering in this way. If it represents new information, the subordinate clause may be positioned finally. In reverse, representing given information, it will be positioned initially. Put it in another way, the heavier information it delivers, the more finally it tends to be positioned. Thus, the final clauses, though main or subordinate, are thought to represent the relatively heavy and new information as a focus.

In conclusion, the simple subordinate clause types, except for *if*-clause, occupy *F* more frequently than *I*. In general, therefore, most of the subordinate clauses may have a relatively heavy meaning weight within complex sentences.

References

- Kennedy G. 1998. *An Introduction to Corpus Linguistics*. Longman.
 McEnery T., and A Wilson. 1996. *Corpus Linguistics*. Edinburgh University Press.
 Quirk, R., S. Greenbaum., G Leech., and J Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman.
 Swan M. 1980. *Practical English Usage*. Oxford University Press.

Corpora

1. The 1,000,000-word Lancaster-Oslo/Bergen (LOB) Corpus.
2. The 1,500,000-word corpus from the British National Corpus (BNC).

Tae-Sook Park
cosmotsp@hanmail.net