A Corpus-based Approach
on the Development of Conjunction *while*

Binna Lee

(Seoul National University)

Lee, Binna. 2006. A Corpus-based Approach on the Development of Conjunction *while.* **SNU Working** *Papers in English Linguistics and Language 5, 7489.* The purpose of this paper is to investigate the development and grammaticalization of the conjunction *while* in English. Based upon the previous studies, this study examines how the Old English phrasal construction p *hwile þ* 'at the time that' was reduced later on to the lexical form *hwile* and ultimately to the conjunction *while.* Subjectification seems to play a major role in this grammaticalization process, that is, the construction, starting as a mere lexical phrase, undergoes the process to have a new grammatical function as well. A quantitative analysis using the Oxford English Dictionary (OED) and the Helsinki Corpus will be conducted for the research of the semantic and morphological change of the use of *while.* The corpus results of this study account for the correlation between the loss of the form *hwile* and the increase of the form *while* with the introductory of the supplementary causation function. **(Seoul National** University)

1. Introduction

1.1 Grammaticalization and corpus linguistics

This paper seeks to examine the development and grarnmaticalization of the English conjunction *while* describing the semantic and morphological change within a corpus-based framework. In recent years, grammaticalization has been considered to be of significant importance with growing body of literature (Traugott and Heine **1992,** Hopper and Traugott 2003). The grammaticalization theory sheds light on the contemporary linguistic

phenomenon that has undergone historical process with functional change across times. Kurylowicz (1965) has given a widely used definition of grammaticalization as follows:

> Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status, e.g. from a derivative formant to an inflectional one.
>
> (Kurylowicz 1965:69)

The general hypothesis made within grammaticalization theory is that function words are said to have derived from content items or phrasal constructions in a preliminary stage on the language as it continues to develop new grammatical functions.

To reiterate, from a grammaticalization perspective, the on-going process of lexical meaning changes can be seen as the "subset of linguistic changes whereby a lexical item or construction in certain uses takes on grammatical characteristics, or through which a grammatical item becomes more grammatical." (Hopper and Traugott 2003:2). Likewise, grammaticalization is of great importance in the field of historical linguistics as it gives a general overview of syntactic and morphological change within a specific language.

As the grammaticalization theory offers a possible linguistic analysis of the evolving process of morphological and semantic change, reflections on the collaboration between grammaticalization and corpus linguistics should be under consideration. As the two fields of study share common ground with similar linguistic aspects, the synthesis of both linguistic approaches has major advantage of revealing the system of language alternation leading to attractive insights about the functional change of language.

1.2 Source of data

A corpus-based approach will be used to investigate the linguistic phenomenon of the grammaticalization of the conjunction *while*. In this paper, I attempt to display the developmental characteristics of

the conjunction *while* with the aid of an extensive use of the historical corpora as it is an effective source of historical data to obtain essential information about the terms that are in need for the research. The most widely known English historical corpus is the Helsinki Corpus. The Helsinki Corpus has historical English data from the early Old English Period to the end of the early Modern English Period with the number of over 1.8 million words.

The additional contribution is based on the Oxford English Dictionary (OED) which will be used as a semantic basis to understand the distinctive use of meanings of *while* that are still prevalent in Modern English. The OED is known as the largest dictionary of the English language. The second edition of OED which had been made in 1994 and available on CD-ROM will be the version of dictionary that is used in the present study. It is said that a dictionary is a characteristic of notions and ideas that are institutionalized by the standard of a culture based on corpus evidence though the corpus data itself is not the input in dictionaries. However, it would be of an interest to collect data from the language as an institutionalized group of words as the grammatical process of conjunction is the goal of this paper.

With the research of grammaticalization process based on the corpus-based framework, I will take Hopper and Traugott's own major study examples, especially the intensive research of the change of conjunction *while,* and compare their account with observable use in frequency in the OED quotation base and the Helsinki Corpus. Frequency seems to play a major role in grammaticalization and it is a significant claim that the present study observes.

The organization of this paper is as follows. The succeeding section introduces previous studies of the conjunction *while* within a grammaticalization framework In section **3,** the corpus-based approach to the grammaticalization process of *while* will be considered in specific details with frequency relevance and corpus results. The final section puts forth the conclusion of this paper with reference to the significance of the findings from the corpus data.

2. Previous studies on the development of while

The developmental process of the English connectives or conjunctions can be shown by the grammaticalization the functional items go through. The preliminary stage or the initial starting point is generally a lexical word or a chunk of phrasal construction. **As** the lexical items or constructions be placed in linguistic contexts, specialized grammatical functions are served while continuously developing new additional grammatical functions. It is generally assumed that grammaticalization process consists of successive periods. The change from content items or lexical phrases to functional elements appears in a relatively initial stage. Later on more or less grammaticalized elements acquire other complex or different functions. This process can be schematized as follows:

(1) phrasal **construction** → lexical element → functional element **1**
    → unctional element 2

In the final stage, a new function is attributed to the preexisting function element which is a new aspect of the grammaticalization framework what is called subjectification introduced by Traugott (1995). The essential **keypoint** of the notion of subjectification in grammaticalization is that two different but parallel linguistic changes undergo during the grammaticalization process. While a linguistic item changes grammatical function from a lexical word its meaning alternates from being objective, propositional and externally based to being subjective, judgmental and speaker-based. In other words, subjectivity presupposes some degree of incorporation of the listener in the description of an entitiy or a certain process. The subjectification hypothesis makes assumption that the degree of perceiver integration can increase diachronically (Cuenca 1997, cited in Traugott 1995).

Furthermore, it is commonly assumed that subjectification is a unidirectional phenomenon, not bidirectional, as the on-going linguistic process throughout the linguistic history gives no possibility of direction changing as linguistic elements become more functional and abstract from a lexical element or a content word to

a grammatical function understood in various ways. To reiterate, if there is an occurrence of a semantic change, a linguistic form develops from less subjective usages to more subjective ones, but never the other way round, in reverse, as it is the reason that "historically, virtually all grammatical markers of clause combining can be shown to have developed out of some earlier more 'objective' function" (Traugott 1995).

It is noteworthy that the assertion of semantic change with periodical stages is understood to account for connectives or conjunctions also. This type of grammaticalization is illustrated by the on-going change that has occurred from early Old English to the early Modern English as we see in the conjunction *while* in Modem English.

The Old English adverbial phrase *þ hwile þ* which had the meaning 'at the time that' (lexical element) eventually developed into the temporal conjunction *wlzile* undergoing the grammaticalization process. Historically, *wlzile* was a noun in Old English with the form *hwile* meaning 'length of time' which is still **perseved** in present day English as in I *was confused for a while.* The complementizer *þ* is lost at the moment and the noun *hwile* is assumed to be the final lexical category (lexical element) to withstand the grammaticalization. The next change is where the form *wlzile* develops from *a* temporal conjunction (functional element 1) into a concessive connective (functional element 2) which can be seen as an example of the final stage in the gramrnaticalization process (Traugott 1995).

There is clear evidence for **subjectification** on examining the development of the English successive connective *while.* As mentioned above, the Old English noun *hwile* 'time' used in adverbial phrase p *hwile þ* 'at the time that' is the focus of lexical word stage but later on, reduced in late OE to the simple conjunction *(h)wile,* which has more subjective meaning than the adverbial phrase "because, instead of profiling a specific time, it now profiles discourse structure" (Traugott 1995). From this grammatical function, the **form** *wlzile* became conventionalized in English in a more subjective usage, more specifically, as a concessive connective or conjunction. The following (2) is a simplification of the whole grammatical development of the

conjunction *while.*

(2) **[Grammaticalization of *while*]**
    þ hwile p 'at the time that' > 'while (temporal)'
    > 'while (concessive)'

As mentioned, the conjunction *while,* originally Old English (OE) *þa hwile þe* 'at the time that', expressed temporal simultaneity. As it later on develops into the conjunction form in late OE, it attracted inferences of causation, as in the quotation taken from Hopper and Traugott (1993).

(3) Thar mycht succed na female, Quhill foundyn mycht be
    ony male.
    'No female was able to succeed while any male could be
    found.'
    (1375, Barbours Bruce 1.60) [OED while 2a]

Additional semantic changes are attributed to the extension of informativeness inferences including the development of concessives or causations, and diverse meaning shifts which may be broadly classed as increase in expressed subjectivity.

    Relevant to subjectification, the semanticization of inferences must play a crucial role in the development of *while* (Hopper & Traugott 2003). This conjunction originated in Old English (OE) in the adverbial phrase þ *hwile* þ consisting of the accusative distal demonstrative, the accusative noun *hwile* 'time', and the invariant subordinator þ a highly explicit coding of simultaneity. Traugott also noted a trend in the usage of *while* when studying examples from 14th century where the temporal frame of reference was dominated by an inference to grounds for the situation. I will not go further into this matter as it is left for further studies in a more detailed quantitative research.

    The earliest evidence of the grammaticalization of *while* is shown in Middle English **with** the change from phrasal construction to lexical form appears in this period. In the next section, the distribution of forms of *while* will be examined with correlation with the facts studied from the grammaticalization process. The

corpus data and quotations from the OED will display the historical change from a phrasal construction to a lexical form with limited scope to the issues of subjectification.

3. Findings from corpus data

Based on the grammaticalization framework, I will examine the quantitative frequency change of the conjunction *while* and make a correlation with the meaning shift to functional purpose. **As** mentioned in the earlier chapter, the Helsinki Corpus will be the source of data collection in this research with the OED as a useful tool for understanding the meaning change of the conjunction *while.* The following table is the tokens and frequencies and percentage by periods from the Helsinki Corpus.

(Table 1> The diachronic information of Helsinki Corpus

Subperiod	Dates	Words	%
Old English			
OE1	-850	2,190	0.5
OE2	850-950	92,050	22.3
OE3	950-1050	251,630	60.9
OE4	1050-1150	67,380	16.3
Middle English			
ME1	1150-1250	113,010	18.6
ME2	1250-1350	97,480	16.0
ME3	1350-1420	184,230	30.3
ME4	1420-1500	213,850	*35.1*
Early Modern English			
EModE1	1500-1570	190,160	34.5
EModE2	1570-1640	189,800	34.5
EModE3	1640-1710	171,040	31.0
Total	-	1,572,820	

(from Kyto's General Introduction to the Helsinki Corpus)

The historical corpus data will illustrate the grammaticalization and various stages of semantic changes of a lexical item or phrasal constructions. As the grammaticalization of the conjunction *wlzile* has undergone meaning or functional change throughout all the subperiods, all data should be considered. A fair amount of the linguistic element *while* was examined in the Helsinki Corpus for individual periods and the tokens are shown in the following tables.

To examine the meaning shift process of the lexical item *while,* other possible candidates of the similar morphological form of *while* should *be* considered. The seven morphological forms listed in Table 2 are from the same etymology of *while* as the source is provided from OED. The tokens and frequencies of the variant forms of *while* by period are displayed in below.

<Table 2> Tokens of the variants of *while* by period

	OE	ME	EModE	Total number of tokens
hwile	166	57	0	223
while	1	90	162	253
whilst	0	0	34	34
whilen	0	1	0	1
whilom	0	4	0	4
hwilem	0	2	0	2
hwilen	0	4	0	4
hwiles	0	2	0	2

As shown in the Table 2, although the form *while* diverged into several forms earlier, the prominent forms of the conjunction *wlzile* in present-day English assumed to be relevant to the gramrnaticalization process are the forms *hwile* and *while*. **As** mentioned in the previous section, the morphological lexical form *hwile* derives from the adverbial phrase *Þ hwile Þ* This construction has been prevalent until the 13th century according to

the OED. And it is evidently shown in the table above that the form *hwile* is no longer in existence after the early Modem English period. **As** the five forms other than *hwile* and *while* do not obtain high frequency in the data, this paper will shed light on the prominent two forms. Although the form *whilst* is prevalent in the present-day English, no tokens are counted in the OE and ME period which the meaning change process occurred, an inference can be made that **the** form *whilst* did not give any effect on the grammaticalization of the conjunction *while*.

The form *hwile* should be considered as of great importance in the grammaticalization of developing into a functional item. A comparison between the two forms will be under consideration additionally. **As** the word *hwile* is separated independently from the cluster of adverbial phrase expressing temporal simultaneity, the linguistic element develops into a lexical form with the lexical meaning 'time' obtaining. The following table shows the frequency of the adverbial phrase construction used throughout the periods.

<Table 3> The occurrences of the adverbial phrase P *while/hwile* P by period (number of tokens)

	Number of tokens of P *while* Þ	Number of tokens of Þ *hwile* P
OE1	0	3
OE2	0	26
OE3	0	23
OE4	1	21
ME1	6	4
ME2	1	0
ME3	0	0
ME4	0	0
EModE	0	0
Total	**8**	**77**

The table **3** supports the evidence that the adverbial phrase

construction _þ hwile þ_ was productive in contexts with high usage in OE while the frequency of token dropped dramatically in the ME period. It seems that the grammaticalization of form alternation from cluster construction to lexical form was build up in this period. And after the period ME2, the phrasal construction does not appear anymore as no token is counted. Later on, this paper will make an assumption that the increase of _while_ is relevant to the loss of the _hwile_ form. The frequency of phrasal construction diminishing and finally loss in the ME should be emphasized in the data.

According to the OED, the adverbial phrase _þ hwile þ_ is considered to be a phrasal construction existing in OE period that has been prevalent until the 13th century where the construction Poses the function of integrity. And the first stage of grammaticalization process to a lexical form is said to be _þ hwile þ_ with the _hwile,_ not the _while_ form.

However, findings from the corpus display the fact that the adverbial phrase _þ while þ_ was additionally used not in great numbers but still in existence. The corpus data show evidence for two forms _while_ and _hwile_ in coexistence in the same phrasal clusters. The phrasal construction containg the form _while_ must have been erased altogether with the _hwile_ form in the ME period where the lexical form becomes a self-reliant form of contemporary conjunction _while._

The phrasal constructions of the _while_ form as in the adverbial phrase '_þ while_ p' can be found from the followings examples in the Helsinki Corpus in ME1 period:

(4) a. ... de half don; after þes kinges dom. pa **while pe** pa king loh; pa ampulle heo ...
  b. ... eora hond. & feden Leir þane king; **þe while pe** he leouede. þaies & nihtes mi ...
  c. ... þanne þeoden. & libben on lisse; **þe while pe** he leouede. pus heo pa ispek ...
  d. ... enske; and bi-tachet hit is childe. **te while þe** he mai hit walden; for ofte hi ...

(quotations from Helsinki Corpus ME1)

From the evidence shown in the data, the assumption that conjunction *while* was used in an adverbial phrase from OE4 to ME2 could be made as the correlation between *hwile* and *while* can be of great interest. As a new function load increases during the ME period, the establishment of a new form rather than *hwile* which was used in a phrasal construction must have been in need. This will be discussed in the later on. The following table clearly shows the quantitative change in the two elements.

<Table 4> The percentage of the tokens of *hwile* and *while*
(tokens in parentheses)

Subperiod	(Number of Tokens) *hwile*	(Number of Tokens) *while*
OE1	1.3% (3)	0% (0)
OE2	22.3% (52)	0% (0)
OE3	33.6% (75)	0% (0)
OE4	16.1% (36)	0.4% (1)
ME1	24.7% (55)	6.8% (17)
ME2	0.9% (2)	5.1% (13)
ME3	0% (0)	19.4% (49)
ME4	0% (0)	4.3% (11)
EModE1	0% (0)	15.8% (40)
EModE2	0% (0)	19.4% (49)
EModE3	0% (0)	28.9% (73)
Total	100% (223)	100% (253)

It is obvious from the numbers shown in the table 4 that there is a mutual relationship between the diminishing of the form *hwile* and the increasing frequency of the **use** of *while* up to the early Modem English period. The alternation from *hw-* to *wh-* form is a widely known common morphological linguistic phenomenon applying to other lexical elements during the ME period. The *hwile* form is shown in high frequency from the early OE until ME3 where the

use drops dramatically in ME2 until it comes to a loss in ME3. The form does not appeared until the end of EModE. In the following table, the frequency of the two forms *while* and *hwile* are displayed in numbers by period.
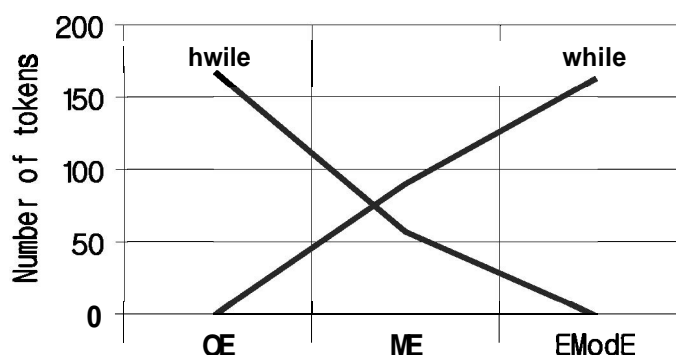
<Table 5> Frequency of the form hwile and while out of total
number by period (tokens in parentheses)

	OE	ME	EModE
hwile	4.01 (166)	0.94 (57)	0.00 (0)
while	0.02 (1)	1.48 (90)	2.94 (162)
Total	4.03 (177)	2.42 (147)	2.94 (162)

(Frequency: out of 10,000 words)

The present form *while* does not appear until the last OE period where only one example was found in the corpus data. Even in this case, the while form was not a lexical element used independently as a nominal but was used to construct the adverbial phrase 'tha while the' as mention in the previous table. The frequency of the form *while* goes up in ME with a dramatic increase in ME4 and a on-going steady number of tokens is displayed in EModE shown in table 5.

<Figure 1> Gradual transition from one pattern to another

The correlation between the two forms, *hwile* and *while,* is clearly illustrated in Figure1. The frequency of the use of *hwile* and *while* is inversely related as shown by the graph. The use of the *hwile* form steadily makes a steep rise until in EModE period, the frequency of the token comes to an end with the loss of the form whereas the *while* form starts from OE and increases until the end of EModE.

A possible assumption relevant to the inverse usage of the two forms is that the grammaticalization process must have given a morphological form shift despite the functional change. The drastic increase of *while* in ME can be said to be the intermediate stage where the lexical form parts from the adverbial phrase construction and becomes an independent linguistic element with semantic reinforcement.

As the grammatical function of *while* was changing in this period throughout EModE, the functional load is understood to have been great as an optional similar form was in need to establish the conjunction or connective function meaning of *while.* After the establishment of the form while, it seemed to rise to the full functional meaning of conjunction. Later on, having secondary meanings added, the temporal meaning shifts to a more abstract concessive conjunctive functional use in EModE.

Focusing on the subjectification of the conjunction *while,* additional examples can be found in the OED offering many definitions including the later developed functional meanings which can be inferred to have the causative function as the following definition describes its use:

> At the same time that (implying opposition or contrast); adversatively, when on the contrary or on the other hand, whereas; concessively, it being granted that; sometimes nearly = although. [OED 2b]

The sentences from (4a) to (4f) are also examples from the OED displaying the meaningful use of concessive conjunction *while.*

(5) a. Painefully to poare vpon a Booke, To seeke the light of truth, while truth the **while** Doth falsely blinde the eye-sight of his looke. (1588 Shakes. L.L.L. i. i. 74)

b.   While they deny a Deity, they assert other things on far less reason. (1719 Watts Ps. xxiii,)

c. There would I find a settled Rest, (While others go and come). (1662 Stillingfl. Orig. Sacrae iii. i.)

d.   White is vulgarly thought to be the most uncompounded of all Colours, while yet it really arises from a certain Proportion of the Seven primary Colours. (1749 Hartley Observ. Man i. i. 75)

e. While the object of the people was to free themselves from the yoke, the object of the nobles was merely to find new sources of excitement. (1857 Buckle Civiliz. I. x. 608)

f.   In rude and unsettled states of society men respect forms and obey facts, while careless of rules and principles. (1864 Bryce Holy Rom. Emp. v. (1875) 52)

g.   While regretting the sorrow which had fallen upon him, Miss Cuthbert was nevertheless glad that her brother was free. (1908 R. Bagot **A.** Cuthbert vi. **51)**

However, it is noteworthy that the causal inference of meaning with the conjunction **while** did not become an objective exemplary criterion of sense in English as full semantic changes of the similar kind are examined in the following examples from Traugott **&** Kö ɪg (1981: 197).

(6) OHG *dia wila so* 'so long as' > German *weil* because'
Latin *dum* 'when, as long as, because'
French *quand* 'when, because'
Finnish *kun* 'when, while, as, since, because'
Estonian *kuna* 'while, as, since, because'

4. Conclusion

This paper investigates the grammaticalization theory and use a corpus-based approach to review the previous studies concerning the development of the conjunction while. As the grammaticalization process in a unidirectional one from a lexical item to a function

element, the same can be said for the conjunctions. Moreover, subjectification, a new aspect of grammaticalization, can be well applied to the phrasal construction breaking into Iexical items and eventually giving semantic abstractedness to the meaning level.

This paper provides evidence from the Helsinki Corpus and Oxford English Dictionary (OED) that the adverbial phrase *Þ hwile Þ* began to lose the cluster effect as the *hwile* form with the meaning 'time' survives independently until ME period. Later on, the form *hwile* undergoes higher functional load as it is utilized as a connective or a conjunction. Eventually, the meaning of temporal simultaneity additionally shifts to successive connective meaning as 'in adverse'. The inversely related use of *hwile* and *while* suggests that as the grammaticalization process undergoes, the morphological form also changes with new functional meanings, the concessive, additionally included.

References

Hopper, Paul and Traugott, Elizabeth. 2003. *Grammaticalization.* Cambridge Press, Cambridge.

Kurylowicz, Jerzy. **1965.** The evolution of grammatical categories. Diogenes. **51:** 55-71.

Rissanen, Matt. **1997.** Grammaticalization from side to side On the development of beside(s). *Corpus Approaches to Grammaticalization in English.* Amsterdam; Philadelphia: John Benjamins. Studies in Corpus Linguistics, **13.** Ljung, M. (ed.). **151-170**

Traugott, Elizabeth. **1995.** Subjectification in Gramrnaticalization. In: Dieter Stein and Susan Wright, (eds), *Subjectivitiy and Subjectivisation. Cambridge: Cambrdige* University Press, **37054**

Traugott, Elizabeth Closs and Bemd Heine. 1991. "Introduction". *In Approaches to Grammaticalization.* Traugott, Elizabeth Closs and Bemd Heine (eds.), John Benjamins B.V.

Helsinki Corpus
Oxford English Dictionary (2nd edition)

Binna Lee
bin4you8@snu.ac.kr