

A qualitative and quantitative analysis of the discourse marker *ati/eti* in Swahili

A corpus-driven approach

Word count: 28 931

Marijke Driesen

Student number: 01407431

Supervisors: Prof. Dr. Koen Bostoën, Prof. Dr. Maud Devos

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of
Master of Arts in African Studies

Academic year: 2018 - 2019

Acknowledgements

Preparing and writing a dissertation is a process in which several people are involved. In the past few months, I received a great deal of support and assistance.

My supervisor Prof. Dr. Koen Bostoën guided me through every stage of the process. I would like to thank him for his insights and valuable feedback on my work. I also thank my co-supervisor Prof. Dr. Maud Devos. She initiated this research by suggesting the subject of this dissertation and provided input and feedback.

I wish to thank Prof. Dr. Gilles-Maurice de Schryver as well. He gave me access to the Swahili corpus and got me started with the research by teaching me how to work with WordSmith Tools.

Lastly, I want to thank Oliva Butoyi, who helped me whenever I had doubts about a Swahili translation and whose enthusiasm kept me motivated.

Table of contents

Acknowledgements	i
List of abbreviations	iv
List of figures	vi
1 Introduction	1
2 Methodology	4
3 Description of the corpus	6
4 Literature review	11
4.1 Background of Swahili	11
4.2 Discourse markers - general introduction	12
4.2.1 Former research	12
4.2.2 Definitions	14
4.3 Discourse markers in Swahili	17
4.4 <i>Ati</i> and <i>eti</i> in the literature	20
4.4.1 Meaning	20
4.4.2 <i>Ati</i> and <i>eti</i> as (in)direct speech indicators	25
4.4.3 Origin	27
5 <i>Ati</i> and <i>eti</i> : a quantitative analysis	32
5.1 Frequency	32
5.2 Collocations	35
6 The pragmatic uses of <i>ati</i> and <i>eti</i>	38
6.1 Sentential complementation	38
6.1.1 Indirect speech	41
6.1.2 Direct Speech	49
6.2 Clause linkage	53
6.3 Interjection	59
6.3.1 In general	59
6.3.2 Conative interjection	61
6.3.2.1 Attention-getting device	61
6.3.3 Expressive interjection	65
6.3.3.1 Contempt	70
6.3.3.2 Surprise	72
6.3.3.3 Doubt	73

6.3.3.4	Other emotions	74
6.3.4	Emphasis	75
6.4	Connotation	77
7	Syntactic aspects of <i>ati</i> and <i>eti</i>	81
8	Pragmatic distinctions between <i>ati</i> and <i>eti</i> ?	84
9	Diachronic scenario for the development of the polyfunctionality of <i>ati</i> and <i>eti</i>	87
10	Conclusion.....	92
	References	95

List of abbreviations

1SG	first person singular
1PL	first person plural
2SG	second person singular
2PL	second person plural
APPL	applicative
CAUS	causative
COMP	complementizer
COND	conditional
CONN	connective
CONS	consecutive
COP	copula
DEM	demonstrative
DIST	distal
FUT	future
FV	final vowel
GEN	generic
HAB	habitual
LOC	locative
NEG	negative
NP ₁	noun prefix of class 1
OP ₂	object prefix of class 2
PART	particle
PASS	passive
PL	plural
POSS	possessive
PP ₃	pronominal prefix of class 3
PRF	perfective
PROG	progressive
PROX	proximal
PRS	present
PST	past

RC₄ referential concord of class 4
RECP reciprocal
REF referential
REFL reflexive
RES resultative
SBJV subjunctive
SG singular
SIT situative
SM stem marker
SP₅ subject prefix of class 5

List of figures

Figure 1. Diagram of the historical distribution of tokens in the Swahili literature corpus.	6
Figure 2. Diagram of the regional distribution of tokens in the Swahili literature corpus.	7
Figure 3. Schema for the grammaticalization relationship between the major functional domains of Shona <i>ti</i> (Güldemann 2002, 282).	29
Figure 4. Number of attestations of <i>ati</i> and <i>eti</i> per 10 000 tokens.	33
Figure 5. Comparison of the <i>ati/eti</i> ratio, according to region and period.	34
Figure 6. Number of <i>ati/eti</i> attestations in the subcategories of the corpus that are used as a complementizer, per million tokens.	41
Figure 7. Number of <i>ati/eti</i> attestations in the subcategories of the corpus that are used as an indirect speech indicator, per million tokens.	49
Figure 8. Number of <i>ati/eti</i> attestations in the subcategories of the corpus that are used as a direct speech indicator, per million tokens.	53
Figure 9. Number of <i>ati/eti</i> attestations in the subcategories of the corpus that are used as an adverbial subordinating conjunction, per million tokens.	59
Figure 10. Number of <i>ati/eti</i> attestations in the subcategories of the corpus that are used as an attention-getting device, per million tokens.	65
Figure 11. Number of <i>ati/eti</i> attestations in the subcategories of the corpus that are used as an expressive interjection, per million tokens.	69
Figure 12. Distribution in the corpus of the emotions conveyed by <i>ati</i> and <i>eti</i> , when employed as expressive interjections.	70
Figure 13. Number of <i>ati/eti</i> attestations in the subcategories of the corpus that are used for emphasis, per million tokens.	77
Figure 14. Number of <i>ati/eti</i> attestations in the subcategories of the corpus that carry an additional emotive connotation (expressive interjections excluded), per million tokens.	80
Figure 15. Percentage of the <i>ati</i> and <i>eti</i> attestations in the corpus that are used for emphasis, as an expressive interjection, attention-getting device, adverbial subordinator or complementizer.	85
Figure 16. Diachronic development of the polyfunctionality of Swahili <i>ati</i> and <i>eti</i>	89
Figure 17. Distribution of the functions of <i>ati/eti</i> in the subcorpora of the early (19 th and early 20 th century) and recent period (1970s until 2010s).	90

Figure 18. Number of *ati/eti* attestations in the subcategories of the corpus that function as an expressive interjection, complementizer, emphasizer, attention-getting device and adverbial subordinator, per million tokens..... 93

1 Introduction

The empirical focus of this dissertation is the particle *ati*, or its variant *eti*, that is used as a discourse marker in Swahili, a Bantu language spoken in Eastern Africa. Discourse markers are those little words or expressions, such as *oh*, *well*, *and*, *now*, *right* and *y'know*, that connect and organise units of talk (Schiffrin 1987, 31). Until now there is no consensus on a definition, nor on terminology. The term discourse marker has different meanings for various researchers and there is a variety of labels used to nominate this category of particles. Alternative names for discourse markers are discourse particles, modal particles, downtoners or pragmatic markers. In this paper, I will continue to refer to these particles as 'discourse markers', because it is the most frequently used term and it seems to be the one with the widest acceptance and the broadest range of application.

More often than not, discourse markers do not have one clear indisputable meaning, but can be used for a variety of functions. This is also true for Swahili *ati/eti*, which has a number of distinct uses, ranging from an interjection to express contempt, over adverbial subordinator, to introducing reported speech. In the dictionaries, *ati* and *eti* are often translated as '(I) say'.

In the history of linguistics, discourse markers have not received a great deal of attention. This changed in the past thirty years, which have seen a rapid development of the study of discourse markers, especially following the publication of Schiffrin's (1987) *Discourse Markers*. However, this relatively recent evolution has not reached the study of African languages, for which analyses of discourse markers are rather scarce. This is also the case for Swahili. Although Swahili is one of the more thoroughly researched African languages, its discourse markers, and *ati* and *eti* in particular, have barely been looked into.

Looking at the available Swahili grammars and dictionaries, one finds limited information on *ati* and *eti*. The little grammatical information that is available only covers reported speech (Massamba 1986), while the available dictionary information focuses on the interjectional properties of the discourse marker. This information is not at all comprehensive, as will be demonstrated in this dissertation.

This paper seeks to show that *ati* and *eti* have a much broader scope of meanings and functions than usually described. It will provide a detailed analysis of all the different contexts in which *ati* and *eti* are used. In doing so, I will take into account the regional and historical distribution of the functions, as well as the syntactic aspects of the discourse marker and a potential distinction between *ati* and *eti*. The diachronic development of the various functions of *ati* and *eti* will also be discussed. I will demonstrate that *ati/eti* has a number of different pragmatic uses that can be classified into two categories. Its first function is that of a subordinating conjunction. *Ati* and *eti* can be used as a complementizer, in a reported speech construction or otherwise, and it can function as an adverbial subordinator, to indicate an interclausal relation of either cause 'because', or purpose 'in order to', 'so that'. The second category is *ati/eti* as an interjection. The particle can be used as a conative interjection to get the hearer's attention, or as an expressive interjection to convey emotions like contempt, surprise or doubt, or to emphasize the discourse prior to it.

The findings in the current paper are based on a corpus of written data, mainly consisting of literary genres such as novels, theatre and poetry. A combination of quantitative and qualitative approaches was used in the data analysis. The quantitative aspect deals with the frequency of *ati* and *eti*, while the qualitative part investigates the various pragmatic functions of the discourse marker, analysing each of its meanings in detail.

The findings should make an important contribution to the field of discourse marker research for African languages, as well as to the field of Bantu corpus linguistics. This research is the first to extensively examine the Swahili particle *ati/eti*, both its quantitative and its qualitative aspects. It provides new insights into the linguistic properties of discourse markers in Swahili, thus filling a knowledge gap in the study of discourse markers in African languages. This work also shows how quantitative and qualitative analyses of large corpora allow us to describe many more meanings and functions than those that are usually recorded in dictionaries.

This paper begins by describing the methodology of this study and the corpus on which this study is based. It will then go on to discuss former research on discourse markers in general, Swahili discourse markers and *ati* and *eti* in particular. In this chapter, special attention will be paid to the etymological origin of *ati* and *eti*. This is followed by a quantitative analysis of *ati* and *eti* in Chapter 5, examining the frequency of the discourse marker across different

regions and time periods. The next part consists of a qualitative analysis, in which I provide a detailed description of all the different pragmatic uses of *ati* and *eti*. In these sections I will also consider the regional and historical distribution of the various functions: when and where are each of the specific functions most frequently used? Next, there will be a short chapter that discusses the syntactic aspects of *ati/eti*: does the intrasentential distribution of the discourse marker have an impact on its meaning? Chapter 8 examines whether there is a difference between the two variations of the discourse marker, *ati* and *eti*. Lastly, I will sketch the diachronic path of development towards the polyfunctionality of *ati* and *eti*.

2 Methodology

The current study is corpus-driven. A major advantage of corpus linguistics is that it allows to examine natural, 'real-life' language. It is a purely descriptive (as opposed to prescriptive) and empirical approach to analyzing language. Traditionally, *ati/eti* has been examined without corpus evidence, which led to incomplete descriptions of the discourse marker.

The corpus-driven approach should not be confused with a corpus-based approach. The distinction between these two methods was introduced by Tognini-Bonelli (2001, 84–85), who defined corpus-based studies as using "*corpus data in order to explore a theory or hypothesis, aiming to validate it, refute it or refine it.*" Corpus-driven linguistics, on the other hand, "*claims instead that the corpus itself should be the sole source of our hypotheses about language.*" It is the latter that is used in this research, allowing to look at the corpus evidence with an open mind, without anticipating the results.

The main tool for carrying out investigations on the corpus is a software called WordSmith Tools 6. This tool makes it possible to search for words in large amounts of texts and to create a wordlist of all the types that are used in the corpus (WordList). It allows then to display every attestation of a selected type in their surrounding context (Concordance). WordSmith Tools 6 also calculates quantitative data about the corpus.

In order to identify all the attestations of *ati* and *eti* in the corpus, I first of all needed to be sure of its orthography. Are they always written disjunctively or do they occur conjunctively as well? Does an alternative spelling exist? To answer these questions I queried **ati, ati**, *eti** and *eti**, as well as some alternative spellings, such as *atii/etii*, *atie/etie*, *hati/heti* and *atia/etia*. However, these queries did not produce any relevant results. Now knowing the exact orthography of the discourse marker, I searched the corpus for all the attestations of *ati* and *eti*, resulting in 331 and 66 hits respectively. After having excluded the irrelevant hits, 376 hits remained (310 for *ati* and 66 for *eti*).

A close examination of the remaining hits was then carried out using Concordance, taking into account the historical and geographical distribution of the texts, *ati/eti*'s collocates, the intrasentential distribution and, most importantly, the pragmatic use of the discourse marker.

When the pragmatic use of *ati/eti* was not clear to me, I received help from a Tanzanian native speaker of Swahili living in Belgium, Oliva Butoyi, who provided clarification.

The fact that I am not a native speaker of Swahili, is a source of uncertainty in this study. After all, I as a non-native speaker, am not fully capable to grasp all the layers of the Swahili language and thus might have missed some meanings or misinterpreted others. In addition, even for native speakers it can be hard to judge what the exact meaning is of *ati/eti* in a particular context (especially when it concerns connotation, cf. Section 6.4). Firstly because one needs a very good contextual knowledge to comprehend all the nuances of the discourse marker in its context. Secondly because of the subjectivity of the individual speaker: their interpretation might be influenced by their own personal views, which might differ from other native speakers' interpretations.

3 Description of the corpus

The corpus used in this study is part of a larger Swahili corpus created by Prof. Dr. Gilles-Maurice de Schryver. The part used here is compiled by Prof. Dr. Maud Devos. It is composed of 53 literary texts written in Swahili between 1895 and 2011 (a list of which can be found at the end of this chapter). Among the texts are mainly novels, theatre and poetry, and to a lesser extent traditional stories, autobiographies and religious texts. The corpus consists of 2 186 147 tokens and 177 779 types. The lexical variation of the corpus is rather high, with a standardized type/token ratio (STTR) of 49,74 %.

The texts in the corpus can be divided into three categories, according to the period they were written in: early, middle and recent. All texts from the nineteenth and the early twentieth century are categorized as early. The middle period contains the texts from the early twentieth century until the 1960s. Finally, all texts written from the 1970s onwards are considered recent. The biggest part of the corpus consists of recent texts: 1 073 088 tokens or 49,09 %. The proportion of tokens from the early period is a little lower, with 953 910 tokens or 43,63 %. The middle period is absolutely the smallest part of the corpus and consists of only 159 149 tokens, which is equivalent to 7,28 % of the corpus.

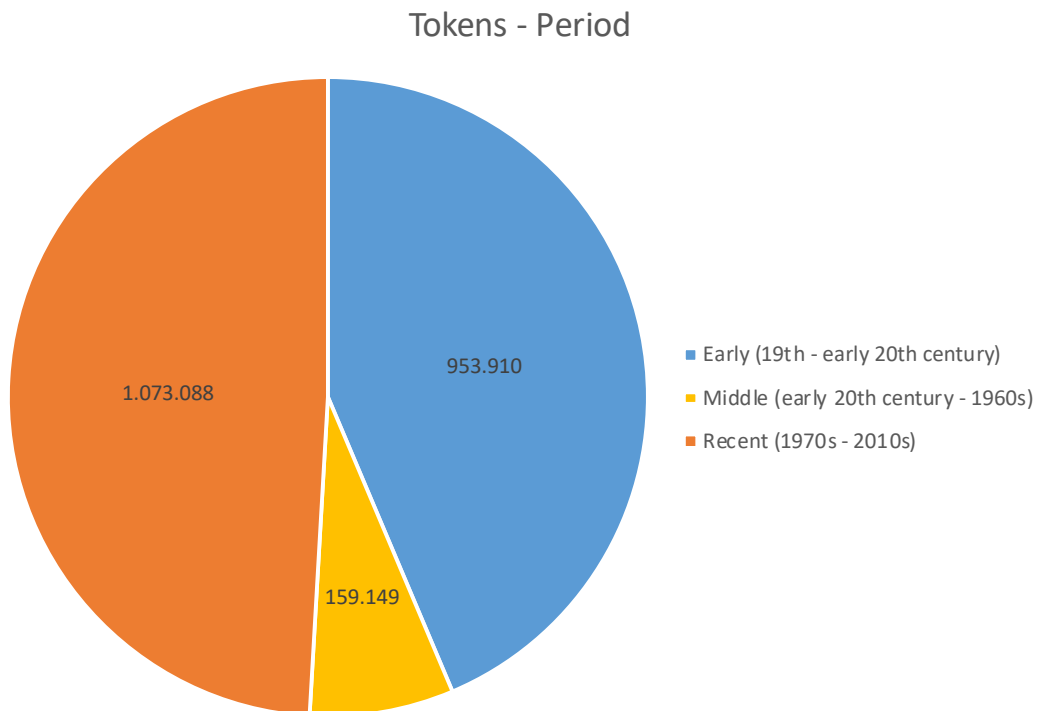


Figure 1. Diagram of the historical distribution of tokens in the Swahili literature corpus.

Apart from historical distribution, the texts can also be classified according to their geographical distribution. They can be categorized by country, Tanzania or Kenya, or by region, inland or coast. With these two factors, four distinct geographical categories can be created: Tanzania inland, Tanzania coast, Kenya inland and Kenya coast. Tanzania is the country that is most represented in the corpus, with a total of 1 187 186 tokens, which is equivalent to 54,30 % of the total corpus (cf. all the orange in Figure 2). Among the Tanzanian texts, 979 402 tokens are from the coast (44,80 % of the total corpus) and 207 784 tokens are from inland Tanzania (9,50 % of the total corpus). Kenya on the other hand is less represented, with 713 394 tokens or 32,63 % of the tokens in the total corpus (cf. all the blue in Figure 2). The texts from Kenya coast contain 376 687 tokens and Kenya inland 320 071 (respectively 17,23 % and 14,64 % of the total corpus). This means that the coastal regions are better represented than the inland regions, with each 62,03 % and 24,15 %. The geographical origin of the remaining 13,06 % is unknown.

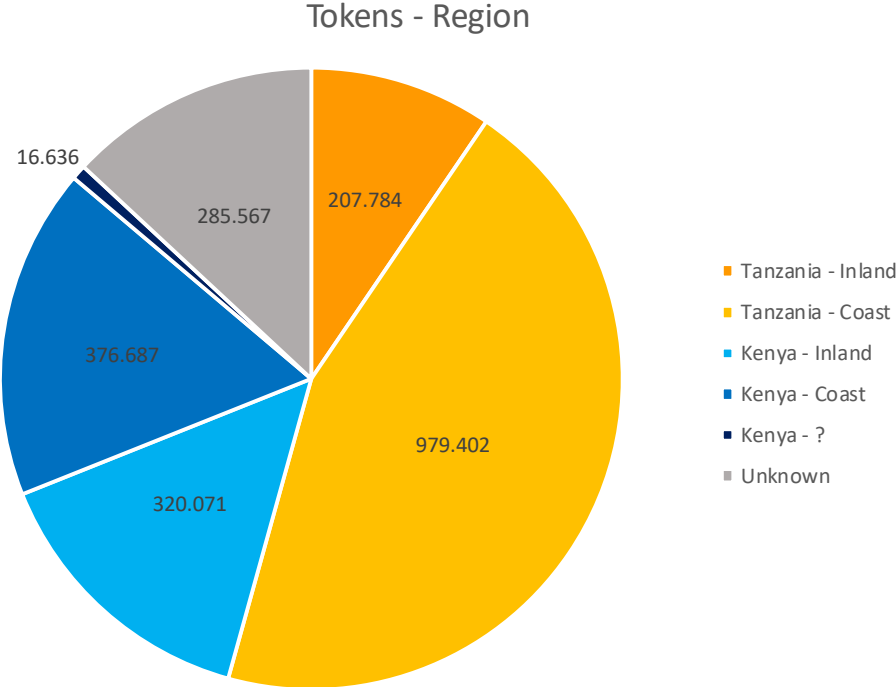


Figure 2. Diagram of the regional distribution of tokens in the Swahili literature corpus.

The corpus that was used in this study is relatively well balanced, both in terms of geographical and historical distribution. There are slightly more data available for Tanzania and the coastal regions. The only significant gap is found in the historical distribution of the

corpus. The middle period is considerably underrepresented, as data from the first half of the twentieth century compose only 7,28 % of the corpus.

Furthermore, the corpus consists only of written data. Oral data are not part of it. Considering discourse markers appear mostly in spoken language (Bayer & Struckmeier 2017, 3), I deliberately selected a literary corpus because the texts include a lot of dialogue. In novels and theatre, it is often envisaged to make the dialogue resemble authentic spoken language as much as possible. By using these specific literary genres, I tried to circumvent the problem of the absence of spoken data. Although the inclusion of oral data in the corpus would still be preferred, its absence is thus no longer an insurmountable obstacle for this study.

Content of the corpus

- Abdallah, bin Hemedi'lAjjemy. 1895. *Habari za Wakilindi*. Dar es Salaam; Nairobi; Kampala: East African Literature Bureau.
- Abdulla, Muhammed Said. 1973. *Duniani kuna Watu*. Nairobi: Heinemann Kenya.
- Bacuez, Pascal. 2000. *Contes Swahili de Kilwa/Hadithi za Kiswahili kutoka Kilwa*. Paris: L'Harmattan.
- Bacuez, Pascal. 2001. *Les ruses de la malice. Contes Swahili*. Paris: L'Harmattan.
- Burhani, Z. 1981. *Mali ya Maskini*. Nairobi: Longhorn Kenya.
- Chachage, Chachage Seithy L. 1991. *Almasi za Bandia*. Nairobi: Vide-Muwe Publishers.
- Chachage, Chachage Seithy L. 2005. *Makuadi wa Soko Huria*. Dar es Salaam: E & D Limited.
- Hichens, W. 1940. *Diwani*. Johannesburg: University of the Witwatersrand Press.
- Hussein, Ebrahim. 1971. *Mashetani*. Nairobi; Dar es Salaam: Oxford University Press.
- Hussein, Ebrahim. 1988. *Kwenye ukingo wa Thim*. Nairobi: Oxford University Press.
- Johnson, Frederick. 1930. *Zamani mpaka Siku Hizi*. London: The Sheldon Press.
- Kezilahabi, Euphrase. 1971. *Rosa Mistika*. Dar es Salaam: Dar es Salaam University Press.
- Kezilahabi, Euphrase. 1990. *Nagona*. Nairobi: Vide-Muwe Publishers.
- Kezilahabi, Euphrase. 1999. *Kaptula la Marx*. Nairobi: Vide-Muwe Publishers.
- Kezilahabi, Euphrase. 2008. *Dhifa*. Nairobi: Vide-Muwe Publishers.
- Kibao, Salim A. 1975. *Matatu ya Thamani*. Nairobi: Heinemann Educational Books.
- Kitsao, Jay. 1983. *Bibi Arusi*. Nairobi: Oxford University Press.

- Mbatiah, Mwenda. 1999. *Upotevu*. Nairobi: Phoenix Publishers.
- Mbega, Hassani Mwalimu. 1985. *Utenzi wa Mfalme Muadilifu*. Nairobi: Kenya Literature Bureau.
- Mbega, Hassani Mwalimu. 1986. *Wasia Bora na Hadithi Zake*. Nairobi: Kenya Literature Bureau.
- Mbotela, James. 1934. *Uhuru wa Watumwa*. Nairobi: The Eagle Press.
- Mkangi, Katama. 1995. *Walenisi*. Nairobi: East African Publishers.
- Mohamed, Said Ahmed. 1980. *Utengano*. Nairobi: Longhorn Publishers.
- Mohamed, Said Ahmed. 1990. *Kivuli Kinaishi*. Nairobi: Oxford University Press.
- Mpinga, A.S. 1984. *Cheka Ucheke*. Ndanda: Benedictine Publications.
- Muhando, Pennina. 1973. *Tambueni haki zetu*. Dar es Salaam: Tanzania Publishing House.
- Ngahyoma, Ngalimecha. 1973. *Huka*. Dar es Salaam: Tanzania Publishing House.
- Njama, Ali H. 1995. *Haki Haizami*. Nairobi: Jomo Kenyatta Foundation.
- Ole Kulet, Henry R. 1990. *Maisha ya Hatari*. Nairobi: Longhorn Publishers.
- Robert, Shaaban. 1946. *Kufikirika*. Dar es Salaam: Mkuki na Nyota Publishers.
- Robert, Shaaban. 1949 & 1960. *Maisha Yangu na Baada ya Miaka Hamsini*. Dar es Salaam: Mkuki na Nyota Publishers.
- Robert, Shaaban. 1951. *Kusadikika. Nchi Iliyo Angani*. Dar es Salaam: Mkuki na Nyota Publishers.
- Robert, Shaaban. 1952. *Adili na Nduguze*. London: The Macmillan Press.
- Robert, Shaaban. 1960. *Siku ya Watenzi Wote*. Dar es Salaam: Taasisi ya Uchunguzi wa Kiswahili.
- Robert, Shaaban. 1960. *Wasifu wa Siti binti Saad*. Dar es Salaam: Mkuki na Nyota Publishers.
- Robert, Shaaban. 1962. *Kielezo cha Fasili*. Dar es Salaam: Mkuki na Nyota Publishers.
- Rocha, Chimerah. 1995. *Nyongo Mkalia Ini*. Nairobi: Oxford University Press.
- Sacleux, Charles. 1930. *Ngano za Hadisi ya Watakatifu*. Rome: Sodalité de St Pierre Claver.
- Shafi, Shafi Adam. 1999. *Vuta n'Kuvute*. Dar es Salaam: Mkuki na Nyota Publishers.
- Shafi, Shafi Adam. 2007. *Kasri ya Mwinyi Fuad*. Nairobi: Vide-Muwe Publishers.
- SPCK (1909) *Hiki ndicho chuo cha sala ya umoja, na kusongeza sakaramenti*. London: Society for Promoting Christian Knowledge.

- Steere, Edward & Alice Werner. 1922. *Swahili Tales, as told by Natives of Zanzibar*. London: Society for Promoting Christian Knowledge.
- Suleiman Muhamed, Muhamed. 1976. *Nyota ya Rehema*. Nairobi: Oxford University Press.
- Velten, C. 1901. *Safari za Wasuaheli*. Göttingen: Vandenhoe & Ruprecht.
- Velten, C. 1903. *Desturi za Wasuaheli na khabari za desturi za sheria za Waswaheli*. Göttingen: Vandenhoe & Ruprecht.
- Velten, C. 1907. *Prosa und Poesie der Suaheli*. Berlin: self-published.
- Wa Mberia, Kithaka. 1997. *Natala*. Nairobi: Marimba Publications.
- Walibora, Ken. 1996. *Siku Njema*. Nairobi: Longhorn Publishers.
- Walibora, Ken. 2003. *Kufa Kuzikana*. Nairobi: Longhorn Publishers.
- Wamitila, Kyallo Wado. 2007. *Dharau ya Ini*. Nairobi: Vide-Muwe Publishers.
- Wamitila, Kyallo Wado. 2011. *Msichana wa Mbalamwezi*. Nairobi: Vide-Muwe Publishers.
- Whitely, W. H. 1959. *Maisha ya Hamed bin Muhammed el Murjebi yaani Tippu Tip, kwa maneno yake mwenyewe*. Nairobi: East African Literature Bureau.
- Zani, Zachariah & James Kitsao. 1975. *Mafarakano na michezo mingine*. Nairobi: Heinemann Educational Books.

4 Literature review

4.1 *Background of Swahili*

Swahili, called Kiswahili in the Swahili language itself, is the most well-known language of the Sabaki languages. The Sabaki languages is a label for six closely related languages spoken on the coast of East Africa. They form a subbranch of North East Coast Bantu, which belongs to the Narrow Bantu family of the Niger-Congo phylum (Nurse & Hinnebusch 1993, 4). In Maho's Guthrie list, Swahili is classified as G42 (Maho 2009, 49). Swahili is the most widely spoken Bantu language, with over 60 million speakers. It is an official language in Tanzania and Kenya and is spoken in all of East Africa (Comrie 1989, 883–87).

Swahili was originally only spoken by the Waswahili, an ethnic group living on the East coast of Africa. It consists of a dialect continuum that was derived from Proto-Sabaki and has been in place for over a millennium. This Swahili coastal continuum stretches over a thin 1000-mile-long line of Swahili speaking villages and towns on the coastline between southern Somalia and northern Mozambique and the Comoros. The reason for the vastness of this early expansion of the Swahili continuum is that it was driven by trade using boats. During the same time-period, the latter centuries of the first millennium, trade also motivated the Swahili communities to regularly interact with Arabic-speaking visitors and settlers (Nurse & Hinnebusch 1993, 29–30). Because of this, the Arabic language has had a considerable influence on Swahili. This influence is mainly visible in the Swahili lexicon, which includes many Arabic loanwords. Also, Swahili was originally written in a modified form of the Arabic script (Comrie 1989, 886–87).

In the nineteenth century, Swahili spread inland when Swahili-speaking traders travelled westwards with their caravans. Swahili was used as a lingua franca on these trade routes and became a widespread vehicular language in Uganda, Rwanda, Burundi and the eastern parts of the Democratic Republic of Congo. This gave rise to the development of various regiolects due to the influence of local languages. The most notable of these regiolects are those spoken in East Congo and are known collectively as Kingwana (Whiteley 1969, 4–5).

Following the arrival of European missionaries and the colonial administration, the Roman alphabet was introduced as the new script of Swahili (which is also how the standard variety

of Swahili is nowadays written) (Comrie 1989, 886–87). During colonialism, the German and the British chose Swahili as the language for administration and education in Tanzania and Kenya, which solidified the status of Swahili as a lingua franca in this area. In 1930, the British set up an official institute for the standardization of Swahili. When the British established a standard variety of Swahili, Kiswahili Sanifu, they based it on Kiunguja, one of the varieties spoken on Zanzibar (Blommaert 1997, 74–77).

Swahili is a well documented and well described language. Nurse & Hinnebusch (1993, 21) write that the oldest preserved sources of Swahili trace back to approximately 1700. The content of these documents suggests that there had been Swahili documents before this date, but none of these older sources have been found yet. There is a long tradition of poetry and literature in Swahili. Scholarly interest started relatively early and a lot of linguistic research has been conducted. Among the major Swahili linguistic works are the grammars of Ethel O. Ashton (1944), Alfons Loogman (1965) and Edgar C. Polomé (1967).

4.2 Discourse markers - general introduction

4.2.1 Former research

Discourse markers, small words such as *well*, *you know*, *I mean* and *sort of* in English, have received relatively little attention in linguistics and the study of this particular subject has just recently developed. According to Aijmer & Simon-Vandenberg (2011, 223–24), who have provided an overview of the history of this field, interest in discourse markers did not start until the late 1960s and early 1970s. Prior to this, there had been a few works on the subject, such as *The Greek Particles* (Denniston 1934) and *Modal Particles in Russian and German* (Arndt 1960), but these examples remained exceptional.

It was not until 1969 that the real break-through in the studies of discourse markers came, with the publication of *Abtönungspartikel: Die deutschen Modalwörter und ihre französischen Entsprechungen* (Weydt 1969). In fact, most of the early works on discourse markers were focused on the German language. From there, the study of discourse markers exploded and expanded to other languages. Important works include Schiffrin (1987), Fraser (1990, 1996, 1999, 2006), Jucker & Ziv (1998) and Fischer (2006).

Nowadays, research on discourse markers is expanding in various directions. They are studied diachronically, with a focus on the origin of discourse markers and how they evolved through time. They are also studied synchronically, often in combination with a cross-linguistic analysis, as discourse markers are found in many languages. Finally, there is some interest in the sociolinguistic aspect of discourse markers.

Most of the scholarly research on discourse markers is focused on European languages. The following list, compiled by Maschler & Schiffrin (2015, 189) provides some examples of research on markers in several languages. Other than English, markers have been studied in German (Abraham 1991; Barske & Golato 2010; Golato 2010, 2016; Imo 2010), French (Cadiot et al. 1985; Degand & Fagard 2011; Hansen 1998), Finnish (Hakulinen 1998, 2016; Laakso & Sorjonen 2010; Sorjonen 2001), Italian (Bazzanella 1990, 2006; Bruti 1999; Mauri & Ramat 2012; Visconti 2005, 2009), Hungarian (Dér & Markó 2010; Vasko 2000) etcetera. Languages from the rest of the world are to a lesser extent present in the research on discourse markers. Examples of non-European languages that have gotten a lot of attention in the research field of discourse markers are Japanese (Cook 1992; Fuji 2000; Matsumoto 1988; McGloin & Konishi 2010; Onodera 2004), Chinese (Biq 1990; Ljungqvist 2010; Miracle 1989; Wang, Tsai, & Yang 2010), Indonesian (Wouk 1998, 2001), Hebrew (Ariel 1998, 2010; Livnat & Yatsiv 2003; Maschler 1997, 2009, 2012; Miller Shapiro 2012; Ziv 1998, 2008) and sign languages (McKee & Wallingford 2011; Perez Hernandez 2006). Studies on discourse markers in African languages are far less common. In fact, research work on discourse markers in an African language other than Swahili can barely be found. There is a paper on discourse markers in Ewe, a Niger-Congo language, spoken in Ghana (Ameka 1991) and there is Güldemann (2002, 2008) who has done some work on quotative indexes in African languages, which can be seen as a subcategory of discourse markers.

There is also a small number of publications on discourse markers in Swahili. The most notable work here is *The pragmatics of selected discourse markers in Swahili*, written by Andrea S. Dunn (1990). In this PhD dissertation, she lists several Swahili discourse markers (*sasa*, *basi*, *maana*, *haya*, *nanihii* and *kumbe*) and describes them in a pragmatic framework. Nicolle (2000) focuses in particular on the discourse marker *je*. Eastman (1983, 1992) also discusses some discourse markers in two papers that deal with exclamations and interjections in Swahili. These publications will all be discussed in Section 4.3 on discourse markers in Swahili.

As this short overview has shown, interest in the study of discourse markers has only begun in the last decennia. Bayer & Struckmeier (2017, 2–3) list 3 main reasons for the late development of this field. An important reason is that the occurrence of discourse markers differs between languages. In some languages, like German, the use of discourse markers is omnipresent and an important part of linguistic expression. In English however, one of the academically dominant languages, discourse markers are rather absent. Another reason concerns the fact that discourse markers are a prominent feature in spoken language, and occur less in written language. For the field of linguistics is still more focused on written than on spoken language, the study of discourse markers, and other features of spoken language, tend to be neglected. Note that there are some written language styles or genres in which discourse markers are clearly attested, such as drama or prose. A third reason, and according to Bayer & Struckmeier the most important one, is that discourse markers are very complex, because they do not neatly fit into the grammatical categories and framework that linguists are used to work with.

4.2.2 Definitions

More recently, linguists have developed theories and methods to describe the functions and properties of discourse markers. Until now there is not yet a consensus on the terminology, which may lead to confusion as different scholars use different names. Alternative names for discourse markers are discourse particles, modal particles, downtoners or pragmatic markers. In this paper, I will continue to refer to these particles as 'discourse markers', because it is the most frequently used term and it seems to be the one with the widest acceptance and the broadest range of application.

Not only is there discussion about what these particles should be called, but there is also little agreement as to which elements should be included in the category of discourse markers. In fact, there is no generally agreed upon definition of the term 'discourse marker'. The following subsection will highlight several definitions and some of the salient features of discourse markers that were put forward by different scholars. I start with the definition formulated by Deborah Schiffrin, who was the first one to come up with the term 'discourse markers' and whose pioneering work *Discourse Markers* (1987) was a significant contribution to this field of research. Schiffrin (1987, 31) defined discourse markers as "*sequentially dependent elements which bracket units of talk*". She describes them as "*non-obligatory utterance-initial*

items that function in relation to ongoing talk and text" (Maschler & Schiffrin 2015, 191). This means that discourse markers are deictic and that they have an indexical function, namely to point to context features. According to Schiffrin (1987), this context referred to by discourse markers operates on different levels: the ideational structure (ideas and propositions), action structure (the way in which speech acts relate to preceding, following or intended actions), exchange structure (turns), information state (management of knowledge and meta-knowledge) and the participant framework (speaker-hearer relationship) (Aijmer & Simon-Vandenberg 2011, 224–25). Schiffrin further indicates that discourse markers can belong to different word classes, such as conjunctions (e.g. *but, or, and*), interjections (e.g. *oh*), adverbs (e.g. *now, then*) and even lexicalised phrases (e.g. *you know, I mean*) and non-verbal expressions (Maschler & Schiffrin 2015, 191).

Fraser considers discourse markers to be more constrained. According to him and contrary to Schiffrin's definition, interjections and non-verbal expressions are not to be categorized as discourse markers. Fraser (1999, 931) defines discourse markers as follows:

"After reviewing prior theoretical research, I define discourse markers as a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs, and prepositional phrases. With certain exceptions, they signal a relationship between the interpretation of the segment they introduce, S2, and the prior segment, S1. They have a core meaning, which is procedural, not conceptual, and their more specific interpretation is 'negotiated' by the context, both linguistic and conceptual."

Another definition is given by Maschler, who puts the emphasis of discourse markers on the fact that they give a metalingual interpretation. According to her, all discourse markers refer to the linguistic interaction itself, rather than to the extralingual realm. Discourse markers refer to the realm of the text, the interpersonal relations between the participants of the conversation, and to their cognitive processes. In Maschler's definition, for an utterance to be considered a discourse marker, it must fulfil two requirements: a semantic and a structural one. Semantically, the particle must have a metalingual interpretation, as already mentioned. Structurally, "*the utterance must occur at intonation-unit [Chafe 1994] initial position, either at a point of speaker change, or, in same-speaker talk, immediately following any intonation contour other than continuing intonation*" (Maschler & Schiffrin 2015, 194–97).

Instead of formulating a definition, other scholars have attempted to define the term 'discourse markers' by giving a list of characteristics that are typical for these particles. Hölker (1991 as cited in Jucker & Ziv 1998, 3) lists four basic features of what he calls pragmatic markers: *"(1) they do not affect the truth conditions of an utterance; (2) they do not add anything to the propositional content of an utterance; (3) they are related to the speech situation and not to the situation talked about; and (4) they have an emotive, expressive function rather than a referential, denotative, or cognitive function"*. Rühlemann (2007, as cited in Aijmer & Simon-Vandenberg 2011, 225) mentions five characteristics: *"(1) they indicate how discourse relates to other discourse; (2) they do meta-lingual work; (3) they are discourse deictic and indicate how the utterance containing them is a response to preceding discourse; (4) they create discourse coherence and (5) they are oriented to the hearer's needs"*. Another, much longer list of characteristics is provided by Brinton (1996 as cited in Jucker & Ziv 1998, 3).

- Phonological and lexical features:
 - a) They are short and phonologically reduced.
 - b) They form a separate tone group.
 - c) They are marginal forms and hence difficult to place within a traditional word class.
- Syntactic features:
 - d) They are restricted to sentence-initial position.
 - e) They occur outside the syntactic structure or they are only loosely attached to it.
 - f) They are optional.
- Semantic feature:
 - g) They have little or no propositional meaning.
- Functional feature:
 - h) They are multifunctional, operating on several linguistic levels simultaneously.
- Sociolinguistic and stylistic features:
 - i) They are a feature of oral rather than written discourse and are associated with informality.
 - j) They appear with high frequency.
 - k) They are stylistically stigmatised.
 - l) They are gender specific and more typical of women's speech.

Evaluating the many definitions and characteristics of discourse markers mentioned in this subsection, it is clear that there is not yet a scholarly consensus on the exact meaning of the

term. Discourse markers are indeed hard to describe, precisely because they are very polyvalent and subjective. Though many of the features listed here differ from one scholar to another, there is also a great deal of overlap. In spite of the absence of one clear definition, the overview of possible uses of the term 'discourse markers' in the preceding paragraphs, gives a good idea of the content of this category.

4.3 *Discourse markers in Swahili*

As already indicated in the previous section, not much research has been done on discourse markers in Swahili. Recent studies are even harder to find, as most work on discourse markers in Swahili dates twenty to thirty years back (Dunn 1990; Eastman 1983, 1992; Eastman & Omar 1985; Nicolle 2000).

In her doctoral dissertation, Dunn (1990) analyses the pragmatic use and meaning of a number of Swahili discourse markers. She examines three discourse markers in detail: *sasa*, *basi* and *maana*. Another three discourse markers are described less elaborately: *nanii*, *haya* and *kumbe*. *Ati* and *eti* are not mentioned in her work.

Dunn (1990, 2–9) describes discourse markers as words used by the speaker to indicate how the hearer should interpret (part of) a sentence in relation to the rest of the discourse. This process is largely subconscious, both for the speaker and the hearer. Dunn assumes that every discourse marker has a basic sense, which most of the time corresponds with its entry in the dictionary. All the various uses of the discourse marker can be traced back to this basic sense. The hearer uses the basic sense of the discourse marker, together with the content of the utterance and the discourse up to that point, to interpret the intentions of the speaker.

Dunn (1990) uses a corpus-driven approach. Her findings are based on her own corpus (29 644 tokens) that mainly consists of personal experience narratives and topical conversations, recorded in Dar Es Salaam in 1986 and 1987. Additionally she also takes examples from '*Wakati Ukuta*' (1971), a play by the Tanzanian author Ebrahim Hussein (Dunn 1990, 35). By examining how the discourse markers are used in this corpus, Dunn tries to state a basic sense for each discourse marker. Using this basic meaning as a starting point, she examines the other possible ways in which the discourse markers can be interpreted, all related to the proposed basic sense. Each interpretation is then provided with examples from the corpus.

The literal translation of *sasa* is 'now'. Much like 'now' in English, *sasa* has both adverbial and discourse related properties. As an adverbial, it is used for temporal reference to indicate present time. As a discourse marker, *sasa* is used when the speaker wants to indicate that the utterance is a subgoal of a broader ongoing goal that is currently relevant in the discourse. This is the basic sense of *sasa* (Dunn 1990, 38–46).

The basic sense of *basi* is 'enough'. *Basi* is a very expressive discourse marker that has a great variety of uses, all of which are dependent on its basic sense of sufficiency. When *basi* is used, the speaker intends the hearer to understand that the idea expressed in the utterance is sufficient for the accomplishment of (one of) the speaker's goal(s). *Basi* can introduce or follow an utterance (Dunn 1990, 76–82).

Maana or *maanake*, which can be translated as 'meaning' or 'its meaning' respectively, is used to introduce a reason for an idea or action that occurred in the preceding discourse. Dunn describes its basic sense as 'considering'. It can also be used for self-correction, to indicate that the speaker made a mistake in their utterance and would like to start over, like 'I mean' in English (Dunn 1990, 125–36).

The three discourse markers that Dunn describes more briefly are *nanii*, *haya* and *kumbe*. *Nanii* is a contraction of *nani* ('who') and *hii* ('this'). This discourse marker is used when the speaker is hesitating about what to say. It can also be used for self-correction when the speaker has misspoken. Dunn (1990, 153–157) compares this discourse marker with 'uh' and 'whatchamacallit' in English. *Haya* is used when the speaker acknowledges or accepts a (verbal or nonverbal) action that has just been performed. This action can be performed by the speaker themselves or by another speaker. *Haya* can be translated by 'ok'. The last discourse marker that Dunn describes is *kumbe*. The use of this discourse marker indicates that the speaker is surprised by the information that *kumbe* refers to (Dunn 1990, 158–64).

Eastman (1983, 1992) and Eastman & Omar (1985) also conducted research on Swahili discourse markers. Their work is about interjections in Swahili, particles that Eastman considers a subcategory of discourse markers. In fact, Eastman refers to discourse markers as 'pragmatic particles', an alternative name for discourse markers that often recurs in the literature. Eastman (1992, 274) describes interjections as elements that bring about stopping to think about the communicative context of the utterance itself and instead allow the speaker

to communicate the extraverbal context, what is going on around the interlocutors in addition to what they are explicitly talking about. In this sense, interjections link verbal and non-verbal behaviour.

Contrary to Dunn, who describes the content of the discourse markers, Eastman does not restrict herself to the text and focuses on the link with the extraverbal reality. For instance, she discusses the gestures that accompany the interjections. Eastman (1992, 282–285) furthermore addresses the sociocultural context of the discourse markers and points out that communities and subcultures select interjections to express their thoughts and feelings in interaction with their social structure (age, class, gender, ethnicity, status, etc.). Examples of interjections described by Eastman (1992, 277–280) include: *ehee* 'I understand', *shabashi* 'bravo, well done', *kelele* to indicate that someone should stop talking, *ng'o* to indicate refusal, *keba* to show rejoicing, *bwana* to call attention, *je* 'well, tell me', etc. *Ati* nor *eti* are mentioned in her work.

Some ten years after the publication of the work of Dunn and Eastman, Nicolle (2000) wrote 'Markers of general interpretive use in Amharic and Swahili'. In this paper, he discusses the use of markers in Amharic and Swahili, two unrelated African languages, in relation to relevance theory. Relevance theory, a theoretical framework first proposed by Sperber & Wilson (1986), states that language can be used descriptively and interpretively. The former, descriptive language use, is the case when an utterance is a representation of a state of affair. In case of the latter, interpretive language use, the utterance is a representation of other representations, like thoughts or other utterances (Nicolle 2000, 173–75).

Nicolle demonstrates the existence of markers of general interpretive use and specific (e.g. interrogative, exclamative) interpretive use. In order to do so, he discusses the Amharic particles *ete* and *inde* and Swahili particle *je*. *ete* is a particle that occurs at the end of interrogative wh-clauses in Amharic. As the use of this particle is not mandatory in questions, it is not a question marker. *ete* rather has the force of an exclamative and expresses surprise. Amharic *inde* is a much more general marker that also has basic interrogative and exclamatory functions. In contrast to *ete*, that marks the speaker's surprise, *inde* can denote a variety of attitudes, including irony or disapproval (Nicolle 2000, 175–79).

As for Swahili *je*, Nicolle (2000, 180–84) considers its specific use to be the marking of an interrogative, but he argues that its functions can be much more general than that. *Je* encodes an instruction to the hearer to treat the utterance interpretively and to seek additional information about it. Furthermore, *je* is often prosodically separated from the following clause in order to distance the speaker psychologically from the interpreted thought expressed by the rest of the utterance. This psychological distance between the speaker and the content of the utterance can then be interpreted by the hearer as surprise, shock or irony.

4.4 *Ati and eti in the literature*

4.4.1 *Meaning*

The most obvious place to start if one wants to know the meaning of a word is the dictionary. The problem with discourse markers however, is that their meaning often is not straightforward. This is also the case with *ati* and *eti*, for which different definitions and translations can be found in various dictionaries. In many dictionaries, *ati* and *eti* are not even included.

The first dictionary I consulted is the 'Dictionnaire swahili-français', one of the most renowned Swahili dictionaries, written by Charles Sacleux (1939). Sacleux writes that *ati/eti* is an expletive used to address one or more interlocutors and can be translated by 'say' or 'hey you'. The following is the full dictionary entry for *ati* and *eti* in Sacleux (1939). It is followed by my own translation.¹

Ati Mot explétif employé en s'adressant à un ou à plusieurs interlocuteurs: Dis donc, dites donc; ô toi! O vous! Hé toi! Hé vous! Ah toi! Ah vous! *Ndžyoo ati* (ou *ati ndžyoo*), viens donc toi. *Ati, una kwēnda kwa miguu?* Dis-donc, vas-tu à pied? *Ati, rafiki, tukisēma hivyo, nani atatuamini?* He! Ami, si nous parlons ainsi, qui nous croira ? *Una ni umiza, ati!* Tu me blesses, toi! *Ati, wewe usēmaye, u mtu gani?* O toi qui parles, qui es-tu? *Sikupata madyibu yake; ila nimēsikia ya kuwa amēsafiri, nae ati amēandika haḡi,* je n'ai pas eu sa réponse; mais j'ai appris qu'il est parti, puis, ô toi! qu'il a écrit une lettre. *Ati wewe mwōngo?* Dis donc es-tu menteur? *Ati, nili kudya,* He toi! J'étais venu. *Nili mwona, ati!* Je l'avais vu, ô toi ! *Nitakupa fūngu lako, ati!* Je te donnerai ta part, ô toi !

¹ All translations in this chapter are my own.

Mkuki huu hauziki, wanifaa ati! Cette lance n'est pas à vendre, elle me sert, ô toi! — R. *Ati* est apparamment la crase de NGATI. (Sacleux 1939, 75)

(**Ati** Expletive word used to address one or more interlocutors: I say; Oh you! Hey you! Ah you! *Ndžyoo ati* (or *ati ndžyoo*), come then you. *Ati, una kwēnda kwa miguu?* I say, are you going on foot? *Ati, rafiki, tukisèma hivyo, nani atatuamini?* Hey! Friend, if we speak like this, who will believe us? *Una niumiza, ati!* You are hurting me, you! *Ati, wewe usèmaye, u mtu gani?* Oh you the one talking, who are you? *Sikupata madyibu yake; ila nimèsikia ya kuwa amèsafiri, nae ati amèandika haḡi,* I didn't get his answer; but I heard that he left, then, oh you! that he wrote a letter. *Ati wewe mwōngo?* Say, are you a liar? *Ati, nili kudya,* Hey you! I came. *Nili mwona, ati!* I saw him, oh you! *Nitakupa fūngu lako, ati!* I will give you your part, oh you! *Mkuki huu hauziki, wanifaa ati!* This spear is not for sale, it is useful to me, oh you! — R. *Ati* is apparently the crasis of NGATI.)

Eti ! = *ati!* (Sacleux 1939, 208)

Therefore, the primary functions of *ati* and *eti* according to Sacleux can be described as follows. Firstly, *ati/eti* is used to call attention. A second function, closely linked to the first one, is the introduction of speech: the speaker points out to the hearer that they are about to say something. These two functions are also reflected in the following definitions of *ati/eti*:

ati! int. (anche: **eti!**) usata per richiamare l'attenzione. *Ati uende!* ti dico di andare. *Ati nisiwasadiki!* dici dunque che non li creda! (Pick 1964, 14)

(**ati!** int. (also: **eti!**) used to draw attention. *Ati uende!* I say go. *Ati nisiwasadiki!* You say then that I should not believe them!)

ati!, eti! interj A. *Anruf*, um Aufmerksamkeit zu erregen "He!", "He, da!"; ~ *wewe!* "He, du!"; B. *Anruf* als Einleitung einer Rede "ich sage", "höre"; ~ ! **yasemwa watu wale ni wageni!** Höre! Man sagt, die Leute sind Fremde (Höftmann & Mhando 1963, 10)

(**ati!, eti!** Interj A. *Call*, to attract attention "Hey!", "Hey there!"; ~ *wewe!* "Hey, you!"; B. *Call*, to introduce speech "I say", "listen"; ~ ! **yasemwa watu wale ni wageni!** Listen! It is said that those people are strangers)

Another dictionary is the monolingual *Kamusi ya Kiswahili sanifu* of the *Taasisi ya Uchunguzi wa Kiswahili*, or institute of Kiswahili Studies (1981). Additional to the one function of *ati/eti* given by Sacleux, namely calling attention, they list two other functions. In the *Kamusi ya Kiswahili Sanifu*, *ati* is said to be used by the speaker for (1) "making somebody listen", (2) "showing surprise about something", and (3) "showing contempt" (Güldemann 2008, 419). The Swahili-English dictionary of Johnson (1971) highlights the same three functions. Although number 3 is not stated explicitly by Johnson, it can be derived from the examples given.

Ati! also **Eti!** A common int. or expletive expressing surprise, or calling attention.

Generally can be translated by 'I say!', as it is from the verb *ti*, say, found in many Bantu languages. *Ati wewe!* I say you! *Niache ati!* leave me alone, I say! *Ati! husema watu ni wapumbavu.* I say! They say that men are fools. (Johnson 1971, 20)

Eti! int. See **Ati!** (Johnson 1971, 86)

Like Sacleux, Johnson translates *ati/eti* by means of the verb 'say'. This translation is provided by most dictionaries. Here are some other examples:

Ati! look you! I say! (Steere & Madan 1894, 251)

No translation provided for *eti* in this dictionary.

eti pia **ati** ki 1 I say. 2 expression of contempt. (TUKI 2001, 68)

ati! ou **eti!** interj. : [interjection] hé toi, dis donc! (Mertens 2006, 16)

(**ati!** or **eti!** interj. : [interjection] hey you, I say!)

eti! ou **ati!** interj.: dis donc! (Mertens 2006, 29)

(**eti!** or **ati!** interj. : I say!)

Even Google Translate, that only gives one single translation of *ati*, translates this word as 'say!'. *Eti* curiously has a completely different translation: 'supposedly' ('Google Translate' n.d.). This translation is reminiscent of another function that is often assigned to *ati/eti*: expressing doubt. *Ati* and *eti*, as described in Krapf's 'A Dictionary of the Suaheli Language', refer to "a matter which one does not know or which one has not witnessed, but only

supposes" (Krapf 1882, 15). According to Krapf, *ati/eti* can best be translated as 'I suppose' or 'I think'. He continues to give an example:

- *Felani yuwápi?* Where is a certain N. N.?

- *Ati, amekuenda Mvita.* (I do not know), I suppose he went to Mombas.

The monolingual Swahili dictionary of Oxford Dictionaries ('Kamusi Ya Kiswahili | Oxford Living Dictionaries' n.d.) also considers expressing doubt to be one of the functions of *ati/eti*. The Swahili Oxford Dictionary gives the following definition of the words *ati* and *eti*.

ati! pia **eti!** *ki* **1** neno la kumfanya mtu asikilize: ~ *Juma!* **2** neno la kuonyesha shaka juu ya jambo fulani: ~ *kesho sikukuu?* **3** neno la kuonyesha dharau: ~ *naye anajifanya ni askari.*

ati! also **eti!** *interjection* **1** word to make someone listen: ~ *Juma!* **2** word to show doubt about a certain issue: ~ *kesho sikukuu?* Is tomorrow a holiday? **3** word to show contempt: : ~ *naye anajifanya ni askari.* And he is pretending to be a policeman.

Online dictionaries list some more functions and translations of *ati/eti*, that are not found in the paper dictionaries that were discussed above. According to some online dictionaries, *ati* and *eti* are used as a request for repetition or as an exclamation implying disagreement. They also translate *eti* (only *eti*, not *ati*) with the conjunction 'as if' (de Schryver & Joffe 2018; 'Swahili-Engels Woordenboek' n.d.).

A completely different use of *ati* (only *ati*, not *eti*) can be found in Kingwana, a variety of Swahili that is spoken in the eastern part of the Democratic Republic of Congo. In this variety, *ati* is used in front of the name of a deceased person. *Ati baba yangu* would be translated in Kingwana as 'my late father'. Kingwana *ati* is probably derived from the Swahili word *hayati*, which also means 'the late' or 'deceased person'.

ATI !, int. Dis donc!, O toi! *Ati wewe!* Hé, toi! Expression exprimant la surprise ou us. pour attirer l'attention. Kgw.: « *ati* » s'emploie devant le nom d'un défunt. (Cf. MAREHEMU) : « *Ati baba yangu* », mon père défunt, feu mon père. (Cf. HAYA !.) (Lenselaer 1983, 34)

(ATI !, int. I say!, Oh you! *Ati wewe!* Hey, you! Expression to express surprise or used to attract attention. Kgw.: « *ati* » is used in front of the name of a deceased person. (Cf. MAREHEMU): « *Ati baba yangu* », my late father. (Cf. HAYA !.))

This usage of *ati* is also described for Kingwana in Sacleux:

Ati (Ngw.). Feu-e, suivi du nom de la personne: *ati Selemani*, feu Séliman. (Sacleux 1939, 75)

(**Ati** (Ngw.) Late, followed by the name of the person: *ati Selemani*, the late Seliman.)

In *Manuel de Kingwana: le dialect occidental de swahili* (Whitehead & Whitehead 1928), a dictionary specifically concentrated on this particular variety, neither *ati* nor *eti* are included.

Differences between *ati* and *eti* are not discussed anywhere in the literature. Most dictionaries treat *ati* and *eti* as synonyms. *Ati* is the form that occurs the most: in many dictionaries *eti* is not to be found.

To summarize, this is a short list of all the functions or translations of *ati* and *eti* that I discussed in this subsection, ranked according to their recurrence in the dictionaries, from most common to least common.

- '(I) say!' 'Dis donc!': Sacleux 1939, Pick 1964, Höftmann & Mhando 1963, Johnson 1971, Steere 1894, TUKI 2001, Mertens 2006, Google Translate s.d., Krapf 1882, Glosbe s.d., Lenselaer 1983
- Calling attention: Sacleux 1939, Pick 1964, Höftmann & Mhando 1963, Johnson 1971, Oxford Dictionaries s.d., Glosbe s.d., Lenselaer 1983, Taasisi 1981
- Expressing surprise: Johnson 1971, Lenselaer 1983, Taasisi 1981
- Showing contempt: TUKI 2001, Oxford Dictionaries s.d., Taasisi 1981
- Introducing speech: Sacleux 1939, Höftmann & Mhando 1963, Glosbe s.d.
- Expressing doubt: Krapf 1882, Oxford Dictionaries s.d.
- Request for repetition (only for *eti*): Glosbe s.d., de Schryver & Joffe 2004-2018
- 'As if' (only for *eti*): Glosbe s.d., de Schryver & Joffe 2004-2018
- Expressing disagreement: Glosbe s.d.

- (Only in Kingwana) in front of the name of a deceased person: Lenselaer 1983, Sacleux 1939

4.4.2 *Ati and eti as (in)direct speech indicators*

Other than in dictionaries, *ati* and *eti* have not been extensively debated in the literature. Massamba (1986) discusses *eti* in his paper on reported speech in Swahili. He regards *eti* as the standard form and writes that *ati* is less frequent. Massamba (1986, 99) defines reported speech as "a message conveying utterance for which the reporter claims no authorship" and distinguishes three kinds of reported speech. The first one is direct speech, in which the reporter quotes the actual words of the original speaker. He gives the following example:

- (1) **Amina akasema, "nitakwenda mjini kesho."** (Massamba 1986, 100)

Amina	a-ka-sem-a	ni-ta-ku-end-a	m-ji-ni
Amina	SP ₁ -CONS-say-FV	SP ₁ SG-FUT-SM-go-FV	NP ₃ -town-LOC
kesho			
tomorrow			

'Amina said, "I will go to town tomorrow".'²

The second kind of reported speech is indirect speech, whereby the reported does not quote the words of the original speaker, but rather paraphrases them. This means that the wording changes, as they have to fit the new grammatical construction and the point of view of the reporter. In English, the word 'that' is often used as an indirect speech indicator. Indirect speech indicators are "those sentential elements that are used to either introduce or indicate a reported clause" (Massamba 1986, 101). In the following example, the Swahili indirect speech indicator is *kwamba*. However, the use of an indirect speech indicator is not always compulsory.

- (2) **Maganga alisema kwamba baba yake alikuwa mgonjwa.** (Massamba 1986, 100)

Maganga	a-li-sem-a	kwamba	baba	i-ake
Maganga	SP ₁ -PST-say-FV	COMP	father	PP ₉ -POSS ₁
a-li-ku-w-a	m-gonjwa			
SP ₁ -PST-SM-be-FV	NP ₁ -sick			

'Maganga said that his father was sick.'

² All glossing in this paper is my own.

The last kind of reported speech is called free indirect speech and is an intermediate stage between the two previous styles. It is basically indirect speech but it lacks the reporting clause that introduces the shift from narration to reporting. The change in wording of the reported clause that occurs in indirect speech, is not retained in this form of reported speech.

Eti occurs in the second type of reported speech: indirect speech. This particle is an example of an indirect speech indicator, much like 'that' in English. Massamba lists some other indirect speech indicators in Swahili. He first gives a list of indirect speech indicators that were provided by Maw (1969): *kuwa, kwamba, ya kuwa, kana, kana/kama kwamba, eti* and *kuona*. However, Massamba argues that the data used by Maw are mostly derived from a few dialects and that her list of indirect speech indicators is thus not representative of standard Swahili. He then continues to give his own list that slightly differs from Maw's list. According to Massamba, these are the indirect speech indicators in Swahili: *kuwa, ya kuwa, kwamba, ya kwamba, kama, eti* and *sijui*. The last two indicators in this list are a little different from the others. While *kuwa, ya kuwa, kwamba, ya kwamba* and *kama* are neutral, *eti* and *sijui* have an additional connotation. *Sijui* ('I don't know'), as the name implies, expresses doubt. *Eti*, as suggested by most dictionaries, carries an implication of surprise and doubt. Massamba (1986, 102) describes *eti* as follows:

"The form *eti* (sometimes *ati*) means something very close to the English expression 'I say!' (as an expression of surprise). Whenever this indicator is used, the implication is that there is some element of surprise involved; furthermore, the reporter is casting doubt about what was stated by the original speaker."

Massamba gives an example of the use of *eti* as an indirect speech indicator:

(3) **Fatma amesema *eti* hakumwona mgeni.** (Massamba 1986, 102)

Fatma	a-me-sem-a	eti	ha-ku-mu-on-a
Fatma	SP ₁ -PRS.PRF-say-FV	PART NEG.SP ₁ -PST.NEG-OP ₁ -see-FV	
	m-geni		
	NP ₁ -stranger		

'Fatma said that she did not see the stranger.'

Massamba further mentions that some speakers suggest that *kwamba eti* and *eti kwamba* can also be used as indicators. According to Massamba however, their acceptability is rather low.

Besides Massamba (1986), there are no other grammars that discuss the properties of *ati* and *eti* as reported speech indicators in Swahili.

4.4.3 Origin

There is considerably less literature about the etymology of *ati* and *eti*. Krapf (1882) mentions a similar expression in Amharic: *anten*, or *enten*, literally 'the what's his name', but remains silent about whether or not this particle is etymologically related to *ati/eti*. In the Swahili-English dictionary of Madan (1903), *ati* is marked with the asterisk, marking words that are not of Bantu origin. According to Madan, *ati* can not be traced back to Proto-Bantu, (he does not give any alternative though). However this claim seems not to be true at all, as all subsequent literature points to the opposite.

The definition of *ati/eti* according to Johnson (1971) has already been quoted above. This dictionary states that *ati/eti* "[g]enerally can be translated by 'I say!', as it is from the verb *ti*, say, found in many Bantu languages" (Johnson 1971, 20). This hypothesis is supported by (Güldemann 2008, 418–19), who, in his habilitation thesis on quotative indexes in African languages, writes that *ati/eti* is etymologically related to the Common-Bantu quotative verb *ti*.

Bantu possesses a verbal lexeme that is reconstructed by Guthrie (1970) as **-ti*. This lexeme is widespread in the Bantu language family. In many Bantu languages *ti* is a quotative marker or plays another important role in the domain of reported speech (both direct and indirect speech). Apart from reported speech, it is also involved in a great variety of other contexts. Different functions can be found both across languages and within one and the same language. This multiplicity of functions makes it hard to assign a consistent meaning to this lexeme, as already became apparent in the previous subsection on the meaning of *ati/eti* in Swahili.

The multiplicity of functions of the reflexes of **-ti*, can be accounted for in two ways. According to the first one, the different functions emerged from different lexical items. The second approach says that all different uses of *ti* developed from one and the same word. Güldemann argues that the second approach is most probable in the case of *ti*. This means that *ti* should have an original basic meaning from which all the other uses developed.

What is the original basic meaning of *ti* and how did the grammaticalization take place towards the ways in which *ti* is used in the present-day Bantu languages? Güldemann (2002) attempts to answer this question in his paper 'When "say" is not *say*', in which he describes the functional versatility of the Bantu quotative marker *ti*, with special reference to Shona.

In this article, Güldemann writes that wide cross-linguistic evidence suggests a universal path of grammaticalization for quotative markers. This unilinear grammaticalization chain is frequently presented as follows: verb 'say' > quotative marker > complementizer > result/purpose clause marker. According to this model, *ti* should originally be a speech verb with the meaning 'to say'. This is confirmed by Meeussen (1980, 12), who translates **-ti* as 'say'.

Guthrie (1970) does not agree with this idea. He proposes a different original meaning: 'that' or 'namely'. This is what Guthrie writes about **ti*:

C. S. 1727 **-ti* that, namely; say ...

Most of these entries form verbal bases with a limited number of tenses, and from the evidence of their distribution it seems possible that **-TÈ* occurred in P[roto]B[antu]-X, probably with the meaning 'that, namely'. The extension of this to provide an all purpose radical, the meaning of which can perhaps be best expressed in English as 'saying', may have taken place later. (Guthrie 1970, 105)

In contrast with the prevailing idea, Guthrie does not consider *ti* to be a speech verb originally, but he rather sees the function of *ti* as a complementizer ('that, namely') as the original purpose. According to him, it was not until later that this complementizer evolved into a speech verb 'to say'. This reconstruction is motivated by the fact that in many Bantu languages, the complementizer function is the central use of *ti*. However, as Güldemann (2002, 274) points out, a synchronic frequency of a particular function does not confirm a diachronically earlier situation.

Güldemann agrees with Guthrie to a certain extent: he also thinks that *ti* is originally not a speech verb. However, Güldemann's reconstructed meaning of *ti* differs from Guthrie's translation 'that, namely'. Güldemann describes the original meaning like this: "*The verb stem*

ti provides a cataphoric orientation for the hearer towards some subsequently identified information about the entity cross-referenced in its subject concord" (Güldemann 2002, 272). The best equivalent in English would be 'thus'.

Güldemann subsequently constructs a scenario for the grammaticalization of *ti* in Shona. In Shona, *ti* is a verb that is traditionally translated as 'say', 'think' or 'do', but unlike the canonical verb stems it has a wide range of idiomatic and grammatical functions. Its most salient function is the introduction of reported discourse, both quotative constructions and indirect speech. Other functions of *ti* in Shona include clause linkage, introducing nominal lists, embedding ideophones into the discourse and expression of quality and manner. The introductory function of *ti* is therefore crucial, not unlike *ati* and *eti* in Swahili. Furthermore, *ti* often conveys a contrast to the other information in the discourse, which is also similar to the functions of *ati* and *eti* in Swahili (surprise, doubt and contempt all imply contrast). How this multifunctionality of *ti* in Shona developed from Proto-Bantu **ti* 'thus', is illustrated in the following figure (Güldemann 2002, 256–73).

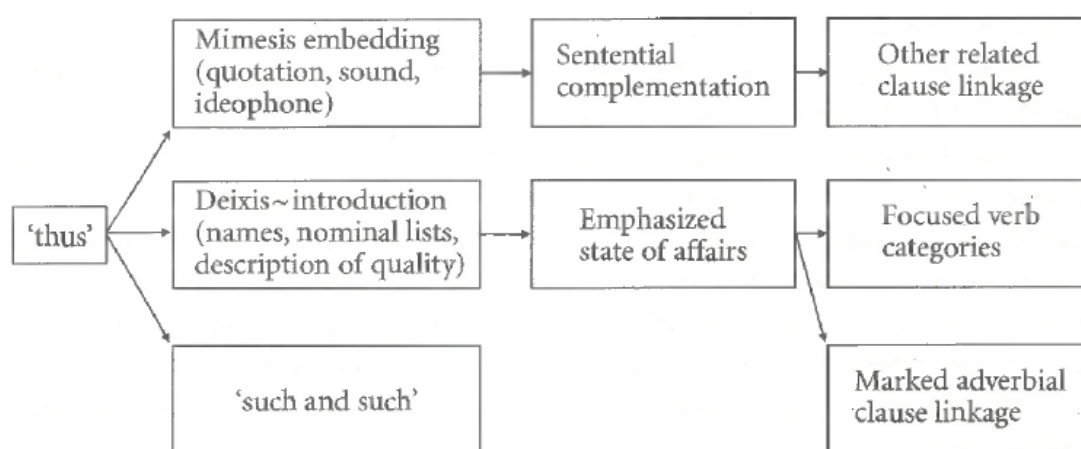


Figure 3. Schema for the grammaticalization relationship between the major functional domains of Shona *ti* (Güldemann 2002, 282).

As Figure 3 illustrates, the original meaning 'thus', proposed by Güldemann, gave rise to two grammatical functions and one specialized meaning. The first grammatical function is embedding mimetic expressions into the discourse. This includes direct reported discourse, ideophones and sound imitations. Indirect reported discourse also falls into this category, although its degree of mimesis is considerably less than is the case with ideophones, sound imitations and direct speech. Precisely because it is the least mimetic, indirect speech lead to

the emergence of two other grammatical functions of *ti* in Shona: sentential complementation and other related clause linkage. The second grammatical function that emerged from 'thus', more specifically from the demonstrative-introductory function of this word, is discourse deixis: identifying nominal lists and introducing expressions of quality and manner. This introducing function of *ti* is the source for emphasizing states of affairs in discourse, which in turn gave rise to the expression of pragmatically marked verbal categories and clause linkage. The third use of *ti* directly linked to its original meaning 'thus', is a lexical meaning that refers to a specific quality or amount. This can be translated as 'such and such', 'particular' or 'some' (Güldemann 2002, 278–82). A similar path of grammaticalization can be expected for *ati* and *eti* in Swahili, even though in Swahili *ati/eti* has a more restricted use than *ti* in Shona. However, there is no literature on the grammaticalization process of *ati* and *eti* in Swahili to confirm or deny this hypothesis.

From the above scheme, a process of increasing 'verbification' can be observed: over time, the Shona particle *ti* progressively adopted the features of a verb. In Shona, this process is not entirely completed (yet), as *ti* has an invariable verb stem. Unlike the other canonical verb stems in Shona, *ti* can not be extended by derivational suffixes and its final vowel can not be changed. Other Bantu languages however, are situated in a further stadium of this process of verbification. Among those are the languages of the Nguni group. This branch of Southern Bantu includes languages like Zulu, Xhosa, Swati and Ndebele. In these languages, the reflexes of **ti* do behave like canonical verb stems, as stem-final changes are possible. This, according to Güldemann (2002), is other evidence of the tendency of **ti* to become a morphologically more canonical verb stem. Other languages, like the languages of zone J, are situated at the other end of the verbification process. These languages still show the assumed original meaning of **ti* 'thus', without the verb-like features that can be found in Shona and the Nguni group (Güldemann 2002, 274–78).

Considering the verbification process of **ti* in many languages related to Swahili, it is natural to assume a similar path of development for *ati* and *eti* in Swahili. According to Güldemann (2002, 278), **ti* was a non-verbal particle that was capable of occurring in a quotative construction of the simple form *subject+ti* (without a tense-aspect marker). Following this line of thought, one could expect the initial vowel of *ati* to be derived from the person marker for the third person singular, namely *a-*. In this scenario, *eti* would be a variation of the original *ati*. However, Güldemann (2008, 419) writes for *ati* and *eti* in Swahili that "there is

no obvious reconstruction of this particle as an earlier finite verb form." He furthermore adds that "*one must remember that ti is originally not a speech verb, but a manner deictic.*"

The languages discussed here, as well as the other languages discussed by Güldemann (2002), indicate a correlation between the degree of verbhood of **ti* and geography. As Güldemann (2002, 284–285) points out, the process of increasing verbification corresponds with a geographical North-South axis. The languages furthest removed from the Bantu homeland in West Africa, such as the Nguni group in Southern Africa, are the most innovative. The reflexes of **ti* in these languages behave most like a canonical verb stem that can be translated as 'say' or 'think'. On the other hand, the cognates in zone J in Central Africa, closer to the place of origin of Bantu, are in a much more conservative stage of the verbification process and sustained the assumed original meaning of **ti* 'thus'. However, Güldemann is very cautious when discussing this hypothesis and states that more comparative research has to be conducted in order to affirm this claim.

Placing Swahili on the North-South axis, it can be observed that Swahili fits this scenario. *Ati/eti* did not obtain the properties of a verb. Instead it is used as a discourse marker that plays an important role in the domain of reported discourse, a function that is much closer to the assumed original meaning of **ti*.

5 *Ati and eti*: a quantitative analysis

This chapter provides a quantitative analysis of the use of *ati* and *eti* in the corpus. The chapter begins by analysing the frequency of the two discourse markers in the corpus, comparing the different geographical areas and periods. It will then go on to discuss clusters that frequently appear with *ati* or *eti*.

5.1 *Frequency*

In the entire corpus, there are 376 attestations of *ati* and *eti*. In a corpus of 2 186 147 tokens, this equates to 1,72 attestations per 10 000 tokens. If we take a look at how these 376 attestations are divided over the various geographical categories of the corpus, we get a better insight in the regional distribution of *ati* and *eti*. Let us first compare the two countries to one another. The Kenyan texts have 182 attestations of *ati* and *eti*, which corresponds to 2,55 attestations per 10 000 tokens in the Kenyan subcorpus. The Tanzanian texts include slightly more attestations of *ati* and *eti*, 192. However, because of the bigger size of the Tanzanian subcorpus, this result is proportionally lower compared to the Kenyan subcorpus, with only 1,62 attestations per 10 000 tokens. Comparing the coastal regions with the inland regions, one can also observe a discrepancy in the frequency of *ati* and *eti*, although it is slightly less significant than the difference between Kenya and Tanzania. In texts that were written in the inland regions of Tanzania and Kenya, 127 attestations of *ati* and *eti* can be observed, or 2,41 per 10 000 tokens. Texts from the coast contain proportionally less attestations of the two particles, namely 1,82 per 10 000 tokens (247 attestations in total).

The 376 attestations of *ati* and *eti* can also be divided according to when the texts they appear in were written. The early period (19th and early 20th century) has a very low number of 20 attestations of *ati* and *eti*. In a subcorpus of 953 910 tokens, this equates to 0,21 attestations per 10 000 tokens. In the middle period (early 20th century until the 1960s), *ati* and *eti* occur 23 times. This period is represented by only 159 149 tokens in the corpus, which gives an average of 1,45 attestations of *ati* and *eti* per 10 000 tokens. In the recent period (1970s until the 2010s) the number of *ati/eti* attestations increases considerably up to 333, which is equivalent to 3,10 attestations per 10 000 tokens (in a subcorpus of 1 073 088 tokens).

These results suggest that the use of *ati* and *eti* is more established in Kenya and in the inland regions, as compared to Tanzania and the coastal regions. The differences are rather big: the

number of attestations of *ati* and *eti* per 10 000 tokens in Tanzania is 1,62, compared to 2,55 in Kenya, and it is 2,41 in the inland regions, compared to 1,82 in the coastal areas. Moreover, the results indicate that there is an increase over time in the use of this discourse marker. The number of *ati/eti* attestations per 10 000 tokens increases from 0,21 in the early period, over 1,45 in the middle period, up to 3,10 in the recent period. However, this increase could possibly be attributed to the text types that are included in the different categories of the corpus. While the subcorpus of the early period consists mainly of traditional stories, poetry and religious texts, the middle and recent period include more novels and plays.

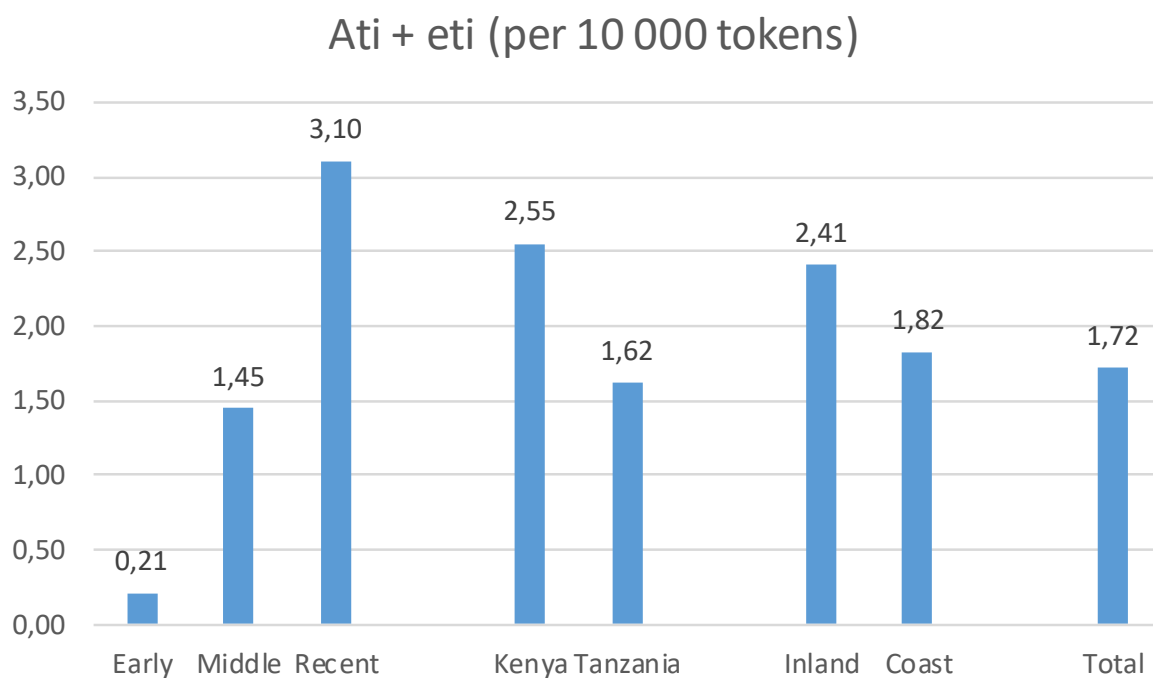


Figure 4. Number of attestations of *ati* and *eti* per 10 000 tokens.

The previous results concern the frequency of *ati* and *eti* without differentiating the two particles. We now analyse the frequency of each particle individually in order to know which version of the discourse marker is preferred: *ati* or *eti*? In the total corpus, there are 310 attestations of *ati* and a mere 66 attestation of *eti*. This means that *ati* occurs 4,70 times more than *eti*. Another way to put this, is that 82,45 % of all the *ati/eti* attestations in the corpus are *ati*, compared to 17,55 % for *eti*.

If we take a look at the geographic categories of the corpus, we see that in the texts from Kenya, there are 167 attestations of *ati*, or 91,76 %. *Eti* on the other hand occurs only 15 times, or 8,24 %. The *ati/eti* ratio in Kenya therefore is very high: *ati* occurs 11,13 times more

than *eti*. In the Tanzanian subcorpus there are 141 attestations of *ati* (or 73,44 %), 51 attestations of *eti* (or 26,56 %), which gives an *ati/eti* ratio of 2,76. The texts that were written at the coast contain 222 attestations of *ati* (89,88 %), 25 attestations of *eti* (10,12 %) and have an *ati/eti* ratio of 8,88. The texts written in the inland regions contain 86 attestations of *ati* (67,72 %), 41 attestations of *eti* (32,28 %) and have an *ati/eti* ratio of 2,10.

There are also some differences in the frequency of *ati* and *eti* through time. In the early period *ati* occurs 20 times, whereas *eti* does not occur at all. The number of attestations of *ati* in the middle period is 18 (78,26 %) and of *eti* 5 (21,74 %). This means that in the middle period *ati* occurs 3,60 times more than *eti*. Lastly, in the subcorpus of the recent period, there are 272 attestations of *ati* (81,68 %) and 61 of *eti* (18,32 %), which gives an *ati/eti* ratio of 4,46. All the above data are summarized in the following chart.

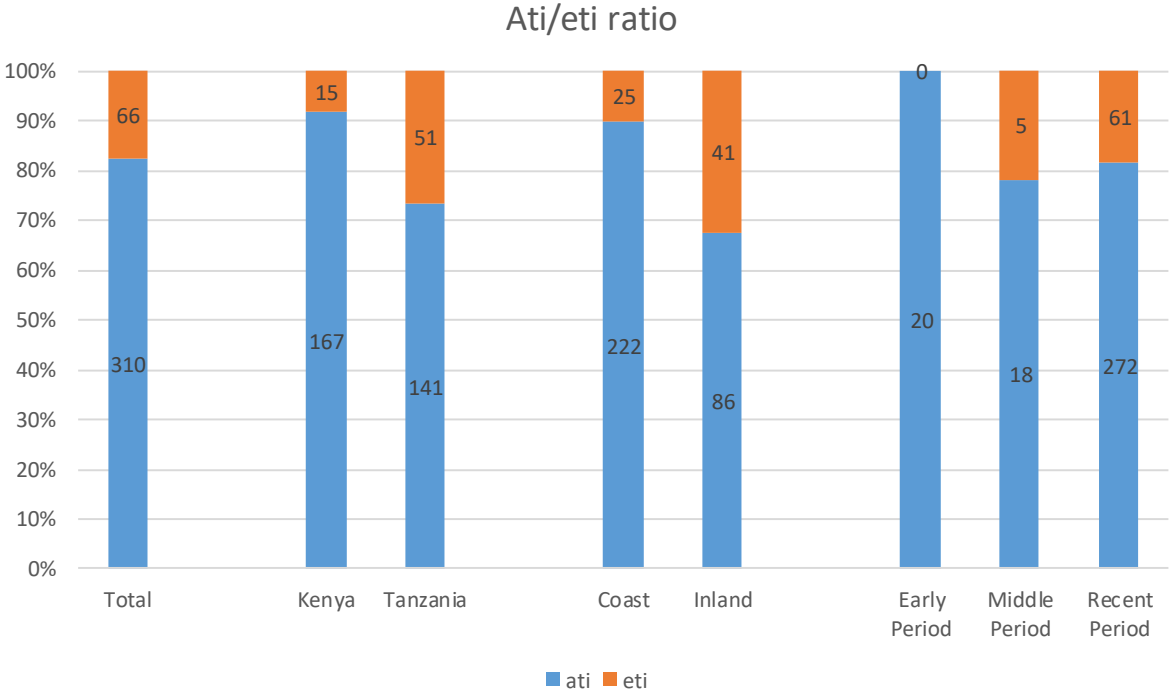


Figure 5. Comparison of the *ati/eti* ratio, according to region and period.

So which version of the discourse marker is preferred: *ati* or *eti*? This graph beautifully visualizes to what extent *ati* and *eti* are represented in the corpus and gives the answer to that question in an instant. It is clear that *ati* occurs much more often than *eti*, 4,70 times to be precise. This difference is bigger in Kenya and in the coastal regions, as compared to Tanzania and the inland regions where *eti* is slightly more popular (though still clearly less frequently used than *ati*).

An evolution over the years can also be observed. In all historical periods, *ati* is more popular than *eti*. Until the early 20th century, *eti* does not even occur once. However, the frequency of *eti* compared to *ati* increases as time goes by. In the texts written from the 1970s until the 2010s, the use of *eti* has gone up to 18,32 %. In the middle period (early 20th century until the 1960s), this percentage is even a little higher (21,74 %). This small peak can probably be attributed to the limited size of the subcorpus of the middle period, only 159 149 tokens, allowing a minor change in absolute numbers to have a major impact on the percentages. Overall, we can conclude that *eti* is gaining ground and starting to catch up with *ati*.

5.2 Collocations

The following section describes the units that frequently collocate with *ati* and *eti*. The first question one might ask in this regard, concerns the syntactic arrangements of the particle: do *ati* and *eti* occur at the beginning, the middle or at the very end of the clause? When *ati/eti* functions as a complementizer and is thus the head of the complement clause, I consider it clause-initial. Furthermore, when *ati/eti* is not the very first word of the clause, but it is part of a clause-initial conjunction (eg. *ili ati* 'so that' or *kwamba eti* 'that'), I also consider it clause-initial.

In the corpus, the majority of the attestations of this discourse marker occur in the clause-initial position: 270 attestations, amounting to 71,81 % of the total number of attestations. The number of *ati/eti* attestations in the clause-final position is much lower: 89 attestations or 23,67 %. *Ati/eti* occurring in the middle field is rather unusual, with only 15 attestations or 3,99 %. There are also two cases in which *ati* or *eti* appear isolated and are the only word in the whole sentence.

Once again, a distinction can be made in these results between *ati* and *eti*. Comparing the two variations of the discourse marker, it can be observed that some differences exist between the two variants. For example, while more than a quarter of the *ati* attestations appear in the clause-final position (84 attestations or 27,10 % of the total number of *ati* attestations in the corpus), this position is much less common for *eti* (only 5 attestations or 7,58 % of the total number of *eti* attestations in the corpus). The clause-initial position on the other hand, is more common with *eti* than with *ati*. *Eti* appears as the first word of the clause in 86,36 % of the total number of *eti* attestations (57 out of 66), while for *ati* this is only 68,71 % (213 out of

311). The clause-internal position is for both variants the least common, with a minimal difference (3,87 % of all *ati* attestations and 4,55 % of all *eti* attestations). The cases in which the discourse marker is the only word in the sentence are exceptional for both *ati* and *eti*, with each only one example in the corpus. Moreover, these two examples could be considered as simply an orthographic variation on the clause-initial position.

I turn now to which units frequently collocate with *ati* and *eti*. The unit that most frequently appears next to the discourse marker is not a word, but a punctuation mark. It is interesting to have a quick look at the kinds of punctuation that can often be found near *ati* and *eti*, in order to get a general idea of the clause types this discourse marker is used in. When *ati* takes the clause-final position, it is most frequently followed by a full stop or comma, thus marking the end of a declarative clause. This is the case in more than half of the clause-final occurrences of *ati*. Otherwise, *ati* is followed twenty times by an exclamation mark (23,81 % of all clause-final occurrences of *ati*) and eight times by a question mark (9,52 %). When *eti* takes the clause-final position, it is most frequently followed by an exclamation mark (four times or 80 % of all clause-final *eti* occurrences), marking the end of an exclamative clause. *Eti* is never followed by a question mark.

Regarding words that frequently collocate with *ati* and *eti*, the cluster that appears most frequently is *ati* in combination with a conjugated form of a reporting verb like *-sema* 'to say' or *-ambia* 'to tell' (*ati* immediately following the verb). This cluster, *-sema ati* and *-ambia ati*, occurs eighteen times in the corpus (5,81 % of all *ati* attestations in the corpus). With *eti*, this combination appears only two times (3,03 % of all *eti* attestations). This seems to confirm Massamba's (Massamba 1986, 101) claim that *ati* and *eti* are used to introduce indirect speech. This will be discussed in greater detail in Subsection 6.1.1.

Another frequently appearing cluster is *ati kwa sababu*, meaning 'because', that occurs ten times in the corpus (3,23 % of all *ati* attestations). A very similar cluster is *ati kwa kuwa*, which also means 'because' and appears seven times (another 2,26 % of all *ati* attestations). *Eti kwa sababu* appears two times (3,03 % of all *eti* attestations), while *eti kwa kuwa* is absent in the corpus. The most frequent cluster with *eti* is *kwamba eti*, which can be translated by the conjunction 'that'. This cluster occurs seven times, which is 10,61 % of all *eti* attestations. *Kwamba ati* on the other hand, appears only three times (0,97 % of all *ati* attestations). More frequent clusters are *ati nini? eti nini?* 'hey what?' (appears fourteen times or 3,72 % of all

ati/eti attestations) and *ati/eti bwana/bibi/dada* 'hey sir/madam/sister' (appears eleven times or 2,93 %).

More analysis about the exact meaning of these clusters follows in the following chapter.

6 The pragmatic uses of *ati* and *eti*

6.1 Sentential complementation

One of the most frequently occurring functions of the discourse marker *ati/eti* is that of a complementizer. According to Nordström (2010, 91–95), a complementizer is a type of subordinating conjunction. Subordinating conjunctions, or subordinators, are used to introduce subordinate clauses. A subordinate clause is defined as "a clause that is (a part of) a constituent in another clause." Nordström distinguishes three types of subordinators, divided according to function: complementizers, relative subordinators and adverbial subordinators. Nordström argues that relative subordinators introduce subordinate clauses that function as modifiers and adverbial subordinators introduce subordinate clauses that function as adverbials. The latter is relevant to *ati* and *eti* as well and will be discussed in Section 6.2.

This section deals with the first type of subordinating conjunction: complementizers. Nordström (2010, 94) describes complementizers as elements that introduce embedded clauses that function as arguments. They are also used to introduce reported speech, which will be addressed in the next two subsections (6.1.1 and 6.1.2). The most common complementizer in English is 'that'.

In Swahili, *kwamba* and *kuwa* (grammaticalized forms of the verbs 'to say' and 'to be') are usually used as complementizers, although the complementizer is often omitted in spoken language (Russell 1992, 126).

(4) **Flora aliona kwamba Charles alikuwa hafahamu vizuri mahali alipokuwa akiishi sasa.** (Kezilahabi 1971, 82)

Flora	a-li-on-a	kwamba	Charles	a-li-ku-w-a
Flora	SP ₁ -PST-see-FV	COMP	Charles	SP ₁ -PST-SM-be-FV
ha-fahamu	vi-zuri	m-ahali	a-li-po-ku-w-a	
NEG.SP ₁ -know	NP ₈ -good	NP ₁₈ -place	SP ₁ -PST-RC ₁₆ -SM-be-FV	
a-ki-ishi	sasa			
SP ₁ -SIT-live	now			

'Flora saw that Charles did not know well the place where he was living now.'

(5) **Siku hiyo ndipo nilipotambua kuwa Munene alikuwa amefariki dunia.**

(Wamitila 2007, 270)

Ø-siku	hi-yo	ndi-po	ni-li-po-tambu-a	
NP ₉ -day	REF.DEM-RC ₉	COP-RC ₁₆	SP _{1SG} -PST-RC ₁₆ -realise-FV	
kuwa	Munene	a-li-ku-w-a	a-me-fariki	Ø-dunia
COMP	Munene	SP ₁ -PST-SM-be-FV	SP ₁ -PRF-separate.from	NP ₉ -world

'It was on that day that I realised that Munene had died.'

In the previous examples, one can clearly distinguish the main clause and the embedded clause, separated by the complementizers *kwamba* and *kuwa*. This role of the complementizer can also be played by *ati* or *eti*, as in (6). However, *ati* and *eti* are pragmatically more marked, whereas *kwamba* and *kuwa* are pragmatically neutral. When *ati* or *eti* are used, this often adds an additional connotation to the utterance. In (6), the speaker uses *eti* to convey an element of contempt towards the interlocutor. More about the connotations of *ati* and *eti* as a complementizer will be given in the subsections on indirect and direct speech (6.1.1 and 6.1.2).

(6) **Hivi wajifanya eti umenisahau?** (Zani & Kitsao 1975, 20)

hi-vi	u-a-ji-fany-a	eti
PROX.DEM-PP ₈	SP _{2SG} -GEN.PRS-REFL-do-FV	PART ³
u-me-ni-sahau		
SP _{2SG} -PRS.PRF-OP _{1SG} -forget		

'Are you pretending that you have forgotten me?'

A combination of *kwamba* or *kuwa* with *ati* or *eti* is also possible. This occurs in 18,52 % of the instances in which *ati* and *eti* are used as a complementizer (6,17 % with *kuwa* and 12,35 % with *kwamba*). If that is the case, *kwamba* or *kuwa* generally precedes *ati* or *eti*. The reverse occurs as well, but less frequently (only once in the entire corpus, or 1,23 % of all *ati/eti* attestations that are used as a complementizer).

(7) **[...] na kujifanya kwamba ati wao ni wanamapinduzi.** (Chachage 1991, 103)

na	ku-ji-fany-a	kwamba	ati	wao	ni	wa-anamapinduzi
----	--------------	--------	-----	-----	----	-----------------

³ Since the uses of *ati* and *eti* are so diverse, they will simply be glossed as PART (particle).

and NP₁₅-REFL-do-FV COMP PART they COP NP₂-revolutionary
[...] and pretending that they are revolutionaries.'

Being employed as a complementizer is one of the main functions of *ati* and *eti*. In the corpus, there are 37,05 attestations of *ati/eti* as a complementizer per million tokens. This means that 21,54 % of all the attestations of *ati* and *eti* in the corpus are used as complementizers. Its frequency increases over the years: the presence of *ati/eti* as a complementizer goes from 2,10 attestations per million tokens in the early period (10,00 % of all *ati/eti* attestations in the early subcorpus) to 63,37 in the recent period (20,42 %). Furthermore, *ati/eti* as a complementizer appears more frequently in Kenya and inland (respectively 64,89 and 68,20 attestations per million tokens), as compared to Tanzania and the coast (26,95 and 31,83 attestations per million tokens).

The difference in frequency between the two variants of the discourse marker, *ati* and *eti*, is minimal and practically nonexistent. There are 30,65 attestations of *ati* and 6,40 attestations of *eti* as a complementizer per million tokens in the corpus. This seems like a big gap, except if one takes into account that there are less attestations of *eti* than of *ati*. 21,61 % of all the *ati* attestations are used as a complementizer, compared to 21,21 % of the *eti* attestations in the corpus. *Eti* as a complementizer occurs most frequently in Tanzania and in the inland regions and is not used until the recent period.

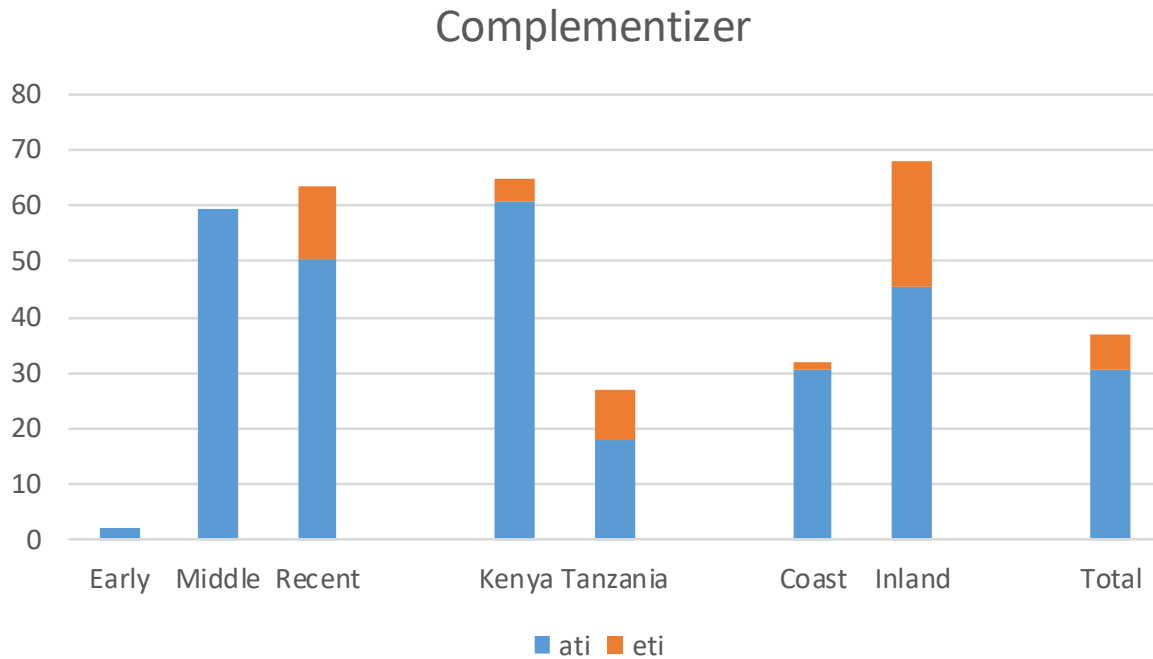


Figure 6. Number of *ati/eti* attestations in the subcategories of the corpus that are used as a complementizer, per million tokens.

The majority of the instances in which *ati* and *eti* function as a complementizer, are in the context of reported speech: 85,19 %. This will be discussed in the next two subsections.

6.1.1 Indirect speech

One of the common specific functions of complementizers is introducing or indicating a reported clause. This may be direct or indirect speech. *Ati* and *eti* function in these cases as reported speech indicators or quotative indexes. Güldemann (2008, 51) describes a quotative index as "a segmentally discrete linguistic expression which is used by the reporter for the orientation of the audience to signal in his/her discourse the occurrence of an adjacent representation of reported discourse." I distinguish two types of quotative indexes, indirect speech indicators and direct speech indicators.

Indirect speech in Swahili breaks down as follows. The sentence contains at least two clauses: the reporting clause and the reported clause. The reporting clause is the main clause, used by the speaker to indicate that the words which follow are being reported and to identify the original speaker. This is achieved by using a reporting verb, for instance *-sema* 'to say'. The reported clause, on the other hand, is a subordinate clause that contains the discourse

paraphrased by the speaker. The reported clause can (but does not have to) be introduced by a quotative index. As mentioned in Subsection 4.4.2, Swahili has several indirect speech indicators, the most essential ones being *kwamba* and *kuwa*, which also function as complementizers apart from reporting indirect speech (cf. supra). *Ati* and *eti* can also be used as markers of indirect speech, as illustrated by the following examples.

- (8) **Anasema *ati* akipita mtihani wake vizuri, ataendelea na masomo ya juu.** (Zani & Kitsao 1975, 29)

a-na-sem-a		ati	a-ki-pit-a		m-tihani		u-ake
SP ₁ -PRS.PROG-say-FV		PART	SP ₁ -COND-pass-FV		NP ₃ -exam		PP ₃ -POSS ₁
vi-zuri	a-ta-endele-a		na	ma-somo	ya-a		juu
NP ₈ -good	SP ₁ -FUT-continue-FV		and	NP ₆ -study	PP ₆ -CONN		up

'He said that if he passes his exam, he will pursue higher education.'

- (9) **[...] mliposema *eti* mizigo yangu haikutumwa, mlikuwa mnamsingizia Merton [...]** (Chachage 1991, 121)

m-li-po-sem-a		eti	mi-zigo		i-angu
SP _{2PL} -PST-RC ₁₆ -say-FV		PART	NP ₄ -luggage		PP ₄ -POSS _{1SG}
ha-i-ku-tum-w-a			m-li-ku-w-a		
NEG-SP ₄ -PST.NEG-send-PASS-FV			SP _{2PL} -PST-SM-be-FV		
m-na-m-singiz-i-a			Merton		
SP _{2PL} -PROG-OP ₁ -accuse.falsely-APPL-FV			Merton		

'[...] when you said that my luggage was not sent, you were falsely accusing Merton [...]'

Massamba (1986, 105–7) writes that with reported speech of questions, there is a different indicator. This is similar to English, where the indirect speech indicator 'that' is replaced by 'if', 'whether' or an interrogative pronoun in questions. In Swahili, when the reported question is a polar interrogative, the indirect speech indicator is *kama* 'like'. In this case, the use of an indicator is obligatory and *kama* can not be omitted. When the reported question is a constituent interrogative, the interrogative pronoun (e.g. *nani* 'who', *nini* 'what', *wapi* 'where', *kwa nini* 'why') takes on the role of the indirect speech indicator. These are the only indirect speech indicators for questions in Swahili. Indeed, when *ati* and *eti* are used in a reported speech construction in the corpus, they never introduce questions.

As Massamba (1986, 119) mentions, *eti* occurs in combination with *kwamba* 'that'. *Kwamba eti* and *eti kwamba* can thus also be used as indirect speech indicators. According to Massamba however, their acceptability is rather low. For *ati* this seems to be the case indeed, as *kwamba ati* as an indirect speech indicator only appears one time in the entire corpus. *Kwamba* in combination with *eti* however, does occur more frequently. *Kwamba eti* can be found six times, which equates to 54,55 % of all the times *eti* is used to indicate indirect speech or 2,74 times per million tokens in the corpus. Contrary to what Massamba argues, *kwamba eti* is actually not that rare. *Eti kwamba*, despite being listed by Massamba, was not found in the corpus.

- (10) **Halafu wanasema kwamba eti mizigo ya John Safari sikupeleka.** (Chachage 1991, 102)

halafu	wa-na-sem-a		kwamba	eti	mi-zigo	i-a
then	SP ₂ -PRS.PROG-say-FV		COMP	PART	NP ₄ -luggage	PP ₄ -CONN
John Safari	si-ku-i-pelek-a					
John Safari	NEG.SP _{1SG} -PST.NEG-OP ₄ -send-FV					

'Then they said that I did not send the luggage of John Safari.'

- (11) **Inasemekana kwamba ati wakubwa wameshauri kuwa asiondoke, akae kama mtego.** (Chachage 2005, 285)

i-na-sem-ikan-a		kwamba	ati	wa-kubwa
SP ₉ -PRS.PROG-say-RES-FV		COMP	PART	NP ₂ -big
wa-me-shauri	kuwa	a-si-ondok-e		a-ka-e
SP ₂ -PRS.PRF-advise	COMP	SP ₁ -NEG-leave-SBJV		SP ₁ -stay-SBJV
kama	m-tego			
like	NP ₄ -trap			

'It is said that the elders have advised that he should not leave, he should stay as a trap.'

Apart from *kwamba*, there is another conjunction that can be found near *ati*: *kuwa*. The literal translation of *kuwa* would be 'to be', but in this context it is a grammaticalized form that can be translated as 'that'. *Kuwa ati* and *ati kuwa* occur 3 times and once respectively. In combination with *eti*, no conjunctions other than *kwamba* are found.

- (12) **Amesema kuwa ati sasa basi, hatakwenda tena maana amechoka sana, anataka kupumzika.** (Burhani 1981, 110)

a-me-sem-a	kuwa	ati	sasa	basi	ha-ta-ku-end-a
SP ₁ -PRS.PRF-say-FV	COMPPART	now	then	NEG.SP ₁ -FUT-SM-go-FV	
tena	Ø-maana	a-me-chok-a	sana	a-na-tak-a	
again	NP ₉ -meaning	SP ₁ -PRS.PRF-be.tired-FV	very	SP ₁ -PRS.PROG-want-FV	
ku-pumzik-a					
NP ₁₅ -rest-FV					

'He said that now, he will not go again, because he is very tired and wants to rest.'

Most of the time, the complementizer *ati* or (*kwamba*) *eti* follows directly on the reporting verb. This is the case 71,19 % of the time (there are no significant differences between *ati* and *eti*). When the reporting verb is not immediately followed by the reported clause, but there are one or more words between them, *ati* and *eti* are used to indicate the start of the reported clause. Otherwise, it might be difficult for the listener to distinguish between the reporting clause and the reported clause. This is the case in example (13), in which nine words appear between the reporting verb and the reported clause.

- (13) **Kaambiwa na wale Wazungu jamaa zao wakaao Kiunga cha Kisauni, ati, pana Mwarabu mmoja kati yetu [...]** (Mbotela 1934, 72)

ka-ambi-w-a	na	wa-le	wa-zungu		
SP ₁ .PRF-tell-PASS-FV	and	PP ₂ -DIST.DEM	NP ₂ -white.person		
Ø-jamaa	zi-ao	wa-kaa-o	ki-unga	ki-a	Kisauni
NP ₁₀ -family	PP ₁₀ -POSS ₂	SP ₂ -live-RC ₂	NP ₇ -suburb	PP ₇ -CONN	Kisauni
ati	pa-na	mu-arabu	m-moja	kati	i-etu
PART	SP ₁₆ -have	NP ₁ -Arab	NP ₁ -one	middle	PP ₉ -POSS _{1PL}

'He was told by these white people whose family lives in the suburb of Kisauni, that there is an Arab amongst us [...]'

The same applies when one reporting clause is followed by several reported clauses. In this case *ati* or *eti* is used to indicate the start of a new reported clause.

- (14) **Wao walicheka, wakasema hawana haja ya kuonyeshwa kulima, ati hiyo ni kazi ya kwao ya asili.** (Mbotela 1934, 60)

wao	wa-li-chek-a	wa-ka-sem-a	ha-wa-na	Ø-haja
they	SP ₂ -PST-laugh-FV	SP ₂ -CONS-say-FV	NEG-SP ₂ -have	NP ₉ -need
i-a	ku-onyesh-w-a	ku-lim-a	ati hi-yo	ni
PP ₉ -CONN	NP ₁₅ -show-PASS-FV	NP ₁₅ -farm-FV	PART REF.DEM-RC ₉	COP
Ø-kazi	i-a	Ø-asili		
NP ₉ -work	PP ₉ -CONN	NP ₉ -origin		

'They laughed, and they said that they did not need to be shown how to farm, that this was their traditional work.'

In (14) the speaker paraphrases not only one, but two phrases, namely *hawana haja ya kuonyeshwa kulima* and *hiyo ni kazi ya kwao ya asili*. The first reported clause is not introduced by a quotative index, because it is not really necessary as the reported clause immediately follows the reporting verb. Thus, no confusion on the part of the hearer is possible. However, this clause is followed by a second reported clause, without using another reporting verb. In order to make clear that this second reported clause is separate from the first one, *ati* is used to indicate where the first reported clause ends and where the second one starts.

The reporting verb that introduces the reported clause varies. The most obvious verb, *-sema* 'to say', proves to be the most popular one. An inflected form of this verb turns out to be the reporting verb in 45,76 % of all the occurrences of indirect speech in the corpus, with no discrepancy between *ati* and *eti*. *-ambia* 'to tell' follows *-sema* and is used as the reporting verb 20,34 % of the time. In the remaining 33,90 % of the cases, the reporting verb can be a variety of other verbs, such as *-jibu* 'to answer', *-danganya* 'to lie', *-dai* 'to claim' or *-imba* 'to sing'.

Massamba (1986, 104) indicates that there is a correlation between the reporting verb and the presence of a direct speech indicator. He claims that reporting verbs can be categorised into two groups. "First, there are those whose occurrence does not necessarily require an 'indicator'." Examples of verbs that fall into this category are *-sema* 'to say' and *-ambia* 'to tell'. Note that, although not required, it is allowed for these verbs to be followed by a quotative index. "Secondly, there are those whose occurrence necessarily requires an 'indicator',

otherwise the sentences so formed will be unacceptable." This category includes verbs like *-eleza* 'to explain' and *-fahamisha* 'to inform'. The basis of this categorization, Massamba admits, is rather vague. Massamba argues that the verbs in the first category express statements of facts, while the verbs in the second group imply something more. After a superficial look at the corpus, the verbs stated in the previous paragraph, *-jibu* 'to answer', *-danganya* 'to lie', *-dai* 'to claim' and *-imba* 'to sing', seem to belong to the second group. When they are used as a reporting verb, they are always accompanied by a quotative index, like *ati/eti* or a different one.

Ati and *eti* as indirect speech indicators are most frequently followed by a clause in the indicative mood. 86,44 % of the attestations of *ati/eti* as indirect speech indicator in the corpus introduce a reported clause in the indicative mood. The subjunctive mood occurs 8,47 % of the time and the conditional appears in only 5,08 % of cases. *Eti* as an indirect speech indicator is only followed by indicative clauses in the corpus.

Massamba (1986, 102) points out that, while most of the indirect speech indicators are neutral, some, like *sijui* and *eti*, have an additional connotation. According to him, *eti*, as also suggested by most dictionaries, carries an implication of surprise and doubt. Based on the corpus used for this research, Massamba's claim appears to be true, although the additional connotation is not always present. This is especially true for *ati* (to which Massamba does not refer), which does not carry this additional connotation as much as *eti*. In 16,67 % of the instances where *ati* is used as an indirect speech indicator, and 27,27 % of the times *eti* is used in that manner, there is an element of surprise involved. This can be observed in the following example, in which the speaker uses *ati* to emphasize the surprising and unexpected content of the words they are paraphrasing.

- (15) **Wengine wakajitia kusema *ati* hata damu zao hazikuwa nyekundu bali buluu!**
(Mkangi 1995, 146)

wa-ingine	wa-ka-ji-ti-a	ku-sem-a	ati	hata	Ø-damu
NP ₂ -other	SP ₂ -CONS-REFL-put-FV	NP ₁₅ -say-FV	PART	even	NP ₁₀ -blood
zi-ao	ha-zi-ku-w-a	N-ekundu	bali	buluu	
PP ₁₀ -POSS ₂	NEG-SP ₁₀ -SM-be-FV	NP ₁₀ -red	but	blue	

'Others pretended to say that even their blood is not red, but blue!'

Massamba furthermore mentions that when the speaker uses *eti* as an indirect speech indicator, "the reporter is casting doubt on what was stated by the original speaker" (Massamba 1986, 102). This aspect is particularly reflected by the fact that more than half (54,55 %) of the instances in which *eti* is used to introduce indirect speech, the speaker indicates that the reported clause is actually a lie. In a minority of these cases, the speaker chooses the reporting verb accordingly and uses verbs like *-danganya* 'to lie' or *-laghai* 'to cheat' to emphasize that the reporter does not agree with the statement of the original speaker. Doubt as a connotation with indirect speech indicator *ati* is far more uncommon and only occurs in 8,33 % of all the times that *ati* is used to introduce indirect speech. In these cases, a reporting verb that expresses 'to lie' is always present. The findings from this corpus therefore suggest that *ati* as an indirect speech indicator does not carry the connotation of doubt/lie as much as *eti*, and the speaker is therefore obliged to use a reporting verb that takes on this role.

(16) **Mtambo aache kudanganya ati yeye mganga.** (Zani & Kitsao 1975, 84)

Mtambo	a-ach-e	ku-dangany-a	ati	yeye	m-ganga
Mtambo	SP ₁ -leave-SBJV	NP ₁₅ -lie-FV	PART	he	NP ₁ -doctor

'Mtambo should stop lying that he is a doctor.'

The reported clause in (16), introduced by *ati*, is preceded by the reporting verb *-danganya*, which makes it very clear that the reported clause is a lie. An example of indirect speech that paraphrases a lie without it being made explicit by the reporting verb can be found in (10). In that example, *eti* in itself is enough to make clear that the reporter believes that the reported clause is a lie (underpinned by the context).

A connotation of *ati* and *eti* not reported by Massamba, is contempt or anger. This additional meaning occurs in 18,75 % of the cases in which *ati* functions as an indirect speech indicator and in 54,55 % of the instances in which *eti* is used as such.

(17) **Hebu nikutobolee, mambo yake tayari yameanza kuni—(Anatikisa kichwa). Amekazania ati kuwa nikwambae wewe!** (Kitsao 1983, 37)

hebu	ni-ku-toboL-e-e	ma-ambo	ya-ake	tayari
hey	SP _{1SG} -OP _{2SG} -tell.frankly-APPL-SBJV	NP ₆ -thing	PP ₆ -POSS ₁	ready
ya-me-anz-a	ku-ni-Ø	a-na-tikis-a	ki-chwa	
SP ₆ -PRS.PRF-start-FV	NP ₁₅ -OP _{1SG} -Ø	SP ₁ -PRS.PROG-shake-FV	NP ₇ -head	

a-me-kazan-i-a ati kuwa ni-ku-amba-e wewe
 SP₁-PRS.PRF-insist-APPL-FV PART COMP SP_{1SG}-OP_{2SG}-avoid-SBJV you
 'Hey, let me tell you frankly, his affairs have already started to— (*He shakes his head*).
 He insisted that I should avoid you!'

In the entire corpus, *ati* and *eti* occur 59 times as an indirect speech indicator (26,99 attestations per million tokens). This means that out of all the attestations of *ati* and *eti* in the corpus, 15,69 % is used as an indirect speech indicator. Indirect speech with *eti* is relatively a little more common than with *ati*. However, the difference is not very big: 16,67 % (or 5,03 per million tokens) of all *eti* attestations is used to introduce indirect speech, compared to 15,48 % for *ati* (21,96 per million tokens).

Ati/eti as an indirect speech indicator is only used once in the entire corpus of the early period (19th and early 20th century), which equates to 1,05 per million tokens. In the corpus of the middle period (early 20th century until the 1960s), it occurs six times (32,35 per million tokens). However, all of these instances are found in the same text, meaning that only one author used *ati* as an indirect speech indicator. It is not until the recent period (1970s until the 2010s) that the use of *ati* and *eti* as an indirect speech indicator becomes more widespread. In the recent period, 48,46 per million tokens are *ati/eti* indirect speech indicators.

In Kenya and in the inland regions of the Swahili-speaking region, *ati* and *eti* are used more frequently to introduce indirect speech, as compared to Tanzania and the coastal regions. In the Kenyan texts, 50,02 per million tokens are attestations of *ati* and *eti* used as indirect speech indicators, compared to only 18,53 per million tokens in the Tanzanian subcorpus. The difference between the inland and the coast is more or less the same, with respectively 51,15 and 23,15 *ati/eti* indirect speech indicators per million tokens.

There is also a difference between *ati* and *eti*. In Kenya, indirect speech is more often used with *ati* than with *eti*: 20,96 % of all *ati* attestation is used as an indirect speech indicator (47,32 per million tokens), compared to 13,33 % of all *eti* attestations (2,70 per million tokens). In Tanzania, the opposite is true and the discrepancy is even more significant. 17,65 % of all *eti* attestations in the Tanzanian part of the corpus are used to introduce indirect speech (7,58 per million tokens), while for *ati* this is only 9,22 % (10,95 per million tokens). This also goes for the inland regions, although the difference there is smaller: 24,39 %

indirect speech for *eti* (18,94 per million tokens), and 19,77 % for *ati* (32,21 per million tokens). In the coastal regions, this is reversed, as only 4 % of all the attestations of *eti* function as an indirect speech indicator (0,72 per million tokens), compared to 13,96 % for *ati* (22,42 per million tokens). All the above data can be observed in Figure 7.

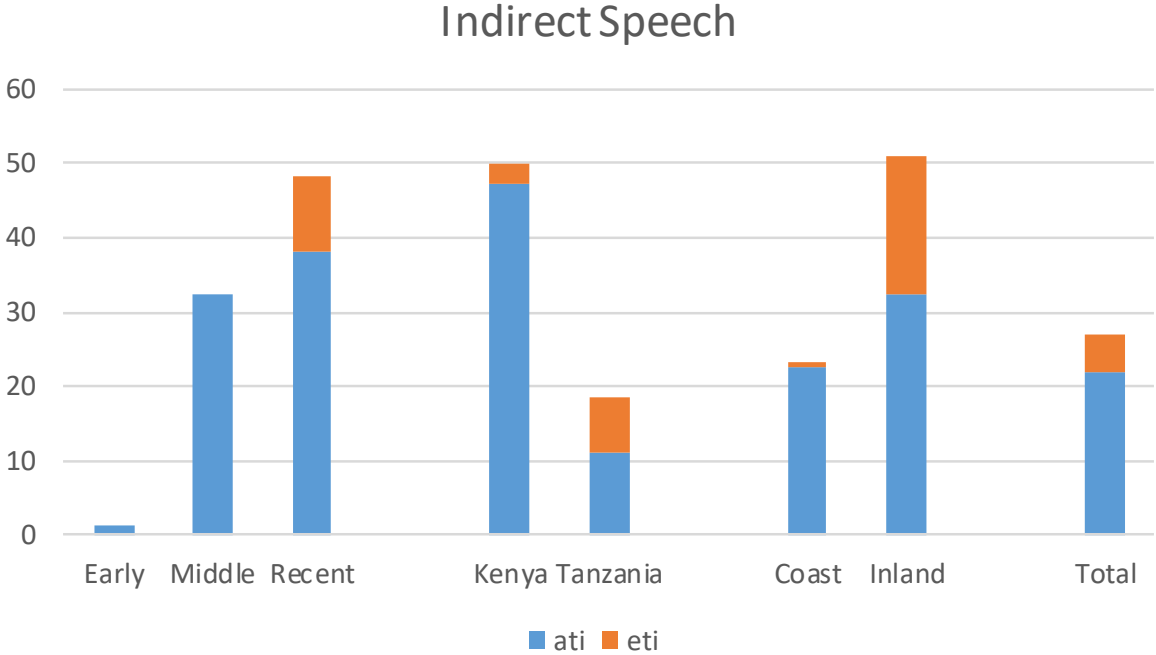


Figure 7. Number of *ati/eti* attestations in the subcategories of the corpus that are used as an indirect speech indicator, per million tokens.

In conclusion, the function of the discourse marker *ati/eti* to indicate indirect speech is more widespread in the recent period, Kenya and the inland regions. In Tanzania and inland, indirect speech occurs proportionately more frequently with *eti* than with *ati*, while the opposite applies to Kenya and the coastal areas.

6.1.2 Direct Speech

Another form of reported speech, besides indirect speech, is direct speech. When the speaker uses direct speech, or quotation, they report discourse by maintaining its original form. Contrary to indirect speech, the form of the verb and pronouns of the reported clause do not undergo any changes. In written language, the reported clause in direct speech is often indicated by quotation marks. A reporting verb can be used to introduce the reported clause, as in the following example.

- (18) **Ndipo alipowauliza hao mabwana zetu Waarabu, "Mwafanyani hapa magengeni?"** (Mbotela 1934, 54)

ndi-po	a-li-po-wa-uliz-a	ha-o	ma-bwana
COP-RC ₁₆	SP ₁ -PST-RC ₁₆ -OP ₂ -ask-FV	REF.DEM-RC ₂	NP ₆ -sir
zi-etu	wa-arabu	mu-a-fany-a-ni[ni]	ha-pa
PP ₁₀ -POSS _{1PL}	NP ₂ -Arab	SP _{2PL} -GEN.PRS-do-FV-what	PROX.DEM-PP ₁₆
ma-genge-ni			
NP ₆ -group-LOC			

"That is when he asked those Arab men, "What are you doing here in groups?"

In Swahili, a quotative index can be added to introduce the reported clause. Note that this is not the case in English, which uses 'that' only to introduce indirect speech, not direct speech. Possible direct speech indicators in Swahili are *kwamba*, *kuwa*, *kama (kwamba)*, *kana (kwamba)* and *ati* or *eti*. The use of a quotative index, however, is not compulsory, as illustrated by (18). Direct speech in Swahili is both possible with and without a quotative index. (19) and (20) are two examples from the corpus where *kwamba* and *kuwa* are used as direct speech indicators. (21) is an example of direct speech with quotative index *ati*.

- (19) **Aliambiwa kwamba "utakwenda kuyajulia mbele".** (Njama 1995, 84)

a-li-ambi-w-a	kwamba	u-ta-ku-end-a
SP ₁ -PST-tell-PASS-FV	COMP	SP _{2SG} -FUT-SM-go-FV
ku-ya-juL-i-a	mbele	
NP ₁₅ -OP ₆ -know-APPL-FV	in.front	

'He was told "you will learn them beforehand."'

- (20) **Alinikumbusha kabla ya kuondoka kuwa "Milima na milima haikutani bali wanadamu hukutana."** (Ole Kulet 1990, 58)

a-li-ni-kumbush-a	kabla ya	ku-ondok-a
SP ₁ -PST-OP _{1SG} -remember.CAUS-FV	before	CONN NP ₁₅ -leave-FV
kuwa mi-lima	na	mi-lima
COMP NP ₄ -mountain and	NP ₄ -mountain	NEG-SP ₄ -meet-RECP-NEG.PRS
bali	wa-anadamu	hu-kut-an-a
but	NP ₂ -human	HAB-meet-RECP-FV

'Before leaving he reminded me "Mountains do not meet, but people do."'

(21) **Nao walimjibu *ati* "Twataka mikoko tupeleke kwetu!"** (Mbotela 1934, 54)

na-o	wa-li-m-jibu	ati	tu-a-tak-a
and-RC ₂	SP ₂ -PST-OP ₁ -answer	PART	SP _{1PL} -GEN.PRS-want-FV
mi-koko	tu-pelek-e	ku-etu	
NP ₄ -mangrove	SP _{1PL} -send-SBJV	PP ₁₇ -POSS _{1PL}	

'And they answered him "We want mangroves to take home!"'

The examples of direct speech that are mentioned here all feature a reporting verb that introduces the quote. However, direct speech without a reporting verb is also possible. In fact, in the corpus used for this research, direct speech occurs more often without than with a reporting verb. Merely three out of ten times that direct speech in the corpus is indicated by *ati* or *eti*, the reported clause is preceded by a reporting verb ((21) being one of them). Most of the time, the reported clause is not introduced by a reporting verb. If direct speech occurs without a reporting verb, the direct speech indicator, in this case *ati/eti*, can be translated by 'she/he was like', as in the following example.

(22) ***Ati* "mimi nimekuja kuchapa kazi, sikuja kuendesha mapenzi."** (Mpinga 1984, 39)

ati	mimi	ni-me-ku-j-a	ku-chap-a	Ø-kazi
PART I	SP _{1SG} -PRS.PRF-SM-come-FV		NP ₁₅ -beat-FV	NP ₉ -work
si-ku-j-a		ku-end-esh-a	ma-penzi	
NEG.SP _{1SG} -PST.NEG-come-FV		NP ₁₅ -go-CAUS-FV	NP ₆ -love	

'He was like "I came here to work hard, not to make love."'

Just like with indirect speech, the use of *ati* or *eti* often adds an extra connotation to the meaning of the sentence, more specifically to how the reporter feels about the reported clause. Most of the time, this additional meaning conveys a feeling of contempt, which is apparent in half of the cases. This feeling of contempt is often expressed through mocking the reported person by repeating their statement in a sarcastic manner. In the following example, mockery and contempt are clearly present.

(23) **Unacheka nini, we Saleh? *Ati* 'hahaha' kama mpumbavu; nini unacheka!**

(Abdulla 1973, 6)

u-na-chek-a nini we Saleh ati hahaha kama m-pumbavu
SP_{2SG}-PRS.PROG-laugh-FV what you Saleh PART hahaha like NP₁-idiot
nini u-na-chek-a
what SP_{2SG}-PRS.PROG-laugh-FV

'What are you laughing at, you Saleh? You're like "hahaha" like an idiot; what are you laughing at?'

Regarding the position of the quotative index with respect to the quote, in African languages the quotative index (both for direct and indirect speech) usually precedes the quote. It is also possible for a language to possess both preposed and postposed quotative indexes (Güldemann 2008, 197–206). According to what we see in the corpus used for this research, this is not the case for Swahili. All quotatives in Swahili precede the quote. Not once do *ati* and *eti* occur in a place other than immediately before the quote.

Direct speech with *ati* or *eti* as a quotative index is not as prevalent as indirect speech introduced by these markers. In the entire corpus, it only occurs ten times. This corresponds to 4,57 attestations per million tokens. Compared with the 26,99 attestations per million tokens of *ati/eti* as indirect speech indicator, this is a very low number. There are no significant differences in the use of *ati/eti* between the various regions. The historical distribution does show some variation. Whereas in the early and the recent period there are only 1,05 and 5,59 occurrences of *ati/eti* direct speech indicators per million tokens, the middle period shows an increase to 16,18 per million tokens. However, this peak might be due to the limited size of that subcorpus. It can furthermore be observed that *eti* as a direct speech marker only occurs in the subcorpora of Tanzania, inland and the recent period.

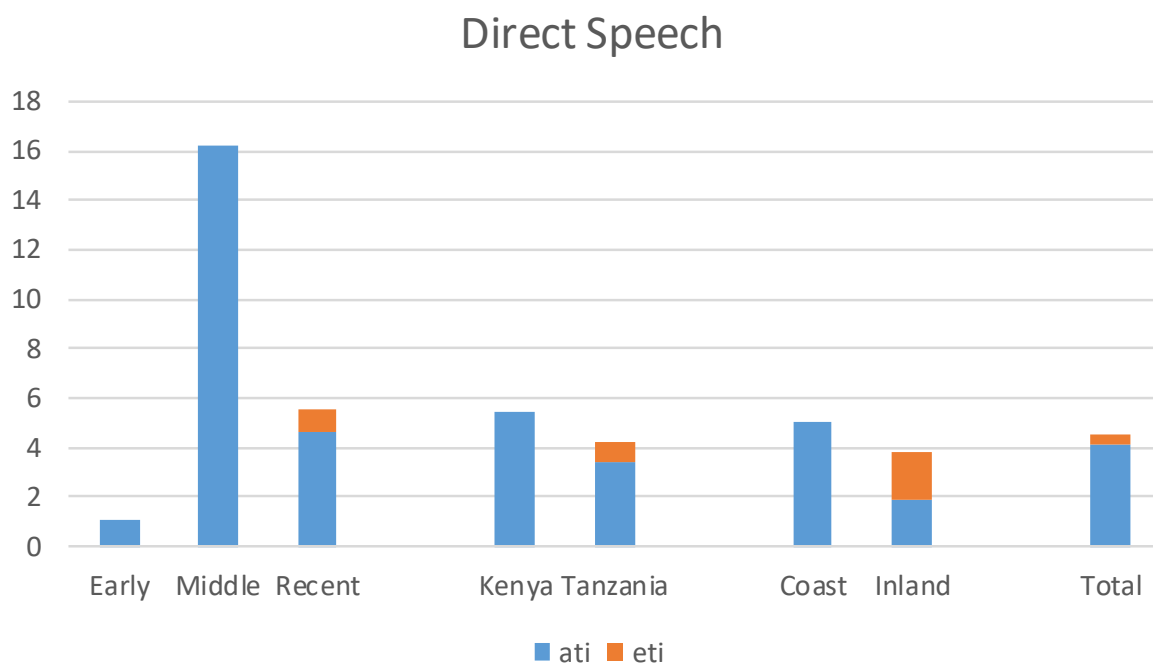


Figure 8. Number of *ati/eti* attestations in the subcategories of the corpus that are used as a direct speech indicator, per million tokens.

6.2 Clause linkage

A function of *ati* and *eti* not addressed in the literature is that of an adverbial subordinator. As mentioned in Section 6.1, adverbial subordinators are one of the three types of subordinators, along with complementizers and relative subordinators. Adverbial subordinators introduce subordinate clauses that function as adverbials (Nordström 2010, 94–95). Another definition of adverbial subordinators is given by Kortmann (1997, 56): "*free forms or bound adverbial morphemes which specify some semantic interclausal (or: circumstantial, adverbial) relation between the subordinate clause over which they operate and the modified matrix clause*".

Adverbial subordinators can be classified into different categories, according to meaning. Kortmann (1997, 84–86) lists 32 types of interclausal relations, among which cause and purpose. When *ati* and *eti* function as an adverbial subordinator, it is always to indicate one of these two types. A subordinate cause clause identifies a cause, reason or motivation for the proposition expressed by the matrix clause. The subordinator that is used to indicate this interclausal relation in English is 'because'. For instance: "*We had an accident because there was oil on the road*". The other interclausal relation marked by *ati* and *eti*, is purpose. In this case, the proposition expressed by the adverbial clause is an intended result or consequence of

the proposition expressed by the matrix clause, that is yet to be achieved. In English, the conjunction that indicates such an interclausal relation is 'in order to'. For example: "*I drove into London in order to buy a piano*".

The most common Swahili conjunctions to express these interclausal relations are *kwa sababu* or *kwa kuwa* (cause) and *ili* (purpose). (24) and (25) are some examples from the corpus to illustrate how these conjunctions are used. The verb in the adverbial clause introduced by *ili*, is often in the subjunctive mood, as in (25). However, this is not required, as *ili* can also be used before an infinitive. When expressing purpose by using a verb in the subjunctive or infinitive mood, *ili* may be omitted.

(24) [...] **Siti alikuwa kitandani bado kwa sababu alikuwa mgonjwa.** (Robert 1960, n.p.)

Siti a-li-ku-w-a ki-tanda-ni bado kwa Ø-sababu
 Siti SP₁-PST-SM-be-FV NP₇-bed-LOC still with NP₉-reason
 a-li-ku-w-a m-gonjwa
 SP₁-PST-SM-be-FV NP₁-sick
 'Siti was still in bed, because he was sick.'

(25) **Rais anapiga filimbi ili wajiandae.** (Kezilahabi 1999, 41)

Ø-rais a-na-pig-a Ø-filimbi ili
 NP₅-president SP₁-PRS.PROG-beat-FV NP₉-whistle so.that
 wa-ji-anda-e
 SP₂-REFL-prepare-SBJV
 'The president blows his whistle so that they make themselves ready.'

Another way to express an interclausal relation of cause, is by using *ati* or *eti*. To mark an interclausal relation of purpose, *ati* or *eti* can be used as well. In the corpus, *ati/eti* is used 47 times as (part of) an adverbial subordinator (21,50 per million tokens). When expressing an interclausal relation of cause, *ati/eti* can be used in combination with another conjunction or in itself. Out of the 35 times that *ati* and *eti* are used as the conjunction 'because' in the corpus, 24 of them (68,57 %) were in combination with *kwa sababu* or *kwa kuwa* (and once with *kwa maana* and *kwa ajili ya*). The combination with *kwa sababu* occurs twice as much as with *kwa kuwa*.

- (26) **Walimu wote walikata shauri ufukuzwe shule, eti kwa sababu ulitoa mimba.**

(Kezilahabi 1971, 52)

wa-alimu wa-ote wa-li-kat-a Ø-shauri u-fukuz-w-e
 NP₂-teacher PP₂-all SP₂-PST-cut-FV NP₆-advice SP_{2SG}-expel-PASS-SBJV
 Ø-shule eti kwa Ø-sababu u-li-to-a
 NP₉-school PART with NP₉-reason SP_{2SG}-PST-take.away-FV
 Ø-mimba
 NP₉-pregnancy

'All the teachers decided that you should be expelled from school, because you had an abortion.'

- (27) **Duniani hakuwa amepata fursa ya kuzuru maktaba yoyote ati kwa kuwa hakuwa amesoma sana.** (Mkangi 1995, 133)

Ø-dunia-ni ha-ku-w-a a-me-pat-a Ø-fursa
 NP₉-world-LOC NEG.SP₁-PST.NEG-be-FV SP₁-PRF-get-FV NP₉-chance
 i-a ku-zuru Ø-maktaba yo-i-ote ati kwa ku-w-a
 PP₉-CONN NP₁₅-visit NP₉-library RC₉-PP₉-any PART with NP₁₅-be-FV
 ha-ku-w-a a-me-som-a sana
 NEG.SP₁-PST.NEG-be-FV SP₁-PRF-read-FV very

'He had never gotten the chance to visit any library on earth, because he did not read much'.

In the previous examples, *ati* and *eti* precede the rest of the conjunction. However, this order can be reversed, although this only happens in a minority of cases (4 out of 24, or 16,76 %). Instead of *ati/eti kwa sababu*, you will get *kwa sababu ati/eti*, as in the following sentence. The internal position of *ati/eti* within the conjunction does not seem to have an influence on the meaning of the phrase.

- (28) **[...] hawakuwa wanauona umuhimu wa kuhudhuria misa kanisani kwa sababu eti walikuwa wakiiona kwenye televisheni zao ibada yote ya misa.** (Mpinga 1984, 64)

ha-wa-ku-w-a wa-na-u-on-a u-muhimu
 NEG-SP₂-PST.NEG-be-FV SP₂-PROG-OP₁₁-see-FV NP₁₁-importance
 u-a ku-hudhuri-a Ø-misa Ø-kanisa-ni kwa

PP₁₁-CONN NP₁₅-attend-FV NP₉-mass NP₅-church-LOC with
 Ø-sababu eti wa-li-ku-w-a wa-ki-i-on-a ku-enye
 NP₉-reason PART SP₂-PST-SM-be-FV SP₂-SIT-OP₉-see-FV PP₁₇-having
 Ø-televisheni zi-ao Ø-ibada i-ote i-a Ø-misa
 NP₁₀-television PP₁₀-POSS₂ NP₉-liturgy PP₉-all PP₉-CONN NP₉-mass
 '[...] they did not see the importance of attending the mass in church, because they
 were watching the whole liturgy of the mass on their televisions.'

In the other eleven instances (31,42 %) of *ati/eti* in the corpus that are used to express a causal relation between two clauses, the particle stands on its own and is not backed up by a conjunction that already has the meaning of 'because'. In this case, *ati/eti* therefore depends much more on the context to convey its meaning.

(29) [...] **ikiwa wao watajiona bora kuliko wengine, *ati* wao ni Wahindi, mimi nawaona wao ni watu kama watu wengine.** (Shafi 1999, 248)

i-ki-w-a wao wa-ta-ji-on-a bora kuliko
 SP₉-COND-be-FV they SP₂-FUT-REFL-see-FV better more.than
 wa-ngine *ati* wao ni wa-hindi mimi
 NP₂-other PART they COP NP₂-Indian.person I
 Ø-na-wa-on-a wao ni wa-tu kama wa-tu
 SP_{1SG}-PRS.PROG-OP₂-see-FV they COP NP₂-person like NP₂-person
 wa-ngine
 NP₂-other

'[...] if they see themselves as better than others because they are Indian, I see them as people who are just like the other people.'

Besides causal interclausal relations, *ati* and *eti* as subordinating conjunctions also express interclausal relations of purpose. When this occurs, *ati* and *eti* can be translated by 'in order to', 'in order that' or 'so that'. To avoid any possible confusion, *ati/eti* can be accompanied by the conjunction *ili* 'so that'. However, this is rather infrequent and appears only twice in the entire corpus. *Ati/eti* can both be preceded or followed by *ili*.

(30) [...] **kupigwa sindano ama kushindiliwa vijiti uzazini *ati* ili wasiweze kutunga mimba.** (Mkangi 1995, 106)

ku-pig-w-a Ø-sindano ama ku-shindili-w-a vi-jiti
 NP₁₅-beat-PASS-FV NP₉-needle or NP₁₅-press-PASS-FV NP₈-small.stick
 u-zazi-ni ati ili wa-si-wez-e ku-tung-a
 NP₁₁-womb-LOC PART so.that SP₂-NEG-can-SBJV NP₁₅-form-FV
 Ø-mimba
 NP₉-pregnancy

'[...] to get an injection or to be stabbed in the womb with little sticks, so that they cannot get pregnant.'

Usually, when *ati* or *eti* is used as a subordinator to express an interclausal relation of purpose, it appears without *ili*. In the majority of cases, *ati/eti* appears with a verb in the subjunctive mood, like in (31). In merely 18,18 % of the instances in which *ati/eti* is used to indicate purpose, it is followed by an infinitive, like in (32).

- (31) **Cynthia umemtorosha hapa *ati* asitahiriwe, wenzake watamwonaje?** (Walibora 2003, 177)

Cynthia u-me-m-torosh-a ha-pa
 Cynthia SP_{2SG}-PRS.PRF-OP₁-escape.CAUS-FV PROX.DEM-PP₁₆
 ati a-si-tahiri-w-e wa-enz-ake
 PART SP₁-NEG-circumcise-PASS-SBJV NP₂-companion-POSS₁
 wa-ta-mu-on-a-je
 SP₂-FUT-OP₁-see-FV-how

'You made Cynthia escape from here so that she would not get circumcised, how will her friends view her now?'

- (32) [...] **walimchukua Mbiju *ati* kumlea kuongeza idadi ya watumishi katika jumba lao.** (Suleiman Muhamed 1976, 22)

wa-li-m-chuku-a Mbiju ati ku-m-le-a ku-ongez-a
 SP₂-PST-OP₁-take-FV Mbiju PART NP₁₅-OP₁-raise-FV NP₁₅-increase-FV
 Ø-idadi i-a wa-tumishi katika ji-umba li-ao
 NP₉-amount PP₉-CONN NP₂-servant in NP₅-mansion PP₅-POSS₂

'[...] they took Mbiju in order to raise him to increase the amount of servants in their house.'

Because of the existence of other adverbial subordinators, *ati* and *eti* are not absolutely necessary to introduce the subordinate clause. This means that when this particle is present, it often adds an extra layer to the meaning of the sentence. In about one third of the attestations of *ati* and *eti* as an adverbial subordinator, it conveys an additional connotation. This extra connotation is mainly anger and contempt, sometimes also surprise. In the following example, the speaker expresses his anger towards the hearer, which is emphasized by *ati*, that is used as an adverbial subordinator.

(33) **Kuniamsha usiku wa manane *ati* nikueleze ujinga, ah!** (Ole Kulet 1990, 113)

ku-ni-am-sh-a	u-siku	u-a	manane
NP ₁₅ -OP _{1SG} -wake.up-CAUS-FV	NP ₁₁ -night	PP ₁₁ -CONN	midnight
<i>ati</i> ni-ku-ele-z-e	u-jinga		ah
PART NP _{1SG} -OP _{2SG} -be.clear-CAUS-SBJV	NP ₁₁ -stupidity		ah

'To wake me up in the middle of the night so that I explain you these stupidities, ah!'

The use of *ati* and *eti* as an adverbial subordinator increases with time. It goes from being absent in the corpus of the early period, over a presence of 5,39 per million tokens in the middle period, to 42,87 attestations per million tokens in the recent period. The use of *ati/eti* as an adverbial subordinator is higher in Kenya (32,45 per million tokens) and in the inland regions (30,31 per million tokens), as compared to Tanzania (19,37 per million tokens) and the coastal regions (22,42 per million tokens). However, comparing its use in the regional subcorpora in proportion with the frequency of *ati* and *eti* in said subcorpora, it can be observed that the geographical differences are minimal. In all regions, the percentage of *ati/eti* adverbial subordinators out of the total number of *ati/eti* attestations in the subcorpora lies around 12 %.

Comparing *ati* and *eti* to each other, *ati* is used more frequently as an adverbial subordinator than *eti*. *Ati* as an adverbial subordinator occurs 18,75 times per million tokens, while with *eti* it appears only 2,74 times per million tokens. Proportionately to the frequency of *ati* and *eti* in the corpus, it is still *ati* that is preferred over *eti* to introduce a subordinate clause (13,23 % of all *ati* attestations are used as an adverbial subordinator, compared to 9,09 of the *eti* attestations). The biggest difference is to be found in Kenya, where *eti* is never used as an adverbial subordinator, compared to 32,45 per million tokens for *ati*.

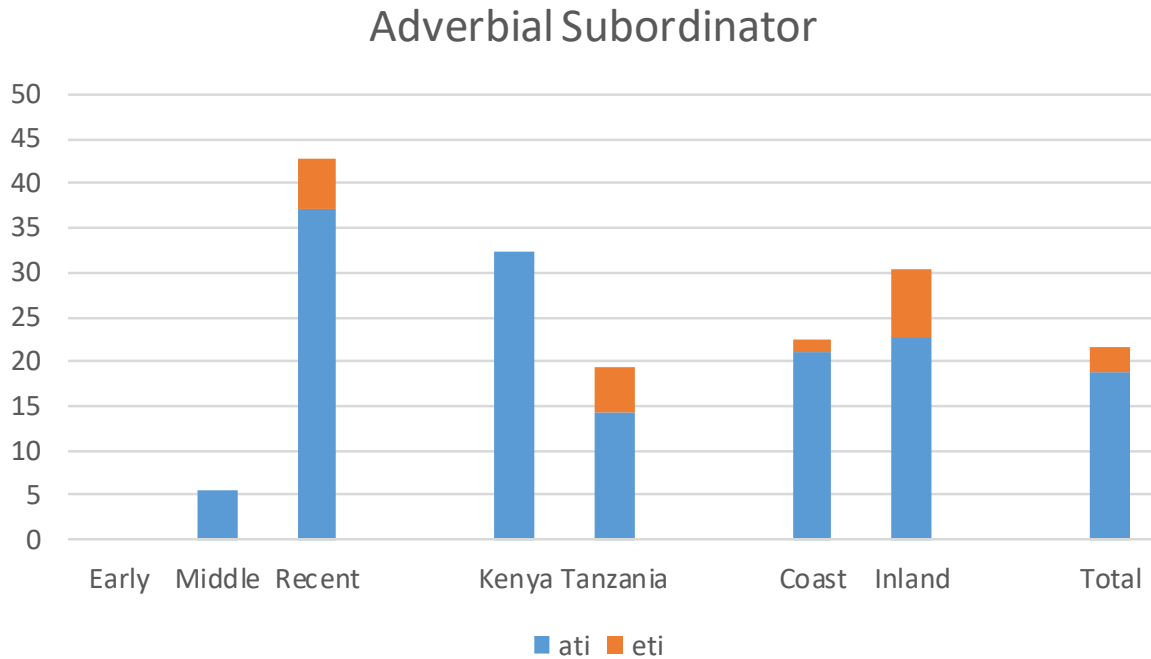


Figure 9. Number of *ati/eti* attestations in the subcategories of the corpus that are used as an adverbial subordinating conjunction, per million tokens.

As mentioned, there are 47 occurrences of *ati/eti* adverbial subordinators in the entire corpus (21,50 per million tokens). 35 or 74,47 % of those are used to express a causal relation, 12 or 25,53 % to indicate purpose. *Ati/eti* as 'because' is therefore far more frequent than *ati/eti* as 'in order to'. Moreover, there is a difference between *ati* and *eti*. While *ati* is used more frequently to express a causal relation (9,65 % of all *ati* attestations, compared to 6,06 % for *eti*), *eti* slightly has the upper hand as a conjunction to express purpose (3,03 % compared to 2,89 %).

6.3 Interjection

6.3.1 In general

Ati and *eti* are often used as interjections to express the attitude of the speaker towards the hearer or the ongoing discourse. This role of the discourse marker is very different from the functions that were discussed in the previous sections, in which *ati* and *eti* had a grammatical function (complementizer or adverbial subordinator).

Ameka (1992, 111) describes interjections as "a class of words which can stand on their own as utterances and which refer to mental acts." He continues that "[f]rom a pragmatic point of

view, interjections may be defined as a subset of items that encode speaker attitudes and communicative intentions and are context-bound" (Ameka 1992, 107). They thus express the speaker's mental state, action, attitude or reaction to a situation.

Furthermore, Ameka considers interjections as a type of discourse marker: "*In this approach interjections are a subclass of a large class of pragmatic markers (cf. Fraser 1990)*" (Ameka 1992, 107). Other scholars, however, believe that that is not the case. Norrick (2009, 889) argues that "*[i]nterjections are too complex and multifunctional to be sensibly listed among the specific classes of pragmatic markers.*"

According to Ameka (1992, 105–111), there are two types of interjections: primary and secondary interjections. Primary interjections are "*little words or non-words which in terms of their distribution can constitute an utterance by themselves and do not normally enter into construction with other word classes*" (Ameka 1992, 105). Examples of primary interjections in English include *Wow!*, *Gee!*, and *Oops!*. They can also co-occur with other utterances, as in *Gee, you look like you had it!* Secondary interjections, on the other hand, are "*those words which have an independent semantic value but which can be used conventionally as utterances by themselves to express a mental attitude or state*", for example *Help!*, *Hell!* or *Boy!* (Ameka 1992, 111). I suggest that *ati* and *eti* are to be included in the latter category of secondary interjections, because those forms also belong to other word classes based on their semantics (*ati* and *eti* as subordinating conjunctions, cf. Sections 6.1 and 6.2) and they originate in Proto-Bantu **ti* that has an independent semantic value 'say' (Güldemann 2008, 418–419).

A different classification proposed by Ameka (1992, 113–114) is based on the specific communicative functions that the interjections fulfil. He distinguishes three categories: expressive, conative and phatic interjections. Expressive interjections vocalise the mental state of the speaker with respect to the emotions or thoughts they are having at the time of utterance. In English, this category includes for example *wow* to express surprise and *ugh* to express disgust. Conative interjections are the ones that are directed towards the hearer and the purpose of which is to get someone's attention or to provoke a reaction on the part of the hearer, for instance *sh!* to ask for silence. Lastly, phatic interjections are used to establish and maintain communication, such as *mhm*, *uh-huh* or *yeah*.

Ameka (1992, 114) stresses however that a particular item may have multiple functions. This is certainly the case for *ati/eti*, which can both be categorised with the expressive and the conative interjections. As a conative interjection, *ati* and *eti* function as an attention-getting device (cf. Subsection 6.3.2.1). As an expressive interjection, they are used to express contempt, surprise, doubt or other emotions (cf. Subsections 6.3.3.1 to 6.3.3.4). Moreover, *ati* and *eti* as interjections can also be used to emphasize the utterance (cf. Section 6.3.4).

For a detailed overview of interjections in Swahili, see Eastman (1983), Eastman & Omar (1985) and Eastman (1992), who do not mention *ati* and *eti*.

6.3.2 Conative interjection

6.3.2.1 Attention-getting device

One of the commonly mentioned functions of the discourse marker *ati/eti*, especially in dictionaries, is calling attention. When *ati* and *eti* are used in this context, its purpose is to draw the hearer's attention towards the speaker. The discourse marker always occurs in the clause-initial position and is usually immediately followed by form of address, such as *bwana* 'sir', *bibi* 'madam', a title, name or personal pronoun. *Ati/eti* in this context can be translated as 'hey', an interjection in English that is used to attract attention.

(34) **KITARU: *Eti* Mfaume.**

MFAUME: Naam baba. (Hussein 1971, 29)

eti Mfaume naam baba

PART Mfaume yes father

'KITARU: Hey Mfaume.

MFAUME: Yes father.'

When *ati* or *eti* is used as an attention-getting device, it is often followed by an interrogative. The speaker wants to attract the hearer's attention and signal that they are about to ask a question. In addition, the use of *ati* and *eti* adds a connotation of doubt or disbelief to the question. For example in (35), the speaker himself does not really believe in the existence of devils. He expresses his doubts about the question by adding *eti*.

(35) ***Eti unaamini kama wapo mashetani?*** (Hussein 1971, 29)

eti u-na-amini

kama wa-po

ma-shetani

PART SP_{2SG}-PRS.PROG-believe like SP₂-LOC.COP NP₆-devil
 'Hey, do you believe that devils exist?'

As mentioned, when *ati/eti* occurs as an attention-getting device, it always takes the clause-initial position. However, there is one exception to this rule. There are some instances in which *ati* and *eti*, in combination with an interrogative, occur in the clause-final position and are preceded by the question. In this case, *ati* and *eti* function as answer-inciting devices. The speaker utters *ati* or *eti* to call the hearer's attention and to provoke an answer on the hearer's part.

(36) **Nini maana ya neno bikira ati?** (Kitsao 1983, 25)

nini Ø-maana i-a Ø-neno Ø-bikira ati
 what NP₉-meaning PP₉-CONN NP₅-word NP₉-virgin PART
 'What is the meaning of the word virgin, tell me?'

It is important to note that *ati* and *eti* are not interrogative particles, although they could be perceived as such because of their frequent collocation with interrogatives. Swahili has another (optional) interrogative particle, *je*, that is the only way to indicate a question, apart from intonation (and interrogative pronouns for constituent questions) (Nicolle 2000, 179). Unlike *je*, *ati/eti* does not have the capability to transform a declarative sentence into an interrogative sentence and can therefore not be classified as an interrogative particle.

Two other functions of *ati* and *eti* that are closely related with its function as an attention-getting device are introducing speech and signalling a topic shift in the ongoing discourse. The former is mentioned in some dictionaries, while the latter is absent from the literature. Both functions do not occur often in the corpus, but as they are closely linked to the function discussed in this section, they are worth mentioning.

When *ati/eti* is used for the introduction of speech, the speaker points out to the hearer that they are about to say something. This is very similar to the role of *ati/eti* as an attention-getting device, but they are not entirely the same. *Ati/eti* as an attention-getting device is used when the speaker does not have the attention of the hearer yet, while in this case the hearer is already listening to the speaker. The speaker was already talking, takes a break and then uses

ati or *eti* to signal the resumption of the speech act. This can be observed in the following example, in which the speaker briefly stops talking, before continuing on the same subject.

- (37) [...] (**anasimama na kutazama jikoni**) **Ati Dewe anajidai kuchukua madaraka ya jikoni hapa.** (Zani & Kitsao 1975, 58)

a-na-simam-a	na	ku-tazam-a	ji-ko-ni
SP ₁ -PRS.PROG-stand-FV	and	NP ₁₅ -watch-FV	NP ₅ -kitchen-LOC
ati	Dewe	a-na-ji-dai	ku-chuku-a ma-daraka
PART	Dewe	SP ₁ -PRS.PROG-REFL-claim	NP ₁₅ -take-FV NP ₆ -responsibility
ya-a	ji-ko-ni	ha-pa	
PP ₆ -CONN	NP ₅ -kitchen-LOC	PROX.DEM-PP ₁₆	

'[...] (he stands still and looks at the kitchen) So Dewe claims to take the responsibilities of this kitchen here.'

When *ati* or *eti* is used to signal a topic shift, the speaker already has the attention of the hearer and uses *ati/eti* to alert the hearer to a sudden change in the subject of the discourse. This function is linked to the role of *ati* and *eti* as an attention-getting device, because the speaker uses the discourse marker to draw the hearer's attention to the fact that there is going to be a change. Fraser (2009, 893-896) calls this type of discourse marker an attention marker. According to him, attention markers "*indicate that a topic change is in the making*" (Fraser 2009, 893). They help the speaker to avoid abruptness in the discourse. Fraser lists several attention markers in English, including 'alright', 'anyway', 'anyhow', 'hey' and 'so'.

- (38) **Nimeshtuliwa na bi-mkubwa wenu. Ati mna jambo muhimu sana mnalotaka kunizungumzia bila kuchelewa.** (Zani & Kitsao 1975, 3)

ni-me-shtuL-iw-a	na	Ø-bi	m-kubwa
SP _{1SG} -PRS.PRF-surprise-PASS-FV	and	NP ₅ -grandmother	NP ₁ -big
u-enu	ati	m-na	ji-ambo muhimu sana
PP ₁ -POSS _{2PL}	PART	SP _{2PL} -have	NP ₅ -thing important very
m-na-lo-tak-a		ku-ni-zungumz-i-a	bila
SP ₁ -PRS.PROG-RC ₅ -want-FV		NP ₁₅ -OP _{1SG} -talk-APPL-FV	without
ku-chelew-a			
NP ₁₅ -be.late-FV			

'I was surprised by the madam. Anyway, you have a very important issue you want to talk to me about without delay.'

Ati/eti as an attention-getting device is furthermore comparable to another Swahili particle: *nasema*. (Güldemann 2008, 418) writes that *nasema* "serves at the onset of a speaker's turn to direct the full attention of the hearer to him/herself. [...] it calls for attention to the fact that [an illocutionary] act is being made immediately." Interestingly, just like *ati/eti*, the particle *nasema* is also derived from a speech verb.

25,62 per million tokens in the corpus are attestations of *ati* and *eti* as an attention-getting device. In the subcorpus of the early period (19th century and early 20th century), *ati* and *eti*'s role of calling attention has a much lower frequency of only 6,29 attestations per million tokens. However, due to the low presence of *ati* and *eti* in the early period, this equates to 30,00 % of all the *ati/eti* attestations in the early subcorpus. This figure evolves over 5,39 in the middle period (4,35 %) to a sudden 45,66 per million tokens in the recent period (14,71 %). In absolute terms, the use of *ati* and *eti* as an attention-getting device has increased over time, but in proportion to the frequency of *ati* and *eti*, its function as an attention-getting device has been halved. It is furthermore remarkable that *eti* does not occur as an attention-getting device until the recent period.

The geographic distribution of this function of *ati* and *eti* is also interesting. Contrary to most of the other functions of this discourse marker, that appear most frequently in the inland regions, the attention-getting function of *ati/eti* is preferred at the coast. 31,11 per million tokens in the subcorpus of the coast is used to attract the hearer's attention, compared to only 24,63 in the inland regions. This is in spite of the low frequency of *ati* and *eti* in the subcorpus of the coast. Comparing the two countries to one another, it is once more Kenya that has the highest number of attestations. There are 35,15 attestations of *ati* and *eti* that are used as an attention-getting device per million tokens in the Kenyan subcorpus, compared to 25,27 in Tanzania. This seems like a substantial difference, but in proportion with the number of *ati/eti* attestations in those subcorpora, the distinction is not that big and Tanzania even slightly has the upper hand (15,63 % of the *ati/eti* attestations in the Tanzanian subcorpus are used as an attention-getting device, compared to 14,29 % in Kenya).

The function of attracting the hearer's attention is proportionately more frequent with *eti* than with *ati*, although the distinction is not that big. 16,67 % of all *eti* attestations in the corpus are used as an attention-getting device (5,03 per million tokens), compared to 14,52 % for *ati* (20,58 per million tokens). *Ati* as an attention-getting device is more prevalent in Kenya and the coast (respectively 32,45 and 26,77 per million tokens) and *eti* as an attention-getting device occurs more in Tanzania and inland (respectively 7,58 and 9,47 per million tokens). This can all be observed in the following graph.

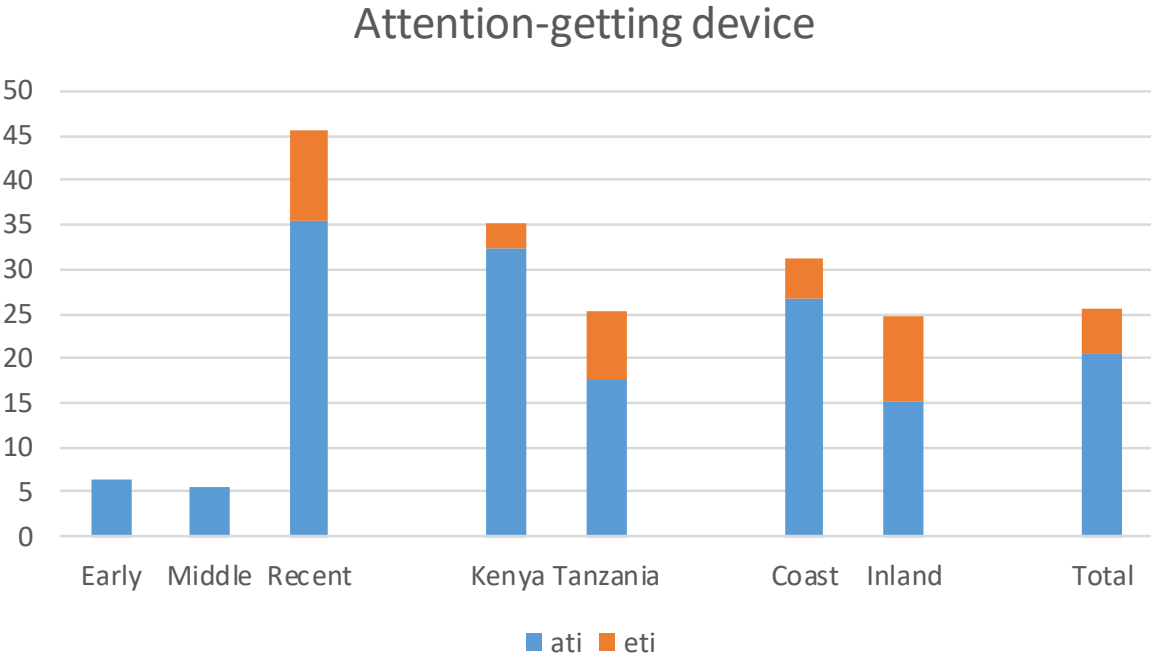


Figure 10. Number of *ati/eti* attestations in the subcategories of the corpus that are used as an attention-getting device, per million tokens.

In the majority of the cases in which *ati* and *eti* are used as attention-getting devices, they appear in combination with an interrogative. There are 19,21 attestations of this combination per million tokens in the corpus.

6.3.3 *Expressive interjection*

As mentioned, expressive interjections signal the speaker's attitude towards the content of the utterance and/or their relation towards the hearer. Their purpose is to express the speaker's feelings at the moment of the utterance and they have an interpersonal and conversational value (Haegeman & Hill 2013, 10–11). The most frequently occurring emotions that can be

conveyed by *ati* and *eti* (contempt, surprise and doubt) are described in the dictionaries. Three dictionaries mention surprise and contempt, two of them mention doubt. Additionally, one dictionary mentions disagreement (cf. p. 24). These can indeed all be found in the corpus.

At first glance, it might seem strange that one word can be interpreted in that many different ways. However, *ati/eti* has a core meaning that connects all the various emotions that it is able to convey. Fraser (1999, 931) writes that discourse markers "*have a core meaning, which is procedural, not conceptual, and their more specific interpretation is 'negotiated' by the context, both linguistic and conceptual.*"

This core meaning of *ati* and *eti* is contrast. All the meanings of *ati/eti* as an expressive interjection can be derived from this single notion of contrast. The core meaning of contrast is most apparent when the discourse marker is used to express an element of surprise. After all, surprise is caused by the unanticipated, by an event that contrasts with one's expectations. Doubt also expresses that the content of the utterance does not entirely correspond with the beliefs of the speaker. The core meaning of contempt essentially comes down to contrast as well, since that feeling is caused by a person or situation that contrasts with the speaker's own values. Even when *ati/eti* is used as an interjection to emphasize part of the discourse, which will be discussed in Subsection 6.3.4, this is connected to this core meaning, because the speaker contrasts part of the utterance with the rest of the discourse.

Dunn (1990, 15), who calls the core meaning the 'basic sense' of the discourse marker, argues that "[w]hen a speaker uses a word as a D[iscourse] M[arker] she mentions or evokes its basic sense, leaving it to the hearer to infer what she intends by its use at the point in the discourse in which she used it." The hearer subsequently interprets the intentions of the speaker by using the core meaning, in combination with the content of the utterance which it accompanies and the discourse up to that point (Dunn 1990, 8–9). The hearer is thus able to interpret which specific emotion (contempt, surprise or doubt) *ati/eti* expresses in a certain context, using its core meaning (contrast) as the starting point. Of course, the process of interpreting the exact meaning of the interjection is more subconscious and less calculated than described here.

The speakers of Swahili that use *ati/eti* are conscious of the emotions and connotations that this word bears. In fact, this interjection may even be used in a metalinguistic way. Although

there are not many examples of this to be found in the corpus, it is definitely worth mentioning. In the following example, the narrator uses the words *bila kutia "ati"* (literally 'without putting "ati"') to indicate that that the subject has no doubt about his feelings.

- (39) **Ingawa yeye mwenyewe Kasim alikuwa hajujui hasa ni nani, ila alijichukulia, kwa kweli hasa, bila kutia "ati" yo yote kuwa yeye ni mwana wa halali wa Bw. Hakimu.** (Abdulla 1973, 95)

ingawa	yeye	mu-enyewe	Kasim	a-li-ku-w-a				
although	he	NP ₁ -himself	Kasim	SP ₁ -PST-SM-be-FV				
ha-ji-ju-i			hasa	ni	nani	ila		
NEG.SP ₁ -REFL-know-NEG.PRS			exactly	COP	who	except		
a-li-ji-chukuL-i-a			kwa	kweli	hasa	bila	ku-ti-a	
SP ₁ -PST-REFL-take-APPL-FV			with	true	exactly	without	NP ₁₅ -put-FV	
ati	yo-i-ote	kuwa	yeye	ni	mu-ana	u-a		
PART	RC ₉ -PP ₉ -any	COMP	he	COP	NP ₁ -child	PP ₁ -CONN		
halali	u-a	bw.	Hakimu					
lawful	PP ₁ -CONN	Mr.	Hakimu					

'Although Kasim himself was not exactly sure who he was, he saw himself, in fact, without any doubt, that he was the legitimate son of Mr. Hakimu.'

It is important to bear in mind that interjections may have multiple functions. Depending on the context, *ati* or *eti* may have a number of meanings. This is for instance the case in the following example, in which the speaker uses *ati* to convey his anger and disagreement, while at the same time emphasizing an element of surprise.

- (40) **"Ati wanataka kumrudisha kwa mumewe," Mwajuma aliingilia kati.** (Shafi 1999, 109)

ati	wa-na-tak-a		ku-m-rudi-sh-a	kwa
PART	SP ₂ -PRS.PROG-want-FV		NP ₁₅ -OP ₁ -return-CAUS-FV	with
m-umewe	Mwajuma	a-li-ingi-li-a	kati	
NP ₁ -her.husband	Mwajuma	SP ₁ -PST-enter-APPL-FV	between	

'"Ah, they want to send her back to her husband," Mwajuma intervened.'

The use of *ati* and *eti* as expressive interjections is the most frequently occurring function of this discourse marker. In the entire corpus, there are 128 instances in which *ati/eti* is used as an expressive interjection (58,55 per million tokens), which is equal to 31,03 % of all the attestations of *ati* and *eti* in the corpus. The function of *ati/eti* as an expressive interjection is therefore far more popular than the other functions of this particle. Moreover, there are no significant differences in the geographical distribution of *ati/eti* as an expressive interjection. The historical distribution follows an upward trend from 5,24 over 59,32 to 104,37 *ati/eti* attestation per million tokens. In proportion to the frequency of *ati* and *eti*, this trend is also visible. 25,00 % of all *ati/eti* attestations in the early subcorpus are used as an expressive interjection, compared to 33,63 % in the recent subcorpus. There is a peak of 47,83 % in the middle period, but this can probably be attributed to the small size of that subcorpus.

There is a remarkable discrepancy between the two variations of the discourse marker. Relatively speaking, *eti* is used significantly more often as an expressive interjection than *ati*. This is not immediately visible by looking at the absolute numbers in Figure 11, but it does become very clear when one looks at the percentages. 48,48 % of all the *eti* attestations in the entire corpus are used as an expressive interjection, compared to only 30,97 % for *ati*. This function is the only one for which *eti* is significantly more popular than *ati*. This gap is the highest in Tanzania and the inland regions.

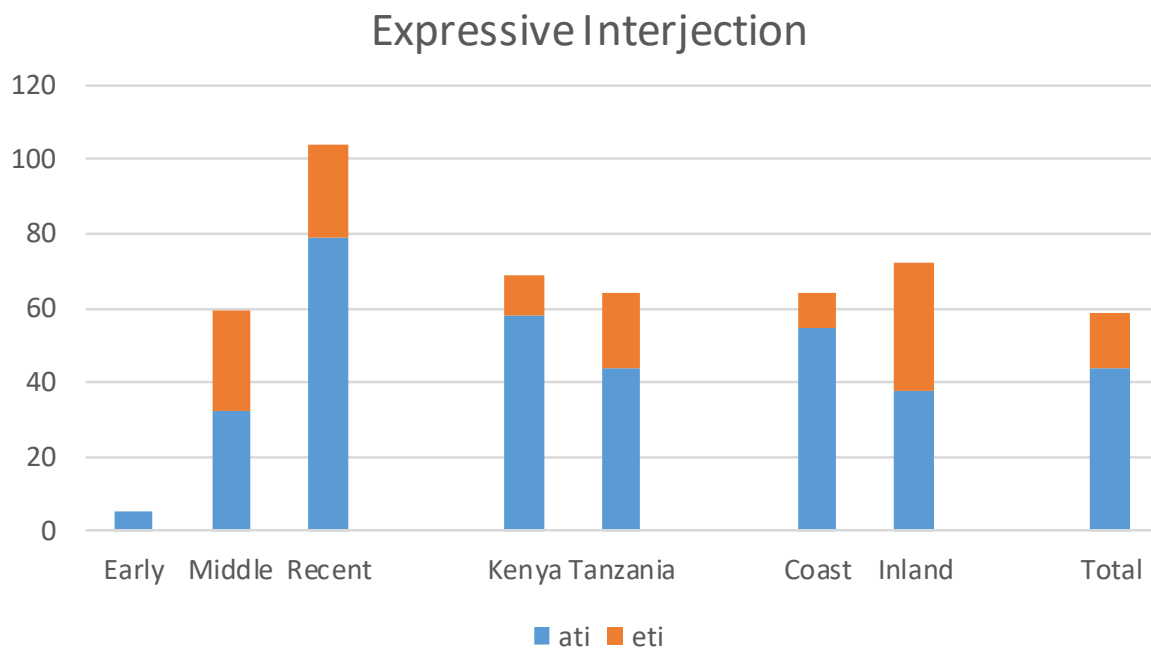


Figure 11. Number of *ati/eti* attestations in the subcategories of the corpus that are used as an expressive interjection, per million tokens.

The interjections mainly take the clause-initial position. 72,66 % of the times *ati/eti* is employed as an expressive interjection, it occurs as the first word of the clause. The clause-final and middle positions are less popular, with respectively 16,41 % and 10,16 % each.

The vast majority of cases in which *ati/eti* functions as an expressive interjection, it is used to express contempt (57,81 %). This is followed by surprise, which is conveyed by over a third of the *ati/eti* expressive interjections (35,16 %). This discrepancy in numbers is mainly caused by *ati*, where the expression of contempt is far more popular than surprise (62,50 % contempt compared to only 31,25 % surprise). For *eti* on the other hand, the frequency of expressing contempt and the frequency of conveying surprise are much more equal. In fact, *eti* as an expressive interjection is even used a little more frequently to express surprise than contempt (respectively 46,88 % and 43,75 %). Another emotion that is frequently expressed by *ati* and *eti*, although significantly less than contempt and surprise, is doubt. Doubt is expressed by 8,59 % of the attestations of *ati/eti* as an expressive interjection. Another 7,81 % conveys other emotions. These figures are visualised in the following chart. Note that one attestation of an expressive interjection can convey several emotions, as in (40), which is why the sum of the above percentages does not equal 100 %.

Emotions expressed by the interjections *ati* & *eti*

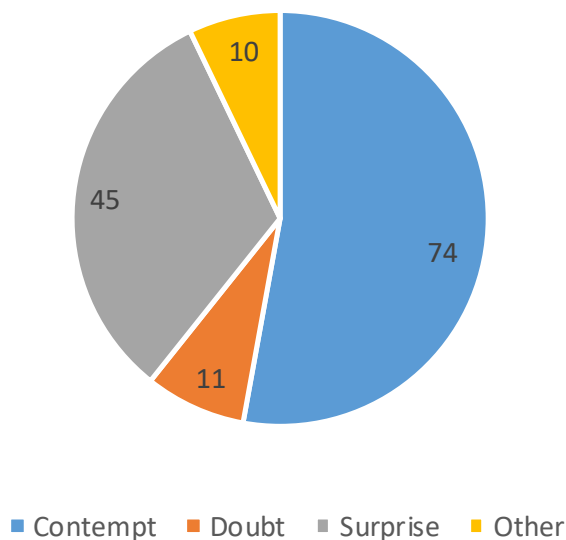


Figure 12. Distribution in the corpus of the emotions conveyed by *ati* and *eti*, when employed as expressive interjections. The data labels indicate the absolute number of *ati/eti* attestations expressing the corresponding emotions.

All the uses of *ati* and *eti* as an expressive interjection will be discussed in greater detail in Subsections 6.3.3.1, 6.3.3.2, 6.3.3.3 and 6.3.3.4.

6.3.3.1 Contempt

Contempt is the number one emotion that can be conveyed by *ati* and *eti*, accounting for 57,81 % of all the attestations of *ati* and *eti* used as expressive interjections in the corpus. As previously stated, *ati* is more frequently used in this context than *eti* (62,50 % compared to 43,75 %). Contempt is used here as an umbrella term that also includes anger, disagreement and annoyance. When *ati/eti* is used in this context, it is comparable to the English interjections *ha*, *bah*, *yech*, *huh*, *puh*, *tsk*, etc. (Jovanović 2004, 22–23).

The speaker's contempt can be referring to the interlocutor or to the discourse itself. In (41), the speaker's contempt is directed towards the discourse, as he expresses anger towards a third party that is not participating in the conversation. However, in (42) the speaker is expressing his contempt towards the hearers of the utterance, with whom he is engaged in an argument.

- (41) **"Ati wanataka wapewe nchi waongoze! Unafikiri wana wanaloliweza hawa? Hawawezi chochote!" Bwana Marjebi alisema kwa hasira.** (Shafi 2007, 35)

ati	wa-na-tak-a		wa-p-ew-e		N-chi
PART	SP ₂ -PRS.PROG-want-FV		SP ₂ -give-PASS-SBJV		NP ₉ -country
wa-ongoz-e	u-na-fikiri		wa-na		
SP ₂ -lead-SBJV	SP _{2SG} -PRS.PROG-think		SP ₂ -have		
wa-na-lo-li-wez-a		ha-wa		ha-wa-wez-i	
SP ₂ -PRS.PROG-RC ₅ -OP ₅ -can-FV	PROX.DEM-PP ₂		NEG-SP ₂ -can-NEG.PRS		
cho-ki-ote	Ø-bwana	Marjebi	a-li-sem-a	kwa	Ø-hasira
RC ₇ -PP ₇ -any	NP ₅ -sir	Marjebi	SP ₁ -PST-say-FV	with	NP ₉ -anger

"Ha, they want to be given a country to lead! Do you think they have what it takes, those people? They cannot do anything!" Mister Marjebi said angrily.'

- (42) **Mnalipwa mishahara minonominono ya bure na kazi hamfanyi; watoto wetu hawafaulu... na kisha mnalalamika ovyo... ati ikifika tarehe pili...** (Mohamed 1990, 21)

m-na-lip-w-a		mi-shahara	mi-nono-mi-nono	i-a
SP _{2PL} -PRS.PROG-pay-PASS-FV		NP ₄ -salary	NP ₄ -fat-NP ₄ -fat	PP ₄ -CONN
bure	na	Ø-kazi	ha-m-fany-i	wa-toto
free	and	NP ₉ -work	NEG-SP _{2PL} -do-NEG.PRS	NP ₂ -child
PP ₂ -POSS _{1PL}				
ha-wa-faulu	na	kisha	m-na-lalamik-a	ovyo
NEG-SP ₂ -succeed	and	then	SP _{2PL} -PRS.PROG-complain-FV	carelessly
ati	i-ki-fik-a		Ø-tarehe	Ø-pili
PART	SP ₉ -COND-arrive-FV		NP ₉ -date	NP ₉ -second

'You are getting paid a big fat salary for nothing and you do not even work; our children are failing... and then you are complaining for no reason... Hell, if the second day comes...'

The intrasentential distribution shows a clear pattern. The clause-initial position is by far preferred over the clause-internal and clause-final position. 67,57 % of the *ati/eti* attestations that are used as interjections to express contempt appear as the first word of the clause. 25,68 % take the clause-final position and 6,76% are to be found in the middle of the clause.

6.3.3.2 Surprise

Ati and *eti* can also be used as an interjection to express an element of surprise. In this context, the speaker uses *ati/eti* to express that a particular proposition has violated their expectations. This particular use of the discourse marker occurs significantly more with *eti* than with *ati*. 46,88 % of the *eti* attestations in the corpus employed as expressive interjections are used to express surprise, compared to only 31,25 % for *ati*. Other interjections in Swahili that express surprise are *kumbe* and *jamani*. Similar expressions in English include *wow*, *gee*, *gosh* and *yikes* (Goddard 2014, 57).

When *ati* and *eti* are expressing surprise, they often appear in exclamative clauses. Norrick (2009, 879) writes that '*[s]ince interjections are essentially expressions of sudden shifts in cognitive states, their occurrence with exclamative clause types might be expected.*'

(43) ***Ati lauliza kama kuna kifaranga cha kuku! Nilistaajabu nakwambia.*** (Zani & Kitsao 1975, 41)

ati	li-a-uliz-a	kama	ku-na	ki-faranga
PART	SP ₅ -GEN.PRS-ask-FV	like	SP ₁₇ -have	NP ₇ -young.bird
ki-a	Ø-kuku	ni-li-staajabu		
PP ₇ -CONN	NP ₉ -chicken	SP _{1SG} -PST-be.surprised		
Ø-na-ku-ambi-a				
SP _{1SG} -PRS.PROG-OP _{2SG} -tell-FV				

'Ha, he asked whether there are chicks here! I was surprised, I'm telling you.'

In the previous example, the speaker does not really pass judgement on the subject of the utterance (other than the fact that it is surprising to him). However, more often than not, the element of surprise is accompanied by a negative assessment of the discourse or the interlocutor. This was the case in (40). A positive connotation as in (44), on the other hand, is also possible with *ati/eti*, but appears notably less than a negative connotation.

(44) ***Ati! Nikauza mavuno yāngu sasa hivi, nikapata faida kubwa, modya kwa tanu.***⁴
(Sacleux 1930, 336)

ati	ni-ka-uz-a	ma-vuno	ya-angu	sasa	hi-vi
-----	------------	---------	---------	------	-------

⁴ Old orthography

PART SP_{1SG}-CONS-sell-FV NP₆-harvest PP₆-POSS_{1SG} now PROX.DEM-PP₈
 ni-ka-pat-a Ø-faida Ø-kubwa Ø-moja kwa Ø-tanu
 SP_{1SG}-CONS-get-FV NP₉-profit NP₉-big NP₉-one with NP₉-tanu
 'Wow! I sold my harvest just now and I made a huge profit, one fifth.'

Ati and *eti* can also be used in combination with the interrogative pronoun *nini* 'what'. Together they form the exclamation *ati nini?* or *eti nini?* It can be translated by 'Wait, what?' or 'What?!' This is a common exclamation to express a more sudden and intense feeling of surprise (more than when *ati* or *eti* stands on its own). There are fifteen attestations of *ati/eti nini?* in the corpus, which corresponds to almost four percent of all *ati/eti* attestations in the corpus (3,98 %). This exclamation only occurs in the recent period (1970s until 2010s).

(45) "***Ati nini?***" Sifuni aliniuliza kwa mshangao. (Chachage 2005, 141)

ati nini Sifuni a-li-ni-uliz-a kwa m-shangao
 PART what Sifuni SP₁-PST-OP_{1SG}-ask-FV with NP₃-surprise
 "'What?!' Sifuni asked in shock.'

Just like with contempt, when *ati/eti* expresses surprise it most likely occurs as the first word of the clause. 86,36 % of the *ati/eti* interjections expressing surprise take the clause-initial position. Furthermore, it seems to be impossible for *ati* or *eti*, when used to express surprise, to appear at the end of the clause, since this never occurs in the entire corpus.

6.3.3.3 Doubt

Another feeling that can be expressed by *ati* and *eti* is doubt. This occurs in 8,59 % of the cases in which *ati* and *eti* are used as an expressive interjection. It is used more often with *ati* than with *eti*: 9,38 % of the *ati* expressive interjections in the corpus are used to convey a feeling of doubt, compared to 6,25 % of the *eti* attestations.

When used in this context, the speaker takes away their responsibility for the authenticity of the proposition of the utterance. *Ati/eti* then expresses something along the lines of 'that's what they say, but I'm not sure whether I should believe it or not', like a verbal form of tracing quotation marks in the air with your fingers. Interjections that express doubt in English include *humph* or *hmm*. Otherwise, it could be expressed by adding the word 'allegedly' to the translation.

- (46) **Naye alidanganywa hivihivi na bwana mmoja toka pwani, eti yeye sijui Mkulugesi wa nini.** (Ngahyoma 1973, 8)

na-ye	a-li-dangany-w-a		hi-vi-hi-vi		na
and-he	SP ₁ -PST-cheat-PASS-FV		PROX.DEM-PP ₈ -PROX.DEM-PP ₈		and
Ø-bwana	m-moja	toka	Ø-pwani	eti	yeye
NP ₅ -sir	NP ₁ -one	from	NP ₉ -coast	PART	he
si-ju-i		m-kulugesi	u-a		nini
NEG.SP _{1SG} -know-NEG.PRS	NP ₁ -manager	PP ₁ -CONN			what

'And she was cheated on just like this by some man from the coast, he allegedly was the manager of I do not know what company.'

The position in the sentence seems to be of importance when expressing doubt since *ati/eti* never occurs at the end of the clause. Furthermore, it appears mostly as the first word of the clause: in 90,91 % of the instances in which *ati/eti* is an interjection expressing doubt. 9,09 % takes the clause-internal position.

6.3.3.4 Other emotions

Other than the three emotions that are most frequently expressed by the interjection *ati/eti* (contempt, surprise and doubt), there are some others that occur less, but deserve to be mentioned nonetheless. Among those are mockery or sarcasm, sadness and confusion. Their frequency in the corpus however is much lower and lies between 3,12 % and 1,56 % of all the expressive interjections. In the following example, *ati* expresses sadness.

- (47) **'Unanambia hivyo ati,' sauti yake Makuudi ilishaanza kuvuta kilio na macho yalishaanza kuiva machozi.** (Mohamed 1980, 107)

u-na-ni-ambi-a		hi-vyo	ati	Ø-sauti
SP _{2SG} -PRS.PROG-OP _{1SG} -tell-FV		REF.DEM-RC ₈	PART	NP ₉ -voice
i-ake	Makuudi	i-li-sha-anz-a	ku-vut-a	ki-lio
PP ₉ -POSS ₁	Makuudi	SP ₉ -PST-already-start-FV	NP ₁₅ -pull-FV	NP ₇ -cry
na	ma-cho	ya-li-sha-anz-a	ku-iv-a	ma-chozi
and	NP ₆ -eye	SP ₆ -PST-already-start-FV	NP ₁₅ -turn.red-FV	NP ₆ -tear

'"You are telling me like this" Makuudi's voice had already started to cry and his eyes started to turn red with tears.'

6.3.4 *Emphasis*

Ati and *eti* can also be used to emphasize a particular part of the discourse. I decided to not classify this use of the interjection under the three categories established by Ameka (1992), expressive, conative and phatic interjections, as it does not entirely fit into any of those categories. In this context, *ati/eti* is neither used to express emotions of the speaker (expressive), nor to provoke a reaction on the part of the listener (conative), nor to maintain the communicative contact (phatic).

When *ati/eti* is used in this manner, it occurs almost always in the clause-final position and it emphasizes the state of affairs that precedes this particle. The speaker pays extra attention to the material that was uttered immediately prior to *ati/eti*. It does not emphasize one particular word, as would be the case for focus particles, but the entire clause preceding the interjection.

This function of the interjection regards only the speaker's own concern with the utterance, without reference to a relationship with the hearer. *Ati* and *eti* is used by the speaker to focus on the linguistic material preceding the particle, but it does not per se instruct the hearer to pay extra attention to the utterance. The speaker does not explicitly expect or request acknowledgement from the hearer. Observe the following example:

(48) **Ngoja niliendee mimi mwenyewe. Litanitambua. Mimi ni "Regional Commissioner", *ati*.** (Mpinga 1984, 26)

ngoja-a	ni-li-end-e-e	mimi	mu-enyewe
wait-FV	SP _{1SG} -OP ₅ -go-APPL-SBJV	I	NP ₁ -self
li-ta-ni-tambu-a	mimi	ni	"Regional Commissioner"
SP ₅ -FUT-OP _{1SG} -recognise-FV	I	COP	regional.commissioner
<i>ati</i>			
PART			

'Wait, let me go to it [the giant] myself. It will recognise me. I am the Regional Commissioner!'

In this example, the speaker concludes his sentence with *ati* to put emphasis on the fact that he is a person with a high status and will therefore certainly be recognised by the giant. However, the speaker does not expect a reaction on the part of the hearer. By using *ati* the speaker merely expresses his certainty about the utterance.

This function of *ati* and *eti* occurs 29,28 times per million tokens in the corpus. *Ati* and *eti* for emphasis occur mostly in Kenya (44,61 per million tokens) and the inland regions (45,47 per million tokens), as compared to Tanzania (26,11 per million tokens) and the coastal areas (28,94 per million tokens). However, in proportion to the frequency of *ati* and *eti* in those subcorpora, the geographic distribution is more even. 18,13 % of the *ati/eti* attestations in the Kenyan subcorpus are used for emphasis, compared to 16,15 % for Tanzania. For the coastal and inland regions, this is respectively 16,19 % and 18,90 %.

Although the geographic discrepancies are minimal, there seems to be an evolution through time. In the early period, there are 7,34 attestations of *ati* and *eti* that are used for emphasis per million tokens in that subcorpus (35,00 % of all *ati/eti* attestations in the early subcorpus). This number increases to 53,12 in the recent period. This equates to only 17,12 % of the *ati/eti* attestations in that corpus, which means that relatively speaking the use of *ati* and *eti* for emphasis has decreased through time. In the middle period, the use of *ati/eti* for emphasis suddenly drops to zero, but this could be a coincidence due to the small size of that category of the corpus.

Moreover, *ati* is much more frequently used for emphasis than *eti*. 19,68 % of all the *ati* attestations are used for emphasis, compared to only 4,55 % for *eti*. The use of *eti* for emphasis is therefore very limited. This function is the only one for which *eti* remarkably occurs more frequently in Kenya (2,70 per million tokens) than in Tanzania (0,84 per million tokens). For all the other functions, the opposite is true and *eti* is used more frequently in Tanzania than in Kenya.

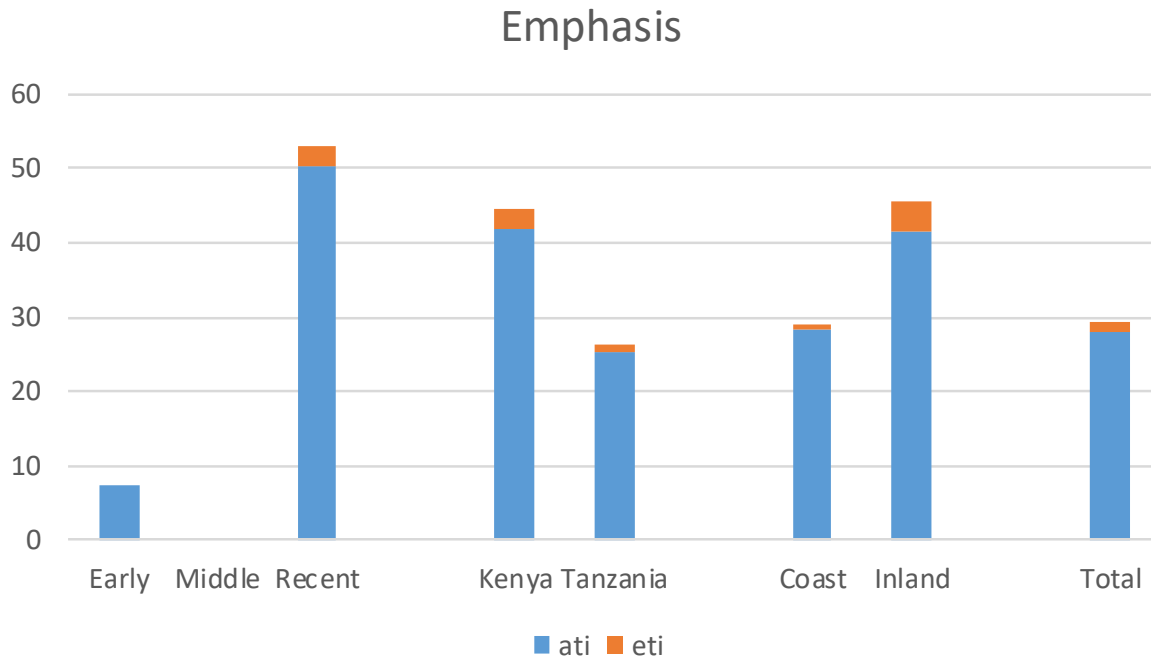


Figure 13. Number of *ati/eti* attestations in the subcategories of the corpus that are used for emphasis, per million tokens.

As already mentioned, the preferred position in the clause for *ati/eti* when used for emphasis, is the clause-final position. 96,88 % of the attestations of *ati* and *eti* used for emphasis appear as the last word of the clause. Since the particle always emphasizes the linguistic material prior to it, it goes without saying that *ati* and *eti* in this context never appear in the clause-initial position. The remaining 3,13 % takes a clause-internal position. If *ati* or *eti* occurs in the middle of the clause, it always highlights the preceding part of the sentence. The part that follows on the particle does not get special attention.

6.4 Connotation

As already shortly mentioned in some of the previous sections, *ati* and *eti* often carry an extra connotation in addition to their primary functions. Connotation and denotation are two kinds of meanings attached to a word. Stubbs (2001, 34–35) writes that denotation is a "*relation between a term in the language and a range of potential referents in the world.*" It is also referred to as cognitive, logical, ideational, propositional and conceptual meaning, or in everyday language the 'literal meaning' of a word. Connotation pertains to the emotional associations that convey the attitude of the speaker. It is sometimes thought of as personal, but this is not true since connotations are widely shared and commonly understood within speech

communities. Connotation is also called affective, associative, attitudinal and emotive meaning.

Some words mainly have a denotational meaning without many connotations. For other words, the connotational meaning is the most important while they express little denotational meaning. *Ati* and *eti* can be considered to belong to the latter group. As the previous chapters have made clear, it is hard to assign a clear and consistent meaning to this little word, as it does not refer to a specific referent in the extralingual world. Instead, it has a wide variety of different functions. Meanwhile there are various connotations that are connected to *ati* and *eti*.

The connotational meanings of words are often seen as secondary and have therefore not been extensively studied. Dictionaries often only focus on the denotational meaning. Work in corpus linguistics can be useful to delineate the connotational meaning of words. In the case of *ati* and *eti*, the dictionaries do address their emotive meanings, since the dictionaries consider their role as an interjection to be the primary (if not only) function of *ati/eti*. The connotations largely correspond to the emotions *ati* and *eti* convey as expressive interjections (cf. 6.3.3).

There are a number of various connotations linked to *ati* and *eti*. Those connotations match the emotions expressed by this discourse marker as an expressive interjection in terms of meaning, but not in terms of frequency. Once again, it is possible for one attestation of *ati* or *eti* to carry multiple connotations (e.g. contempt and surprise). Contrary to the expressive interjections, that mostly convey an emotion of contempt, the emotion that is most frequently expressed by *ati* and *eti* as a connotational meaning is doubt. In 47,00 % of the instances in which *ati* and *eti* carry a connotation, it is used by the speaker to express doubt about the content of the utterance. This includes for instance disbelief/doubt of the reporter about the reported clause, as in (10) and (16), and doubt regarding a question, as in (35). A connotation of doubt is more prevalent with *eti* than with *ati*: 63,64 % of the *eti* attestations that carry a connotation express doubt, compared to 42,31 % of the *ati* attestations.

The second most occurring connotation is contempt, expressed by 43,00 % of the *ati/eti* attestations that carry an additional meaning. To express this emotion, *ati* is used a little more frequently than *eti* (43,59 % for *ati* and 40,91 % for *eti*). An example of a connotation of

contempt can be seen in (33), in which the speaker uses *ati* to introduce a subordinate clause of purpose and at the same time to express his anger towards the hearer.

Another 24,00 % of the *ati/eti* attestations with an additional connotation is used to convey an element of surprise (23,08 % for *ati* and 27,27 % for *eti*). Observe the following example, in which the primary function of *eti* is to introduce the reported clause, but meanwhile it emphasizes the surprising aspect of said reported clause.

(49) [...] **anaimba *eti* mtangazaji wa redio katamka kwamba maana ya Zambe kwa Kilingala ni Mungu.** (Chachage 1991, 58)

a-na-imb-a	eti	m-tangazaji	u-a	Ø-redio
SP ₁ -PRS.PROG-sing-FV	PART	NP ₁ -announcer	PP ₁ -CONN	NP ₉ -radio
ka-tamk-a	kwamba	Ø-maana	i-a	Zambe kwa
SP ₁ .PRF-announce-FV	COMP	NP ₉ -meaning	PP ₉ -CONN	Zambe with
ki-lingala	ni	m-ungu		
NP ₇ -Lingala	COP	NP ₃ -god		

'[...] she is singing that the radio announcer said that the meaning of Zambe in Lingala is God!'

A few other possible connotations that can be expressed by *ati* and *eti* are sadness and mockery, but these do not appear often in the corpus.

There are at least 45,74 *ati/eti* attestations per million tokens in the corpus that carry an additional connotation (expressive interjections not included). I say 'at least', because I, as a non-native speaker, am not fully capable to grasp all the layers of the Swahili language and thus might have missed some of the connotations conveyed by *ati* and *eti*. These additional connotations seem to be far more present in Kenya than in Tanzania. In the Kenyan subcorpus there are twice as many *ati/eti* attestations with an additional emotive meaning than in the Tanzanian subcorpus: 74,36 compared to 37,90 per million tokens. This large discrepancy is also present when comparing the coast and inland. The inland subcorpus has more than twice as many *ati/eti* attestations that carry a connotation than the subcorpus of the coast: 83,36 compared to 40,51 per million tokens. There also seems to be a historical evolution in the occurrence of connotations attached to *ati* and *eti*. *Ati* and *eti* never carry an additional connotation in the subcorpus of the early period. In the middle period, there are 32,35 *ati/eti*

attestation with a connotation per million tokens. In the recent period, this number increases to 87,60 per million tokens.

Comparing the two variations of the discourse marker, connotations are relatively speaking far more frequently present with *eti* than with *ati*. 64,71 % of the attestations of *eti* carry an extra connotation additional to their primary function (10,06 *eti* attestations per million tokens in the corpus). In the case of *ati*, this number is significantly lower: 36,45 % (35,68 per million tokens). This high number of connotations for *eti* is especially notable in the subcorpora of Tanzania and the inland regions, as can be observed in Figure 14.

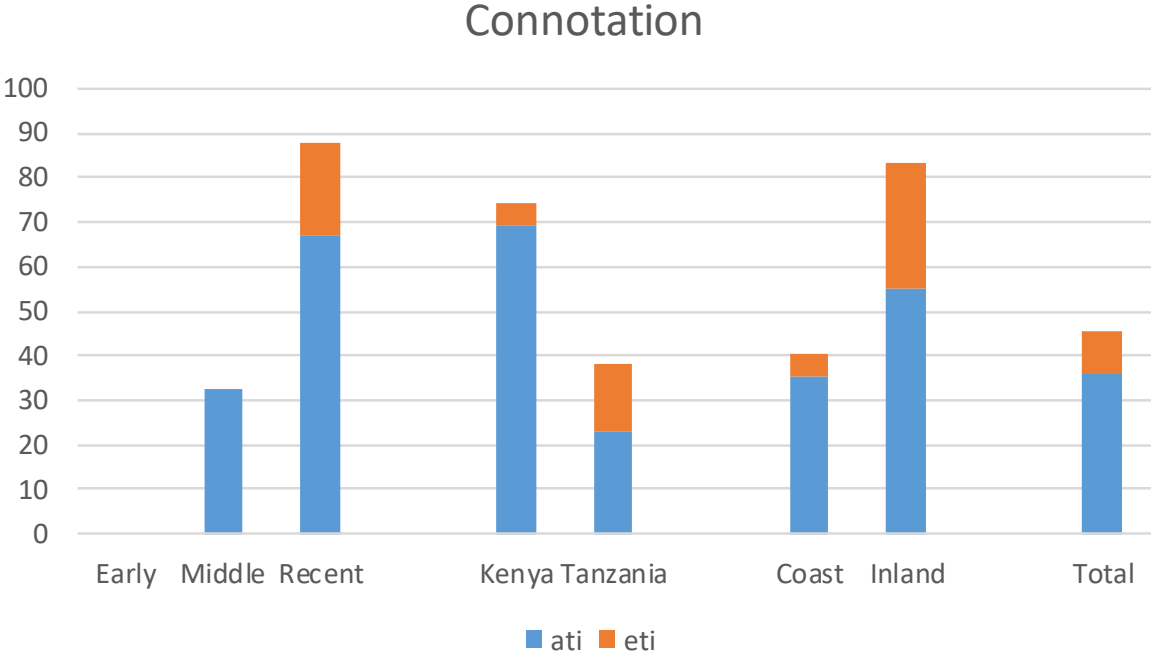


Figure 14. Number of *ati/eti* attestations in the subcategories of the corpus that carry an additional emotive connotation (expressive interjections excluded), per million tokens.

7 Syntactic aspects of *ati* and *eti*

Having defined the pragmatic meanings of *ati/eti*, I will now move on to discuss the syntactic aspects of this particle. Does the intrasentential distribution of the discourse marker have an impact on its meaning?

The markers *ati* and *eti* can take three positions: clause-initial, clause-internal and clause-final. Returning briefly to the frequency of each of these three possibilities, as discussed in Section 5.2, the clause-initial position is the most common. 71,81 % of all the attestations of *ati* and *eti* in the corpus are positioned at the start of the clause. 23,67 % appears in the clause-final position and only 3,99 % is located in the middle field of the clause.

Different from most other parts of speech, *ati/eti* can also stand alone, even in total absence of linguistic context. However, there are only two examples in the corpus. When *ati/eti* occurs isolated, it is always used as an interjection, either as an attention-getting device or as an expressive interjection to express surprise or contempt.

When *ati/eti* takes the clause-final position, it is always used as an interjection. This is not surprising, as it is not possible for adverbial subordinators and complementizers, the two other uses of *ati/eti*, to occur as the final word of a clause. The vast majority of these (69,66 % of all the clause-final occurrences of *ati* and *eti*) are used to emphasize the preceding discourse. 24,72 % are used as expressive interjections, almost all of them to convey an emotion of contempt. Another 6,74 % are used as attention-getting devices to incite an answer on the part of the hearer following a question.

When *ati/eti* occurs on the other end of the clause, in the clause-initial position, it is never used for emphasis. As the first word of the clause, *ati* and *eti* are mainly used as subordinating conjunctions: 30,00 % as a complementizer and another 17,41 % as an adverbial subordinator. *Ati* and *eti* as expressive interjections do also occur often in the clause-initial position (34,44 % of all the clause-initial *ati*'s and *eti*'s). Once again, the emotion that is most frequently expressed by them is contempt, which is expressed by 15,93 % of the clause-initial occurrences of *ati/eti*. However, this time it is followed closely by surprise (14,44 %). Doubt

is only expressed by 3,70 % of the clause-initial attestations of *ati/eti*. Furthermore, 18,15 % is used as an attention-getting device.

In the middle field of the clause, *ati/eti* is mainly used as an expressive interjection (86,67 % of the clause-internal attestations of *ati* and *eti*). Otherwise, when *ati/eti* occurs in the clause-internal position, it can also be used for emphasis. This is the case 13,33 % of the time.

As already noted, it makes sense that adverbial subordinators and complementizers only occur clause-initial. They link the clause they introduce to another clause, which requires them to follow the grammatical rules of that language. Interjections on the other hand, as their name suggests, can be 'thrown into' the sentence and can in principle occur anywhere in the clause, because they are syntactically independent of the other words in the sentence. However, it is not true that interjections are completely beyond the scope of syntactic rules: their meaning has an impact on their position in the clause. The position of *ati/eti* as an (expressive) interjection therefore deserves some extra attention. (Poggi 2009, 173–74)

According to Poggi (2009), interjections are subject to some very peculiar syntactic rules. She writes that "*their position with respect to a sentence is not completely free: their being uttered at the beginning, in the middle or at the end of a sentence, or finally as completely detached from the sentence itself, is determined by their meaning*" (Poggi 2009, 174). Poggi argues that interjections expressing surprise generally precede the clause containing their reference-element. This is true for *ati/eti*, that practically always appears in the clause-initial position and never in the clause-final position when expressing surprise. When expressing contempt, *ati/eti* is also mainly found in the clause-initial position. The expressions conveying an element of doubt, on the other hand, are preferably uttered in the middle of the sentence and are not acceptable in the sentence-final position, according to Poggi. This is only partially true for *ati* and *eti*. *Ati* and *eti* indeed never occur at the end of the clause or sentence when they are used to express doubt. Also true is that *ati/eti* expressing doubt often occurs in the middle of the sentence (36,36 %). Contrary to Poggi's claim however, this is not the preferred position. The sentence-initial position is more frequent (63,64 %). The clause-internal position is even less frequent, only occurring in 9,09 % of the attestations in which *ati/eti* expresses doubt.

It has been claimed that discourse markers "*require information focus to their right, and focussed elements to the left of the [discourse marker] have to be interpreted contrastively*" (Egg & Mursell 2017, 19). In the case of *ati* and *eti*, the observations suggest that this is the other way around. Disregarding *ati/eti* as a subordinator, *ati* and *eti* generally require information focus on their left (emphasis) and contrastive focus on their right (surprise and contempt).

8 Pragmatic distinctions between *ati* and *eti*?

Apart from the differences between the frequency of the two variants of the discourse marker, that has already been discussed in Section 5.1, there are also some slight differences in meaning between *ati* and *eti*. A more detailed account of the pragmatic distinctions between *ati* and *eti* is given in the following chapter.

All the various functions of *ati/eti* exist both with *ati* and with *eti*. A real difference in meaning between the two varieties therefore does not exist. Neither the literature on *ati/eti* nor the dictionaries mention a difference between the two. Most of the time, *eti* is merely seen as a variation on *ati*, the latter being the standard form. Nevertheless, there are some uses of the discourse marker that are more often used with *ati* and others with *eti*. The distinction is roughly as follows: when the discourse marker has a grammatical function in the sentence (complementizer or adverbial subordinator), it is more likely to be expressed by *ati*. When the discourse marker is used as an interjection (attention-getting device or expressive interjection), it is *eti* that proportionately has the upper hand (in absolute terms however, *ati* is still more frequent than *eti*). The one exception to this rule is when *ati/eti* is used as an interjection for emphasis, because in that case it is not *eti* but *ati* that is clearly more popular. Furthermore, when the discourse marker carries an additional emotive connotation, chances are higher that the speaker chooses *eti* over *ati*.

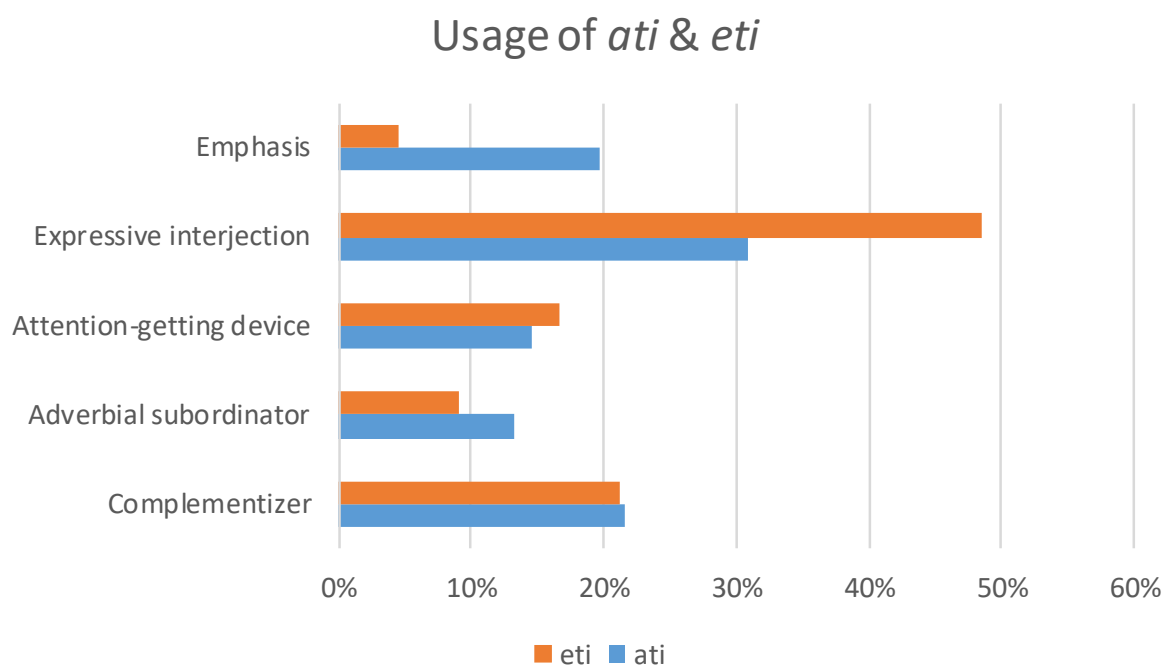


Figure 15. Percentage of the *ati* and *eti* attestations in the corpus that are used for emphasis, as an expressive interjection, attention-getting device, adverbial subordinator or complementizer.

These numbers illustrate the general trends, but do not point at concrete differences between *ati* and *eti*. One might think that these trends can be accounted for by the regional differences in frequency of the two varieties (*ati* is more frequently used in Kenya and in the coastal regions, while *eti* is more popular in Tanzania and in the inland regions). Although this might be part of the explanation, it is definitely not the full picture. If the distinctions between *ati* and *eti* were exclusively regional, the authors that are included in the corpus would be expected to choose one variety and consistently stick to that one option. However, this is not the case, as a number of texts in the corpus feature both *ati* and *eti*.

The texts in question are: Chachage 1991, Chachage 2005, Mkangi 1995, Zani & Kitsao 1975, Ole Kulet 1990, Walibora 2003, Mohamed 1990, Mpinga 1984 and Shafi 1999. They are all part of the recent subcorpus and the geographic distribution is more or less equal. In all of these texts *eti* is clearly used less frequently than *ati*, except for Chachage (inland Tanzania) where the distribution is almost fifty-fifty. Even within these texts, there are no consistent distinctions between the two varieties in the pragmatic use of the discourse marker. The general trends described in the second paragraph of this chapter are still present, so *ati* is more

used in a grammatical way and *eti* more in an emotive way. However, other than that it seems like the authors more or less selected *ati* and *eti* at random for the sole purpose of creating variation in their texts.

9 Diachronic scenario for the development of the polyfunctionality of *ati* and *eti*

In this short chapter, I attempt to reconstruct the path of development of *ati* and *eti*, from its Proto-Bantu origin as a manner deictic 'thus', as proposed by Güldemann (2002, 272), to all the present-day uses of *ati* and *eti* in Swahili, described in Chapter 6.

As mentioned in Section 4.4.3, Güldemann (2002, 278) assumes that Proto-Bantu **ti* was originally a manner deictic that, in spite of its non-verbal nature, was capable of taking a subject prefix. *Ati* might have emerged from this construction, *a-* being the person marker for the third person singular in Swahili. However, *ati* in Swahili is a petrified form that was never used in the way that Proto-Bantu **ti* functioned, i.e. as a manner deictic. Assuming that *ati*'s first vowel indeed is a subject marker, *ati*'s original meaning might have been 'he/she/it was like'. *Eti* would then be a later regional development derived from *ati*, as evident from the *ati/eti* ratio in the corpus according to period and region (cf. Figure 5), which illustrates that *eti* does not occur in the early period and is preferred in Tanzania and the inland regions.

Following Güldemann's schema for the grammaticalization of *ti* in Shona (cf. Figure 3), the original meaning of *ti* can be traced back to 'thus', which subsequently gave rise to some other grammatical and specified functions of this particle. I propose that the development of *ati* and *eti* in Swahili went down in a similar manner. The original meaning 'thus' in Proto-Bantu, that is not attested in present-day Swahili, gave rise to two new functions of *ati* and *eti* in Swahili. The first function is to embed mimetic expressions into the discourse. In Shona, this includes ideophones and sound imitations, but the only function of *ati* and *eti* in Swahili as a subdomain of mimesis is introducing direct reported speech or quotation. This in turn gave rise to the introduction of indirect speech, for which the degree of mimesis is already considerably less than is the case for direct speech. Because it is the least mimetic, indirect speech led to the emergence of two other grammatical functions of *ati* and *eti*: sentential complementation and (cause and purpose) clause linkage.

Güldemann (2008, 462–469), without mentioning Swahili, points out that clause linkage can be a functional extension of reported speech constructions. He argues that reported discourse constructions often are sources for a variety of other functions, such as clause linkage, but also naming, illocution reinforcement, internal awareness, etc. These linguistic expressions

share certain linguistic material with reported discourse constructions, but deviate from the functional aspects of reported speech. Güldemann (2008, 463–64) explains the similarities as follows:

"An essential requirement for the development of purpose clause linkage is the routinized combination of a main clause and an irrealis clause expressing the internal motivation/objective/purpose for the former. With respect to the use of R[eported] D[iscourse]-constructions in this domain, the quote becomes the dependent clause and the Q[uotative] I[ndex] or an important element thereof serves as the semantic and syntactic pivot between main and dependent clause. [...] The only structural adjustment required is that the predicative nucleus of the RD-construction, namely the QI-predicate, is conventionally linked to the main clause with the result that the entire modal expression is transformed into an adverbial clause."

This scenario for purpose clause linkage is found in a very similar way with reason clauses. Güldemann adds that the purpose and reason clauses in question are derived directly from the reported discourse construction and do not depend on a function as a general complementizer (Güldemann 2008, 467–68).

The second function that emerged from 'thus', more specifically from the demonstrative-introductory function of this word, is discourse deixis. An explicit deictic use of *ati* and *eti* in Swahili is not (or no longer) in use. The deictic function of *ti* to orient the hearer toward a particular piece of information within the discourse makes it a suitable device to emphasize states of affairs in the discourse (Güldemann 2002, 281). This is how the function of *ati* and *eti* in Swahili to emphasize the preceding discourse has evolved. In Shona, this emphasizing function was the source of a marked form of adverbial clause linkage, both to draw the hearer's attention to the following clause and to contrast a state of affairs against the information in an adjacent clause (Güldemann 2002, 263–65). In Swahili, *ati* and *eti* perform the same pragmatic functions of calling attention and expressing contrast. However, contrary to Shona, *ati* and *eti* perform these functions in the form of an interjection. *Ati/eti*'s function to express contrast has evolved to an expressive interjection that conveys an emotion of contempt, surprise or doubt.

In summary, my hypothesis suggests that Proto-Bantu **ti* 'thus' gave rise to two grammatical functions, mimesis and deixis. The former, mimesis, becomes increasingly less mimetic and is the source for all the grammatical functions of *ati* and *eti* in Swahili: sentential complementation (including reported speech) and clause linkage (purpose and cause). The latter, deixis, developed from a grammatical construction to lexical properties and led to the emergence of *ati* and *eti* as an interjection (emphasis, attention-getting device and expressive interjection). This scenario is illustrated in a schematic way in Figure 16. A note of caution is due here: since I do not have sufficient historical evidence, this scenario is merely hypothetical.

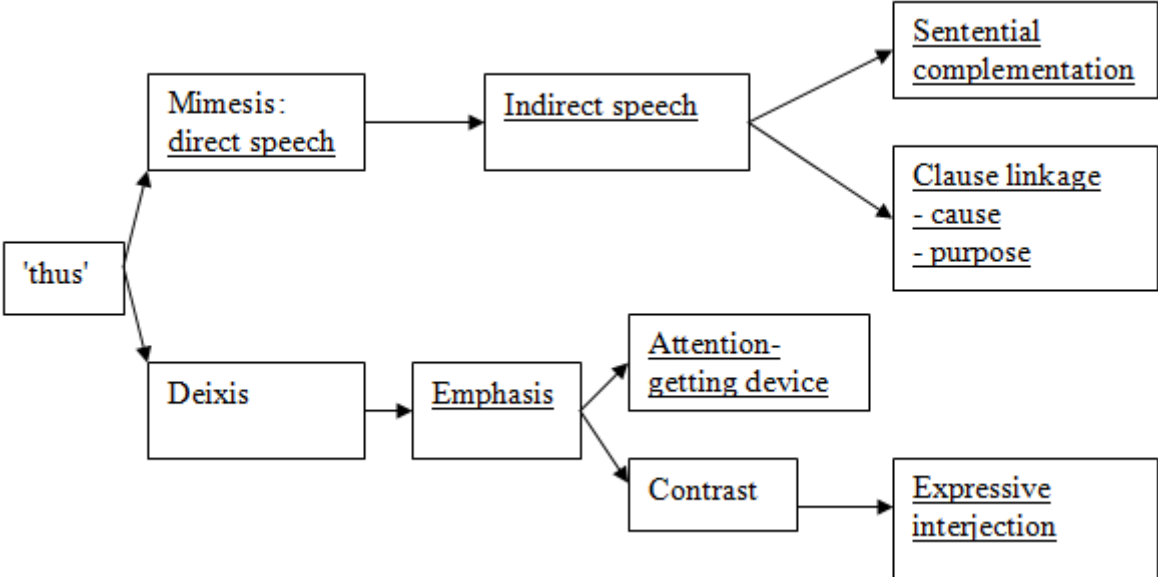


Figure 16. Diachronic development of the polyfunctionality of Swahili *ati* and *eti*. The underlined functions are the ones that are attested in the corpus.

This path of development is supported by diachronic corpus evidence. The historical distribution of the various functions of *ati* and *eti* in the corpus show in which time periods the various functions of the discourse marker are most prevalent. This is illustrated in Figure 17. I will only consider the early (19th and early 20th century) and recent (1970s until 2010s) period here, because the subcorpus of the middle period (early 20th century until 1960s) is too small (only 185 446 tokens) to contribute to a reliable conclusion.

Historical distribution of the functions of *ati* & *eti*

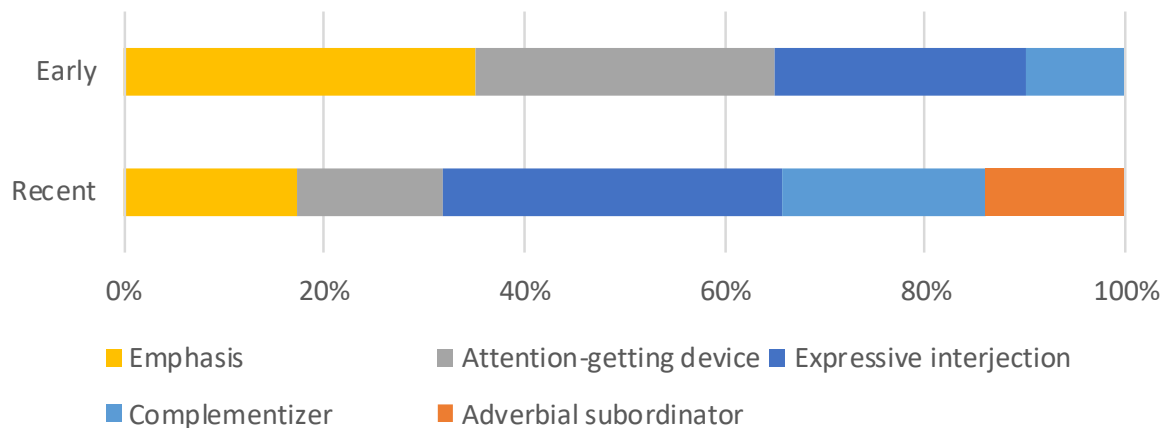


Figure 17. Distribution of the functions of *ati/eti* in the subcorpora of the early (19th and early 20th century) and recent period (1970s until 2010s).

Let us start with the bottom half of the schema in Figure 16, the interjectional uses of *ati* and *eti*. Figure 17 shows that the uses of *ati* and *eti* as a device to call attention and to emphasize the utterance were most popular in the early period, where they make up for respectively 30,00 % and 35,00 % of the *ati/eti* attestations in that subcorpus. Their frequency drops in the subcorpus of the recent period, to only 14,71 % for attention and 17,12 % for emphasis. *Ati/eti* as an expressive interjection on the other hand, is used less frequently than the other two functions in the early period, 25,00 % of all *ati/eti* attestation in that subcorpus, but its use increases considerably to 33,63 % in the recent subcorpus. This supports the hypothesis that attention and emphasis are among the older functions of the discourse marker, of which one (emphasis) later gave rise to the emergence of *ati* and *eti* as an expressive interjection to express contrast.

Let us now consider the upper half of the schema in Figure 16. The historical distribution of the grammatical functions of *ati* and *eti* (adverbial subordinator and complementizer) in the corpus show that these uses of the discourse marker were not widespread in the early period. *Ati/eti* as a complementizer (including reported speech) occurs only in 10,00 % of the *ati/eti* attestations in the early subcorpus and *ati/eti* as an adverbial subordinator does not occur at all. However, in the recent subcorpus, they account for almost 35 % of all the *ati/eti* attestations: 20,42 % for *ati/eti* as a complementizer and 13,81 % for *ati/eti* as an adverbial subordinator. With regard to *ati/eti*'s function as a complementizer, one can also observe that its use

becomes increasingly less mimetic over time. *Ati/eti*'s use as a direct speech indicator decreases (from 5,00 % of all *ati/eti* attestations in the early subcorpus to 1,84 % in the recent subcorpus), whereas its use as an indirect speech indicator and as a complementizer without reported speech function increases (respectively from 5,00 % to 15,62 % and from absent to 3,00 %).

The data in Figure 17 also point to a chronological order between the two branches of the schema in Figure 16. The interjectional uses of *ati* and *eti* (attention, emphasis and expressive interjection) make up 90,00 % of the *ati/eti* attestations in the early subcorpus. The grammatical functions (complementizer and adverbial subordinator) account for only 10,00 % in the early period. In the recent period, the frequency of *ati/eti* as an interjection dropped to 65,77 %, whereas the frequency of *ati/eti* as a grammatical particle increased to 34,23 %. This diachronic corpus evidence therefore suggests that the development of *ati/eti* as a grammatical particle (the upper branch in Figure 16) happened in a later stage than its interjectional functions (the bottom branch in Figure 16).

There are two approaches available in order to account for the connotational meanings of *ati* and *eti*. The first is a process of semantic bleaching, in which *ati/eti* used to have connotational meanings of contempt, surprise and doubt, but gradually lost these meanings over time. However, this approach is not supported by diachronic corpus evidence. The historical distribution in the corpus of *ati/eti* with an additional connotational meaning (absent in the subcorpus of the early period to 87,60 attestations per million tokens in the recent subcorpus) points towards the opposite. I suggest that the emergence of *ati* and *eti* as an expressive interjection lead to a semantic shift in the other uses of *ati* and *eti*. They kept their functions as adverbial subordinator, complementizer or attention-getting device, and additionally adopted the denotational meaning of *ati/eti* as an expressive interjection (i.e. contrast) as a connotational meaning. One could say that *ati/eti*'s meanings as an expressive interjection 'rubbed off' on its other functions. This is reflected in the historical distribution of *ati* and *eti* with a connotation in the corpus (cf. Figure 14).

10 Conclusion

The aim of the present research was to examine the pragmatic uses of the discourse marker *ati/eti* in Swahili. Employing both quantitative and qualitative methods, I described the various synchronic functions of *ati* and *eti* and provided a hypothesis for the diachronic development of this discourse marker.

The most obvious finding to emerge from this study concerns the multifunctionality of discourse markers. *Ati/eti* is a discourse marker with multiple uses depending on the context. This study has identified five main functions of *ati* and *eti*, that all developed from Proto-Bantu **ti* 'thus' and can be classified into two categories. On the one hand, Proto-Bantu **ti* 'thus' underwent a process of grammaticalization which gave rise to Swahili *ati* and *eti* with functions in clause linkage and sentential complementation. *Ati* and *eti* can be used as a complementizer (including in the context of reported speech) and as an adverbial subordinator (for cause and purpose clause linkage). In this category the use of *ati* is more widespread than *eti*. On the other hand, Proto-Bantu **ti* developed into Swahili *ati/eti* as an interjection. This includes the use of *ati* and *eti* for emphasis, to call attention and as an expressive interjection to convey an emotion of contempt, surprise or doubt. In this category it is *eti* that is, relatively speaking, more prevalent than *ati*, except when it is used for emphasis, in which case *ati* is used more frequently. The frequency of these functions can be observed in Figure 18.

A particular characteristic of *ati/eti* is that it often carries an additional connotational meaning. This connotation can be contempt, doubt or surprise; the same emotions that are conveyed by *ati/eti* as an expressive interjection. The denotations of *ati* and *eti* as an expressive interjection are thus present as a connotation in all the other uses of the discourse marker.

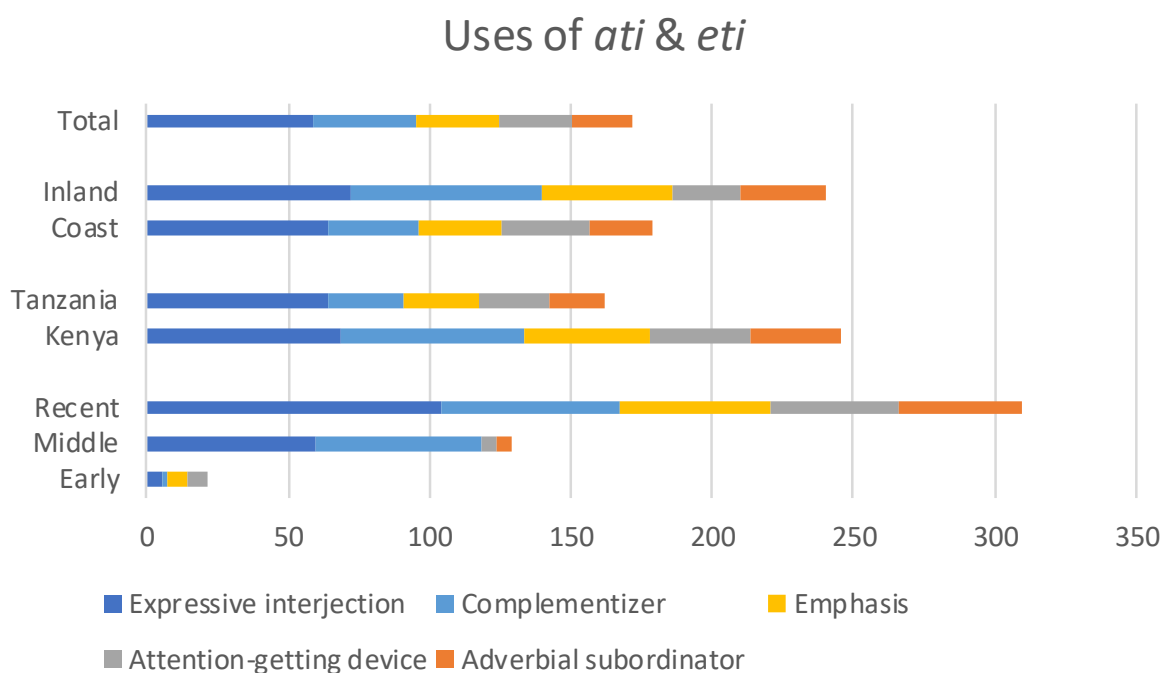


Figure 18. Number of *ati/eti* attestations in the subcategories of the corpus that function as an expressive interjection, complementizer, emphasizer, attention-getting device and adverbial subordinator, per million tokens.

The historical distribution of the attestations of *ati* and *eti* in the corpus show that the use of this discourse marker has increased over time. Moreover, the results from the corpus suggest that the use of *ati* and *eti* is more established in Kenya and in the inland regions, as compared to Tanzania and the coastal regions. Regarding a distinction between the two variants of the discourse marker, *ati* occurs much more frequently than *eti*, especially in Kenya and the coastal regions. However, *eti* seems to be winning ground and is used more and more frequently as time goes by.

The present study has been the first to thoroughly examine the use of *ati* and *eti* in Swahili. It contributes to our understanding of discourse markers in Swahili and the results add to the young but rapidly expanding field of Bantu corpus linguistics. However, this study was limited by the absence of oral data in the corpus (relevant given the fact that discourse markers appear mostly in spoken language).

Further research on *ati* and *eti* should therefore be carried out on a corpus of oral data. This could involve a comparative analysis between the use of discourse markers in spoken

language and in written language. Follow-up research that includes the judgements of native speakers is also needed to provide more definitive evidence for the findings that emerged from this study. A further study could also assess the sociolinguistic aspects of the use of *ati* and *eti*: who (age, gender, status, etc.) uses *ati/eti* and in what situations (formal, informal)? Lastly, a natural progression of this work would be to expand the research focus to include other discourse markers in Swahili, such as *mbona* 'why', *nasema* 'I say', *jamani* 'hey!', *je* 'well?', or *ewe* 'hey there'.

References

- Abraham, Werner, ed. 1991. *Discourse Particles: Descriptive and Theoretical Investigations on the Logical, Syntactic and Pragmatic Properties of Discourse Particles in German*. Amsterdam: Benjamins.
- Aijmer, Karin, and Anne-Marie Simon-Vandenberg. 2011. 'Pragmatic Markers'. In *Discursive Pragmatics*, edited by Jan Zienkowski, Jan-Ola Östman, and Jef Verschueren, 223–47. Amsterdam; Philadelphia: John Benjamins Pub. Co.
- Ameka, Felix. 1991. 'How Discourse Particles Mean: The Case of the Ewe "Terminal" Particles'. *Journal of African Languages and Linguistics* 12: 143–70.
- . 1992. 'Interjections: The Universal yet Neglected Part of Speech'. *Journal of Pragmatics* 18 (2): 101–18.
- Ariel, Mira. 1998. 'Discourse Markers and Form–Function Correlations'. In *Discourse Markers: Descriptions and Theory*, edited by Andreas Jucker H. and Yael Ziv, 223–60. *Pragmatics & Beyond New Series* 57. Amsterdam, Philadelphia: John Benjamins.
- . 2010. *Defining Pragmatics*. Cambridge: Cambridge University Press.
- Arndt, Walter. 1960. 'Modal Particles in Russian and German'. *Word* 16 (3): 323–36.
- Ashton, E. O. 1944. *Swahili Grammar*. London: Longmans.
- Barske, Tobias, and Andrea Golato. 2010. 'German so: Managing Sequence and Action'. *Text & Talk* 30 (3): 245–66.
- Bayer, Josef, and Volker Struckmeier. 2017. 'The Status Quo of Research on Discourse Particles in Syntax and Semantics'. In *Discourse Particles: Formal Approaches to Their Syntax and Semantics*, edited by Josef Bayer and Volker Struckmeier, 1–14. Berlin: Walter de Gruyter GmbH.
- Bazzanella, Carla. 1990. 'Phatic Connectives as Interactional Cues in Contemporary Spoken Italian'. *Journal of Pragmatics* 14 (4): 629–47.
- . 2006. 'Discourse Markers in Italian: Compositional Meaning'. In *Approaches to Discourse Particles*, edited by Kerstin Fischer, 449–64. Amsterdam; Boston; London: Elsevier.
- Biq, Yung-O. 1990. 'Conversation, continuation, and connectives'. *Text - Interdisciplinary Journal for the Study of Discourse* 10 (3): 187–208.
- Blommaert, Jan. 1997. *State Ideology and Language: The Politics of Swahili in Tanzania*. Duisburg: Gerhard-Mercator-Universität - Gesamthochschule Duisburg.

- Brinton, Laurel J. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin: Mouton de Gruyter.
- Bruti, Silvia. 1999. 'In Fact and Infatti: The Same, Similar or Different'. *Pragmatics* 9 (4): 519–33.
- Cadiot, Anne, Oswald Ducrot, Bernard Fradin, and Thanh Binh Nguyen. 1985. 'Enfin, marqueur metalinguistique'. *Journal of Pragmatics* 9 (2): 199–239.
- Comrie, Bernard. 1989. *The World's Major Languages*. Repr. with rev. and additions. London: Routledge.
- Cook, Haruko Minegishi. 1992. 'Meanings of non-referential indexes: A case study of the Japanese sentence-final particle ne'. *Text - Interdisciplinary Journal for the Study of Discourse* 12 (4): 507–39.
- De Schryver, Gilles-Maurice, and David Joffe. 2018. 'Swahili - English Dictionary'. *African Languages*. 2018. <https://africanlanguages.com/swahili/>.
- Degand, L., and B. Fagard. 2011. 'Alors between Discourse and Grammar: The Role of Syntactic Position'. *Functions of Language* 18 (1): 29–56.
- Denniston, John Dewar. 1934. *The Greek Particles*. Oxford: Clarendon press.
- Dér, Csilla, and Alexandra Markó. 2010. 'A Pilot Study of Hungarian Discourse Markers'. *Language and Speech* 53 (2): 135–80.
- Dunn, Andrea Susan. 1990. 'The Pragmatics of Selected Discourse Markers in Swahili'. *Ann Arbor (Mich.): UMI*.
- Eastman, Carol M. 1983. 'Exclamations in Standard Swahili as Cultural Communication'. *Journal of African Languages and Linguistics* 5 (2): 157–80.
- . 1992. 'Swahili Interjections: Blurring Language-Use/Gesture-Use Boundaries'. *Journal of Pragmatics* 18 (2): 273–87.
- Eastman, Carol M., and Yahya Ali Omar. 1985. 'Swahili Gestures: Comments (Vielezi) and Exclamations (Viingizi)'. *Bulletin of the School of Oriental and African Studies* 48 (2): 321–32.
- Egg, Markus, and Johannes Mursell. 2017. 'The Syntax and Semantics of Discourse Particles'. In *Discourse Particles: Formal Approaches to Their Syntax and Semantics*, edited by Josef Bayer and Volker Struckmeier, 15–48. Berlin: Walter de Gruyter GmbH.
- Fischer, Kerstin, ed. 2006. *Approaches to Discourse Particles*. Amsterdam: Elsevier.
- Fraser, Bruce. 1990. 'An Approach to Discourse Markers'. *Journal of Pragmatics* 14 (3): 383–98.
- . 1996. 'Pragmatic Markers'. *Pragmatics* 6 (2): 167–90.

- . 1999. ‘What Are Discourse Markers?’ *Journal of Pragmatics* 31 (7): 931–52.
- . 2006. ‘Towards a Theory of Discourse Markers’. In *Approaches to Discourse Particles*, edited by Kerstin Fischer, 189–204. Amsterdam; Boston; London: Elsevier.
- . 2009. ‘Topic Orientation Markers’. *Journal of Pragmatics* 41 (5): 892–98.
- Fuji, S. 2000. ‘Incipient Decategorization of Mono and Grammaticalization of Speaker Attitude in Japanese Discourse’. In *Pragmatic Markers and Propositional Attitude*, edited by Gisle Andersen and Thorstein Fretheim, 85–118. Amsterdam, Philadelphia: John Benjamins.
- Goddard, Cliff. 2014. ‘Interjections and Emotion (with Special Reference to “Surprise” and “Disgust”)’. *Emotion Review* 6 (1): 53–63.
- Golato, Andrea. 2010. ‘Marking Understanding versus Receipting Information in Talk: Achso and Ach in German Interaction’. *Discourse Studies* 12 (2): 147–76.
- . 2016. ‘Nu(n) in Standard German: Objections and Overcoming Objections’. In *NU/NÅ: A Family of Discourse Markers Across the Languages of Europe and Beyond*, edited by Peter Auer and Yael Maschler. *Linguae & Litterae* 58. Berlin, Boston: De Gruyter.
- ‘Google Translate’. n.d. Accessed 26 November 2018. <https://translate.google.com/?hl=nl#view=home&op=translate&sl=sw&tl=en>.
- Güldemann, Tom. 2002. ‘When “say” Is Not Say : The Functional Versatility of the Bantu Quotative Marker Ti with Special Reference to Shona’. In *Reported Discourse: A Meeting Ground for Different Linguistic Domains*, edited by Tom Güldemann and Manfred von Roncador, 253–87. Amsterdam: John Benjamins Publishing Company.
- . 2008. *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey*. Berlin, New York: Mouton de Gruyter.
- Guthrie, Malcolm. 1970. *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages*. Vol. 4. Farnborough: Gregg International Publishers Ltd.
- Haegeman, Liliane, and Virginia Hill. 2013. ‘The Syntacticization of Discourse’. In *Syntax and Its Limits*, edited by Raffaella Folli, Christina Sevdali, and Robert Truswell, 370–90. Oxford: Oxford University Press.
- Hakulinen, Auli. 1998. ‘The Use of Finnish *Nyt* as a Discourse Particle’. In *Discourse Markers: Descriptions and Theory*, edited by Andreas Jucker H. and Yael Ziv, 83–97. *Pragmatics & Beyond New Series* 57. Amsterdam, Philadelphia: John Benjamins.

- . 2016. 'The Word Ny(t) as an Adverb and a Particle in Finnish'. In *NU/NÅ: A Family of Discourse Markers Across the Languages of Europe and Beyond*, edited by Peter Auer and Yael Maschler. *Linguae & Litterae* 58. Berlin, Boston: De Gruyter.
- Hansen, Mosegaard Maj-Britt. 1998. 'The Function of Discourse Particles: A Study with Special Reference to Spoken Standard French'. Amsterdam: John Benjamins.
- Höftmann, Hildegard, and Stephen Mhando. 1963. *Suahili-deutsches Wörterbuch*. Leipzig: Verlag Enzyklopädie.
- Hölker, Klaus. 1991. 'Französisch: Partikelforschung'. In *Lexikon der Romanistischen Linguistik*, edited by G. Holtus, 77–88. Tübingen: Niemeyer.
- Imo, Wolfgang. 2010. 'The Adverb Jetzt ("Now") between Temporal and Conversational Deixis'. *Zeitschrift Für Germanistische Linguistik* 38 (1): 25–58.
- Johnson, Frederick. 1971. *A standard English-Swahili dictionary (founded on Madan's English-Swahili dictionary)*. Oxford: Oxford university press.
- Jovanović, Vladimir. 2004. 'The Form, Position and Meaning of Interjections in English'. *Facta Universitatis: Series Linguistics and Literature* 3 (1): 17–28.
- Jucker, Andreas, H., and Yael Ziv, eds. 1998. *Discourse Markers: Descriptions and Theory. Pragmatics & Beyond New Series 57*. Amsterdam, Philadelphia: John Benjamins.
- 'Kamusi Ya Kiswahili | Oxford Living Dictionaries'. n.d. Oxford Dictionaries | Kiswahili. Accessed 26 April 2019. <https://sw.oxforddictionaries.com/>.
- Kortmann, Bernd. 1997. *Adverbial Subordination: A Typology and History of Adverbial Subordinators Based on European Languages*. Berlin: Mouton de Gruyter.
- Krapf, Johann Ludwig. 1882. *A dictionary of the Suahili language*. London: Trübner and Co.
- Laakso, Mina, and Marja-Leena Sorjonen. 2010. 'Cut-off, the Particle Eiku and Other Practices for Initiating Self-Repair, and the Interactional Functions of Self-Repair'. *Virittaja* 109 (2): 244–71.
- Lenselaer, Alphonse. 1983. *Dictionnaire Swahili-Français : partiellement adapté du Standard Swahili-English dictionary de Frederick Johnson avec l'autorisation d'Oxford university press*. Paris: Karthala.
- Livnat, Zohar, and Ilil Yatsiv. 2003. 'Causality and Justification: The Causal Marker Ki in Spoken Hebrew'. *Revue de Semantique et Pragmatique* 13: 99–119.
- Ljungqvist, Marita. 2010. 'A Relevance-Theoretic Analysis of the Pragmatic Marker Ba in Mandarin Chinese'. *Journal of Chinese Linguistics* 38 (2): 261–87.
- Loogman, Alfons. 1965. *Swahili Grammar and Syntax*. Pittsburgh: Duquesne University Press.

- Madan, A. C. 1903. *Swahili-English Dictionary*. Oxford: Clarendon Press.
- Maho, Jouni Filip. 2009. 'NUGL Online: The Online Version of the New Updated Guthrie List, a Referential Classification of the Bantu Languages'. https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf.
- Maschler, Yael. 1997. 'Discourse Markers at Frame Shifts in Israeli Hebrew Talk-in-Interaction'. *Pragmatics* 7 (2): 183–211.
- . 2009. *Metalanguage in Interaction: Hebrew Discourse Markers*. Amsterdam: John Benjamins Pub. Co.
- . 2012. 'Emergent Projecting Constructions the Case of Hebrew Yada ('know')'. *Studies in Language* 36 (4): 785–847.
- Maschler, Yael, and Deborah Schiffrin. 2015. 'Discourse Markers: Language, Meaning, and Context'. In *The Handbook of Discourse Analysis. Vol. I*, edited by Deborah Tannen, Heidi Ehernberger Hamilton, and Deborah Schiffrin, 189–221. Malden: John Wiley & Sons, Inc.
- Massamba, David P.B. 1986. 'Reported Speech in Swahili'. In *Direct and Indirect Speech*, edited by Florian Coulmas, 99–119. *Trends and Linguistics: Studies and Monographs* 31. Berlin: Mouton de Gruyter.
- Matsumoto, Yo. 1988. 'From Bound Grammatical Markers to Free Discourse Markers: History of Some Japanese Connectives'. In *BLS Annual Meeting of the Berkeley Linguistics Society*, edited by Shelley Axmaker, Annie Jaissner, and Helen Singmaster, 14:340–51. Berkeley: University of California.
- Mauri, Caterina, and Anna Giacalone Ramat. 2012. 'The Development of Adversative Connectives in Italian: Stages and Factors at Play'. *Linguistics* 50 (2): 191–239.
- Maw, Joan. 1969. *Sentences in Swahili: A Study of Their Internal Relationships*. London: School of Oriental & African Studies.
- McGloin, Naomi H., and Yumiko Konishi. 2010. 'From Connective Particle to Sentence-Final Particle: A Usage-Based Analysis of Shi ("and") in Japanese'. *Language Sciences* 32 (5): 563–78.
- McKee, Rachel L., and Sophia Wallingford. 2011. "'So, Well, Whatever": Discourse Functions of Palm-up in New Zealand Sign Language'. *Sign Language and Linguistics* 14 (2): 213–47.
- Meeussen, A. E. 1980. *Bantu Lexical Reconstructions*. Tervuren: Koninklijk Museum voor Midden-Afrika.

- Mertens, Georges. 2006. *Dictionnaire kiswahili-français et français-kiswahili*. Paris: Karthala.
- Miller Shapiro, Carmit. 2012. 'Hamaba Naxonbasiax Ha'ivri Hayom-Yomi Hadavur'. *Balshanut 'Ivrit* 66: 107–30.
- Miracle, W. Charles. 1989. 'Hao: A Chinese Discourse Marker'. *Chicago Linguistic Society* 25: 213–27.
- Nicolle, Steve. 2000. 'Markers of General Interpretive Use in Amharic and Swahili'. In *Pragmatic Markers and Propositional Attitude*, edited by Gisle Andersen and Thorstein Fretheim, 173–88. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Nordström, Jackie. 2010. *Modality and Subordinators*. Amsterdam: Benjamins.
- Norrick, Neal R. 2009. 'Interjections as Pragmatic Markers'. *Journal of Pragmatics* 41 (5): 866–91.
- Nurse, Derek, and Thomas J. Hinnebusch. 1993. *Swahili and Sabaki: A Linguistic History*. Berkeley: University of California Press.
- Onodera, Noriko. 2004. *Japanese Discourse Markers: Synchronic and Diachronic Discourse Analysis*. Amsterdam: John Benjamin Publishing Company.
- Perez Hernandez, Yolanda. 2006. 'Manual Markers in Narrative Discourse in Venezuelan Sign Language'. *Letras* 1 (72): 157–208.
- Pick, Vittorio Merlo. 1964. *Vocabolario swahili-italiano e italiano-swahili*. Torino: Missioni consolata.
- Poggi, Isabella. 2009. 'The Language of Interjections'. In *Multimodal Signals: Cognitive and Algorithmic Issues*, edited by Anna Esposito, Amir Hussain, Maria Marinaro, and Raffaele Martone, 170–86. Berlin, Heidelberg: Springer.
- Polomé, Edgar C. 1967. *Swahili Language Handbook*. Washington: Center for applied linguistics.
- Rühlemann, Christoph. 2007. *Conversation in Context: A Corpus-Driven Approach*. London; New York: Continuum.
- Russell, Joan. 1992. 'From Reanalysis to Convergence: Swahili-Amba'. Edited by S.J. Harlow and A.R. Warner. *York Papers in Linguistics* 16: 121–38.
- Sacleux, Charles. 1939. *Dictionnaire Swahili-Français*. Paris: Institut d'ethnologie.
- Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.
- Sorjonen, Marja-Leena. 2001. *Responding in Conversation: A Study of Response Particles in Finnish*. Philadelphia: John Benjamins.

- Sperber, Dan, and Deirdre Wilson. 1986. *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Steere, Edward, and A. C. Madan. 1894. *A handbook of the Swahili language, as spoken at Zanzibar*. London: Society for Promoting Christian Knowledge.
- Stubbs, Michael. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- ‘Swahili-Engels Woordenboek’. n.d. Glosbe. Accessed 26 April 2019. <https://glosbe.com>.
- Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work. Studies in Corpus Linguistics 6*. Amsterdam: John Benjamins Publishing Company.
- TUKI. 2001. *Kamusi Ya Kiswahili-Kiingereza*. Dar es Salaam: Taasisi ya Uchunguzi wa Kiswahili, Chuo Kikuu.
- Vasko, I. 2000. ‘The Interplay of Hungarian *de* (but) and *Is* (Too, Either)’. In *Pragmatic Markers and Propositional Attitude*, edited by Gisle Andersen and Thorstein Fretheim, 255–64. Amsterdam, Philadelphia: John Benjamins.
- Visconti, Jacqueline. 2005. ‘On the origins of scalar particles in Italian’. *Journal of Historical Pragmatics* 6 (2): 237–61.
- . 2009. ‘From Textual to Interpersonal: On the Diachrony of the Italian Particle *Mica*’. *Journal of Pragmatics* 41 (5): 937–50.
- Wang, Yu-Fang, Pi-Hua Tsai, and Ya-Ting Yang. 2010. ‘Objectivity, Subjectivity and Intersubjectivity: Evidence from *Qishi* (‘actually’) and *Shishishang* (‘in fact’) in Spoken Chinese’. *Journal of Pragmatics* 42 (3): 705–27.
- Weydt, Harald. 1969. *Abtönungspartikel: die deutschen Modalwörter und ihre französischen Entsprechungen*. Bad Homburg: Gehlen.
- Whitehead, John, and L. F. Whitehead. 1928. *Manuel de Kingwana: le dialecte occidental de swahili*. Wayika: La Mission.
- Whiteley, Wilfred Howell. 1969. *Swahili: The Rise of a National Language*. London: Methuen & Co.
- Wouk, Fay. 1998. ‘Solidarity in Indonesian conversation: The discourse marker *kan*’. *Multilingua* 17 (4): 379–406.
- . 2001. ‘Solidarity in Indonesian Conversation: The Discourse Marker *Ya*’. *Journal of Pragmatics* 33 (2): 171–91.
- Ziv, Yael. 1998. ‘Hebrew *Kaze* as Discourse Marker and Lexical Hedge: Conceptual and Procedural Properties’. In *Discourse Markers: Descriptions and Theory*, edited by

Andreas Jucker H. and Yael Ziv, 203–22. *Pragmatics & Beyond New Series 57*. Amsterdam, Philadelphia: John Benjamins.

———. 2008. ‘Codifying Apparent Inconsistencies in Discourse: The Case of Hebrew Ma’. In *Generative Approaches to Hebrew Linguistics*, edited by Sharon Armon-Lotem, Gabi Danon, and Susan Rothstein, 353–88. Amsterdam: John Benjamins.