

SVEUČILIŠTE U ZAGREBU

FILOZOFSKI FAKULTET

Odsjek za anglistiku

Hermina Reicher

PRAGMATIC MARKERS IN ENGLISH AS A LINGUA FRANCA:

A CORPUS STUDY

Diplomski rad

Mentor: dr. sc. Mateusz-Milan Stanojević, doc.

Zagreb, listopad 2014.

UNIVERSITY OF ZAGREB
FACULTY OF HUMANITIES AND SOCIAL SCIENCES
Department of English

Hermina Reicher

PRAGMATIC MARKERS IN ENGLISH AS A LINGUA FRANCA:

A CORPUS STUDY

Diploma thesis

Advisor: Assistant Professor Mateusz-Milan Stanojević

Zagreb, October 2014

Povjerenstvo:

Committee in charge:

dr. sc. Višnja Josipović Smojver, red prof. voditelj

Professor Višnja Josipović Smojver, Chair

dr. sc. Anđel Starčević, asistent

Teaching Assistant Anđel Starčević

dr. sc. Mateusz-Milan Stanojević, doc.

Assistant Professor Mateusz-Milan Stanojević

Contents

1 Introduction.....	1
2 Theoretical framework.....	4
2.1 Subjectivity, intersubjectivity and grammaticalization.....	4
2.2 Past research on <i>like</i>	7
3 Methodology.....	14
4 Results.....	15
4.1 Classification of the pragmatic marker <i>like</i>	15
4.1.1 Hedges.....	17
4.1.2 Illustratives.....	20
4.1.3 Quotatives.....	22
4.1.4 Fillers.....	23
4.1.5 Independent use of <i>like</i>	25
4.2 <i>Like</i> in numbers.....	27
4.2.1 Hedges in numbers.....	30
4.2.2 Illustratives in numbers.....	31
4.2.3 Quotatives in numbers.....	33
4.2.4 Fillers in numbers.....	34
4.2.5 Independent use of <i>like</i> in numbers.....	35
5 Discussion and conclusion.....	37
6 Literature.....	40
7 Appendix.....	42

1 Introduction

The basic meaning of the English word *like* is to signal comparison, as is evident from a variety of dictionaries of English. However, it is also multi-functional, and in addition to its basic propositional function, *like* may also serve a variety of pragmatic functions. Numerous authors dealt with pragmatic functions of *like*, for instance, Underhill (1988) in *Like Is, like, Focus* (1988), Blyth, Recktenwald and Wang (1990) in *I'm like, "Say What?!": A New Quotative in American Oral Narrative*, Andersen (1998) in *The Pragmatic Marker Like From a Relevance-theoretic Perspective*, Jucker and Smith (1998) in *And people just you know like 'wow'. Discourse Markers as Negotiating Strategies*, and others, which will be discussed later in more detail.

Pragmatic markers normally develop by the process of (inter)subjectification and grammaticalization, which some authors also refer to as the pragmaticization. Mulder and Thompson (2008, 198) in their study of grammaticalization of *but* as a final particle in English conversation define pragmaticization or ‘pragmatic particle-ization’ as a process similar to grammaticalization, in which the morphemes in question retain both lexical and grammatical functions in the early stages of change. In this process, their meaning increasingly loses its ‘objective’ propositional content, becoming more pragmatic and intersubjective. For example, this happens when *like* is no longer used as a preposition which serves to compare two types of content, but, for instance, as an introducer of reported speech:

➤ *Example 1*¹

¹ The exact spelling from the Vienna Oxford International Corpus of English (VOICE) corpus was used in the examples provided in this paper. It follows a specific set of notational conventions, such as no capital letters except for marking emphasis, no commas, words spoken with rising intonation are followed by a question mark, words spoken with falling intonation are followed by a full stop, pauses are marked with a full stop in parentheses, and so on (VOICE Project. 2007. VOICE Transcription Conventions [2.1]. http://www.univie.ac.at/voice/voice.php?page=transcription_general_information. 15/6/2014).

they needed people so badly that they were like (.) hey you can come on the night shift and i was like (.) but i'm not educated i don't want to.

These instances of *like* exhibit distancing from the primary meaning and a case of grammaticalization in progress. In this example *like* does not signal an objective, but a subjective comparison to something that was said (or thought).

It is reasonable to assume that, because of the specific characteristics of ELF discourse, the use of *like* as a pragmatic marker might differ in native English discourse and English as a Lingua Franca. Such pragmatic differences were found in the analysis of collocations *I think* and *I don't know* as markers of stance-taking by native and non-native speakers of English in L1 and English as a Lingua Franca interaction by Baumgarten and House (2010). Since “in ELF communication, speakers of different L1s with potentially differing conventionalized patterns of stance-marking and stance-taking and differently diversified L2 varieties of English interact”, it is possible that the frequency of stance-taking in ELF discourse differs from the communicative style found in comparable native English interactions (Baumgarten and House 2010, 1184). Culture- and community specific conventions regulate the linguistic pattern of each speaker, so, in comparison with L1 speakers, the individual participants in ELF interaction “operate on a much smaller common ground of mutually expectable communicative behaviour” (Baumgarten and House 2010, 1185). Baumgarten and House (2010) found that the L2 speakers tend to use the collocations more frequently in their prototypical meaning, whereas the L1 speakers use the expressions in a predominantly hearer-oriented way, i.e. in their grammaticalized and pragmaticalized form. Since the L2 speakers operate on the basis of a reduced shared knowledge, the grammaticalized forms seem to be “not as immediately present for the L2 speakers as they are for the L1 speakers” (Baumgarten and House 2010, 1198).

The aim of this paper is, therefore, to investigate whether the pragmatic marker *like* displays a certain level of grammaticality in ELF discourse, as it does in native English interactions, and, additionally, to explore the potential different roles of the pragmatic marker *like* in English as a Lingua Franca communication. The examples of the pragmatic marker were taken from the Vienna Oxford International Corpus of English (VOICE) and the findings lead to the conclusion that *like* has indeed grammaticalized in English as a Lingua Franca. Considering similar past research and current observations in the VOICE corpus, five different categories of the marker *like* were established on the basis of functions it performs in a given example: hedges, illustratives, quotatives, fillers and independent use of *like*. Each of the categories was then analysed in further detail considering the setting in which it occurs, as well as the age and gender distribution of speakers using it.

This paper is organized as follows. The following section gives the theoretical framework dealing with (inter)subjectivity and grammaticalization, as well as the past research on the pragmatic marker *like*. Section 3 briefly describes the methodology and some general characteristics of the VOICE corpus. Section 4 gives deals with the results of the study, explaining the different categories of pragmatic marker and the frequencies of use of each individual category, as well as the setting in which *like* was used. The paper ends with a discussion and conclusion.

2 Theoretical framework

2.1 Subjectivity, intersubjectivity and grammaticalization

In order to discuss the role of the pragmatic marker *like* in English as a Lingua Franca, we must first turn to some of the key terms relevant in the process of *like* gaining its new meanings and use.

Subjectification is an important notion in the process of words gaining new functions and meaning. The simplest definition would be that “subjectification (and intersubjectification) denote processes whereby a linguistic element acquires increased (inter)subjective meaning“ (Davidse, Vandelanotte, and Cuyckens 2010, 1). Traugott (2010) dealt extensively with these terms looking back on her research and updating her work on subjectivity and intersubjectivity. In her view, “subjectification and intersubjectification are the mechanisms by which: a. meanings are recruited by the speaker to encode and regulate attitudes and beliefs (subjectification), and, b. once subjectified, may be recruited to encode meanings centred on the addressee (intersubjectification)“ (Traugott 2010, 35).

If we say that “‘subjectivity’ subsumes various ways in which an expression may involve speaker-reference“ (Davidse, Vandelanotte, and Cuyckens 2010, 8), it could also be concluded that language altogether entails subjectivity: “In a general sense, then, language can be said to be strongly marked by subjectivity in that any selection from the lexical and/or grammatical repertoire passes through the speaker“ (Davidse, Vandelanotte, and Cuyckens 2010, 9).

Considering this to be true, the concept of subjectivity is then not a matter of either-or, but a matter of degree, that is, some expressions can exhibit more subjectivity than others. On the other hand, intersubjectivity can then also be viewed as a characteristic of language in general: “Just as all language use can be said to be subjective in a general sense, it is intersubjective in a

general sense, reflecting the impact of the speech situation which not just involves a speaker but a communicative relationship between speaker and hearer“ (Davidse, Vandelanotte, and Cuyckens 2010, 13).

As far as discourse markers go, as they are the object of this study, subjectivity and intersubjectivity are important factors in the process of their formation and development. In her various case studies of subjectification, Traugott lists discourse markers as one of the main grammatical expressions of speaker-stance and characterizes them as “explicit markers of speaker attitude towards discourse structure, i.e. towards the relationship between what precedes and what follows, or of the connectivity between propositions“ (qtd. In Davidse, Vandelanotte, and Cuyckens 2010, 11). This and other characteristics of discourse markers will be dealt with in the following chapter.

Subjectivity and, in the case of discourse markers dealt with within this study, intersubjectivity are the main motivational forces for the process of grammaticalization. Lehmann (2002, 10) defines grammaticalization as “a process which may not only change a lexical into a grammatical item, but may also shift an item from a less grammatical to a more grammatical status”. Grammaticalization is not necessarily linked to subjectivity and intersubjectivity: “(Inter)subjectification often involves grammaticalization, but they are different types of changes which may occur independently of each other“ (Davidse, Vandelanotte, and Cuyckens 2010, 6). However, considering discourse markers, these terms appear to be closely intertwined. It is precisely because of their increased (inter)subjectivity that the discourse markers embark on the process of change into (more) grammatical units.

The term “grammaticalization“ is problematic and it is not easy to strictly define the boundaries of this process, so there are examples where it is not entirely clear whether a particular element is involved in the process of grammaticalization or not. Lehmann (2002,

146) tries to define more precisely what constitutes grammaticalization and provides six parameters: integrity, paradigmaticity, paradigmatic variability, structural scope, bondedness and syntagmatic variability. He emphasizes that these parameters do not identify grammaticalization, but the grammaticality of the sign, i.e. “the degree to which it is grammaticalized“ and concludes that “we may say that grammaticalization as a process consists in a correlative increase or decrease — as the case may be — of all the six parameters taken together“ (Lehmann 2002, 111).

Discourse markers do not fulfil all of Lehmann's criteria for grammaticalization, at least in the case of structural scope and syntagmatic variability. Rather than decrease in their scope and fixation, they show an increase in structural scope and looser syntactic connections. Traugott (2010, 41) also refers to Lehmann in her work and criticizes his criteria: “Limiting grammaticalization to reduction and condensation appears to be too restrictive, however. They pertain to certain domains of grammaticalization such as the development of case and tense, but not to other domains such as epistemic modality, connectives, discourse markers, etc., where scope increase is typical of grammaticalization“. However, Lehmann is also aware of the fact that his theory and correlation between all of the six parameters does not apply to all possible examples of grammaticalization. He believes that the parameters are theoretically dependent on each other, but also states that “the theoretical basis has not been made fully explicit, nor can it be, at least in this study. There are therefore no theoretical grounds on which to expect a 100% correlation between them“ (Lehmann 2002, 111).

Lehmann (2002, 11) emphasizes that grammaticalization is “a process of gradual change, and that its products may have different degrees of grammaticality. If grammaticalization is not a binary, but a gradual change of state, then we must face the problem that it may be an open-ended process“. He also stresses that “grammaticalization is not a process restricted to some

particular part of the grammatical system, but that it asserts itself everywhere between discourse structure and morphonology“ (Lehmann 2002, 107).

2.2 Past research on *like*

Over the years an increasing number of studies have dealt with *like* in its newly noticed role as a discourse marker. It became clear that *like* has indeed grammaticalized and developed new pragmatic functions.

Underhill (1988) in his *Like Is, like, Focus* offers one of the earliest descriptions of these functions. He refers to it as “the intrusive *like* that is entirely ungrammatical in standard English and makes sentences seem disjointed to many listeners“ (Underhill 1988, 234). His hypothesis is that “that nonstandard *like* is neither random nor mindless. Instead, it functions with great reliability as a marker of new information and focus“ (Underhill 1988, 234), whereas he defines focus as “the most significant new information in a sentence — often, the point of the sentence“ (Underhill 1988, 238). Underhill (1988, 236-242) goes on to explain the functions of *like*: to introduce new concepts or identities; to mark focused information; to mark the focus in questions; marking the focus in answers to questions; as a hedge; to set off unusual notions or a stereotyped expression, i.e., content which is not meant to be taken literally.

Some authors focused on one particular function of the pragmatic marker *like*. Blyth, Recktenwald and Wang (1990, 215) contrasted the quotative function of the phrase *be+like* with quotative verbs *say* and *go* and concluded that *like* as a quotative finds itself between direct and indirect speech: “Whereas most quotatives introduce either inner monologue or direct speech, the new quotative *be like* can introduce both kinds of reported speech, thus allowing the speaker to express an attitude, reaction, or thought, as well as something actually

said“. The use of *be+like* allows the speaker to express himself/herself in a way that was not possible with existing quotatives, and, therefore, it brings about a new discourse function: “*Be like*, however, performs discourse functions which *go* cannot. Whereas the quote which is introduced by *go* is necessarily interpreted as an utterance, the quote which follows *be like* may represent a thought, a state of mind, or inner monologue and therefore may be interpreted as never having been uttered“ (Blyth, Recktenwald and Wang 1988, 222).

One of the most prominent research on the subject was done by Romaine and Lange (1991), who also deal with *like* as an introducer of reported speech and the process of grammaticalization. Similar to Blyth, Recktenwald and Wang (1990), Romaine and Lange (1991, 238) compare the use of *like* with that of quotative verbs *go* and *say* and find that only *like* can be used to report on someone's thoughts. The aim of their study is to show “a case of ongoing grammaticalization involving the use of *like* in American English to mark reported speech and thought“ (Romaine and Lange 1991, 227). In their research they conclude that *be+like* has grammaticalized and became a fixed unit, which they called a “quotative complementizer“ (Romaine and Lange 1991, 248).

Andersen (1998) also researched the various uses of *like* as a pragmatic marker. The aim of his analysis of the pragmatic marker *like* within the framework of relevance theory was to prove that various uses of *like* can actually be subsumed under a single definition: “*like* can be used to signal looseness in the use of a wide range of different class elements and of whole propositions“ (Andersen 1998, 153). He lists examples which illustrate the different uses of the marker, such as indicating approximation, suggesting an alternative, marking reported speech and so on, but despite the versatility of functions that marker *like* can perform, Andersen (1998, 155) claims that the common denominator is “indicating loose interpretation of the speaker's belief“. Andersen deals extensively with this discrepancy between language and thought in

connection with pragmatic markers. He emphasizes that there is no one-to-one relation between thoughts and language and concludes that *like* serves as a signal of looseness in such cases:

“The function of *like* is precisely to signal that the speaker is opting for a loose interpretation of her beliefs, thus *like* can be considered a looseness marker. *Like* appears to provide an explicit signal of a discrepancy between the propositional form of the utterance and the thought it represents“ (Andersen 1998, 153).

In *And people just you know like 'wow'. Discourse Markers as Negotiating Strategies*, Jucker and Smith (1998) analyse discourse markers in the speech of young adults from Southern California. In their analysis they emphasize the importance of common ground and therefore they take discourse markers to be “one type of cue that conversationalists use to negotiate their common ground“ (Jucker and Smith 1998, 172). They present evidence which suggests that “representations of common ground that are consulted during a conversation are dynamic, i.e. they include assumptions not only about explicitly stated information that is assumed to be shared but also about the inferences the partner is expected to draw from this information“ (Jucker and Smith 1998, 172-173). Jucker and Smith analyse three most frequent discourse markers in their data: *yeah, like and you know*, and characterize the discourse marker *like* as information-centered presentation marker, which means that it “accompanies and modifies the speaker's own information“ (Jucker and Smith 1998, 174). In other words, marker *like* in this case serves as a cue to the hearer, signalling how the content to which it refers should be interpreted. In their data analysis, they distinguished between several functions of *like* as a discourse marker: approximator, exemplifier, hedge and *like* used in quotation (Jucker and Smith 1998, 183). Similar to Andersen (1998), Jucker and Smith (1998, 191) also emphasize the discrepancy between the utterance and the thought it represents, and conclude that “all its uses can be subsumed under its core function of flagging linguistic expressions, i.e.

words, phrases, clauses or entire utterances, as less than literal representations of the thoughts they are intended to represent“.

The research done by Cameron and Deignan (2003) shows that the pragmatic marker *like* in its function as an indicator of less-than-literal rendering of speaker's thoughts can also be viewed as a type of “tuning device“, since it signals to the hearer how to interpret certain expressions and statements. Cameron and Deignan research these expressions, which they call tuning devices, used in connection with metaphors. Taking into consideration both the hearer's and the speaker's needs, they find that these expressions perform important functions in the discourse. Tuning devices “appear to serve a range of pragmatic microfunctions, loosely described as offering cues to how the speaker intends metaphorically used language to be interpreted“ (Cameron and Deignan 2003, 153). They conclude that these markers of looseness “serve to alert an interlocutor to unexpectedness in the discourse, to direct his or her interpretation, and to adjust the strength and emphasis of a metaphor“ (Cameron and Deignan 2003, 159).

Another study crucial in the context of English as a Lingua Franca is Mauranen's study of metadiscourse, or discourse reflexivity as “a fundamental property of human communication“ (Mauranen 2010, 13). She assumed that findings from the ELFA (English as a Lingua Franca in Academic Settings) corpus could be useful for establishing potentially universal aspects of discourse, because the participants of lingua franca interaction “must adapt to highly diverse and unpredictable circumstances while maintaining communicative efficiency“ (Mauranen 2010, 13). In her work she defines reflexive discourse as “discourse about the ongoing discourse“ (Mauranen 2010, 16) and although she sees metadiscourse as roughly synonymous to reflexivity, she emphasizes that it has “become to be used for a wider range of phenomena than reflexive discourse“ (Mauranen 2010, 17). Reflexive language is, according to Mauranen

(2010, 16), highly interactive, as it “sharpens conversationalists’ mutual understanding of how to relate to the discourse that is being co-constructed. In this way, it contributes to the two fundamental uses that language has: sharing experience and negotiating interaction“. The results reveal that discourse reflexivity is not equally salient in all circumstances of native and ELF discourse, but nevertheless present in both alike, which suggests that reflexivity may indeed be a “discourse universal“ (Mauranen 2010, 21). Another finding was that metadiscourse tends to collocate with hedges, which Mauranen (2010, 24) calls “discourse collocations“. Since metadiscourse “implies an imposition of the speaker’s perspective on the discourse, and in so doing reduces the negotiability of the dialogic perspective“ (Mauranen 2010, 24), it has much in common with the study of pragmatic functions of markers in ELF, as they both deal with indicating awareness of our talk and signalling the interlocutors how to interpret it. The fact that this concept appears in ELF communication suggests that it is “not confined to Anglo-American discourse conventions but is a more fundamental feature of discourse“ (Mauranen 2010, 24).

As for the sociolinguistic properties of the pragmatic marker *like*, Dailey O-Cain (2000) comes to interesting conclusions in her study of the quotative *like*. She considers the age and gender distribution of *like* in a corpus of informal U.S. English, as well as perceived age and gender distribution and analyzes specific sociolinguistic stereotypes associated with this usage. She finds that younger people use both kinds of *like* more often than older people do, and that men and women use it approximately equally often (Dailey O-Cain 2000, 60). As for the attitudes towards *like*, its use is perceived as mixed, or as both positive and negative at the same time, which could be explained with the help of “solidarity“ and “status“ traits – Dailey O-Cain (2000, 76) elaborates that the use of *like* can “have negative connotations to a listener in

terms of status“, but it can also „emphasize solidarity between the speaker and the hearer“ and thus be positively perceived.

Past research on the subject of various uses of *like* shows that the grammaticalization of the pragmatic marker *like* has its place in native English discourse. Although English as a Lingua Franca is expected to show some similarities with native English discourse, it is nevertheless a very specific type of discourse, occurring between speakers with different languages and cultural background, which have to build their common ground as they interact. House (2010) dealt with the subject extensively in her *The Pragmatics of English as a lingua franca*. She describes ELF communication as multi-voiced, produced by (mostly) multilingual speakers who develop their own discourse strategies (for instance, the “let-it-pass principle”), and claims that it should be regarded as “a third way, a crossing of borders, a hybrid language” (House 2010, 382). House (2010, 365) emphasizes that ELF communication has “no pre-fixed norm, and therefore lingua franca speakers must always work out a new joint linguistic, intercultural and behavioral basis for their communication in different communities of practice”.

“Communities of practice” is a term also used by Cogo (2010, 296), which she defines as speech communities in which “the norms are not preestablished, and they are not ex normatively imposed, but they are negotiated by its users (‘mutual engagement’) for specific purposes (‘joint enterprise’) by making use of the members’ lingua-cultural resources (‘shared repertoire’)”.

The lack of shared knowledge between ELF speakers of different linguistic and cultural backgrounds can lead to misunderstanding and communication difficulties, but research shows that these cases are actually rare because “speakers show that they can use ELF in their own ways by also drawing on their shared multilingual repertoires” (Cogo 2010, 309). ELF speakers appropriate English in their own way and “effective interactional work is carried out through

various strategies in a supportive manner, so that meaning is explored, clarified and eventually understanding is promoted” (Cogo 2010, 309). Although the cases of misunderstanding are not so frequent, the reduced shared repertoire still plays a crucial role in the ELF discourse, since “the ELF speakers, with their individual socio-cultural background, need to jointly develop and continuously negotiate a repertoire of resources which cannot be taken for granted or assumed a priori” (Cogo 2010, 296). Because of the lack of shared knowledge, the ELF speakers might not exploit the whole range of pragmatic functions that pragmatic markers can perform in discourse.

Subjectivity and intersubjectivity play an important role in the grammaticalization of pragmatic markers and the roles which *like* performs in the native English discourse exhibit not only subjective, but also intersubjective meaning. The pragmatic functions of the marker *like* can be said to be heavily hearer-oriented, as they tune the content and signal to the hearer how to interpret them. It will be interesting to see whether the pragmatic functions of *like* in ELF discourse show a certain degree of grammaticality and similar patterns as in native English discourse. If so, the similarities could prove that these pragmatic functions are fundamental to discourse, as Mauranen (2010) concluded in her study of discourse reflexivity in ELF. Another potential goal is to find whether the functions of *like* in ELF share enough of the same properties to be subsumed under a single definition, similar to the ideas proposed by Jucker and Smith (1998) and Andersen (1998).

3 Methodology

The examples for this study were taken from the Vienna-Oxford International Corpus of English (VOICE), version VOICE POS Online 2.0. The corpus comprises transcripts of naturally occurring, non-scripted face-to-face interactions in English as a Lingua Franca, and at the present moment VOICE contains over 1 million words of spoken ELF interactions. Over 1250 ELF speakers are currently recorded in VOICE, with approximately 50 different first languages, including native speakers of English (VOICE. 2013. The Vienna-Oxford International Corpus of English. 15/6/14).

A random sample of roughly 500 examples (534, to be exact) of the discourse marker *like* was extracted from the corpus and analysed according to their pragmatic function. The interactions from which the data were collected were recorded in different domains (educational, leisure, professional) and they represent different speech event types (conversation, interview, meeting, panel, press conference, question-answer session, seminar discussion, service encounter, working group discussion, workshop discussion). Therefore, these factors were also taken into account in the data analysis, as well as the speaker's age and gender.

4 Results

4.1 Classification of the discourse marker *like*

The division of categories presented here is based on the categories found in the research presented above, which were combined with my current observations of examples from the VOICE corpus. The criterion for this classification is based on the pragmatic function that *like* performs in a given example, the context and the structural scope of the marker. Taking all of this into consideration, I found five major categories, some of which are closely intertwined: hedges, illustratives, quotatives, fillers and independent use of *like*. Each of the categories will be presented in more detail in the following section. The results are organized according to the frequency of each category, starting with the most frequent one.

4.1.1 Hedges

Hedges, as the word itself suggests, enable the speaker to distance himself/herself from the spoken content. Since one can never articulate concepts and ideas exactly as they appear in the mind, hedges provide the speaker with a “safety net” against the hearer's potential false interpretation and decrease the speaker's responsibility for the expressed facts. By using hedges, the speaker also avoids the danger of being taken too literally. These expressions could even be seen as a type of euphemisms, since they make context, or at least the expression to which they refer, more innocuous.

In this case *like* as a pragmatic marker serves as a signal which could be paraphrased as “my thoughts are something like the following“. Since *like* is defined in dictionaries as “similar to something or someone“, “comparable to or close to something“ (Merriam-Webster.com. 2014. 15/6/14), this use of pragmatic *like* is actually not so far from its primary meaning.

According to their specific function and collocations, I further divided the category of hedges into four subcategories: approximative, descriptive, metalinguistic and emphasizing hedges. They all share most of their characteristics, but differ in certain aspects.

a) approximatives

The pragmatic marker *like* is in this case used before numerical expressions, referring to time, percentages, measures, etc., in order to avoid unnecessary precision. Andersen (1998, 152) gives an example of such a hedge, saying that if someone is asked for the time and gives an answer ten thirty, although the watch shows “10:31:04“, this speaker is “offering a loose interpretation of what she believes to be true“. The general idea and content are conveyed without being too detailed or specific, which is unnecessary at that point in the discourse.

➤ Example 2

okay let's move to the second half question because we have like three minutes left for it

It is quite possible that they have perhaps two or four minutes left to finish their work, but the speaker just wishes to point out that there is a fairly small amount of time left. The notion of vouching for the exact number of minutes left is not so important.

➤ Example 3

we went to liverpool. we were like nineteen or something? a:nd (1) yeah we just drove into the city. didn't know where we were going (just) (.) with a renault

Another example where the exact age is not crucial for the content of the story. The idea of “nineteen or something” provides the hearer with enough information to set the story in context.

➤ *Example 4*

yeah this is sort of one one thing that we (1) er caught as er (.) as something really good for consumer er at least in our country (.) was that er we claimed somewhere there that er erm it is like five per cent less (.) er of electricity you use.

The speaker cannot warrant that it was exactly “five per cent”, but the general idea is nevertheless successfully conveyed.

b) descriptives

This category is similar to the previous one, but is often used while retelling and describing. The speaker is searching for the right word and is not sure if s/he found the best expression and if the content is conveyed successfully, hence s/he is distancing himself/herself from the expression used. It can mostly be paraphrased with “sort of”.

➤ *Example 5*

S1: er: the grand master used (.) er he had a as like a summer residence there you know (.)

S2: yes

S1: and he built like a small forest to be able to do hunting you know (.)

In these examples the speaker is not quite sure if “summer residence” and “small forest” are the best possible phrases, but the general idea and understanding is achieved.

➤ *Example 6*

S1: mhm. (2) <smacks lips> li:ke a typical dish that my MOther prepares when it's co:ld is like this erm <Llita> come si chiama {what's it called} </Llita><smacks lips><soft> god i forgot </soft> (2) <Llita> pizzoccheri {type of pasta} </Llita> hh AND there's like (.)

*LOADS of BUTTer this CHEESE from the north of italy? hh and then: potatoes? and s- e:rm
<smacks lips><Llita> erbette {beet tops} </Llita> like erm <smacks lips> they're like er (.)
like spinach leaves. but they're not spinach? like really: big leaves? that you COOK hh AND
it's thick pasta (.) which is made of e:rm (.) a special flour (1) and you cook it ALL together?
(3) and it becomes like a s:- cre:amy: chee:sy pasta you put it in the oven i'm sure there is
something corresponding (.) like =*

S2: like (.) in a white (.) thingy in the ov- like (.) it has =

*S1: it's more like erm <smacks lips> yeah i mean: VARious ingredients. and then you put it a
bit in the oven to: mend the whole thing together (.) but you don't really (.)*

The speaker is experiencing some planning difficulties while trying to describe a specific meal and recipe. Some of these pragmatic markers are on the verge between their function as a filler and their function as a hedge. The examples “and it becomes like a s:- cre:amy: chee:sy pasta” or “like (.) in a white (.) thingy” are descriptive hedges, the speaker can't decide or can't think of the expression she would like to use so settles for this one, for the sake of getting her idea across.

c) metalinguistic hedges

This type of hedge is the most similar to the descriptive kind, but instead of distancing himself/herself from the content of the expression, the speaker is distancing himself/herself from the style. However, the metalinguistic hedge seldom appears in connection with planning difficulties, as it was the case with the descriptives. Andersen (2000, 30) mentions this kind of pragmatic markers in his *The Role of the pragmatic marker like in utterance interpretation* : “the speaker implicitly suggests that there may be alternative expressions that might be just as fitting or perhaps more fitting than the one she chooses“.

➤ *Example 7*

when they want to remain in power when they like quarrel with students or with pupils

The student is talking to his professor, and is therefore insecure about using the right style. Perhaps the speaker believes that the verb is “too strong” for the setting, but can’t think of a more fitting one.

➤ *Example 8*

S1: what about i took a course of like

S2: of

S1: economics as (.) like in lingua franca

Students are having a relaxed conversation in a bar. The speaker is distancing himself from the content because he is perhaps not sure whether the course he took was called exactly “economics” (maybe it was “economy”, “business” or something else), whereas “lingua franca” probably refers to the fact that the course was in English. As opposed to descriptive hedges, there is no ambiguity about the content conveyed, just the possibility of it being stylistically awkward.

d) emphasizing hedge

In its function as an emphaser, the marker *like* is used to highlight, emphasize or signal that the speaker is exaggerating and is not to be taken literally. The content to which *like* refers is in this case is often superlative, hyperbolic or metaphoric.

➤ *Example 9*

that's like my favorite line in the movie...

The speaker wishes to emphasize how much she likes a certain movie line.

➤ *Example 10*

and he said i'll get you something (real) drink (.) but you HAVE to drink it as a shot (.) you MUST drink it as a shot (1) and (.) and he said it like ten times and ...

The speaker wants to emphasize the fact that the person was very persistent, but it is very unlikely that the person in question said it literally ten times.

Hedges are divided here into four subcategories: approximative, descriptive, metalinguistic and emphasizing. They all signal to the hearer that “the following is not to be taken literally”, but differ in the elements which they modify. Approximatives refer to numbers, descriptives distance the speaker from the content, metalinguistic hedges from the style of the expression and emphasizing hedges serve as boosters. *Like* in its hedging function could be even considered an example of language economy, since it enables the speaker to convey the most content with the fewest words, without any long and perhaps unnecessary explanations. Andersen (1998, 152) offered a similar explanation of this function: “(...) loose use of language is a result of the speaker aiming at optional relevance by providing an answer which requires less processing effort than a strictly literal one would do“.

4.1.2 Illustratives

Although often used, this function of *like* is not so often discussed nor researched. *Like* in its illustrative function introduces some information which provides more detailed information on

the element it precedes. The data in the corpus show that the category can be divided further into two subcategories: exemplary and explanatory illustratives.

a) exemplary illustratives

Exemplary illustratives introduce a concrete example, usually of something already mentioned, immediately preceding *like*, and they are often followed by an enumeration of examples. An exemplary illustrative could be paraphrased with “for example“ or “for instance“, which in fact sometimes follow it.

➤ Example 11

and this as we sai- as us as us e:r as i said (.) at the level of [org4] this regional level (.) the biggest (.) documents communication documents like message truck (.) if you like (.)like q and a-s (.) like scientific dossier like l- legal erm e:r like legal er description of the project and so on so forth.

These instances of *like* show that the speaker is actually enumerating examples in order to illustrate his thoughts more precisely and get his idea across. All the notions introduced by *like* are examples of different types of “documents”.

➤ Example 12

yeah just i mean because like for instance in germany there is not only german

The meaning of *like* here is paraphrasable with “for example” or “for instance” and could actually be left out of the sentence since it is immediately followed by “for instance”. In this particular example it could also be said, since it is surrounded by other elements serving discourse functions: “yeah”, “just”, “I mean”, that it serves to keep a speaker’s turn and in that case its function is not that different from a filler word.

b) explanatory illustratives

The explanatory illustrative *like* specifies an already mentioned expression. It does not introduce a concrete example but rather a more elaborate explanation or a specification of the previously mentioned idea.

➤ Example 13

and then (1) when you have internet fora erm (1) personalize and er have er CONcrete have concrete persons for the students e:r like if we strengthen the role of the chairs then this would be the case.

In order to illustrate more specifically what is meant by *have concrete persons for the students*, the speaker elaborates, and also actually gives a potential example: *like if we strengthen the role of the chairs*.

Both types of illustrative markers serve to give additional information about the preceding element in order to specify an expression and avoid potential misunderstanding, but they differ in the content of that information – whereas an exemplary illustrative provides a concrete example, the explanatory one offers a more elaborate description of the conveyed idea.

4.1.3 Quotatives

Much research has been done on the quotative function of *like*, whether dealing with the process of its grammaticalization in progress (Blyth, Recktenwald and Wang 1990; Romaine and Lange 1991), or studying the sociolinguistic distribution and perception of the quotative *like* (Dailey-O'Cain 2000). Past research on the subject shows that *like* is increasingly used as an introducer of reported speech and that its grammaticalization cannot be denied.

Both Romaine and Lange (1991) as well as Blyth, Recktenwald and Wang (1990) agree that *like* as a quotative enables the speaker to introduce both speech and thoughts and therefore offers new possibilities of expression. The notion of hedging is not to be disregarded here, since the quotative *like* also relieves the speaker from responsibility of vouching for the absolute truth of the presented content. The content introduced by the quotative *like* may have not actually happened, but rather serves as a story-telling tool.

➤ *Example 14*

erm (1) and (.) there was like there were (.) one person on a night shift (.) and you're supposed to be educated to be on a night shift (.) hh but they needed people so badly that they were like (.) hey you can come on the night shift and i was like (.) but i'm not educated i don't want to (.)but you can (just get some) training and it's okay i was like (.) no i don't want to i don't want to have (.) i don't want to have the responsibility for twenty-eight old people they could (.) all collapse and die on my shift

This example illustrates the blurred distinction between speech and thought. The use of the quotative *like* creates a vivid retelling effect, although it is unlikely that the speaker actually said the things she reported saying. According to Romaine and Lange (1991, 263), *like* only creates „an example of something that could have been said or thought without implying the kind of commitment that *say* does“.

4.1.4 Fillers

In the process of online planning, the interactants build their discourse as they go along, which may sometimes result in hesitation, pauses or repetitions. Although not always noticeable in spontaneous speech, hesitation actually serves an important pragmatic function, allowing the speaker to plan the content of his/her speech while keeping his/her turn in

conversation. Among other various elements used to fill the pauses when the speaker is having planning difficulties, the pragmatic marker *like* is also found to function as a filler. In my data analysis, I have categorized *like* as a filler in cases where it was surrounded by other hesitation markers, such as silent and filled pauses (“erm”, “uh”) or repetitions.

➤ *Example 15*

it's very good and er you you eat salad a:nd or then or you (.) you also: have some like erm (com commo se pronunce accetato?)

In this example it is obvious that the speaker is having planning difficulties and can't think of the right expression in English. The sentence contains repetitions, pauses and filled pauses (“erm”) and even code switching to native tongue (“com commo se pronunce accetato”).

➤ *Example 16*

yeah yeah yeah yeah hh there's supposed to be a like er er erm (.) a room you can stay in? (.) erm (.) at my place like with a (.) telly or something (.) but it's locked? and i'm and my key doesn't work through the door?

The first instance of *like* in this example is followed by a longer filled pause, so I characterized it as a filler. However, it is clear that the speaker can't think of the right expression: “like (...) a room you can stay in?” so this example could also be interpreted as a descriptive hedge.

➤ *Example 17*

but there can also be er like erm a th- a third partner. (.) not necessarily a northern

This is another example of speaker's hesitation, where the speaker is insecure about using the term “third partner”. Therefore, this instance could also be understood as descriptive or

metalinguistic hedge, although the use of filled pauses around *like* signal that it is used here to keep his turn while planning the content of his speech.

With examples of *like* in its function as a filler, it is often not clear whether the speaker is having planning difficulties and wants to buy himself more time to think, is he just searching for the right word, or not sure if the following expression is the most appropriate one. *Like* in this function is often very similar to hedges, and it would perhaps not be wrong to make fillers a yet another subcategory of hedges, although the categorization of fillers was based more on formal criteria, i.e., hesitation marks. In this case, hedges and fillers find themselves in a continuum, with no clear boundaries between them.

4.1.5 Independent use of *like*

The term “independent“ refers to the fact that a single element, phrase or subordinate clause which *like* refers to could not be identified. Instead it has a much wider structural scope. This function of *like* could also be understood as type of hedge, since it contains the element of looseness, however, it is hard to identify a single element to which it refers. Andersen (1998, 156) mentions similar type of markers and claims that “the same analysis can be applied in these examples, the only difference being that *like* does not have narrow scope over a particular sentence component, but it imposes some element of looseness on the propositions at large“.

➤ Example 18

i mean but you have (.) like which part is yours?

In this example it would be logical to assume that *like* refers to “which part is yours”, however there is no uncertainty about the choice of content or style which would require this particular part of the utterance to be modified.

➤ *Example 19*

S1: it was kind of funny too it was something different like and

S2: yeah

S1: and a lot of (gay) people watched it =

S2: o:h okay

Although the distribution of *like* was not a criterion in this study, it is noticeable that it modifies the element it immediately precedes or follows. Therefore, it is not clear in this example which element *like* refers to. If we restate the example as “it was *like* something different”, we could say that *like* would function as a hedge modifying “something different”. However, as it actually follows that phrase, we might assume that it functions as an illustrative explanatory hedge, i.e. that “a lot of gay people watched it” is actually an explanation of why was it “something different”. This is also not the case since the conjunction “and” is found in between.

Since it was not clearly identifiable which element the marker *like* refers to, it is possible that *like* in these examples has hedging function with the entire utterance in its scope, signalling looseness and informality of speech.

4.2 *Like* in numbers

As already mentioned, apart from the functions that *like* performs, different domains and speech events in which these instances took place were also taken into account.

The VOICE corpus compilers distinguish between three domains. The educational domain includes all social situations connected with institutions or people involved in teaching, training or studying, and accounts for 25.51% of all data in the VOICE corpus. 9.89% of all data in the corpus belongs to the leisure domain, which refers to social situations spent doing something one chooses to do when not working or studying. And, lastly, the professional domain includes social situations connected with an activity that in some way needs special expertise. The professional domain is further subdivided into professional business (19.88%), professional organizational (34.66%) and professional research and science (10.06%), which altogether make up 64.6% of the transcribed data in the corpus.

Furthermore, the VOICE compilers have classified their data according to speech event types, which are defined on the basis of purpose, type and number of participants, and are also represented by different percentages in the corpus data: meetings (26.73%), working group discussions (17.7%), conversations (15.45%), workshop discussions (15.43%), panel discussions (9.06%), seminar discussions (6.22%), interviews (3.55%), question-answer sessions (2.69%), press conferences (1.72%) and service encounters (1.46%).

As for the data extracted for the purposes of this study, which consists of 534 randomly chosen instances of the pragmatic marker *like*, the percentages differ. Although the largest proportion of the data in the VOICE corpus belongs to the professional domain, *like* was most frequently found in the domain of education (35.02%) and leisure (32.77%) (see Table 1 in the Appendix), whereas the majority of examples belong to speech events such as conversation

(40.07%), workshop discussion (21.72%) and working group discussion (17.04%) (see Table 2 in the Appendix). These differences could be explained by the fact that *like* as a pragmatic marker signals loose and informal speech. In the corpus, conversation is defined as a speech event at which people interact without a predefined purpose, whereas workshop and working group discussion both refer to speech events in which a group of people (of the same rank) discusses a problem or exchanges views, ideas and information on a particular topic. They can therefore be understood as occurring in informal contexts, where the informants speak among themselves in a relaxed manner. The professional domain is the predominating one in the VOICE corpus, and workshop discussions and working group discussions are the most informal speech events belonging to that domain. Apart from that, these speech events are also found in the educational domain.

The lowest number of pragmatic markers was found in the professional business domain (8.61%) and the professional research and science domain (3.18%) (see Table 1 in the Appendix). These two provide more formal settings than the professional organizational domain, which refers to social situations connected with activities of international organizations that are not doing research or business, and accounts for 20.41% of instances of *like* in this study.

Press conferences (0.19%), service encounters (0.37%), question-and-answer sessions (1.31%) and panel discussions (2.06%) are the least represented speech acts with respect to the pragmatic marker *like* (see Table 2 in the Appendix). Service encounters in the VOICE corpus refer to speech events at which someone seeks a service provided by someone else, so it presumably, most of the time, does not refer to an informal setting with well acquainted informants. Question-answer sessions are speech events at which an audience asks specialist speakers questions, whereas panels, similarly, refer to an event at which specialists give their

advice and opinion on a specified topic to an audience. Since they both refer to a very formal setting, the pragmatic marker *like* is not very common at these speech events.

As for the age and gender of speakers, the fact that they are not evenly represented in the VOICE corpus must also be taken into account when analysing the data. Out of 1270 speakers in the corpus, 55.66% are female and 43.38% male (with 0.15% unknown). In the past research done by Blyth, Recktenwald and Wang (1990) and Dailey O-Cain (2000), the findings have shown that men use *like* even more often than women, although the difference is negligible (O-Cain 2000, 75), while Romaine and Lange (1991, 151) found that the majority (83%) of their examples of quotative *like* were produced by women. The findings differ, even when comparing research dealing with the same, quotative function of the pragmatic marker.

In my data women were predominant with 64.79%, while male speakers accounted for 35.21% of the uses of *like*. Considering that the VOICE corpus consists of slightly more female speakers than male and that not all instances of the pragmatic marker *like* in the corpus were analysed, this difference in the percentage of women and men in the data cannot be taken as relevant.

The distribution of speakers by age groups in the corpus is as follows: 40.62% speakers aged 17-24, 17.48% speakers aged 25-34, 24.09% speakers aged 35-49 and 9.68% speakers aged 50 and over (with 7.32% unknown).

Considering my examples of pragmatic marker *like*, 62.92% of speakers belong to the youngest age group (17-24) and 19.66% to those aged 25-34 (see Table 14 in the Appendix). Only 13.67% of speakers in the data were aged 35-49, although it is the second largest age group represented in the corpus. 3.75% of instances of *like* were used by speakers aged 50 and older. These results show, similarly to those of Blyth, Recktenwald and Wang (1990) and

Dailey O-Cain (2000), that younger people do use *like* more often than older people. Even when accounting for the fact that the majority of speakers in the VOICE corpus are young, it could still be concluded that age gradation is a noticeable phenomenon when dealing with the sociolinguistic distribution of the pragmatic marker *like*.

In the following section each category of *like*, i.e. hedges, illustratives, quotatives, fillers and independent use of *like*, will be analysed in relation to the domain, speech event type, and age and gender of the speakers.

4.2.1 Hedges in numbers

The results reveal that the most prominent category of the pragmatic marker *like* is the category of hedges. With 249 occurrences out of the total of 534 coded examples, hedges make up 46.63% of the data extracted from the VOICE corpus (see Table 3 in the Appendix).

Hedges were most frequently used in the leisure (17.42%) and educational domain (17.23%) (see Table 4 in the Appendix). 8.05% of hedges were found in the professional organizational domain, in the professional business domain 2.62% and only 1.31% in the domain of professional research and science, which accounts for a total of 11.98% in the professional domain.

As for speech event types, 21.72% of hedges were used during conversations, 9.36% during workshop discussions, and 7.12% during working group discussions (see Table 5 in the Appendix). The fewest number of hedges were used during panel discussions (0.37%) and service encounters (0.37%).

Speakers aged 17-24 used hedges most often (29.59%), as opposed to 10.49% of speakers aged 25-34, 5.62% of speakers aged 35-49 and only 0.94% of speakers aged 50 and older (see

Table 14 in the Appendix). Although speakers aged 17-24 are the most numerous group in the corpus, it can still be said that younger speakers tend to use hedges more often than older ones.

Hedges in women's speech make up 29.96% of the sample, 16.67% in men's speech, which does not seem relevant since female speakers outnumber male speakers in this corpus.

All of these numbers are actually consistent with the overall frequency of the marker *like* in the corpus, which can again be explained by the fact that *like* as a pragmatic marker signals loose and informal speech and is therefore more frequent in the more informal domains and speech event types.

Since speakers with different L1s may have difficulties establishing common ground, it is not surprising that hedges are found to be the most frequent function of *like* used in this data. Hedges enable distancing from the spoken content and protect the speaker in case of possible misunderstanding or false interpretation. They decrease speaker's responsibility for the proposed content and allow him/her to express insecurity or just signal that the following statement is to be interpreted loosely. Therefore, they serve as an important tool in negotiating common ground and make the communication between speakers with different L1s easier.

4.2.2 Illustratives in numbers

Illustratives are the second most frequent category and they make up for 29.03% of the pragmatic marker *like* in the data (see Table 3 in the Appendix).

11.05% of markers used as illustratives are found in the educational domain and 6.37% in the professional organizational domain, followed by 5.24% in the professional business domain and 5.06% in the leisure domain (see Table 6 in the Appendix). The least represented domain is

the professional research and science domain with only 1.31%, which is also the least represented domain in the entire corpus.

Illustratives were mostly used during workshop discussions (7.12%), conversations (6.55%), working group discussions (5.62%) and meetings (5.43%) (see Table 7 in the Appendix). The fewest number of illustratives were found in the question-answer sessions (0.37%), press conferences (0.19%) and seminar discussions (1.12%).

It is noticeable that *like* in its function as an illustrative marker is used at both informal and formal occasions. Illustratives are not limited to mostly informal speech, as was the case with hedges.

Illustratives were most frequently used by speakers aged 17-24 (16.10%) (see Table 14 in the Appendix). Interestingly, speakers aged 35-49 were found to use illustratives more often (5.81%) than speakers aged 25-34 (4.68%), which are the more numerous group in the corpus. 2.43% of illustratives were produced by speakers 50 and older.

The age gradation is noticeable, but it is not as prominent as in the case of other categories of pragmatic marker *like*. Although the numbers in question do not warrant definite conclusions, it appears that this particular function is less exclusively linked to younger generations of speakers.

As for the gender aspect, numbers are consistent with the overall statistics of the corpus, i.e. the number of female speakers is slightly higher than the number of male speakers. 18.43% of illustratives were produced by female speakers and 10.67% by male speakers.

Illustratives seem to be a versatile category in terms of speakers and setting in which they appear. They are not limited to a purely informal setting and they are used more frequently than the other categories of *like* by speakers older than 25.

4.2.3 Quotatives in numbers

The quotative category makes up 10.49% of the data collected and is the third most frequent category of the marker *like* in this data (see Table 3 in the Appendix).

7.12% of quotatives are found in the leisure domain, followed by 2.06% in the educational domain (see Table 8 in the Appendix). The fewest number of quotatives are found in the professional business (0.19%), professional research and science (0.19%) and the professional organizational domain (0.94%), accounting for only 1.32% in the professional domain altogether.

As for speech event types, the majority of quotative markers were found during conversations (7.68%), whereas other speech event types are represented by numbers below one percent (see Table 9 in the Appendix). These percentages indicate that the quotative *like* tends to be used in informal and relaxed speech, and less so in more formal contexts.

8.80% of quotatives in the data were produced by speakers aged 17-24, whereas other age groups are represented by numbers below one percent, or none in the case of speakers aged 50 and older (see Table 14 in the Appendix). Quotatives tend to be used by younger speakers, even more so than the other types of the pragmatic marker *like*. Clear age gradation in the case of the quotative *like* was mentioned by Blyth, Recktenwald and Wang (1990, 219), as they found in their corpus that it was often used by younger speakers, but never by speakers older than 38. However, they dealt with native speakers of English, so the correlation in terms of age gradation between native English and ELF speakers is of great importance.

8.05% of quotatives were produced by female speakers, and 2.45% by male, which is consistent with the statistics of other categories, and considering the predominance of female speakers in the VOICE corpus, the numbers do not seem relevant.

Quotative *like* seems to be limited to the informal context and relaxed speech. It is also almost exclusively used by the youngest group of speakers in our data, which is consistent with the findings by Blyth, Recktenwald and Wang (1990) in native English discourse. Quotative *like* remains a pragmatic marker preferred by younger generations of speakers, and it exemplifies age gradation better than any other category in the data.

4.2.4 Fillers in numbers

Fillers are the fourth most frequent category of marker *like* in this data and they make up 9.36% of pragmatic markers analysed in this study (see Table 3 in the Appendix).

Fillers were most frequently found in the professional organization (4,12%) and educational domain (3.18%) (see Table 10 in the Appendix). They were least frequent in the domains of professional research (0,37%) and professional business (0.37%).

As for the speech event types, fillers were mostly found at workshop discussions (3.93%), working group discussions (2.43%) and conversations (1.87%), and the fewest fillers were found at panels (0.19%) and question-answer sessions (0.37%) (see Table 11 in the Appendix).

As in the case of other categories of *like*, fillers are most commonly found in the speech of the youngest group of speakers (17-24) with 5.24% (see Table 14 in the Appendix). Speakers aged 25-34 account for 2.62% of fillers, speakers aged 35-49 only 1.12% and speakers aged 50 and older just 0.37%.

Female speakers are more frequent with 5.43%, as opposed to male with 3.93%, which is consistent with gender distribution of the other categories of *like*.

Considering the domains and speech event types, the results are mixed and there is no clear indicator that fillers are used strictly in informal settings. The reason for that may be that fillers,

as indicators of hesitation, depend on the individual speaker's confidence more than the informality or formality of the context.

4.2.5 Independent use of *like* in numbers

The least frequent category of marker *like* in this data is the independent category, which makes up 4.49% of the analysed markers (see Table 3 in the Appendix).

It was mostly found in the domains of leisure (1.87%) and education (1.50%) (see Table 12 in the Appendix). Only 0.19% of this type of markers were found in the professional business domain and none in the professional research and science domain.

Independent markers were most frequently used during conversations (2.25%), whereas other speech event types are represented by numbers below one per cent (see Table 13 in the Appendix). These percentages show that the independent marker tends to be used more frequently in informal contexts.

As for the age, speakers aged 17-24 account for 3.18% of independent markers in the analysed data, speakers aged 25-34 make up 0.94%, speakers aged 35-49 only 0.37%, and no example was found by speakers aged 50 and older (see Table 14 in the Appendix). These numbers are consistent with the age gradation shown in the case of other categories of marker *like*.

Gender distribution is also similar to other categories, i.e. female speakers account for 3% of independent markers and male for 1.5%.

The independent marker *like* is supposed to contain an element of looseness and informality, so it was expected to be found in more informal contexts such as leisure domain and conversations. This category is not as represented in the corpus as the other functions of

pragmatic marker *like*, but it shares their overall statistics, such as age gradation and a higher number of female speakers.

5. Discussion and conclusion

The analysis of the pragmatic marker *like* in the VOICE corpus has shown that it is used by a large number of ELF speakers in various contexts. Since these instances of pragmatic marker *like* are actually hearer-oriented, signalling how the content to which they refer should be interpreted by the addressee, it can be said that they definitely exhibit a certain degree of intersubjectivity. As far as grammaticalization is concerned, Lehmann (2002, 11) emphasized that it is a gradual change of state and possibly an open-ended process. The different pragmatic functions of *like* found in the corpus, show that *like* in ELF is undergoing the same process as *like* in the native English discourse, and a certain degree of grammaticalization of *like* in ELF is evident. More specifically, the illustrative function of the pragmatic marker *like* seems to be the closest to the primary meaning of *like* as a preposition signalling comparison, since it enables the speaker to bring up examples that conform to a standard that is mentioned in the discourse itself. Thus, when the speaker in example 11 mentions "scientific dossier" and "description of the project" as types of "documents", he is actually saying that they conform to the standard of documents. The remaining functions found in the random sample show a clear removal from the primary meaning and are more grammaticalized, as they exhibit a switch from an "objective" comparison (where the two types of content being compared are actually present) to a more "subjective" comparison (where the speaker signals that the expression used with *like* is comparable to or approximates a certain standard, but without mentioning this standard). These standards may include numbers or imprecise expressions (hedges), or one's thoughts (quotatives). The hedging and quotative functions are more grammatical (i.e. grammaticalized) because they are parallel to expressing grammatical function such as modal certainty. Finally, the category of fillers shows complete grammaticalization, where only the intersubjective content of keepings one's turn remains, and the meaning is almost entirely bleached.

It can be concluded that *like* is used in ELF as a pragmatic marker, and that it indeed also performs a variety of pragmatic functions. In the data analysis I distinguished between five different categories of *like*: hedges, illustratives, quotatives, fillers and independent. Although they each serve a particular purpose in discourse, these functions can actually be subsumed under one definition, similar to what Andersen (1998) proposed in his study of the pragmatic marker *like* in native English discourse. All categories share the property of signalling loose and informal talk, decreasing the speaker's responsibility for the spoken content and offering cues to the hearer not to take the spoken content too literally. Those are actually the most salient characteristics of the category of hedges, which was also the most frequent one in my data. It can therefore be concluded that hedging is the most prototypical function of *like* in ELF, and all the other categories could possibly be understood as subcategories of hedges.

Furthermore, I looked into the domains and speech event types where the different categories of *like* appear most frequently. Although the illustrative and filler categories showed mixed results, it was found that *like* mostly tends to be used in informal contexts, such as leisure domain and conversations, which was as expected.

As for the speakers age, a case of age gradation is noticeable. However, the sample consists of mostly younger speakers, so it is difficult to make any definitive conclusions. Gender distribution analysis also did not reveal any significant patterns. In the data extracted from the corpus, women used *like* in 64.79%, and men used it in 35.21% of the cases. The number of female speakers was consistently higher than male in each of the categories, so it can be said that there were no major differences in the use between female and male speakers.

Similar functions of the pragmatic marker *like* were found in the native English discourse and the data extracted from the VOICE corpus, which consists of non-native speakers of English with varying degree of language competence and a small number of native speakers

(9.36%). So it could also be concluded that the pragmatic use of the marker *like* is, similar to Mauranen's reflexives, not restricted to native speaker's English discourse, and therefore a “fundamental feature of discourse“ (Mauranen 2010, 24).

Regardless of whether these features of *like* are fundamental or not, they do fulfill important pragmatic functions in discourse. The hedging function is of particular relevance in ELF communication, considering that the speakers may have difficulties building common ground. With its element of looseness and informality and decreasing the speaker's responsibility for the expressed facts while signalling to the hearer how to interpret the spoken content, the pragmatic marker *like* in its broad hedging function seems to be an important tool in ELF communication.

6 Literature

- Andersen, Gisle. 1998. "The pragmatic marker like from a relevance-theoretic perspective." In *Discourse Markers: Descriptions and Theory*, edited by Andreas H. Jucker and Yael Ziv. 147-70. Amsterdam: John Benjamins.
- Andersen, Gisle. 2000. "The role of the pragmatic marker like in utterance interpretation." In *Pragmatic Markers and Propositional Attitude*, edited by Gisle Andersen and Thorsten Fretheim. 17-38. Amsterdam: John Benjamins.
- Baumgarten, Nicole, and Juliane House. 2010. "I think and I don't know in English as lingua franca and native English discourse". *Journal of Pragmatics*, Vol. 42: 1184-1200.
- Blyth, Carl Jr., Sigrid Recktenwald and Jenny Wang. 1990. "I'm like, 'Say What?!': A New Quotative in American Oral Narrative". *American Speech*, Vol. 65 (3): 215-227.
- Cameron, Lynne, and Alice Deignan. 2003. "Combining Large and Small Corpora to Investigate Tuning Devices Around Metaphor in Spoken Discourse." *Metaphor and Symbol*, Vol. 18 (3): 149-160.
- Cogo, Alessia. 2010. "Strategic Use and Perceptions of English as a Lingua Franca." *Poznań Studies in Contemporary Linguistics* 46 (3): 295-312.
- Dailey-O'Cain, Jennifer. 2000. "The sociolinguistic distribution of and attitudes toward focuser like and quotative like." *Journal of Sociolinguistics*, Vol. 4 (1): 60-80.
- Davidse, Kristin, Lieven Vandelanotte, and Hubert Cuyckens. 2010. "Subjectification, Intersubjectification and Grammaticalization." Berlin: De Gruyter Mouton.
- House, Juliane. 2010. "The Pragmatics of English as a Lingua Franca". In *Pragmatics across Languages and Cultures*, edited by Anna Trosborg. 363-387. Berlin: De Gruyter Mouton.

- Jucker, Andreas H. and Sara W. Smith. 1998. "And people just you know like 'wow': Discourse markers as negotiating strategies." In *Discourse Markers: Descriptions and Theory*, edited by Andreas H. Jucker and Yael Ziv. 171-201. Amsterdam: John Benjamins.
- Lehmann, Christian. 2002. "Thoughts on Grammaticalization". *Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt*, Nr. 9. Philosophische Fakultät. Universität Erfurt.
- Mauranen, Anna. 2010. "Discourse reflexivity-a discourse universal? The case of ELF." *Nordic Journal of English Studies*, Vol. 9 (2): 13-40.
- Mulder, Jean and Sandra A. Thompson. 2008. "The Grammaticalization of but as a Final Particle in English Conversation". In *Crosslinguistic Studies of Clause Combining: The Multifunctionality of Conjunctions*, edited by Ritva Laury. 179-204. Amsterdam: John Benjamins.
- Romaine, Suzanne, and Deborah Lange. 1991. "The Use of like as a Marker of Reported Speech and Thought: A Case of Grammaticalization in Progress". *American Speech*, Vol. 66 (3): 227-279.
- Traugott, Elizabeth. 2010. "(Inter)subjectivity and (Inter)subjectification: A Reassessment." In *Subjectification, Intersubjectification and Grammaticalization*, edited by Kristin Davidse, Lieven Vandelanotte, and Hubert Cuyckens, 29-71. Berlin: De Gruyter Mouton.
- Underhill, Robert. 1988. "Like Is, like, Focus". *American Speech*, Vol. 63 (3): 234-246.
- VOICE. 2013. *The Vienna-Oxford International Corpus of English* (version 2.0 online).
 Director: Barbara Seidlhofer; Researchers: Angelika Breiteneder, Theresa Klimpfinger, Stefan Majewski, Ruth Osimk-Teasdale, Marie-Luise Pitzl, Michael Radeka.
<http://voice.univie.ac.at> (15/6/2014).

7 Appendix

Table 1

Pragmatic marker *like* in different domains

Domains	Number of examples	Percentage
Education	187	35.02%
Leisure	175	32.77%
Professional organizational	109	8.61%
Professional business	46	20.41%
Professional research and science	17	3.18%

Table 2

Pragmatic marker *like* in different speech event types

Speech event	Number of examples	Percentage
Conversation	214	40.07%
Workshop discussion	116	21.72%
Working group discussion	91	17.04%
Meeting	54	10.11%
Seminar discussion	23	4.31%
Interview	15	2.81%
Panel	11	2.06%
Question-answer session	7	1.31%
Service encounter	2	0.37%
Press conference	1	0.19%

Table 3

Categories of *like* and their frequency in the data

Category	Number of examples	Percentage
Hedge	249	46.63%
Illustratives	155	29.03%
Quotatives	56	10.49%
Fillers	50	9.36%
Independent use of <i>like</i>	24	4.49%

Table 4
Hedges in different domains

Domain	Number of examples	Percentage
Leisure	93	17.42%
Education	92	17.23%
Professional organizational	43	8.05%
Professional business	14	2.62%
Professional research and science	7	1.31%

Table 5
Hedges in different speech event types

Speech event	Number of examples	Percentage
Conversation	116	21.72%
Workshop discussion	50	9.36%
Working group discussion	38	7.12%
Meeting	18	3.37%
Seminar discussion	13	2.43%
Interview	7	1.31%
Question-answer session	3	0.56%
Panel	2	0.37%
Service encounter	2	0.37%

Table 6
Illustratives in different domains

Domain	Number of examples	Percentage
Education	59	11.05%
Professional organizational	34	6.37%
Professional business	28	5.24%
Leisure	27	5.06%
Professional research and science	7	1.31%

Table 7**Illustratives in different speech event types**

Speech event	Number of examples	Percentage
Workshop discussion	38	7.12%
Conversation	35	6.55%
Working group discussion	30	5.62%
Meeting	29	5.43%
Interview	7	1.31%
Panel	7	1.31%
Seminar discussion	6	1.12%
Question-answer session	2	0.37%
Press conference	1	0.19%

Table 8**Quotatives in different domains**

Domain	Number of examples	Percentage
Leisure	38	7.12%
Education	11	2.06%
Professional organizational	5	0.94%
Professional research and science	1	0.19%
Professional business	1	0.19%

Table 9**Quotatives in different speech event types**

Speech event	Number of examples	Percentage
Conversation	41	7.68%
Working group discussion	5	0.94%
Workshop discussion	4	0.75%
Meeting	2	0.37%
Seminar discussion	2	0.37%
Interview	1	0.19%
Panel	1	0.19%

Table 10**Fillers in different domains**

Domain	Number of examples	Percentage
Professional organizational	22	4.12%
Education	17	3.18%
Leisure	7	1.31%
Professional research and science	2	0.37%
Professional business	2	0.37%

Table 11**Fillers in different speech event types**

Speech event	Number of examples	Percentage
Workshop discussion	21	3.93%
Working group discussion	13	2.43%
Conversation	10	1.87%
Meeting	3	0.56%
Question-answer session	2	0.37%
Panel	1	0.19%

Table 12**Independent use of *like* in different domains**

Domain	Number of examples	Percentage
Leisure	10	1.87%
Education	8	1.50%
Professional organizational	5	0.94%
Professional business	1	0.19%

Table 13**Independent use of *like* in different speech events**

Speech event	Number of examples	Percentage
Conversation	12	2.25%
Working group discussion	5	0.94%
Workshop discussion	3	0.56%
Meeting	2	0.37%

Table 14

Age groups and their usage of *like*

	Number of examples	Percentage
Speakers aged 17-24	336	62.92%
Hedge	158	29.59%
Illustratives	86	16.10%
Quotatives	47	8.80%
Fillers	28	5.24%
Independent use of <i>like</i>	17	3.18%
Speakers aged 25-34	105	19.66%
Hedges	56	10.49%
Illustratives	25	4.68%
Fillers	14	2.62%
Independent use of <i>like</i>	5	0.94%
Quotatives	5	0.94%
Speakers aged 35-49	73	13.67%
Illustratives	31	5.81%
Hedges	30	5.62%
Fillers	6	1.12%
Quotative	4	0.75%
Independent use of <i>like</i>	2	0.37%
Speakers aged 50 and older	20	3.75%
Illustratives	13	2.43%
Hedges	5	0.94%
Fillers	2	0.37%