

## 2. Literature review

### 2.1. Quotatives

Quotatives are markers used to start a quotation or quote, which are utterances repeated word for word and attributed to another speaker. They are used for dramatic effect (Blyth, Recktenwald, & Wang, 1990) and to introduce gestures, nonlexicalized sounds, and reported speech (either direct speech or inner monologue). While the verb *to say* was the prototypical traditional quotative in the English language (Wikström, 2014, p. 83), in the last decades several quotatives have appeared in English and other languages. Speakers who use a quotative also express what the probability of the quote having been actually uttered is. This means that certain quotatives express certain degrees of hypotheticality (i.e. the probability of the realization of a state of affairs, event, or action).

Regarding the categorization of quotatives, there has been a number of proposals to classify the various quotatives in the English language. For example, a number of authors use chronological criteria to talk about ‘innovative’ and ‘new’ quotatives (Blyth, Recktenwald, & Wang, 1990; Romaine and Lange, 1991; Ferrara and Bell, 1995, as cited by Tagliamonte, & Hudson, 1999). A much clearer distinction can be made observing their structure. Following this idea, Buchstaller (2015, p. 462) proposes a characterization of quotatives with straightforward examples, making a clear distinction between old forms (*to say*, *to think*, and *to decide*) and new forms (*be like*, *kind of*, and *zero*):

(1) Old forms: NOUN PHRASE + TRANSITIVE VERB OF REPORTING +  
QUOTE

- a) If you touched a one they would say “wey you’re on”
- b) And I thought “Well we need some more popcorn”.
- c) And then I decided “oh I want to do art after all”

(2) New forms: NOUN PHRASE + COPULA + (DISCOURSE MARKER) +  
QUOTE

- a) She's like "This is ridiculous"
- b) My mum feels kinda "bleugh"
- c) He goes Ø "[makes face]"

However, what sets the new forms of quotatives apart is that they allow the speaker to dramatize narration (especially of personal experiences) and summarize certain parts, such as dialogue (Winter, 2002, p. 7). Another characteristic is their variable hypotheticality level. Following Buchstaller, this is

“the probability of the realization of a state of affairs, event, or action. In the case of reported speech, which is the reiteration of words, sounds, and gestures that have already been produced, there is a relationship between the quote and the original utterance. We can distinguish between utterances that could have been spoken out loud (with various degrees of probability) and those that were spoken out loud” (2002, p. 5).

This allows the speaker to specify when the actual utterance existed, and when there was not necessarily one but there was something to report, such as gestures, the thoughts of the speaker, of the other participants, or even of the narrator. From the traditional quotatives, *think* had this function specifically, but the new quotatives take this function to a different level. “Contrarily to *think*, *like* and *go* theatricalize inner speech by outwardly displaying it in vivid, emotionally heightened output. And, in contrast to *say*, they do not pin down a quote as to its hypotheticality level. The newly grammaticalized quotatives *go* and *like* are

distinguished from the old quotatives *say* and *think* by their function as mimesis markers”

(Buchstaller, 2002, p. 15).

For the English language, the quotatives *be like*, *go*, *kinda*, *sorta*, *all like*, *go totally*, and *be all* (Alphen & Buchstaller, 2012) are just some of the newer ones compared to *think* and *say*, and many other equivalents have been found in other languages. In the table below, we can observe the semantic category several quotatives belong to. The first and second category are very relevant for this study, the reasons for which will be explained below.

Table 1. (New) quotatives and their semantic sources (Alphen & Buchstaller, 2012)	
i. Comparative (similarity / approximation):	
<ul style="list-style-type: none"> <li>• Afrikaans soos ‘so + as’</li> <li>• Czech jako (že) ‘as’</li> <li>• Buang (na) be ‘like’</li> <li>• Danish ligesom ‘like + as’</li> <li>• Dutch van ‘like’</li> <li>• English like</li> <li>• Estonian nagu ‘like’</li> <li>• Finnish niinku (niin kuin) ‘as if’, ihan et ‘like’</li> <li>• French comme ‘like’ genre ‘kind (of)’, style ‘style’</li> <li>• Frisian fan ‘like’</li> </ul>	<ul style="list-style-type: none"> <li>• Greek tipou ‘type’, genere ‘kind’, Japanese mitai-na ‘like’</li> <li>• Croatian kao ‘like’, tipa ‘type’</li> <li>• Swedish typ ‘type’, liksom ‘like + as’</li> <li>• Norwegian typ ‘type’, liksom ‘like + as’</li> <li>• Polish typu ‘type’</li> <li>• Portuguese tipo ‘type’</li> <li>• Brazilian Portuguese tipo + assim ‘type + so’</li> <li>• Russian tipa ‘type’</li> <li>• Spanish como ‘like, as’</li> <li>• Thai bæ:p ‘like’</li> </ul>
ii. Demonstrative deictic:	
<ul style="list-style-type: none"> <li>• Afrikaans soos ‘so + as’</li> <li>• Czech na to ‘on this’</li> <li>• Danish sådan ‘such + like + this’</li> <li>• Dutch zo ‘so’</li> <li>• Estonian nii et ‘so that’</li> <li>• London English this/here is NP</li> </ul>	<ul style="list-style-type: none"> <li>• Hebrew kaze ‘like + this’</li> <li>• Croatian ono ‘that’, ono kao ‘that + like’</li> <li>• Brazilian Portuguese (tipo+) assim ‘(type+) so’</li> <li>• Russian ta-koj ‘such + like + this/that’</li> <li>• Spanish asi ‘so’</li> </ul>

<ul style="list-style-type: none"> <li>• Finnish et (tä) ‘that’</li> <li>• German so ‘so’</li> </ul>	<ul style="list-style-type: none"> <li>• Norwegian sånn ‘such + like + this/that’</li> <li>• Swedish såhär (sär) ‘such + like + this/that’</li> </ul>
iii. Quantifiers:	
<ul style="list-style-type: none"> <li>• Danish bare ‘just, only’</li> <li>• Dutch helemaal ‘all’</li> <li>• English all</li> <li>• Estonian täiega ‘to-tally’</li> </ul>	<ul style="list-style-type: none"> <li>• Finnish vaa(n) ‘just’</li> <li>• Icelandic bara ‘just, only’</li> <li>• Norwegian bare ‘just, only’</li> <li>• Swedish ba(ra) ‘just, only’</li> </ul>
iv. Generic verbs of motion and action:	
<ul style="list-style-type: none"> <li>• English <i>go</i></li> <li>• Dutch komen ‘to come’</li> </ul>	<ul style="list-style-type: none"> <li>• Greek kano ‘do’</li> <li>• Puerto Rican Spanish hacer ‘do, make’</li> </ul>

The first category, which denotes similarity or comparison, allows the speakers to point out it is possible the quotation is not likely to be 100% accurate to the original, due to factors such as their own memory. This allows them to have a justification and not have to compromise themselves by declaring something is an accurate report when it is possible it's not. Fox Tree & Tomlinson (2008, p. 95) also mention this for *be like* in particular, (which will be more relevant in the sub-chapter ***Be like***), which may explain why so many languages (especially Indo-European) have a quotative that uses *like*.

Quotatives which have a demonstrative deictic function focus instead on the performative aspect of quotation: rather than suggest the quote is not accurate, the users allow listeners to see and hear for themselves how the original utterance was emitted. As the examples in the table suggest, most forms involve variations of *like this*.

The third category, which involves quantifiers or quantificational semantics, allows speakers to express their position or opinion with respect to the quote. For example, the use of *all* in English and *täiega* (to-tally) in Estonian expresses a characteristic that permeated the reported quote, while the use of *bare* in Danish and *helemaal* in Dutch, (both which translate to ‘just, only’) expresses a weak characteristic barely present in the quote.

The fourth category describes verbs of motion and action which have been connected to quotatives through the use of metaphors (Sweetser, 1987, as mentioned by Alphen & Buchstaller, 2012). A conversation is something that has to be maintained by its speakers, making it an effort that can be easily related to verbs of motion and action.

Much research has been done regarding these modern quotatives in English. Singler (2001, p. 257) mentions that the quotative *go* has been around since the 40’s and 50’s, while Buchstaller (2006b, p. 373) comments that “its first attestation in the OED (in 1791) was after the war of independence in a British English source”. Unfortunately, there has not been much proper research on the origins of *go* as a quotative compared to its popular American counterpart to further explore its early expansion.

According to Butters (1980, 1982, as mentioned in Macaulay, 2001, p. 5) the quotative *go* has been around since 1980, and *be like* appears two years later. *Go* is used solely to introduce direct speech, both in American English (Schourup, 1982, as mentioned in Blyth, Recktenwald, & Wang, 1990, p. 216) and in British English, but is favored with internal dialogue in Canada (Tagliamonte & Hudson, 1999, p. 164). It became common among adolescents and younger speakers, its place rapidly taken by *be like* in recent decades (Palacios, 2014).

### **2.1.1 Constructed Dialogue and Historical Present**

In storytelling, the way one recalls or reports the dialogue of the involved characters depends on the preferences of the narrator. These decisions include the degree of accuracy and embellishment, and the use of language to involve the reader. According to Tannen (1986, p. 311), taking into account human memory, it would be a misnomer to brand the recalling of previous dialogue as ‘reported speech’, as not only it is very likely the quote is not accurate, it may have never been spoken at all. In its place, what a speaker does is *reconstruct* this dialogue.

For Wolfson (1982, p. 3), “the historical present tense, known also as the dramatic or the narrative present tense, is the use of the present tense, in narrative to refer to events which began and ended at some time previous to the moment at which the narrative itself is told.” The use of a different tense for the purposes of better storytelling (Schiffrin 1981, and Johnstone, 1987, as mentioned by Blyth, Recktenwald, & Wang, 1990) from the point of view of the narrator is similar to the use of direct speech instead of the more appropriate indirect speech. While Tannen (1986) holds every instance of reported speech as nothing but constructed dialogue, the probability of an actual speech having taken place will be the focus of this study. Similarly, the use of present tense by speakers will be assumed to be historical present tense.

### **2.1.2 *Be like***

While *be like* as a quotative is relatively new compared to its fellow innovative quotatives, it has characteristics that make it stand out from the others. While *go* and *say* usually report an utterance that truly happened, *be like* can be used to represent a thought, a state of mind, an inner monologue, or any other sort of utterance that actually took place while retaining

its vividness. (Blyth, Recktenwald, & Wang, 1990; Romaine and Lange, 1991). Related to this, the use of *be like* allows speakers to disentangle themselves from the responsibility of being perfectly accurate to the source reports. This peculiarity is explained by the *approximation theory* (Andersen, 1998, 2000, Buchstaller, 2002, and Romaine and Lange, 1991, as mentioned by Fox Tree, & Tomlinson, 2008, p. 87). Similarly, the *demonstration theory* suggests certain aspects of the original utterance are given priority so they can successfully carry the core meaning or to give the listener the opportunity to understand what it was like to be there and hear it (Clark & Gerrig, 1990, and Wade & Clark, 1993, as mentioned by Fox Tree & Tomlinson, 2008, p. 88). These aspects may be the content of the utterance that have the most emotional impact, or they may be certain characteristics such as tone of voice and the speaker's accent, which can have a strong impact on the listener, which contrasts with the use of *to say*, which signifies a strong confidence in the relative accuracy of a quote that was actually uttered (Blyth, Recktenwald, & Wang, 1990, p. 215; Fox Tree & Tomlinson, 2008, p.88). Alongside *go*, for Blyth, Recktenwald, & Wang, these quotatives

“typically introduce a perfective (i.e. completed and punctual) speech act, whereas *be like* may be either perfective or imperfective according to its discourse function. Thus when *be like* is interpreted as imperfective it introduces a thought, inner monologue, or a gestalt which summarizes the speaker's frame of mind; when perfective it introduces direct speech” (1990, p. 222).

This all-in-one flexibility of *be like* may be what makes it stand out from other quotatives, as shown in several studies that compare it with both old and new (Ferrara and Bell, 1994,

as mentioned by Tagliamonte, & Hudson, 1999; Winter, 2002). Probably also as result of this flexibility, these studies show that a significant number of speakers used this quotative.

The increase of speakers that use *be like* can also be attributed to its fast spread thanks to a “generalization and loss of selectional restrictions of *like*” (Buchstaller, 2001, p. 27). This means that, instead of being used for the speaker’s own speech, and the other verbs of saying for the interlocutor or others as in Romaine and Lange’s study in 1991 (p.242), it is now used for the narrator and other speakers (Buchstaller, 2001).

Another quotative is *all*, which also had its origins in 80s California. In the early 90s, high school students used this quotative above the others, but its use did not carry over to college-age speakers. It was used to introduce reported speech with present-tense verbs. Ten years later, it had fallen into disuse compared to its other quotatives, and now the college-age speakers used it more than high school students. This represents a classic case of retraction (Haspelmath, 2004, as mentioned by Rickford, Wasow, Zwicky, & Buchstaller, 2007, p. 22), where new variants start being used but are eventually replaced by other, more dominant variants. Above, *go* was this newcomer variant that enjoyed certain use for some time, until speakers replaced it with *be like*.

This is a more straightforward way of showing *be like*’s dominance as a quotative, and with it, its importance in modern and future speech.

#### **2.1.2.1 Who supposedly uses it?**

Research on attitudes done in the past two decades (Blyth, Recktenwald, & Wang, 1990; Buchstaller, 2006b; Ferrara and Bell, 1995, as mentioned by Dailey-O’Cain, 2000; Romaine and Lange, 1991) suggests that most people (both the interviewed and the researchers) believed women, young people, and working-class people were the ones who used *be like*



the most, both as a focuser (which emphasizes a certain part of the utterance) and as a quotative. For some of British respondents of Buchstaller's study (2006b, p. 374), *be like* came from the U.S. This supposed relationship between women and *like* has its origins in the Valley Girl stereotype, which includes so-called air-headed Californian girls who were seen as having the tendency to include the discourse filler *like* in their speech (Blyth, Recktenwald, & Wang, 1990, p. 224). In this case, *be like* as a quotative is not the primary target of this association, but any use of the word *like* over the average. For comparison, quotative *go* was an indicative of "uneducated, lower-class males and the use of *be like* indicative of middle-class teenage girls. Typical epithets to describe users of *go* were 'jocks', 'blue-collar', 'men like Rocky.'

#### **2.1.2.2 Opinions**

As teenagers tend to be the primary indicators of language change, they are usually pointed out by language purists as the primary offenders of degenerating the language. For this reason, and because the general opinion has teenagers as the primary users of *be like*, this quotative has been the target of negative opinions. "A particularly interesting aspect of the social context of language change is that from a diachronic perspective, the cumulative effects of change are unexceptional, yet in synchronic time individual changes are synonymous with degradation." (D'Arcy, 2007, p. 387). From the perspective of these critiques, any sort of change is seen as an attack on the language by either ignorant or lazy young people who add new functions to existing lexicon that does not need them.

### 2.1.2.3 Who truly uses them?

The earliest mention of *be like* and its use appears in Tannen (1986, p. 324), where it is believed to be “casual speech of middle class American teenagers.” In a study done in 1990, the results showed that younger speakers were the primary users, but that the use “dropped off sharply after the age of 25 and disappeared altogether at the age of 38” (Blyth, Recktenwald, & Wang, p. 219).

Ferrara and Bell (1995, as mentioned by Tagliamonte & Hudson, 1999, p. 151) observed that in their data from 1990, *be like* was favored by women, which was also the same result Romaine and Lange saw (1991), but by 1992, Ferrara and Bell (as mentioned by Buchstaller, 2001, p. 39) found both genders used it equally, possibly by a neutralization of the sex effect due to increasing expansion. In the same study, it was found that “that the oldest speaker that actually used it as a discourse marker was 39”. To further ‘legitimize’ this innovative quotative, the Random House Webster dictionary (1999, as mentioned by Buchstaller, 2001, p. 2) also added the use of *like* as a quotative, noting this new use as “Informal (used esp. after forms of ‘to be’ to introduce reported speech or thought)”. In a later study, teenagers continued to use *be like* more times than the other groups (Dailey-O’Cain, 2000, p. 67; Cukor-Avila, 2002), but by 2006, this was completely inverted compared to what was reported in Blyth, Recktenwald, & Wang’s 1990 study, as it started high in 14-19, higher in 20-25, before [...] taking a plunge. However, its use stayed at the same level between 38 and 56, before disappearing (Buchstaller, 2006a, p. 9). Similarly, men were found to use *like* both as a focuser and as a quotative more than women (Dailey-O’Cain, 2000, p. 75), going against Ferrara and Bell’s theory of a neutralization of the sex effect due to expansion. Barbieri’s 2007 study shows a more balanced pattern of use, where

the quotative *be like* is the favored choice for young women in their late teens and early to mid-20s (ten times more frequently than their male peers do), but its use decreases dramatically among women in their late 20s. Men display a reversed pattern of use: Teenagers and college-aged men rely overwhelmingly on the traditional quotative *say*, using the new quotatives *be like*, *go*, *be all* relatively infrequently. In contrast, men in their late 20s and in their 30s favor *be like* and *go* (p. 26).

Similarly, in Barbieri (2007) women in general were found to use quotatives more than males (four times more for *go*, and twice as much for *be all*). Men between the ages of 27 and 40 use newer quotatives, though. But from all the quotatives, *be like* is the one most used by both sexes. A hypothesis developed by D'Arcy (2007, p. 397) suggests that the different uses of *like* are used more frequently by a particular sex: women use it as a quotative and as a marker, men use it as a particle, and both use it as an adverb.

In yet another study about the evolution of *be like*, Barbieri (2009) compared the results of a corpus from the middle 90's to one from 10 years later. This way, she was able to observe the development of *be like* in American English through time. Her conclusion was that *be like* was now the most preferred quotative by both male and female speakers under the age of 40. Thanks to her analysis, she also showed that speakers who used *be like* when they were teenagers (the demographic that used it the most), not only kept using them when they became adults ten years later, but also increased their use. In 2004-2005, *be like* was favored by teenagers below age 16, and in this group, girls used it more, which suggests they are the ones at the front of the changes *be like* undergoes. In adults, women maintained their use of *be like* but in a lesser frequency.

On the other side of the pond, the quotative *be like* also increased rapidly in use. “Among York undergraduates, *be like* has become the most frequently used quotative, increasing from 19 percent in 1996 to 68 percent in 2006. This increase comes at the expense of *be like*’s main competitors- *say*, *think*, and, to some extent, *go*- all of which show lower rates of use in the 2006 sample than in the 1996 data.” (Durham, Haddican, Zweig, Johnson, Baker, Cockeram, Danks, & Tyler, p.323). Before this increase, the dominant quotative (at least among working class girls) in 1997 was *go*, which was used 26% of the time. *Be like* was barely used 14% of the time. However, middle-class girls used *be like* 56% of the time. Adults (especially women) used traditional quotatives such as *say* and *think* (Macaulay, 2001). Later studies showed that, not unlike the case in American English, the use as a quotative was high in 14-16 year old speakers, higher in 17-19 year olds/o, before taking a big dip in from 20+ onwards (Buchstaller, 2006a).

#### **2.1.2.4 Other uses of *like***

Searching in the Merriam-Webster online dictionary (2016) for the word *like* shows that it has a wide range of uses: as a verb (2 uses), a noun (4 uses), an adjective (2 uses), a preposition (2 uses), an adverb, a conjunction (2 uses, one which includes its quotative use), a verbal auxiliary, an adjective combining form, and as an adjective suffix.

Buchstaller (2001) claims that the many functions of *be like* are actually interrelated to each other, such as:

- *From the comparative meaning to a hedge*

A: If someone slips on the ice outside a building, they could almost sue the architect for not having put ice-melting equipment in the sidewalk

B: Yeah, into the sidewalk like heaters. But then...

- *From a comparative to a filler*

A: So what else do you have apart from history

B: I have, I take, um, I'm taking math, and it's like, it's called Sequential Bias, which next year I take Precalculus

- *From comparison/approximation to focus*

A: And he was like put up in a house, Cambridge and everything, was just amazing.

This 'evolution' also applies to *be like* as a quotative, via a process known as grammaticalization or morphologization: "a unidirectional historical process whereby lexical items acquire a new status as grammatical or morphosyntactic forms" (Romaine and Lange, 1991, p. 257). Blyth, Recktenwald, and Wang expand on Underhill's (1988, as mentioned in their 1990 study) analysis of the use of *like* as a discourse filler and suggest *be like* is a variant of it. As the primary use of the focus *like* is to bring attention to the next sentences as if they were significant, *be like* as a focus quotative may be used as "a quotative which introduces a particularly salient piece of information packaged in the form of reported speech." (Blyth, Recktenwald, and Wang, 1990, p. 225) Because of this, "*Be like* is functionally versatile and therefore may have more staying power in the lexicon."

Regarding this last study, Blyth, Recktenwald, and Wang notice that there are only few uses where the PERSON + VERB TO BE + *LIKE* formula is possible. So for these authors, the following uses are the most relevant.

- Preposition: Similar to (something or someone), typical of (someone), comparable to or close to (something). This is the most common and well known use of *like*, for example in ‘Tell us what you spoke about and what she was like?’
- Adverb: Nearly, approximately, used interjectionally in informal speech with expressions of measurement. Example:
  - “[A:] She was like two?
  - [B:] Yeah, it was before her second birthday. She was one.” (Davies, 2008)
- Focuser: Used interjectionally in informal speech often to emphasize a word or phrase or for an apologetic, vague, or unassertive effect. It is a “discourse or pragmatic marker, similar to *you know* or *well*. These markers are considered to be optional rather than required, to have little inherent meaning or at least meanings which cannot be easily specified lexically, and to also have no clear grammatical function” (Schiffrin, 1986, as cited by O’Cain, 2000, p. 61).
- Quotative: Used interjectionally in informal speech often with the verb *be* to introduce a quotation, paraphrase, or thought expressed by or imputed to the subject of the verb, or with *it’s* to report a generally held opinion.
- Response cry: Highly conventionalized utterance which makes the inner state transparent to the audience (Goffman, 1978, p. 283, as cited by Fox & Robles, 2010, p. 717), or a non-lexicalized interjection (p. 800) which typically consists of stock phrases. In this study, common stock phrases were *well*, *huh*, *what*, *wow*,

*okay, what the hell?, alright, oh jeez, oh okay, whoa, ow, oh boy, oh my gosh, oh man, yeah, hm, holy crap, oh yea, oh great, oh, yay!, oh oh, yeah right, etc.*

## **2.2 Corpus linguistics**

Since most of the studies just mentioned in former sections use corpora to explore the evolution of *be like* a short note on this approach for the study of language is due. Rather than being specific to linguistics, the use of corpora in linguistics is “a methodology comprising a large number of related methods which can be used by scholars of many different theoretical leanings” (Lindquist, 2009, p. 1) In the study of language, it involves the use of large bodies of texts of authentic and language representative texts (the aforementioned corpora) to observe varied linguistic phenomena such as language change, overarching themes, and characteristics in people’s speech. In modern times, the use of computer corpora adds a bit to the basic definition: a corpus now is made “in machine-readable form which may be annotated with various forms of linguistic information” (McEnery, Xiao, & Tono, 2006, p. 4). Other uses include the development of grammar books and dictionaries. In the past, the text collecting was complicated work that took a long time to finish, but nowadays with the help of computers it is possible to both create corpora and make use of them more easily. *A Modern English Grammar on Historical Principles* (published between 1909 and 1949), a multi-volume work by Otto Jespersen, was created thanks to a corpus made from texts from English literature, painstakingly collected by its author, whose work was finished posthumously (Lindquist, 2009). On the other hand, the Brown corpus, which compiled the English spoken in 1961, was ready by 1964. This was the first machine-readable corpus to be completed. Whereas at that time this

1-million-word corpus was considered a huge success, modern corpora can hold several hundreds of millions of words.

The Brown corpus was also historically important in that it is a good example of a balanced corpus, a corpus that “usually covers a wide range of text categories which are supposed to be representative of the language or language variety under consideration.” (McEnery, Xiao, & Tono, 2006, p. 16). It also set the standard for modern corpus-based research, including the principle of free access, the structure used, and the method used to select the samples it contains (Kennedy, 1998, p. 26-27).