

IT'S *LIKE* 120 MILLISECONDS: A SEARCH FOR GRAMMATICALIZATION IN THE
DURATION OF *LIKE* IN FIVE FUNCTIONS

by

FRANCES R. GRAY

(Under the Direction of Don R. McCreary)

ABSTRACT

This thesis addresses the discourse and quotative marker functions of *like* from a grammaticalization perspective. Many of the functions share a polysemous thread of meaning because of their historical relatedness; in particular, the marker functions grammaticalized from the function words. Grammaticalization has certain features which characterize it, including phonological reduction. Though the marker functions of *like* are not yet segmentally changed from the functions from which they grammaticalized, this study hypothesizes that the durations of the grammaticalized morphemes will be shorter than the nongrammaticalized forms. To test this hypothesis, the speech of four young-adult speakers from the Buckeye corpus was analyzed regarding the durations of five functions of *like*, though the hypothesis was not upheld. In fact, the marker functions were found to have the longest mean durations.

INDEX WORDS: Grammaticalization, discourse markers, quotative markers, Buckeye corpus.

IT'S *LIKE* 120 MILLISECONDS: A SEARCH FOR GRAMMATICALIZATION IN THE
DURATION OF *LIKE* IN FIVE FUNCTIONS

by

FRANCES R. GRAY

Shepherd University, Shepherdstown, West Virginia, 2008

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2010

© 2010

Frances R. Gray

All Rights Reserved

IT'S *LIKE* 120 MILLISECONDS: A SEARCH FOR GRAMMATICALIZATION IN THE
DURATION OF *LIKE* IN FIVE FUNCTIONS

by

FRANCES R. GRAY

Major Professor: Don R. McCreary

Committee: Lewis C. Howe
Keith Langston

Electronic version :

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2010

ACKNOWLEDGEMENTS

I have so much gratitude for the people who have helped guide me through this process, both academically and personally. First, I would like to thank Dr. McCreary for his guidance and helpful advice. I would never have finished this project without his willingness to stick with me. I also want to thank Dr. Howe for suggesting this project that made me excited to get to work. Like Dr. McCreary, he has also been very patient with me during this process. Dr. Langston and Dr. Anya Lunden have also been very valuable to this process, and I want to thank them both for their quick willingness to respond to my requests for help. Additionally, my parents, Bob and Susan Gray, and my grandparents, Bill and Edna Gray, have always encouraged me and given me more support than I could have asked for, both tangible and intangible. Finally, I have to thank my fellow graduate students Magdalene Jacobs and Franchesca Judd for making graduate school a worthwhile pursuit in so much more than academics.

TABLE OF CONTENTS

| | Page |
|--|------|
| ACKNOWLEDGEMENTS..... | iv |
| CHAPTER 1: INTRODUCTION..... | 1 |
| CHAPTER 2: REVIEW OF LITERATURE ON FUNCTION AND GRAMMATICALIZATION OF <i>LIKE</i> | 4 |
| 2.1 THE FUNCTIONS OF <i>LIKE</i> | 4 |
| 2.2 GRAMMATICALIZATION OF <i>LIKE</i> | 23 |
| 2.3 DURATION AND ATTRITION..... | 28 |
| CHAPTER 3: DESCRIPTION OF THE RESEARCH AND PROCEDURE | 32 |
| 3.1 CORPUS..... | 32 |
| 3.2 SELECTED SPEAKERS | 33 |
| 3.3 DURATION DETERMINATION AND EXTRACTION | 34 |
| 3.4 DETERMINATION OF FUNCTION | 35 |
| 3.5 INDETERMINACY OF FUNCTION..... | 39 |
| CHAPTER 4: RESULTS..... | 41 |
| 4.1 SELECTION OF SPEAKERS..... | 41 |
| 4.2 STATISTICAL RESULTS FOR SPEAKER VARIABLE | 42 |
| 4.3 STATISTICAL RESULTS FOR FUNCTION VARIABLE | 44 |
| CHAPTER 5: CONCLUSIONS | 46 |
| 5.1 DISCUSSION OF THE RESULTS..... | 46 |
| 5.2 AMBIGUITY OF DURATION..... | 48 |

| | |
|--|----|
| 5.3 SOCIOLINGUISTIC OBSERVATIONS | 50 |
| 5.4 OTHER CONCERNS..... | 51 |
| 5.5 CONCLUSION..... | 52 |
| REFERENCES | 54 |
| APPENDIX I | 58 |
| APPENDIX II | 60 |

CHAPTER 1: INTRODUCTION

Grammaticalization is a process of semantic change wherein a lexical morpheme takes on grammatical function (Hopper & Traugott 1993:4). This process of language change is important in the study of linguistics because it explains much of language development. Language change in general is as old as language itself, though speakers can be as theoretical opposed to language change as they are intolerant of language variation, from which language change is advanced. When groups of speakers, divided by age or gender or some other social category, begin to use some variable differently, they are often ridiculed for its use. The discourse marker and quotative marker functions of *like* are a choice example of this linguistic innovation and subsequent scorn; these markers are as salient in the speech of young Americans as they are annoying to some. Yet the young speakers who employ these markers are adding to the grammatical system of the English of the future.

Units of language undergoing the process of grammaticalization show in microcosm the developmental processes of language writ large (Romaine & Lange 1991:258). Specifically, since discourse markers “play a crucial role in the theory of grammaticalization,” the marker functions of the English morpheme *like* make an ideal example of grammaticalization in progress (Andersen 2001:33). The term *grammaticalization* also includes the process of grammatical forms becoming “more grammatical” as well as lexical entries becoming grammatical morphemes (Romaine & Lange 1991:257). In the case of *like*, a grammatical lexeme functioning as a preposition or conjunction becomes more grammaticalized when it begins to be used for discourse and quotative functions (Romaine & Lange 1991:261). Some traits of grammaticalization are immediately evident in the profile of *like*, such as its wide range of

functions and meanings and the fact that it is a short, single-syllable morpheme, which is more likely to undergo grammaticalization.

Grammaticalization, which is a synchronic as well as a diachronic phenomenon, is more often studied from a diachronic perspective. Hence grammaticalization is usually relatively clear because it is seen at completely separate points in time, such as in studies of one possible consequence of grammaticalization, phonological reduction (Romaine & Lange 1991:258). The phonological reduction from *going to* to *gonna*, for example, is studied now after grammaticalization of the informal form *gonna* is complete, including obvious reduction (Hopper & Traugott 1993:2). However, this study argues that a synchronic examination of more subtle phonological reduction from grammaticalization in process, specifically shorter duration, can reveal on a smaller scale the same effects that are seen in a diachronic look at grammaticalization's beginning (*going to*) and end phases (*gonna*).

In contrast with the perspective from which grammaticalization is typically studied, the “diachronic development has been largely ignored” for the marker *like* (Brinton 1996:7-8). *Like* has received much attention in the literature, however, for its synchronic points of interest. Its innovative use by younger American speakers serves many discursive functions depending on context and speaker intent, making it an interesting challenge to unpack the incentive for *like*'s use as it develops and evolves. This study seeks to observe evidence of the diachronic development of the marker functions of *like* through a study of synchronic effects of grammaticalization. The effects of the grammaticalization that *like* is undergoing should be visible synchronically through slight phonetic attrition or weakening of the grammaticalizing forms, i.e. the marker functions of *like*.

If *like* is in the process of grammaticalization, the grammaticalizing functions should show phonological attrition in the form of shorter duration. Operating under the hypothesis that the grammaticalizing marker functions should show grammaticalization in progress by exemplifying shorter durations than the non-grammaticalized functions, this study examines durations of five different functions of *like* found in a sample taken from the Buckeye Corpus (Pitt et al. 2007).

This study hypothesizes that *like* will occur with different durations depending on its function. Gahl (2008) and Gosy & Horvath (2010) show that when homophones have differing frequencies within a language, the different functions manifest differing durations corresponding to their frequency. In other words, the various functions of *like* have different frequencies, so their duration will vary depending on their function: more frequent, more grammaticalized marker uses will have shorter mean durations and less frequent, less grammaticalized uses will be longer in duration.

CHAPTER 2: REVIEW OF LITERATURE ON FUNCTION AND THE GRAMMATICALIZATION OF *LIKE*

2.1 THE FUNCTIONS OF *LIKE*

Andersen colorfully calls *like* “notorious for its functional complexity and distributional versatility” (2001:210); *like* has a particularly rich array of functions in American English, including lexical items, function words, and marker forms. This diversity of use is a feature of forms in the process of grammaticalization: a morpheme’s assorted forms coexist in various states of change (Romaine & Lange 1991:258).

2.1.1. Lexical:

a. Noun

b. Verb

2.1.2. Functional:

a. Suffix

b. Conjunction

c. Preposition

2.1.3. Marker:

a. Discourse Marker

b. Quotative Marker

2.1.1. Lexical uses

a. Noun

Like has only two usages that are content words: the noun and verb forms. The noun function is used typically only in the plural and collocates with *dislikes* (OED 1989).

(1) Her odd *likes* and dislikes. (OED 1989)

The noun function of *like* is a rarity compared to the other usages of *like*. A more frequent lexical function of *like* is the verbal usage.

b. Verb

Like can be used as a transitive verb meaning “to have a taste or fondness for; be pleased with; have a preference for; enjoy,” as seen below (Neufeldt 1988:783). Example sentences which include more than one token of *like* are not uncommon. In these cases, the token in italics is the type of *like* under examination.

(2) Not that that's a big issue but I don't *like* the way he addressed it, yknow?
(s30m)¹

(3) I, *like*, told on him and that made him not *like* me for a while. (s04f)

Verbal *like* can also mean “to want or wish” (Neufeldt 1988:783).

(4) She did all that stuff that I would *like* to be done. (s04f)

(5) I'd *like* to ge-- I don't know how I would get involved in something like that but I would *like* to. (s30m)

The noun and verb functions are the only salient lexical function of *like*. The other usages of interest are all grammatical function words or grammaticalized markers.

2.1.2. Functional uses

a. Suffix

The suffix, conjunction, and preposition uses are function words, grammatical words from which the marker functions grammaticalized. The suffix form of *like* is used to form adjectives or adverbs (Neufeldt 1988:783).

¹The labels such as s30m correspond partially to the labels used in the Buckeye corpus. Here ‘s’ refers to speaker and 30 is the number assigned to the particular speaker out of 40 total. See Appendix II for the distribution of data on all the Buckeye corpus speakers. ‘m’ indicates that this speaker is male, an addition made to the Buckeye’s notation for the purposes of this study.

(6) He brought along things for the picnic, food, drinks, and such-*like*. (Romaine & Lange 1991:244)

(7) The sculpture looked quite human-*like*. (Romaine & Lange 1991:244)

The suffix function has served an important role in the development of the more grammaticalized functions of *like*, which will be addressed below.

b. Conjunction

This function of *like* works syntactically as a complementizer, introducing clauses (Adger 2003:290). The conjunction form of *like* is synonymous with *as* and *as if* (Neufeldt 1988:783).

(8) There's a –seems *like* there's a lot more to do in Columbus. (s30m)

(9) I feel *like* you don't have to go to church to believe. (s04f)

The conjunction usage of *like* has a contentious history with language prescriptivists. Though *like* has been used as a conjunction since the fourteenth century, this usage in American English has been labeled by some as nonstandard and colloquial (Romaine & Lange 1991:271). The most famous example of prescriptive derision for this usage comes from the 1950s Winston advertisement “Winston tastes good *like* a cigarette should” (Romaine & Lange 1991:271). Some grammarians at the time claimed that *like* should have been replaced with *as*, the so-called correct expression (Romaine & Lange 1991:244). Analogy with the preposition *like* has aided in establishing the conjunction *like* as a standard usage (Romaine & Lange 1991:244).

c. Preposition

Like is also used as a preposition meaning “similar to; somewhat resembling” (Neufeldt 1988:783).

(10) Y'know to look at him and go, ‘Oop, my water broke!’ Ø ‘Oh no!’ And watch him flip out y'know *like* on movies. (s21f)

(11) I can't say I'd support anyone *like* that. (s40m)

However, the Oxford English Dictionary classifies this usage as adjectival. Likewise, Meehan (1991:39) calls it specifically “an adverbial extension of the adjectival use.” In Webster’s, the adjectival function is defined as “having almost or exactly the same characteristics, etc.; similar; equal,” the essential meaning of which is also included under the definition of the prepositional form: “in a manner characteristic of; similarly to” (Neufeldt 1988:783). This study follows the analysis of Romaine & Lange (1991) and Webster’s (1988) and refers to this use of *like* as a preposition, especially since this usage is able to assign objective case.

Additionally, prepositional *like* is also defined as “for example” (Neufeldt 1988:783). Usages preceding a list or specification can be paraphrased as “for example” or “such as” (Meehan 1991:42):

(12) Marriage and kids and stuff like that with them. Um *like* they’re living together, things like that. (s40m)

According to Meehan, this prepositional use of *like* began to be used in the 19th century, whereas the original prepositional function is first observed in the 14th century (1991:42, 39).

2.1.2. Marker uses

a. Discourse marker

In addition to the lexical and grammatical usages of *like*, two marker functions have grammaticalized further: a discourse marker and a quotative marker. The discourse marker function is demonstrated here:

(13) I mean I even had problems *like* on the school bus and stuff, yknow? (s04f)
 (14) He ended up *like* running away from home, basically. (s40m)

The definition of this function is somewhat elusive; there is some disparity in how it is referred to in the literature. How the word functions pragmatically in the sentence in which it appears

changes the classification, so this variety of categorization can be attributed to the perspective of the researchers and the pragmatic purpose of interest. For the purposes of this study, this function of *like* will be referred to as a discourse marker. Some researchers name this function of *like* a discourse marker (also known as a pragmatic marker, among other terms), a filler, or a focus marker.

Discourse markers are very frequent in spoken language (Brinton 1996:6), especially in informal conversation (Brinton 1996:33). Discourse markers are “short words or phrases” inserted into spoken discourse (Brinton 1996:6). These markers are also syntactically free (Romaine & Lange 1991:252), so they do not serve a grammatical function in the sentence in which they occur (Brinton 1996:34). Discourse markers can be extricated from the sentence without affecting the propositional content of the statement (Brinton 1996:35). That is not to say, however, that these markers serve no communicative purpose in discourse.

Discourse markers have reduced lexical meaning, but as they grammaticalize from entries of the lexicon, they gain more grammatical and pragmatic meaning (Romaine & Lange 1991:266). *Like* as a discourse marker affects the pragmatic interpretation of whatever it modifies, be it a word, phrase, or an entire utterance (Jespersen 1942:418). Its purpose is to “guide the hearer in utterance interpretation and constrain the identification of the intended explicit and implicit meaning of an utterance” (Andersen 2001:41). The increased pragmatic meaning that *like* gained as it grammaticalized into a discourse marker communicates the speaker’s attitude to the utterance, which will be discussed in more detail below (Romaine & Lange 1991:266).

To complicate the characterization of this function of *like* even further, discourse markers themselves are a class diverse in form, function, and source, so they are difficult to define and

categorize completely, resulting in some inconsistency in the research (Brinton 1996:37). In fact, this diversity of use is a central characteristic of discourse markers: discourse markers are multifunctional; “they may be used in several different discourse functions simultaneously” (Romaine & Lange 1991:250). Stenstrom, et al., for example, find the discourse marker *like* used with an “approximative, exemplificatory, metalinguistic, [or] hesitational/linking” sense in their data (Stenstrom, et al 2002:117). Also, as Underhill (1988) argues, this function of *like* can be used as a focus marker, discussed below.

As a result of this multifunctionality, each researcher tends to define and inventory discourse markers according to his or her own research perspective (Brinton 1996:30). The view that discourse markers lack propositional meaning is one of the only areas of general agreement in the definition of discourse markers as a class (Andersen 2001:40). For example, Schiffrin (1986:42, 47, 62) explicates the “meaning-minimalist view,” wherein discourse markers are specifically identified as having no propositional meaning, although that does not imply they have no meaning at all: “[t]he minimalist view reduces the signaling load of the referential meaning of a particular form, and increases the role of pragmatic principles governing use of that form in context.” (47). In other words, there is less “value attributed to semantic meaning, as compared to pragmatic function, in an analysis of the total communicative force of an utterance” (Schiffrin 1986:47).

Schourup (1999:227) also writes that discourse markers are “thought not to affect the propositional content of utterances in which they occur,” a viewpoint that is uncontroversial with most researchers. However, Gisle Andersen’s findings (2001) suggest the opposite is the case. Andersen’s perspective is an outlier, contrasting with other researchers’ views. Andersen asserts that propositional meaning is communicated not only through semantic conduits but also through

pragmatic implicature (2001:44). This supports his claim that discourse markers, widely considered semantically void but pragmatically rich, do contribute to propositional meaning, since he asserts that pragmatic meaning plays an important role in utterance interpretation.

Admitting that some discourse markers do seem to lack propositional meaning (40), Andersen (2001) specifically proposes that discourse markers as a class may not be “necessarily external to propositions” (38). Andersen proposes a continuum of uses of discourse markers that do or do not have propositional meaning based on their stage of grammaticalization (2001:40-41). For example, discourse markers that “originate in expressions with inherent conceptual meanings and which have not been fully grammaticalized” are more prone to have a more complex effect on the propositional meaning of an utterance (Andersen 2001:55). Thus, a discourse marker such as *like* has some propositional meaning due to the grammaticalization principle of persistence (Andersen 2001:41), wherein grammaticalized forms tend to retain some of their “original lexical” meaning (Hopper 1991:22).

Because *like* grammaticalized from the prepositional function, a lexical entry which indicates approximation, the grammaticalized forms of *like* retain some of that vestigial semanticity, granting polysemy to the historically related but syntactically different homophones. Jespersen (1942) describes the pragmatic function of the discourse marker *like* as a humble indication that the speaker’s “choice of words was not, perhaps, quite felicitous” (418). Andersen (2001) also argues that *like* “instructs the hearer to opt for a less-than-literal interpretation of the utterance, and it may signal the speaker’s mildly dissociative attitude towards a chosen expression” (228).

Another recurring instance in which this function of *like* indicates that an inexact interpretation is most appropriate is when *like* precedes “numerical expressions” (Underhill 1988:240) or “quantity phrases” (Meehan 1991:40).

(15) Yknow, I was with his father for *like* nine years, since I was sixteen. (s04f)

(16) They *like* overdosed him *like* twice in a day. (s21f)

(17) I think they even might have found *like* one or two guns. (s30m)

This function of *like* is even found in a news article found on National Public Radio’s website, npr.org, to signify an imprecise reading of what is most likely hyperbole:

(18) Her song "Whip My Hair" hit urban and pop radio last month and now it's on *like* every ten minutes. (Kelley 2010)

In cases such as these, Underhill paraphrases *like* as ‘approximately’ (1988:240), and Meehan also notes that it can have the meaning of ‘similar to’ (1991:40). The function that the discourse marker *like* serves here is an indication that the number following *like* should not be taken as an exact account but rather a rough estimation (Andersen 2001:234). This is a model example of *like*’s indication of approximation of interpretation.

Another perspective on the function of the discourse marker *like* is put forth by Robert Underhill (1988). He writes that *like* “functions with great reliability as a marker of new information and focus” (234). This discourse function of *like* is said to mark new, significant information, usually the point of the sentence (Underhill 1988:235, 238).

Meehan agrees that in her research *like* does regularly precede new, significant information, thereby operating in a focusing capacity (1991:43). However, Meehan takes some issue with this classification: “*like* still has enough meaning that it cannot be considered simply a marker of information organization” (1991:43). Underhill’s (1988) analysis of *like* as a focus marker may therefore be oversimplified, ignoring the approximating pragmatic meanings inherent in the discourse marker function of *like* that developed as it grammaticalized from the

prepositional function. Meehan argues that the discourse marker *like* usually cannot be removed from the sentence without altering the interpretation of the sentence because of its pragmatic value that began historically as semantic meaning (1991:45).

However, unlike other uses of the discourse marker *like*, when *like* does precede new, significant information in the utterance, it can be removed without a change in semanticity (Meehan 1991:45). A central feature of discourse markers as a class is multifunctionality, and *like* is no exception to that rule. Hence, perhaps in addition to the pragmatically salient discourse marker function of *like* there is a separate discourse marker function which marks focus and is being reanalyzed and grammaticalized further. Meehan finally concludes that in “focus constructions, *like* is semantically bleached enough that it is really more of a grammatical item marking focus” (Meehan1991:45).

(19) I mean, I even had problems *like* on the school bus and stuff, yknow?

On the other hand, in addition to appearing in questions, answers to questions, and demarcating expressions, Underhill finds *like* to draw focus to “ideas that are not meant to be taken literally” (1988:241). Even this more semantically bleached discourse marker usage, therefore, can be used to distance the interpretation from the literal meaning of the utterance. In other words, it still shares the polysemous meaning of approximation with the other functions. Focus *like* and the more general discourse marker *like* are therefore both treated as discourse functions.

Traditional dictionaries present still another view of the function of this marker *like* in discourse: Webster’s determines *like* is “used without meaning or syntactic function, as in casual talk” (Neufeldt 1988:783). Though *like* may be more salient in casual discourse, it should not be assumed that the discourse function of *like* has no meaning at all, as Webster’s suggests.

Andersen (2001:21) reminds readers that discourse markers are not “void of semantic content”; this function of *like* has pragmatic rather than lexical meaning.

Similarly, the Collins COBUILD English language dictionary classifies *like* in this way: “[i]n spoken English, people sometimes say *like* when they are hesitating or when they are thinking about what to say next” (Sinclair 1987:842). Discourse markers such as *um*, for example, are sometimes used as a result of planning difficulties (Brinton 1996:33), either to retain the speaker’s turn while he or she searches for the desired word or to show that the speaker’s poor planning has forced the speaker to rephrase and choose words other than the exact ones intended.

The conception of *like* in particular being used as a hesitational device as a result of “planning difficulties” is a common view; *like* is sometimes used in this way to allow the speaker to pause for time to think, especially when found between clauses (Andersen 2000:18). However, Andersen argues against naming *like* a hesitational device devoid of meaning, reminding readers that “the occurrence of this marker in an utterance can crucially constrain its communicative import” (Andersen 2001:216).

Filler, the term for this type of hesitational device, is the traditional term for the discourse marker function of *like*. This name carries traditional prescriptivist stigma and implies that the words are inserted without any purpose or meaning (Brinton 1996:6). Gosy & Horvath (2010) use the term *filler* for the markers they analyze, which are similar to *like* in semantic meaning and functional diversity. Gosy & Horvath (2010: 137, 138) recognize that the particles they analyze have pragmatic meaning, yet they still refer to them as fillers, justifying their choice of terminology by emphasizing the occurrence of pragmatic markers as a result of speech planning

problems. However, Gosy & Horvath may choose the term *filler* simply because they narrowly define discourse markers as boundary markers that connect utterances (2010:137).

It is true that *like* can serve as a stalling device when the speaker is engaged in thinking; it is possible for *like* to be a filler or hesitational device, but because practically any word can allow a speaker to hesitate, one must avoid the temptation to leave the analysis of *like* there and ignore its deeper pragmatic meaning. Having retained some semantic meaning in its grammaticalization from the preposition function, *like*'s distribution "is not as random as that of hesitational devices like *eh* and *uhm*" (Andersen 2001:229). After all, "an overwhelming number of tokens of the pragmatic marker *like* occur where neither speed of production nor discourse coherence suggest that there are any planning difficulties involved," as Andersen (2001:227) points out.

Aside from disagreements over the meanings and exact functions of the discourse marker *like*, the history of this marker is also a point of confusion in the literature. According to the Dictionary of American Slang published in 1960, the origin of *like* as a discourse marker is in the culture of "jazz, cool, and beat groups, especially in New York City" (Wentworth & Flexner 1960:319). Then, by 1973, Mario Pei was describing this marker as "the intrusive '*like*' . . . which has spread to the point of appearing almost universally in the speech of younger-generation [speakers] who have no intellectual pretensions, and even some who do" (125). Since then, of course, *like* has spread throughout the English-speaking world, now appearing in corpora from Canada and the United Kingdom (Andersen 2001).

In contrast, other sources suggest this use of *like* is much older. Romaine & Lange suggest "discourse *like* has probably been used for over a century" (1991:270). Jespersen admits he has not found examples of use of this *like* "before Dickens" (Jespersen 1942:418).

(20) They say she was out of her mind *like* for six weeks. (Jespersen 1942:418)

However, as this admission implies, he does have examples from just after Dickens' era. Interestingly, this older use of discourse marker *like* modifies whatever it follows, similar to the suffixal *like*. In contemporary American English, discourse *like* precedes what it modifies.

The actual origins of discourse *like* are thus somewhat debated. Evidence such as Jespersen's instances of *like*, found in England years before it is supposed to have developed in the United States, makes determining exactly when and how this usage developed and spread more challenging. Andersen discusses the quandary that British English *like* predates the American marker:

[I]t is . . . clear that the word *like* has been used as a pragmatic marker especially in northern dialects of the British Isles for at least two centuries. It seems unlikely that the 'northern' usage should have expanded into London English, since the usual spread of linguistic innovations is in the opposite direction, that is, outward from large urban centres. This is not to say that the marker did not appear in Southern British English before the onset of this apparent "mega-trend" [from the U.S.]. But in a London context, the use of *like* as a multifunctional marker whose capacity includes the quotative and [discourse marker] functions is, to the best of my knowledge, first attested in . . . 1997. (Andersen 2001:221)

Therefore, *like* as a discourse marker may have developed independently in the two countries at different time periods, then did in fact spread from the United States back into England for a different social and pragmatic purpose than it originally developed there (Andersen 2001:221).

As Andersen references, one often-suggested place of origin for discourse *like* is the Valley girl culture in California (Andersen 2001:221). Perhaps because of this possible origin, discourse *like* is often associated with female speakers. On the other hand, the association of the discourse marker *like* with women may not be only a perception. Women have been found to use discourse markers more often than men because women are more likely to "express tentativeness or powerlessness" (Ragan 1983:166-167). Underhill points out how, even when it

is not possible to confidently interpret a particular usage of *like* “as an approximator,” *like* seems “to leave the statement slightly open,” so that it still has a hedging effect (Underhill 1988:241). This hedging effect coincides with the politeness associated with general female pragmatic tendencies (Romaine & Lange 1991:269).

Female speech, it is generally agreed, is at least where the use of *like* gets its start; however, there are some exceptions (Andersen 2000:287, see Dailey-O’Cain 2000). For the most part, *like* has been spread by females, both in the “American pattern” and in Andersen’s analysis of data from The Bergen corpus of London teenage language (Andersen 2000:287, 1997). Ferrara & Bell studied speakers in Texas and found that by 1992 *like* had begun to lose “its gender-marking,” occurring universally in young male and female speech (1995:271). This gender leveling is discussed in more detail with respect to the quotative *like* below.

The use of *like* as a hedge which Underhill describes above is associated with female speakers, at least in part, because it communicates an air of politeness to the listener (Underhill 1988:241). Similarly, Jespersen claims *like* is commonly “used by inferiors addressing superiors,” but does not elaborate as to who are these supposed “inferiors” and “superiors” (Jespersen 1942:418). In traditional gender roles, certainly those in place in 1942, Jespersen could be referring to women hedging in addressing men.

Socioeconomic class is also a plausible interpretation of Jespersen’s inferior/superior distinction. Class is not typically accounted for in contemporary studies of *like* as a discourse marker. However, Romaine & Lange (1991:269) suggest changes in progress, such as the grammaticalization of *like*, are usually advanced by lower social classes, therefore so-called social “inferiors,” which may support Jespersen’s claim (Jespersen 1942:418).

This suggestion that *like* is used to show deference also relates to the polysemous meaning of approximation “as it provides speakers with a tool for not sounding too assertive but expressing themselves with a tentative attitude” since whatever they say may not be exactly what they mean (Andersen 2001:229). This polysemy connecting the central functions of *like* is evidence that these homophones are historically related. Just as the discourse marker function of *like* developed from the grammatical functions, so did the final usage addressed in this study: the quotative marker function.

b. Quotative marker

The quotative function of *like* is a nonstandard quotative which entered the lexicon of younger American speakers in the 1980’s and has since spread geographically, if not generationally (Tagliamonte & Hudson 1999:148; Romaine & Lange 1991:236).

(21) If he didn't do it he might as well just be *like*, “No, I didn't do it.” (s30m)

(22) When you're a teenager so you're *like*, “So what?” (s04f)

This quotative marker is unique within the English quotative system. *Like* can introduce a more diverse range of quotations than the other English quotatives: “[m]ost quotatives are associated with either direct speech or inner monologue but rarely with both” (Blyth et al. 1990:215). For example, the unmarked choice *say* and the nonstandard more recent innovation *go* can only be used to demarcate utterances that were spoken aloud, whether direct or indirect speech (Blyth et al. 1990:215).

(23) When she said that, I *said*, “Well, is that in California?” ’cause I wasn’t sure if it was in California.

And she goes, “Yes.” (Blyth et al. 1990:215)

(24) He *said* you'd have people with like knives and you'd have people with fist fights. (s30m)

On the other end of the traditional American English quotative spectrum, *think* can also introduce “inner monologue” that is not actually uttered aloud (Blyth et al. 1990:215). In

contrast with these other quotatives, however, *be like* can be used not only for one kind of utterance. In fact, *be like* is used for both of the contexts that *say*, *go*, and *think* cover: speech and thought. Additionally, *be like* can also introduce nonlexical vocalizations, explained below.

The quotative marker *be like*, together with *say* and *go*, comprise “the three most frequent” quotatives in American English, each capable of introducing audible speech (Blyth et al. 1990:215). It is interesting to examine whether the newest innovation, *be like*, introduces only direct speech, as *go* does, or both indirect and direct speech as *say* does. Romaine & Lange, on one hand, suggest *be like* is a “new kind of construction which represents a compromise between the direct and indirect mode” (1991:258). However Blyth et al. (1990) point out that, as is the case with *go*, introducing indirect speech with *be like* would actually be ungrammatical: “He’s *like/goes*, ‘I saw her!’ versus *He’s *like/goes* that he saw her” (222). Likewise, Meehan (1991:48) also characterizes *be like* as direct quotation.

Although *be like* cannot introduce indirect speech as the indirect mode is defined for other, traditional quotatives, *be like*’s distinctive application does to some extent behave as if presenting indirect speech. Romaine & Lange (1991) argue that the pragmatic implication when *be like* is the choice for quotation is “that the speaker stands in a relation of reduced responsibility and commitment to the truth of the report,” as is the case with indirect quotations (263). Similar to indirect quotations, *be like* does not vouch for the exactness of the quote. In other words, though constructions with *be like* appear on the surface to be identical to direct quotations, they are at least pragmatically similar to indirect quotations. *Be like* conveys the narrative “vividness” and syntactic simplicity of direct speech without the pragmatic implication that the quotation is verbatim (Romaine & Lange 1991:264). Romaine & Lange (1991:264) “regard the grammaticalization of *like* as a quotative complementizer to be a natural historical

development for the spoken channel” because of the usefulness of this intermediate quotative option that stands somewhere between direct and indirect mode.

In addition to thought and audible quotations (direct and indirect mode aside), *like* can also introduce nonlexical vocalizations (Romaine & Lange 1991:241), as this study’s corpus demonstrates:

- (25) My boyfriend would love to mess with this. I just want you to know that he’d be *like* ‘Eeee,’ takin it apart, puttin it back together. (s21f)
 (26) So I’m *like* <INHALE>. (s04f)

Blyth et al. borrowed “the category *interjection*” to classify these exclamations, which they define as utterances not included in “the conventional American English lexicon,” that is, utterances that are not words (1990:217).

The common link between all these functions for the quotative *like* is that *be like*, in all these contexts, is communicating essentially the general feeling or emotion of the party to which it is attributed. Tannen (1986) posits that the original meaning of *like* influences the interpretation of the quotative in that “the dialogue is not being quoted but simply represents the *kind of thing* that character was saying or thinking” (321, emphasis added). As with the discourse marker function of *like*, the quotative marker developed from the original lexical entries, the preposition and conjunction functions, and also retains the vestigial meaning of approximation: “[s]emantically, it is because *like* has the referential meanings of ‘comparison,’ ‘for example,’ ‘as if,’ and so on, that it is suitable for use in a construction reporting hypothetical discourse or thought” (Romaine & Lange 1991: 245, 259). The grammaticalization feature of persistence is also evident in the quotative marker *like*, as it retains this approximation of meaning from the original lexical entries (Hopper 1991:22).

Rather than being used for explicit speech or expressions, quotative *like* conveys merely the general feeling of the character whose contextual reaction is being captured, be it overtly stated or not (Romaine & Lange 1991:241). Since quotative *like* makes no commitment to the reported speech being verbatim or even having been out loud, this pragmatically allows a speaker to be less accountable for the quotation being verbatim, thus allowing the speaker to save face in case he or she has misrepresented the original meaning (Romaine & Lange 1991:241, 238, 243). Quotative *like* is at least not yet semantically bleached; its implication of approximation of meaning allows it to be applied diversely and defined loosely (Romaine & Lange 1991:246). The grammaticalization from the original lexical entry lends this pragmatic meaning to quotative *like* to allow it to fill in this gap in the English lexicon.

Since *be like* can introduce several different types of quotation, it can be difficult to determine which kind of representation is intended at the time. Returning to the example given above from Blyth et al. (1990:215), when the speaker uses *be like* to introduce “Oh,” it is not clear whether this is uttered aloud or simply represents inner thought:

(27) When she said that, I *said*, “Well, is that in California?” ’cause I wasn’t sure if it was in California. And she goes, “Yes.” And I’m *like*, “Oh.” (Blyth et al. 1990:215)

Additionally, *like* sometimes has an ambiguous scope so that it is also difficult, if not impossible, to determine where the quotation ends and the speaker’s narration resumes (Romaine & Lange 1991:248).

There is another question regarding the syntax of *be like*, specifically, how to refer to the role of *be like* in the syntax of a sentence. Romaine & Lange stipulate that “where *like* occurs with *be*, it has not become a verb of saying but retains its function as complementizer” (1991:248). It is for this reason that Romaine & Lange refer to *be like* as a “quotative

complementizer” rather than a quotative (1991:248). However, *like* can function as a quotative without *be* (Stenstrom et al. 2002:116-117):

(28) I would just call my mom *like*, “Hey what can I do?” and she just *like*, “Calm down Rachel; just calm down.” (s04f)

(29) I mean, yknow, I have little thoughts of *like*, “I like to design.” (s04f)

(30) Makes you think *like*, “Right now I'm cold.” (s21f)

Stenstrom et al. find in the COLT corpus instances of: *like* alone, go *like*, say *like*, other verb + *like*, and it's *like*, which is classified separately from *be like* since *be like* follows a personal subject pronoun, which is replaced by *it* in this case (2002:117). For the purposes of this study, *be like* is referred to with the standard term *quotative marker*.

The treatment of *be like* by dictionaries and popular perceptions of *be like* are both interesting alternative perspectives. The discourse marker usage of *like* may be treated with some dismissal in Webster's, but the quotative marker function is not treated at all. Romaine & Lange (1991) found the same conspicuous absence in *The American Heritage Dictionary* (1969:757) and speculate about the reason: at the time the dictionary was compiled, “this usage was either nonexistent, too infrequent to have attracted notice, and/or so colloquial (and thus, unacceptable) that it was simply ignored” (245). Regardless of the exclusion from some traditional dictionaries, the quotative function has been analyzed more than any other usage of *like* in linguistic research (Andersen 2001:217).

In the early 1990's, Blyth et al. (1990:224) found a stigma attached to the use of this quotative marker in an attitudinal study they conducted as part of their research. They found *be like* to be stereotypically associated with “middle-class teenage girls” who were assumed to be “‘vacuous,’ ‘silly,’ ‘air-headed,’” and, geographically speaking, associated with California. Romaine & Lange (1991) likewise report that *be like* is “judged by teenagers of both sexes as being characteristic of female speech” (Lange 1986, cited in Romaine & Lange 1991:255).

Perceptions of *be like* may or may not correspond to these actual sociolinguistic distributions. Both sexes produced *be like* in the Lange (1986) study (Romaine & Lange 1991:255). Likewise, Blyth et al. (1990) found empirical evidence that directly contradicted their attitudinal study: in their data, males were more likely (0.534) than females (0.466) to use *be like* with a p-value of less than 0.006 (1990:221).

However, not all studies on *be like* found males to be more prevalent users of this variable. Romaine & Lange (1991:251) found 83% of their examples of *be like* in the speech of female speakers rather than male. Romaine & Lange also point out that perhaps *like* is simply “more salient as a stereotypical marker of female speech style though its use to discuss topics which women favor more than men” since women are more likely to discuss interpersonal relationships and employ dialogue as a narrative device (Romaine & Lange 1991:256). In addition, the approximating value of *be like* that it retains as polysemy with other functions of *like* could make it a more appealing variable choice for women since “it is consistent with many of the goals and values associated with the female communicative style” in its ability to communicate emotion without committing to a precise account of another person’s state of mind or exact utterance (Romaine & Lange 1991:269).

Aside from gender, the above perceptions of *be like* also make predictions about the age and socioeconomic class of the speakers who are more likely to use this quotative. In the case of age, the perceptions that teenagers, or at least younger speakers, are more likely to use *be like* may be in line with reality. Romaine & Lange (1991) point out that this quotative “tends to occur in the speech of those under 30” (236). This assessment is confirmed by results from Romaine & Lange’s (1991) experiment as well as Blyth et al. (1990). Romaine & Lange (1991) write “[a]lmost all of our examples come from young people” (251). Blyth et al. found similar

age gradation in their results: past 25-years-old, use of *be like* “dropped off sharply” (1990:219). For Blyth et al., quotative *like* is “often used by younger speakers but never (in our corpus) by speakers older than 38” (1990:219). Andersen (2001) also finds *be like* is used very little by speakers older than 45 (225).

More particularly, Blyth et al. found quotative *like* to be associated with teenage speech. This specific association also seems to be a fair assessment of reality. Blyth et al. write, “teenage and college-age speakers use *be like* and *go* much more frequently than speakers from the older age groups” (1990:219). Andersen (2001) also confirms that quotative *like* is common in teenage discourse (224).

Some of these sociolinguistic observations about *be like* and the discourse function of *like* will be explored below. The crucial observation about *be like* is its ability to introduce speech tentatively, the same way the discourse marker usage can communicate a tenuous connection between the literal utterance and the general thought expressed. This polysemous connection that these more grammaticalized functions of *like* share with the preposition and conjunction usages is only one indication of the grammaticalization that allowed these marker functions to develop.

2.2 GRAMMATICALIZATION OF *LIKE*

Grammaticalization is a process of language change in which a lexical morpheme, or a grammatical one such as a preposition or conjunction, begins to be used in a more grammatical sense, such as the discourse and quotative functions of *like* (Kurylowicz 1965:52). These marker functions developed from the functional uses of *like* through grammaticalization. Because grammaticalization is such a complex, long-term process, it is sometimes difficult to identify indisputably, and there is sometimes controversy as to whether a particular language change is

grammaticalization or some other diachronic linguistic process. Thus, Hopper (1991) suggests a rubric of five principles of grammaticalization meant to help identify forms in the process of grammaticalization (21). Most of these principles, all of which are described below, can be applied to *like*.

a. Layering

“Within a broad functional domain, new layers are continually emerging. As this happens, the older layers are not necessarily discarded, but may remain to coexist with and interact with the newer layers” (Hopper 1991:22). Layering can be seen in the development of *like* in that *like* has several different functions at one synchronic point in the language, described in detail above.

b. Divergence

“When a lexical form undergoes grammaticalization to a clitic or affix, the original lexical form may remain as an autonomous element and undergo the same changes as ordinary lexical items” (Hopper 1991:22). Though *like* has not grammaticalized to the point of becoming “a clitic or affix,” this study hypothesizes that the different functions of *like* are autonomous in how their frequency affects their phonetic duration (Hopper 1991:22).

c. Specialization

“Within a functional domain, at one stage a variety of forms with different semantic nuances may be possible; as grammaticalization takes place, this variety of formal choices narrows and the smaller number of forms selected assume more general grammatical meanings” (Hopper 1991:22). Hopper & Traugott (1993:114) point out that specialization “may be manifested simply as textual preferences, conditioned by semantic types, sociolinguistic contexts, discourse genres, and other factors” which is certainly true of the marker functions of

like since their sociolinguistic use is not universal (i.e., it is age-graded) and context formality is equally limited (i.e., used in informal discourse only). Also, one result of specialization is the grammaticalized form becoming more frequent, which is true of the marker functions of *like* (Brinton 1996:51).

d. Persistence

“When a form undergoes grammaticalization from a lexical to a grammatical function, so long as it is grammatically viable some traces of its original lexical history may be reflected in constraints on its grammatical distribution” (Hopper 1991:22). Persistence is a salient feature of the grammaticalization of *like*. Both the discourse marker and quotative marker functions retain enough semanticity from their original lexical entry to indicate approximation of meaning, not to be synonymous with *um* or *say*, respectively, and to constrain the pragmatic meaning of an utterance.

All of the above types of *like* do share a common thread of meaning. This polysemy comes in two forms: comparison and approximation. Romaine & Lange (1991:246) posit that these sundry functions tend to imply “a comparison, either actual or hypothetical,” between two subjects. Andersen describes the relation between the various usages of *like* as a “non-literal resemblance between an utterance and the underlying thought” (2001:210). In addition to comparison, the grammaticalized marker usages of *like* share a polysemous relationship with the preposition and conjunction functions that the meaning expressed is merely an approximation of the speaker’s intended meaning because the marker functions of *like* originally evolved from these lexical items which characterize analogous but distinct states (Andersen 2000).

Though the central meaning of the prepositional usage is of comparison, this is still not distinct from the other meanings that the functions of *like* share. The preposition function of *like*

may be defined “of *approximately* identical shape, size, colour, character, etc., with something else” (OED online, emphasis added). When a speaker makes a comparison using *like*, after all, he or she is not asserting that the two things being compared are exactly alike, but are instead similar. The marker functions are related to this meaning of ‘similar to’ as is exemplified in their “lack of exactness” or commitment (Meehan 1991:50).

e. De-categorialization

“Forms undergoing grammaticalization tend to lose or neutralize the morphological markers and syntactic privileges characteristic of the full categories Noun and Verb, and to assume attributes characteristic of secondary categories such as Adjective, Particle, Preposition, etc” (Hopper 1991:22). *Like* in its marker functions cannot take morphological marking and, as a discourse marker, is indeed syntactically free, a feature more characteristic of functional words than content words.

As to the specific development of the marker functions of *like*, Meehan suggests that the grammaticalized functions probably developed from the preposition ‘similar to’ since that was the first meaning (Meehan 1991:50). Table 2.1 shows Meehan’s chronological representation of the development of *like* in what she considers the salient functions.

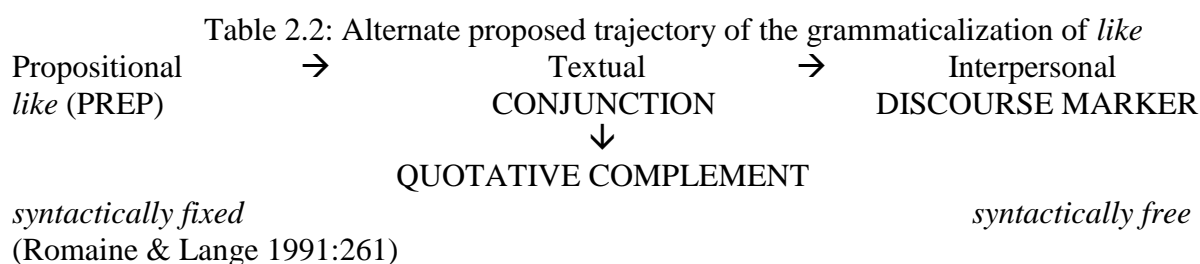
Table 2.1: Proposed trajectory of the grammaticalization of *like*

| 1400 | 1800 | 1950+ | |
|--|---------------|--------------------------------------|---------------------|
| More Lexical | | | More Grammatical |
| -----> | | | |
| ‘Similar to’ ‘Approximately’ ‘As if’ | | Discourse Marker Quotative Marker | |
| | ‘For example’ | | |

(adapted from Meehan 1991:49, 51)

On the other hand, Romaine & Lange (1991) agree that the preposition is the original function of *like* but argue that the marker functions probably actually developed from the conjunction usage

(261). Table 2.2 represents Romaine & Lange's (1991) theory for the development of the marker functions.



This theory is more complex than the simple chronological view put forth by Meehan (1991). Romaine & Lange (1991) acknowledge that the marker functions of *like* developed from either the conjunction or preposition usages, but specifically identify this process as “specialization within the textual component,” under which they include the conjunction usage. Both Meehan (1991) and Romaine & Lange (1991) propose valuable representations of the history of *like* and its multiple functions and meanings, but Meehan's (1991) is more helpful in the chronological development whereas Romaine & Lange (1991) more convincingly represent the actual development of *like*.

Jespersen (1942) proposes a model for the development of *like* from lexical entry to grammaticalized marker functions that relies on a function of *like* not found in this study's corpus: the suffix. The suffix form of *like* is cognate with the suffix *-ly* which is a phonologically condensed form of *liche* meaning ‘body;’ this is another example of grammaticalization from a lexical item, in this case including complete semantic bleaching (Romaine & Lange 1991:245). In fact, this case of grammaticalization is a prime example of grammaticalization in an advanced stage: affixation (Heine, Claudi, & Hünemeyer 1991:12).

Romaine & Lange (1991) explain the importance of the suffix form in the development of the marker functions: the preposition and suffix functions together made grammaticalization

syntactically possible by providing a position “from which *like* can be extended to focus on a whole clause, sentence, or chunk of discourse” (Romaine & Lange 1991:261). In other words, when “*like* can appear as a suffix following an item . . . as well as precede a clause or sentence, it can be reanalyzed as a discourse marker, which shows syntactic detachability and positional mobility” (Romaine & Lange 1991:261). Syntactic freedom is an important element in the development of *like* as a discourse marker, and it was enabled in part by this now relatively rare suffix function.

In addition to the suffix and prepositional uses of *like*, the conjunction usage has contributed to the development of the marker functions. Specifically, the quotative marker can easily develop from the conjunction’s presentation of a clause “for comparison or exemplification so that it can be construed as a report of speech or thought” (Romaine & Lange 1991:262). In addition to the preposition and conjunction functions of *like* and their entailed approximation and comparison, the discourse use of *like* as a marker of focus has aided in the construction of both marker functions (Romaine & Lange 1991:245).

With regard to the marker functions individually, the quotative function is generally considered to predate the discourse function. Since the quotative marker function of *like* has a more fixed syntax than the discourse marker, the quotative seems to be more advanced in the process of grammaticalization (Meehan 1991:46). Andersen (2001) posits that “the quotative construction is a relatively late stage in the grammaticalization process and . . . the occurrence of other less syntactically fixed uses precede the quotative use” (Andersen 2001:218). When a grammaticalized morpheme is syntactically fixed, “it is more likely to become an affix as it passes through the stages of grammaticalization” (Meehan 1991:47). In other words, the

quotative marker is more likely to advance to the next stage of grammaticalization, if it ever comes to that.

2.3 DURATION AND ATTRITION

Forms that are grammaticalized, as the above description establishes *like* to have been, frequently undergo “[p]honological ‘attrition’ or ‘erosion,’” wherein the morpheme loses “phonological substance, from full to reduced form and eventually to zero” (Brinton 1996:52). Andersen writes that most cases of grammaticalization of discourse markers are “accompanied by a reduction of phonological salience” (Andersen 2001:55).

Duration in this case is the amount of time in milliseconds a speaker takes to articulate a word. According to Gahl, words become shorter with “articulatory routinization,” wherein the brain and the articulators of the vocal tract become so used to the same pattern of configurations that they speed up the articulation (2008:474). Bybee explains that “the general tendencies affecting speech production are the same as those affecting other instances of repeated motor behavior. With repetition, neuromotor routines become more compressed and more reduced” (Bybee 2001:78). In other words, more frequent words are often shorter in duration. The shortening process is partly brought about by the frequency of these words, and since these marker functions are even more frequent, this study hypothesizes that they will be shorter in duration (Gahl 2008:480).

Gahl also points out that, in her data from the Switchboard corpus of American English telephone conversations, “[f]unction words are approximately 60 milliseconds shorter on average than their content-word twins (210 ms vs. 273 ms),” by which Gahl means homophones such as *in* and *inn* (2008:480, 481). Gahl found, first, that “the lower-frequency words were longer than their high-frequency counterparts,” which is consistent with Bybee’s (2001) discussion of

articulatory routinization (2008:481). Second, and more central to her research question, Gahl found that “high-frequency words like *time* are significantly shorter than their low-frequency homophones like *thyme*,” though depending on one’s model for language, the more frequent homophone could have exerted a phonological influence on the less frequent one (2008:474). This study likewise expects to find the homophones of *like* to have differing durations.

Gosy & Horvath (2010) also examine differing durations of homophones, in this case homophones with different functions, the Hungarian *ilyen* ‘like this’ and *tehát* ‘that is’ and ‘consequently’ (135). *Ilyen* and *tehát* are currently “undergoing a functional change, acquiring the function of fillers, while retaining their former lexical function,” which make these words very similar in type to *like*, though in an earlier stage of the change (Gosy & Horvath 2010:135). Gosy & Horvath found that both of these pairs of homophones “showed significant differences in duration depending on function” (2010:135). The same result is expected from *like* in its various functions.

Specifically, Gosy & Horvath (2010) examine *tehát* for phonetic attrition in the loss of the word-medial consonant and a syllable (147). Consistent with the assertion that frequent words often lose speech sounds (2010:147), the filler function of *tehát*, the most frequently used function in Gosy & Horvath’s data (146), is also the least likely to retain both syllables (148). Gosy & Horvath do not find a shortening effect for *ilyen*, but do observe vowel centralization in the filler function of *ilyen* (147). Given the shortening of *tehát* and vowel centralization in *ilyen*, Gosy & Horvath conclude that the speakers in their data “reduce articulatory efforts when the words do not have lexical status,” such as grammaticalized fillers or markers (2010:152).

Gosy & Horvath (2010) also measure durations of the functions they examine. In the case of *tehát*, its articulation “in the filler function was completed in a shorter time than the

articulation” of the other functions (149). By contrast, the filler function is longer in duration than the one other function, though the filler function also displays greater variation in the duration lengths (Gosy & Horvath 2010:154). According to Gosy & Horvath, this “durational data support our assumption that a synchronic change is taking place in the function of *ilyen*, and this functional change is inducing a change in the temporal patterns associated with the word” (2010:155). This result is the same as this study is searching for in the case of the English *like*.

CHAPTER 3: DESCRIPTION OF THE RESEARCH AND PROCEDURE

The current study examines the durations of instances of *like* from a grammaticalization perspective, using data from the Buckeye corpus of conversational speech. Four speakers were selected from the subcategories of speakers in the Buckeye corpus, and their durations of *like* in all five functions were analyzed and compared with respect to function, speaker, and sex of the speaker (Pitt 2007).

3.1 CORPUS

The Buckeye corpus of conversational speech was collected in 1999 “to create a database of approximately 300,000 words of conversational speech by native central Ohio speakers” (Kiesling et al. 2006:2, 3). The speakers represent “natives of Central Ohio (i.e., born in or near Columbus, or moved there no later than age 10)” and are limited to “middle-class Caucasians” (Kiesling et al. 2006:3). Of the 40 interviews included in the corpus, 20 are from females and 20 from males (Kiesling et al. 2006:2). Of these males and females, 20 were interviewed by a graduate student interviewer of the same sex and 20 by an interviewer of the opposite sex (Kiesling et al. 2006:2). The informants were likewise segmented into two groups by age: one group of 20 “older” speakers over the age of 40 and another group of the same size of “younger” speakers under the age of 30 (Kiesling et al. 2006:2, 3).² The metadata is not more specific in participants’ ages than these groupings, limiting the accuracy of age-related observations.

² See Appendix II.

3.2 SELECTED SPEAKERS

For the current study, four speakers were selected from the Buckeye Corpus. The speakers examined in this study are 2 males and 2 females from the younger group of informants. Speakers were selected at random to fill one speaker per cell for informant and interviewer sex, as represented in Table 3.1.

Table 3.1: Distribution of selected speakers

| | Male Interviewer | Female Interviewer |
|------------------|------------------|--------------------|
| Male Informant | Speaker 30m | Speaker 40m |
| Female Informant | Speaker 21f | Speaker 04f |

However, interviewer sex was not accounted for in the analysis. Because of this small sample size, speakers were only chosen from the younger group in hopes of being able to make generalizations across a smaller community rather than attempt to account for a larger community and lose the ability to generalize (Kiesling 2006:3).

For the first file (speaker 40m), the procedure differed from the rest of the speakers' files, but this inconsistency was rectified later. This first file was analyzed only by listening in Praat to the Buckeye corpus audio files, tagging, and marking the boundaries of each instance of *like* (Boersma & Weenink 2004). This method proved imperfect, however, since only 172 out of 192 tokens were identified on the first try.

From each interview, an Antconc concordance of all the instances of *like* was compiled in order to count the tokens of *like* and show them in context (Anthony 2010). This concordance was therefore used to identify and locate the missed instances of *like* in this first interview, and thereafter the procedure was modified to control for the human-error prone method of listening for *like* in the interviews.

For the rest of the files, the Antconc concordance was opened alongside the Praat audio file and each instance in the concordance was located in chronological order within the audio. The Buckeye corpus provides several separate zipped files for the audio and transcription of each speaker, the number of files depending on the length of the interview. The transcript files were unzipped and combined, and the combined file for each speaker was used to create an individual Antconc wordlist for each participant. Interestingly, for each of the four speakers analyzed, *like* was among the 10 most frequent words used in the interview. From these Antconc wordlists, a concordance for *like* was compiled for each speaker, organized chronologically within the interview.

Using these concordances to locate each instance of *like*, the unzipped audio files were opened individually (rather than combined as the transcript files were) in Praat. Within the Praat file, once a token of *like* was located according to the corresponding Antconc concordance, its phonetic boundaries were determined in order to extract the duration.

3.3 DURATION DETERMINATION AND EXTRACTION

In phonetically transcribing the Buckeye Corpus, “[h]uman labelers used speech spectrogram and waveform displays . . . as well as auditory perceptual evaluation in determining phone labels” (Kiesling et al. 2006:10). Likewise, in this study, in order to determine the boundaries of the token word, both the waveform and spectrogram were consulted and the audio was listened to. In cases of more difficult discrimination, at least two out of three indicators were required to be in agreement for the boundary to be marked. “In some cases (depending on the phonetic context), other cues, such as formant discontinuity or differences in intensity, were also used as segmentation cues” for Gosy & Horvath’s study and in this study (Gosy & Horvath 2010:139).

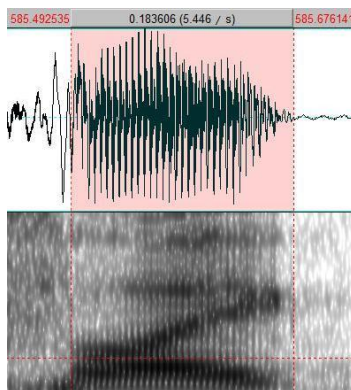


Figure 3.1: An example of *like* and its boundaries at the left and right edges.

Duration was captured by locating the left and right edges of each occurrence of *like*. On the left edge, the entire phonetic realization of the [l] in *like* was captured by locating the formants that characterize [l], 250, 1200, and 2400 Hz, or a close approximation of those formants where possible, again listening and reading both the waveform and spectrogram (Ladefoged 2006:195). For optimum consistency across all tokens, the right edge of *like* was captured at the closure for [k], since in some cases the [k] is released and in others it is not.

When all tokens of *like* were located and boundaries demarcated, durations could be extracted in Praat. First, all labeled intervals, i.e. demarcations of instances of *like*, were extracted in the Praat Objects window. Then each extracted interval was queried for total duration in seconds, the value of which was copied into an Excel file, retaining the file name and function of the token so that each instance and all functions could be compared and evaluated.

3.4 DETERMINATION OF FUNCTION

The functions found in the data used in this study are verb (V), preposition (P), conjunction (C), discourse marker (DM), and quotative marker (QM). Function was determined for each individual token by a careful and consistent paraphrasing process. The verb function can be characterized by its syntax, occurring at or near the beginning of a verb phrase, so that it

follows a nominal noun phrase as in example (30), auxiliary or modal verb as in example (31), or *do* plus negation as in example (32).

- (31) I *like* to do office work. (s04f)
 (32) I'd *like* to think I don't have trouble making friends. (s30m)
 (33) I don't *like* . . . like, I *like* business. (s30m)

The verbal *like* found in this data is transitive. This function of *like* is identified by the replacement test: if another verb can replace *like* in the sentence, the token is a verb (Adger 2003:63).

The conjunction usage of *like* is identified by paraphrase. The definition of the conjunction usage given in chapter 2 states that this type of *like* is synonymous with *as* and *as if* (Neufeldt 1988:783). Thus, if a particular instance of *like* can be paraphrased as *as* or *as if* while still preserving the basic meaning of the utterance, the token is labeled as a conjunction.

- (34) My mom's kind of weird *like* that. (s40m)
 (35) I worked with, um, adults that act *like* we were in high school. (s04f)

It was by this paraphrase that *like I said*, *feel like*, and *seem like* are all also considered uses of the conjunction function. *I feel like* may be paraphrased as *I feel as if*:

- (36) *I feel like* he's a grown man. (s04f)

Seems like may be paraphrased as *seems as if*.

- (37) It *seems like* there's a lot more to do in Columbus, obviously. (s30m)

Like I said may be paraphrased as *I said*:

- (38) It's just having it forced on me . . . *like I said* just sort of gives me a different opinion. (s40m)

Like I said as a collocation is addressed again in chapter 5.

Prepositional *like* is identified by paraphrasing the token as “similar to; somewhat resembling” (Neufeldt 1988:783).

(39) I mean I . . . I don't have [morning sickness] *like* my sister: she's puking every day. (s21f)

(40) I did have a job *like* that when I worked full time. (s04f)

Alternatively, usages preceding a list or specification can be paraphrased as “for example” or “such as” (Meehan 1991:42):

(41) Marriage and kids and stuff *like* that with them. Um *like* they're living together, things *like* that. (s40m)

As with the collocations mentioned above for the conjunction usage, the phrases *something like that*, *things like that*, and *stuff like that* are also categorized as prepositional usages.

(42) I don't know how I would get involved in *something like that*, but I would like to. (s30m)

(43) I don't have a set schedule of when I have to work, *things like that*. (s40m)

(44) I've seen a few of like infomercials on *stuff like that*. (s40m)

In each of these instances, the collocation may be paraphrased by replacing *like* with “similar to; somewhat resembling” and without a change in meaning (Neufeldt 1988:783).

The discourse marker function of *like* is identified by several factors. First, discourse markers can be extricated from the sentence in which they occur without affecting the propositional content of the statement (Brinton 1996:35). Thus, if *like* can be removed from the sentence without affecting the semantic meaning of the statement, the token is labeled a discourse marker.

(45) It went on for *like* two months where he just cried. (s04f)

(46) It would take *like* years I would think to implement the programs. (s30m)

Gosy & Horvath (2010) likewise “regard an item as a filler when there is no grammatical or semantic reason for its presence” (137).

Substitution can serve as another diagnostic tool for identifying discourse *like*: if another discourse marker such as *um* can replace *like*, its function in the sentence is pragmatic:

(47) Yeah . . . and that . . . *like* . . . I don't know. (s30m)

Collocation with other discourse markers such as *I mean*, *you know*, or *um* also suggests that the token is a discourse function:

- (48) I mean other than *like* I mean obviously like guns and things like that. (s40m)
 (49) Carryable and concealable I mean . . . *like* you know back in the sixties stuff like that I mean . . . *like* you know back in the sixties stuff like that I mean they had switchblades things like that. (s40m)
 (50) Y'know . . . um . . . so *like* that's why we don't do nothing. (s04f)

Collocation with planning problems paired with prosodic cues also indicates the token is a discourse marker:

- (51) I don't --*like* the problem, I think, with universal healthcare would be everyone --don't you think *like* everyone with the littlest scratch would go in? (s30m)

Discourse *like* can also be identified, as mentioned above, by its lack of lexical or propositional content. If a token cannot be identified as another function of *like*, it is examined as a possible discourse marker.

The quotative marker *be like* can be paraphrased as *say*, *think*, or *feel* wherever it precedes something that can be interpreted as a hypothetical or real statement uttered by a speaker or a general emotion or opinion felt by someone or attributed to him or her.

- (52) One of my friends was *like*, "What'd you say?" (s40m)

In example sentence (52), 'was *like*' can be paraphrased as *said* (s40m). Sentence (53) is the slightly more complex "it's *like*," rather than simply *be like*, but it can still be interpreted to mean 'I thought' or 'I felt.'

- (53) All of a sudden it's *like*, "I want to eat that and that and that." (s21f)

Finally, the italicized token in example (54) clearly precedes an entity's general attitude about the subject being discussed, in this case the death penalty.

- (54) That's, *like*, such a disturbing fact to me that [the state] could take someone's life and be *like* "Oh our bad." (s30m)

3.5 INDETERMINACY OF FUNCTION

In some cases, the function of a particular instance of *like* is indeterminate (Romaine & Lange 1991:253). Function is sometimes difficult to determine because grammaticalization is ongoing, so various functions of *like* in various stages of development co-exist synchronically (Romaine & Lange 1991: 258). Likewise, Meehan argues that these difficult-to-analyze instances suggest “the meaning is becoming more generalized and . . . usage is becoming more grammatical than lexical” (1991:44). Also, because of the in-progress nature of grammaticalization, the demarcation between more and less grammaticalized usages is sometimes tenuous (Andersen 2001:54).

Sometimes “the indeterminacy [of function] is due to *like* appearing in a fractured sentence which may be the result of planning difficulties” (Andersen 2001:212). However, “prosodic features crucially constrain the interpretation of *like* as a marker or a non-marker” so that “a brief pause immediately following *like* . . . fairly strongly suggests that *like* is a pragmatic marker” (Andersen 2001:213). This use of *like* as a hesitational device, especially preceding a pause to retain one’s turn while the speaker searches for the desired expression, may account for the long mean duration of the discourse *like*. Perhaps because *like* is meant to buy the speaker time to arrange his or her thoughts, a longer duration would be valuable, negating the effects of phonological reduction from grammaticalization.

With regard to the indeterminacy of functions of *like*, prepositional *like* is also used in this study’s sample to mean “for example” (Neufeldt 1988:783). Usages preceding a list or specification can be paraphrased as ‘for example’ or ‘such as’ (Meehan 1991:42) and were therefore tagged as prepositions unless intonation, false starts, pauses, or collocation with other discourse markers suggested *like* functioned as a discourse marker. In most cases, as Andersen

writes, “I listened repeatedly to the examples during the analysis, and many of the examples could then be disambiguated” (Andersen 2001:213). Most cases could be effectively identified if listened to.

However, the following examples were still ambiguous in function:

(55) Marriage and kids and stuff like that with them. Um *like* they're living together, things like that. (s40m)

(56) The business school is coming up, *like* they're building a new building. (s30m)

(57) She more believes in the religion aspects of it, not so much as what the Catholics believe about it, um *like* their views on outside of religious matters. (s40m)

(58) I mean other than-like I mean obviously *like* guns and things like that are being around. (s40m)

These instances could be interpreted as prepositions or discourse markers. Each of these functions was tagged as a prepositional usage because, as discussed in section 3.4 above, the token can be paraphrased as *for example*.

CHAPTER 4: RESULTS

4.1 SELECTION OF SPEAKERS

Only speakers from the Buckeye corpus subcategory of younger speakers were chosen for this study. This selection of younger speakers stems from the assumption that the marker functions of *like* tend to be used by younger speakers. Table 4.1 shows the raw numbers of tokens of *like* in all its functions as spoken by both males and females from the younger and older groups. Though this table represents raw numbers of all functions of *like* rather than only the innovative functions, based on this cursory analysis, older speakers do appear to use *like* less often than younger speakers: 310 total tokens to the younger group's total of 1215. Even divided by sex, the older speakers do not use half the number of tokens as the younger speakers.

Table 4.1: Total Buckeye corpus rates of *like*, all functions

| Number | Sex | Old | Young | Total |
|--------|------------|------------------|-------------------|-------------------|
| 20 | Males | 191 / 105609 .2% | 739 / 82060 .9% | 930 / 198843 .5% |
| 20 | Females | 119 / 99626 .1% | 476 / 85042 .6% | 595 / 184668 .3% |
| 40 | Both Sexes | 310 / 205235 .2% | 1215 / 167102 .7% | 1525 / 383511 .4% |

Considering the saliency of the marker functions of *like* found in this study's sample, and the previous research showing these marker functions to be age graded, it is entirely plausible that this disproportionate distribution is in fact due to the innovative discourse and quotative markers.

This small sample is justified in the large number of tokens (798) that even this small sample yielded. Though Gosy & Horvath (2010:139), for example, examined speech from 20 different speakers, they extracted only 195 and 535 tokens for each of their Hungarian morphemes of interest which, similar to *like*, have more than one function (141). The interviews in this study lasted an average of 52 minutes, with specific interview times detailed in table 4.2.

Table 4.2: Interview times for each speaker

| | |
|--------------------|---------------------------------|
| All Speakers Total | 208.45 minutes or 3.47 hours |
| Speaker 04f | 60.67 minutes |
| Speaker 21f | 39.42 minutes |
| Speaker 30m | 48.65 minutes |
| Speaker 40m | 59.71 minutes |

4.2 STATISTICAL RESULTS FOR SPEAKER VARIABLE

The four speakers studied do show some variation among themselves with respect to their rates of function use.

Table 4.5: Rates of function use per speaker

| Function | Speakers | | | | Total Functions |
|-------------------|----------|-----|------------|-----|-----------------|
| | 04f | 21f | 30m | 40m | |
| Verb | 18 | 9 | 18 | 11 | 56 |
| | 16% | 6% | 5% | 6% | 7% |
| Conjunction | 13 | 4 | 14 | 25 | 56 |
| | 12% | 2% | 5% | 13% | 7% |
| Preposition | 25 | 14 | 27 | 77 | 143 |
| | 22% | 9% | 8% | 40% | 18% |
| Discourse Marker | 36 | 78 | 258 | 74 | 446 |
| | 32% | 50% | 77% | 38% | 56% |
| Quotative Marker | 20 | 52 | 18 | 5 | 95 |
| | 18% | 33% | 5% | 3% | 12% |
| TOTAL Per Speaker | 112 | 157 | 335 | 192 | TOTAL 796 |

In particular, speaker 30m uses the discourse marker function much more heavily than the other speakers: 258 instances, which constitutes 77% of his total use of the morpheme *like*. This outlying result makes comparing results across speakers more challenging and extrapolation more complex. This unusual rate is the most extreme outlier in the data. Other than speaker 30m's extremely heavy use of discourse marker *like*, the speakers' rates of use are relatively similar.

The relationship between speaker, function, and duration was measured using a 2-way ANOVA, performed once using unbalanced data and once with balanced data. Because the

functions evaluated here are extracted from several different speakers, interference from differences in speaker idiolects is possible. Table 4.3 shows the statistical significance of speaker, function, and the interaction between those two variables.

Table 4.3: Results from unbalanced ANOVA

| Variable | Degrees of Freedom | F Value | P-Value |
|------------------------------|--------------------|---------|--------------------|
| Speaker | 3 | 2.88 | 0.0350 |
| Function | 4 | 6.77 | < 0.0001 |
| Speaker*Function Interaction | 12 | 2.16 | 0.0122 |

The significance level is $p < .05$. At a p-value of 0.035, the speaker is significant. Function is also significant, at a p-value of less than 0.0001. Thus, the duration of any given instance of *like* is determined first by the function that the token is performing in this context, and then by the speaker. The interaction between the two variables is also significant at a p-value of 0.0122. Therefore, not only does each speaker show his or her own overall duration difference from other speakers, but also his or her own functions are predictive of the length of duration of articulation.

However, when the results are balanced to control for a speaker's heavy use of a particular function, speaker is not significant. Balanced results account for differing function frequencies for different speakers. Balanced results behave as if each speaker had used the same amount of tokens for each function. When results are balanced, one speaker using more tokens of a particular function is removed from the effect. It is not actually the speaker that effects the duration difference, causing an interaction between speaker and function, but his or her status as a heavy user of a function, since frequent use of a token affects (presumably, shortens) duration (Gahl 2008:474).

Table 4.4: Results from balanced ANOVA

| Variable | Degrees of Freedom | F Value | P-Value |
|------------------------------|--------------------|---------|---------------|
| Speaker | 3 | 1.08 | 0.3573 |
| Function | 4 | 3.95 | 0.0035 |
| Speaker*Function Interaction | 12 | 1.89 | 0.0325 |

In Table 4.4, in which the results are balanced, speaker has a p-value of 0.3573, so speaker no longer has a statistically significant effect on duration. Balancing results to account for some speakers being heavy users of certain functions, speaker turns out not to be a relevant variable in determining the durations of *like*. This is an expected finding, since function, rather than speaker, should be the determining factor for the duration of a particular instance of *like*.

4.3 STATISTICAL RESULTS FOR FUNCTION VARIABLE

Table 4.6 shows the mean durations of each salient function of *like* across speakers. Contrary to predictions, the marker functions were not significantly shorter in duration than the ungrammaticalized usages. In fact, the opposite of this expected result is found: the marker functions show the longest mean duration of all the functions. As Table 4.6 shows, the conjunction function (C) is significantly shorter than the verb (V), discourse marker (DM), and quotative marker (QM). The preposition usage (P) is significantly shorter than the discourse (DM) and quotative (QM) markers. Members of the same group (A, B, or C) are not significantly different from each other.

Table 4.6: Mean durations in milliseconds: Shortest to longest

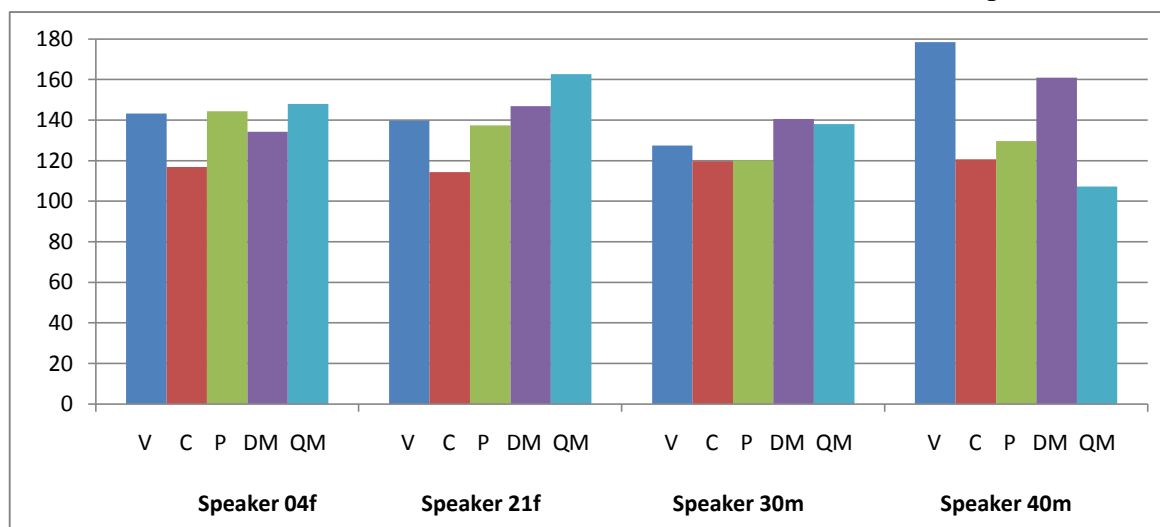
| Function | Mean Duration | Groups not significantly different |
|----------|---------------|------------------------------------|
| C | 118 | A |
| P | 129 | A B |
| V | 146 | B C |
| DM | 148 | C |
| QM | 151 | C |

Even though speaker is not a significant variable when results are balanced to account for heavy use of a function, duration has to be compared within speakers rather than across speakers. This within-speaker comparison of function durations is necessary in order to control for overall speech rate. Some speakers may have an overall rate which is faster or slower than others, but within a single session interview, in which the conditions of the interview did not change, the

speaker's individual speech rate should remain about the same, making comparisons of function durations within one speaker's interview possible.

Comparing individual speaker's durations in Table 4.7, it is clear that, contrary to the prediction, the marker functions of *like* are not shorter in duration than the other functions. Speaker 30m stands out from the other informants because of his much heavier use of *like* as a discourse marker (see Table 4.5). Intriguingly, speaker 30m, who is by far the heaviest user of the discourse marker *like*, has a duration of the discourse marker (DM) which is slightly longer than any of his other functions.

Table 4.7: Durations in milliseconds of each function for each speaker



Though the grammaticalized marker functions of *like* are among the most frequent functions of *like*, and certainly the discourse marker function is by far the most frequent, the frequency effects these functions are undergoing are not what was expected, based on past studies and the common attrition effect of grammaticalization.

CHAPTER 5: CONCLUSIONS

5.1 DISCUSSION OF THE RESULTS

As Andersen points out in his description of the classification of marker *like*, “a brief pause immediately following *like* . . . fairly strongly suggests that *like* is a pragmatic marker” (2001:213). This use of *like* as a hesitational device, especially preceding a pause to retain one’s turn while the speaker searches for the desired expression, may account for the long mean duration of the discourse *like*. Perhaps because *like* is meant to buy the speaker time to arrange his or her thoughts, a longer duration would be valuable, negating the effects of phonological reduction from grammaticalization.

Since discourse *like* is thus positioned, preceding a pause and allowing a speaker to hold his or her turn and listener’s attention while planning the rest of the utterance, it is syntactically distinct from the rest of the sentence (Traugott 1995:14). However, “an overwhelming number of tokens of the pragmatic marker *like* occur where neither speed of production nor discourse coherence suggest that there are any planning difficulties involved,” as Andersen (2001:227) points out. Nevertheless, discourse markers can be identified by their ability to be removed from the sentence without a change in the overall propositional content. These markers are syntactically free (Romaine & Lange 1991:252), so they do not serve a grammatical function in the sentence in which they occur (Brinton 1996:34). Discourse markers can be extricated from the sentence without affecting the propositional content of the statement (Brinton 1996:35). Thus syntactically separate, perhaps discourse *like* is retaining prosodic difference at the same time.

Likewise, quotative *like* is more syntactically separate from the rest of the sentence than the other, nongrammaticalized functions, in that it precedes a quotation. The quotation may only represent a vague impression of the party's general attitude, but it is nevertheless syntactically distinct from the rest of the speaker's narrative.

This syntactic distinctness may be used to explain the marker functions' longer overall duration, compared to the other functions of *like*. The preposition and conjunction uses of *like* in particular are often engulfed in a phrase they cannot be described as distinct from, such as *things like that* and *like I said*.

Like I said is a good illustration of another possible explanation for the unexpected duration results. Interestingly, speaker 40m has a very low standard deviation for the conjunction function (0.0189, compared to the overall mean standard deviation 0.0378), most likely because he uses the collocation *like I said* with unusually high frequency and prosodic consistency.

(58) It's just having it forced on me *like I said* just sort of gives me a different opinion. (s40m)

Speaker 40 uses *like I said* so frequently, its use suggests it may be employed as a phrasal discourse marker.

Since the frequent collocates *like I said* and prepositional phrases such as *something like that*, *things like that*, and *stuff like that* are categorized here as less grammaticalized functions, their durations are not included in the discourse marker durations. However, *like I said* collocates with other discourse markers:

(59) Akron/Cuyahoga Falls area but *like I said* I my *uh* -- I had apartment down there. (s30m)

(60) It --it was known that he was the father, *like I said*, *I mean*, he --he was. (s40m)

These instances of *like I said* also seem to co-occur with planning difficulties, which also suggests they are used as discourse markers. Perhaps this frequent use of *like I said* as a phrasal discourse marker which functions as a quickly-articulated single unit is articulatory routine for speakers such as speaker 40. This frequent collocation of *like* in a phrase may account for its showing the shortest overall duration.

Another possibility is that, though the marker functions of *like* grammaticalized from the other lexical entries discussed in Chapter 2, they are not the only functions here that may have grammaticalized as part of their history. Table 2.2 shows Romaine & Lange's (1991) theory for the development of the marker functions, in which the conjunction usage of *like* first grammaticalized from the preposition function and then the marker functions grammaticalized from the conjunction. Table 4.7 shows that for each speaker, the conjunction usage is the shortest in duration. Thus, perhaps the conjunction form of *like* is simply more grammaticalized than the marker functions, thus displaying a shorter duration.

An alternative explanation for the unexpected results is that high frequency of use can produce either phonetic reduction or fossilization thereof. Just as the English copulative, the most frequently employed verb in the language, is also the most suppleted, sometimes very frequent forms actually retain phonetic difference rather than weaken it. It is also plausible that because the discourse and quotative functions of *like* are so liberally applied in the speech of these young speakers, their phonetic weight is more devotedly retained than in the less salient, less grammaticalized functions.

5.2 AMBIGUITY OF DURATION

Another challenge of disambiguation in the data analysis was finding the boundaries of *like* in certain contexts of connected speech. Demarcating the right side of the word was usually

simple since the right boundary was determined at the closure rather than the release of the [k] and the closure was easy to find, as can be seen in Figures 3.1 and 5.1 and in Appendix I. Depending on the articulation, the velar pinch that characterizes the formants as they transition from the diphthong to the final consonant sometimes ends abruptly and sometimes lengthens into a long tail. Initially, this variation caused some difficulty in determining where to mark the right edge. However, after all the speakers were analyzed and the possible range of variation was more fully grasped, the right edges were re-checked, and in some cases re-captured, and maximum consistency was achieved.

On the other hand, the left side of the word was somewhat more difficult to consistently capture, as the [l] was sometimes assimilated into the preceding word, such as in the collocation “I *like*,” shown in Figure 5.1.

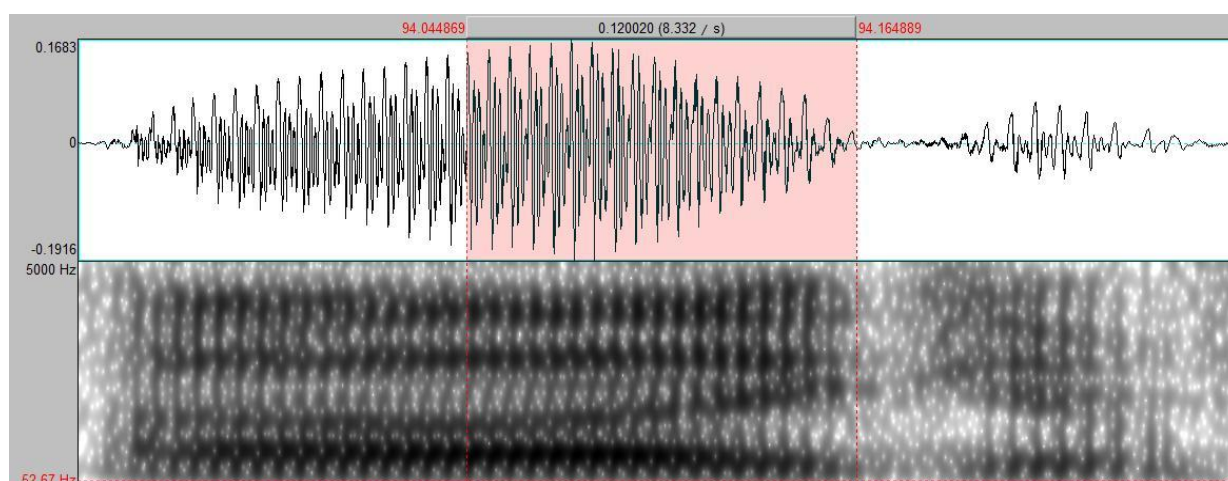


Figure 5.1: “I like”

In these challenging cases, the border of the word was located by confirming two out of three of the visual and auditory indicators: the wave form, spectrogram, and repetition of the audio while experimenting with moving the boundary one wave form at a time. In Figure 5.1, *like* was finally demarcated where the highlighting indicates.

This difficulty in finding the edges of these tokens of *like* may have affected the final results of durations. Perhaps in future studies this particular shortcoming of this pilot can be resolved by discarding these ambiguous cases or by finding a more reliable method of disambiguation. As it stands, this study has aimed for consistency in bearing out these difficult cases by the method described above.

5.3 SOCIOLINGUISTIC OBSERVATIONS

Due to the limited sample examined in this study, sociolinguistic observations cannot really be generalized. Nevertheless, there are some interesting patterns visible in this preliminary study. The discourse marker *like* is generally found to be associated with female speakers (Andersen 2000:287). However, in this data, the male speakers use more than twice as many tokens of *like* as a discourse marker than the female speakers.

With respect to the quotative marker function, Romaine & Lange (1991:251) found 83% of their examples of *be like* in the speech of female speakers rather than male. Consistent with Romaine & Lange's findings, the female speakers in this study were more likely to use *be like* than the males were. This is in contrast with the findings in the Blyth et al. study, where males were more likely (0.534) than females (0.466) to use *be like* with a p-value of less than 0.006 (1990:221).

Table 5.1: χ^2 female versus male function uses

| | Females | | Males | | Total |
|-----------|--------------|----------|--------------|----------|--------------|
| | Observations | Expected | Observations | Expected | Observations |
| V | 27 | 18.9 | 29 | 37.1 | 56 |
| C | 17 | 18.9 | 39 | 37.1 | 56 |
| P | 39 | 48.3 | 104 | 94.7 | 143 |
| DM | 114 | 150.7 | 332 | 295.3 | 446 |
| QM | 72 | 32.1 | 23 | 62.9 | 95 |
| Total | 269 | | 527 | | 796 |

$$\chi^2=96.535$$

However, it is important to recall that the small number of speakers sampled here makes true extrapolation from these results unviable. The high rate of use of the discourse marker function by males, for example, may be due entirely to one high-rate user, speaker 30m, who used 258 tokens to the other speakers' average of 62.6 (see Table 4.5). The other male speaker, speaker 40m, is in fact outperformed in use of the discourse marker function by one of the female speakers, speaker 21f.

Table 5.2: Individual speakers' rate of use of discourse marker *like*

| | DM |
|-------------|-----|
| Speaker 04f | 36 |
| Speaker 21f | 78 |
| Speaker 30m | 258 |
| Speaker 40m | 74 |

The only consistent indicator of sociolinguistic distribution that can be taken from these χ^2 results is that females did outperform males in quotative marker usage (Reeves, personal communication).

5.4 OTHER CONCERNS

Some tokens were located in the audio which were missed in the compilation of the Antconc files, and there are still some inconsistencies in the number of tokens extracted and the number found by Antconc, but the differences are considered negligible and analysis was nevertheless conducted, despite the slight disparity.

Table 5.3: Concordance and final count discrepancies

| Speaker | Total Durations Extracted | Concordance Total |
|---------|---------------------------|-------------------|
| 04f | 112 | 107 |
| 21f | 158 | 160 |
| 30m | 336 | 339 |
| 40m | 192 | 194 |
| TOTAL | 798 | 800 |

There were also some outlying tokens with extremely long durations, i.e., 0.30 seconds and higher. Speaker 40m has a comparatively high rate of such long durations. For speaker 40m, 6 out of 192 tokens were 0.30 seconds or longer, so these durations were not considered outliers and were not excluded. However, for two other speakers, these long durations were actually exceptional, and so they were excluded: one of speaker 30m's verbs lasted 0.367 seconds and speaker 21f laughed during articulation of a discourse marker function so that it had a duration of 0.333 seconds. Both of these tokens were excluded from analysis, though they are included in the durations frequencies compiled in Table 4.8.

Also, the rate of speech in which each token occurs was not measured. Rather, the rate of speech for each speaker was assumed mostly consistent, or consistent enough to make comparisons within each speaker's use. If, as in the Gahl (2008:482) study, the ambient speech rate had been measured to normalize durations, the results may have differed. In addition to speech rate, future studies could control for "segmental context, pitch accent, and contextual predictability," which can all affect duration (Gahl 2008:478).

5.5 CONCLUSION

According to the literature, the multifunctional English morpheme *like* has grammaticalized into two marker functions: a discourse marker and a quotative marker. These salient functions are used with great frequency in the speech of young Americans of both sexes. However, despite their frequency, these markers do not show the phonological reduction that is somewhat typical of grammaticalized morphemes. The results do not bear out this study's hypothesis that the grammaticalization of *like* would be evident in the marker functions' durations; in fact, the opposite was found to be true: the marker functions had the longest mean durations.

The hypothesis regarding the frequency effects of grammaticalization must therefore be reformulated. It is not entirely clear whether it is syntactic separation, phonological preservation through frequent use, or a combination thereof that allows the grammaticalized forms of *like* to appear with longer durations than the original lexicon entries from which they developed. It is clear, though, that these marker functions have grammaticalized and have developed or retained an articulatory duration length that separates them from the other functions of *like*.

REFERENCES

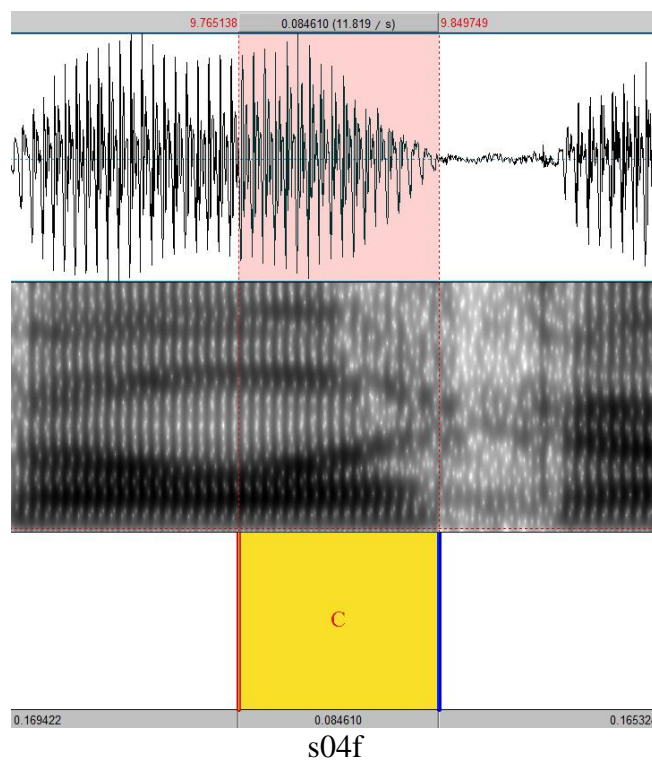
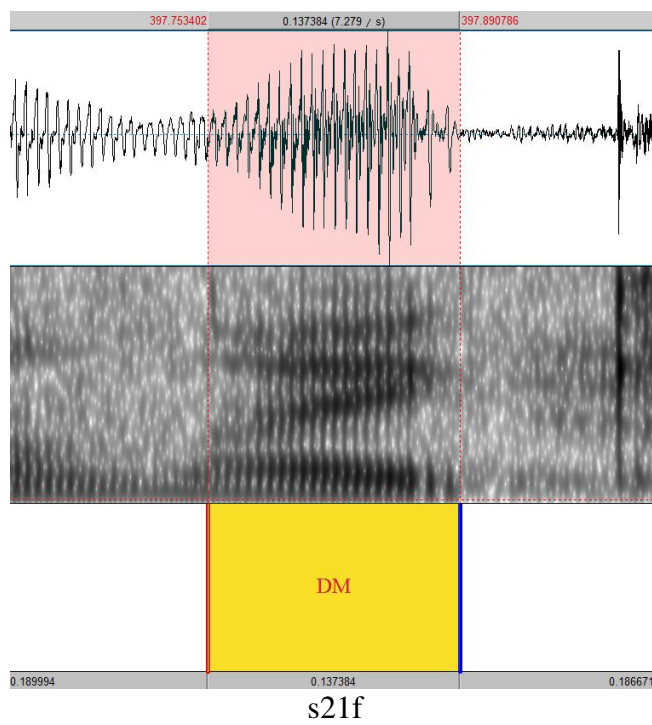
- Andersen, Gisle. 2000. The role of the pragmatic marker *like* in utterance interpretation. *Pragmatic markers and propositional attitude*, ed. By Gisle Andersen & Thorstein Fretheim. Amsterdam: John Benjamins Publishing Company. 17-38.
- . 2001. *Pragmatic markers and sociolinguistic variation: A relevance-theoretic approach to the language of adolescents*. Amsterdam: John Benjamins Publishing Company.
- Anthony, L. 2010. AntConc (Version 3.2.1) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.antlab.sci.waseda.ac.jp/>
- Blyth, Jr., Carl, Sigrid Recktenwald & Jenny Wang. 1990. I'm like, "Say what?!": A new quotative in American oral narrative. *American Speech* 65.3. 215-227.
- Boersma, Paul & David Weenink. 2004. Praat: Doing phonetics by computer (version 4.4). <http://www.praat.org>.
- Brinton, Laurel J. 1996. *Pragmatic markers in English: Grammaticalization and discourse functions*. Berlin: Mouton de Gruyter.
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Dailey-O'Cain, Jennifer. 2000. The sociolinguistic distribution of and attitudes toward focuser 'like' and quotative 'like'. *Journal of Sociolinguistics* 4, 1: 60-80.
- Ferrara, Kathleen & Bell, Barbara. 1995. Sociolinguistic variation and discourse function of constructed dialogue introducers: The case of be + like. *American Speech* 70.3. 265-290.
- Gahl, Suzanne. 2008. *Time and thyme* are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84.3, 474-496.

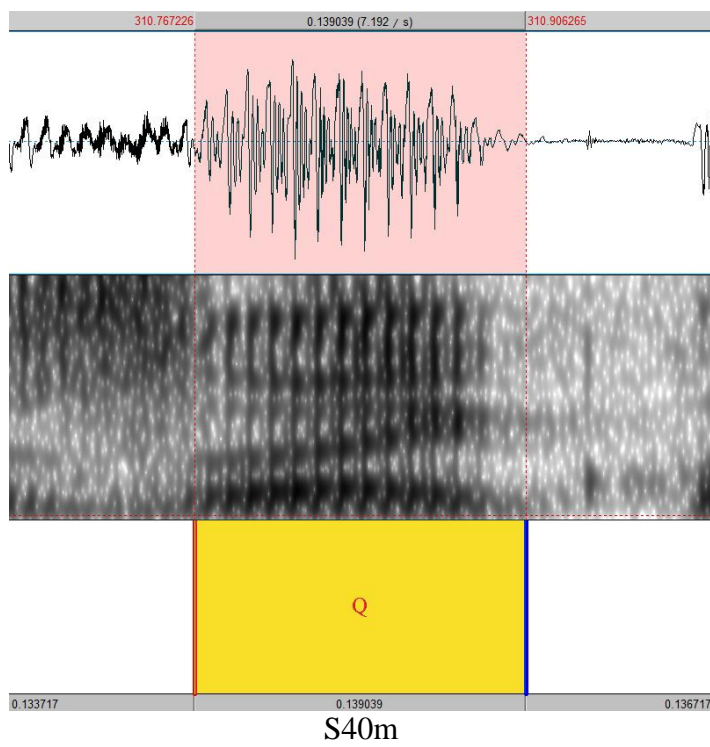
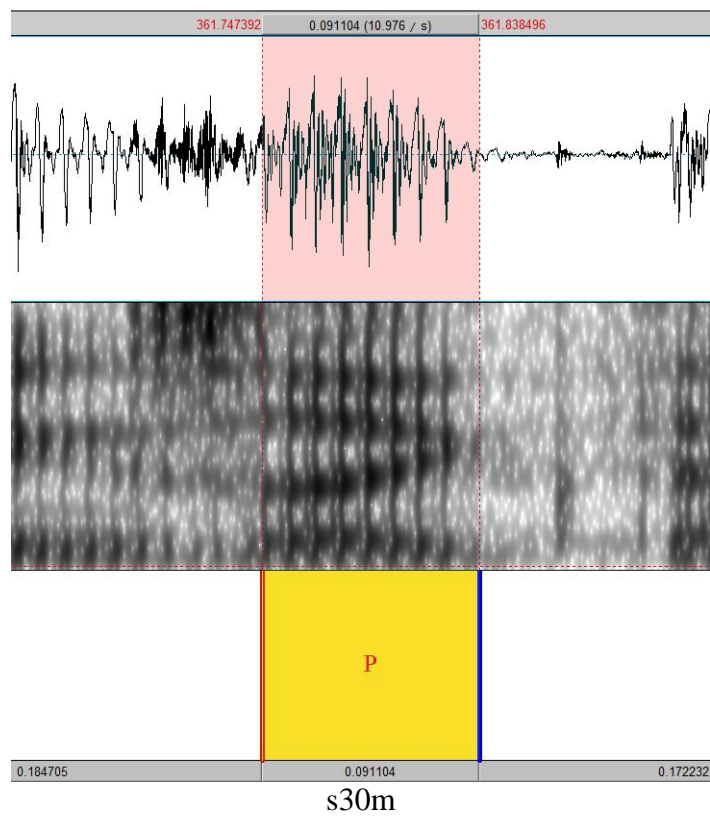
- Gósy, Mária & Viktória Horváth. 2010. Changes in articulation accompanying functional changes in word usage. *Journal of the International Phonetic Association* 40.2. 135-158.
- Heine, Bernd, Ulrike Claudi, & Friederike Hünemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: The University of Chicago Press.
- Holmes, Janet. 1986. Functions of you know in women's and men's speech. *Language in Society* 15.1. 1-21.
- Hopper, Paul J. 1991. On some principles of grammaticalization. *Approaches to grammaticalization*, vol. 1, ed. by Elizabeth Closs Traugott & Bernd Heine. Amsterdam: John Benjamins Publishing Company, 17-35.
- Hopper, Paul J. & Elizabeth Closs Traugott. 1993. *Grammaticalization*. Great Britain: Cambridge University Press.
- Jespersen, Otto. 1942. *A modern English grammar on historical principles, part VI morphology*. Copenhagen: Ejnar Munksgaard, 417-418.
- Kelley, Frannie. 2010. 'Whip My Hair' And Sesame Street Song Rare Moments Of Positivity In Viral Video. October 21, 2010. <http://www.npr.org/blogs/therecord>
- Kiesling, Scott, Laura Dilley, & William D. Raymond. 2006. The Variation in Conversation (ViC) project: Creation of the Buckeye Corpus of Conversational Speech.
- Kuryłowicz, Jerzy. 1965. The evolution of grammatical categories. *Diogenes* 51:55-71.
- Ladefoged, Peter. 2006. *A course in phonetics*, 5th ed. Australia: Thomson Wadsworth.
- Lange, Deborah. 1986. Attitudes of teenagers toward sex-marked language. Unpublished ms. Department of Linguistics, Georgetown University.

- "like" 1989. *The Oxford English Dictionary*, 2nd ed. *OED Online*. Oxford University Press. October 2010 <<http://dictionary.oed.com>>.
- Meehan, Teresa. 1991. It's like, 'what's happening in the evolution of like?': A theory of grammaticalization. *Kansas working papers in linguistics*, 16. 37-51.
- Neufeldt, Victoria. 1988. *Webster's New World Dictionary of American English*, 3rd college edition, 783. New York: Prentice Hall.
- Ocampo, Francisco. 2006. Movement towards discourse is not grammaticalization: The evolution of *claro* from adjective to discourse particle in spoken Spanish. In N. Sagarra and A. J. Toribio, editors, *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, pages 308–319. Cascadilla Proceedings Project, Somerville, MA.
- Pei, Mario. 1973. *Double-speak in America*. New York, Hawthorn Books.
- Pitt, M.A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E. and Fosler-Lussier, E. 2007. Buckeye Corpus of Conversational Speech (2nd release) [www.buckeyecorpus.osu.edu] Columbus, OH: Department of Psychology, Ohio State University (Distributor).
- Ragan, Sandra L. 1983. "Alignment and conversational coherence." *Conversational coherence: Form, structure and strategy*, ed. by Robert T. Craig and Karen Tracy, 157-171. Beverly Hills: Sage Publications.
- Reeves, Jaxk. 2010. Personal communication. Statistical consultation, University of Georgia.
- 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*, 66.3. 227-279.
- Schiffrin, Deborah. 1987. *Discourse Markers*. Great Britain: Cambridge University Press.
- Schourup, Lawrence. 1999. Discourse markers. *Lingua* 107. 227-265.

- Sinclair, John. (ed.) 1987. *Collins COBUILD English language dictionary*. Great Britain: Harper Collins Publishers.
- Stenstrom, Anna-Brita, Gisle Andersen, & Ingrid Kristine Hasund. 2002. *Trends in Teenage Talk: Corpus compilation, analysis and findings*. Amsterdam: John Benjamins Publishing Company.
- Tannen, Deborah. 1986. "Introducing constructed dialogue in Greek and American conversation and literary narrative." *Coulmas* (1986c), 311-332.
- Tagliamonte, Sali & Rachel Hudson. *Be like et al. beyond America: The quotative system in British and Canadian youth*. *Journal of sociolinguistics* 3.2. 147-172.
- Traugott, Elizabeth C. 1995. The role of the development of discourse markers in a theory of grammaticalization. Paper presented at ICHL XII.
- Underhill, Robert. 1988. *Like is, like, focus*. *American speech* 63.3. 234-245.
- Watts, Richard J. 1989. Taking the pitcher to the 'well': Native speakers' perception of their use of discourse markers in conversation. *Journal of pragmatics* 13.2. 203-237.
- Wentworth, Harold & Stuart Berg Flexner. (eds.) 1960. *Dictionary of American slang*. New York: Crowell.

APPENDIX I





APPENDIX II

Buckeye Corpus Metadata

| SPEAKER | SPEAKER'S GENDER | SPEAKER'S AGE | INTERVIEWER GENDER |
|----------------|-------------------------|----------------------|---------------------------|
| S01 | f | y | f |
| S02 | f | o | m |
| S03 | m | o | m |
| S04 | f | y | f |
| S05 | f | o | f |
| S06 | m | y | f |
| S07 | f | o | f |
| S08 | f | y | f |
| S09 | f | y | f |
| S10 | m | o | f |
| S11 | m | y | m |
| S12 | f | y | m |
| S13 | m | y | f |
| S14 | f | o | f |
| S15 | m | y | m |
| S16 | f | o | m |
| S17 | f | o | m |
| S18 | f | o | f |
| S19 | m | o | f |
| S20 | f | o | f |
| S21 | f | y | m |
| S22 | m | o | f |
| S23 | m | o | m |
| S24 | m | o | m |
| S25 | f | o | m |
| S26 | f | y | f |
| S27 | f | o | m |
| S28 | m | y | m |
| S29 | m | o | f |
| S30 | m | y | m |
| S31 | f | y | m |
| S32 | m | y | f |
| S33 | m | y | f |
| S34 | m | y | m |
| S35 | m | o | m |
| S36 | m | o | f |

| | | | |
|-----|---|---|---|
| S37 | f | y | m |
| S38 | m | o | m |
| S39 | f | y | m |
| S40 | m | y | f |

(Kiesling 2006:4)