

## Frequency and variation in the community grammar: Tracking a new change through the generations

SALI A. TAGLIAMONTE

*University of Toronto*

ALEXANDRA D'ARCY

*University of Canterbury*

### ABSTRACT

In this article we perform a quantitative analysis of verbs of quotation in a cohesive speech community. The incoming form *be like* overshadows all other quotative verbs among speakers under 30. This telescoped rate of change provides an opportunity to investigate the actuation problem as well as to probe the underlying mechanism of change in the contrasting variable grammars across generations. Multivariate analyses of factors conditioning *be like* (content of the quote, grammatical person, sex) reveal stability in the significance of constraints, however the rankings and relative strengths reveal subtle ongoing changes in the system. Interpreting these in sociocultural context, we suggest that *be like* is an innovation that arose out of a preexisting niche in the grammar. It accelerated during the 1980s due to its preppy associations, later specializing as a marker of narrative present. In accounting for these findings, we are led to contrast generational and communal change and to question what it means to 'participate' in linguistic change.

Quotative *be like*, exemplified in (1), is a vigorous change in contemporary English. Among Canadians in their twenties, it has risen from an incipient phase of 13% in 1995 (Tagliamonte & Hudson, 1999) to accounting for 58% of all quotatives in 2002 (Tagliamonte & D'Arcy, 2004:501, Table 2).

- (1) a. I'd *be like*, "Hey." I'd be with a friend of mine.  
They're *like*, "Who is that person?"  
I'm *like*, "I don't know." (N/f/f/32)<sup>1</sup>

The first author gratefully acknowledges the support of the Social Sciences and Humanities Research Council of Canada (SSHRC) for grant #410-2003-0005 "Linguistic changes in Canada entering the 21st century" and the Research Opportunities Program at the University of Toronto (ROP). Special thanks are owed to the research assistants on these projects who have been instrumental in transcribing, and in some cases extracting, the data: Stephanie Cali, Jonille Clemente, Derek Denis, Louis Filipe, Kevin Galloway, Marion Hau, Matthew King, Megan Lyon, Marion Macdonald, Sonja Molfenter, Jenny Seppänen, Karina von Stedingk, Muhammad Velji, Jessica Wertman, Lina Yang, and Stacy Yeh. An earlier version of this article was presented at N.WAVE 33 (Ann Arbor, October 2004), where feedback from our audience was instrumental. We are also grateful to Nathalie Dion and Shana Poplack for their discerning comments on an earlier written version.

	Sample Size	Age Range	Speech Community	Total <i>N</i>
Blyth et al., 1990	30	20–72	Cornell University	
Ferrara & Bell, 1995	115	18–40+	Texas	485
Dailey-O’Cain, 2000	30	14–69	southeastern Michigan	
Singler, 2001		9–51	“no geographic limitations”	5,898
Cukor-Avila, 2002	14	14–89	Springville	3,203
Buchstaller, 2004	136	–/+ 45	“all areas of the US”	1,371
	64		Derby, Newcastle	2,064
Tagliamonte & Hudson, 1999	44	18–28	York	665
	23		Ottawa	612
Tagliamonte & D’Arcy, 2004	44	10–19	Toronto	2,058
D’Arcy, 2004	16	8–16	St. John’s	184

Note. Shading denotes that the information is not recoverable from the text.

FIGURE 1. Details of apparent time studies of the quotative system.

- b. *I’m like*, “Oh my God! Did you make pee-pee?”  
*He’s like*, “Yeah.”  
 So then I *was like*, “Oh no.” (2/a/f/16)
- c. *We’re like*, “How was it?”  
 And they’*re like*, “Oh, it was different.”  
*We’re like*, “‘Good’ different?”  
 They’*re like*, “You’ll see, you’ll see.” (N/ð/m/26)

The sheer rate of change makes it plausible that this might be the upswing of the classic S-curve of linguistic change (Bailey, 1973; Weinreich, Labov, & Herzog, 1968). If so, we have an opportunity to investigate the incrementation problem, “the mechanism by which changes advance in a stepwise fashion” (Labov, 2001:446).

Real time data provides the best material for tracking the progression of change, yet Cukor-Avila (2002), which follows three African American speakers, and Ferrara and Bell (1995), which includes a trend component, are the only real time studies of the quotative system of which we are aware. Lacking this type of data, the apparent time construct provides the next best tool, and is most effective when the speakers under consideration share the same background. Thus, Chambers and Trudgill (1998:151) wrote:

Studying the diffusion of innovations in apparent time involves surveying the differences between the speech of people of different ages in the same community, while controlling the other variables such as sex, social class and ethnicity.

To this point, there has been ample research on *be like*, but none has been based on a large-scale socially stratified sample from a single, cohesive, speech community. Consider Figure 1. The earliest study, Blyth, Recktenwald, and Wang

(1990:216), included the full spectrum of the adult population, but there are just 30 speakers from a range of backgrounds; most were affiliated with Cornell University in “some capacity.” Ferrara and Bell (1995) was primarily a real time study, but the distribution by age also enabled an apparent time comparison of quotative use in each of their three corpora, collected in 1990, 1992, and 1994. However, the corpora represent assorted speech communities. Corpus Two, for example, came from a “diverse population” in Texas, including Dallas, Houston, San Antonio, and their suburbs (1995:273; fn.10, 277). Moreover, there is no way to recuperate ethnicity or other social factors. The data presented in Dailey-O’Cain (2000:64) spanned a broad age range, 14–69, but in this case the sample was relatively small: 30 speakers. Singler (2001) has been the most ambitious effort to date. His corpus totalled nearly 6,000 quotatives from speakers between the ages of 9 and 51, but there were “no geographic limitations as to where in the United States the speakers came from” (2001:259). Cukor-Avila (2002:6) included an apparent time study, but it was based on a modest 14 speakers in Texas, aged 14–89. Buchstaller (2004:Appendix 1, 296–297) is a comprehensive cross-variety comparison. However, here too there is a mix between broad-based sampling and/or restricted age distributions. In sum, these works are either too narrowly circumscribed (number of speakers) or too broad to be representative (regionality). Other studies, such as Tagliamonte and Hudson (1999), Tagliamonte and D’Arcy (2004), and D’Arcy (2004), worked with tightly defined regional varieties, but focused uniquely on younger members of the population. Thus, despite the wealth of investigations focusing on the quotative system and those on *be like* in particular, there remains little information regarding its diffusion at the community level.

Critical to the exploration we undertake here are the issues surrounding different types of linguistic change, namely age-grading versus generational change versus communal change. These are approachable most felicitously with the perspective of a speech community as a whole, so that the regularities of the variation within can be viewed as with a microscope (see Labov, 2001:39).

Consider Labov’s definitions in (2).

(2) Patterns of Linguistic Change (Labov, 1994:84)

*Age-grading*

“If individuals change their linguistic behaviour throughout their lifetimes, but the community as a whole does not change, the pattern can be characterized as one of age-grading.”

*Generational change*

“Individual speakers enter the community with a characteristic frequency for a particular variable, maintained throughout their lives; but regular increases in the values adopted by individuals, often incremented by generations, lead to linguistic change for the community.”

*Communal change*

“In communal change all members of the community alter their frequencies together or acquire new forms simultaneously.”

*Age-grading* refers to the association of some feature with a particular life stage that repeats in every generation. This is the type of change that is often associated with adolescents. Whenever novel linguistic features are reported in this population, they tend to be considered age related and are expected to recede as speakers age, yet it is important to bear in mind that very few changes of this type have been reported (Chambers, 2003:206). *Generational change* is the standard type of linguistic change in which we observe incremental shifts in the frequency of forms from one age group to another. Generational change is based on the assumption that once stabilized, an individual's linguistic usage does not change over the course of his or her lifetime (Labov, 2001:454; Milroy & Gordon, 2003:36). Finally, there is *communal change*. This is when individuals themselves shift the frequency of linguistic features over their lifetimes alongside incremental change from one age group to another. Labov (2001:447) hinted that the assumption of stability, particularly for young adults, may have to be revised. Indeed, there is building evidence that young adults continue to advance ongoing linguistic changes well into adulthood (Nordberg & Sundgren, 1998; Nevalainen & Raumolin-Brunberg, 2003). In other words, the most recent research on linguistic change suggests that generational change and communal change can progress simultaneously. As we shall see, we believe that there are subtle differences that distinguish the progression of these types of change among the adult population.

Our points of departure in this analysis are: (1) to attempt to uncover the origins of *be like*; (2) to consider its developmental trajectory and the underlying mechanisms that may be guiding its pathway; and (3) to explore whether or not we can use the diffusion of *be like* to understand how the proposed types of linguistic change unfold in the community. A different way to frame this query, perhaps, is to think about when individuals in a community begin participating in a linguistic change and when—indeed if—they stop.

#### METHOD

To explore these issues, we performed a quantitative analysis of the quotative system in the Toronto English Corpus. This 1.5 million-word archive is comprised of sociolinguistic interviews collected using a combination of quota-based random sampling and social networking.<sup>2</sup> As a fundamental sampling criterion, participants must have been born and raised in Toronto. The current data set includes 199 speakers, aged 9 to 87, from whom we have extracted over 6,300 quotative tokens. The sample details are outlined in Table 1.<sup>3</sup>

Each token was coded for the operation of internal and external constraints. The two linguistic factors that have received the most attention in quotative analyses are grammatical person and pragmatic content. The grammatical person constraint operates so that *be like* is favored with first person subjects, as in (3). The consistency of this effect across studies (e.g., Blyth et al., 1990; Cukor-Avila,

TABLE 1. *The Toronto quotative corpus*

Age	N speakers	N quotatives
9–12	14	396
13–14	7	204
15–16	12	505
17–19	54	1,992
20–29	38	1,138
30–39	13	524
40–49	25	550
50–59	16	503
> 60	20	552
Total	199	6,364

2002; Ferrara & Bell, 1995; Tagliamonte & Hudson, 1999) led Tagliamonte and D’Arcy (2004:509) to suggest that it is a defining characteristic of *be like*.

- (3) a. *I’m like*, “Okay, what are you going to say next?”  
*I’m like*, “Now you’re going to say something bad.” (3/T/f/18)
- b. *We’re like*, “No, don’t spray us!”  
*We are like*, “No, don’t hit us!” (3/U/f/12)

The content of the quote has also been a fundamental aspect of *be like*, which seems to have originated as a marker of internal dialogue and nonlexicalized sounds, as in (4) (Butters, 1982; Tannen, 1986). However, where previously the strongest constraint on *be like* (e.g., Cukor-Avila, 2002; Tagliamonte & Hudson, 1999), recent results have revealed that the pragmatic effect is the lowest ranked constraint among 17–19 year olds, suggesting that this effect may be weakening (Tagliamonte & D’Arcy, 2004:507).

- (4) a. *I was like*, “Man, what else am I going to do?” (1/%/f/27)
- b. *I’m like*, “Oh my God, what do I say? This is bad.” (2/k/m/11)
- c. *You’re just like*, “(weird sound).” (N/r/f/22)
- d. He didn’t know how to place his hands, and he *was like*, “(noises).” (2/a/f/16)

In addition to the classic factors of grammatical person and content of the quote, we consider two other characteristics of the grammar: (1) surface morphology and (2) temporal reference. The correlation of *be like* with present tense has been noted since the earliest research on this quotative form. Both Blyth et al. (1990) and Romaine and Lange (1991) reported that *be like* is more likely to occur in the present tense than in the past tense. More recently, Singler (2001:272–273) tested for a tense effect and found that tense was statistically significant with present tense favoring *be like*. However, occurrences of the present tense in his corpora were “almost always” instances of Historical Present (HP). The HP has a number of stylistic functions in narrative structure that are independent of the

variation that concerns us here (see Schiffrin, 1981; Wolfson, 1978, 1979). The relevant point for this analysis is simply that the HP is a specific and highly circumscribed construction in which surface morphology is present tense but the context of temporal reference is past tense. An analysis based solely on the HP is not sufficient for disentangling which underlying factor—tense or specifically HP—is contributing to the choice of forms. In other words, a contrasting comparative context is necessary. The Toronto materials provided sufficient data to test for a three-way distinction, enabling us to disambiguate the effects of tense and temporal reference. Consequently, data were coded for past morphology with past temporal reference (5a), present tense with present temporal reference (5b), and finally, present morphology with past temporal reference, the HP (5c).

- (5) a. And then after she *was like*, “Hi Tootsie-Roll.” (2/r/f/11)  
 b. Now it’s *like*, “I have to pay more than thirty-five dollars for jeans?” (1/@/f/19)  
 c. She’s *like*, “Put this in your eye.”  
*I’m like*, “Oh God.” (2/o/m/13)

In summary, the state of research on quotative *be like* provides a number of solid foundations on which to base a large-scale community study. With these in mind, we turn to the data.

## RESULTS

Figure 2 displays the frequency of each of the main quotatives according to speaker age. This graphic representation of the Toronto speech community reveals a dramatic division in the existing population. *Be like* overshadows all other forms among speakers under age 30, with the predicted “adolescent peak” appearing among the 17–19-year-olds (Labov, 2001:517). Conversely, *say* is by far the front runner among those over 40, but in the younger age groups its use steadily declines. This is where *be like* is supreme. Taking a broad view, observe the cross-over as one form appears to replace the other. In the middle, among the 30-year-olds, we find a generation in flux where the frequency of *be like* is virtually equal to that of *say*. This is the only group of speakers for whom this is the case.

Putting these findings in context with earlier research (Blyth et al., 1990; Cukor-Avila, 2002; Ferrara & Bell, 1995; Romaine & Lange, 1991; Singler, 2001; Tagliamonte & D’Arcy, 2004; Tagliamonte & Hudson, 1999), *be like* was in its incipient phase in the early 1980s (Buchstaller, 2004; Butters, 1982; Tannen, 1986). At that time the 30-year-olds in the Toronto corpus were teenagers. We would like to suggest that this group may well be the first generation of native users of *be like*. Observe the catastrophic shift between speakers in their 30s and 40s; we will return to this later.

The next question is, “Do these Torontonians share the same grammar for *be like*?” To explore this issue, we tested for the operation of the constraints discussed earlier. Table 2 provides a consistent variable rule analysis for each age group for whom *be like* is robust: the 9–14-year-olds, the 15–16-year-olds, the 17–19-year-olds, the 20–29-year-olds, and the 30–39-year-olds.<sup>4</sup>

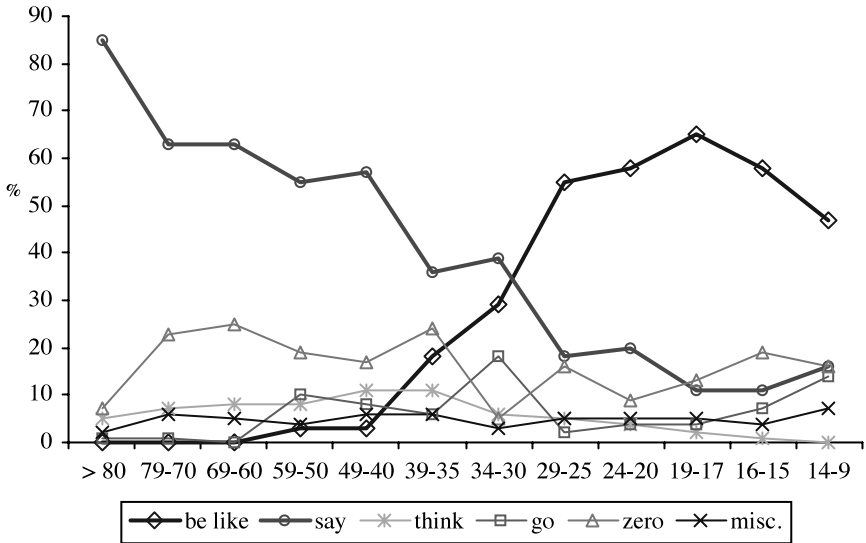


FIGURE 2. Overall distribution of quotatives across the generations in Toronto English.

Let us simply consider the statistical significance of factors. All the constraints are in place and they are all significant. Moreover, for each factor group that has previously been studied the direction of effects is as predicted. Internal dialogue consistently favors *be like* over direct speech.<sup>5</sup> The grammatical person constraint is stable with first-person subjects favoring *be like* whereas third persons disfavor it. Females favor *be like* more than males, except among the 30-year-olds. The most explanatory factor by far, however, is the intersection of surface morphology and tense. The HP is the most important context for the use of *be like* in every age group. The factor group morphology/tense is the strongest for all age groups, displaying the highest range, and the HP is the most favorable environment within this factor group for each age cohort.

The consistency in statistical significance of factors, as well as in the constraint hierarchies for each age group, indicates remarkable stability across the population of *be like* users, suggesting that *be like* is firmly entrenched in their community grammar. Observe also that this form is used well into adulthood (the 30-year-olds), although Figure 2 clearly corroborates earlier research heralding that this as an under-40s phenomenon (Ferrara & Bell, 1995:286).

Going beyond these initial observations, let us probe the constraints in more detail to increase our understanding of the underlying system. First, we consider morphology/tense. Earlier research, relying largely on HP contexts, suggested that present tense constrains the use of *be like* (Blyth et al., 1990; Singler, 2001). Blyth et al. (1990:218) interpreted this to suggest that *be like* may serve an evaluative function, heightening the dramatic effect of complicating action sequences. Singler (2001:273–274), on the other hand, proposed that the HP and *be like*

TABLE 2. *Contribution of external and internal factors on the use of be like by age in Toronto*

Input	9–14 years			15–16 years			17–19 years			20–29 years			30–39 years		
	FW	%	N	FW	%	N	FW	%	N	FW	%	N	FW	%	N
Total N	600			505			1,992			1,138			524		
Morphology/Tense															
HP	.66	77	237	.70	90	144	.67	90	708	.73	86	307	.74	53	76
Present	.31	41	111	.28	63	73	.44	75	276	.50	66	111	.68	45	40
Past	.32	47	87	.38	75	108	.32	70	546	.34	58	445	.39	24	250
range	35			42			35			39			35		
Content															
Thought	.70	64	50	.73	66	87	.54	66	408	.55	60	265	.70	39	138
Direct speech	.48	50	479	.44	58	351	.49	67	1,443	.48	57	775	.41	21	327
range	22			29			5			7			29		
Person															
First	.56	65	170	.55	76	163	.55	80	804	.56	69	460	.51	29	232
Third	.47	52	316	.46	70	230	.45	71	905	.44	61	499	.49	27	191
range	9			9			10			12			2		
Sex															
Female	.53	51	423	.55	64	340	.56	72	1,518	.52	59	732	.48	22	346
Male	.44	37	177	.41	46	165	.33	45	473	.47	51	406	.53	28	178
range	9			14			23			5			5		



co-occur because both are “characteristically informal and colloquial.” This assumes that the use of *be like* falls out from the use of an informal stylistic option (D’Arcy, 2004:328), more like a chance coincidence than part of a systematic pattern. However, this prediction does not account for the full range of variability observable in the data. In Tagliamonte and Hudson (1999: cf. 158, 166), for example, the rates of HP do not correlate with the rates of *be like*. A comparison between the British and Canadian narrative data revealed that where the use of the HP is more frequent, the use of *be like* is less frequent, and vice versa. Thus, a more elaborated test is necessary to capture the regular tendency of *be like* to appear in the present tense (see also D’Arcy, 2004:329).

In Table 2, where a clear distinction is made between present tense and HP, *be like* is not favored for present tense for any of the speakers under 20. In contrast, among the 30-year-olds, *be like* is favored for *both* the HP and present tense morphology; in other words, all cases of present tense. The 20-year-olds are hovering in between. Here, the overarching correlation with present morphology does not exist. In fact, present tense contexts are precisely on the median, at .50. In short, the factor group morphology/tense splits the population. The present tense effect has shifted towards favoring for the 20–29-year-olds and is the prevailing effect among the 30–39-year-olds, but for the younger speakers, it is clearly HP.

Next, we consider the content of the quote. While the correlation with inner thought is stable across the board, its strength differs depending on speaker age. It stands as a clear second ranked constraint for the 9–14 and 15–16 years olds, with ranges of 22 and 29, respectively. It is also one of the top-ranked constraints among the 30-year-olds. In fact, with a range of 29, its strength is nearly equal to that of morphology/tense. But, among the 17–19-year-olds and the 20–29-year-olds its effect is very weak. We return to this observation later when we discuss the possible grammar-internal mechanism behind the rise of *be like*.

There is little to add about the first versus third person constraint, except to note that it is quite weak, particularly among the 30-year-olds, who barely have it.

Finally, we consider the sex effect. It peaks in strength among the 17–19-year-olds, where it is second only to the strong correlation with HP. Interestingly, among the 30–39-year-olds, it operates in the opposite direction, albeit weakly; males, not females, favor *be like*.

In sum, despite the apparent stability of *be like* in terms of the significance of factors, the detailed evidence from the constraint rankings and their relative strength shows us that there are subtle changes going on in the system. At this point, a model of what the underlying mechanism of this system is might be needed. Recall that based on age and frequency, we hypothesize that the 30-year-olds were the first generation of *be like* users. If so, then the speakers in their 20s represent the second generation. These age groups can be positioned in the model as Stage 1 and Stage 2. Labov’s (1994, 2001:447) account of sound change sets the age of stabilization for *phonological* features at about 17. In these data, not only do the 17–19-year-olds have the highest frequency for *be like*, but they are

TABLE 3. *Contribution of external and internal factors on the use of be like in Toronto English: Three stages [c. 2004]*

	Stage 1 30–39 years	Stage 2 20–29 years	Stage 3 17–19 years
Corrected mean	.31	.72	.82
Total <i>N</i>	524	1,138	1,992
	FW	FW	FW
<hr/>			
Tense			
HP	.74	.73	.67
Present	.68	.50	.44
Past	.39	.34	.32
range	35	39	35
Content			
Thought	.70	.55	.54
Direct speech	.41	.48	.49
range	29	7	5
Person			
First	.51	.56	.55
Third	.49	.44	.45
range	2	12	10
Sex			
Female	.48	.52	.56
Male	.53	.47	.33
range	5	5	23

also the group where the sex effect is strongest, an indication that this feature has developed vital social differentiation. We therefore take them to represent Stage 3.

Table 3 reconfigures the results from Table 2, grouping the data into the three developmental stages that we have just outlined.

At Stage 1, *be like* introduces inner thought. In fact, most studies concur that *be like* entered the quotative system carrying with it this strong pragmatic correlation. Our evidence supports this perspective because the content of the quote is strongest among the 30-year-olds, the early adopters of *be like* in the 1980s. If so, then we would expect the pragmatic constraint to be salient for them, and it is.

The correlation of *be like* with first person is likely a by-product of the primordial semantic/pragmatic effect, narrators quoting themselves, that is, first person. Our evidence in support of this position is that the grammatical person constraint reaches statistical significance for every age group. It too is present from the beginning, and is somewhat stronger at Stage 2 and beyond.

As far as the sex effect is concerned, we believe that it must have started out quite minor. This may explain why early studies were inconsistent in their findings for sex (cf. Blyth et al., 1990; Dailey-O'Cain, 2000; Ferrara & Bell, 1995; Tagliamonte & Hudson, 1999) and makes us unwilling to attribute much importance in the marginal effect for males that shows up here among the 30-year-olds. Furthermore, as *be like* rises in frequency, it begins to develop a correlation with

Constraints	Stage 1 (30-year-olds)	Stage 2 (20-year-olds)	Stage 3 (17–19-year-olds)
Content	Inner thought favors	Leveling trend visible	Leveling of the content constraint
Person	1st persons favor	1st persons favor	1st persons continue to favor
Morphology/ Tense	Present tense favors	HP, narrative breaks away	HP highly favors
Sex	Minimal sex effect	Sex effect still weak, but females take lead	Strong female lead

FIGURE 3. Developmental trajectory for *be like*.

female speakers, until this becomes one of the strongest constraints on its use. This is precisely what we would predict, if, as Labov (2001:308) suggested, sociolinguistic variation is parasitic on foundational linguistic variation. Indeed, the pattern observed for the sex effect falls directly in line with Labov’s (2001:307–308) discussion regarding the development of gender asymmetry. Once a change becomes associated with a particular gender (i.e. Stage 2 in our model), what Labov calls gender specialization, men either retreat from or resist the change (i.e. Stage 3 in our model), causing a gender split.

However, it is the association of *be like* with morphology/tense which presents the most interesting nuance for the incrementation problem. The evidence from this constraint suggests that *be like* is associated with the HP, as pointed out by Singler (2001:272–273), although this observation was made without the added perspective of present tense more generally. Discussed earlier, *be like* may be lexicalized with present tense morphology regardless of orientation in time. Our data suggest that developmentally, a distinction between morphology and temporal reference is critical. *Be like* initially carries present tense morphology, but this more general pattern is localized to the 30-year-olds in Toronto, that is, Stage 1. At Stage 2, however, the correlation with the simple present weakens and the HP begins to break away until it stands apart as *the* favoring factor at .67. This developmental path for *be like* is summarized in Figure 3.

DISCUSSION

We have presented an account of how *be like* has developed in Toronto based on a series of regularities interpreted in a sociocultural context. It now remains to offer our explanation as to why these changes have taken place. The benefit of data from many generations within the same community provides insight into the elusive actuation problem (Weinreich et al., 1968). Recall Figure 2, which displays the frequency of quotatives across all age groups. This trajectory looks very much like weak complementarity (Sankoff & Thibault, 1981; Sankoff, 1982), with *be like* replacing *say*. However, the results reported in Tables 2 and 3 revealed

TABLE 4. *Factors on the use of be like in speakers under 40 years old and say in speakers over 40 years old*

Corrected mean	<40 years old			>40 years old		
	<i>be like</i>			<i>say</i>		
Total <i>N</i>	4,759			1,605		
	FW	%	<i>N</i>	FW	%	<i>N</i>
Tense						
HP	.70	85	1,472	.33	75	199
Present	.43	64	611	.47	67	94
Past	.32	57	1,436	.54	80	871
range	.38			.21		
Content						
Thought	.59	60	948	.06	13	261
Direct speech	.48	57	3,375	.65	75	1,228
range	.11			.59		
Person						
First	.55	69	1,829	.44	66	621
Third	.46	62	2,141	.55	85	683
range	.09			.11		
Sex						
Female	.54	60	3,359	.51	65	878
Male	.40	44	1,400	.48	62	727
range	.14			.03		

Note. Shading denotes favoring factor groups.

subtle shifts in the constraint rankings and relative strength of factors across age groups that hint at reorganization of the system. We therefore hypothesize that *be like* is more than simply a new lexical item within the English quotative cohort. If this were true, then we would expect that the constraints on *be like*, which seem to be pervasive across all the speakers who use it, will differ from the underlying organization of the system that existed previously.

Table 4 shows two independent multivariate analyses of our corpus, split at the watershed: the under 40 year-olds, in contrast to the over 40 year-olds.<sup>6</sup> The two sets of results could not be more different; *say* has a completely different profile than *be like*. It is favored in past and present tense (not HP), for direct speech (not thought), and for third person (not first). The only similarity across age groups is that the women of both populations are more likely to use these forms. The apparent time picture made it appear like the substitution of one form for another, but it is not. *Be like* is not a replacement but an innovation in a restructured quotative system.<sup>7</sup>

The pointed question then becomes, “What happened that would have led to this remarkable exchange of form and reorganization of function?” Among the constraints that figure in the development of *be like*, it is the pragmatic effect contrasting thought with direct speech that is most baffling. When did narrators

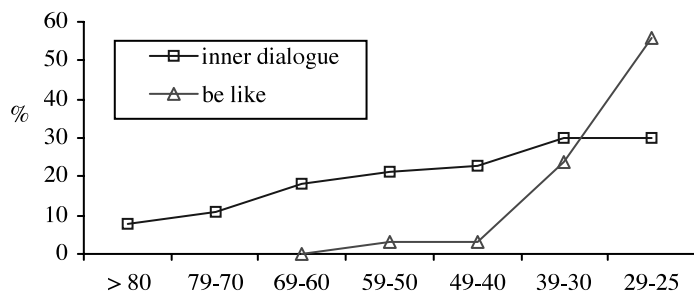


FIGURE 4. Proportion of inner dialogue and *be like* across the generations in Toronto English.

begin telling stories by regaling their audience to a running stream of their own inner thought processes? Consider the examples in (6), typical passages from the under 20 year-olds in our corpus,<sup>8</sup> those at Stage 3 in the development of *be like*.

- (6) a. I'm like, "Oh my God!"  
 I'm like, "He's going to get wet, I give it ten seconds."  
 ... sure enough, Ø "Boom!"  
 I was just like, "Oh man, you guys are so predictable." (I!/f/19)
- b. I'm like, "Shit, my parents are going to wake up 'cause of that."  
 I'm like, "Oh shit."  
 I'm like, "Okay, I gotta pretend like something happened, think something, think something," right? (S/W/f/17)

In fact, according to our data, this is quite a recent phenomenon and one that has undergone a dramatic increase over the past 65 years.

Figure 4 plots the overall proportion of quotations that are internal dialogue, beginning with the oldest speakers in our sample and tracking them down to the early adopters of *be like*. Among speakers in their 80s, quoted thought represents just 8% of all direct quotations (*N* = 363). However, this proportion increases steadily in apparent time and peaks among the 30-year-olds. In other words, Figure 4 captures a change in how narrators construct their stories. Notably, this development immediately precedes the acceleration of *be like* among the 30-year-olds. The explanation we would like to offer is that the rising stylistic option of inner monologue in narratives of personal experience was expanding before *be like* entered the system, setting the scene for a new development. Thus, *be like* filled a niche that already existed, only later specializing into a device to grandstand the narrator as a participant in his or her own stories.<sup>9</sup> From there, the path was laid for it to develop into a narrative present marker in its own right.<sup>10</sup>

Interestingly, this developmental reshuffling may be reconstructed from the multivariate results in Table 2. As we have noted, *be like* is associated with first persons from the outset, Stage 1, in the 30–39-year-olds. Its specialization as a narrative marker occurs at Stages 2 and 3, among the 20–29-year-olds and the 17–19-year-olds. These are also the two groups where the content of the quote

constraint is most weakened. In Stage 1, this factor is the second ranked constraint on *be like*. In Stage 2, where the effect of the present tense first weakens, the relative weight of content of the quote drops to third. In Stage 3, where the present tense begins to disfavor, the content constraint again drops in the rankings, this time becoming the weakest of the four. It is only among the subsequent generations—the 9–16-year-olds—that the variable grammar stabilizes and the effect of content of the quote is again eminent in the system. We suggest that these correspondences between the factor groups morphology/tense and the content of the quote are not coincidental. We propose the following scenario. *Be like* entered the quotative paradigm marking morphologically present internal dialogue. This is consistent with the view that it was filling a niche in the grammar, that is, speaker thought processing. Crucially, *be like* was not restricted to this context, nor is it now. However, once it had a foothold in the system it was then free to specialize. This is where the rising stylistic option of internal dialogue becomes pivotal in the further development of *be like*. A defining property of the narrative is that it retells some event that occurred in the past, yet many complicating action clauses are encoded with the HP (Schiffrin, 1981; Wolfson, 1978, 1979), present tense morphology to refer to past actions and, more and more frequently, past thoughts about those actions, as well. Thrust into narrative by its initial association with monologue, *be like* then latches on to the archetypal narrative device: the HP. During this period, the correlation with internal dialogue weakens due to the increasing importance of the HP. However, once entrenched as a narrative marker, the correlation of *be like* with internal dialogue, always present, again comes to the fore.

The issue remains, however, as to what catapulted *be like* into the adolescent vernaculars of the 1980s. Labov (2001:462) observed that the “acceleration of linguistic change logically begins when the incipient change is attached to or is associated with a particular style or social group: A social category like *burnout* or a neighborhood like *Kensington*.” In this case, we would like to suggest that the associated social category was *Valley Girl* and the place was *California*. Indeed, such an association is part of the received wisdom surrounding *be like* (see, e.g., Blyth et al., 1990:224; Dailey-O’Cain, 2000:76). As part of the “preppie” movement of the 1980s, *be like* gained prestige as a trendy and socially desirable way to voice a speaker’s inner experience. In other words, linguistic change begins with a hospitable grammatical environment, but requires a social force to drive it forward.

## CONCLUSION

A recurrent topic in the literature is whether *be like* represents age-grading or generational linguistic change (e.g., Cukor-Avila, 2002; Ferrara & Bell, 1995; Tagliamonte & D’Arcy, 2004). In Toronto, the adolescent peak makes it appear as though *be like* is age-graded. However, such a peak is predicted by the logistic model of linguistic change proposed by Labov (2001). Indeed, the perspective from the community as a whole reveals the classic S-curve of linguistic change. Interpreting the crossroads in the contemporary population of Toronto is made

more complex by the fact that *be like* was only incipient in the system prior to the 1980s. Speakers over the age of 40 never had it in the first place because the point at which their grammar stabilized either predated or coincided with the introduction of *be like* to the quotative repertoire. The 30-year-olds are on the frontier, and from that point onward there is ongoing generational change. There is simply too much stability across the population for *be like* to be a passing fancy. However, is there evidence for communal change as well?

The central assumption of the apparent time construct is that once out of their teens, speakers no longer participate in the advancement of change (Labov, 2001:454). The relevant implication here is that the 30-year-olds have continued to use *be like* at the same rate throughout their lifetimes, after initially adopting it in the 1980s. However, there are indications that these frontliners have also been participants in ongoing change. Recall that the 30-year-olds are using *be like* a full third of the time. It is unlikely that they had this rate in their late teens. Instead, it is more probable that they have increased their use over the last fifteen years. Support for this conjecture comes from a comparison with data from Canadian English in 1995 in which the frequency of *be like* was 13% (Tagliamonte & Hudson, 1999). These speakers would now be between the ages of 25 and 34. If we compare the 1995 rate of *be like* with these two age groups in our corpus, their frequencies of *be like* are 58% and 31%, respectively. This is well beyond the 13% of 7 years ago.<sup>11</sup> In other words, the evidence suggests a profile of communal change, with speakers increasing their use of *be like* throughout their lifetime.

Discussed at the outset, a growing body of work suggests that individuals do not cease to participate in linguistic change post-adolescence, but may indeed contribute to its advancement into adulthood (e.g., Nevalainen & Raumolin-Brunberg, 2003; Nordberg & Sundgren, 1998). The Toronto results thus lend support to Labov's (2001:447) observation that adults may not be as stable as previously thought. Yet the 30-year-olds, while evidently augmenting the *rate* at which they use *be like* since their adolescence, do not exhibit the advanced stages of development with regard to the constraints (Table 2). Instead, it appears that speakers can enhance their use of an incoming feature by increasing its frequency, but they do not show signs of the more advanced state of the variable grammar. In other words, it seems that what it means to 'participate' in a linguistic change requires elaboration. Is frequency sufficient, or does participation in a change require the full composite of constraints? Our findings suggest that adult frequencies of linguistic forms are labile (Labov 2001:447), but the grammar underlying them is not. This observation provides us with a tantalizing starting point for future research.

#### NOTES

1. Examples from the Toronto corpora, housed in the Sociolinguistics Laboratory at the University of Toronto, are identified by corpus, followed by the individual speaker's code, sex, and age. The corpora are coded as follows: 2 = ROP 2002, 3 = ROP 2003, I = IN-TO-VATION 2003, N = IN-TO-VATION 2004, S = Story-telling 2002, \$ = Story-telling 2003.

2. For discussion of the sampling methodology for the Toronto sociolinguistic projects, see Tagliamonte and D'Arcy (2004:497), Tagliamonte (2006), and Tagliamonte (2003–2005).

3. Tagliamonte and D'Arcy (2004) focused on 10–19-year-olds only and at that time, the data comprised 44 speakers and a total of 2,058 tokens. In contrast, this age group in the current analysis is represented by a substantially larger data set: 87 speakers and 3,097 tokens. The heightened numbers in the 17–29-year-old age group is due to the inclusion of two corpora which were collected independently as part of course work for HUM199Y, *The Linguistics of Story-telling*, in 2002 and 2003. These data represent considerable speaker numbers, but smaller total tokens per speaker.

4. In Table 2, the number of tokens within each factor group does not add up to the total number of tokens for each age group. Due to our focus on the major constraints on *be like*, we have abstracted away from a number of other nuances in the system, necessitating a number of exclusions. First, within the content of the quote factor group, all hypothetical, written, internet dialogue, gestures, and nonlexicalized sounds have been excluded. Second, within the grammatical person factor group all tokens of existential *there*, zero, and second person singular subjects were excluded. Third, within the morphology/tense factor group, less frequent tense and aspectual constructions were excluded (e.g., progressives, habituais, futures, etc.). The influence of all these categories on the use of *be like* certainly bears investigation. There is a growing range of quotation types, and a contrast between full noun phrases (NPs) and pronouns has been in evidence throughout our research. To facilitate ongoing study we have included a comparison of marginals of these distributions in Tables 5–7 in the Appendix. We also note that co-occurring adverbs (particularly *just*, but also *all*, *so*, etc.) overwhelmingly occur with *be like*. However, all such adverbs combined represent only 3.7% of all tokens ( $N = 235/6364$ ) and only 8.8% of all *be like* ( $N = 234/2670$ ) tokens in these data.

5. As far as we are aware, the only time this direction of effect has not appeared is among the 17–19-year-old cohort in Tagliamonte and D'Arcy (2004), in which a moderate effect in the opposite direction was in evidence. Furthermore, it must be noted that Table 4 in Tagliamonte and D'Arcy (2004:504) contains an error. There were 667 tokens of direct speech and 214 tokens of internal dialogue, not the other way around, as reported.

6. Recall that the nonapplication contexts in both runs are virtually the same (see Figure 2).

7. It is not simply that *be like* appears where *say* is disfavored. An obvious contradiction to this is the effect of speaker sex, since both *be like* and *say* are (significantly) disfavored among males. It is important to bear in mind that these forms are just two of numerous quotative choices, each of which possesses its own profile with respect to conditions on use. Within the repertoire, *say* is not unique in having constraints which pattern in opposition to those on *be like*. However, Table 4 contrasts these two forms in order to demonstrate that despite their weak complementarity in apparent time (Figure 2), *be like* is not replacing *say*. Thus, even though *say* has receded dramatically in these data, *be like* is not stepping into the niche that *say* is leaving behind. In other words, the rise of *be like* is not due to lexical replacement.

8. An interesting aspect of *be like* is its perceived tendency to encode routinized expressions (e.g., *Oh my God*, *okay*, *oh*, *whatever*) in putatively “performative” dialogue. However, these “highly conventionalized exclamation[s],” which have been a part of the *be like* repertoire since its inception and characterize an early stage in its use (Ferrara & Bell, 1995:283), account for a negligible proportion of instances of *be like* overall (6.5%;  $N = 2,670$ ). In fact, the particular collocation of *be like* + *Oh my God*—without further elaboration within the quotation frame, as in (6a)—comprises just 1% ( $28/2,670$ ) of the data for *be like*.

9. We wonder why the extant form *think* was not recycled for the rising option of quoting inner monologue. As Figure 2 demonstrates, however, *think* has always been a marginal contender in the quotative system, at least in Toronto. We leave this issue for future research.

10. Such qualitative changes may simply reflect the fact that people of different ages tell stories in diverse ways. This notwithstanding, if the ongoing change in the rates of inner monologue were simply age-related behavior then we might not expect the regular, step-wise progression seen in Figure 4. That is, even though apparent time distributions are not sufficient for disentangling age-grading and generational change (Labov, 1994:84), age-graded phenomena are typically associated with a distinct developmental stage (Chambers, 2003:206), most often adolescence, and as far as we are aware, are not generally thought of as gradient, spanning the full life cycle. Moreover, prototypical age-grading pertains to aspects of language that involve a high degree of social awareness and as such, would be “more readily available for conscious manipulation” (Milroy & Gordon, 2003:36–37). This is not the profile of a feature like inner monologue. Nevertheless, there is much that could be done with a more qualitative analysis of narrative structure across the life span.

11. Interestingly, in a study conducted by students of Shana Poplack in Ottawa in the autumn of 2004, the 18–35-year-old age group had a rate of *be like* of 41% ( $N = 1,607/3,940$ ) (Dion & Poplack, 2005), a remarkably comparable result to our 25–34-year-olds.



## REFERENCES

- Bailey, Charles-James. (1973). *Variation and linguistic theory*. Washington, DC: Center for Applied Linguistics.
- Blyth, Carl, Recktenwald, Sigrid, & Wang, Jenny. (1990). *I'm like*, "say what?!": A new quotative in American oral narrative. *American Speech* 65:215–227.
- Buchstaller, Isa. (2004). *The sociolinguistic constraints on the quotative system: British English and US English compared*. Doctoral dissertation, University of Edinburgh.
- Butters, Ronald. (1982). Editor's note [on 'be + like']. *American Speech* 57:149.
- Chambers, J. K. (2003). *Sociolinguistic theory* (2nd ed.). Oxford: Blackwell.
- Chambers, J. K., & Trudgill, Peter. (1998). *Dialectology*. New York: Cambridge University Press.
- Cukor-Avila, Patricia. (2002). *She say, She go, She be like*: Verbs of quotation over time in African American Vernacular English. *American Speech* 77:3–31.
- Dailey-O'Cain, Jennifer. (2000). The sociolinguistic distribution of and attitudes toward focuser *like* and quotative *like*. *Journal of Sociolinguistics* 4:60–80.
- D'Arcy, Alex. (2004). Contextualizing St. John's Youth English within the Canadian quotative system. *Journal of English Linguistics* 32:323–345.
- Dion, Nathalie, & Poplack, Shana. (2005). "I'm like: 'I know, I do the same thing'": Can minority speakers keep up with the Joneses? Paper presented at NWAVE-34, New York University.
- Ferrara, Kathleen, & Bell, Barbara. (1995). Sociolinguistic variation and discourse function of constructed dialogue introducers: The case of be+like. *American Speech* 70:265–289.
- Labov, William. (1994). *Principles of linguistic change: Internal factors*. Oxford: Basil Blackwell.
- \_\_\_\_\_. (2001). *Principles of linguistic change: Social factors*. Oxford: Blackwell.
- Milroy, Lesley, & Gordon, Matthew. (2003). *Sociolinguistics: Method and interpretation*. Oxford: Blackwell.
- Nevalainen, Terttu, & Raumolin-Brunberg, Helena. (2003). *Historical sociolinguistics: Language change in Tudor and Stuart England*. Harlow, UK: Pearson Education Limited.
- Nordberg, Bengt, & Sundgren, Eva. (1998). On observing real-time language change: A Swedish case study. FUMS report no. 190. Uppsala: Enheten för sociolinguistik, Institutionen för Nordiska Språk vid Uppsala Universitet.
- Romaine, Susanne, & Lange, Deborah. (1991). The use of *like* as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech* 66:227–279.
- Sankoff, David. (1982). Sociolinguistic method and linguistic theory. In L. Cohen, J. Los, H. Pfeiffer, & K. Podewski (eds.), *Logic, methodology, philosophy of science VI*. Amsterdam: North Holland; and Warsaw: Polish Scientific.
- Sankoff, David, & Thibault, Pierette. (1981). Weak complementarity: Tense and aspect in Montreal French. In B. Johns & D. Strong (eds.), *Syntactic change*. Ann Arbor: University of Michigan Press.
- Schiffrin, Deborah. (1981). Tense variation in narrative. *Language* 57:45–62.
- Singler, John. (2001). Why you can't do a VARBRUL study of quotatives and what such a study can show us. *University of Pennsylvania Working Papers in Linguistics* 7:257–278.
- Tagliamonte, Sali A. (2003–2005). *Linguistic changes in Canada entering the 21st century*. SSHRC grant no. 410-2003-0005.
- \_\_\_\_\_. (2006). "So cool, right?": Canadian English entering the 21st century." In P. Avery, A. D'Arcy, E. Gold, & K. Rice (eds.), *Canadian English in a Global Context. Theme issue of Canadian Journal of Linguistics*.
- Tagliamonte, Sali A., & D'Arcy, Alex. (2004). "He's like, she's like": The quotative system in Canadian youth. *Journal of Sociolinguistics* 8:493–514.
- Tagliamonte, Sali A., & Hudson, Rachel. (1999). Be like et al. beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics* 3:147–172.
- Tannen, Deborah. (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In F. Coulmas (ed.), *Direct and indirect speech*. Amsterdam: de Gruyter. 311–332.
- Weinreich, Uriel, Labov, William, & Herzog, Marvin. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (eds.), *Directions for historical linguistics*. Austin: University of Texas Press. 97–188.
- Wolfson, Neesa. (1978). A feature of performed narrative: The conversational historical present. *Language in Society* 7:215–237.
- \_\_\_\_\_. (1979). The conversational historical present alternation. *Language* 55:168–182.

## APPENDIX

TABLE 5. *Distribution of quotatives by content of the quote*

	<i>be like</i>		<i>say</i>		<i>think</i>		<i>go</i>		<i>zero</i>		<i>misc.</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
Internal dialogue	583	48	46	4	272	23	46	4	193	16	68	5
Direct speech	1937	42	1596	35	–	–	293	6	576	13	196	4
Sound	39	36	2	2	–	–	32	30	25	23	10	9
Writing	14	14	47	49	–	–	2	2	26	27	8	8
Gesture	4	44	1	11	–	–	3	33	1	11	0	0
Hypothetical	76	25	96	32	5	2	16	5	92	31	16	5
Internet dialogue	17	46	14	38	–	–	2	5	3	8	2	3
Total	2670	42	1802	28	277	4	394	6	916	14	300	5

Note. 0 = no data, – = n/a.

TABLE 6. *Distribution of quotatives by grammatical person*

	<i>be like</i>		<i>say</i>		<i>think</i>		<i>go</i>		<i>zero</i>		<i>misc.</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
Ø	0	0	2	0	0	0	0	0	916	100	0	0
1st person singular	1103	50	646	29	215	10	119	5	–	–	116	5
2nd person singular	53	31	57	34	19	11	19	11	–	–	22	13
3rd person singular	816	50	600	36	8	1	155	9	–	–	69	4
1st person plural	177	70	41	16	10	4	14	6	–	–	10	4
3rd person plural	213	51	152	37	6	1	20	5	–	–	24	6
Indefinite pronoun	18	29	31	50	0	0	4	7	–	–	9	15
Referential <i>it</i>	11	33	16	49	0	0	6	18	–	–	0	0
Existential <i>there</i>	1	50	–	–	–	–	–	–	–	–	1	50
NP singular	206	42	200	41	7	1	38	8	–	–	37	8
NP plural	72	42	57	33	12	7	19	11	–	–	12	7
Total	2670	42	1802	28	277	4	394	6	916	14	300	5

Note. 0 = no data, – = n/a.

TABLE 7. *Distribution of quotatives by tense and aspect*

	<i>be like</i>		<i>say</i>		<i>think</i>		<i>go</i>		<i>zero</i>		<i>misc.</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
N/A	–	–	–	–	–	–	–	–	906	100	1	0
Simple past	839	36	1120	49	163	7	46	2	–	0	136	6
Unambiguous simple present with past temporal reference	1261	76	210	13	5	0	150	9	–	0	41	3
Unambiguous simple present with present temporal reference	401	56	166	23	32	4	78	11	4	1	41	6
Future, <i>be going to</i>	7	37	6	32	0	0	4	21	0	0	2	11
Future, <i>will</i>	26	37	32	46	3	4	4	6	0	0	5	7
Past progressive	1	2	24	44	12	22	4	7	–	–	13	24
Present progressive	0	0	25	25	36	35	28	28	–	–	13	13
Past habitual	0	0	10	77	0	0	0	0	0	0	3	23
Habitual <i>would</i>	56	42	50	38	2	2	17	13	1	1	6	5
Modal	56	38	53	36	4	3	16	11	3	2	15	10
Infinitive, <i>to</i>	12	21	29	52	1	2	7	13	–	–	7	13
Participle, <i>-ing</i>	4	3	49	42	17	15	38	33	–	–	6	7
Other	7	14	28	54	2	4	2	4	2	4	11	21
Total	2670	42	1802	28	277	4	394	6	916	14	300	5

Note. 0 = no data, – = n/a.