Is one innovation enough? Leaders, covariation and language change[*]

Cathleen Waters

University of Leicester, UK

Sali A. Tagliamonte

University of Toronto, Cda

Abstract:

Do the people who lead in one linguistic change, lead in others? Previous work
has suggested that they do not, but the topic has not been addressed extensively
with non-phonological, spoken data. In this paper, we answer this question
through an examination of lexical, morphosyntactic and discourse-pragmatic
changes in progress in Canadian English as spoken in the largest urban center of
the country, Toronto. Close scrutiny of the behavior of individuals across multiple
linguistic variables, i.e. covariation, and using the Pearson product-moment
correlation coefficient enables us to test the use of incoming variants both by the
community of speakers as a whole and by those who are leading change. The
innovative variants of quotatives, i.e. *be like*, intensifiers, i.e. *really, so,* deontic
modality, i.e. *have to*, stative possession, i.e. *have* and general extenders, i.e. *and
stuff,* demonstrate that the leaders of these multiple linguistic changes have
common social characteristics (e.g. women lead more than one change), but it is
not the case that any one individual in a community will be at the forefront of
more than one change.

The study of COVARIATION, the degree of consistency of behavior by individuals across multiple linguistic variables, has received increasing interest in the past twenty years. Covariation has been examined in a range of languages and varieties, with synchronic studies tending to focus primarily, though not exclusively, on phonetic or phonological variation and change (Chambers 2009; Guy 2013; Labov 2001; Maclagan et al. 1999; Oushiro 2016; Oushiro and Guy 2015; Stuart-Smith and Timmins 2010; Thorburn 2014; and contributions in Hinskens and Guy 2016). Covariation across multiple morphosyntactic and lexical variables, including change in progress, has been examined more extensively in historical, diachronic work (Nevalainen and Raumolin-Brunberg 2003; Nevalainen et al. 2011) than in contemporary varieties.

The goal of this study, therefore, was to examine lexical, morphosyntactic and discourse-pragmatic (henceforth, nonphonological) changes in progress with a synchronic corpus of a variety of contemporary English. The data comprises vernacular, spoken language and innovative uses of quotatives, intensification, general extenders, stative possession and deontic modality. We explored covariation in the use of incoming variants in the community as a whole, and then focused more closely on those who are leading change, to determine if the leaders of one change are the leaders of multiple changes. We assessed leadership in a change based on the frequency of incoming variants out of the total number of variants used for the same meaning. As we will demonstrate, our findings are consistent with historical studies of nonphonological changes and contemporary phonological changes, namely, that there is little correlation among the innovative

2

forms in the community as a whole, and even less correlation among the leaders. Drawing on previous quantitative studies of covariation and our own results and analysis we will argue that, although the leaders of these multiple nonphonological linguistic changes may have common social characteristics (e.g. women may lead change), any one individual is unlikely to be at the forefront of more than one change.

The paper is organized as follows. In the next section, we discuss stages of change and categories of individuals during change, including some ideas from diffusion theory[i] (Rogers 2003) that were central to our conceptualization of leaders. We then present an overview of the literature on covariation for changes in progress. The subsequent section describes our data, the linguistic variables in detail, and our methodology, both for determining who we categorized as leaders and for the correlation calculations. The presentation of our results, first for all speakers in the study, and then only for those we identified as leaders, follows. We demonstrate extremely limited correlation in the use of more than one incoming variant by individuals, and the correlations that we do find resist straightforward explanation, leading us to conclude that the evidence from linguistic changes beyond phonology in contemporary, spoken language also shows that leaders of one change are unlikely to be the leaders of others.

BACKGROUND

STAGES OF CHANGE. Linguistic change has been described as having five stages: incipient, new and vigorous, mid-range, nearing completion, and

3

completed (Labov 1994, 63-64, 83-84; Nevalainen and Raumolin-Brunberg 2003, 55). Similarly, and drawing from diffusion theory (Rogers 2003, 22, 267-299), speakers have been described as belonging to one of five groups: innovators, early adopters, early majority, late majority, and laggards (Stuart-Smith and Timmins 2010, 43-51). However, as other linguists using the concepts of diffusion theory have noted (Milroy 1992, 183-184; Stuart-Smith and Timmins 2010, 45), innovations in language are different from some other types of change that diffusion theory seeks to explain, such as the implementation of a cultural or industrial practice. As diffusion theory focuses on a binary contrast between "*adoption*, a decision to make full use of an innovation as the best course of action available, or *rejection*, a decision not to adopt an innovation" (italics original, Rogers 2003, 21), diffusion theory is an oversimplification for linguistic change, which is characterized by a (sometimes lengthy) period of variability during which both an older and a newer form coexist (Hopper 1991, 22).

Categorization of individual linguistic behavior, therefore, requires a relative measure of frequency (Nevalainen et al. 2011, 5); it is necessary to consider the individual with respect to the rest of population at the same point in time. Thus, a study of how much an individual participates in an on-going change must first determine some measure relevant for study. A linguistic phenomenon undergoing change presents an ideal litmus test for such an endeavor since it can be counted and compared (e.g., second formant of a vowel, proportion of innovative to conservative features, etc.). Individuals are then assigned to categories based on these relative criteria, resulting in divisions such as leader

versus laggard, high versus low rate of use, etc. (Guy 2013, 70; Maclagan et al. 1999, 33; Nevalainen et al. 2011, 26; Stuart-Smith and Timmins 2010, 49-50; Thorburn 2014, 258). Although five stages of change were listed above, in practice, the categorization of linguistic behavior is usually operationalized into a ternary system such as progressive/in-between/conservative (e.g., Nevalainen et al. 2011), high/middling/low (Guy 2013) or innovative/neutral/conservative (e.g., Maclagan et al. 1999). Even with a less granular scale, it is not always a straightforward process to assign an individual to a group (Stuart-Smith and Timmins 2010, 53). Similar to patterns observed for other social and linguistic polarities (e.g., formal/informal), the patterns of innovative use by speakers participating in a linguistic change fall on a continuum from innovative to conservative with few speakers at either pole (Labov 2001, 463; Nevalainen et al. 2011, 1-2). For our purposes, we established a binary division of leaders versus all others. We describe the process for that categorization in the Data, Variables and Methods section below.

COVARIATION OF INNOVATIVE FORMS. There is some limited evidence for covariation of incoming forms with phonological change.[ii] For instance, Labov (2001, 371-373) calculated correlations by using Pearson's correlation coefficient (a.k.a. Pearson's *r*), a measurement of linear relationship between two variables, for 86 individuals in Philadelphia focusing on phonological changes in progress. Some changes were highly correlated with each other (Pearson's *r* greater than 0.6) while other changes showed no significant correlations.

Similarly, in a study of phonological change in New Zealand English, Maclagan et al. (1999) contrasted the behavior of over two hundred individuals across five phonological features (three front vowels and two diphthongs), and examined the behavior of the leaders of change more closely. Across their data, some individuals were consistently innovative, some were consistently conservative, and some individuals exhibited conservative behavior for one set of features and innovative behavior for the other (Maclagan et al. 1999, 34).

Thus, an individual's use of an innovative (or conservative) variant for one phonological change in a community may, or may not, be correlated with that speaker's use of other innovative (or conservative) phonological forms. As Guy (2013, 64) observed, the existence of phenomena such as chain shifts implies that correlation of incoming forms must be possible in some circumstances, and thus some covariation may be structural. However, even when phonological changes covary in a community, not all phonological changes are correlated (Labov 2001, 372f), and the behavior of individuals, including leaders of change, is not always consistent across phonological variables (Guy 2013, 70; Maclagan et al. 1999, 33-34; Oushiro and Guy 2015, 165; Stuart-Smith and Timmins 2010, 49-50; Thorburn 2014, 263).

As is the case in variationist work more generally, there have been more studies of phonological covariation than covariation at other levels of the grammar, though some previous work on the leaders of nonphonological change has been undertaken. Nevalainen et al. (2011) used historical written data to explore real-time morphosyntactic change in English through an investigation of

six changes that are now completed, such as the shift in second person subject pronoun *ye* to *you* and the change in the third person singular suffix *-th* to *-s* (*goeth* to *goes*). Nevalainen et al. (2011, 26) sought to "focus on all the six changes simultaneously and ask whether it is possible to find people who are linguistically more progressive or conservative than others" based on data from letters written from 1410 to 1681 (the Corpus of Early English Correspondence).[iii] Through the method of repeated sampling with replacement to estimate frequency of use, they calculated average usage of a variant for a given time period (Nevalainen et al. 2011, 14ff.). They then determined the leaders of change by comparing the rates of individuals to the averages (ibid). For a subset of their data, the period of 1500 to 1619, Nevalainen et al. (2011, 36) examined the behavior of individuals across real-time and uncovered limited consistency across variables in general. Nevalainen et al. (2011, 25-27) found no speaker who they categorized as "progressive" in all changes, and they further noted that some speakers who were categorized as progressive for some changes were conservative for others. Nevalainen et al. (2011, 29-30) did identify two individuals that they categorized as being leaders more generally, by virtue of their higher rates of innovative forms in five of the six variables. As they note, however, one of these individuals was Elizabeth I, who "most likely promoted the diffusion of the variants she adopted" (Nevalainen et al. 2001, 29).

In summary, previous investigation into the correlation of individual behavior across variables has revealed some instances of covariation, however individual behavior, particularly with respect to nonphonological change remains

in question. Indeed, there are few studies that combine contemporary, spoken data and changes beyond phonology. The combination of vernacular speech, non-phonological variables and statistical methods we present in this study enable us to expand the study of covariation into new territory.

DATA, VARIABLES AND METHOD

The data come from the two-million-word Toronto English Archive (henceforth, TEA), a sociolinguistic corpus of vernacular speech that was collected by members of the community from 2002 to 2004 and stratified by social factors (Tagliamonte 2003-6, 2006). Crucially, the corpus has sufficient data to adequately assess variability in multiple linguistic features for many speakers, and in some cases, many leaders. The infrequent nature of mophosyntactic variables necessitated the inclusion of some individuals with a low number of tokens, and we address our strategies for managing this below. However, for each pair of variables, we were able to look at the correlation in behavior for between 30 and 84 individuals (adults) overall, though the number of individuals who could be categorized as leaders, as we discuss below, was dependent on the variable. Further, we specifically chose linguistic variables that had previously been studied in the community because "the behavior of the individual speaker cannot be understood until the sociolinguistic pattern of the community as a whole is delineated" (Labov 1994, 33). This strategy also allowed us to account for linguistic-internal factors in our calculations of individuals' propensity for use of the innovative variants. One trade-off of using this type of large sociolinguistic

corpus, however, is our inability to establish reliable information on indexicality (Eckert 2008) for the variants we examined, and we are therefore cautious in our interpretation of social meaning in our discussion.

We illustrate the five variables we examined below, with a reference to the previous study of them using TEA, and providing examples[iv] from the corpus:

(1) QUOTATIVES to introduce direct speech (Tagliamonte and D'Arcy 2004)

a. She's like, "That's your boyfriend?" (3T)

b. And she said, "What are you doing?" (2e)

c. People will think, "This is a good, like, advantage." (2n)

(2) INTENSIFIERS preceding adjectives (Tagliamonte 2008)

a. My mom said that it's really dangerous for me. (I4)

b. It seemed a very pleasant place. (6)

c. The guys are so different! (ND)

(3) GENERAL EXTENDERS (Tagliamonte and Denis 2010)

a. I have to carry like a map around campus and stuff (4e)

b. I heard that underneath the Taj-Mahal, they bury like the kid or something. (2p)

c. … supplies and things like that (N<)

(4) STATIVE POSSESSION (Tagliamonte, D'Arcy and Jankowski 2010)

a. It has some strength (I2)

b. And I've got one [kid] in my house. (NT)

c. I got a test in math and law coming up. (3m)

(5) DEONTIC MODALITY to express necessity or obligation (Tagliamonte and

D'Arcy 2007b)

a. Things change. And you have to change with it. (Nr)

b. So it's pretty understood on both sides what one must do to get the

other's attention, you know. (3D)

c. We told her owner "You've got to get control of that dog." (Ir)

d. You haven't seen *Orange County*? You need to see that movie. (3S)

      These five variables comprise lexical, morphosyntactic, and discourse-

pragmatic levels of the grammar. As the earlier studies of these variables in

Toronto (cited above) have shown, all five demonstrate both synchronic

variability and change over apparent time in Toronto English. The next step in our

study was to identify a replicable, consistent and objective criterion for

determining which variant was on the leading edge of change. We categorized as

"incoming" that variant which was used most commonly by speakers under the

age of 30, although we also discuss some results based on alternative criteria

below. Based on the previous studies of these features, the incoming variants

were, quotative *be like*, intensifier *really*, general extender (GE) *and stuff*, stative

have and deontic *have to*; each incoming form is illustrated in the examples

labelled (a).

Although all the variables are established as changes in progress in the community (Tagliamonte and D'Arcy 2009), some of the changes are long-term developments and some of them are relatively recent innovations. Quotative *be like* is a recent innovation in Toronto (Tagliamonte and D'Arcy 2004), and, based on previous work in the TEA, we also identified as more recent innovations intensifier *really* and GE *and stuff* as their upward trajectory in the TEA data only began in earnest with 40-60 year olds (Tagliamonte 2008, 372; Tagliamonte and Denis 2010, 358). On the other hand, the preference for stative *have* and deontic *have to* was established less recently as they are the most frequently used variant for all ages of speakers in Toronto (Tagliamonte and D'Arcy 2007b; Tagliamonte et al. 2010). Despite these differences in terms of time-depth, as we will demonstrate, recentness cannot explain the correlations we see in the data.

After establishing the incoming variant for each variable, we then determined a rate of use for each speaker across the phenomena. To do this, we conducted a fixed effects logistic regression using Goldvarb X (Sankoff, Tagliamonte and Smith 2005) for all the speakers who had been included in the original studies. We used the incoming variant as the application value in each case (see Guy 2013, 66 for a similar approach). This provided a series of individual factor weights (i.e., relative probabilities, following Guy 2013) to express the likelihood that an individual speaker would use a particular variant. For example, using the data for quotatives, we calculated a factor weight for each speaker for the use of *be like*. Then, using the data for intensifiers, we calculated a factor weight for each speaker for *really*, and so on for the remaining variables.

In general, this approach yielded meaningful factor weight values, though it was necessary to consider two categories of exceptional cases. In the first category were the situations in which we were unable to adequately determine a speaker's behavior for a particular variable. These cases arose either because we did not have any tokens at all from a speaker for one of the variables, or because we had only a single token for a speaker for that variable. As we were unable to confidently determine an appropriate factor weight in these cases, the missing values were excluded (pair-wise) in our analysis. Thus, in the tables below, the number of matched speakers is not identical across all the variables. The second category of exceptional cases were speakers for whom we had two or more tokens but who were either categorical users (those who used only the incoming variant) or categorical non-users (those who never used the incoming variant). These speakers were assigned factor weights of 1 and 0, respectively in these cases.[v]

Once we had a set of factor weights for the incoming variants and had matched them by speaker, we were then able to investigate patterns across variables. Due to the fact that we were evaluating a hypothesis that a leader of one change would also be a leader of other changes, we needed a statistical approach that would test a linear relationship. Following Labov (2001) and Guy (2013), and in keeping with much other recent work in the area (see Guy and Hinskens 2016, 6 for discussion), we used the Pearson product-moment correlation coefficient (Pearson's r) as our statistic. Pearson's r was appropriate for our data because we had multiple scores (factor weights) for each speaker, and because the Pearson statistic tests how well the data fit a linear model (values fall between minus one

and positive one, with values near zero indicating no relationship). We were expecting positive values, since we hypothesized that, if there were a relationship between innovative forms, the use of one incoming variant to be positively correlated with another incoming variant. We therefore used the one-tailed[vi] significance test. Statistical tests were executed using the *cor.test* function in *R* (R Development Core Team 2010). To allow a comparison of our results with previous studies, we calculated correlations by pairs of variables for all speakers first. We then performed a second round of calculations looking only at the leaders.

In any study of innovation, determining who to categorize as a leader is a non-trivial exercise, which can be further complicated by low token counts per individual (Nevalainen et al. 2011, 14). Those categorized as "innovators" in diffusion theory typically make up only 2.5 percent of any population (Rogers 2003, 281), making a statistical study of them problematic given the usual size of sociolinguistic corpora. Nevalainen et al. (2011, 7) cited "sparse" data as the motivation for not looking at the "incipient" stage of change in their historical data. The group we examined as leaders also included not only the 2.5 percent who would be innovators but also those who would be categorized as early adopters, and some of the early majority speakers. To do this, we set our criterion for leadership of a change as a factor weight of .70 or higher for the incoming variant. This gave us groups of leaders that ranged from 12 individuals (intensifier *really*) to 35 (deontic *have to*).

As we identified the leaders for each variable, it was soon apparent that the population of leaders for one variable was not necessarily the population of leaders for the other variables.[vii] The lack of consistency in which individuals were leading across variables required a careful approach for calculating correlations for our identified leaders. Given that the leaders were not consistently the same individuals across variables, we adopted a strategy for the analysis in which we undertook a *series*[viii] of correlation calculations rather than simultaneously examining everyone who was a leader of any of the changes. That is, we first identified the leaders of one variable, then calculated the correlation coefficients across the four other variables for the leader subset of the first variable. We subsequently repeated the process for the (different) subset of speakers who were leaders for each change. For example, starting with quotative *be like*, we identified the individual speakers who met the .70 criterion (i.e., the leaders of *be like*). Next, we calculated the correlation values for only those individuals (the leaders of *be like*) by looking at their use of *be like* with intensifier *really*, GE *and stuff*, stative *have* and deontic *have to*, respectively. We then repeated the process for the leaders of the other variables.

This leaders-by-variable strategy also allowed us to examine different criteria for leadership for different variables in our subsequent analyses. We were able to calculate correlations for leaders using a higher factor weight criterion with the two longer established changes, stative *have* and deontic *have to*, and a lower factor weight criterion with the newer changes such as quotatives and intensifiers,

14

as we discuss further below. Having now described our methods, we turn to the results of the analysis.

RESULTS AND DISCUSSION

ALL SPEAKERS. As a first line of enquiry, we calculated correlations in the use of the incoming variants for all speakers for whom we had data (not only the leaders). This information is displayed in Table 1 (where a correlation value of '1' represents the correlation of each variable with itself). In some cases, for instance quotative *be like* versus deontic *have*, we did not find a statistically significant relationship in the use of innovative forms at all; these are indicated with a dash in the table. When we did have a significant correlation ($p < 0.01$), we generally found low to moderate relationships, with the greatest correlation being between quotative *be like* and intensifier *really* ($r = 0.47$, df $= 51$). Other significant correlations ($p < 0.01$) included the use of quotative *be like* and GE *and stuff* ($r = 0.36$, df $= 46$), quotative *be like* and the use of stative *have* ($r = 0.26$, df $= 82$), and the use of stative *have* and deontic *have to* ($r = 0.34$, df $= 75$). Moreover, these findings corroborate the results of previous similar studies: some correlations can be observed between variables, but most of the time there is none.

Table 1: Statistically significant correlations across incoming variants for ALL speakers ($p < 0.01$; n = Number of speakers)					
	Quotative *be like*	Intensifier *really*	GE *and stuff*	Stative *have*	Deontic *have to*

15

	r	n	r	n	r	n	r	n	r
Quotative *be like*	1								
Intensifier *really*	.47	53	1						
Extender *and stuff*	.36	48	-	31	1				
Stative *have*	.26	84	-	49	-	54	1		
Deontic *have to*	-	70	-	44	-	44	.34	77	1

We will examine the correlations for leaders in a moment, but the patterns of correlation for all speakers were intriguing, and we propose some possible explanations for those results first. Stuart-Smith and Timmins (2010, 53) found that speakers were more consistent in their use of the more recent phonological changes they examined, so we hypothesized that the nonphonological variants that reached their leading status in our data more recently might also be correlated. As noted earlier in the discussion of the variables, *really* and *and stuff* are more recent innovations (Tagliamonte 2008; Tagliamonte and Denis 2010) and thus might be expected to pattern together. However, the use of *really* is not correlated with the use of GE *and stuff*, yet both of these are correlated with the recent quotative variant, *be like*. Similarly, we see inconsistent patterns across variants that have reached their leading status less recently. Consider the correlation between stative *have* and deontic *have to*. As mentioned earlier, stative *have* and deontic *have to* are the most commonly used variants of speakers across all age groups in Toronto (Tagliamonte and D'Arcy 2007b, 71, Tagliamonte et al. 2010, 159).[ix] While this correlation would seem to suggest a role for the time depth of a variant in a

community, other evidence indicates that age of the change alone cannot explain the results. First, the use of stative *have* is correlated with use of quotative *be like*, a variant which has only superseded its predecessor (quotative *say*) among the under 30 year olds (Tagliamonte and D'Arcy 2007a, 205), therefore it cannot be the case that older innovations are only correlated with each other. In addition, the use deontic *have to* is <u>not</u> correlated with quotative *be like*, so we see an asymmetrical pattern of correlations for the two older changes. Thus, the patterns of correlation defy any simple explanation related to the age of the change or structural features of the variants.

We note that, aside from the correlation between stative *have* and deontic *have to*, all the statistically significant correlations involve quotative *be like*. Tagliamonte, D'Arcy and Rodríguez Louro (2016, passim) have recently demonstrated, through cross-variety data, that *be like* is an exceptional linguistic form. In studies of multiple variables, the rise of *be like* has also been observed to occur alongside changes in more local variables (e.g. Cheshire, Kerswill, Fox and Torgersen 2011 for Multicultural London English; Hazen, Butcher and King 2010 for West Virginia English), though the social variables with which *be like* correlates may not be consistent across communities (Buchstaller and D'Arcy 2009). We suggest that its tendency to be correlated with other innovative forms is an additional aspect of its unique development.

LEADERS. We were particularly interested in using our contemporary data to examine potential covariation among the leaders of change. As described above,

17

we used a factor weight value of .70 as the criterion for being a leader in our first analysis, and we looked at leaders of one change versus the other changes in a series of tests, as the leaders were not a uniform group across all the linguistic features. At the end of the first series of analyses, we did not have even one significant correlation (at the 0.01 level) in the use of innovative forms by the leaders of *any* of the changes we examined. That is, when looking at only the leaders of any one variable, there were no significant correlations with other variables. The leaders of one change were not leaders of changes more generally. Moreover, our vernacular spoken data showed even less of a relationship across variables than Nevalainen et al. (2011) had found in their historical written data.

Our means of identifying leaders was in line with other studies (i.e., we included more than the putative 2.5 percent of speakers who are at the cutting-edge of change) and our results were generally in keeping with previous findings. Nevertheless, Nevalainen et al. (2011, 36) noted that in change over time, "the more protracted the process, the fewer individuals [who] can be labeled as progressive or conservative, whereas in rapid changes there are fewer in-betweens." Therefore, it was possible that the original factor weight criterion of .70 had not correctly identified leaders across all the variables, and we established variable-specific criteria for leadership. We conducted a subsequent series of correlation calculations for leaders, using a leadership criterion of a factor weight of .80 for older changes (namely, stative *have* and deontic *have to* where we had also identified a large group of leaders in the first analysis) and a lower leadership

criterion of .60 for the other three changes.[x] However, this strategy also showed no correlations among the adjusted groups of leaders.

To be thorough, we explored the possibility that we had not targeted an early enough stage of linguistic change. As described above, the strongest correlation we observed across all speakers was between quotative *be like* and intensifier *really.* Given that the introduction of a new intensifier form may proceed rapidly (Tagliamonte 2008, 362), we speculated that if we looked at an incipient intensifier, rather than *really*, we might catch linguistic change at an earlier stage. For speakers under 30, the second most commonly used intensifier in Toronto is *so* (Tagliamonte 2008, 372), exemplified in (2c). We therefore repeated our correlation calculations looking at i) the use of intensifier *so* and the other variables for all speakers, not only leaders, ii) the use of the other variables by the leaders of intensifier *so*, and, iii) the use of intensifier *so* by the leaders of the other variables.[xi]

For *all* speakers, the only correlation that is statistically significant was between quotative *be like* and intensifier *so*, with a Pearson's *r* value of 0.35 (df = 51), and this value is lower than the correlation value for intensifier *really* (*r* = 0.47) displayed in Table 1. Table 2 shows the results of correlations for leaders in intensifier *so* and the other variables we examined. We again see that the values are not statistically significant, except in one key pair: the use of quotative *be like* and intensifier *so* are highly correlated. For the leaders of quotative *be like*, the correlation between quotative *be like* and intensifier *so* is 0.63 (df = 9), much higher than any of the correlations across all speakers. For the leaders of

intensifier *so,* the correlation between quotative *be like* and intensifier *so* is a remarkable 0.80 (df = 5), a very strong correlation by any measure.

Table 2: Correlations with intensifier *so* (Values in brackets not significant at p < 0.05)			
Leading in	Correlated with	*r*	n
Intensifier *so*	Quotative *be like*	.80	7
Intensifier *so*	GE *and stuff*	[-.28]	3
Intensifier *so*	Stative *have*	[.26]	6
Intensifier *so*	Deontic *have to*	[-.21]	5
Quotative *be like*	Intensifier *so*	.63	9
GE *and stuff*	Intensifier *so*	[-.52]	12
Stative *have*	Intensifier *so*	[-.21]	19
Deontic *have to*	Intensifier *so*	[.29]	27

We now probe the relationship between quotative *be like* and intensifier *so* in detail, including any confounding effects. First, quotative *be like* and intensifier *so* are strongly female-led changes in Toronto (Tagliamonte and D'Arcy 2004, 504; Tagliamonte 2008, 384), therefore it would be reasonable to expect that these two variables were correlated because they are both used by young women. In order to explore the role of speaker sex, we divided the leaders into male and female, and calculated the correlations for male and female leaders separately. We did not get statistically significant relationships; however, this partitioning of the data led to some very low speaker numbers (less than five) which may be the root

cause for this result. There *might* be an explanation related to innovative variants and speaker sex, similar to the observation in previous studies of stable variation (Guy 2013, 69; Thorburn 2014, 261) that women were slightly more consistent in their choices than men. However, even if the correlations between quotative *be like* and intensifier *so* had been statistically significant when segregated by speaker sex, the correlations (not shown here) would have been positive for both male and female speakers. That is, the higher use of one incoming variant would have been correlated with the higher use of the other incoming variant for both male and female leaders. In addition, examining correlations for other variables with the speakers separated by sex did not produce significant correlations. Therefore, younger female speakers are not always consistent in their use of incoming variants. Moreover, previous work by Tagliamonte and D'Arcy (2007b, 78) on deontic modality in Toronto English demonstrated a sex effect, with women between 17 and 30 using (statistically significantly) more *have to* than their male counterparts, but we did not see statistically significant correlations for deontic *have to* and either *be like* or *so*. All this evidence suggests that the correlation between quotative *be like* and intensifier *so* is not simply a confounding effect of speaker sex.

Other similarities between *be like* and *so* also could not explain their relationship. First, we returned to the possibility that the recentness of the variant might be relevant, as it was for the phonological variables studied by Stuart-Smith and Timmins (2010). Quotative *be like* is definitely an innovation of the past 30 years, including in Toronto (Tagliamonte and D'Arcy 2004, 495). However, the

use of *so* as an intensifier is attested in the *TEA* data even among the oldest

speakers (Tagliamonte 2008, 372). Thus, recentness of the innovation is again not

what is at work here either. Second, we noted that *be like* and *so*, despite the

difference in when they originally arose, have both recently been adopted with

vigor by younger speakers. That is, they are both seemingly trendy variants. We

speculated that they are correlated because they are both markers of young,

innovative speech, particularly in contrast to the incoming variants in the stative

and deontic systems. However, the use of the GE *and stuff* and intensifier *really*

are also associated with younger speakers (Tagliamonte and Denis 2010, 360) so

the connection cannot be only the youthfulness of its users either. Third, in an

examination of covariation in stable variables of Brazilian Portuguese, Oushiro

and Guy (2015, 165) noted that unmarked forms "are subject to more automatic

and consistent treatment" and are more likely to covary. However, the incoming

variants examined here are likely to be more marked because they are innovations,

and thus we would expect they would be less likely to be correlated. Therefore, we

are inclined to come back to the idea that the relationship between *be like* and *so*

may be due to the distinctive nature of *be like.*

Taken together, the all the evidence we have presented leads us to

conclude that individuals who lead in one (nonphonological) linguistic change are

unlikely to be leading in other(s) in progress at the same time. As a practical

example, consider the behavior (i.e., factor weights) of two individuals. The

(pseudonymous) speakers shown in Table 3 are two people who had high factor

weights for intensifier *so*. As you can see, they are both women, and very close in

age. However, they do not display consistent behavior with *be like,* with Mary being a higher user of it than Fiona.

Table 3: Comparison of factor weights for two leaders of intensifier *so*				
	Mary (19, female)		Fiona (21, female)	
	FW	n	FW	n
Intensifier *so*	.81	189	.71	102
Quotative *be like*	.63	192	.42	21

Guy and Hinskens (2016, 5) noted that "on-going language change is a systemic disrupter for community coherence", and it may be that the disruptive nature of innovation makes it nearly impossible to identify correlations across individuals synchronically.

To sum up, some individuals do participate actively in more than one change, but we do not see *consistent* behavior across all of the leaders for most variables. This result is in keeping with previous synchronic studies of phonological change (Maclagan et al. 1999; Stuart-Smith and Timmins 2010) and diachronic study of nonphonological change (Nevalainen et al. 2011), and holds regardless of what methodological approach is used to examine the data. This result corroborates the idea that speakers are doing what sociolinguists have observed for decades: they are making use of the linguistic resources that are available in context (Eckert 2008). Nothing we have found contradicts previous claims about what social groups leaders may come from or what personal traits leaders may have (e.g., Chambers 2009, 92-114; Labov 2001, 409-411; Milroy

1992, 164-205; Nevalainen et al. 2011, 27-34; Rogers 2003, 282-285). What these findings reveal is that, when there are changes in progress in a community, speakers who have similar social or personal characteristics may lead these changes, but the particular individuals within that group may well be on the forefront of *different* changes.

Nevalainen et al. (2011, 27) explained some of the inconsistency in their results by appealing to "varying social trajectories of the changes." Similarly, Oushiro (2016, 125) noted that variants were more likely to pattern together if there was community "agreement on their social meaning". While objectively determining community meaning for our variables retrospectively is not possible (Eckert 2008), it may be a fruitful avenue for future study of covariation of incoming changes that can be predicted to diffuse in the speech community.

CONCLUSION

We have now put the covariation hypotheses to the test using contemporary, spoken data by examining a series of changes in progress in a cohesive speech community among a substantive data set of (non-phonological) variables among leaders and non-leaders. While it might seem intuitive to think that the leaders of one change will also the leaders in others, this is not the case. The findings we have presented demonstrate, similar to previous studies that were more focused on phonological variables (Guy 2013, 70; Maclagan et al. 1999, 33-34; Oushiro and Guy 2015, 165; Stuart-Smith and Timmins 2010, 49-50; Thorburn 2014, 263), and to Nevalainen et al.'s (2011) examination of historical morphosyntactic data, that

individuals who are leading one change in a speech community are not necessarily

leading other, concurrent changes. These building findings lead us to speculate

that the definition of a "leader of linguistic change" may not rest with linguistic

criteria. Instead, a speaker's use of one frequently used innovation may be enough

as long as it operates in tandem with other concordant social characteristics.

ENDNOTES:

[i] We refer here to the diffusion of innovations more generally and not the specific mechanisms of
linguistic diffusion (e.g. Labov 2007).

[ii] See Guy and Hinskens (2016) for an overview of covariation with stable variation.

[iii] https://www.helsinki.fi/en/researchgroups/varieng/corpus-of-early-english-correspondence

[iv] Speaker codes from the TEA are given in brackets after each example.

[v] We considered the possibility that the inclusion of categorical users of incoming variants as
leaders might interfere with the linearity of any relationship. However, recalculating the
correlations for leaders of change with categorical users excluded from the analysis did not
change our findings. We note that Nevalainen et al. (2011, 6) also observed and included
writers with categorical use and non-use of features.

[vi] We also examined the data with the two-tailed significance level and the results are the same.

vii Although this suggested that we would not have strong correlations across variables, it did not rule out the possibility that we might discern some consistency in behavior using statistical methods. It was still conceivable that the higher use of one incoming variant might be positively correlated with a higher use of another, even if only one of them was used often enough to meet the criterion for leadership.

viii For the sake of thoroughness, we also used *R*'s *cor* function to execute a multiple correlation calculation where we considered, in a single analysis, correlations across variables for anyone who was a leader in anything; the results were even lower correlations than those shown in Table 1 for all users.

ix At the outset, we considered, as Guy (2009: 8) did, the possibility that there might be some kind of language-internal reason why we might get parallel behavior in the variables. Deontic *have to* and stative *have* are similar in morphological form, but the correlation between the morphologically different stative *have* and quotative *be like* in Table 1 argues against a morphological explanation.

x We also examined the higher and lower cut-off points across *all* variables. However, this approach yielded unreasonable groups of leaders. For instance, a factor weight of .60 categorized over half of the speakers for whom we had data on deontic *have to* as "leaders".

xi Given the much lower number of individuals who used intensifier *so*, we also revised our significance level upward from 0.01 to 0.05, although we acknowledge that this does allow greater scope for Type I error.

REFERENCES

Buchstaller, Isabelle and Alexandra D'Arcy. 2009. "Localized Globalization: A

      Multi-local, Multivariate Investigation of Quotative *be like*." *Journal of*

      Sociolinguistics 13: 291–331.

Chambers, J.K. 2009. *Sociolinguistic Theory*. Revised Edition. Oxford: Blackwell.

Cheshire, Jenny, Paul Kerswill, Sue Fox and Eivind Torgersen. 2011. "Contact,

      the Feature Pool and the Speech Community: The Emergence of

      Multicultural London English." *Journal of Sociolinguistics* 15: 151–196.

Eckert, Penelope. 2008. "Variation and the Indexical Field." *Journal of*

      Sociolinguistics 12: 453–476.

Guy, Gregory R. 2013. "The Cognitive Coherence of Sociolects: How Do

      Speakers Handle Multiple Sociolinguistic Variables?" *Journal of*

      Pragmatics 52: 63-71.

Guy, Gregory R. and Frans Hinskens. 2016. "Linguistic Coherence: Systems,

      Repertoires and Speech Communities." *Lingua* 172-173: 1-9.

Hazen, Kirk, Paige Butcher and Ashley King. 2010. "Unvernacular Appalachia:

      An Empirical Perspective on West Virginia Dialect Variation." *English*

      Today 104: 13-22.Hinskens, Frans and Gregory R.Guy (eds). 2016.

      Coherence, Covariation and Bricolage. Various Approaches to the

      Systematicity of Language Variation. *Special Issue of Lingua* 172-173: 1-

      146.

Hopper, Paul. 1991. "On Some Principles of Grammaticalization." In Elizabeth

    Closs Traugott & Bernd Heine (Eds.), *Approaches to Grammaticalization,*

    Volume I. Amsterdam: John Benjamins: 17-36.

Labov, William. 1994. *Principles of Linguistic Change, Volume 1: Internal*

    Factors. Malden: Blackwell.

-----. 2001. *Principles of Linguistic Change, Volume 2: Social Factors*. Malden:

    Blackwell.

-----. 2007. "Transmission and Diffusion." *Language* 83: 344-387.

Maclagan, Margaret A., Elizabeth Gordon and Gillian Lewis. 1999. "Women and

    Sound Change: Conservative and Innovative Behavior by the Same

    Speakers." *Language Variation and Change* 11: 19-41.

Milroy, James. 1992. *Linguistic Variation and Change.* Oxford: Blackwell.

Nevalainen, Terttu, Helena Raumolin-Brunberg and Heikki Mannila. 2011. "The

    Diffusion of Language Change in Real Time: Progressive and

    Conservative Individuals and the Time Depth of Change." *Language*

    Variation and Change 23: 1-43.

Nevalainen, Terttu and Helena Raumolin-Brunberg. 2003. *Historical*

    Sociolinguistics. London: Longman.

Oushiro, Livia. 2016. "Social and Structural Constraints in Lectal Cohesion."

    Lingua 172-173: 116-130.

Oushiro, Livia and Gregory R. Guy. 2015. "The Effect of Salience on Co-

    variation in Brazilian Portuguese." *University of Pennsylvania Working*

Papers in Linguistics: Selected Papers from New Ways of Analyzing Variation (NWAW) 43, 21: 157-166.

R Development Core Team. 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Rogers, Everett M. 2003. *Diffusion of Innovations*. Fifth edition. New York: Simon and Schuster.

Sankoff, David, Sali A.Tagliamonte and Eric Smith. 2005. Goldvarb X. Department of Linguistics, University of Toronto, Toronto, Canada. http://individual.utoronto.ca/tagliamonte/goldvarb.htm

Stuart-Smith, Jane and Claire Timmins. 2010. "The Role of the Individual in Language Variation and Change." In Carmen Llamas and Dominic Watt (eds.) *Language and Identities*. Edinburgh University Press: 43-51.

Tagliamonte, Sali A. 2003-6. "Linguistic Changes in Canada Entering the 21st Century." Research grant, Social Sciences and Humanities Research Council of Canada (SSHRCC), #410-2003-005.

-----. 2006. "So cool, right? Canadian English Entering the 21st Century." *Canadian English in a Global Context: Theme Issue of Canadian Journal of Linguistics* 51:309-331.

-----. 2008. "*So* Different and *Pretty* Cool! Recycling Intensifiers in Canadian English." *Special issue of English Language and Linguistics, Intensifiers* 12: 361-394.

-----. 2009. "Historical Doublets in Synchronic Perspective: The English Indefinite

    Pronouns." Paper presented to SHEL6, Banff, Canada.

Tagliamonte, Sali A. and Alexandra D'Arcy. 2004. "He's Like, She's Like: The

    Quotative System in Canadian Youth." *Journal of Sociolinguistics* 8: 493-

    514.

-----. 2007a. "Frequency and Variation in the Community Grammar: Tracking a

    New Change Through the Generations." *Language Variation and Change*

    19: 1-19.

-----. 2007b. "The Modals of Obligation/Necessity in Canadian Perspective."

    English World-Wide 28: 47-87.

-----. 2009. "Peaks Beyond Phonology: Adolescence, Incrementation, and

    Language Change." *Language* 85: 58-108.

Tagliamonte, Sali A., Alexandra D'Arcy and Bridget Jankowski. 2010. "Social

    Work and Linguistic Systems: Marking Possession in Canadian English."

    Language Variation and Change, 22: 149-173.

Tagliamonte, Sali A., Alexandra D'Arcy and Celeste Rodríguez Louro. 2016.

    "Outliers, Impact, and Rationalization in Linguistic Change." *Language*

    92: 824-849.

Tagliamonte, Sali A. and Derek Denis. 2010. "The Stuff of Change: General

    Extenders in Toronto, Canada." *Journal of English Linguistics,* 38: 335-

    368.

Thorburn, Jennifer A. 2014. "Dialect Development in Nain, Nunatsiavut: Emerging English in a Canadian Aboriginal Community". PhD diss., Memorial University.