So who? Like how? Just what?
Discourse markers in the conversations of Young Canadians

Sali Tagliamonte

University of Toronto, Department of Linguistics, 130 St George Street, Toronto, Ont., Canada M5S 3H1

Received 19 November 2003; received in revised form 10 March 2004; accepted 5 February 2005

Abstract

In this paper I mine the rich resources of a large corpus of conversational data from young Canadians between 10 and 19 collected in 2002–2003 by participant observers from the same community. The spoken discourse of these speakers is replete with forms such as like, just, so, etc. as in I’m just like so happy’. What is the nature of these linguistic features? Where did they come from? How are they spreading and by whom? This pilot investigation reveals a concentration of these forms amongst the 15- to 16-year olds generally, and female speakers in particular. This result corroborates many other findings suggesting the tremendous influence of the peer group in the middle teenage years, the correlation of adolescence with dramatic linguistic differentiation and that females lead linguistic change. Further, despite the received wisdom, at least in the media, these features are not haphazard, random or indiscriminate. Instead, their patterns of use are quite circumscribed and linguistically defined. Indeed, their contrasting linguistic profiles viewed across different age cohorts (tweens, teens, and young adults) suggest that they are undergoing different types of change. More generally, the findings highlight the extent of linguistic innovation among young people in contemporary urban speech communities.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Language change; Innovation; Teenagers; Like; Intensifiers

This paper was presented for the Approaches to Spoken Interaction Panel organized by Anna-Brita Stenström and Karin Aijmer at the International Pragmatics Association Meeting in Toronto Canada on July 15th 2003.

E-mail address: sali.tagliamonte@utoronto.ca.
1. Introduction

Innovation in language affects all areas of society. A case in point is a number of dramatic ‘new’ discourse/pragmatic markers in the English language which have gained considerable high-profile attention in recent years, from the media, educationalists and linguists alike. The innovative features are highly conspicuous and typically associated with the younger generation. Some of these are forms such as like, just and so, etc. as in (1) through (5):¹

1. I’m just like so there, you know?
2. Like, that’s what I like told you.
3. I just decided and just went.
4. She’s so not cool.
5. You so want it.

Explanations for such uses are many, including pause fillers, hedges, highlighting or focus devices, as well as indicators of vagueness or ‘loose language’ (e.g., Underhill, 1988; Miller and Weinert, 1995; Andersen, 1996, 1997a, 1997b, 1998, 2000). Whatever the function, young people are unanimously singled out as the innovators (Andersen, 1996, 1997b, 2000; Buchstaller, 2001; Erman, 2001; Siegel, 2002). However, to date, very little research has explored these particular phenomena in Canada, nor more importantly, explored the origin, pathways and impacts of such changes in the English language amongst the youngest generations of speakers where they are apparently innovating from. Thus, little is known about where they are developing in the speech community, or about their trajectory of linguistic change. Indeed, another important question is whether these linguistic items really represent linguistic change or whether they are simply items that distinguish teen-talk from the language of both younger and older speakers. Of course, this question can only be answered by tracking these changes into the future; however, the results I will report on here lay the foundation for a more longitudinal study.

Although there are many claims about the use of these features by ‘young people’ — to which young people does this refer? There is actually a dearth of conversational data from speakers in the late primary (10- to 12-year old) and middle school (13- to 14-year olds) age brackets. The linguistic features in question may be diffusing to these age cohorts, or maybe even have originated there, long before later adolescence in high school (the 15- to 17-year olds).

In this paper, I provide an empirical account of a number of salient discourse markers, like and just, as well as intensifiers (really, very, so), in the spoken interaction of these

¹ I do not discuss here the use of like as a quotative, e.g., I’m like, “Hello”, and She’s like, “Hi”, in this paper (e.g., Blyth et al., 1990; Buchstaller, 2001, in press; Dailey-O’Cain, 1996; Romaine and Lange, 1991; Tagliamonte and Hudson, 1999). This function of like in this corpus is analyzed in Tagliamonte and D’Arcy (2004).
younger sectors of the population. The data comes from a new research project on the language of tweens, teens and 20 ‘somethings’ in Toronto, Canada. The method I employ is quantitative and within the traditional of sociolinguistic variation theory (e.g., Labov, 1972; Sankoff, 1988, 1980b). Although the study of discourse markers does not belong to the fields traditionally associated with this approach, I attempt to adapt the methodology in order to tap into these innovative features of language, which are undergoing change in Canadian English. One of the aims of this research is to track the new developments in English more generally.

2. The project

The project is part of an ongoing research program being conducted within the Research Opportunities Program (ROP) for second-year students at the University of Toronto. This research project focuses on ‘new’ features of Canadian English that are reported to be rapidly increasing in the speech of the younger generation, particularly in Western English-speaking countries. Admittedly, the size of this corpus is small in comparison to other teen language corpora, with only 26 speakers overall and 4–5 members of each age group. In addition, it is somewhat focused on middle-class speakers. However, as the corpus is replicated over the years it will gradually build a more socially complex distribution. For example, the recently completed 2004 corpus has now added another 24 speakers. Thus, the combined corpus currently has a break down of 8–12 members per age group. It must also be said that the data is narrow geographically, but at the outset it is worthwhile to circumscribe the corpus in this way so as not to create too diffuse a sample. For now, I restrict the investigation to the urban context of Toronto and to speakers born and raised in the city. Thus, my interpretation of the data must be necessarily cautious and limited to the particular data involved.

The long-term goal of this project is to investigate the origin of these features in the Toronto speech community more generally (i.e., across age groups), and further to explore their history in the English language. The project aims to contribute new insights into the social and linguistic motivations of language change, particularly the influence of youth. Moreover, the fact that the ROP project is repeated in each academic year also means that the corpora will eventually provide a longitudinal view of change in progress, in essence a typical ‘trend’ study of linguistic change (e.g., Chambers, 1995).

In order to ‘tap into’ innovative changes in Canadian English in Toronto, members of the ROP research team interview members of their own social and familial networks between the ages of 10 and 20 (e.g., their sisters, cousins, neighbours, friends, etc.). Thus, the

\footnote{The term ‘tweens’ refers to young people between the ages of 8 or 9 and 12. These speakers are typically referred to as pre-adolescents in the literature.}

\footnote{COLT, for example, has 500,000 words.}

\footnote{Special thanks to Jonille Clemente, Marion Hau, Madeleine Macdonald, Muhammad Velji, and Jessica Wertman — students of the University of Toronto Faculty of Arts and Science — for their work in ROP 299Y1 2002–2003 section L0101, Project 37. They interviewed the Toronto natives for this study, and extracted and coded the data.}
Interviewers, aged 20–22, are not simply participant observers but actual members of the same community. This makes this corpus one of the few data sets collected by in-group members of the current youth culture. The approach the interviewers take is simply to sit down and talk to the young people for about an hour. In each case a small, unobtrusive tape-recorder is employed, with a lavaliere (tie-clip) microphone, which is clipped within 20 cm of the mouth. The ‘interviews’ focus on informal topics such as school activities, hobbies, sports, friends, and lots of commiseration about problems with their parents, boyfriends and girlfriends, etc.

3. Data

The details of the 2002 corpus are shown in Table 1. It comprises 26 speakers between the ages of 10–20 and amounts to about 200,000 machine-readable words. The age groupings correspond to the different school levels in Canada, primary school (grades 1–5); middle school (grades 6–8), and secondary school (grades 9–12). At the outset our plans were to study the discourse features that characterize the language of young people today. However, studies of other levels of linguistic analysis are also underway: in morpho-syntax, e.g. deontic modality, have to/gotta/must, possessive have, future going to and in phonology, e.g. variable (ing), (t,d), etc.

In this paper I introduce the corpus and target for examination three salient and frequent forms in the data — like, just and so. The following excerpts from the transcriptions illustrate each of the targeted features.

3.1. Like

We have five teachers. Like, they- I don’t know- they’re not exactly- some of them are really nice. Like you really like them- Like one of my teachers, she’s amazing. Like I love what she teaches. And it’s a really- like the way that she presents like the class and the

---

5 It must be kept in mind that the interviewers are not necessarily middle-class, but come from a broad range of educational and socio-economic and socio-cultural backgrounds. This is, in fact, a particular socio-cultural fact about Toronto.
subject is really great. And I’m really understanding like everything she’s saying. And some of the other teachers are from another planet. Like can’t think properly, like.

Amanda Levy, 12, Female

3.2. Just

I’m there, I’m like, just playing around, doing nothing. [1] Same thing. [03] Same thing over and over again. And then, every once-in-a-while, me and my friend who plays the bass. Like, whenever somebody would come in, we’ll like, just stop and play Another-One-Bites-the-Dust. [1] Oh yeah. [03] It’s really funny. And then we’re going-to try and learn, like just to piss him off, really. We just do, ’cause it’s more fun.

Antonio Silvaggio, 16, Male

3.3. Intensifiers

She was actually like, really grateful and like, ’Cause like, I thought she would just be-like, “Euh! These are so small! Oh, couldn’t you find more?” Or like, something like gay. So then, I was happy that she was-like, “Aw, thank you so much like, oh, these pictures are so good!” Like, I got like, really good pictures’.

Clara Felipe, 16, Female

In this preliminary report, I explore the nature and character of these features. On the hypothesis that an increase in an item’s text frequency is an important indication of ongoing linguistic change, I focus on broad distributional facts, and examine how age and sex interact with the linguistic patterns of occurrence.

4. Results

4.1. Like

The most frequent of the targeted features is the ubiquitous form like, as in the examples in (6). In its non-verb, non-conjunction use (i.e., non-standard) it occurs 9739 times in this corpus. For many speakers, this use of like is more frequent than the word and, and often occurs multiple times in a single sentence, as in (6b–e).

6. a. Cause we made like a video.
   b. Now like my plan is to go like with one friend to the movies and stuff like that.
   c. Like I know when like wherever I start like going- buying my own clothes or whatever.

---

6 All names are pseudonyms.
d. I can read **like**, a whole day and like, “Hmm.”

e. She’d always wear **like** those tight **like** leather or **like** really tight cloth

Whether **like** can occur anywhere or is in some way rule-governed is much debated in the literature, undoubtedly due to the fact that it is difficult to define syntactically (Schiffrin, 1987). Romaine and Lange (1991) claim that **like** can occur before prepositions, adjectives and adverbs as well sentence initially. Andersen’s (2000) seminal work on **like** reveals that it can occur in a wide variety of contexts, from whole propositions to single terms. Further, it can be syntactically bound (clause internal) or unbound. Other accounts suggest that **like** can be placed anywhere in the sentence. For example the Webster’s New World Guide to Current American Usage says “some speakers use **like** without meaning or syntactical function almost anywhere in a sentence” [emphasis mine] (Randall, 1988: 206). In the popular press **like** is widely regarded as a ‘space filler’, which is apparently used “when the person speaking cannot come up with anything else to say” (Body Berg, The Exponent, U. Wisconsin-Platteville, 22 April 2002).

However, there are clear hints that **like** may have a certain systematicity in sentence structure. Meehan (1991: 40) suggests that “the positioning of **like** in the sentences is never random, although to some hearers it may seem so”. Underhill (1988: 243) claims that **like** is ‘rule governed’ nearly always introducing a constituent, occurring most often before a verb phrase, but also before a noun phrase or between the auxiliary and the rest of the verb phrase. Romaine and Lange report that sentence initial position represents the most frequent (50%) of all tokens and Andersen (2000) suggests that it is less likely to occur within phrases with high syntactic fixedness, and more likely to occur immediately before the lexical material of a phrase, as opposed to grammatical words. In much more recent research, Wolgemuth (2003) showed that in Canadian English (c. 1995) the most frequent positions for **like** were before a noun phrase (34%, N = 68) and before a sentence (23%, N = 47).7

Using these observations as a guide, let us now turn to the data to determine how often and where **like** is occurring in Toronto Youth English. Fig. 1 shows the frequency of **like** according to the following grammatical category.8

First of all, these data indicate that **like** may not occur just anywhere. What we find is that the vast majority of the 9739 tokens of **like** in the corpus occur in only two or three locations in spoken interaction — 30% of all the **like** tokens occur before an noun phrase [\_NP], as in (7). Note that the noun phrase may be full noun or indefinite pronoun9, and can be bare, determined, or quantified. Another 23% of the **like** tokens occur at the beginning of a sentence [\_S], as in (8).10 With respect to this context, it is clear that **like** can occur after discourse markers (e.g., so) and/or temporal adverbs (e.g., then). The

---

7 Wolgemuth (2003) also found that **like** never occurred within well-established compounds, or within proverbial expressions. A similar observation was made by a reviewer of this paper.

8 The categories in Figure 1 refer to the post-like grammatical context. For example, \_S refers to a following sentence, \_INT refers to a following intensifier and \_DM refers to a following discourse marker. Further, \_SUB refers to a following subject, \_PREP refers to a following preposition, and \_ADV, a following adverb.

9 Recent findings reveal that **like** does not occur before a personal pronoun (Wolgemuth, 2003).

10 The notion of sentence is problematic in conversational studies, since it is difficult to categorize spoken language, which is filled with one word or one phrase utterances, cut-off sentences and the like. Our approach was simply to categorize the following context as a sentence whenever it contained a main clause determiner phrase, whether this was a fragment or not.
next most frequent context for *like* (9%) is before a verb, as in (9). Interestingly, a cursory examination of the verbs reveals that they are virtually all encoded with preterit morphology.

7. a. You’re trying to get attention in *like* [the wrong way]. (p)
   b. So then when he had the speech, he had *like* [everything] on *like* [paper]. (a)
   c. You can get sued for *like* [a million dollars]. (g)
   d. So you get *like* [a worm] and then you catch *like* [a little tiny stupid little dumb fish]... (t)
   e. Oh yeah, yeah. You get *like* [three passes]. (p)
   f. I found *like* [a hundred pictures] for her already, ‘cause she needs *like* [six hundred].

8. a. So then *like* he posted up signs around the school, (a)
   b. *Like* it’s where you photocopy things and open floppyies and print things for them. (s)
   c. For messing up the computer, apparently *like* the whole system was ruined. (q)
   d. *Like* it makes me feel good, it makes me feel better. (t)

9. a. So then he *like* [whipped] it. (a)
   b. Halfway up the hallway I just *like* [dropped] it and *like* [walked] over to a place and bought water. (k)
   c. It had a fish right in its mouth and we *like* stepped right on over it. (e)
   d. I think he’s just *like* changed. (q)

This tabulation shows that *like* is not evenly distributed across sentence structure. Indeed, it corroborates earlier research in demonstrating the propensity of *like* in pre-sentential and pre-noun phrase position. Interesting is also the fact that the frequency of *like* in pre-noun
Figure 2 shows the distribution of *like* and *and* tokens by individual arranged from youngest to oldest. The distribution of *like* in each of eighteen individual speakers in our corpus processed at the time of writing. The individuals are arranged in age from youngest to oldest and separated by male (the striped columns) and female (the dark columns).

Fig. 2 also shows that females tend to use more *like* in every age group. Moreover note that for many of these speakers (*N* = 8) the frequency of *like* is greater than *and*.
Fig. 3 shows the distribution of *like* across age groups. In this tabulation, the proportion of *like* has been calculated out of the total number of words provided by each age group. Fig. 3 reveals a very dramatic pattern. The 15- to 16-year olds are using more tokens of *like* than any other age group. At the two ends of the scale — the 10- to 12-year olds and the 17- to 19-year olds, however, the use of *like* is much lower. The frequency of *like* across age groups suggests that Toronto youth acquire the use of *like* in primary school, develop its use in middle school, but it is not until secondary school that it is used in abundance. Note the drop in frequency at the university level, where the use of *like* declines substantially. This trajectory resembles the classic pattern for *age-grading* — a change correlated with “a particular time of life” (Chambers, 1995: 164), rather than a change that is percolating successively through the generations, as would be expected with an incoming grammatical change.

4.2. *Just*

While an inordinate amount of media attention as well as academic research has been devoted to *like*, the use of *just* is barely mentioned. However, this form has also been increasing in recent years and has apparently garnered the same type of stigma as *like*. Speakers are admonished to reduce their use of *just* and the form itself is referred to as “the whiny excuse word” (Bickel, 2002).

Many of the current uses appear unremarkable in isolation, as in (10a–c), where their function has its standard adverbial meaning of *simply* or *only*. The uses in (10d and e), however, show that while the same reading is present, the context of use, particularly in (10e) is somewhat strange. This suggests that *just* may be expanding beyond its standard function.

10. a. That’s *just* textbook work.
   b. Well, I dunno, they’re *just* kissing.
   c. Some of them are cute, some of them are *just* weird.
   d. So I can *just* go on the Internet, look up some cheats and then *just* play.
   e. I *just* stayed home ‘cause someone was taking care of me. And then I was *just* watching T-V. And I *just* took a nap.

Indeed, when we examine our corpus, we find that *just* is one of the most frequent forms used among the young people. In this corpus it occurs 1738 times. Once again, we can begin with the question — does *just* occur just anywhere? Fig. 4 shows a tabulation of the frequency of *just* tokens according to the following grammatical category.

---

11 As part of the preliminary assessment of these data, we calculated the total number of words and the total vocabulary for each speaker. Individual word counts were tabulated using CONCORDER 2.0 (Rand and Patera, 1992) which easily calculates the frequency of each word spoken by each speaker in the data. What counts as a single word as opposed to a compound is another problematic issue. We follow a strict transcription protocol so that corpus-internal consistency is maintained. This ensures that word counts are comparable across speakers.

12 This count represents all orthographic instances of *just*.
As we observed with like, just does not occur anywhere either. Indeed, it is even more highly circumscribed than like. The majority of forms — a full 46% (N = 760) — occur before a verb, as in (11). Note too that the verb may be simple present, present progressive or simple past.

11. a. I just know that an older sister would be awful. (f)
    b. Oh, wouldn’t it be so nasty if we were just talking to him and then they just fell off? (a)
    c. Just practicing for like the school. We just put them on, we didn’t care. (e)
    d. Two minutes into it I just forgot and just started talking. (k)

As far as I am aware, this propensity of just to occur before a verb has not been reported before. Here too, the shear frequency of the collocation in the data is suggestive of ongoing linguistic change.
However, does this feature show the same trajectory of change across the generations as we saw for like? Fig. 5 shows the distribution of just across age groups. In this tabulation, as with like in Fig. 3, the proportion of just has been calculated out of the total number of words provided by each age group.

Here, unlike the distribution for like we see a regular progression such that the use of just increases incrementally from youngest to oldest speakers. This type of distributional pattern is characteristic of ‘real’ change in progress — one that is percolating successively through the generations in the community (e.g., Labov, 1972; Chambers, 1995).

If the rise in just is an indication of ongoing change we might expect to find, consistent with earlier research, that females lead across the speech community. Fig. 6 tests this possibility by comparing the male and female speakers. In this analysis, the proportion of just has been calculated out of the total number of words for each of male and female speakers in each age group.

Fig. 6 reveals that female speakers are indeed way ahead of the males — clearly leading the change in most age groups. This correlation is also a hallmark of change in progress. However, notice the reversal amongst the university students. Here, males use more just than females. This suggests that the use and trajectory of just across age groups does not adhere entirely to all the characteristics of ongoing change.

In sum, the feature just provokes a number of research questions. Firstly, what kind of change is this? Further, what is just doing in the grammar? In other words, what are the young people doing with just? These are the types of questions I will be turning to in the next phase of this research. In particular, we will need to calculate which of the multiple functions of just (Erman, 1997) are represented in each age group and whether or not these are changing in apparent time. Further, we will have to examine the pre-VP position in more detail by examining how the verbs used with just compare with all the other verbs in the data.
4.3. Intensifiers

The final feature I consider is the use of intensifiers, as in (12).\(^{13}\)

12. a. It was, really, really, really bad.
   b. She’s so completely overprotective.
   c. It was very pretty . . . I loved her hair, it was amazing . . . And it was really pretty.
   d. Yeah, that was extremely nice . . that’s really nice.
   e. That must be pretty freaky in the winter time.

The corpus presents a particularly good opportunity to study this feature as young people are notorious for an overabundance of intensification in general (Stenström, 2000). Moreover, intensifiers are thought to be increasing in frequency in recent times (e.g., Ito and Tagliamonte, 2003). Further, intensifiers are associated with rapid turnover and lexical renewal and thus are thought to present an excellent means to track language change as well as to tap into current trends in contemporary English (Ito and Tagliamonte, 2003; Stenström, 2000).

I focus here on the types of intensifiers that occur with adjectives (Stenström, 1999: 70) and follow earlier methodological practice in extracting them from the data (see Ito and Tagliamonte, 2003). Thus, in contrast to the preceding analyses on like and just, here I take into account not simply the forms under investigation, but also the contexts in which they could have occurred but did not (Labov, 1972; Guy, 1993). One of the advantages of this method is that it enables me to compare the findings with other studies of the same feature.

Table 2 shows the overall distribution of adjectival intensification in the data. Of the 4296 adjectives which could be ‘boosted’ with an intensifying adverb, 1396 were, making the overall propensity of intensification 32.5%. This is quite high — indeed it is

\(^{13}\)Only forms functioning as intensifiers are included in this analysis. Other uses of the same forms, e.g., really as an adverb or pretty as an adjective, have been excluded.
substantially higher than previous reports. For example, in contemporary British English, there is 24% intensification overall (Ito and Tagliamonte, 2003) and in the television series *Friends* Tagliamonte and Roberts (2004) find 20% intensification. In these data, the rate of intensification is 10% higher.

Given the well-known tendency for intensifiers to cycle and recycle over time, an important question to ask is whether or not this overall rate of intensification changes according to speaker age. Fig. 7 shows the distribution of the most frequent intensifiers — *really, so, pretty, very* — by age.

Fig. 7 reveals that the high average rate of over 30% intensification by age is remarkably consistent across all age groups. Thus, there is no evidence for increasing intensification according to speaker age, at least not amongst these younger sectors of the population.

Table 2 also provided a breakdown of the intensifiers that occur more than 10 times in the corpus. Although there are a variety of different intensifiers, most of these occur very infrequently. The vast majority are represented by a very small group — *really, very, pretty* and *so*. It is widely known that intensifiers change constantly and rapidly in form and
frequency from one generation to the next. This means that the young people may not be using the same set intensifiers in each age bracket. Thus, the next question that arises is: which intensifiers are used by the different age groups?

Fig. 8 separates out three of the most frequent intensifiers in the data — very, really and so — and displays which of these intensifiers are being used in each age groups and how often.

Fig. 8 shows that very is marginal across the board. However, this is not surprising because very is the older, more formal, form (see Ito and Tagliamonte, 2003). Really, the intensifier that is the most frequent amongst British youth (Stenström, 1999; Ito and Tagliamonte, 2003) and thought to be the most frequent and ubiquitous intensifier in North America more generally (Labov, 1985) — is indeed the most frequent intensifier in most age groups. These results for the distribution of really and very mirror these contemporary reports and reveal that the young speakers in our corpus are picking up the typical norms of their community. However, notice that so is a close contender. The examples in (13) show some typical uses:

13. a. I feel so out of style! (b)
   b. Oh, I’m so jealous! (w)
   c. Lord-Of-The-Rings was so boring; it was talk talk talk talk talk. (d)

Fig. 8 shows that the use of so increases from the 10- to 12-year olds to the 13- to 14-year olds, to the 15- to 16-year olds where it is most frequent. However, notice that use of so recedes amongst the university students. Taken along with the results for like and just, it appears that there is a blossoming of the ‘new’ intensifier so at the secondary school level; while the two ends of the pole are more conservative.

As I considered earlier, if this is truly a linguistic change in progress, we might expect to find that females are leading the change. Indeed, earlier research (Stenström, 1999: 77; Ito and Tagliamonte, 2003) showed a tendency for females to use more intensification in general, and recent research (Tagliamonte and Roberts, 2004) has demonstrated a female
preference for *so* in particular. Fig. 9 tests whether there is a difference in the behaviour of male and female speakers for this intensifier in this corpus.

Fig. 9 shows that females are overwhelmingly leading in the use of *so* from middle school (the 13- to 14-year olds) onwards.

Thus, for this feature there is evidence of age grading as well as ongoing linguistic change. On one hand, *so* is the favoured intensifier amongst the secondary school cohort and recedes as the speakers get older, suggesting age-grading. On the other hand, there is dramatic sex differentiation across the board. Females clearly lead in every age group beyond the tweens, the 10- to 12-year olds (the youngest cohort in our sample). This is a pattern typical of incoming linguistic change.

5. Discussion

I now consider the general trends that have become apparent with regard to the use of *like*, *just* and intensifier *so* in these data.

First of all, the use of *like* and *just* in the English of Toronto youth between 10 and 20 is exceptionally frequent. *Like* can occur up to once every 20 words; while *just* occurs up to once every 50 or 60 words. As far as intensification is concerned, the rate is higher than in previous studies. Taken together these facts highlight the innovative nature of these features in this variety of English. There is an acceleration of frequency for all of them.

However, when the results for age distributions for the three different features *like*, *just* and intensifier *so* are viewed in tandem, each one provides a different perspective on linguistic change. Taken together, they provide a composite picture of the overall nature of ongoing linguistic innovation in these data.

5.1. Like

Of the three features — *like*, *just* and intensifier *so* — only *like* shows a clear and unique concentration amongst the 15- to 16-year olds. Indeed, the shifting frequencies of its use by different age groups shows a trend that looks like classic age grading. The youngest speakers participate in the use of these forms, but they are the most conservative users. There is a trend towards greater use amongst the next age cohort, the tweens (the 13- to 14-year olds), and again with the 15- to 16-year olds. However, the oldest speakers (the 1st year university students) show a trend toward more conservatism, patterning along with the youngest age group. This marks the type of retrenchment of conservative norms identified by Chambers (1995: 170) and pinpoints the direction of linguistic change towards standard (mainstream) norms as adolescents enter young adulthood. This trajectory of change for *like* suggests age grading. Of course the next question is: what is it about the 15- to 16-year olds that lends them to an overabundant use of *like*? Further, what functional and/or pragmatic features are observable that may be connected with

---

14 The conservatism of the youngest age group may be interpreted as reflecting the degree of influence of parents during this phase of development. Indeed Chambers (1995: 158) argues that the first formative period in language development is characterized, in part, by the influence of family.
adolescence in general, and their linguistic behaviour in particular. It is to these issues that we can now turn, framed by the overall distributional patterns in the material.

5.2. Just

The use of just contrasts with the trajectory of development visible for like. Use of just, overall, appears to be increasing incrementally from one age group to the next. Furthermore, sex differentiation is visible from the youngest age group right through. Here then, the trajectory of change bears the hallmarks of ongoing linguistic change.

5.3. Intensifiers

As far as the use of intensifiers is concerned the corpus reveals a high frequency of intensification. Like like, the incoming feature, so is concentrated amongst the 15- to 16-year olds; however, with so there is no evidence for a reversal back towards the more traditional intensifiers amongst the oldest age groups. Instead, so seems to be taking over as the new intensifier across all age groups. However, we do see that amongst the first-year university students there is a reversal of sex differences. This suggests a reallocation of the social value of intensification with so amongst young adults. Where once so was associated with females, it now appears to be used by males as well, if not more so. I suggest therefore that as so diffuses across the population it may be expanding to include both male and female speakers equally.

6. Summary

Table 3 summarizes the findings for like, just and so.

Like, just and so are all salient features of Toronto Youth English, concentrated amongst 15- to 16-year olds generally and female speakers in particular. However, only like declines by age 20, while just and so prevail, perhaps heralding their status as new features of English. These overall findings suggest that both just and so represent bona fide changes in progress. Further research is needed in order to clarify precisely how, or if, these features are changing in this speech community.

Taken together, the findings support the tremendous influence of the peer-group in the middle teenage years — high school. The correlation of adolescents with dramatic linguistic differentiation and innovation is not new. This relationship has been a prominent
finding in the sociolinguistic literature (Chambers, 1995), particularly in the seminal work of Eckert (e.g., 1988, 1997, 2000). It is confirmed here in graphic detail for the Toronto youth for features within the domain of discourse/pragmatics.\footnote{Eckert’s (e.g., 1988) studies focussed on sound change only.}

However, the contrasting social profiles of like, so and just as viewed across the different age groups (tween, teen, young adult) suggest that these features are undergoing different types of change. Thus, one of the advantages of this cross-sectional sample of young people is that it provides a much more detailed picture of different linguistic changes within this critical age group in a speech community. Eckert (1988: 206) argued that adolescence, “... provides greater motivation than at any other time in life to adapt linguistic patterns to community structure. This age group, therefore, provides an important key to the study of the mechanisms of such adaptation.” I suggest that the adaptation process is developmental, beginning in pre-adolescence and extending through the teenage years (high school). Further, the trajectory of change across these ages leads to an enhanced picture of what the socio-symbolic value of linguistic features may have as they develop.

On a more methodological note, these results highlight the value of pursuing a quantitative analysis of proportion and distribution when it comes to innovating features, even when they may have a number of different functions in the grammar. It is only when the high frequencies of individual forms are calculated from the total number of words spoken by individuals or groups (or some other normalizing measure) that number of forms can be compared accountably (whether across different sub-groups of the speech community or across studies). In this way stability or change can be identified in usage patterns.

It remains to determine whether the patterns visible from the simple frequency counts of overt forms across linguistic contexts of use will hold when every instance of each of the relevant grammatical categories in the data is taken into account. Such an analysis would have required counting every single instance of nouns, verbs and other categories in the data and was beyond the scope of this preliminary paper. However, see D’Arcy (2005). It is only when the high frequencies of these individual forms are calculated from the total number of functional categories of the same type, that these raw frequencies can be compared accountably as well. Thus, for example, the high frequency of like in pre-noun phrase position may simply be the result of a higher frequency of noun phrases in the data. The high frequency of just in pre-verbal position simply the result of a higher frequency of verbs. I will turn to this type of analysis in future research (e.g., Tagliamonte, 2004).

Finally, in all of these analyses, our research has demonstrated that the females are often well ahead of the males. This too is a consistent finding throughout the literature on linguistic change (e.g., Labov, 1972; Eckert, 1988; Chambers, 1995). While these findings reveal that the correlation of linguistic change with females does operate in these very young sectors of the population, it does not do so consistently. Male/female differences do not pattern similarly in all age groups. Amongst the youngest cohort (the 10- to 12-year olds) sex differences are either non-existent or marginal. It is only amongst the older teen cohorts that the features become strongly correlated with females. These trends show that sex differences — at least with respect to these discourse/pragmatic features — are
developmental, and are learned. They do not appear to be endemic to the features themselves, but are created in the speech community, within the peer group.

Finally, and perhaps most surprisingly given the received wisdom, at least in the media, the use of *like* and *just* are not haphazard, random, or indiscriminate. Indeed, even *like* — so notorious for its ability to turn up just anywhere — may well be quite circumscribed. So far, these preliminary results show that *like* occurs before nouns, before full sentences, and before verbs, but rarely elsewhere, while *just* occurs overwhelmingly before verbs. Such focussing, or co-occurrence pattern is precisely the type of trend that would be predicted if these forms were becoming meaningful parts of the grammar. In other words, strong contextual constraints on language use, e.g., particular positions in sentence structures hint at deeper changes in the language. For example, the fact that *like* is occurring so often before a noun may mean that it is developing into some kind of nominal marker. Indeed, one of the main claims about its function is that it is a focuser (Underhill, 1988). While focusing particles are not part of standard English grammar, they are quite common and indeed elaborate in other languages. Similarly, the fact that *just* is occurring so often just before a verb may mean that it is developing into some kind of verbal marker. In fact, research has shown that English is developing a system of aspectual particles, a feature which adds additional information to the verb (e.g., Brinton, 1988). In other words, these forms, which have been thought of as being simply discourse or pragmatic markers, may well be in the process of becoming part of the grammar of a new variety of English. A trajectory of change from discourse to grammar has been noted in linguistic studies previously (e.g., Sankoff, 1980a), thus providing a model for extending the current study. In future research we will be turning to explore these types of hypotheses, and to testing the findings I have reported here in larger and more socially complex material.

For now, however, I would like to join with other researchers of teen language (e.g., Eckert, 1988, 2000, 1997; Stenström and Andersen, 1996; Andersen, 1997b, 2000; Erman, 2001) to emphasize the tremendous breeding ground for linguistic innovation that exists amongst the tweens and teens in contemporary, urban speech communities. The language of these speakers is a gold mine for innovative linguistic features, revealing evidence for both grammatical, as well as sociolinguistic change. However, which of the changes are here to stay, and which are subject to fashion, will have to await further research.

**Acknowledgements**

I would like to thank the enthusiastic audience for their many comments and suggestions for this research.

**References**


Buchstaller, Isabelle. He goes, and I’m like: the new quotatives re-visited. American Speech, in press.


Dailey-O’Cain, Jennifer, 1996. The sociolinguistic distribution of and subjective attitudes toward the discourse marker ‘like’ in the midwestern United States. Paper presented at Cardiff, Wales, UK.


Sali Tagliamonte is an Associate Professor of Linguistics at the University of Toronto, Canada. Her research focuses on morpho-syntactic variation and change in the evolution of English. She has worked extensively on North American varieties of English and is the co-author of “African American English in the Diaspora: Tense and Aspect” (Blackwell Publishers). Currently, she is engaged in multiple research projects. One investigates British dialects and involves cross-variety comparisons amongst British and North American dialects, another explores the nature of linguistic innovations amongst contemporary youth, and another focuses on changes in apparent time in Toronto English.