Yod-dropping in English – A World-Wide Web Survey

by Neil Wick for J.K. Chambers, March 2000

How can we learn about language change and variation over a wide area?

Since the study of language change became a common subject of study during the last century, many methods have been used to collect data. The first dialect survey with national scope used a postal questionnaire which was distributed beginning in 1876 (see Chambers 1998). Since that time, field workers have largely replaced written questionnaires for this purpose. In recent years, postal questionnaires have become the primary gathering tool in the various regions of Canada where the Dialect Topography project (Chambers 1994) has been working.

The Internet has spread rapidly in recent years and processes that would not have been practical in former times are now possible. The use of the Web made it possible in the space of three weeks to gather responses from English-speaking countries spread from the northern hemisphere to the south with coverage encompassing most regions of North America. This was done at almost no cost, although an incentive in the form of a $20.00 gift certificate to a lucky winner did form part of the strategy for attracting respondents.

Recruiting respondents

The contest was advertised on my Web page, which is affiliated with the online bookstore, Chapters.ca, so a gift certificate for Chapters was a good incentive for these people. This resulted in a 10% response rate from this source. Since people on this page are generally anxious to get shopping, such a contest

will just be a distraction for most of them. Two links were added to the top of this page to check on which link placement was more likely to be followed.

Offering a prize also allowed the contest to be featured on such sites as contestworld.com and treasurehunt.com. These sites were valuable, since they attracted not only casual visitors to those respective pages, but also subscribers to the newsletters of those sites, who were able to learn about the contest from their regular weekly email list.

The use of the contest also served another purpose; it tended to attract a broader range of visitors that just ones for whom language matters is a special interest.

Newsgroups are discussion groups where people can publish public questions and viewpoints on a particular area of interest. It is important to be sure that any message posted to these newsgroups is relevant to the topic at hand. At best, an off-topic post will simply be ignored. At worst, complaints against posters who post with no regard for the topic of the newsgroup could lose their Internet accounts due to complaints from people objecting to "spam." Infamous cases of the mid-1990s involved the posting of the same messages to hundreds or thousands of these groups. As much as this may bring results, it is not to be recommended, since it would not create a good image for the poster.

Other respondents were obtained from personal contacts and the members of the *Linguist List*, an e-mail discussion group with members world-wide who research various topics in linguistics.

Survey Design

There were several technical hurdles involved in getting the desired form for the survey. The problem with computers is that not every computer has the

same configuration or capabilities. The survey was designed to minimise the variables in presentation as much as possible.

Most web browsers currently in use have the ability to use a scripting language called JavaScript. The survey page was configured such that users with JavaScript capabilities received the actual survey on a new window of predetermined size. This window contained an instruction to reposition itself 20 pixels from the top of the screen and 20 pixels from the left. This was done to make it obvious to users that this window was new and that they could safely close it after they were done with it.

Resizing the new window helped to fit a similar number of words on one line for users who had the same font size. Some of the newest browsers will permit a Web page to specify and exact font size in pixels, but we can expect the user to have their default font sized in an appropriate size for them, and I deemed it best not to tamper with this. On an average-sized monitor set to a very high resolution, a 10-pixel font can be difficult to read, whereas the same 10-pixel font will be quite readable on a lower-resolution set-up where the actual pixel size is bigger.

Font style is one aspect of web pages which is impossible to completely control. Most browsers presently in use will accept suggestions as to which font they should use, but if the font is not installed on the user's computer, the suggestion cannot be carried out. The technology exists to have fonts automatically downloaded to a user's computer with the page, but not all users are currently able to take advantage of this. It also introduces an extra file, leading to a longer wait before the user can see the page as intended. The survey form was left with the user's default font, which is most commonly some form of Times Roman.

Another problem with fonts concerns underlining. Normally it's best to avoid underlining, but part of words were underlined on the original source survey which I wanted to replicate. I considered putting these parts of words in

capital letters, but I was afraid of these being interpreted as stress indicators. On some computers, bold characters are not differentiated in smaller text sizes. Since the underlining was relatively clear when tested on a Windows computer with Times Roman, I elected to stay with the underlines, but this caused problems with some Macintosh users, since underlines do not extend to letters with descenders.

Input types

Web forms can have several different types of inputs. These fall into two major categories: pre-determined selections, and free-form text. The first type allows only a limited number of pre-selected choices.

Selection types

With radio buttons, it is only possible to choose one of the possible choices. This type of input was used for the "Sex" question.

A second type of selection input is the checkbox type. In this case, the user is able to choose more that one selection for a single question. This type of input was used for the phonological questions. Although respondents were not told that they could pick two answers to the same question, I wanted to leave it open to them to give more than one response if they felt it was necessary. This form of selection also made the questions look more similar to the printed Dialect Topography questions I was trying to compare.

The third type of selection input is a drop-down menu. This was used on the attitudes section. It allowed the selected option to appear on the same line as the question, so that the respondent could read the response in context

Text inputs

Text inputs are slightly more difficult for the user, since most users will need to use a mouse to position the cursor on the text box and then switch back

to the keyboard to type the text. There were three required text inputs, which were all grouped together to isolate the keyboard activity to one area of the form.

I decided to use a text box for the "Age" question, since this allowed more flexibility in working with the data than a selection of age ranges would give. Any kind of list of all the possible selections for age seemed unwieldy. The text box for this question proved to be easy to deal with in processing the data, with the exception of a few respondents who declined to answer this question.

For the questions about place of origin and current address, I considered selection inputs. This would have allowed the forms to be processed more easily, since all of the responses would have been standardised in spelling and format. However, there are several drawbacks to this approach. Every country is divided up differently in terms of states, provinces or counties, and some people have lived in diverse places during ages 8 and 18, which could not easily be indicated using a selection input. The main drawback of this approach is on the part of the researcher. A great advantage to the collection of the data in electronic form is that it saves time inputting information into a database, but having the places in this form defies any form of electronic processing. The countries, states, and provinces had to be hand-encoded into each response.

Automatic inputs

A few useful pieces of information were obtainable with no intervention by the respondents. The domain name of the respondent's computer, the time and date, the browser information, and the referrer were recorded for each survey response.

The combination of the browser information and the domain name are fairly unique for each respondent and help to separate one respondent from another.

The domain name is also useful to disambiguate some of the responses to the address questions. Despite instructions to the contrary, some respondents gave less than complete information about their place of origin and place of residence. For example, if a respondent says he grew up in "Cambridge" and now lives in "London," we can't tell whether he's from Canada or the United Kingdom. However, if his domain name ends in "uwo.ca," we know he's using a computer at the University of Western Ontario in Canada and we can fairly safely assume that he's from Canada.

Most respondents were provided with a URL containing a question mark followed by two letters at the end. This two-letter "query string" served to measure the effectiveness of recruitment methods and could be used to separate the responses from certain identifiable groups. For example the URL given to members of the Linguist List contained a "?ll" at the end, making it possible to separate their responses from the others.

The questions

The main part of the survey consisted of seven pronunciation preference questions and four demographic questions. Three of these questions were designed to replicate and extend the results of the Dialect Topography survey of Canada. A fourth question from that survey was modified slightly to accommodate palatalisation responses which were expected from the United Kingdom and the southern hemisphere countries.

An optional second set of questions was offered to most respondents and consisted of four two-part questions on pronunciation attitudes. These questions were not based on any previous work; their design will be discussed along with the results.

The respondents

The most important geographical question was the place where respondents were raised between the ages of 8 and 18. This question was based on one of the questions in the Dialect Topography survey, in order to provide a sound basis for comparison. Where respondents indicated several places of residence during this period of their lives, the earliest was chosen as a basis for analysis in the present paper. Various factors influence an individual's pronunciation, including the place where they first learned to speak and the provenance of their parents. There is no doubt, however, that peers exert considerable influence on children during their time in school.

The responses for this childhood place of residence included more than 20 countries, with the largest number of responses coming from the United States and Canada. The respondents represent nine of Canada's ten provinces and 42 of the 50 states of the U.S. Smaller but significant numbers of people from the United Kingdom and from New Zealand also responded with the remainder from a scattering of English-speaking and even a few non-English-speaking countries.

Analysis of results

The first question in analysing the results concerns assigning indices to the variants whitch follow the alveolar oral stops /t d/. The fricative which sometimes follows these stops is analysed here as articulatory gesture intermediate to the stop which precedes it and the glide which may follow it. The distribution of variants in the data lends support to this analysis:

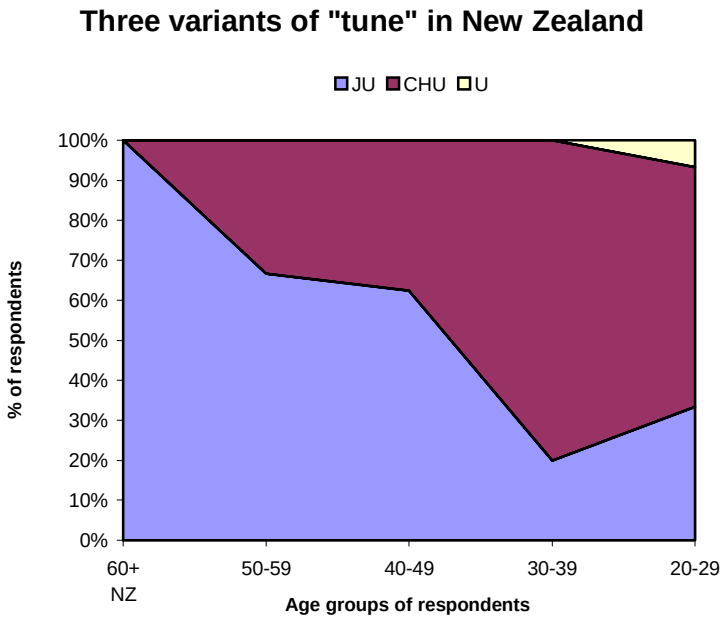

Figure 1

All of the alveolar oral stop data supports this ranking. The oldest respondents in Figure 1 have 100% /tjun/, the next two age groups have a gradually increasing proportion of /tʃun/, and only in the youngest group is the consonant simplified completely to /tun/. This ranking is repeated quite consistently thoughout the /t d/ data in this survey.

A spectrographic analysis of a Canadian subject saying /tjuzdɛj/ also seems to indicate a fricative following the stop. In figure 2, random high-frequency energy is clearly visible between the beginning of the word and the diverging formant contours indicating the glide from /j/ to /u/.

This analysis could possibly be extended to other coronals, including /n l/. The most gradual way of dropping the glide is to transfer the glide to a secondary palatalisation feature on the consonant causing frication, and this segment can then be reduced in length until it is gone.

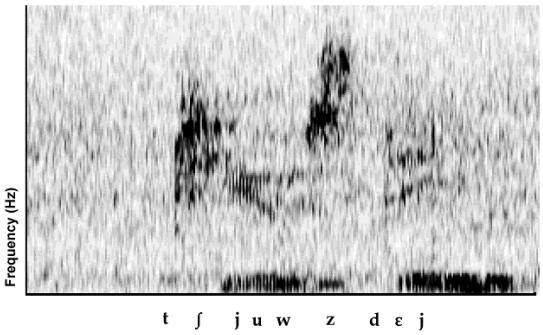

Figure 2

Given this analysis, indices were assigned as follows: /Cj/ - 0, /CF/ - 1, /C/ - 2. For example, (tu) - /tju/ - 0, /tʃu/ - 1, /tu/ - 2. For consistency, the /n/ and /l/ variables were analysed the same way, although no intermediate variable was tested for. One linguist who responded commented that he has a palatalised /ɲ/ rather than either /nj/ or /n/ in "news."

Respondents who gave two variants to a single question were treated as if they used each variant 50% of the time, leading to an index that averaged the two possibilities for that respondent.

Index of avenue		Country	Sex			
		CA	NZ	UK	US	Grand Total
AgeGrp	avenue					
50+		14	0	0	121	71
40-49		14	0	0	141	59
30-39		25	0	0	159	75
10-29		59	0	0	161	89
Grand Total		30	0	0	145	74

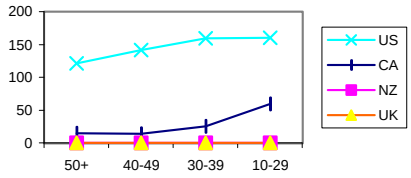

Figure 3a

Index of tune		Country	Sex			
		CA	NZ	UK	US	Grand Total
AgeGrp	tune					
50+		145	20	23	176	141
40-49		172	38	64	189	161
30-39		185	80	53	194	171
10-29		172	73	53	192	159
Grand Total		171	61	46	187	159

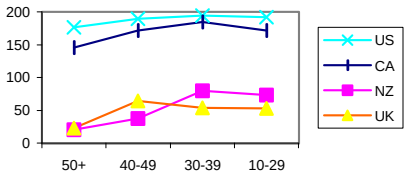

Figure 3b

Index of student		Country	Sex			
		CA	NZ	UK	US	Grand Total
AgeGrp	student					
50+		127	40	18	158	126
40-49		153	13	40	185	147
30-39		156	70	33	187	153
10-29		155	60	35	166	139
Grand Total		150	50	31	173	141

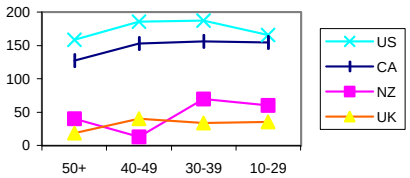

Figure 3c

Index of dune		Country	Sex			
		CA	NZ	UK	US	Grand Total
AgeGrp	dune					
50+		135	0	27	168	133
40-49		147	38	47	174	141
30-39		148	100	53	176	149
10-29		155	53	50	175	144
Grand Total		148	55	43	173	142

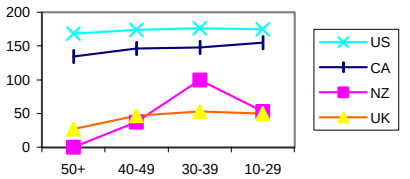

Figure 3d

Index of news		Country	Sex			
		CA	NZ	UK	US	Grand Total
AgeGrp	news					
50+		117	40	9	171	129
40-49		159	0	0	190	148
30-39		157	20	13	183	147
10-29		163	13	10	191	146
Grand Total		153	16	8	183	143

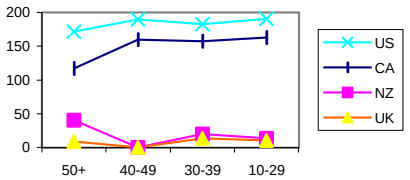

Figure 3e

Index of lube		Country	Sex			
		CA	NZ	UK	US	Grand Total
AgeGrp	lube					
50+		166	200	141	183	173
40-49		161	200	138	183	170
30-39		141	200	160	177	161
10-29		160	173	180	176	170
Grand Total		156	189	156	180	168

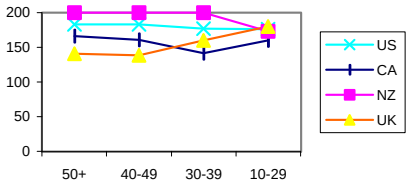

Figure 3f

The variable showing the strongest apparent-time change is the unstressed /u/ of "avenue" (Figure 3a). All respondents in the U.K. and New Zealand report 100% retention of the [ju] diphthong, but North Americans are currently in the process of dropping the [j].

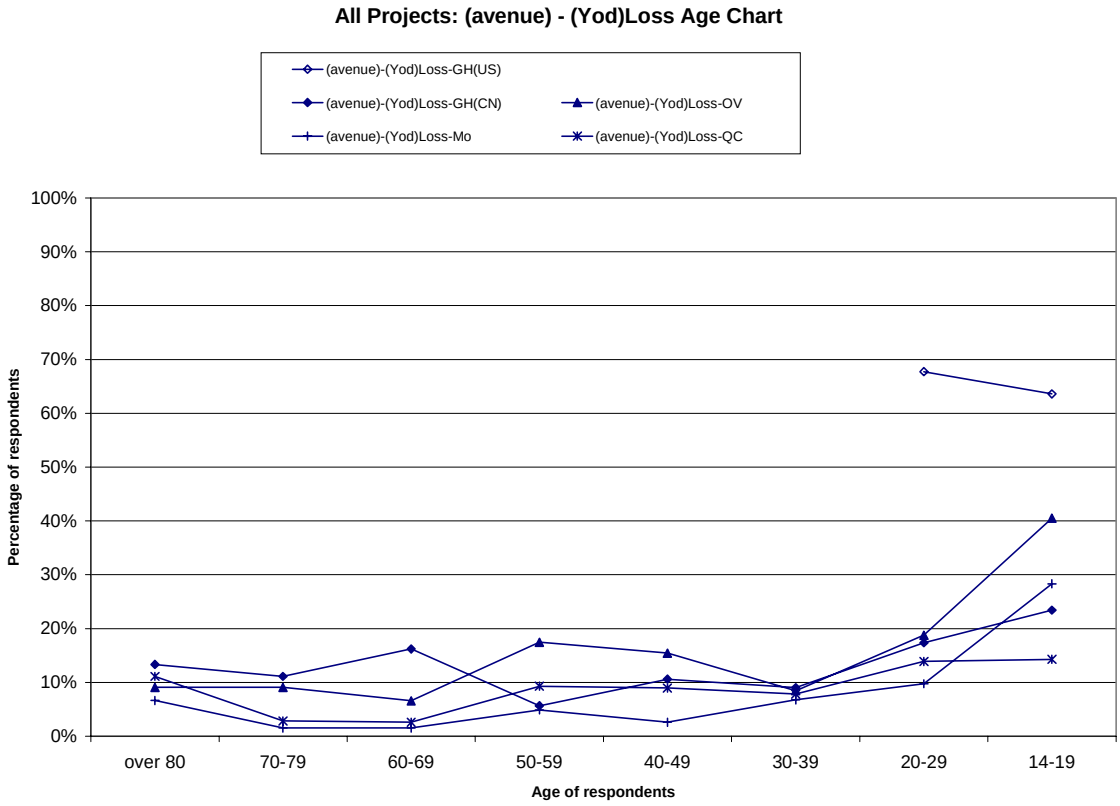

Figure 4 – Yod-dropping in "avenue" from the Dialect Topography project (MacKeracher 2000)

The data from the Dialect Topography project in Figure 4 is very similar to my findings in Figure 3a. Although the variant appears stable across most of the graph, the beginnings of an s-curve are apparent and show up even more clearly in the data from the Web survey.

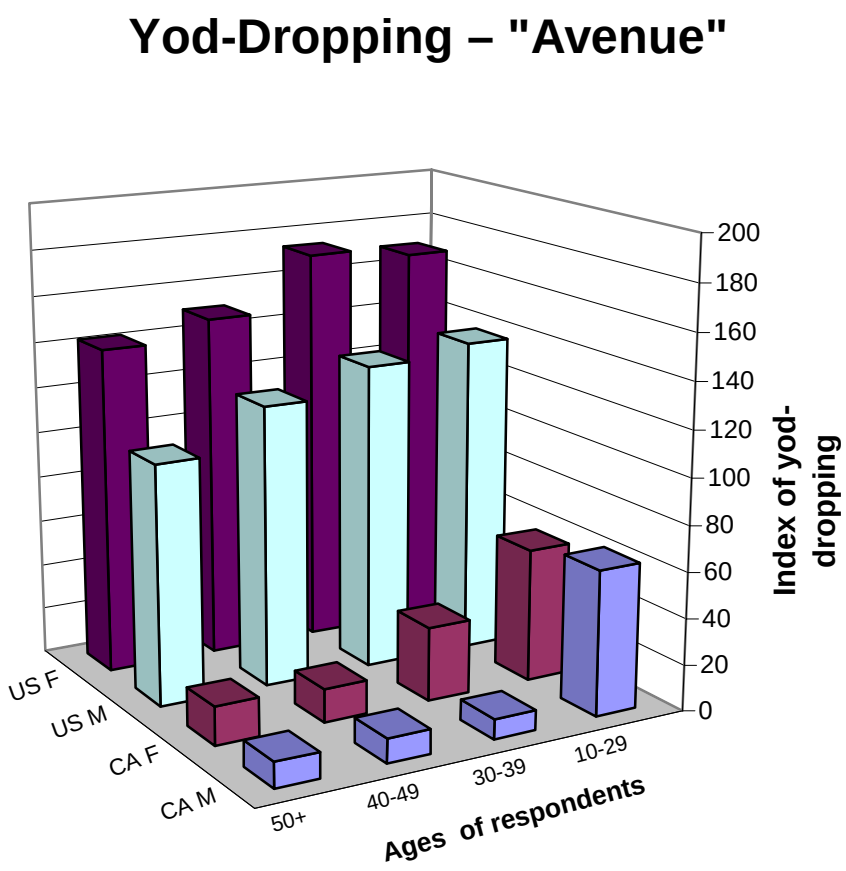

Figure 5 – Comparison of U.S. and Canadian males and females in yod-dropping of “avenue”

As figure 5 indicates, females seem to be leading the change in this variable in both the United States and in Canada. Although Canadians are at a much less advanced stage than Americans, it does appear that there may be influence from the United States in this variable.

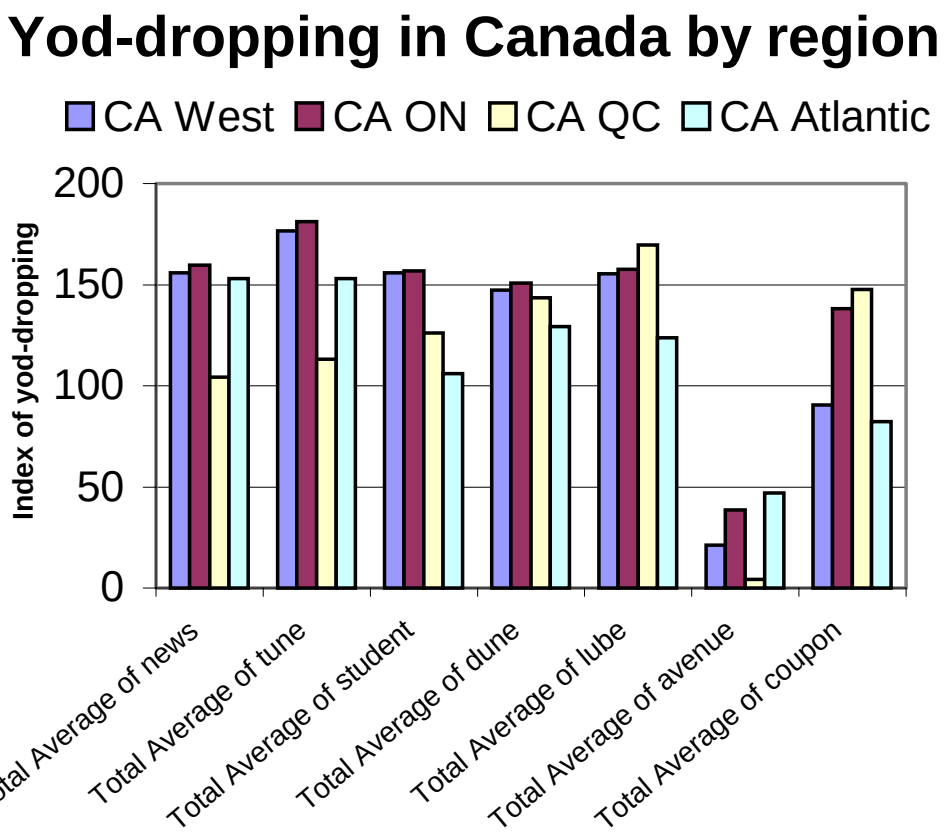

Figure 6

Following the regions of Canada from west to east, the yod-dropping in "avenue" increases with one anomaly. Quebec is not participating in this trend. This would be probably due to the influence of the French language in that province. Looking back at figure 4, you can see that the French-speaking areas are again the least affected by this change.

References cited:

Chambers, J.K. (1998) "Social Embedding of Changes in Progress." *Journal of English Linguistics*. 26: 5-36.

Chambers, J.K. (1994) "An Introduction to Dialect Topography" *English World-Wide*. 15: 35-53.

MacKeracher, Mary S. (2000) Unpublished research. University of Toronto.