Applied catastrophe theory in dialectology – evidence from Canada

Neil Wick
For J.K. Chambers
LIN 452

Neil Wick, April 16, 2001

Since mathematician René Thom first proposed catastrophe theory in the 1960s and 1970s, it has been applied to many fields. Catastrophe theory is a mathematical model which describes discontinuous leaps in dynamic systems. It has been proposed by Schneider (1997) that many of the concepts of chaos and catastrophe theory appear to correspond to properties of dialect variability and change. The purpose of this paper is the application of catastrophe theory to some aspects of the data collected in the Canadian Topography project.

Dialect Topography

The Dialect Topography project (Chambers 1994) consists of several regional surveys of Canadian English using postal questionnaires. Data from the first four project regions were available for the present analysis. Two of these regions are located in the Province of Ontario, where the majority of the population speak English and the other two regions are located in the Province of Quebec, where French is dominant and the English speakers are a minority.

The first project region, known as the Golden Horseshoe, is ideal in many ways for the application of catastrophe theory. The region extends along the north shore of Lake Ontario from Oshawa, just east of Canada's largest city, Toronto, and continues around the western end of the lake and across the Niagara River on the south side of the lake into the United States. Since the population density tends to be higher closer to the lake, this region can be viewed as largely a one-dimensional continuum with minimal influence from outside areas.

The other interesting aspect of this region is the border with the United States. The Niagara River forms a political boundary as well as a formidable physical barrier. Since the river has formed a deep gorge for half of its length and contains a large waterfall, making nearby river navigation hazardous, the only practical means of crossing the river for most people is by way of several bridges.

In addition, the area east and west of the river is a peninsula bounded by Lake Ontario in the north and Lake Erie to the south, so it forms a bottleneck with no other easy access across the border for several hundred kilometres to the east and west. The Golden Horseshoe survey was completed in 1991-92, with 935 respondents from the Canadian side and 80 from the U.S.

The second Ontario region to be surveyed was the Ottawa valley. Located along the border with Quebec and Canada's bilingual capital city, the Ottawa valley contains a more significant French presence than the Golden Horseshoe, though the majority of the population are still English-speakers. 681 questionnaires were collected from this region in 1997.

The Quebec City area project was completed in 1994 with 307 questionnaires received. Although the Anglophones form a small minority in this city of French culture, they have a significant amount of history in the area (see Chambers & Heisler 1999 for details).

The newest survey region analyzed in this paper is Montreal, with 500 questionnaires received during 1998-99. Montreal, Canada's second-largest city, has a sizeable English-speaking population, although the number of Anglophones has declined in recent years, largely due to the language policies of the Quebec provincial government, which are designed to protect the French language in the province.

Respondents were asked 11 demographic questions (age, sex, education, place of residence, etc.) In all regions except the Golden Horseshoe, there was an additional four-part question concerning English usage frequency. Some of these questions will be discussed later in this paper.

Following the demographic questions, the questionnaire contains questions on 81 linguistic items, involving pronunciations, vocabulary, morphology, syntax, and usage (Chambers & Heisler 1999:26). The responses that proved to be most interesting from a catastrophe theory perspective involved items which vary regionally, although other independent variables from the personal demographic questions are also involved.

The Basics of Catastrophe Theory

Catastrophe theory contains terminology and concepts which are likely to be unfamiliar to the linguist, so a detailed introduction is in order.

Fundamental to catastrophe theory is the concept of stability. A system can exist in a stable state, and semi-stable state, or under other conditions, no stability at all (Woodcock & Davis 1978:41). Picture a ball-point stick pen. A typical one has a hexagonal cross-section, a triangular point (ending in a ball) at the writing end, and is fairly flat on the top end. It is perfectly symmetrical in the longitudinal direction, so it should be possible to place it on a table resting on the writing point, so long as it is perfectly balanced. If it is not perfectly balanced, the condition is unstable. The pen can never remain in this position without support and will fall rapidly to a horizontal position. Even though it is theoretically possible to balance the pen on its point, if you ever accomplished it, the condition would not exist for long, since the slightest disturbance would knock it over – it is semi-stable. It is probably possible to place such a pen on its other end, especially if that end is fairly flat, but this position is also only semi-stable. It takes very little force to knock it over. Lying on its side on a level table, the pen is stable. Only a strong external force will move it from its place.

Such stable and unstable positions are commonplace in language. Features tend to be attracted to certain stable positions known in dynamic systems theory as *attractors*. For example, one parameter differentiating vowels is tongue-height, but certain points on the continuum between high and low constitute attractors with the possible tongue-heights defining a certain vowel being restricted to only three or four.

Many mathematicians have found it helpful to visualize the results of differential equations, so the science of topology has become important to types of theories we're discussing here. "The history of any dynamical system can be mapped as a trajectory in a multidimensional space ('phase space') where each point in the space represents a possible system-state" (Lass 1997:293). Each dimension of the phase space represents a parameter or constraint which affects the state of the system.

A system of two dimensions (two "degrees of freedom") can be mapped onto a plane surface. Typically, only "four types of behaviour are possible in a system of differential equations in the plane ... exceptions are infinitely rare" (Stewart 1989:103). These four types of behaviour revolve around four types of features.

The first type of feature is known as a *sink*, which consists of a single, stable point. If you start the system at some point near to the sink, it will move towards it (often in a spiral). When it reaches the sink, it stays there. A linguistic example might be the merger of the vowels in "cot" and "caught." Once they are fully merged, there is no possible way for them to be unmerged and restored to their former distinctions. Briggs and Peat (1989:151 quoted by Schneider 1997:25) call this the *irreversibility of processes*.

The second type of feature is a *source*. This is a single point, like the sink, but it is unstable. Any system that starts at or near a source tends to move away from it. Even though the state of the source may be steady and unmoving, the behaviour of nearby flows causes points near the source to be unstable.

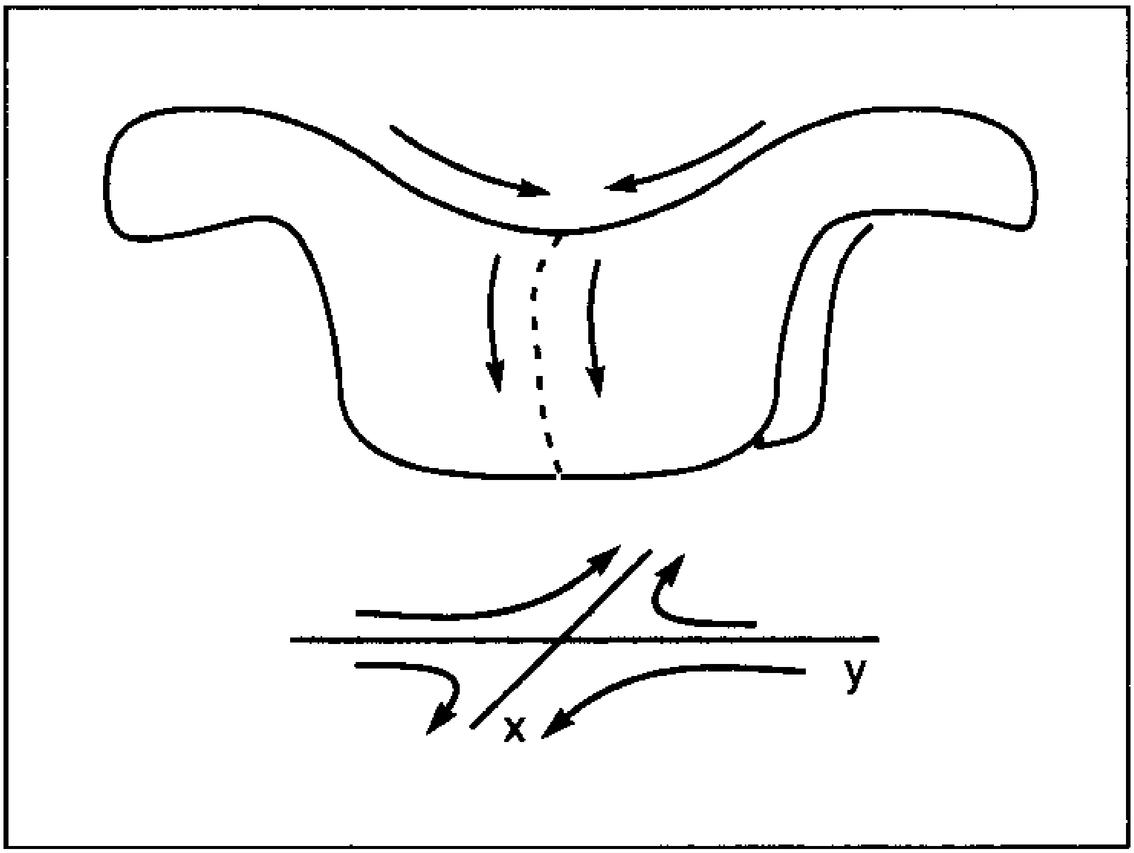

Figure 1: The saddle (from Alligood et al 1997:64)

The third type of feature is a *saddle* (Figure 1). This combines some features of the source and the sink. It is stable in some directions and unstable in others. Two flow lines which meet at the saddlepoint are known as the separatices of the saddle, one directing the flow towards the saddlepoint and the other one directing the flow away from it. A flow which follows alongside the incoming separatrix turns towards the direction of the outgoing one as it approaches the saddlepoint. The saddlepoint is the third type of steady state.

The fourth and final feature is known as a *limit cycle*. Limit cycles are a little more complicated that the other three feature type, because they don't feature a steady-state point. Instead, they consist of a closed loop. If you start on a limit cycle, you go around forever in a periodic motion. There are two basic types of limit cycles. The stable type attracts nearby points, while the unstable type repels them.

Some examples of linguistic processes which seem to represent periodic cycles are given by Lass (1997:299-300). One example is "the constant emergence and loss of front rounded vowels in the history of English. Neither Proto-Germanic nor Northwest Germanic had /y, ø/; but these emerged in prehistoric Old English as the result of *i*-umlaut of */u, o/ ..., were lost during Old English by merger with /i/ and /e/, re-emerged in Middle English, were lost in the southern standard varieties in late Middle English, and are becoming increasingly common in modern dialects." Lass takes front-rounding as a limit cycle value, giving a pattern of recurrence.

Thom's "classification theorem" of 1965 stated that when a system is governed by no more than four different factors, there are only seven structurally stable ways for it to change discontinuously. Each of these seven elementary catastrophes can be depicted on a graph that has one dimension, or axis, for each *control factor* that determines a system's behaviour and an additional axis or two to represent the behaviour itself (Woodcock & Davis 1978: 43). The simplest of these elementary catastrophes has one control axis and one behaviour axis and is two-dimensional, while the most complex has four control axes and two behaviour axes, producing a six-dimensional result (Table 1).

Control variables	**State variables**	
Number of control factors (codimension)	**One behaviour axis (cuspoids, corank 1)**	**Two behaviour axes (corank 2)**
1	Fold	---
2	Cusp	---
3	Swallowtail	hyperbolic umbilic, elliptic umbilic
4	Butterfly	parabolic umbilic

Table 1: The seven elementary catastrophes (after Woodcock & Davis 1978:43, with alternative terminology from Saunders 1980:31-32)

These are "all the different cases which can arise with codimension no greater than 4" (Saunders 1980:31). You might ask how different cases are determined. Saunders says, "Two catastrophes are equivalent if one can transformed into the other by

(i) a diffeomorphism of the control variables, and
(ii) at each point in the control space a diffeomorphism of the state variable."

In general, two objects are *homeomorphic* (topologically equivalent) "if one can be continuously transformed into the other without any tearing or pasting together." Two objects are *diffeomorphic* if they are homeomorphic *and* the transformation involves no creasing or flattening of creases. A sphere, an ellipsoid, and a cube are all homeomorphic, but only the sphere and the ellipsoid are diffeomorphic.

The four canonical cuspoids can be described by the following "universal unfolding" equations:

Codimension	Name	Universal unfoldings
1	Fold	$x^3 + ux$
2	Cusp	$x^4 + ux^2 + vx$
3	Swallowtail	$x^5 + ux^3 + vx^2 + wx$
4	Butterfly	$x^6 + tx^4 + ux^3 + vx^2 + wx$

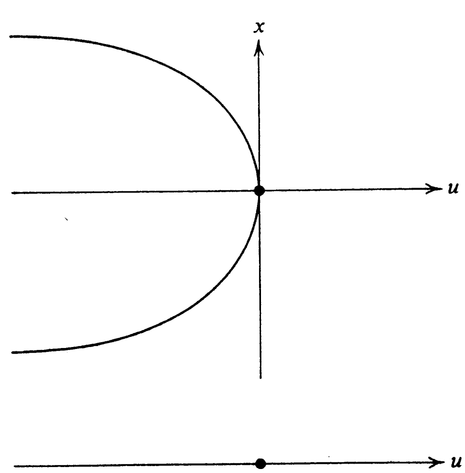

Figure 2: The fold catastrophe (from Saunders 1980:42)

Only the fold catastrophe and the cusp catastrophe are considered in this paper, so these will now be defined in greater detail.

In a fold catastrophe (Figure 2), the potential is $V(x)=x^3+ux$, so the phase space is two-dimensional. *M* is a manifold smooth surface which contains all of the points of equilibrium. The equilibrium surface *M* is the curve $3x^2+u=0$. If x>0 then the manifold represents a local minimum which attracts nearby points. If x<0 then the manifold represents a local maximum which repels nearby points. This part of the curve is considered inaccessible.

The bifurcation set is the place where jump behaviour occurs. The bifurcation set for a fold catastrophe is the single point $u=0$ (one-dimensional). If $u<0$ then there are two equilibria, one stable and one unstable. If $u>0$, then solving for equilibrium results in no real solution – no real equilibrium exists. In practice, there is usually another nearby equilibrium (such as a line parallel to the u vector) that the system can jump to.

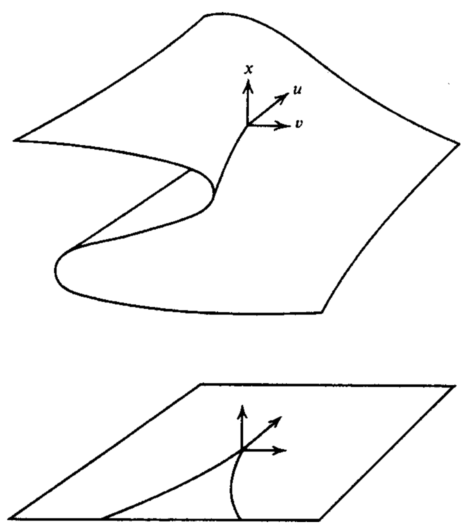

Figure 3: The cusp catastrophe (from Saunders 1980:43)

In a cusp catastrophe, the potential is $V(x)=x^4+ux^2+vx$ so the phase space is three-dimensional. It looks like a rubber sheet with a kind of a rounded z-shaped profile at the front edge. Travelling around the flatter back part of the sheet allows us to go from the upper to lower surface smoothly, but moving in the direction of the *v* vector near the front of the sheet causes a discontinuous jump from the upper to lower part of the surface. The equilibrium surface *M* is $x^3+2ux+v=0$. The top and bottom of the z-shape are local minimums and attract nearby points. The centre part corresponding to the diagonal part of the Z is a local maximum and repels nearby points. The region under the curves is inaccessible. The curve is shaped like two fold catastrophes interlocked.

The bifurcation set $8u^3+27v^2=0$ is two-dimensional. The projection of the bifurcation points onto a flat sheet outlines the triangular area with arcs on either side that gives the cusp catastrophe its name.

When we decrease *u* to negative values discontinuities in *x* can occur.

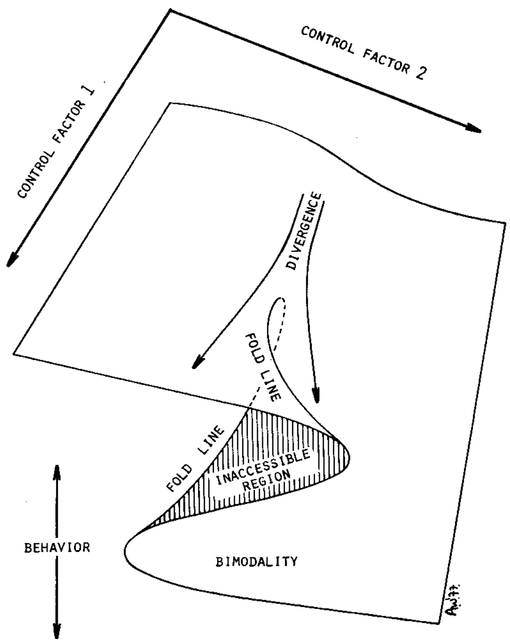

Figure 4: Divergence and the cusp catastrophe (Woodcock & Davis 1978:46)

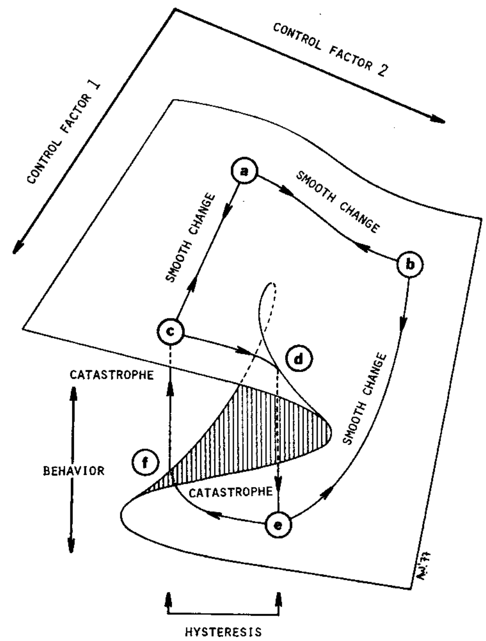

Figure 5: Paths on the cusp (Woodcock & Davis 48)

A point approaching *u=0* parallel to control factor 1 (Figure 4) will diverge towards the upper or lower surface of the manifold surface, depending on which side of the bifurcation point it approaches.

In Figure 5, various smooth paths are shown, along with two places where jump behaviour may occur. Paths travelling towards the left of the diagram jump at a different place than path travelling to the right. Paths travelling to the right jump at location d, whereas paths travelling to the left along the lower sheet will jump at location f. This delay gap is known as *hysteresis* (from a Greek word meaning “to be behind”).

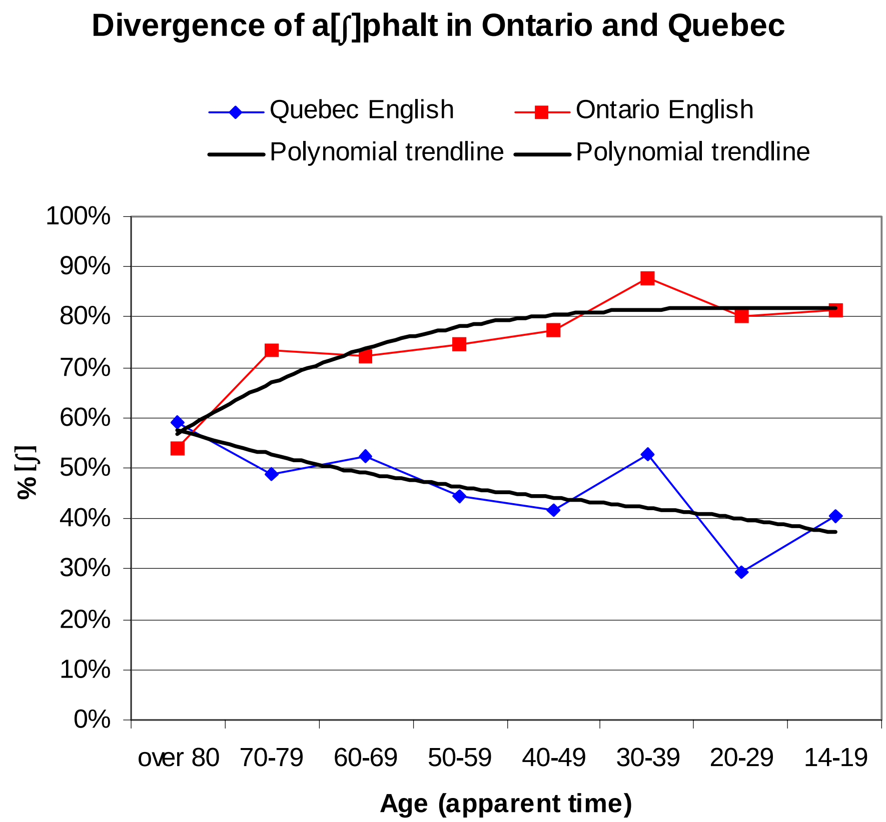

Figure 6

The variation in pronunciation of (asphalt) in the four earliest study areas of the Dialect Topography project is a good example of a divergence. MacKeracher (2001: 114-121) has noted that trends in Quebec City and Montreal are diverging from those in Ottawa and the Golden Horseshoe. We can conveniently group these into responses from the Province of Quebec (Montreal and Quebec City) and responses from the Province of Ontario (Ottawa and the Golden Horseshoe) (Figure 6). There is not only a sharp stratification between the two Ontario samples and the two Quebec samples – the Ontario usage of [ʃ] in this word is trending upward, while Quebec usage is in a decline. Using an apparent-time scale, it is easy to see that the two lines started out basically congruent – there was no difference between the prevalence of [ʃ] in the province of Ontario and in the province of Quebec. It is clear from the graph in Figure 1that these two populations are now drifting apart with regard to this variable. The cubic least squares fit trendlines in figure 6 bear a striking resemblance to the lines of divergence illustrated in figure 7. The cause of the split may very well be related to the prevalence of English in the areas under study as we shall see shortly.

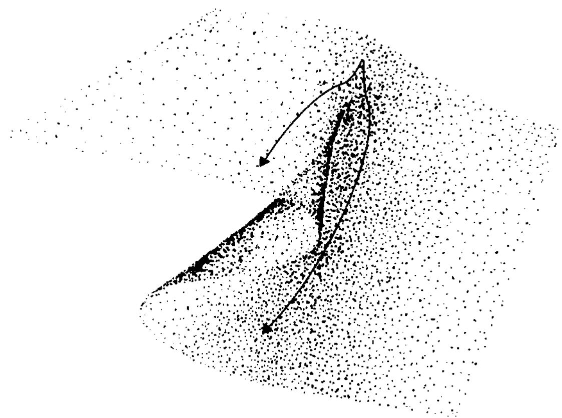

Figure 7: Divergence and the cusp catastrophe (from Poston & Stewart 1978:85)

The trendlines emphasize the general trend in each case, but it will be observed that the fit of the data to these trendlines is not very good, particularly in the case of the Quebec one. The Ontario trendline ($y = 0.001x^3 - 0.0215x^2 + 0.1576x + 0.4321$) is reasonably reliable, with an r^2 value of 0.8244 (where 1.0 represents a perfect fit) but the Quebec trendline ($y = -0.0006x^3 + 0.0104x^2 - 0.0767x + 0.6436$) has an r^2 value of 0.5421 – a particularly bad fit.

Chambers & Heisler (1999:29) noted this erratic pattern in the Quebec City portion of the data (Figure 8): "Age is incoherent as a determinant, with inexplicable differences between contiguous age groups, such as 60-year-olds and 50-year-olds

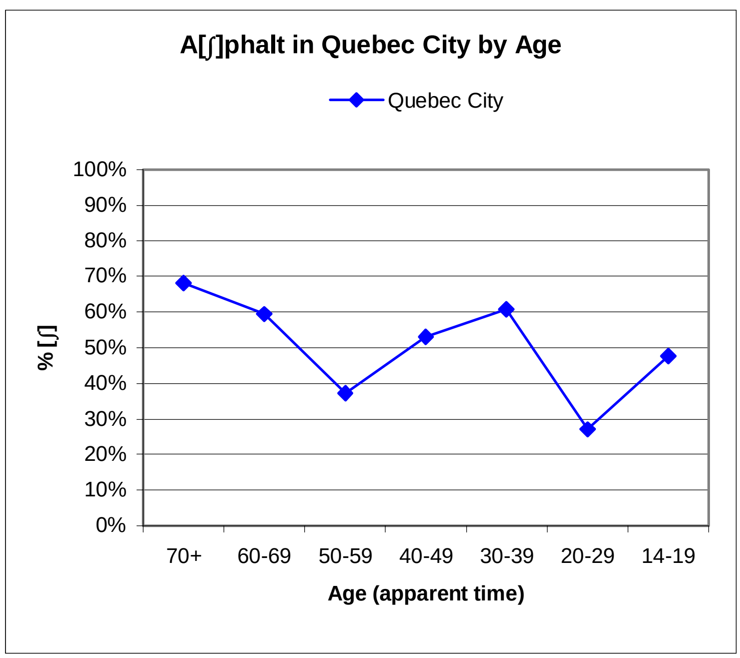

Figure 8

(59 per cent to 37 per cent) and 30-year-olds and 20-year-olds (62 per cent to 27 per cent)." While it is certainly true that these differences do not contribute to any kind of smooth function, there does seem to be a pattern to the oscillations, and it is possible that chaos and catastrophe theory may offer some explanation.

Analysis of the data for the whole province of Quebec with the GoldVarb program (Rand & Sankoff 1990) shows that four variables are statistically significant with respect to (asphalt): language use index, education, regionality index, and sex.

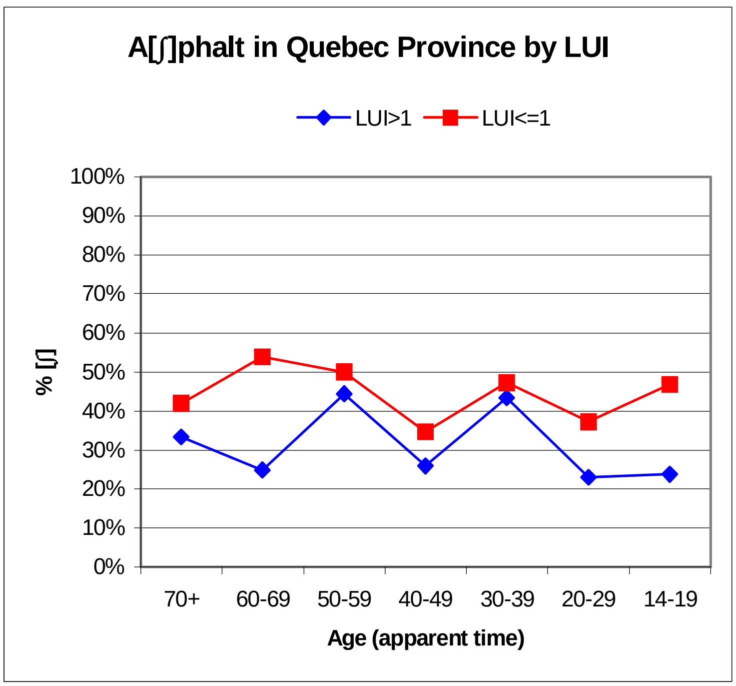

Figure 9

Figure 9 shows that (in every age group) there is a higher incidence of [ʃ] in asphalt among Quebeckers who use English frequently than there is among speakers who use English less frequently. The Language Use Index or "LUI" is higher for speakers who use English in fewer domains. For details, see Chambers & Heisler (1999). There were no speakers over the age of 79 with an LUI of greater than 1, so they were combined into a 70 and over group, but it's interesting to note that the two LUI groups follow the same zigzag path across most of the graph. With the exception of the 60-69 age group, when one line turns upwards, they both turn upwards and when one line turns down, they both turn down. The total sample size for 60-69-year-olds with LUI greater than 1 is only 12 people, so the margin for error is quite high in this group.

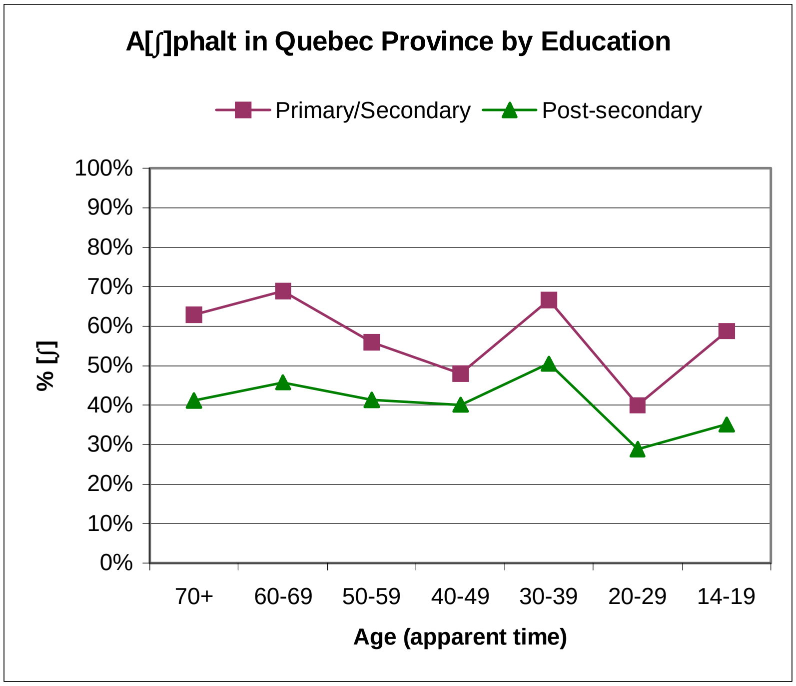

Figure 10

Looking at education (Figure 10), the data make a consistent oscillating pattern similar to the LUI pattern. Again, when one portion of the population goes up, the other goes up at the same time, and when one goes down, they both go down. The steepest jumps are in the youngest four age groups and these are also the most consistent. It is possible that the distribution of respondents is not ideal in the upper age groups, although this doesn't seem to have any significant effect on this education graph.

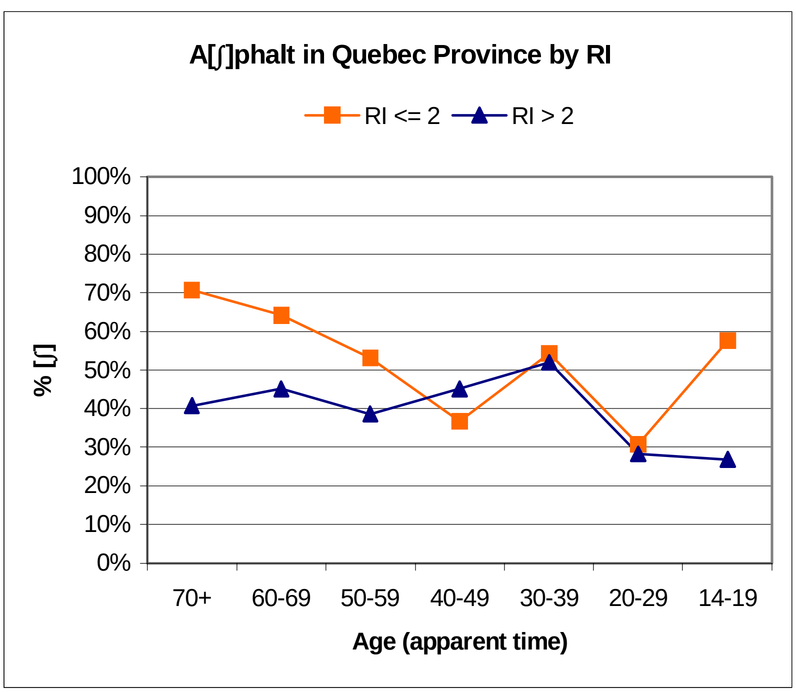

Figure 11

Regionality index (Figure 11) was a third variable identified by GoldVarb as significant for this question. The stratification in this case is not as clean as in the two cases we previously looked at, but there is only one cross-over – the 40-49-year-olds. The trends are still fairly consistent, with the oscillations in both groups of regionality indices following similar paths.

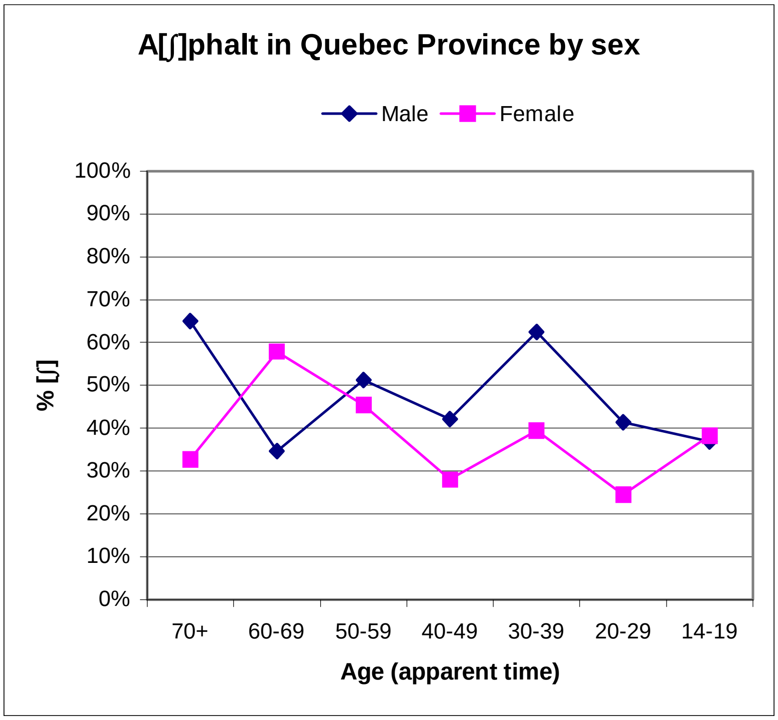

Figure 12

The last significant variable to be considered is that of sex (Figure 12). In this case the youngest four age groups are again affected similarly by the oscillations, but there is a strong cross-over in the 60-69 age group. There is also a small cross-over in the youngest age group, but there are only 19 males in this group, so one changed response in this group (e.g. one person changing from "no" to "yes") would make a more than a 10 per cent difference in the result.

MacKeracher (2000:116) solved the problems of the large fluctuations by grouping responses into four age groups rather than eight. As you can imagine, seeing that many of the vectors in the graphs we've just seen seem to change direction every ten years, grouping responses into 20-year blocks helps greatly to smooth out the fluctuations. The question is whether smoothing out the fluctuations is really what we want to do in this case.

MacKeracher says that (asphalt) is changing, but not in the way of a standard S-curve diffusion pattern. She concludes (117) that "the changes to (asphalt) appear to be languidly drifting, as if the Canadian population is indecisive and reluctant to change." While I agree that the populace is indecisive on this change, it seems from the evidence in figures 9 through 12 that (asphalt) is doing anything but "languidly drifting." I suggest that it is jumping back and forth between two extremes, particularly for speakers in a Francophone environment who have frequent occasion to use French. Note in figure 9 that the oscillations are much sharper for speakers who use less English in their daily lives.

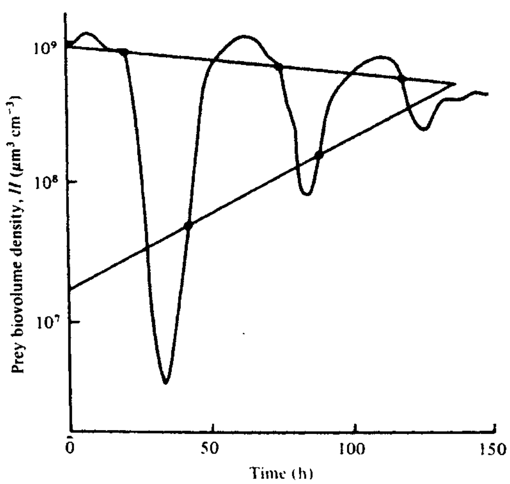

Figure 13: Prey biovolume (Saunders 1980:108)

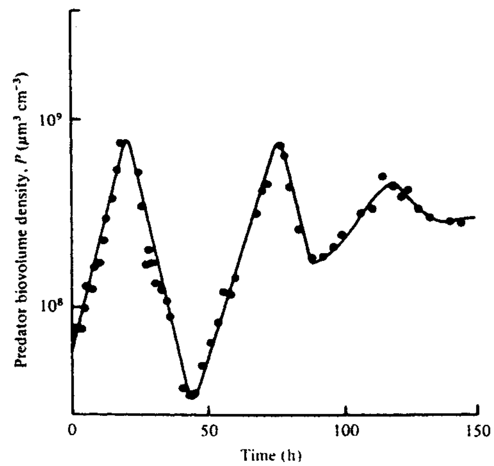

Figure 14: Predator biovolume (Saunders 1980:109)

This type of zigzag oscillation is classic behaviour in a cusp catastrophe, according to some theorists. An experiment which measured volumes of the bacterium *E. coli* (Figure 13), being preyed on in a laboratory by the amoeba *Dictyostelium discoideum* (Figure 14), showed sudden jumps across a triangular cusp-like area. (For details, see Saunders 1980:106-111 and references therein). Particularly striking is the very constant growth rate exhibited by the prey (Figure 14), which hits a certain point and then changes abruptly in a straight line in another direction. As time progresses, the jumps become progressively smaller until they disappear altogether as the system finally reaches a stable equilibrium.

Turning from biology to more familiar ground in the social sciences, we can look at the study of prison disturbances in 1972, carried out by E.C. Zeeman *et al.* and published in 1976. (See also the discussion in Poston & Stewart 1978:416-418, summarized here.) The relevant factors which eventually led to a prison riot were divided into two groups: tension (frustration, distress) and alienation (division, lack of communication, polarization). These two factor groups were set up appropriately on the catastrophe manifold (Figure 15a) and *feedback flows* (indicated by arrows) were added to fix the dynamical behaviour of the system. For low levels of alienation, the system tends towards stable position of neutral disorder, but as alienation increases, it oscillates back and forth within the bifurcation set, jumping from the top sheet to the bottom and back. Statistical factor analysis gave objective measures of tension, such as the number of men reporting sick each week, and of alienation, such as the numbers of inmates in the segregation room.

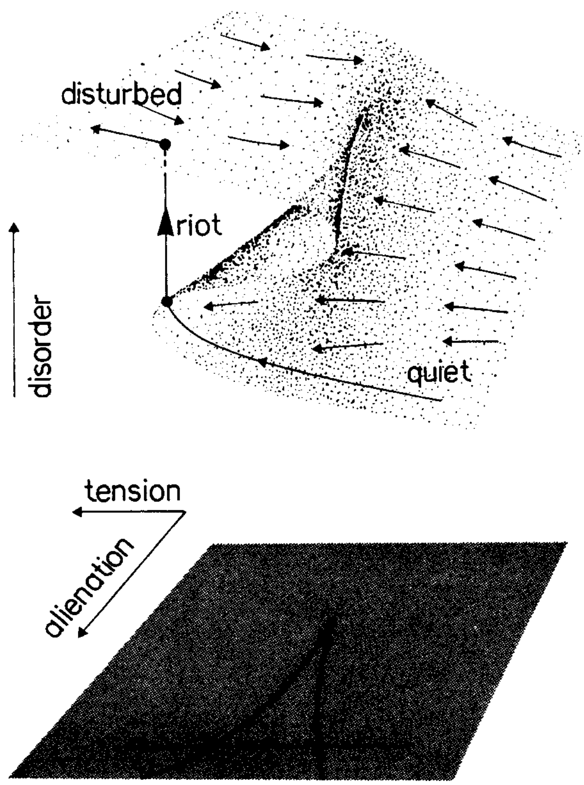

Figure 14a: The cusp catastrophe and the Gartree prison riot (Poston & Stewart 1978:417)

The resulting plot (Figure 15b) gives a clear view of oscillations within the cusp area (actually two cusp areas, for reasons explained in the original paper). When the track finally crossed the cusp line, a full-scale riot broke out.

We might imagine a similar oscillation for our (asphalt) data. As a certain segment of the population tries to accommodate to the other members of the community, the behaviours of the members of the group jump back and forth across the hysteresis gap seeking a stable resting point.

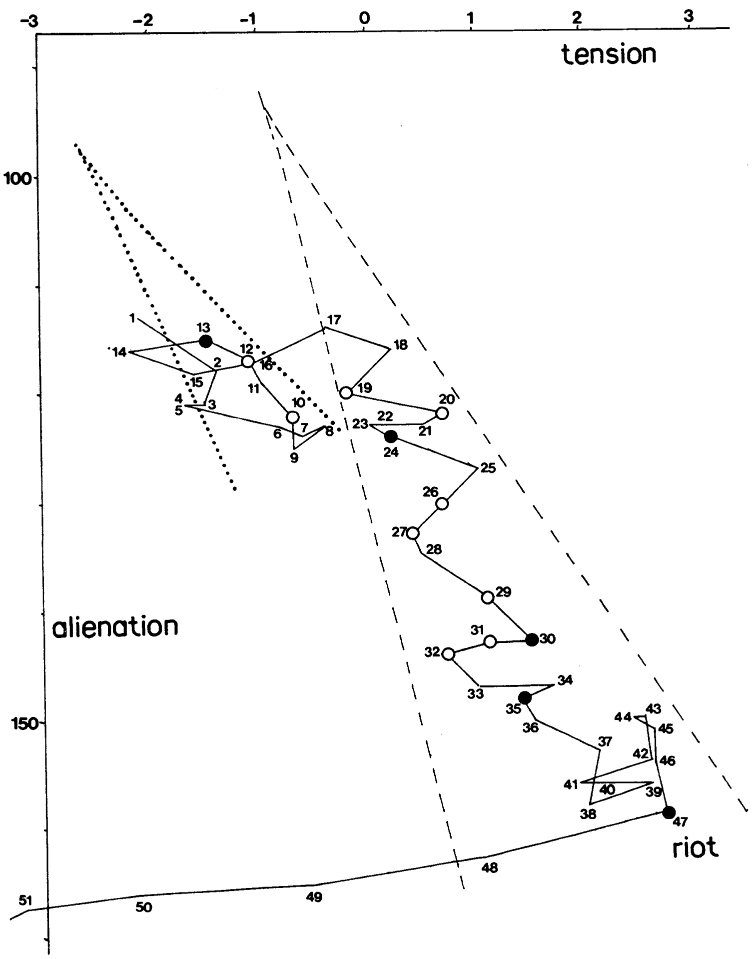

Figure 15b: Data plot for the prison riot (Poston & Stewart 1978:418)

Unfortunately, the granularity of the age groups is not fine enough to confirm the significance of the jump behaviour, but if we assume oscillations, there is no particular reason that they should happen on a ten-year cycle as you might infer from the graphs. Also, it is worth pointing out that points we are plotting are the average usage of the variable for each ten-year slice of time. Averaging each ten-year period has no noticeable effect on a smooth function, but averaging violent oscillations tends to smooth them out, so we can assume that if oscillations exists, they could be much deeper than the ones on the graph in figure 9.

It is interesting to note that there could finally a split developing in the two youngest age groups. The group with LUIs larger than one appears to have stabilized somewhat in the youngest age group without the sharp upturn characteristic of many of the preceding points on the graph in figure 9. This would be consistent with the general divergence happening between English-speaking Ontario and bilingual Quebec.

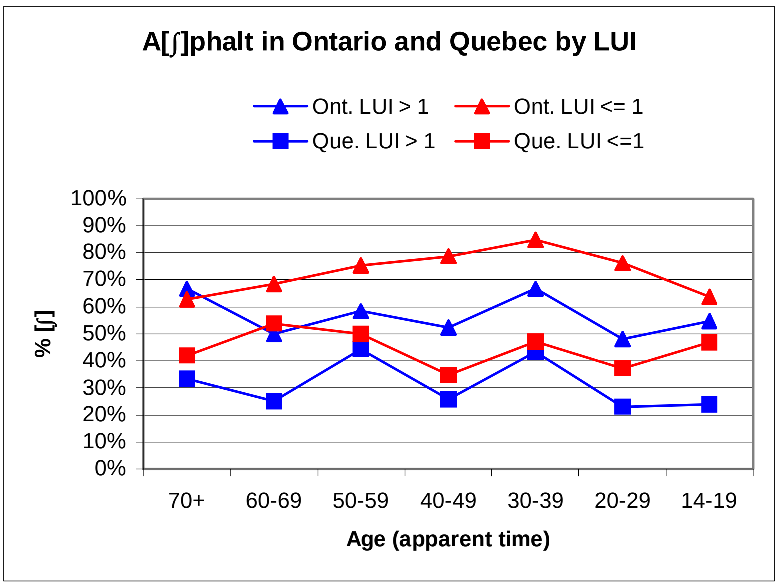

Figure 16

When we split the Ontario speakers by LUI (Figure 16), we see that the fluctuations of the respondents with the lowest use of English mimics those of speakers in the province of Quebec. (Golden Horseshoe respondents were not asked about language use, but it's fairly safe to assume that few of them use French on a regular basis, so they have all been placed in the low LUI group. Statistics Canada figures from 1996 show that only 0.64 per cent of Toronto residents speak French at home.) The Ontario respondents with an LUI of 0 or 1 defy explanation, but it is safe to say that they are not following the other groups shown on the graph.

The Golden Horseshoe and U.S. Border

Having examined a case of divergence over time, we now turn to case of stable variation. It has been observed at various times that Canadian speech differs from that of the United States in various ways. While some variables are changing in the same way on both sides of the Canada/U.S. border, in other cases stable differences are maintained. The question to be examined is how the Canadian/U.S. border is formed linguistically. Basically stable variables were chosen for analysis.

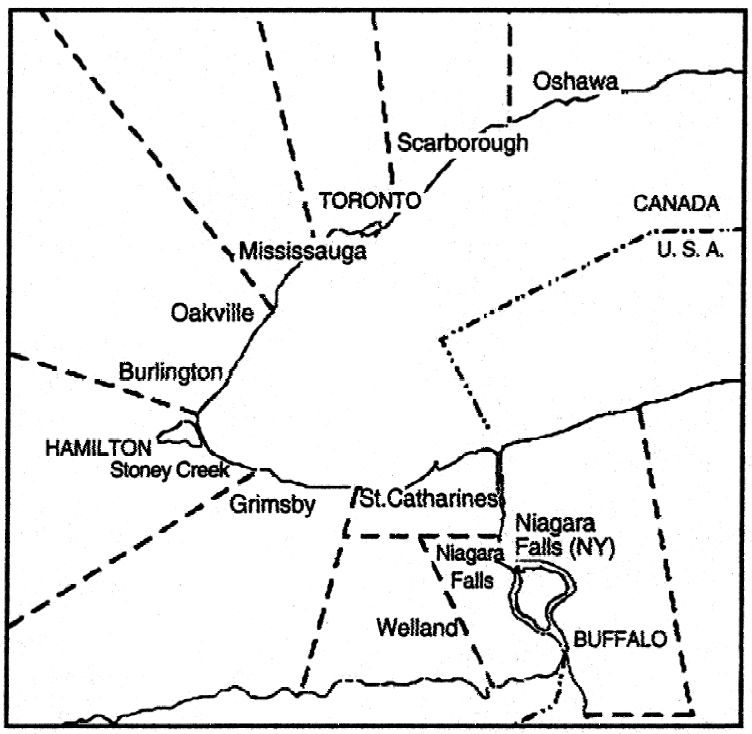

Figure 17: The Golden Horseshoe study area

The main geographical variable in the Golden Horseshoe area is the continuum around Lake Ontario. This can be considered very close to a single dimension of variation with minimal interference from other directions, as mentioned earlier. The goal in this case is to model a system with only one major control factor.

Schneider says, “real-life systems are defined not by a single variable but by an interaction of several variables, some of which are essential for the structural stability, while others are not” (1997:27). Although it is undeniable that several variables interact to produce the variation that we find in the study area, in the linguistic variables under consideration, the geographical variable is overwhelmingly dominant, so that variation caused by other factors such as age, social class, etc., can be largely considered as noise.

A total of 22 variables from the Golden Horseshoe survey were examined for clear differentiation in usage between the Canadian and U.S. portions of the study area. These will henceforth be referred to as Canadian and U.S. variants without any implication that these variants occur or do not occur in areas of Canada or the U.S. outside the study area.

Table 2: Age distribution of the sample

Age	Canada	U.S.
14-19	64	33
20-29	297	31
30-39	166	2
40-49	151	2
50-59	106	5
60-69	37	5
70-79	36	2
over 80	78	
Grand Total	935	80

According to Wildgen (1987:260), one of the basic requirements for application of catastrophe theory to linguistics is interplay between discontinuous discrete phenomena and continuity. In the model under discussion, the continuity is the geographical location in a line parallel to the lakeshore (the control variable), and the discrete phenomenon is the choice between variants in each of the variables in question (the state variable). All of the variables examined have two major variants making up more than 90 per cent of responses, creating a binary choice for most speakers between one and the other. A potential complicating factor in the data is that the U.S. responses are weighted heavily towards the younger age groups, with respondents over the age of 29

Table 3: U.S. Canadian shibboleths at the Niagara River border (respondents aged 14 to 29).

Question #/Desc.	Canadian variant	Can	US	Diff.
39: Athletic shoes	runn- (vs. sneak-)	91%	0%	91%
43: Shone	[a] (vs. [o])	85%	2%	83%
5: Garden knob	tap (vs. faucet)	89%	6%	83%
4: Sink knob	tap (vs. faucet)	84%	5%	79%
58: Anti	tee (vs. tie)	86%	16%	70%
8: Vase	ause/ays (vs. ace)	76%	7%	69%
57: Semi	me (vs. my)	89%	25%	64%
62: Z	zed (vs. zee)	64%	5%	59%
6: Cloth for face	facecloth (vs. washcloth)	66%	11%	55%
40: wants (to go) out	out (vs. to go out)	61%	8%	53%
37: Asphalt has [sh]	sh (vs. z)	80%	27%	53%
35: Lever	[eaver] (vs. [ever])	66%	16%	50%
36: Avenue	you (vs. oo)	82%	34%	48%
16: Mom	um (vs. om)	46%	3%	43%
11: Soda pop	pop (vs. soda)	94%	53%	41%
19: Evening meal	supper (vs. dinner)	51%	20%	31%
64: Progress	go (vs. got)	49%	19%	30%

being poorly represented among U.S. responses (Table 2). The result is that responses favoured by younger respondents can appear to be favoured by U.S. respondents when the entire database is analyzed. In order to minimize such misleading results, only responses from respondents aged 14 to 29 will be considered.

In each variable, the two major variants were determined and the regions were grouped into Canadian and U.S. clusters to determine the percentage of usage of the two major variants in each country. In table 3, the results are shown in terms of usage of the "Canadian variant" in each country. Since two variants were considered in each case, the U.S. variant is the difference between the figure shown and 100 per cent. The degree of divergence between the Canadian and U.S. responses was determined by calculating the difference between usage of the Canadian-favoured variant in each of the two countries. Although the calculation would be different using the U.S. variants, the result would be the same. For example, in the first line of table 1, using the figures for Canadian variants shown in the chart produces a difference of 91 per cent – 0 per cent = 91 per cent, while the U.S. figures would be 9 per cent – 100 per cent = 91 per cent difference.

The results shown in table 3 are ranked by the difference in usage between the two countries. Variables with differences of 50 per cent or greater were accepted for further study.

Each variable was analyzed by GoldVarb to determine statistically significant factors. Regionally, the factors were recoded into Canadian and U.S. groups to see the weight contributed by each of the two national areas. These will be termed "nationality" in the discussion that follows. As in the other analysis presented here, only the 14-29-year-old age groups were considered.

The variables

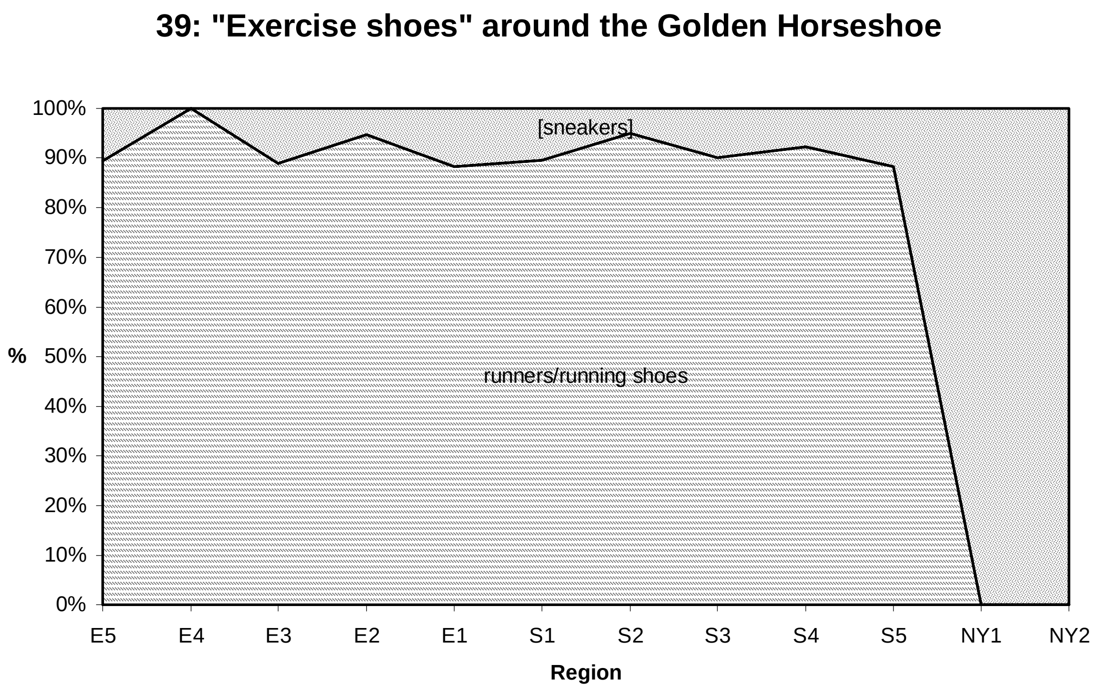

Figure 18

The most dramatic change at the border is in question 39, "What do you call the rubber-soled shoes you'd wear with [casual wear for exercise]?" with 91 per cent of the Canadians responding with "runners" or "running shoes" and 100 per cent of the U.S. responses being "sneakers." A large number of miscellaneous variants account for about 11 per cent all responses.

GoldVarb found nationality, education, and regionality index to be significant to this variable.

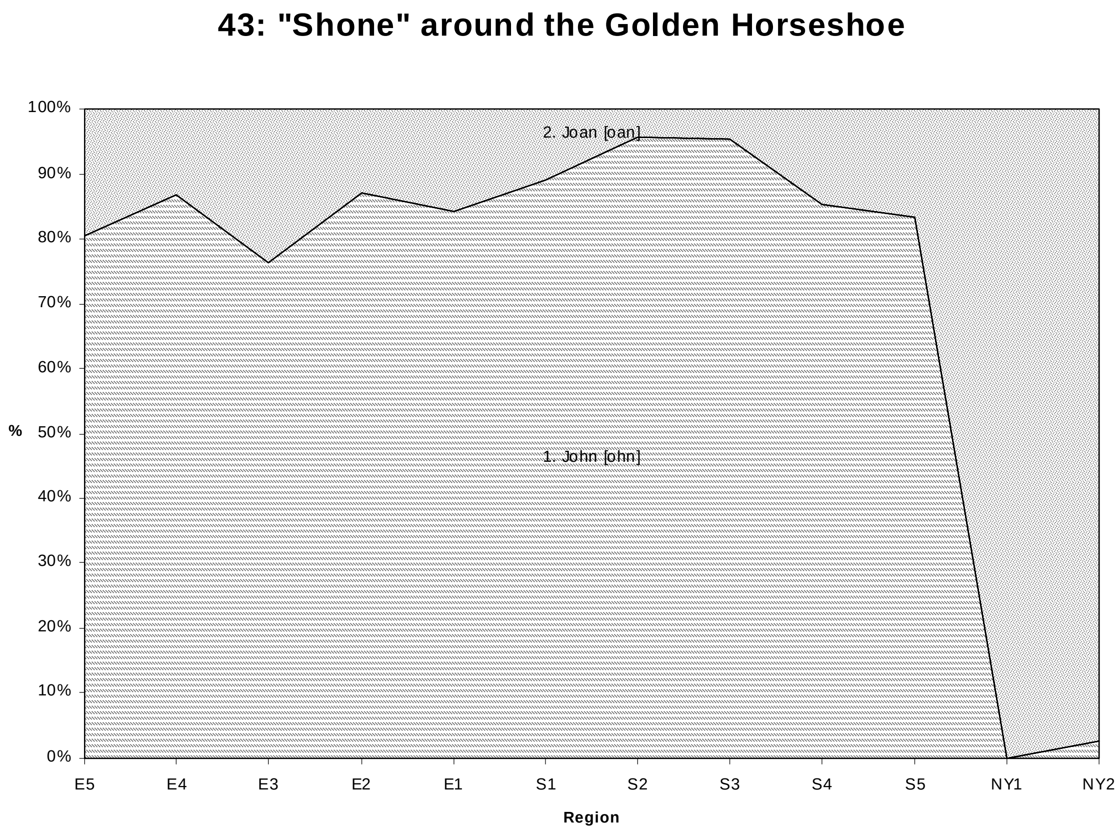

Figure 19

In the case of the pronunciation of “shone” (figure 2) there is a sharp division in the pronunciation of the letter “o,” with Americans favouring a rhyme with “Joan” almost 100 per cent and Canadians almost all rhyming it with “John.” Regional variation is fairly level throughout the Canadian regions, with the highest usage of the Canadian variant coming at the west end of Lake Ontario.

GoldVarb found nationality and age to be significant for this variable. It does appear that younger Canadians are adopting the U.S. variant.

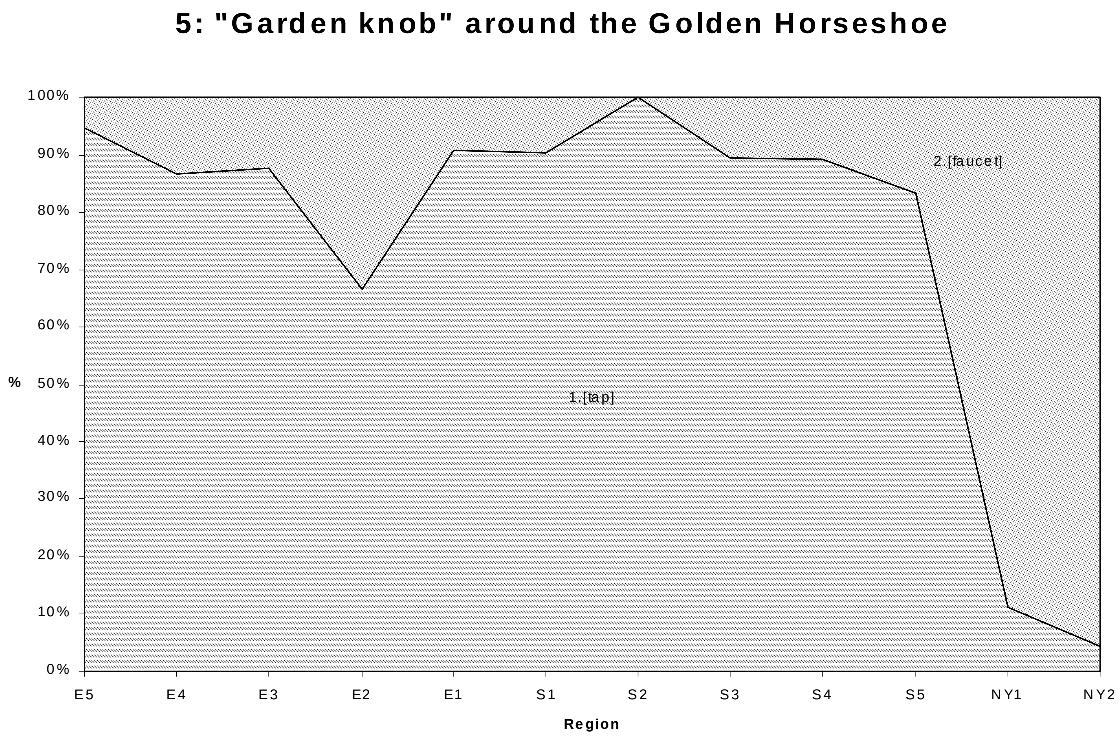

Figure 20

Question 5, “What do you call the knob you turn to get water outdoors or in the garden?” (figure 3) has two major variants “tap” and “faucet” with many minor responses making up 17 per cent of the total, the largest being “hose” at 9 per cent. This variable is the farthest from having only binary variants, although it could be argued that a hose is not any kind of knob. Although several variables are aberrant in the Toronto/Mississauga area, this one shows the most dramatic dip in the highly urban western Toronto/Mississauga region (E2). This aberration aside, the variable is quite level across the Canadian regions.

GoldVarb found that nationality and age were significant for this variable.

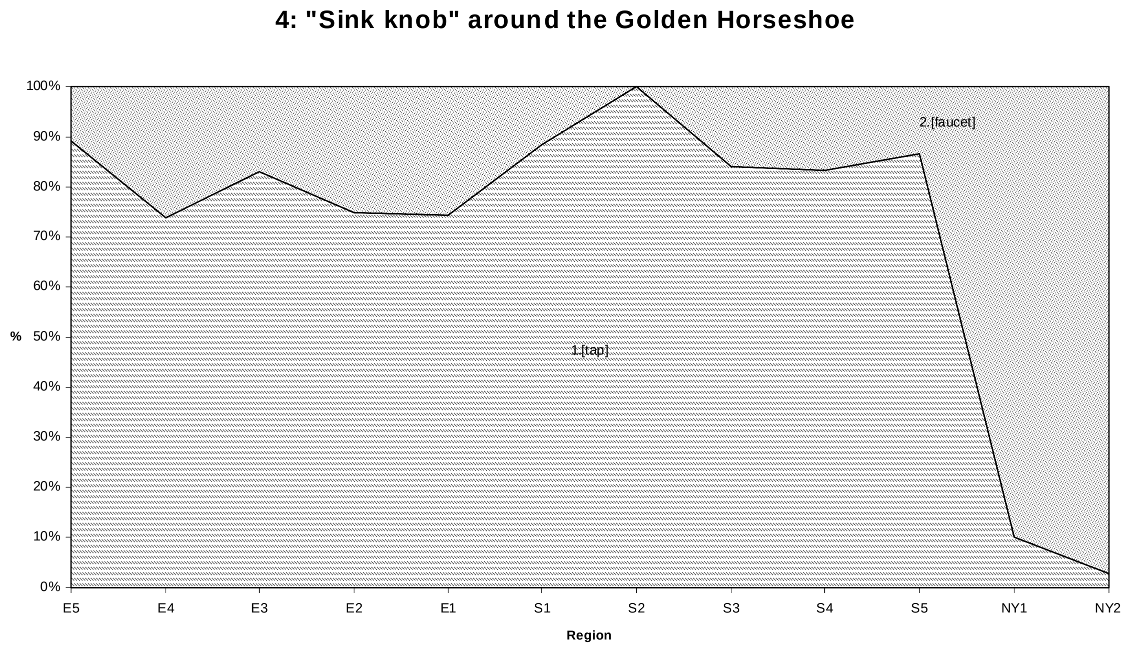

Figure 21

Moving indoors to question 4, "What do you call the knob you turn to get water in a sink," the results are very similar to question 5, with only a 4 per cent change in the difference between Canadian and U.S. responses to this question compared to the difference seen in question 5. In this question, the large aberration in region E3 is not seen, although the categorical response of people in S2 could be slightly problematic.

GoldVarb found that nationality was the only significant factor for this variable.

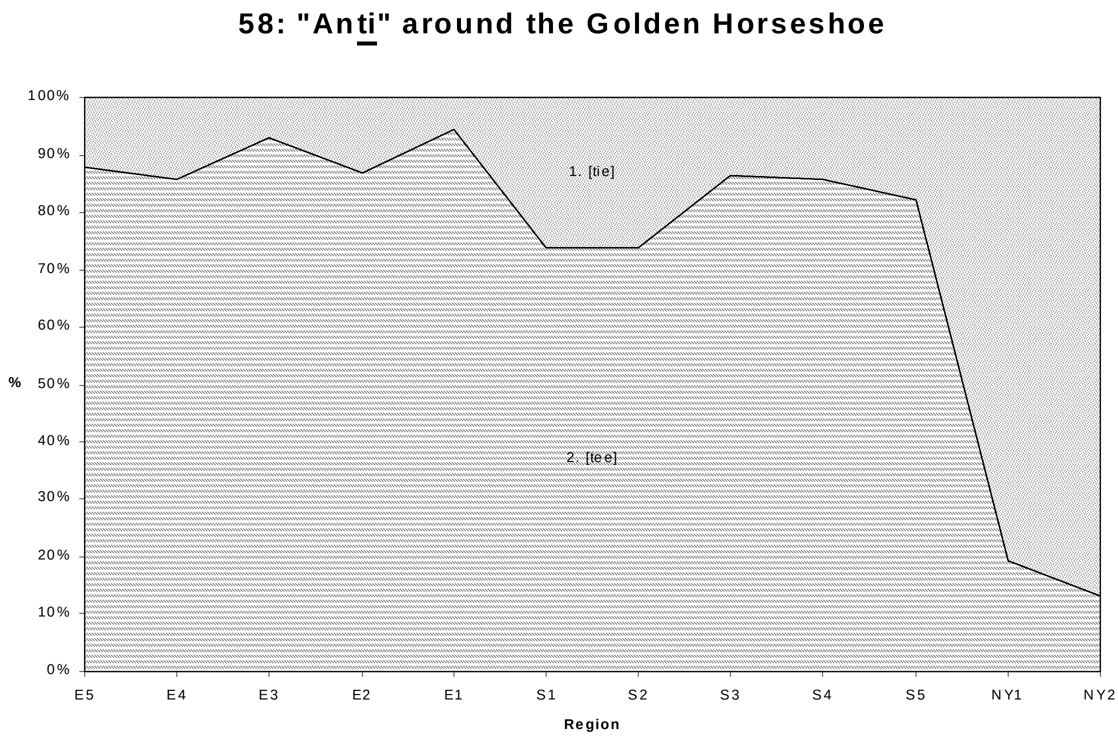

Figure 22

The pronunciation of “anti-” shows a dip in a slightly different place than in the graphs we’ve already seen, although it is still obvious that the important difference comes between the Canadian regions and the ones in the U.S.

GoldVarb found that nationality and class were significant for this variable.

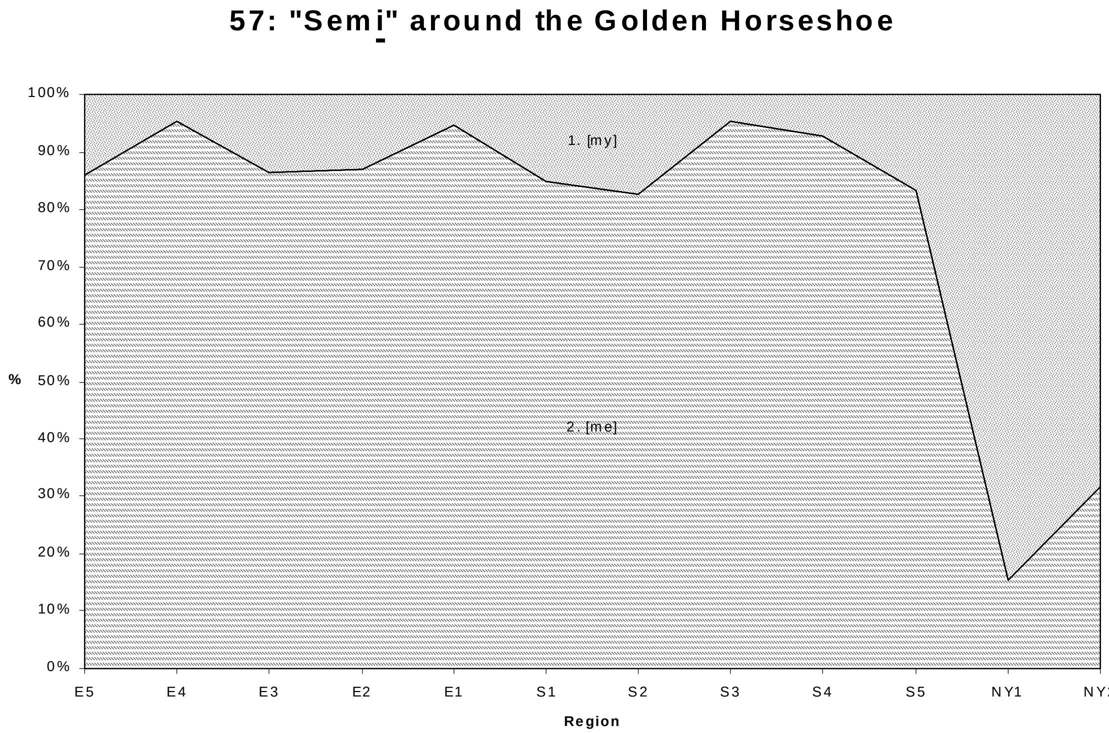

Figure 23

The pronunciation of the last vowel in “semi-,” while slightly less dramatic than “anti-,” shows a similar pattern, with the majority of Americans rhyming it with “my” rather than “me.” The Canadian/American difference is 64 per cent for this variable, compared to 70 per cent for “anti-.”

Only nationality was found significant for this variable by GoldVarb. Unlike the case of “anti-,” class was selected when stepping up in a binomial varbrul run, but eliminated when stepping down.

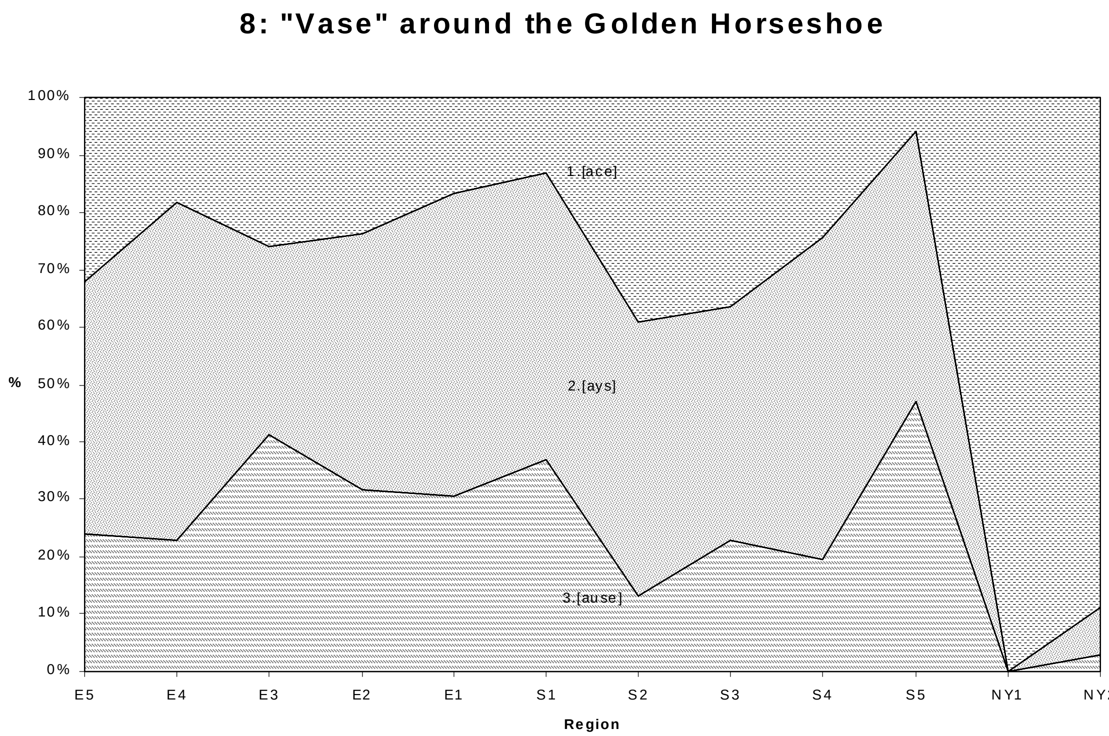

Figure 24

The pronunciation of "vase" is a little more complicated (figure 6), with three common variants. There is variation in the pronunciation of the vowel, but the important split between Canadian and American pronunciations is in the final consonant. Canadians favour a voiced [z], while the Americans in this study overwhelmingly use an unvoiced [s] pronunciation. A small number of "rhymes with 'has'" responses (2 per cent of the total) could also be added to the other [z] variants. There were very few multiple responses.

GoldVarb found nationality, age, and education to be significant with this variable.

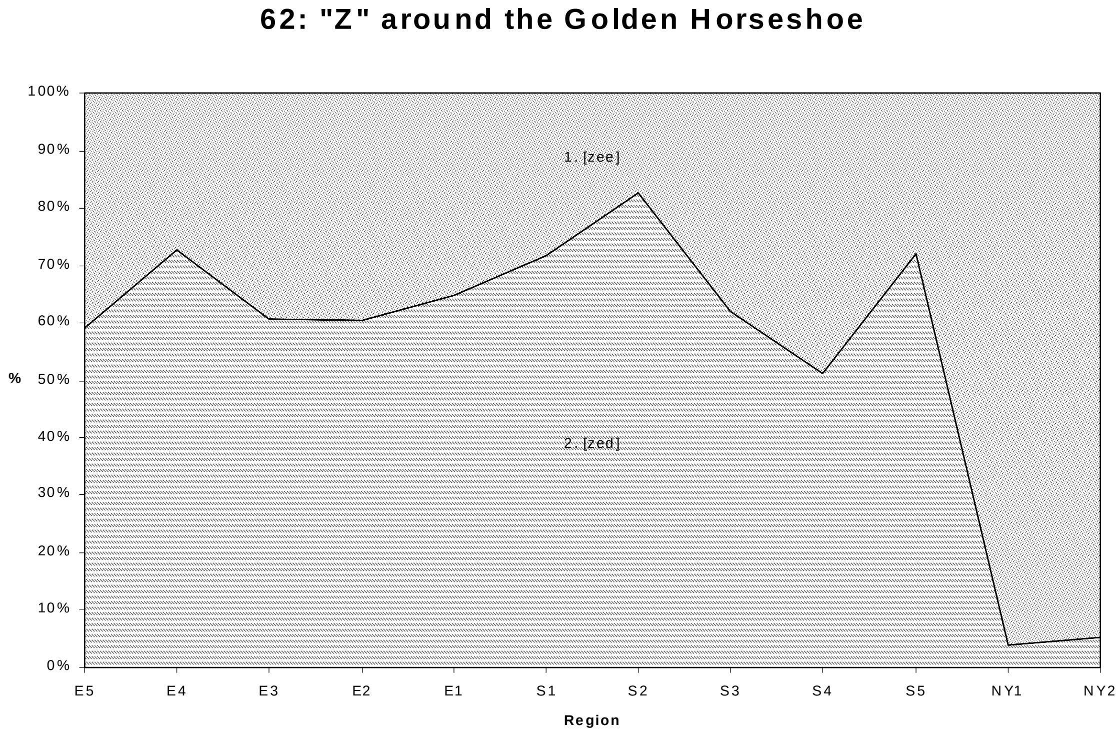

Figure 25

The pronunciation of the letter “Z” is a little less stable than the variables discussed thus far, but there is a clear division at the Canadian/U.S. border. Respondents giving both answers were about 1 per cent of the total.

Nationality, age, and education are significant, according to GoldVarb.

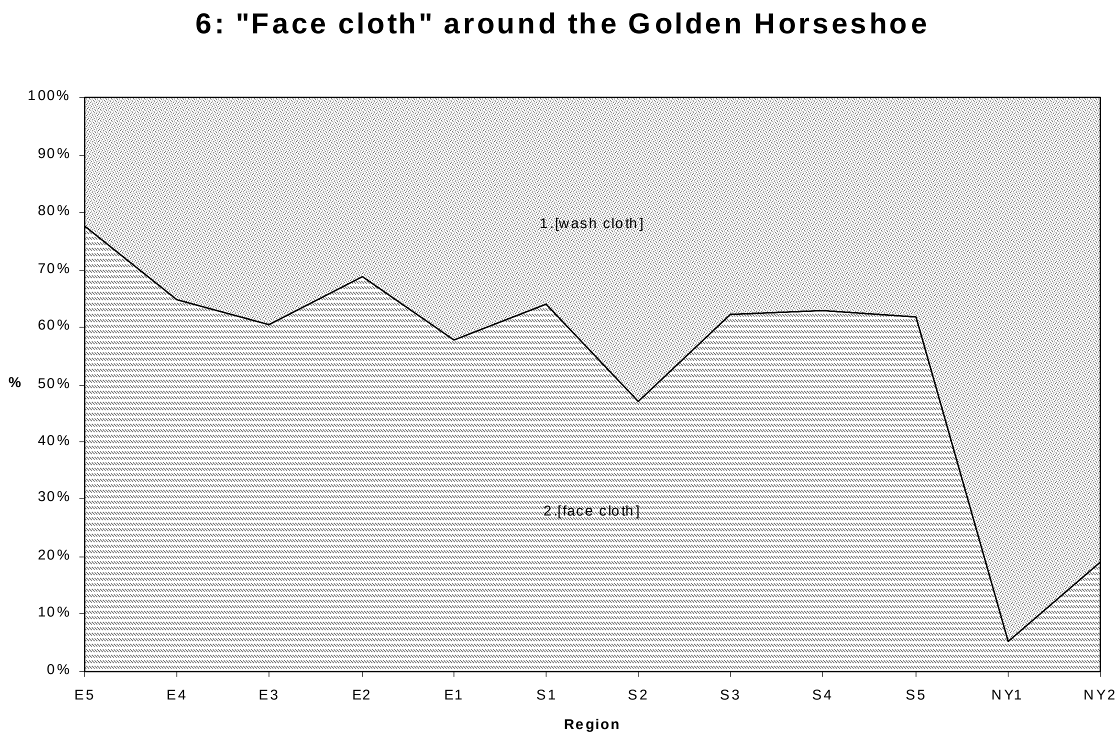

Figure 26

For question 6, “What do you call the small cloth you use for washing your face?” the Canadian preference is for “face cloth,” with all Canadian regions close to the 50 to 70 per cent range for this variant. In U.S. responses, “face cloth” occurred only about 11 per cent of the time. A number of minor variants not shown account for about 4 per cent of the total, and dual responses, almost all containing washcloth and one other variant, account for another 2 per cent.

GoldVarb selects nationality and education as significant factors for this variable.

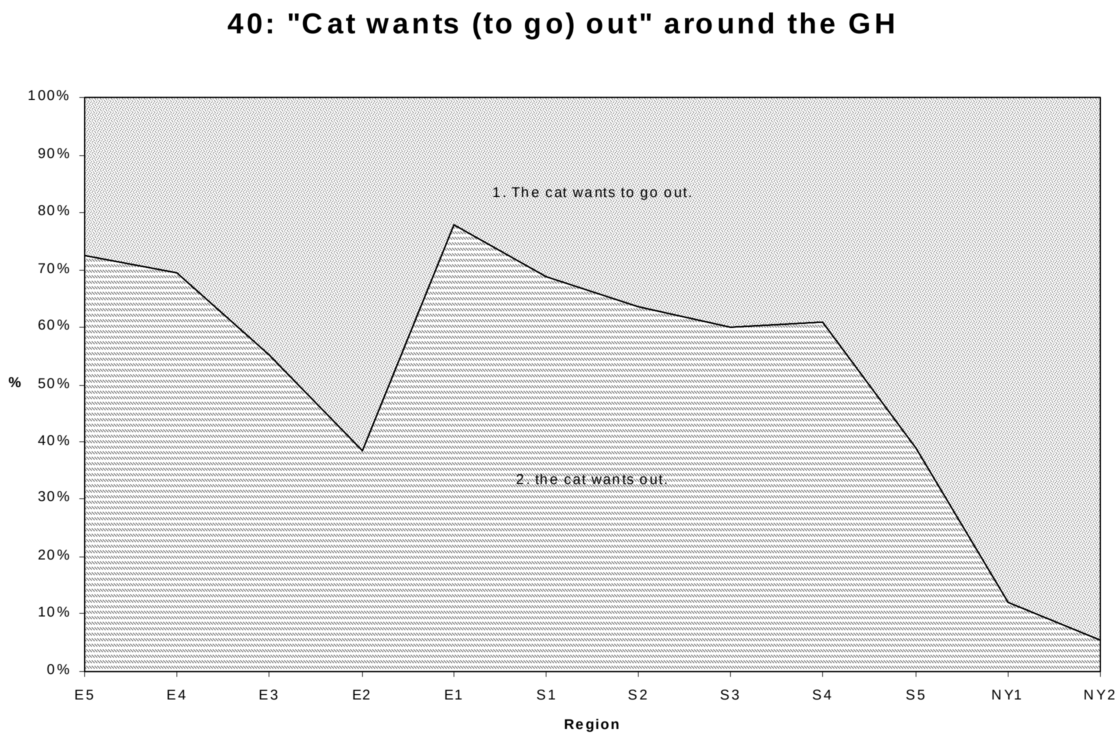

Figure 27

In question 40, respondents were asked to choose between "The cat wants out" and "The cat wants to go out." Multiple responses only account for about 1 per cent of the total. For this question, there is a sharp dip in region E2, as we've seen before in the discussion of question 5 ("garden knob"). This dip also shows up to a lesser extent in the adjoining region, E3 (Toronto). These aberrations aside, there seems to be a gradual slope in the variant as we move towards the U.S. border, but there is still a fairly sharp difference between the Canadians as a whole and the Americans.

GoldVarb found nationality, age, education, and occupational mobility index to be significant.

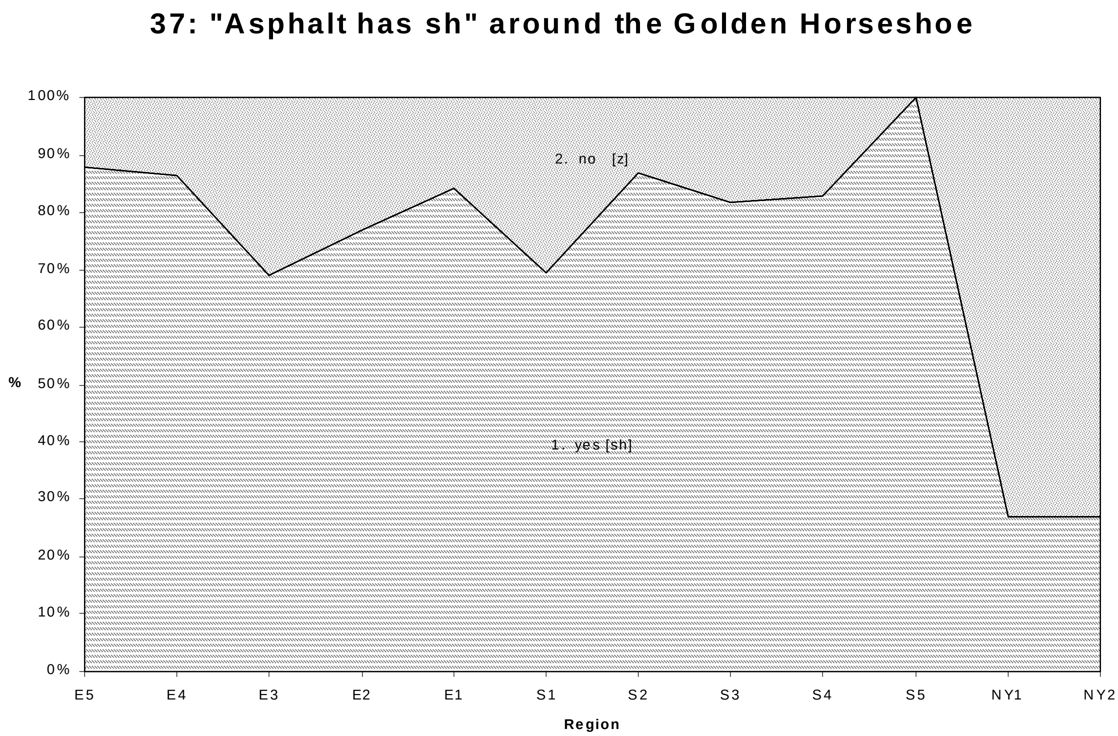

Figure 28

In question 37, "Does the S [in ASPHALT] sound like <u>sh</u>?" the division between Canada and the U.S. is less clear than in some of the others, but the U.S. regions do not come anywhere near to the range of the Canadian regions and there is a difference of 53 per cent between the U.S. speakers and the Canadians, with only 27 per cent of the Americans in these age groups using a "sh."

GoldVarb found nationality, class, and regionality index to be significant.

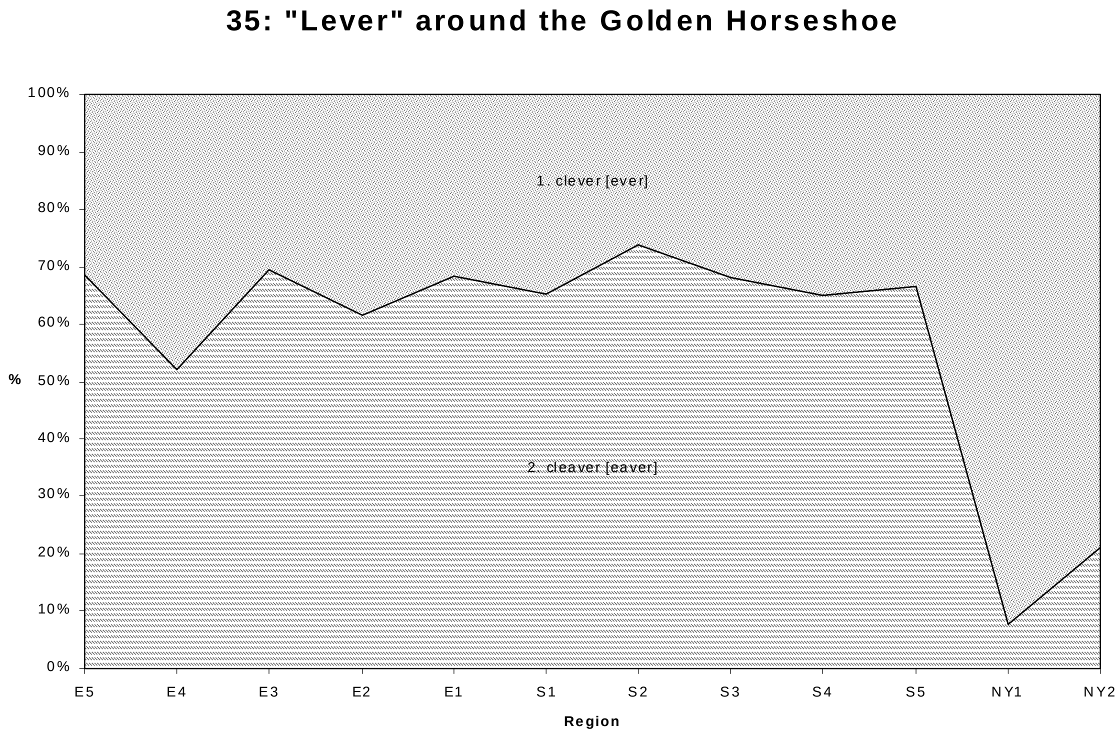

Figure 29

The final variable under consideration is the pronunciation of "lever" (figure 12), with the majority of the Canadians pronouncing it to rhyme with "cleaver" and the majority of the Americans rhyming it with "clever."

GoldVarb found that significant factors for "lever" were nationality, age, sex, class, and regionality index.

Having selected variables with clear Canadian and American variants, the questions now are whether there is an overall pattern to the variation from region to region and whether that pattern is consistent with catastrophe theory. There is often an ill-defined "transition zone" between dialect regions. Do these variables gradually transition from one region to another?

To attempt to establish a pattern, all of the variables discussed thus far were plotted onto a single graph (figure 30). Still, the pattern is not clear, so these plot lines were averaged to produce figure 31 in which a clear trend emerges.

The result is that the variants are remarkably stable across the regions in each country. The exceptions are the E3 and E2 regions previously discussed, which do exhibit some aberration from the Canadian norm.

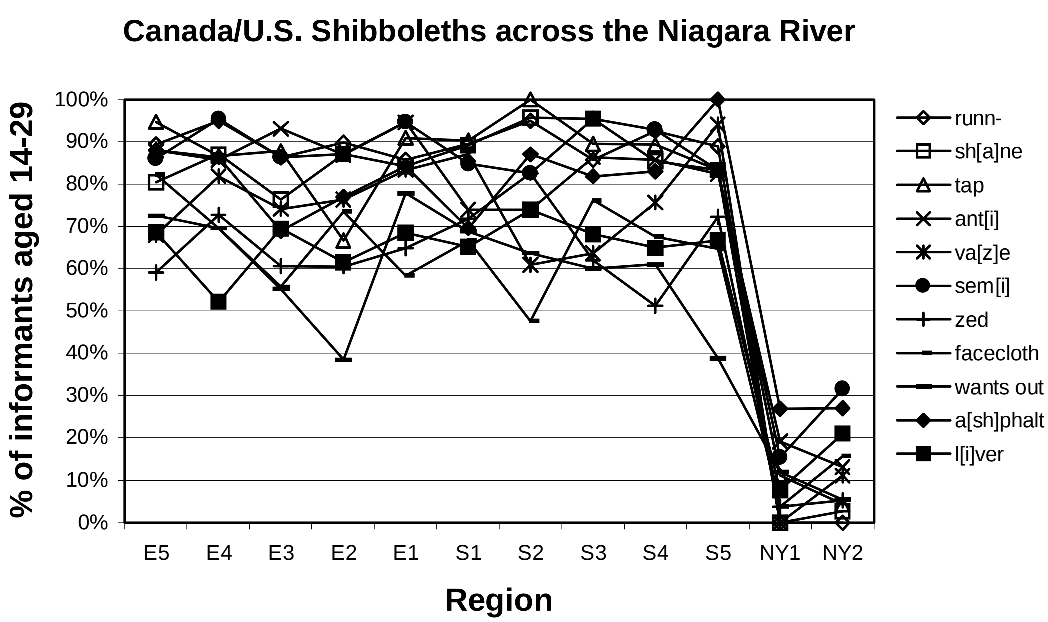

Figure 30

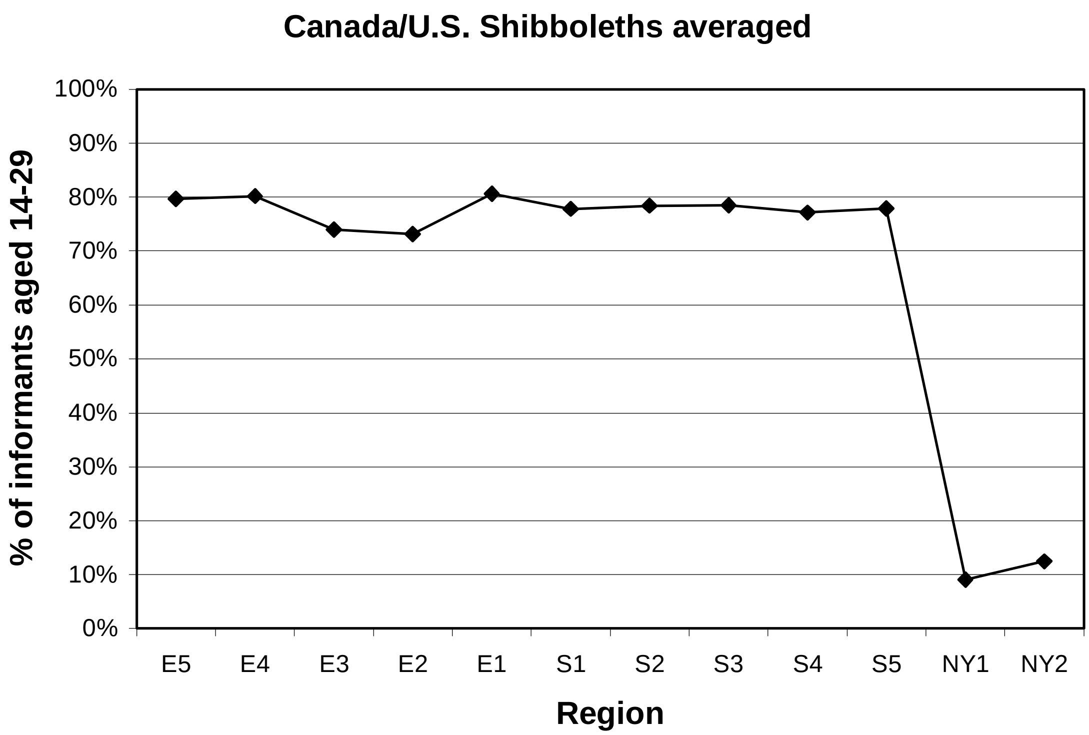

Figure 31

Does this result fit with catastrophe theory? Certainly, insofar as a small movement in one direction (geographically towards Canada or the U.S.) creates a large effect in something else (the Canadian and U.S. variants in this case), a basic requirement of a catastrophe theory is met. Another characteristic feature of a catastrophe is hysteresis, the lag effect where the transition in one direction differs from the transition point in the other direction.

To find hysteresis at the dialect boundary, we would need to find some evidence of dual responses. Unfortunately, dual responses are very rare in the Dialect Topography databases, due to the instructions on the questionnaire. The questionnaire explicitly discourages dual responses when it says, "Your first answer is likely to be the best one." By this instruction, respondents are encouraged to answer the first thing they think of and disregard any variation they may have in their own speech.

A study by Christine Zeller (1990, 1993), which slightly pre-dates the original Dialect Topography survey, got very different results with respect to dual responses. Zeller's study, with written questionnaires distributed in a corridor going from Toronto, Canada, to Milwaukee in the U.S., received many more multiple responses. She asked the informants to "feel free and give as much information as you want" and "if you have more than one word in answer to a question, please list them all, but indicate (if possible) the one most often used by you and your friends by underlining it" (Zeller 1990:7).

Unfortunately, it is not easy to create tables of raw numbers from Zeller's graphs, but her general conclusions do support a type of hysteresis across the border. Zeller divides her results into two main patterns, symmetrical and asymmetrical (1993:196). The symmetrical overlap would be similar to a cusp catastrophe behaviour.

In asymmetrical distributions, "a predominant American response is also regularly found across the border, but the popular Canadian response does not surface in the U.S." (1993:196-97).

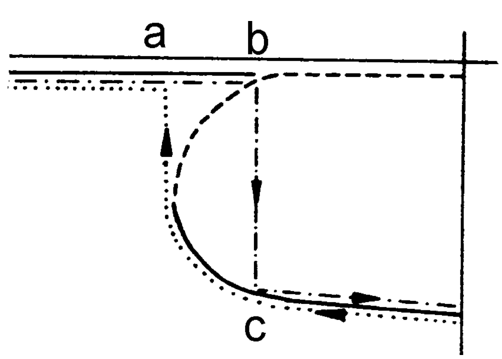

Figure 32: Hysteresis on the fold catastrophe (adapted from Wilson 1981:81)

The fold catastrophe, with only one critical variable, seems like a good fit to this behaviour and the behaviour we've seen in the Dialect Topography data. Figure 32 has been adapted from Wilson (1981:81), but has been turned 180 degrees to better match the orientation of the Golden Horseshoe graphs just presented here. In figure 32, the Canada/U.S. border would be located at the line bc. As a variable moves to the right in figure 32, it jumps from the horizontal attractor across the top of the graph onto the curve at the bottom. The portion of the curve to the right of point b is inaccessible. Variables coming from the right, approaching point c, continue along the attractor through the hysteresis until they reach the bifurcation point, at which point, they jump to point a. This way, points between a and b have two attractors to choose from, while those to the right of the border have only one possible stable resting place. The question of which of the two attractors wins out for an individual speaker depends on the source of the original linguistic input.

Some speakers will possess two or more individual grammars (Lightfoot 1999:92 ff.). "Change proceeds via competition between grammatically incompatible options which substitute for one another in usage" (Kroch 1994:180 quoted in Lightfoot 1999:94). Lightfoot says that grammars unify alternating forms, not from free variance, but oscillation between two (or more) fixed points.

Conclusions

` In order to apply chaos and catastrophe theory to data of dialect areas and linguistic variation, a new way of thinking has to be applied, but if we do this we can certainly make ourselves open to new insights about ways in which language can vary in time and space, among many other factors which affect language in society. This field is only slowly and tentatively being applied to linguistic problems. Many new insights have yet to be made, once the basics are in place.

References:

Alligood, Kathleen T., Tim Sauer, and James A. York. 1997. *Chaos – an introduction to dynamical systems*. New York: Springer-Verlag.
Briggs, J. and F.D. Peat. 1989. *Turbulent mirror. An illustrated guide to chaos theory and the science of wholeness*. New York: Harper and Row.
Campbell, Lyle. 1998. *Historical linguistics: an introduction*. Cambridge MA: The MIT Press.
Chambers, J.K. 1994. An introduction to dialect topography. *English World-Wide* 15:35-53.
Chambers, J.K. 1995. *Sociolinguistic Theory*. Cambridge MA: Blackwell.
Chambers, J.K. and Troy Heisler. 1999. Dialect Topography of Quebec City English. *Canadian Journal of Linguistics* 44(1):23-48.
Kroch, A. 1994. "Morphosyntactic variation." In K. Beals et al. (eds.), *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: parasession on variation and liguistics theory*. Chicago: Chicago Linguistics Society.
Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.
Lightfoot, David. 1999. *The Development of Language*. Malden MA: Blackwell.
MacKeracher, Mary S. 2001. *Language Stability in a World of Language Innovation*. Generals paper, Department of Linguistics, University of Toronto.
Rand, David, and David Sankoff. 1990. *GoldVarb version 2: A variable rule application for Macintosh*. Centre de recherches mathématiques, Université de Montréal.
Saunders, P.T. 1980. *An introduction to catastrophe theory*. Cambridge: Cambridge University Press.
Schneider, Edgar W. 1997. Chaos theory as a model for dialect variability and change? *Issues and Methods in Dialectology,* ed. Alan R. Thomas. Bangor: University of Wales.
Statistics Canada. 1996. Population by home language, 1996 Census, census metropolitan areas. [`http://www.statisticscanada.ca/english/Pgdb/People/Population/demo29f.htm`]
Stewart, Ian. 1989. *Does God play dice?: the mathematics of chaos*. Cambridge MA: Blackwell.
Wildgen, Wolfgang. 1987. Catastrophe Theory as a Basic Tool in Theoretical Linguistics. *Theoretical Lingusitics* v. 14. 259-94.
Wilson, A.G. 1981. *Catastrophe theory and bifurcation*. Berkeley/Los Angeles: University of California Press.
Woodcock, Alexander and Monte Davis. 1978. *Catastrophe theory*. New York: E.P. Dutton.
Zeller, Christine. 1990. *Dialect Variants from Toronto to Milwaukee*. M.A. thesis, Linguistics Department, University of Toronto.

Zeller, Christine. 1993. "Linguistic symmetries, asymmetries, and border effects within a Canadian/American sample." In *Focus on Canada*, ed. Sandra Clarke. Amsterdam/Philadelphia: John Benjamins. 179-99.