ANALYSIS OF U.S. PRESIDENTIAL CANDIDATES' SPEECHES USING WORDSTAT 3.0

Paper presented by Normand Péladeau
at the Computer Assisted Content Analysis (CATA) workshop
51th Annual Conference of the International Communication Association
Washington, May 2001

SOFTWARE USED: WordStat v3.05

SAMPLING: Thirty one speeches on the only two topics covered by all six candidates:

- Candidature announcement
- Foreign policy

TEXT PREPROCESSING:

- No lemmatization or spell checking
- Titles and interviewers' questions were put between braces in order to removed them from the analysis.

CATEGORIZATION USED:

- Most frequent words (at least 50 times).
- Lasswell Value Dictionary
- Martindale's Regressive Imagery Dictionary (RID)
- Pennebaker's LIWC dictionary

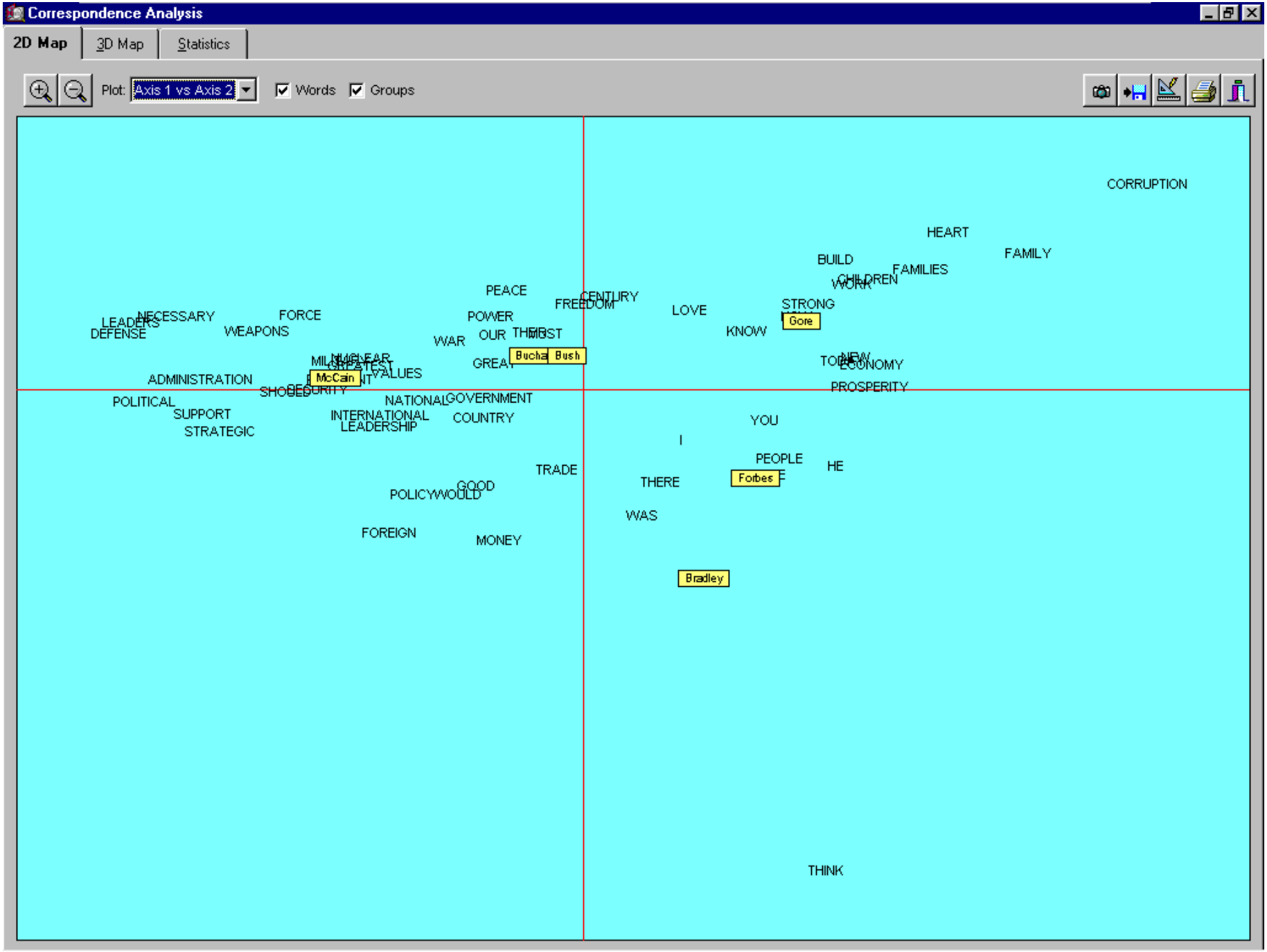
ANALYSIS FOR COMPARISON:

- Crosstabulation (with computation of F statistics)
- Correspondence Analysis
- Charting: barchart and joint plot.

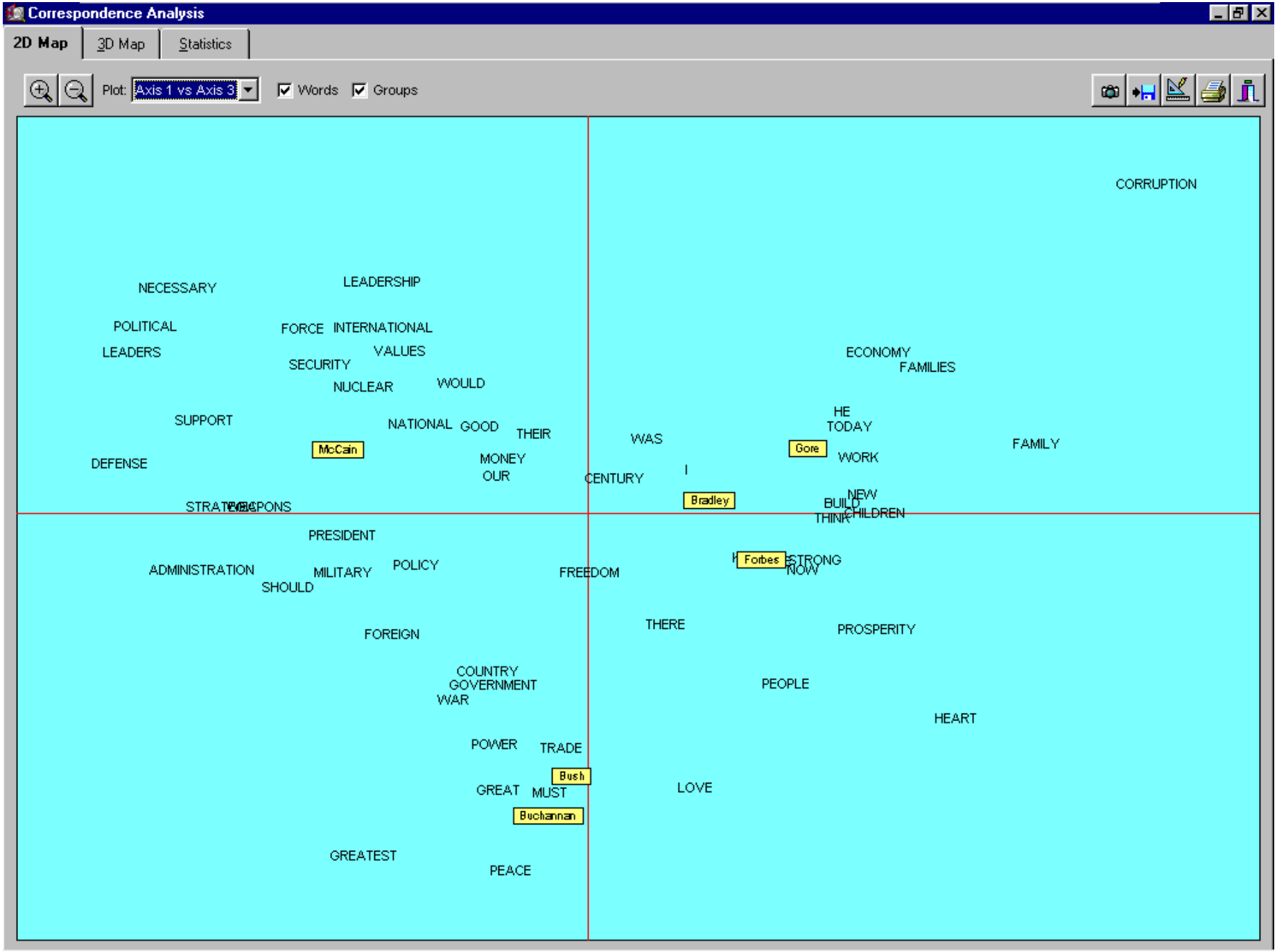
VALIDITY CHECK:

- Only a few KWIC listings were performed. More validation work is needed to assess the validity of the conclusions.

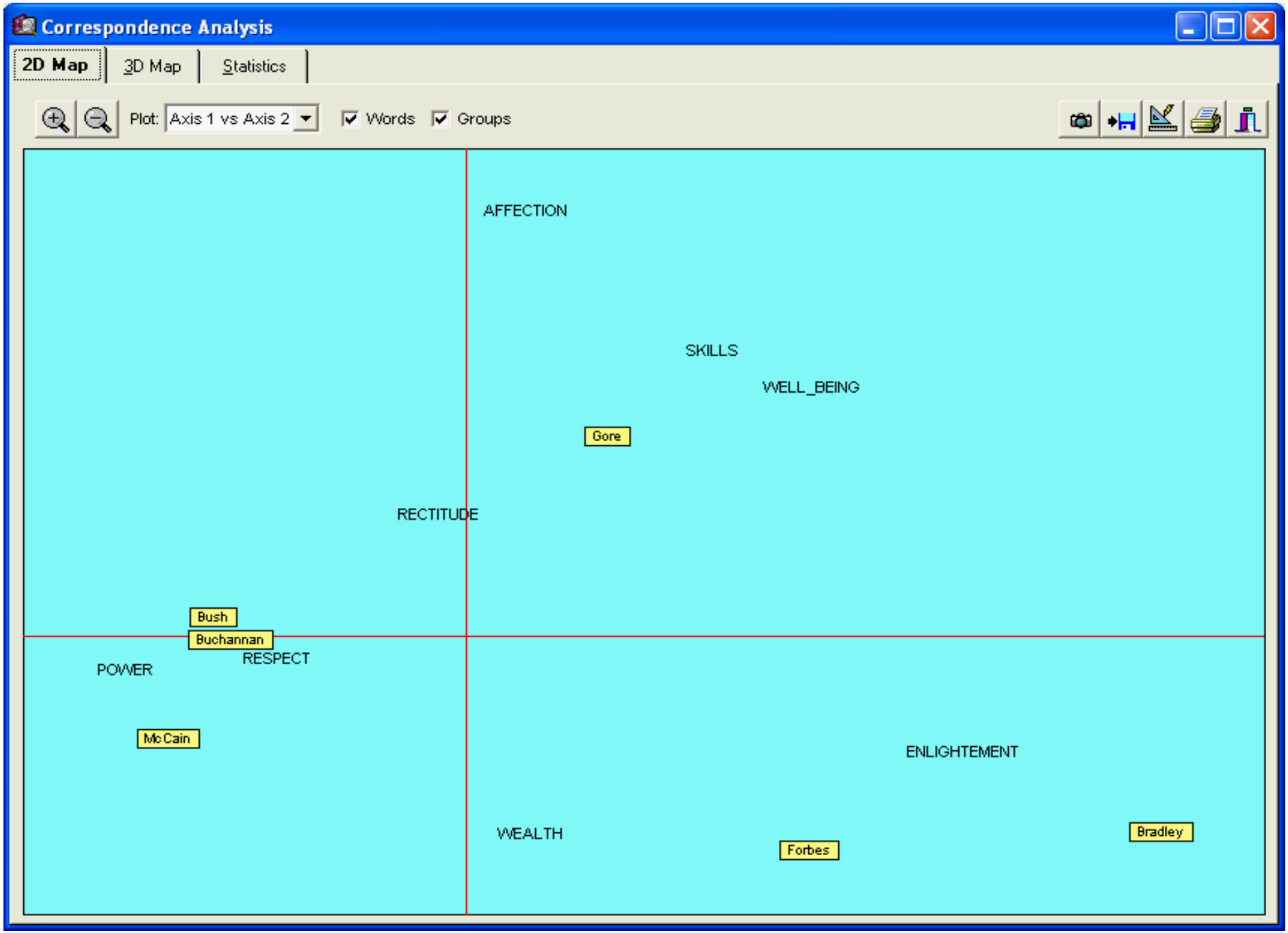
Correspondence Plot of Frequent Words and Presidential Candidates (axis 1 & 2)



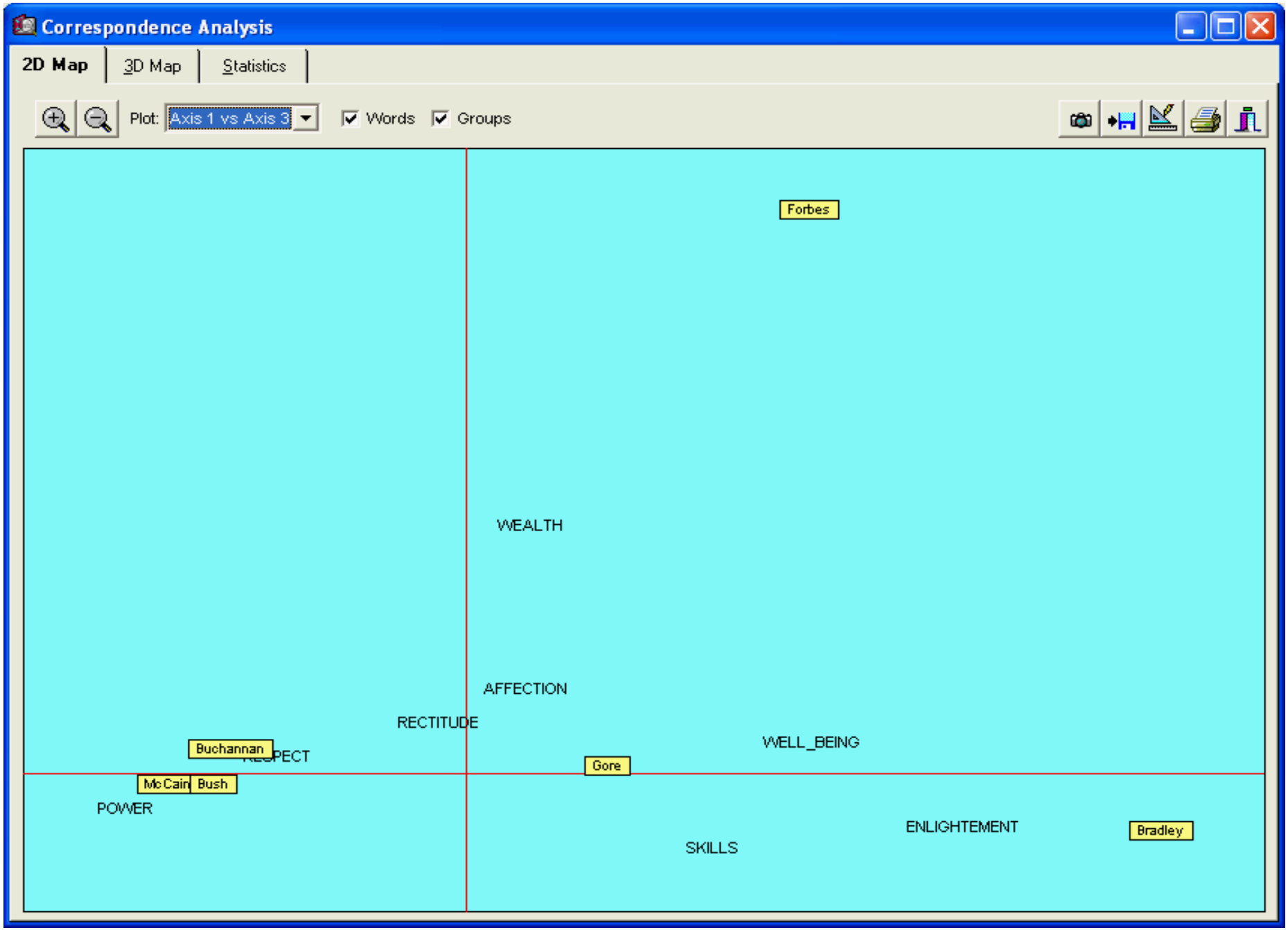
Correspondence Plot of Frequent Words and Presidential Candidates (axis 1 & 3)



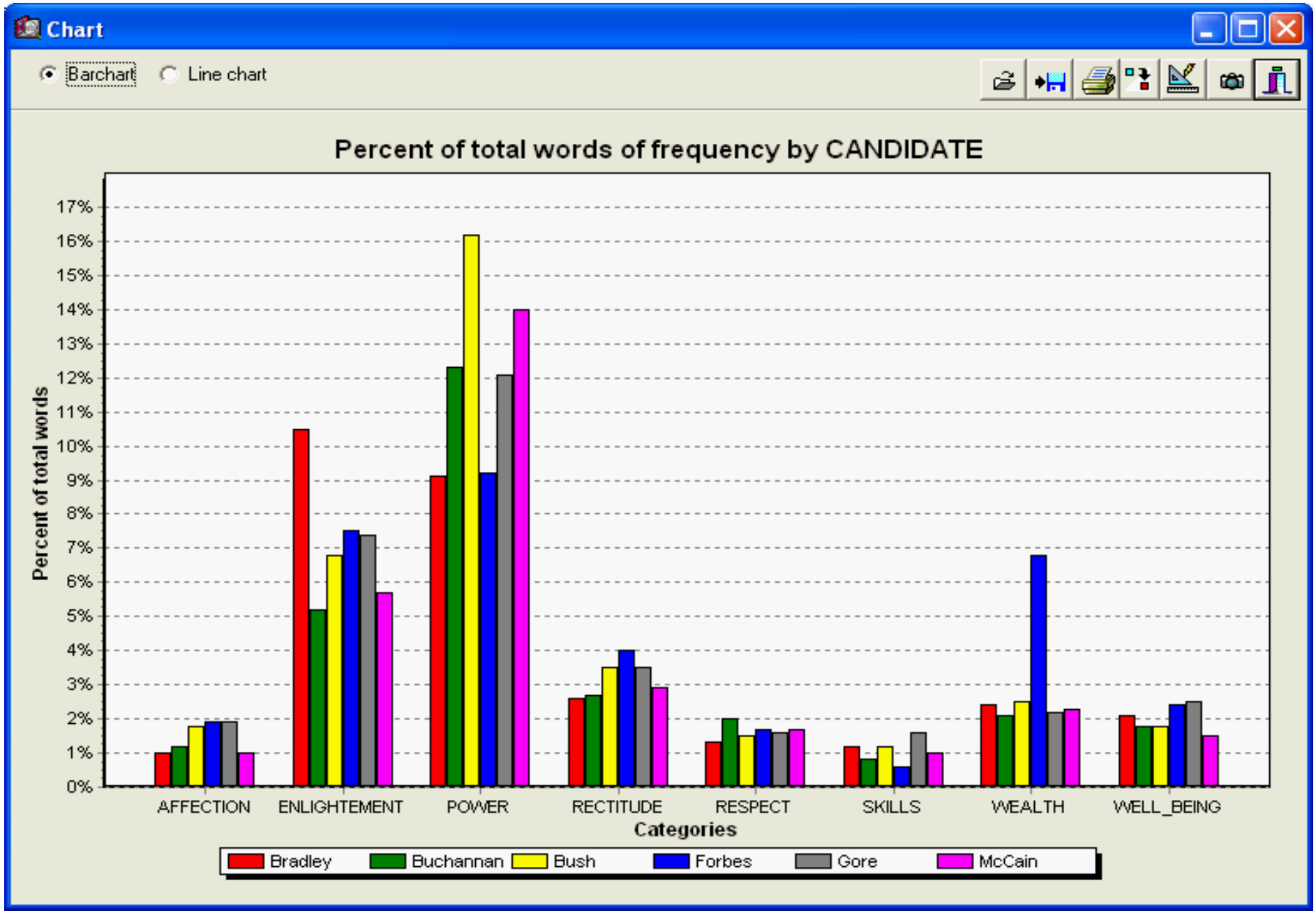
Correspondence Plot of Lasswell Values Dictionary and Presidential Candidates (axis 1 & 2)



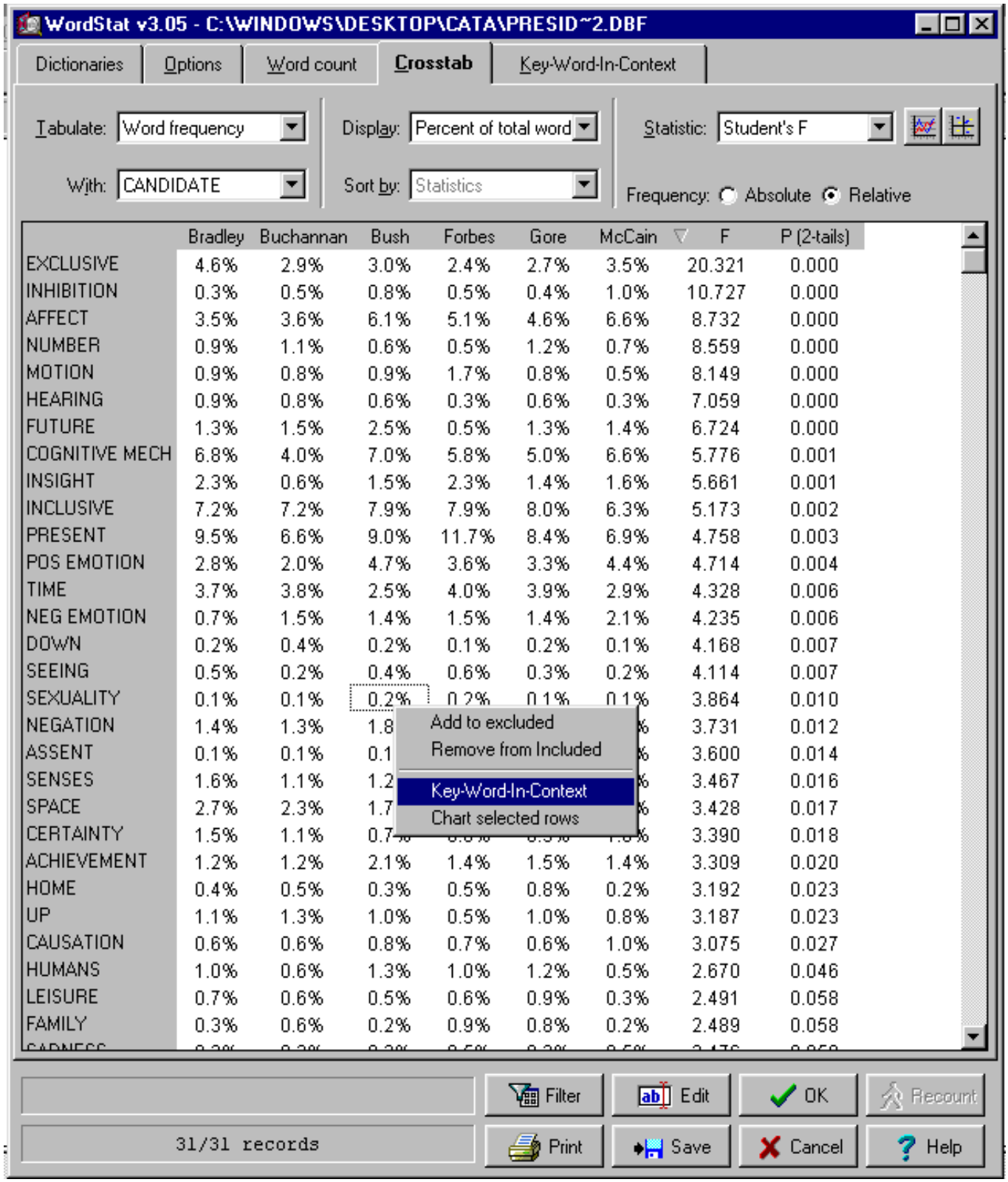
Correspondence Plot of Lasswell Values Dictionary and Presidential Candidates (axis 1 & 3)



Relative Importance of Eight Categories of Lasswell Values Dictionary among Presidential Candidates



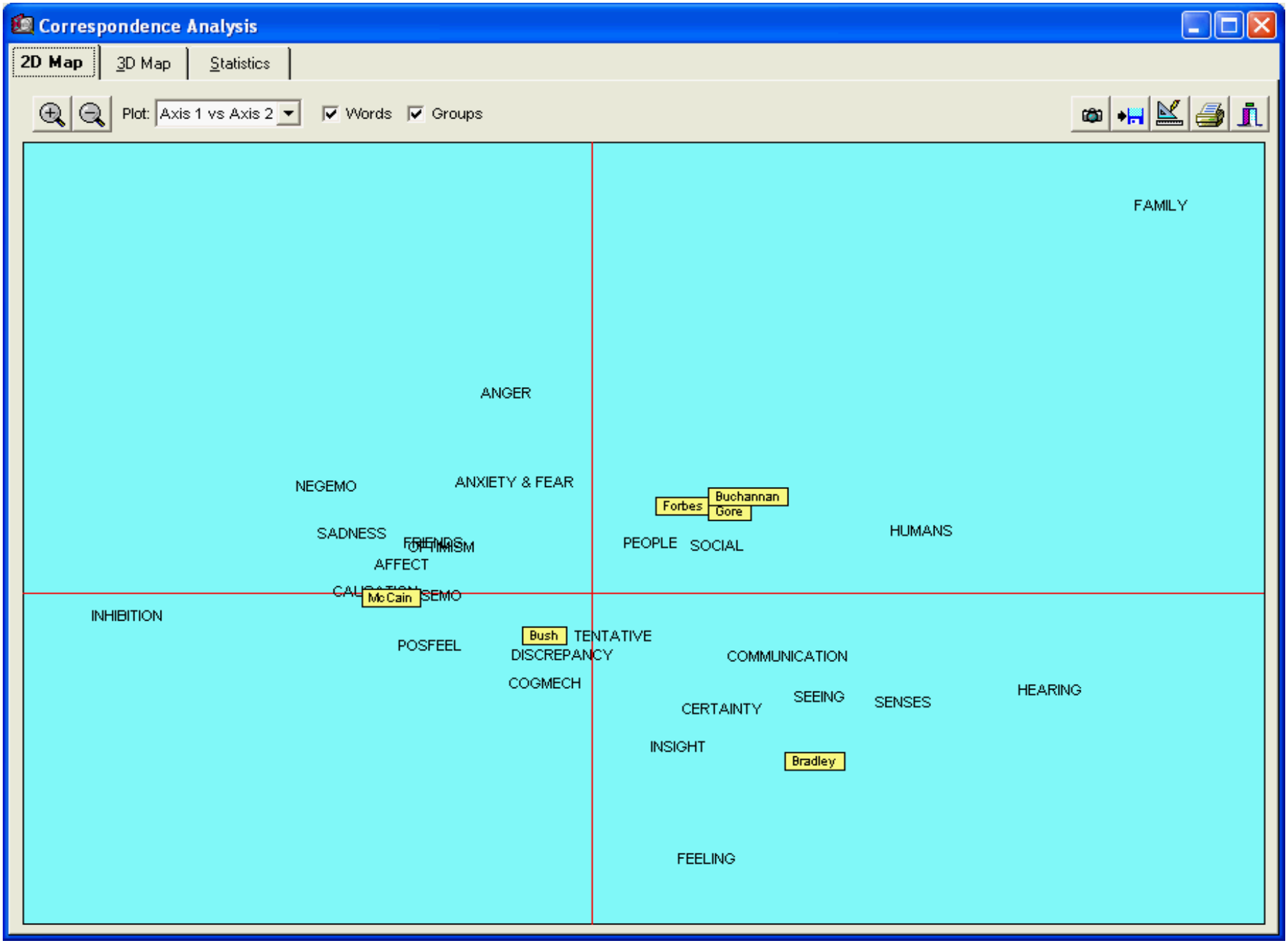
Relationship between LIWC Categories and Presidential Candidates Sorted in Descending order of F Values



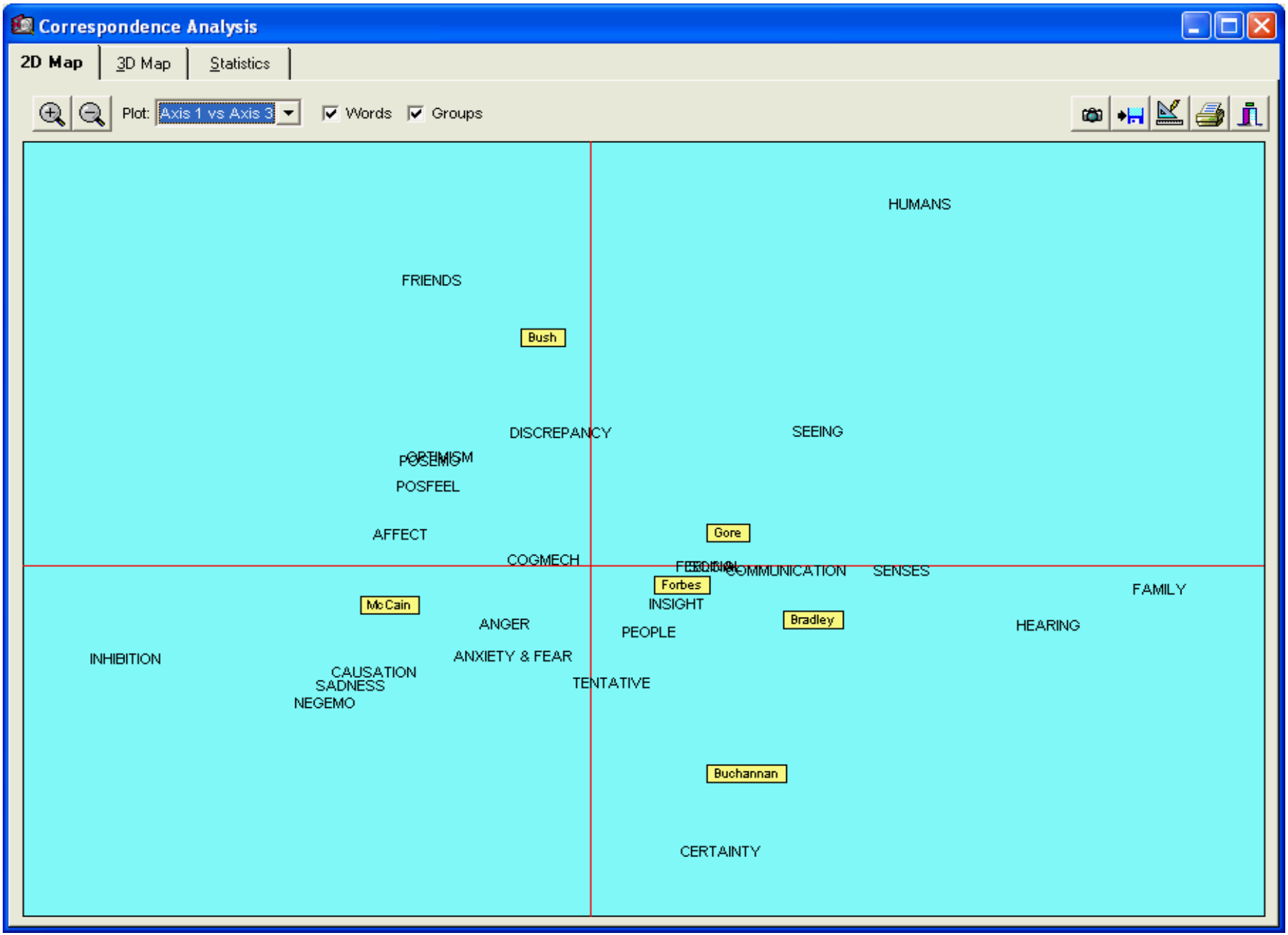
	Bradley	Buchannan	Bush	Forbes	Gore	McCain	F	P (2-tails)
EXCLUSIVE	4.6%	2.9%	3.0%	2.4%	2.7%	3.5%	20.321	0.000
INHIBITION	0.3%	0.5%	0.8%	0.5%	0.4%	1.0%	10.727	0.000
AFFECT	3.5%	3.6%	6.1%	5.1%	4.6%	6.6%	8.732	0.000
NUMBER	0.9%	1.1%	0.6%	0.5%	1.2%	0.7%	8.559	0.000
MOTION	0.9%	0.8%	0.9%	1.7%	0.8%	0.5%	8.149	0.000
HEARING	0.9%	0.8%	0.6%	0.3%	0.6%	0.3%	7.059	0.000
FUTURE	1.3%	1.5%	2.5%	0.5%	1.3%	1.4%	6.724	0.000
COGNITIVE MECH	6.8%	4.0%	7.0%	5.8%	5.0%	6.6%	5.776	0.001
INSIGHT	2.3%	0.6%	1.5%	2.3%	1.4%	1.6%	5.661	0.001
INCLUSIVE	7.2%	7.2%	7.9%	7.9%	8.0%	6.3%	5.173	0.002
PRESENT	9.5%	6.6%	9.0%	11.7%	8.4%	6.9%	4.758	0.003
POS EMOTION	2.8%	2.0%	4.7%	3.6%	3.3%	4.4%	4.714	0.004
TIME	3.7%	3.8%	2.5%	4.0%	3.9%	2.9%	4.328	0.006
NEG EMOTION	0.7%	1.5%	1.4%	1.5%	1.4%	2.1%	4.235	0.006
DOWN	0.2%	0.4%	0.2%	0.1%	0.2%	0.1%	4.168	0.007
SEEING	0.5%	0.2%	0.4%	0.6%	0.3%	0.2%	4.114	0.007
SEXUALITY	0.1%	0.1%	0.2%	0.2%	0.1%	0.1%	3.864	0.010
NEGATION	1.4%	1.3%	1.8			%	3.731	0.012
ASSENT	0.1%	0.1%	0.1			%	3.600	0.014
SENSES	1.6%	1.1%	1.2			%	3.467	0.016
SPACE	2.7%	2.3%	1.7			%	3.428	0.017
CERTAINTY	1.5%	1.1%	0.7			%	3.390	0.018
ACHIEVEMENT	1.2%	1.2%	2.1%	1.4%	1.5%	1.4%	3.309	0.020
HOME	0.4%	0.5%	0.3%	0.5%	0.8%	0.2%	3.192	0.023
UP	1.1%	1.3%	1.0%	0.5%	1.0%	0.8%	3.187	0.023
CAUSATION	0.6%	0.6%	0.8%	0.7%	0.6%	1.0%	3.075	0.027
HUMANS	1.0%	0.6%	1.3%	1.0%	1.2%	0.5%	2.670	0.046
LEISURE	0.7%	0.6%	0.5%	0.6%	0.9%	0.3%	2.491	0.058
FAMILY	0.3%	0.6%	0.2%	0.9%	0.8%	0.2%	2.489	0.058

Note: Popup menu showed above allows one to move words in the crosstab table, jump to a KWIC list of the selected category or chart

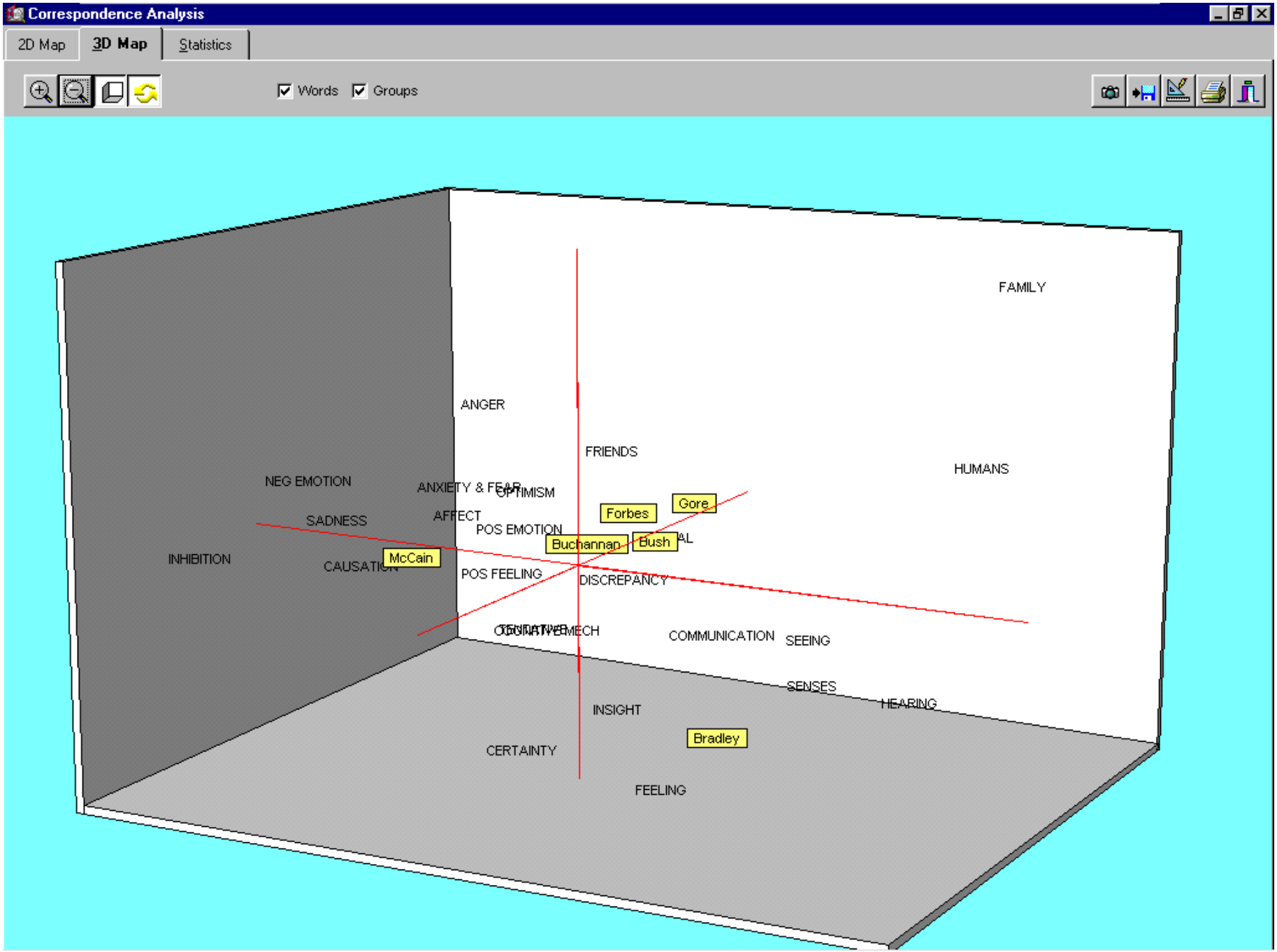
Correspondence Plot of Pennebaker's LIWC Psychological Categories and Presidential Candidates (axis 1 & 2)



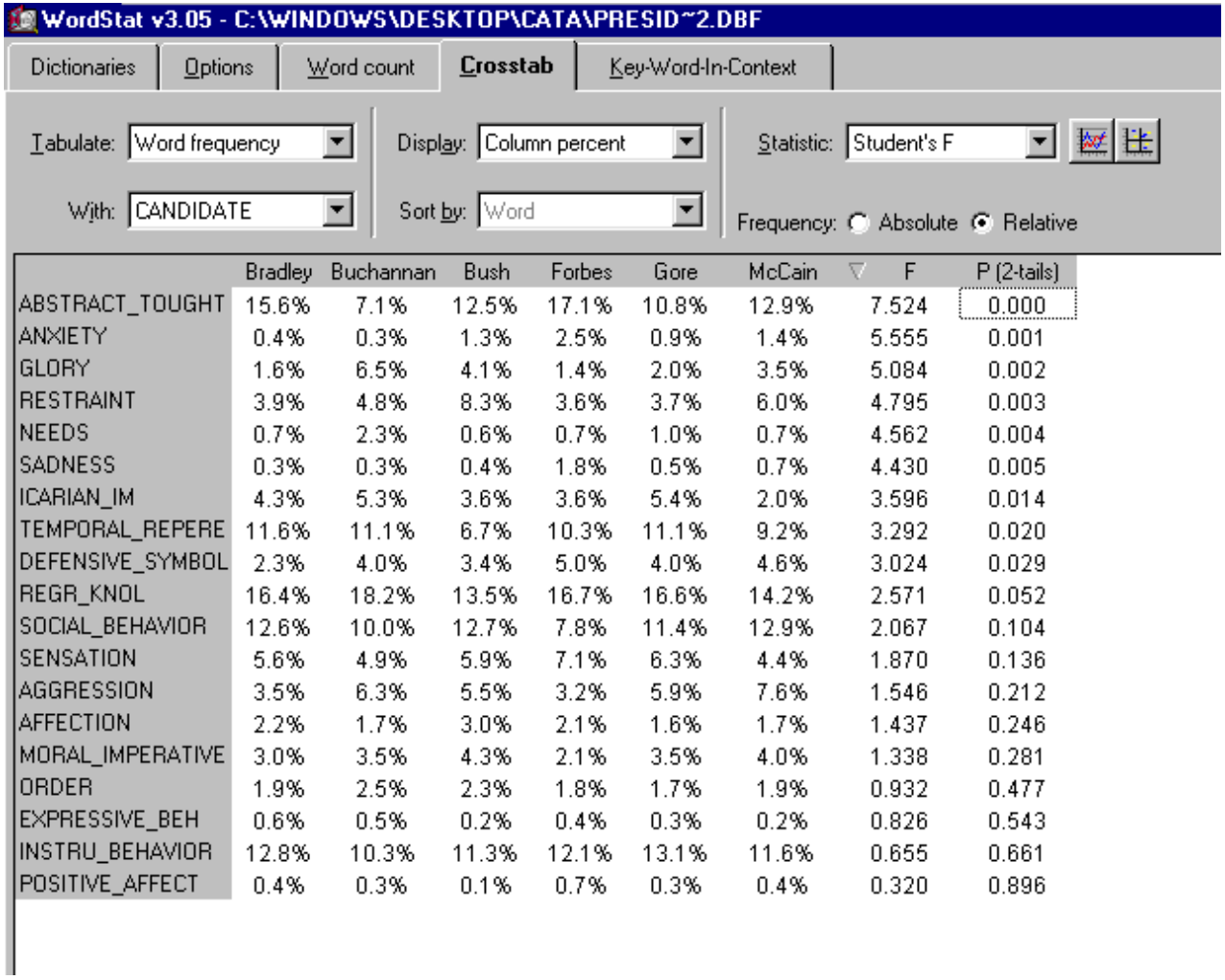
Correspondence Plot of Pennebaker's LIWC Psychological Categories and Presidential Candidates (axis 1 & 3)



3D Correspondence Plot of Pennebaker LIWC Categories and Presidential Candidates



Sample Crosstabulation Table: Relationship between RID Categories and Presidential Candidates (column percent and F test)



WordStat v3.05 - C:\WINDOWS\DESKTOP\CATA\PRESID~2.DBF

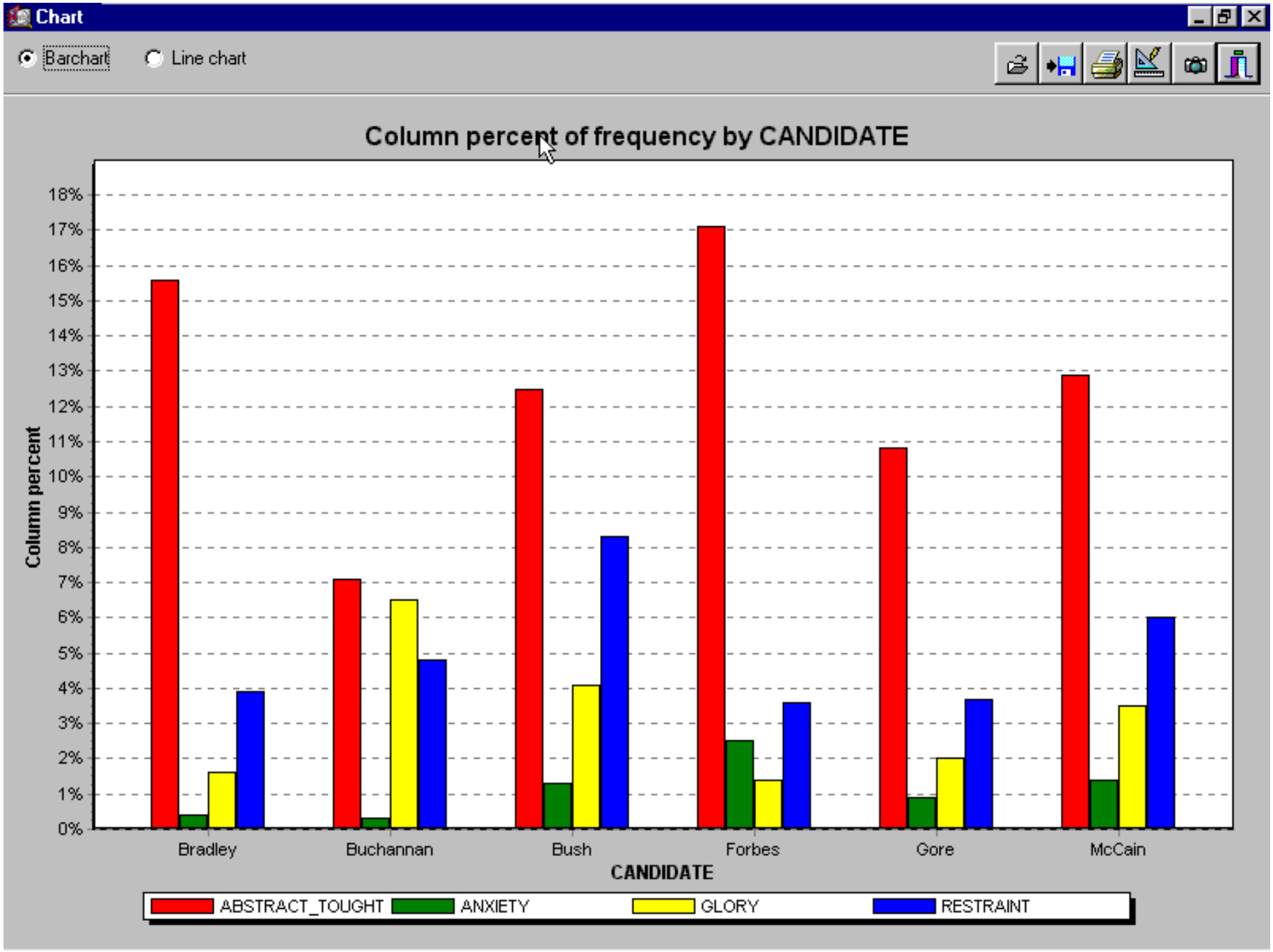
Dictionaries | Options | Word count | **Crosstab** | Key-Word-In-Context

Tabulate: Word frequency Display: Column percent Statistic: Student's F

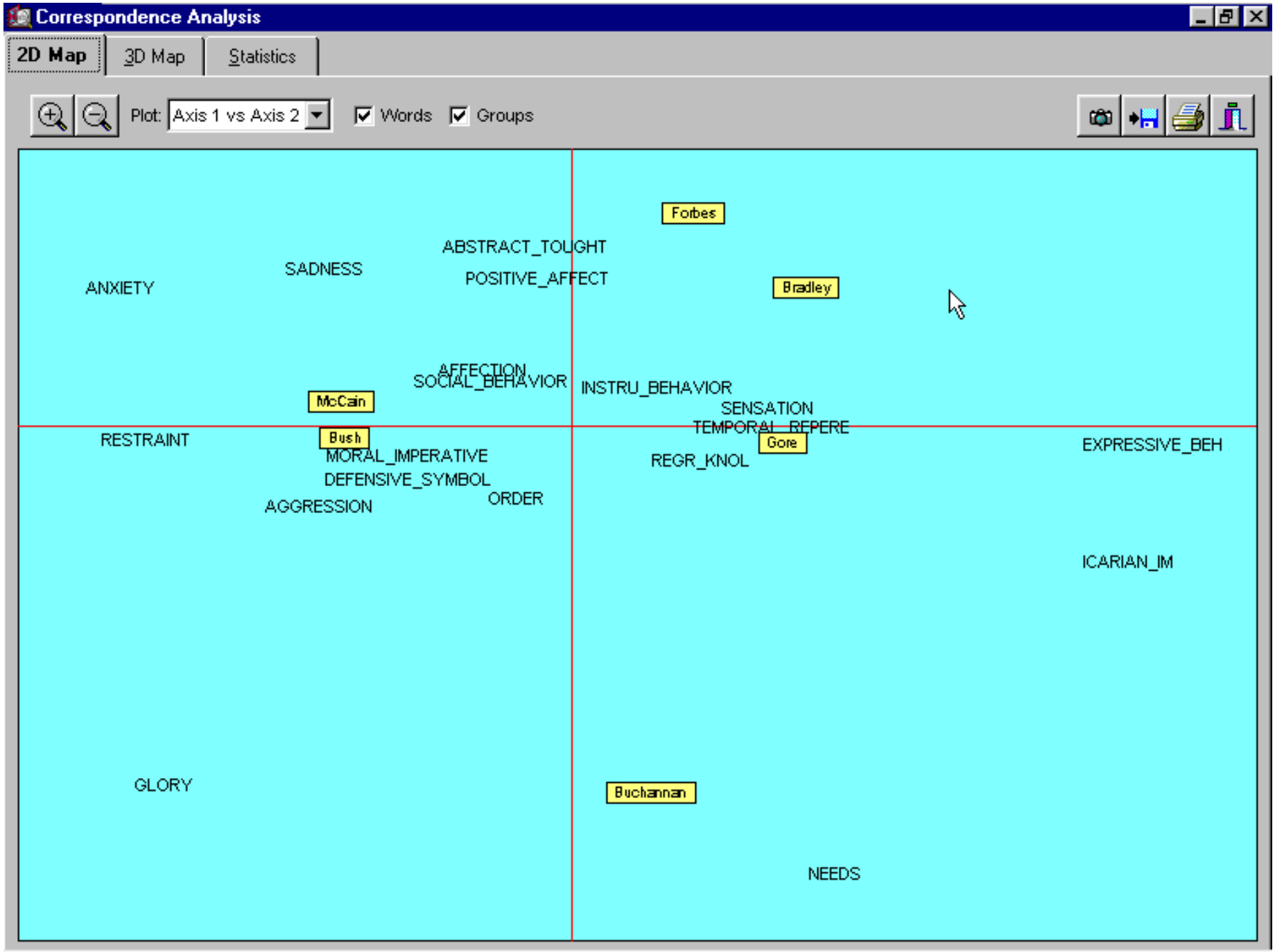
With: CANDIDATE Sort by: Word Frequency: Absolute Relative

	Bradley	Buchannan	Bush	Forbes	Gore	McCain	F	P (2-tails)
ABSTRACT_TOUGHT	15.6%	7.1%	12.5%	17.1%	10.8%	12.9%	7.524	0.000
ANXIETY	0.4%	0.3%	1.3%	2.5%	0.9%	1.4%	5.555	0.001
GLORY	1.6%	6.5%	4.1%	1.4%	2.0%	3.5%	5.084	0.002
RESTRAINT	3.9%	4.8%	8.3%	3.6%	3.7%	6.0%	4.795	0.003
NEEDS	0.7%	2.3%	0.6%	0.7%	1.0%	0.7%	4.562	0.004
SADNESS	0.3%	0.3%	0.4%	1.8%	0.5%	0.7%	4.430	0.005
ICARIAN_IM	4.3%	5.3%	3.6%	3.6%	5.4%	2.0%	3.596	0.014
TEMPORAL_REPERE	11.6%	11.1%	6.7%	10.3%	11.1%	9.2%	3.292	0.020
DEFENSIVE_SYMBOL	2.3%	4.0%	3.4%	5.0%	4.0%	4.6%	3.024	0.029
REGR_KNOL	16.4%	18.2%	13.5%	16.7%	16.6%	14.2%	2.571	0.052
SOCIAL_BEHAVIOR	12.6%	10.0%	12.7%	7.8%	11.4%	12.9%	2.067	0.104
SENSATION	5.6%	4.9%	5.9%	7.1%	6.3%	4.4%	1.870	0.136
AGGRESSION	3.5%	6.3%	5.5%	3.2%	5.9%	7.6%	1.546	0.212
AFFECTION	2.2%	1.7%	3.0%	2.1%	1.6%	1.7%	1.437	0.246
MORAL_IMPERATIVE	3.0%	3.5%	4.3%	2.1%	3.5%	4.0%	1.338	0.281
ORDER	1.9%	2.5%	2.3%	1.8%	1.7%	1.9%	0.932	0.477
EXPRESSIVE_BEH	0.6%	0.5%	0.2%	0.4%	0.3%	0.2%	0.826	0.543
INSTRU_BEHAVIOR	12.8%	10.3%	11.3%	12.1%	13.1%	11.6%	0.655	0.661
POSITIVE_AFFECT	0.4%	0.3%	0.1%	0.7%	0.3%	0.4%	0.320	0.896

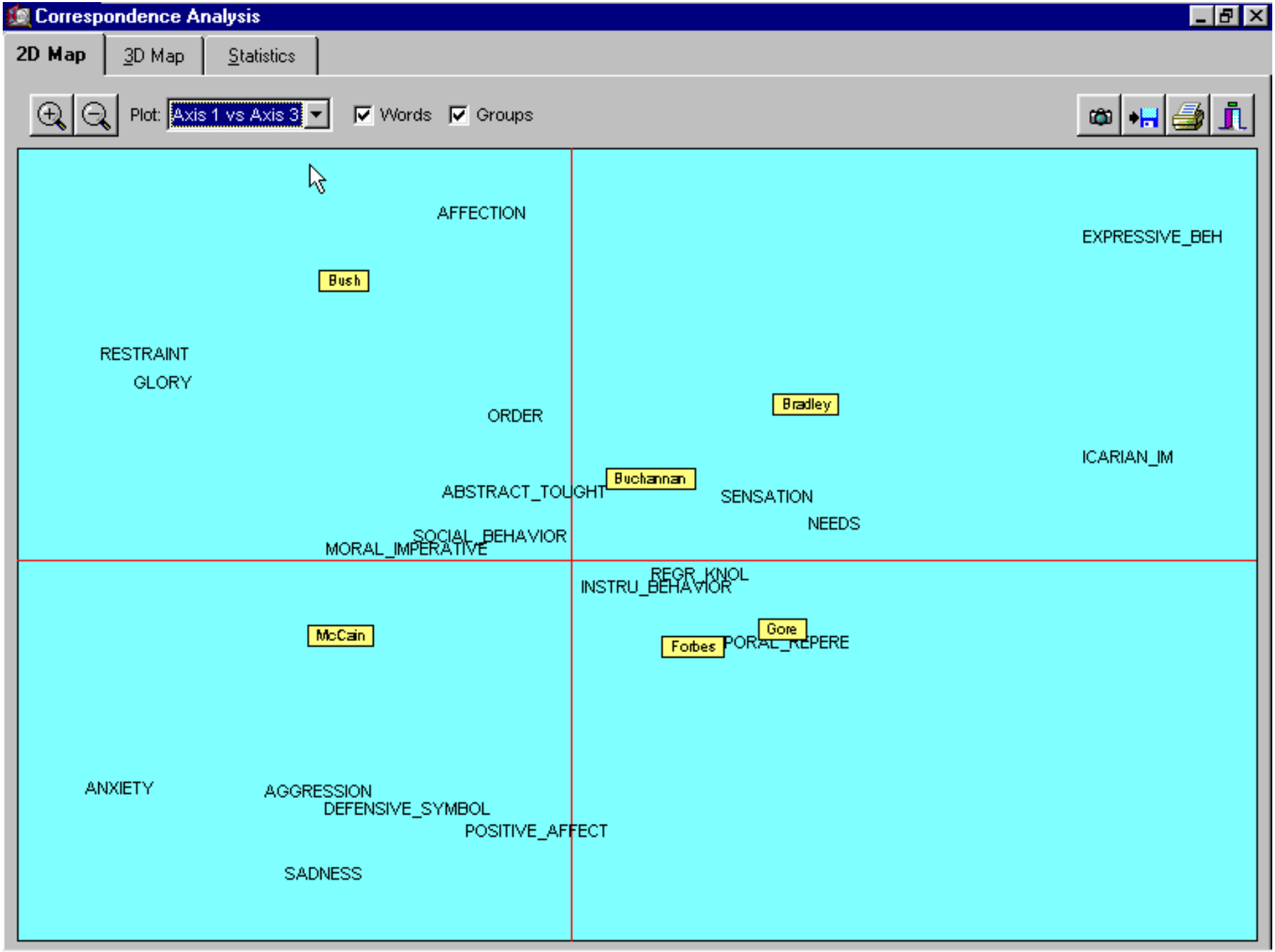
Relative Importance of Four Categories of Martindale RID among Presidential Candidates



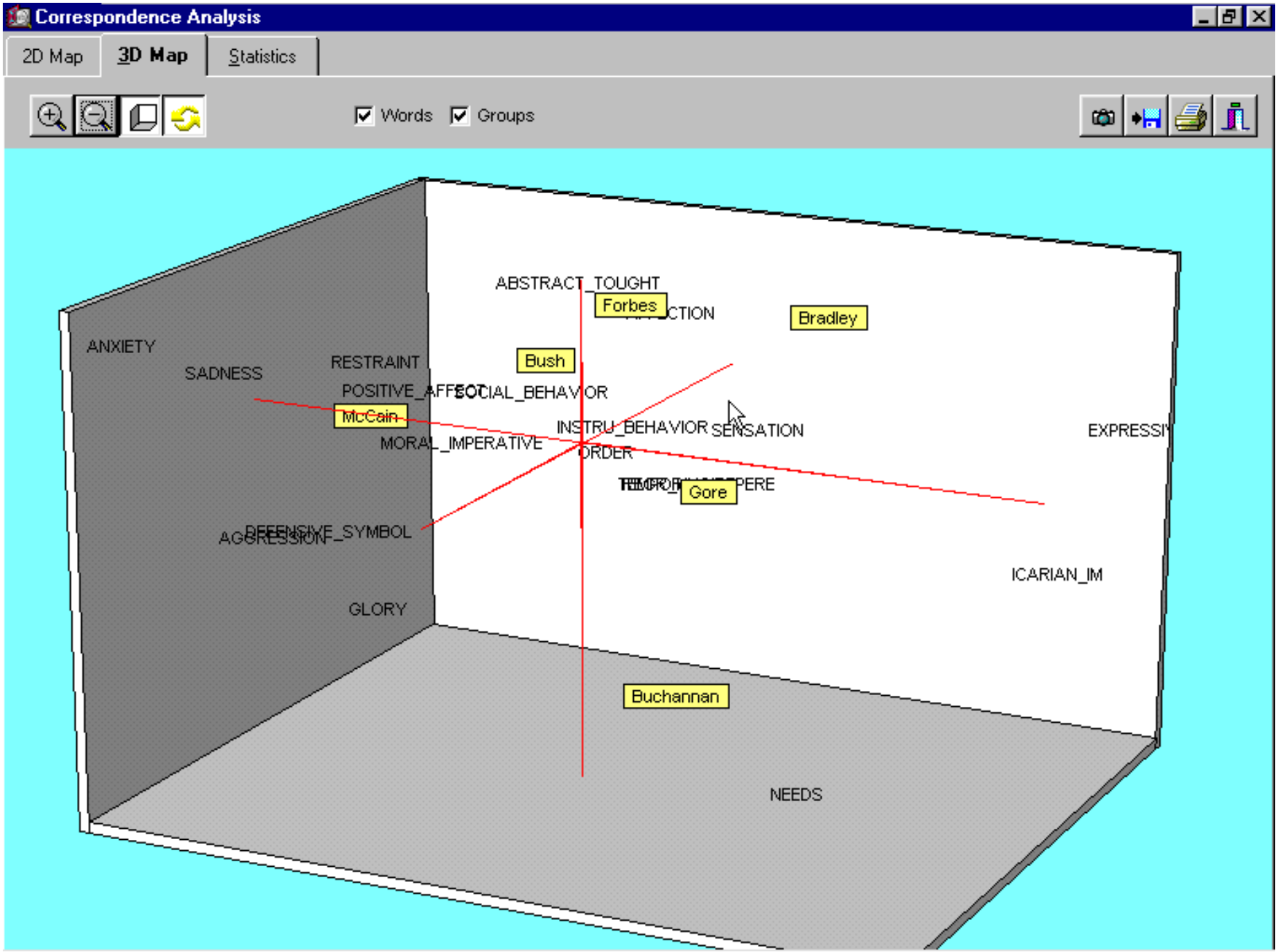
Correspondence Plot of Martindale RID Categories and Presidential Candidates (axis 1 & 2)



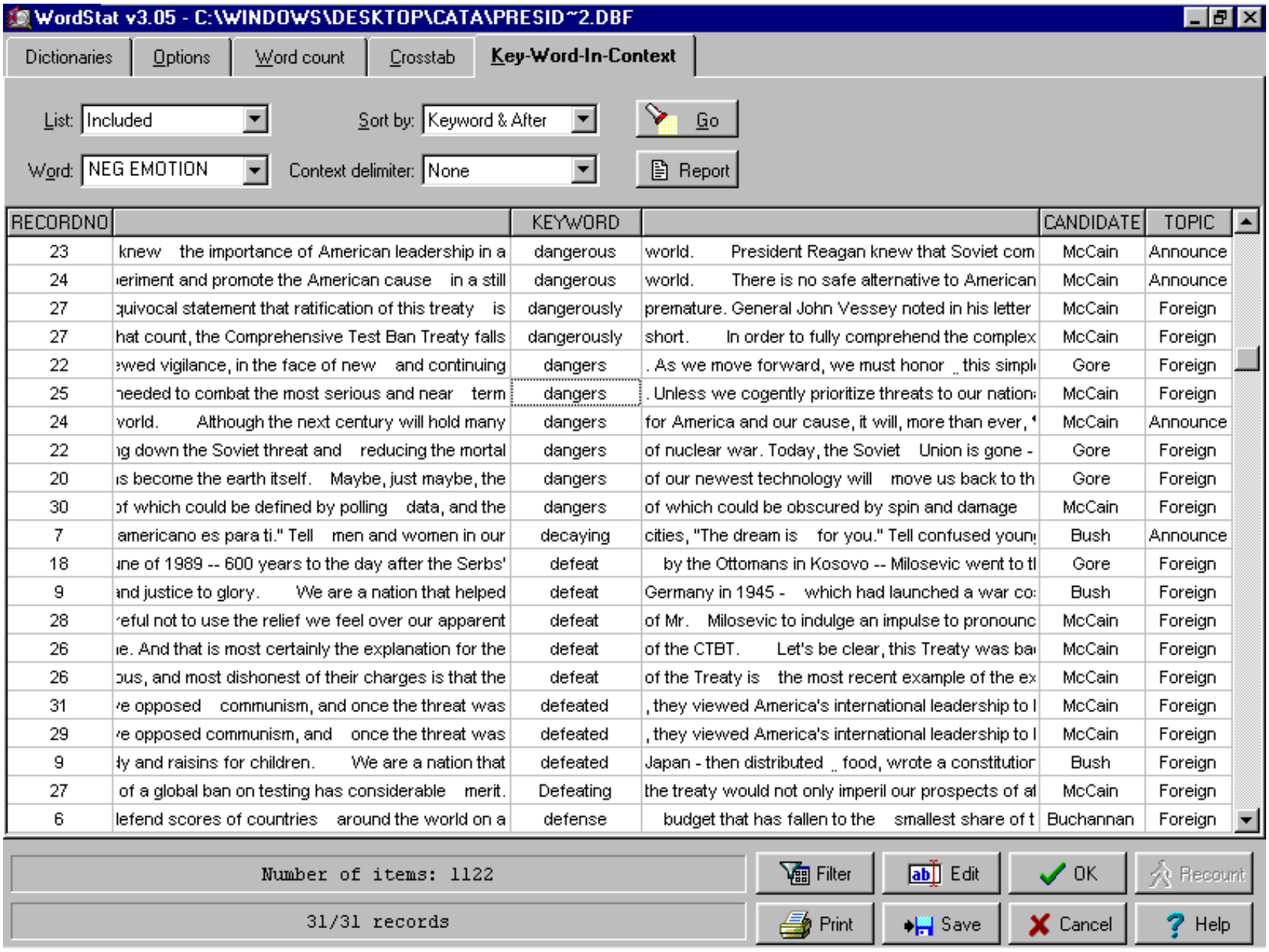
Correspondence Plot of Martindale RID Categories and Presidential Candidates (axis 1 & 3)



3D Correspondence Plot of Martindale RID Categories and Presidential Candidates

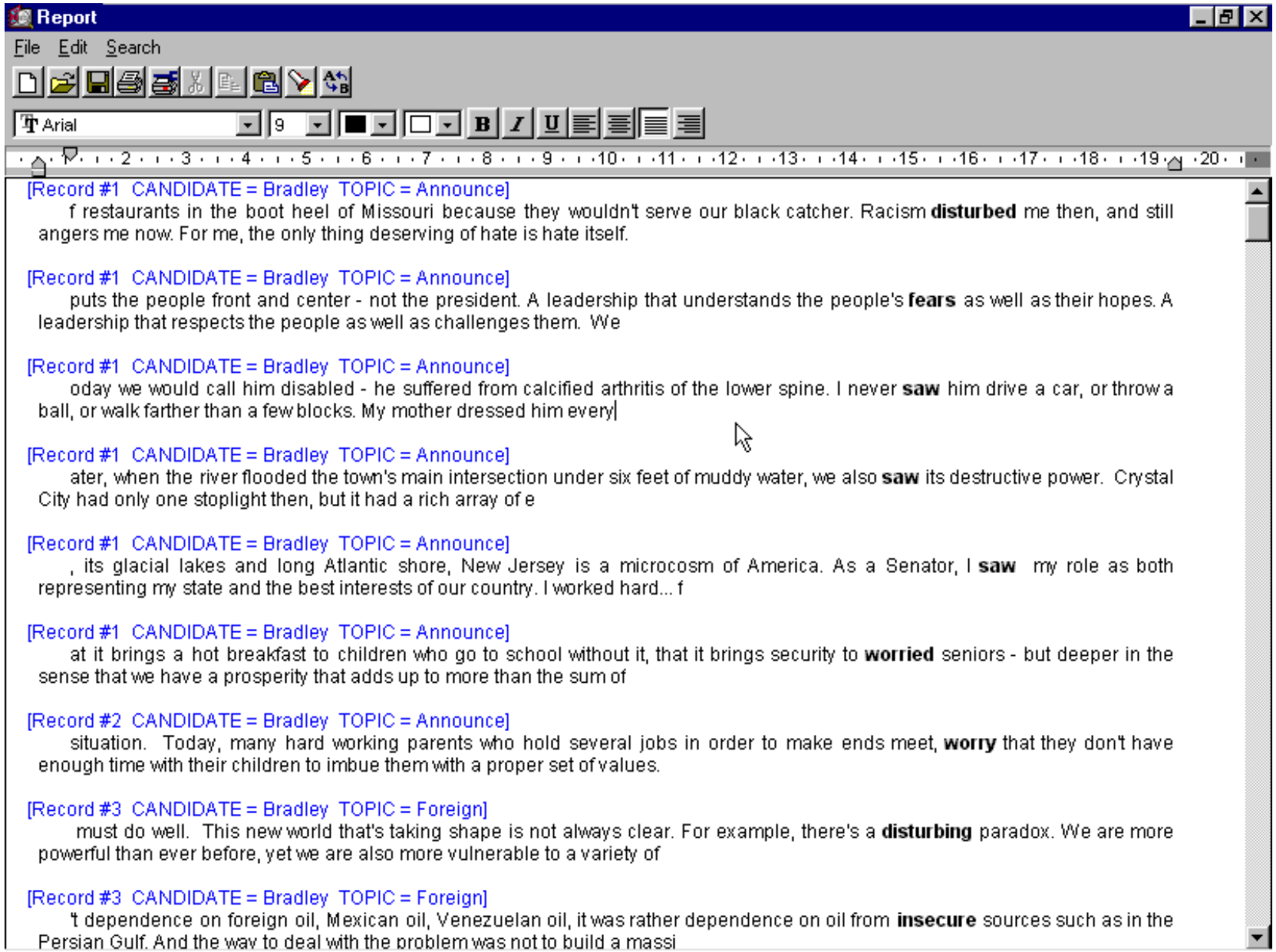


Sample KWIC List of Negative Emotion (LIWC) Words

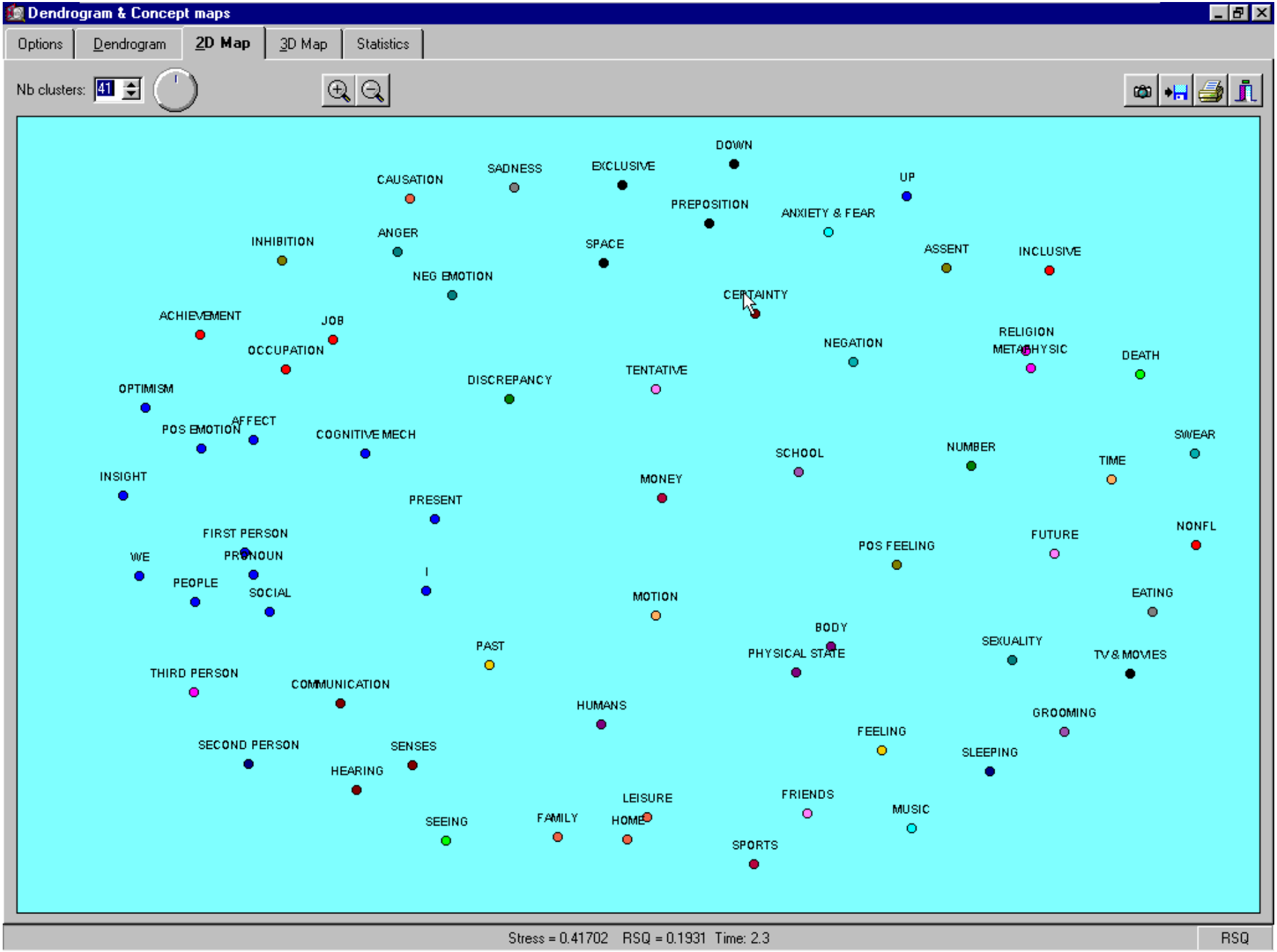


RECORDNO		KEYWORD		CANDIDATE	TOPIC
23	knew the importance of American leadership in a	dangerous	world. President Reagan knew that Soviet com	McCain	Announce
24	ıeriment and promote the American cause in a still	dangerous	world. There is no safe alternative to American	McCain	Announce
27	quivocal statement that ratification of this treaty is	dangerously	premature. General John Vessey noted in his letter	McCain	Foreign
27	hat count, the Comprehensive Test Ban Treaty falls	dangerously	short. In order to fully comprehend the complex	McCain	Foreign
22	ewed vigilance, in the face of new and continuing	dangers	. As we move forward, we must honor this simpl	Gore	Foreign
25	ıeeded to combat the most serious and near term	dangers	. Unless we cogently prioritize threats to our nation	McCain	Foreign
24	vorld. Although the next century will hold many	dangers	for America and our cause, it will, more than ever,	McCain	Announce
22	ıg down the Soviet threat and reducing the mortal	dangers	of nuclear war. Today, the Soviet Union is gone -	Gore	Foreign
20	ıs become the earth itself. Maybe, just maybe, the	dangers	of our newest technology will move us back to th	Gore	Foreign
30	of which could be defined by polling data, and the	dangers	of which could be obscured by spin and damage	McCain	Foreign
7	americano es para ti." Tell men and women in our	decaying	cities, "The dream is for you." Tell confused youn	Bush	Announce
18	ıne of 1989 -- 600 years to the day after the Serbs'	defeat	by the Ottomans in Kosovo -- Milosevic went to tl	Gore	Foreign
9	and justice to glory. We are a nation that helped	defeat	Germany in 1945 - which had launched a war co	Bush	Foreign
28	reful not to use the relief we feel over our apparent	defeat	of Mr. Milosevic to indulge an impulse to pronounc	McCain	Foreign
26	ıe. And that is most certainly the explanation for the	defeat	of the CTBT. Let's be clear, this Treaty was ba	McCain	Foreign
26	ous, and most dishonest of their charges is that the	defeat	of the Treaty is the most recent example of the ex	McCain	Foreign
31	ve opposed communism, and once the threat was	defeated	, they viewed America's international leadership to l	McCain	Foreign
29	ve opposed communism, and once the threat was	defeated	, they viewed America's international leadership to l	McCain	Foreign
9	ly and raisins for children. We are a nation that	defeated	Japan - then distributed food, wrote a constitutior	Bush	Foreign
27	of a global ban on testing has considerable merit.	Defeating	the treaty would not only imperil our prospects of al	McCain	Foreign
6	lefend scores of countries around the world on a	defense	budget that has fallen to the smallest share of t	Buchannan	Foreign

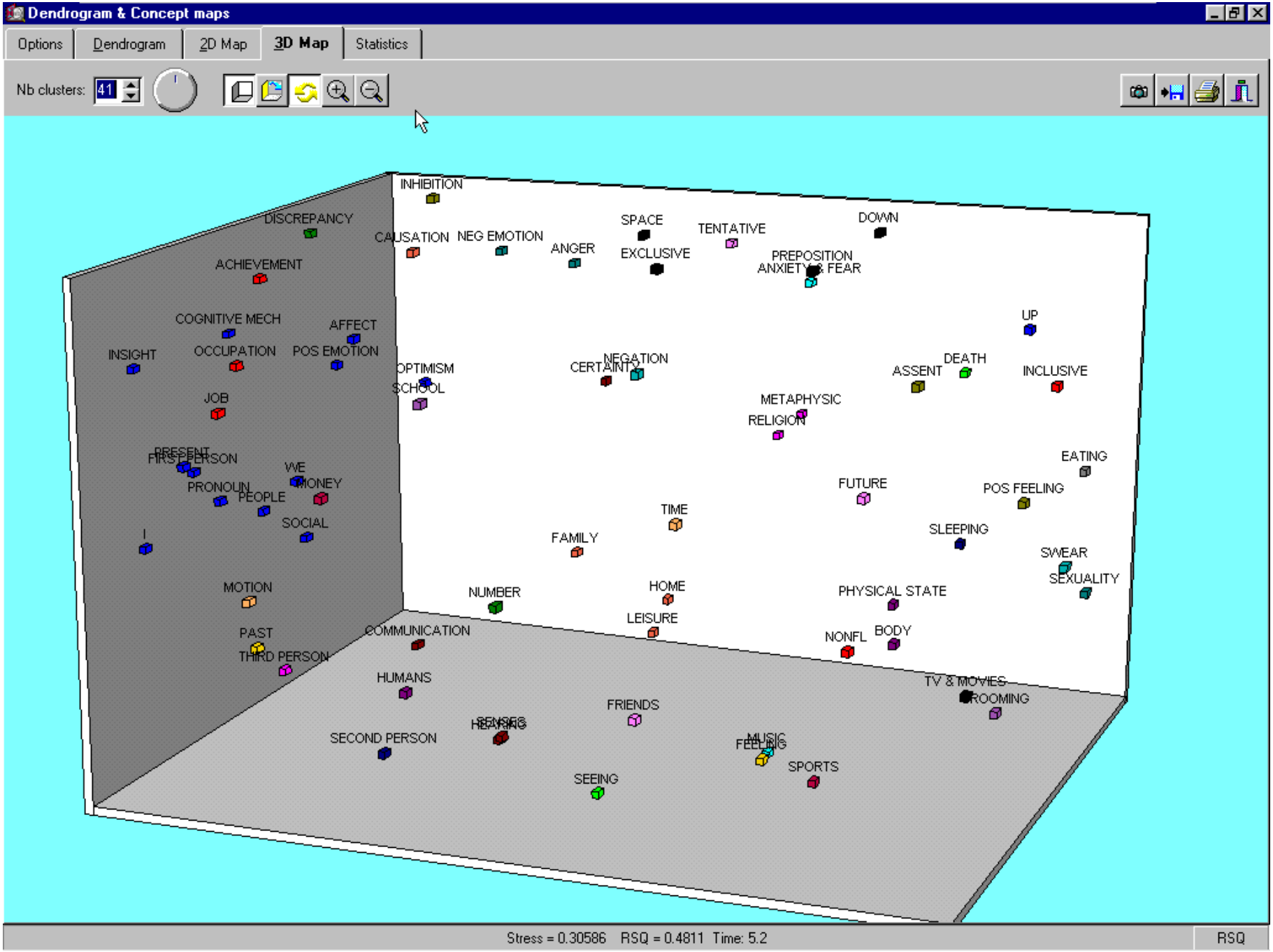
Sample Concordance Report of Negative Emotion Words (LIWC)



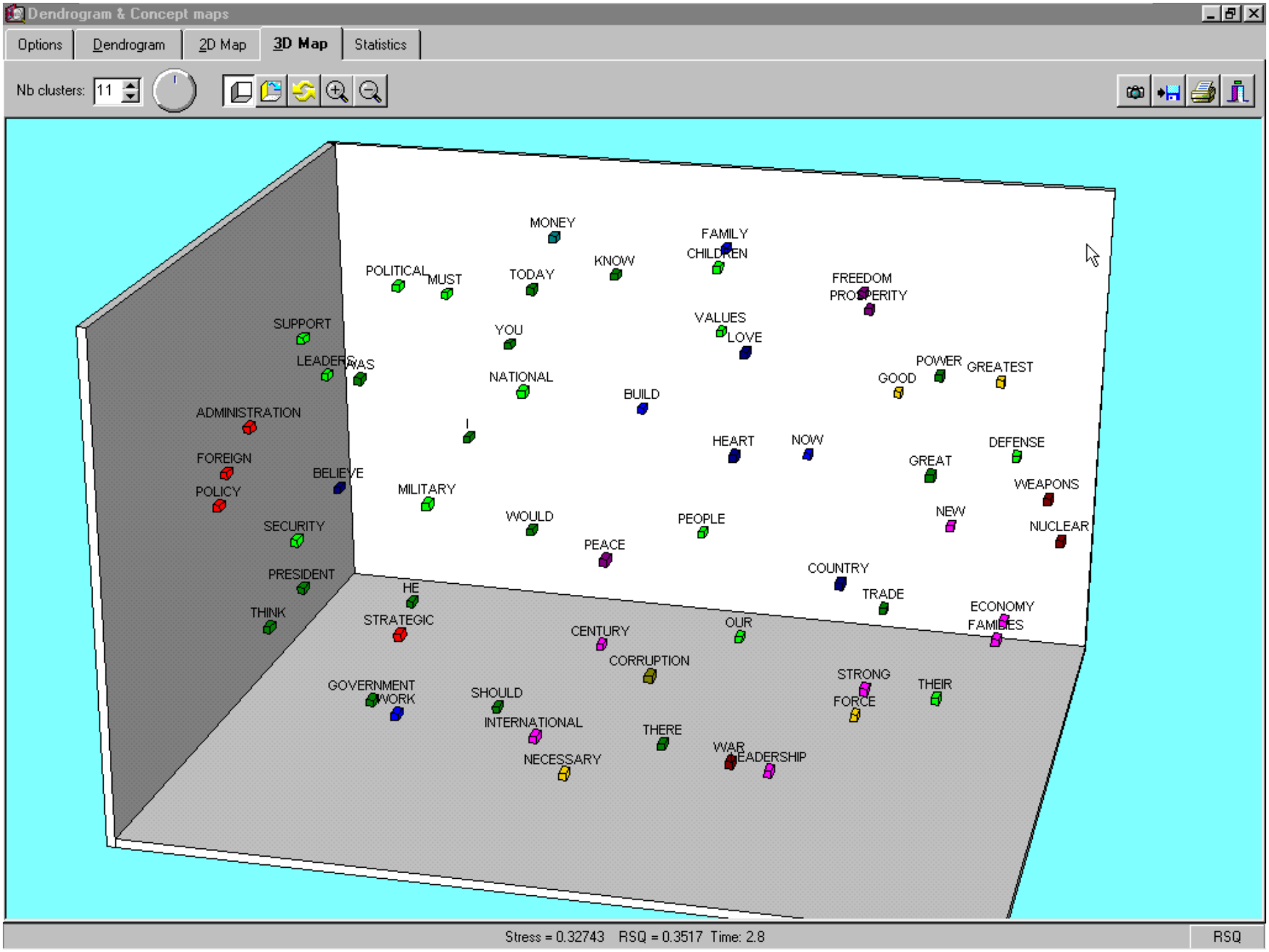
Sample 2D Multidimentional Scaling of LIWC categories



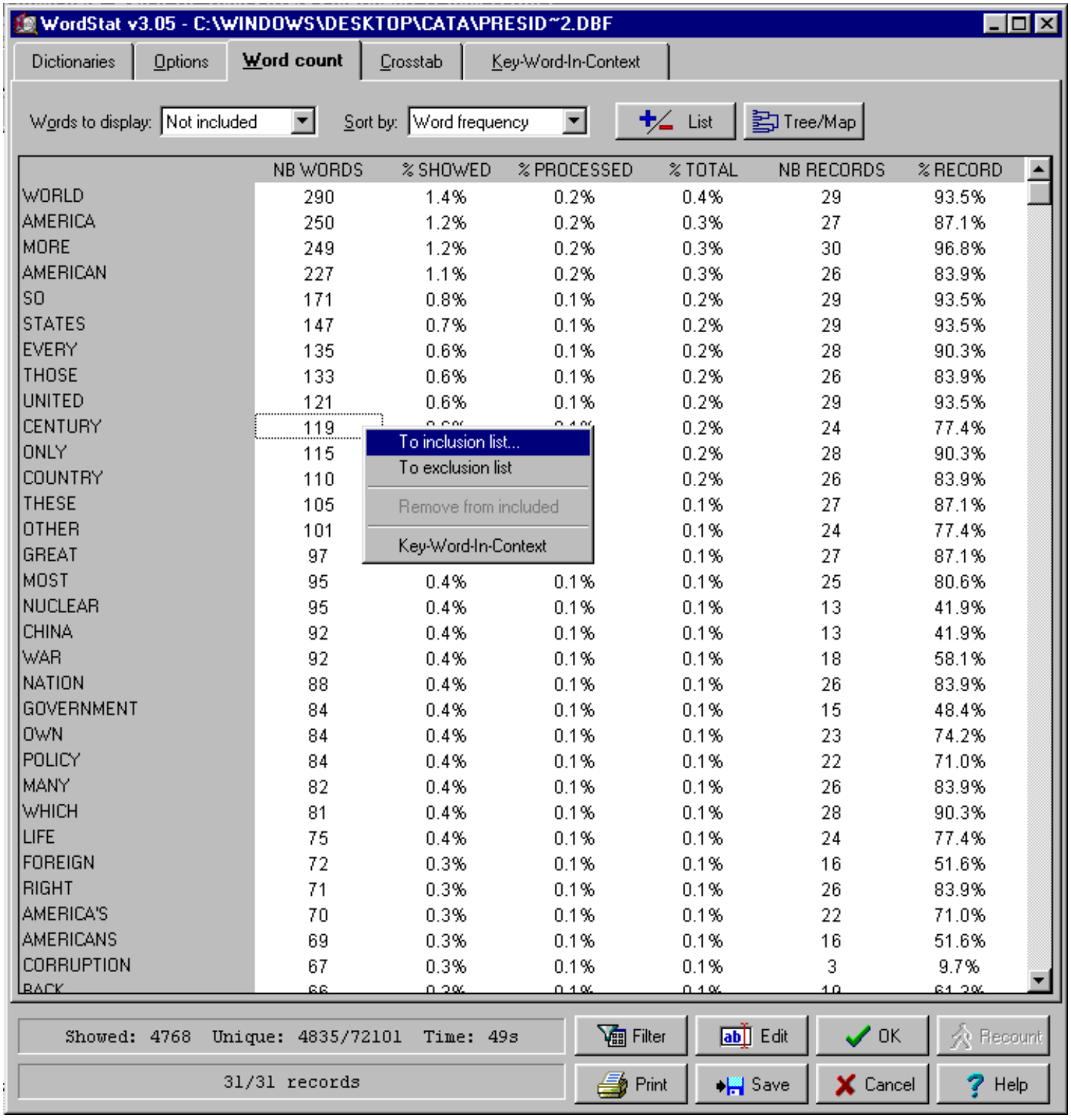
Sample 3D Multidimensional Scaling of LIWC Categories



Sample 3D Multidimensional Scaling of Frequent Words Used by Presidential Candidates

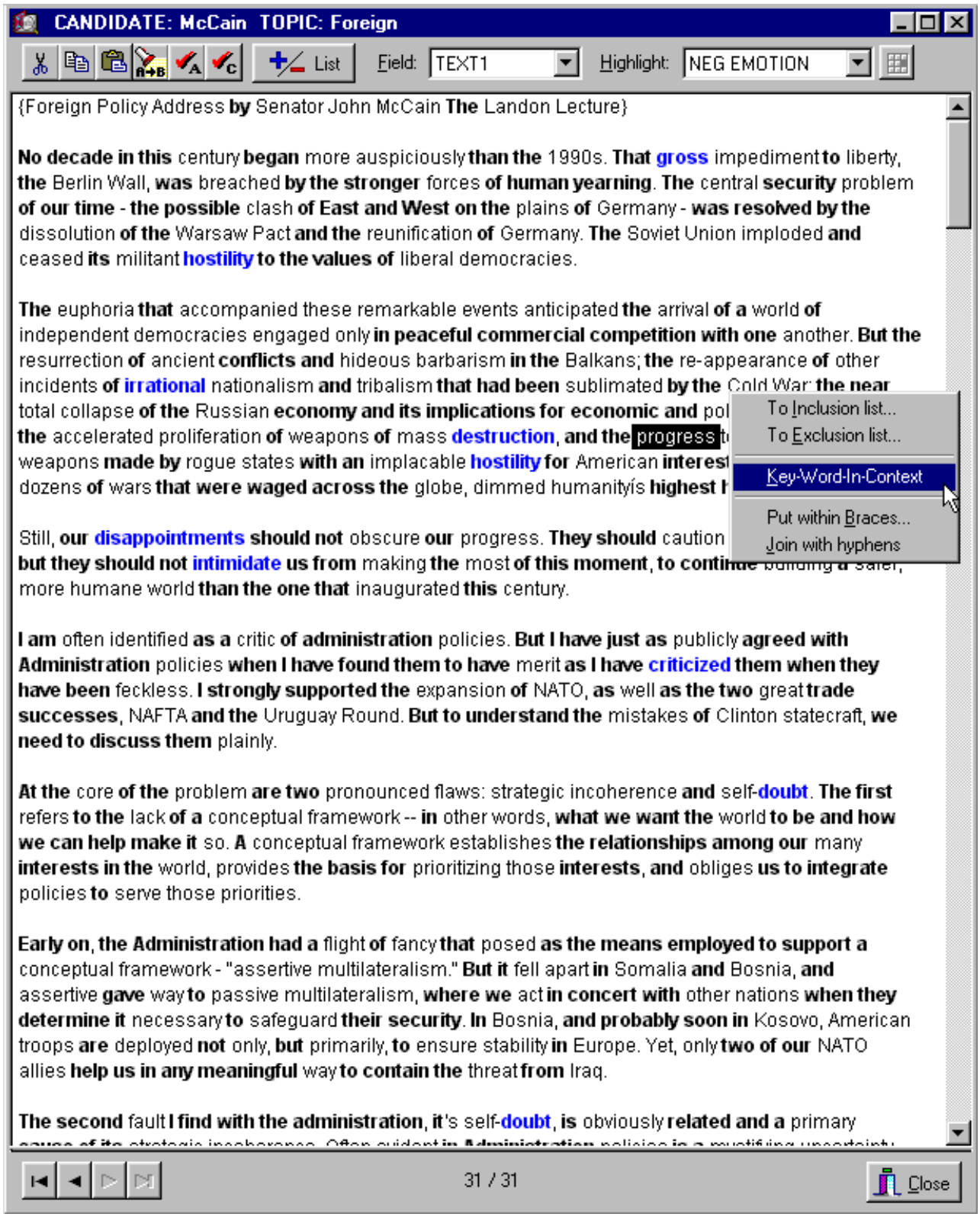


Frequency Table of Left Over Words (LIWC) in Descending Order of Frequency



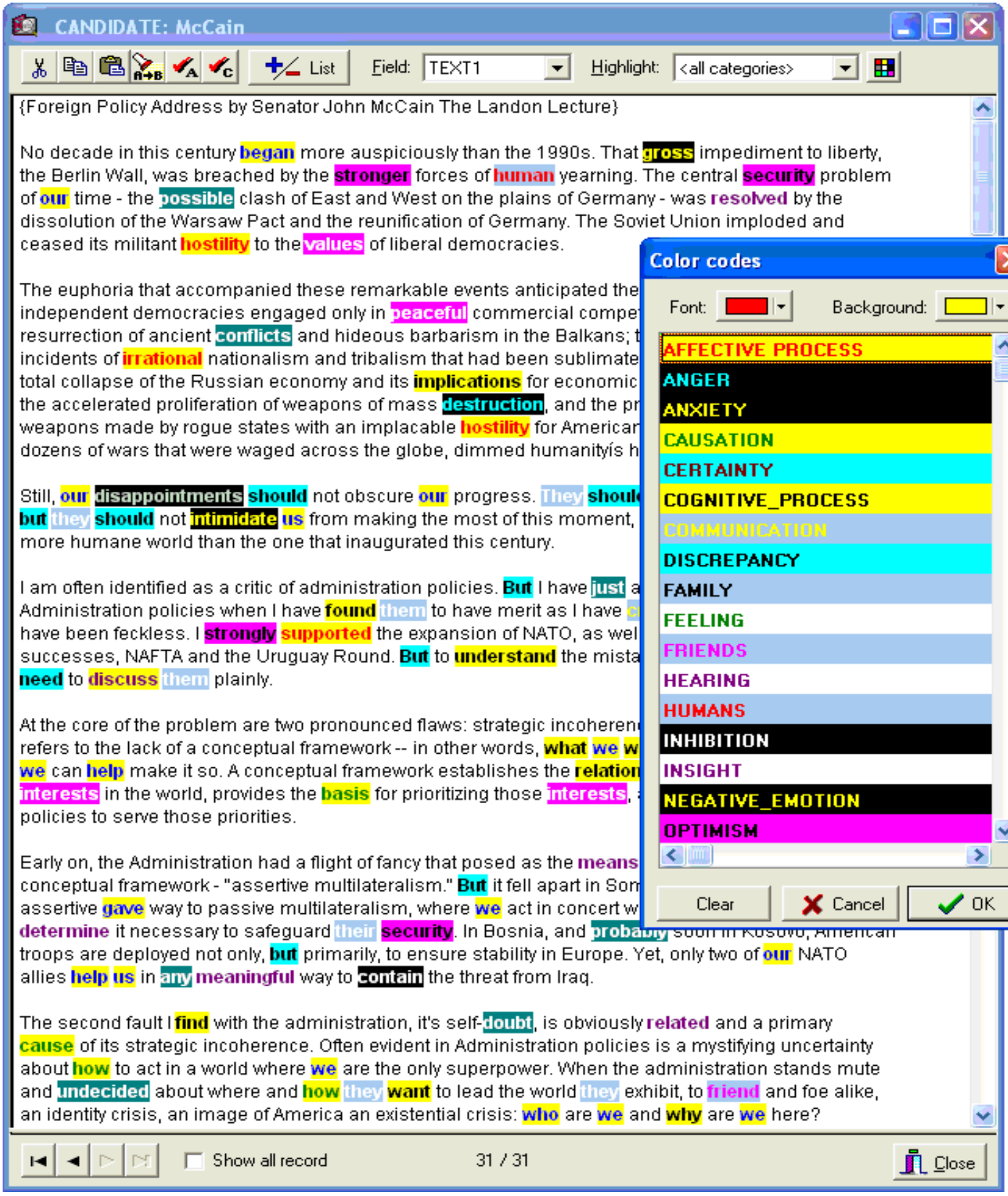
Note: The popup menu showed above allows one to add any left over word to an existing category, to the exclusion list or to quickly obtain a KIWC list of it.

Text Editor View of a Speech with all Coded Words in Bold and Negative Emotion Words (LIWC) in Blue.

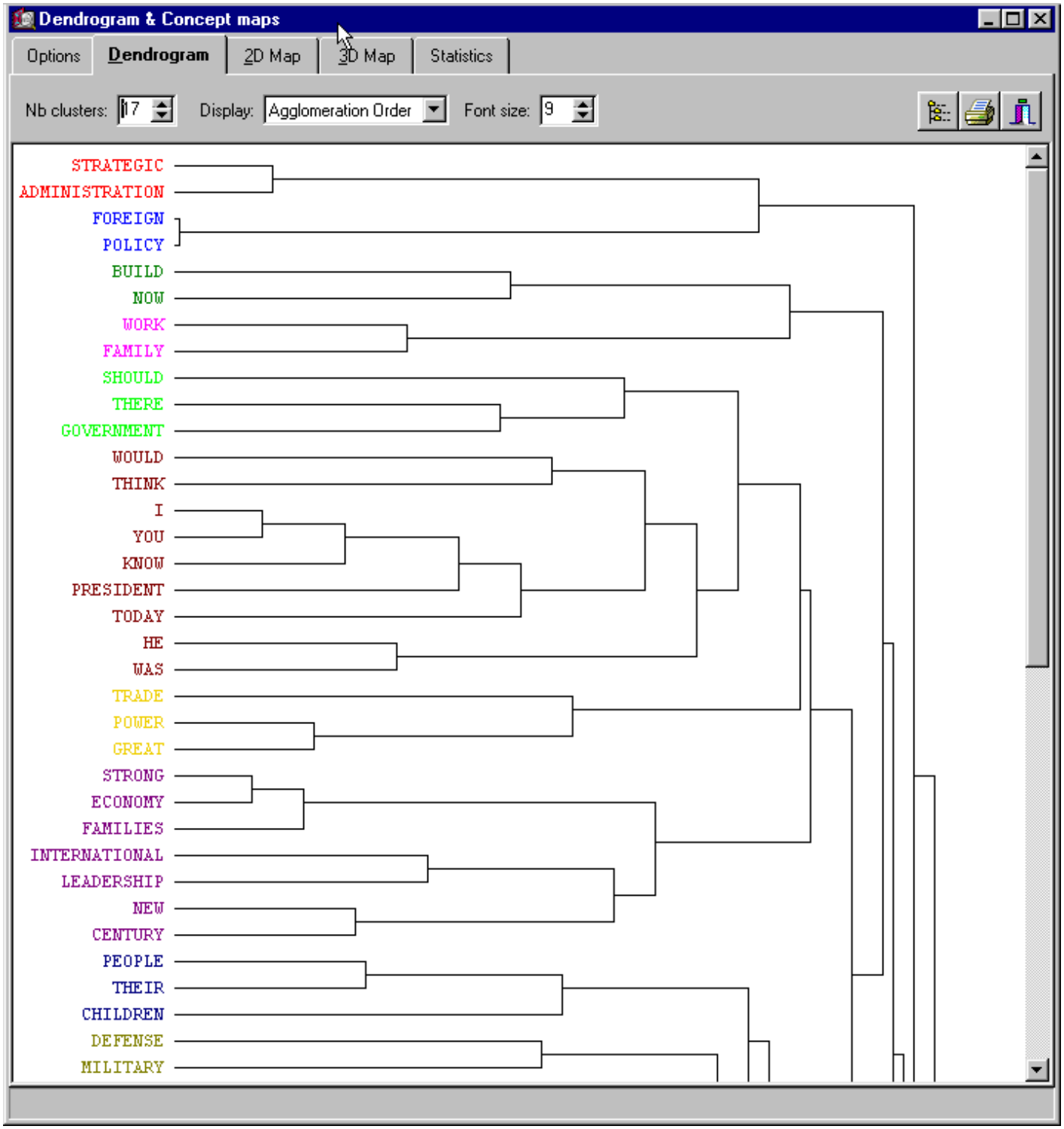


NOTE: The popup menu showed above allows interactive coding while in the editor (adding left over words in a category or to the exclusion dictionary).

Text Editor View of a Speech with Color Coded Words.



Sample Dendrogram of Frequent Words used by Presidential Candidates Based on a Cluster Analysis of Cooccurrence within Sentences.



Sample Dendrogram of LIWC Categories Based on a Cluster Analysis of Cooccurrence within Sentences.