

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

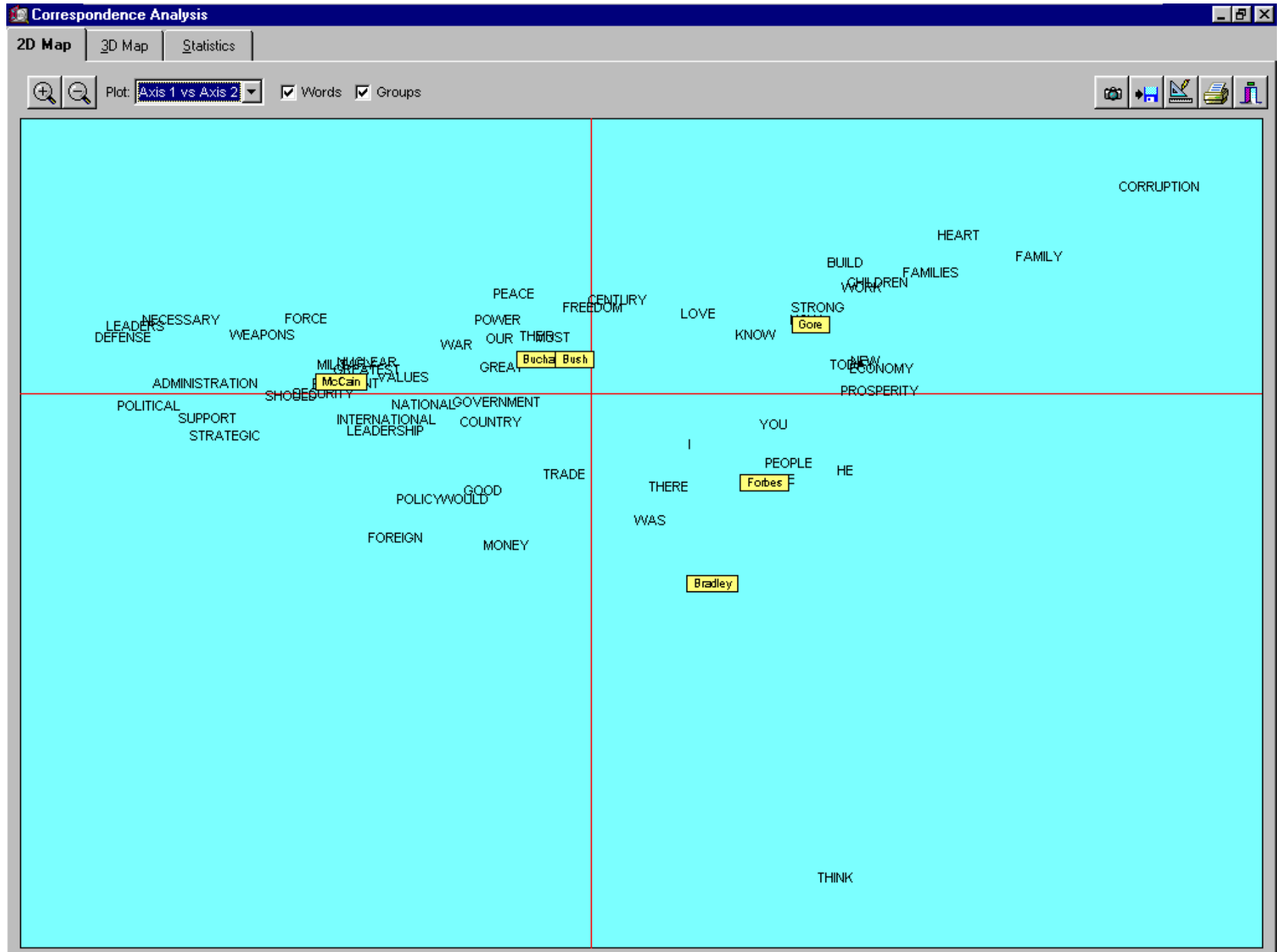
ANALYSIS FOR COMPARISON:

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

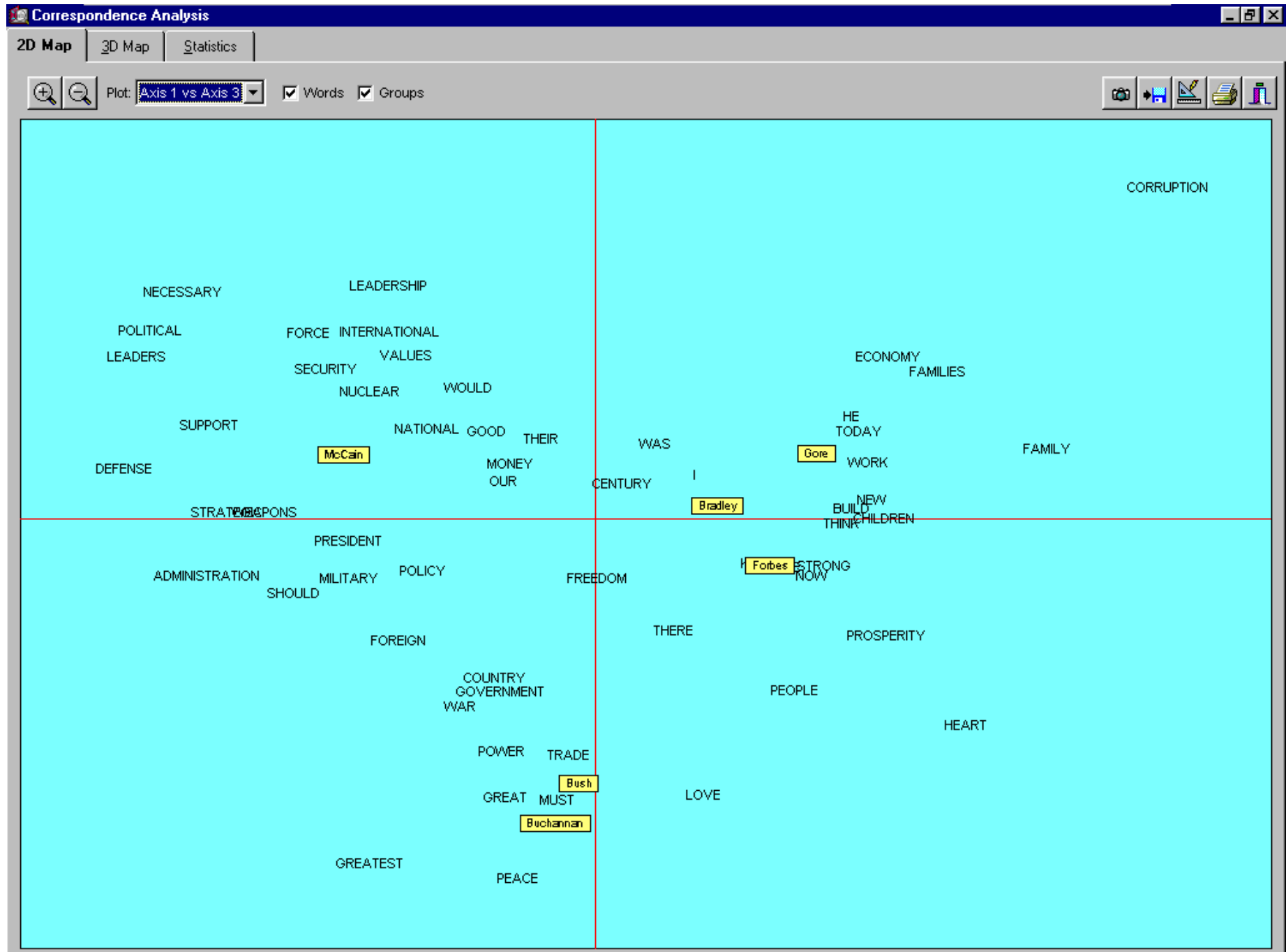
VALIDITY CHECK:

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

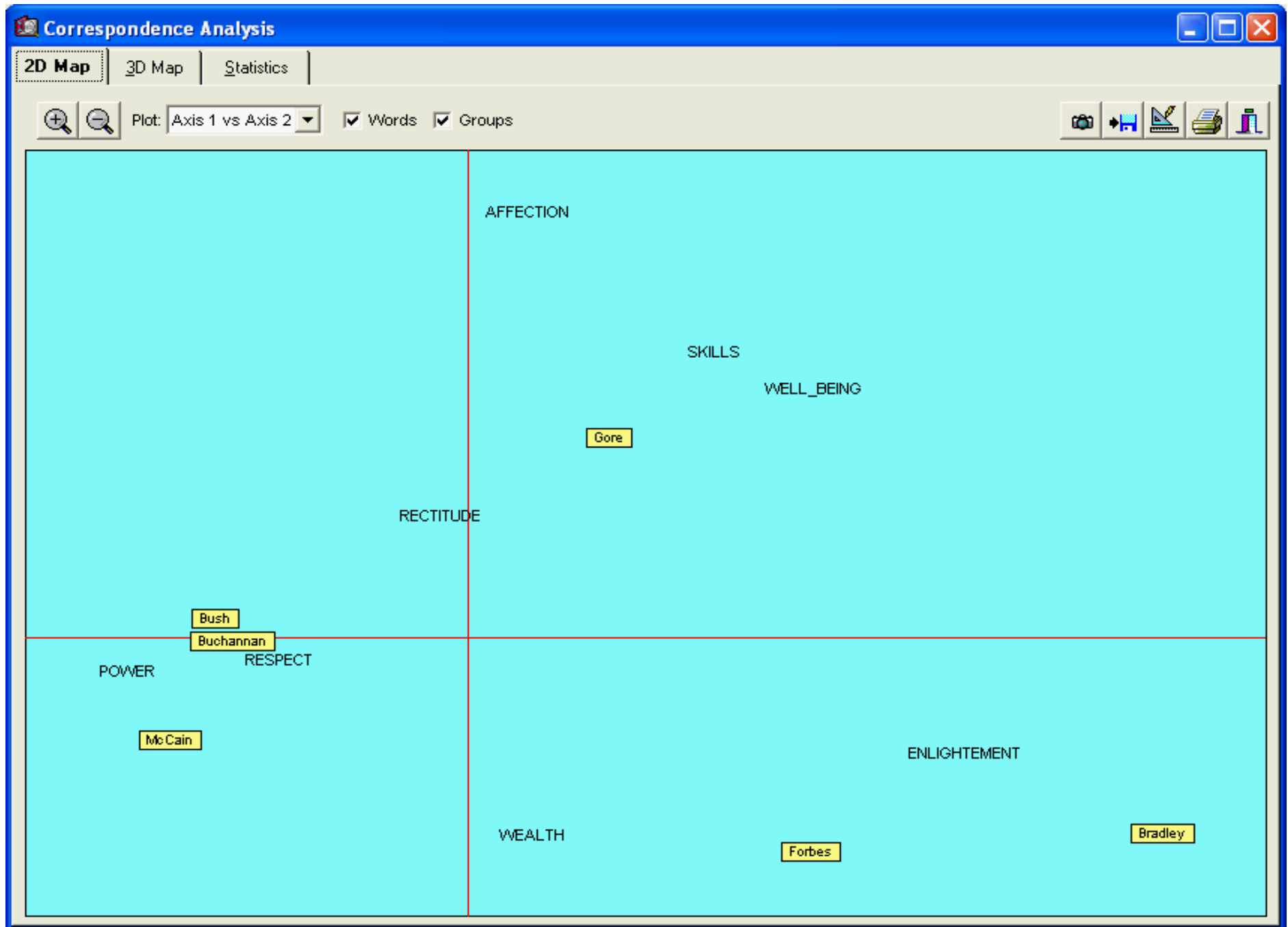
Correspondence Plot of Frequent Words and Presidential Candidates (axis 1 & 2)



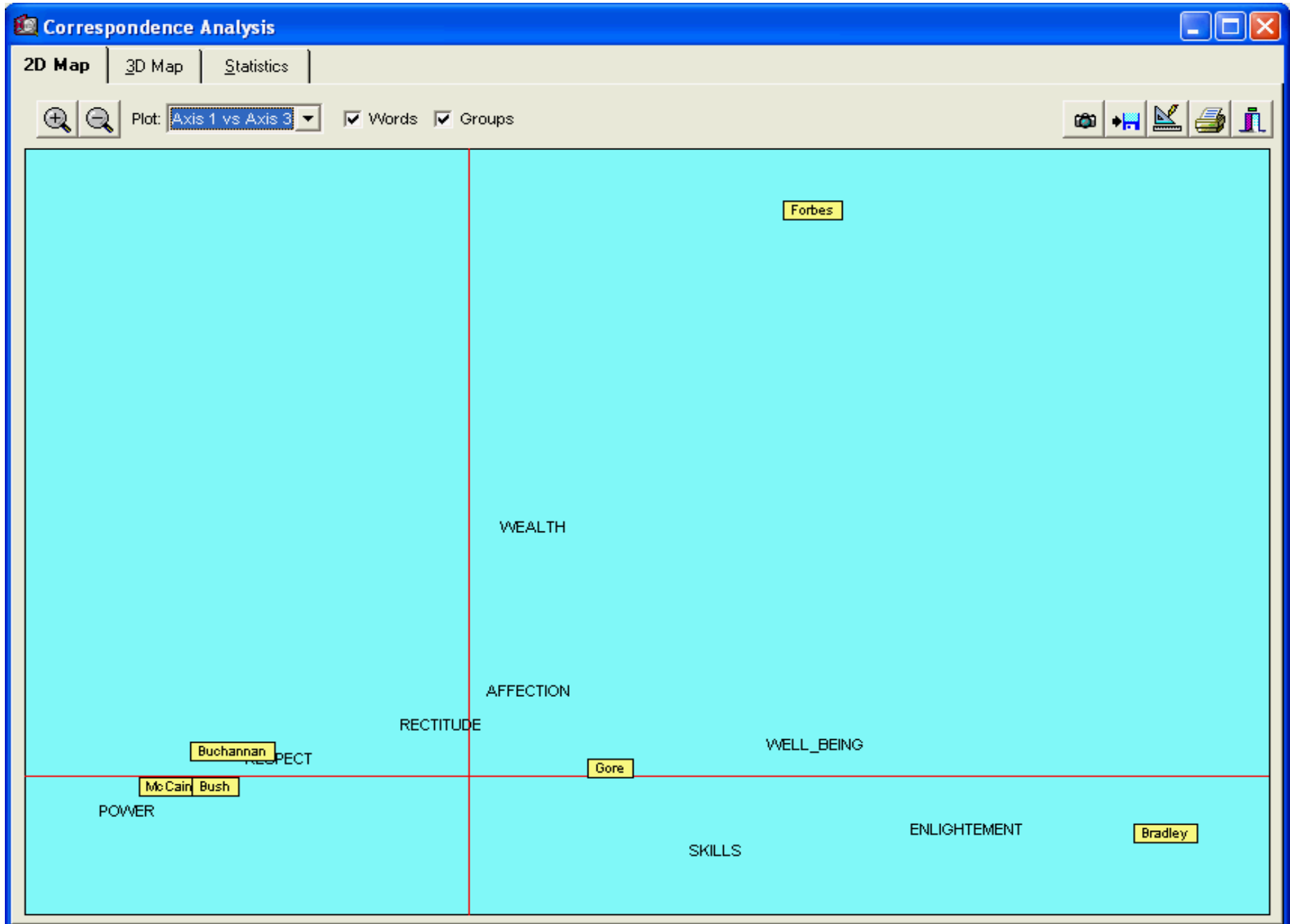
Correspondence Plot of Frequent Words and Presidential Candidates (axis 1 & 3)



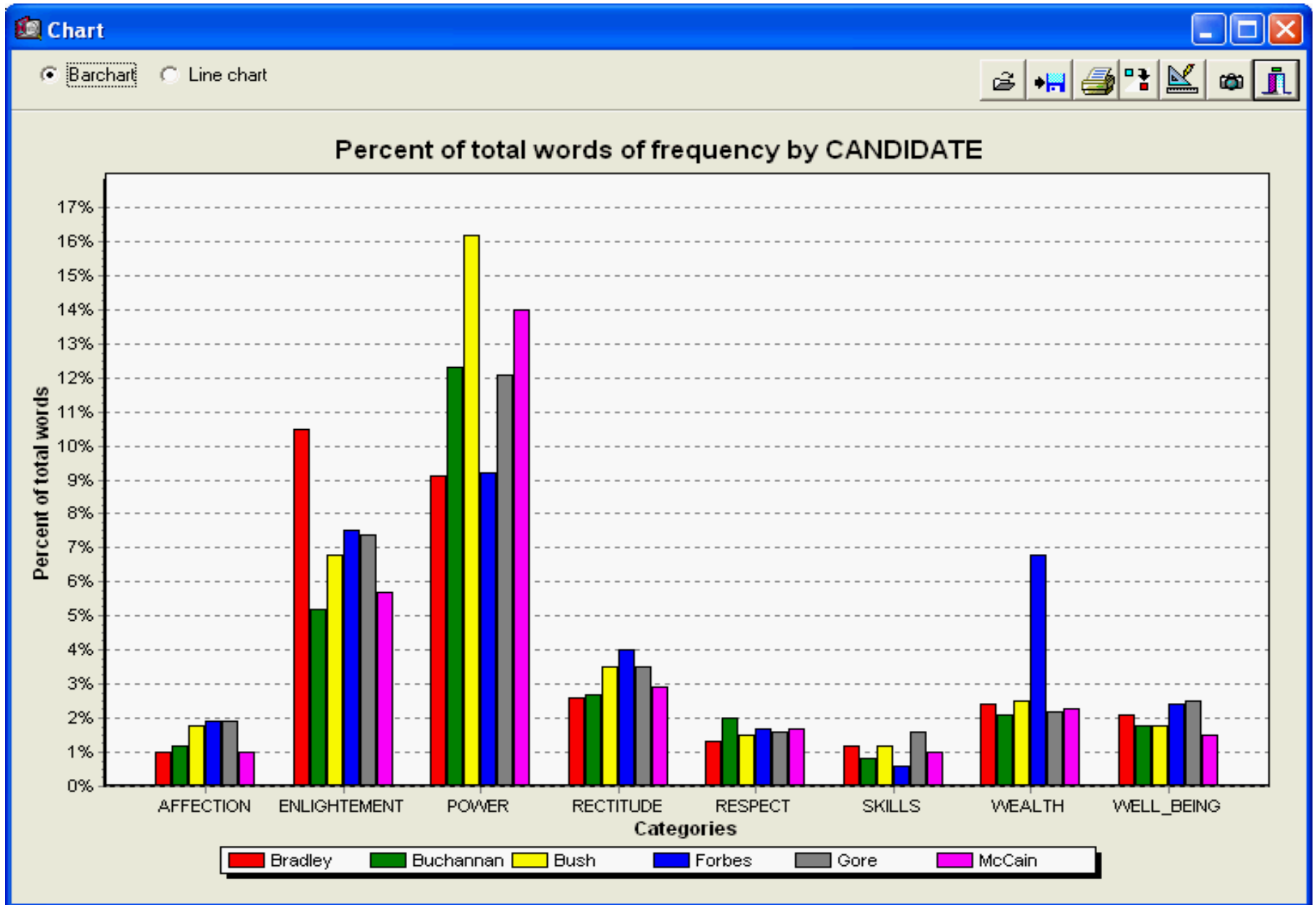
Correspondence Plot of Lasswell Values Dictionary and Presidential Candidates (axis 1 & 2)



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

The screenshot shows the WordStat v3.05 Crosstab window. The window title is "WordStat v3.05 - C:\WINDOWS\DESKTOP\CATA\PRESID~2.DBF". The "Crosstab" tab is active. The "Tabulate" dropdown is set to "Word frequency", "Display" is "Percent of total word", and "Statistic" is "Student's F". The "With" dropdown is set to "CANDIDATE" and "Sort by" is "Statistics". The "Frequency" options are "Absolute" (unselected) and "Relative" (selected).

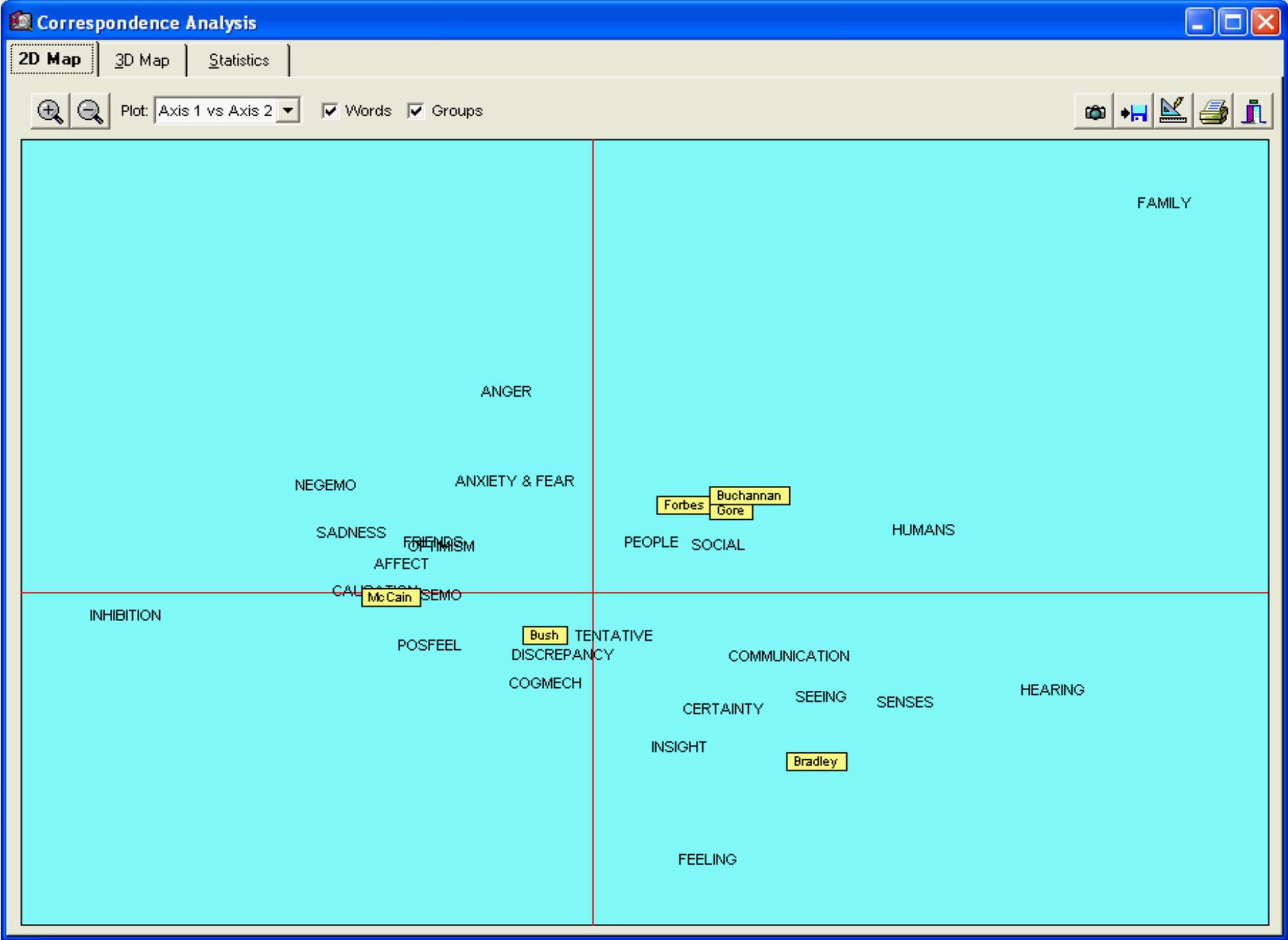
	Bradley	Buchanan	Bush	Forbes	Gore	McCain	F	P (2-tails)
EXCLUSIVE	4.6%	2.9%	3.0%	2.4%	2.7%	3.5%	20.321	0.000
INHIBITION	0.3%	0.5%	0.8%	0.5%	0.4%	1.0%	10.727	0.000
AFFECT	3.5%	3.6%	6.1%	5.1%	4.6%	6.6%	8.732	0.000
NUMBER	0.9%	1.1%	0.6%	0.5%	1.2%	0.7%	8.559	0.000
MOTION	0.9%	0.8%	0.9%	1.7%	0.8%	0.5%	8.149	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%	1.8%	1.8%	1.8%	3.731	0.012
ASSENT	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	3.600	0.014
SENSES	1.6%	1.1%	1.2%	1.2%	1.2%	1.2%	3.467	0.016
SPACE	2.7%	2.3%	1.7%	1.7%	1.7%	1.7%	3.428	0.017
CERTAINTY	1.5%	1.1%	0.7%	0.7%	0.7%	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
CADRECC	0.3%	0.3%	0.3%	0.5%	0.3%	0.5%	2.476	0.058

A context menu is open over the "SEXUALITY" row, showing options: "Add to excluded", "Remove from Included", "Key-Word-In-Context", and "Chart selected rows".

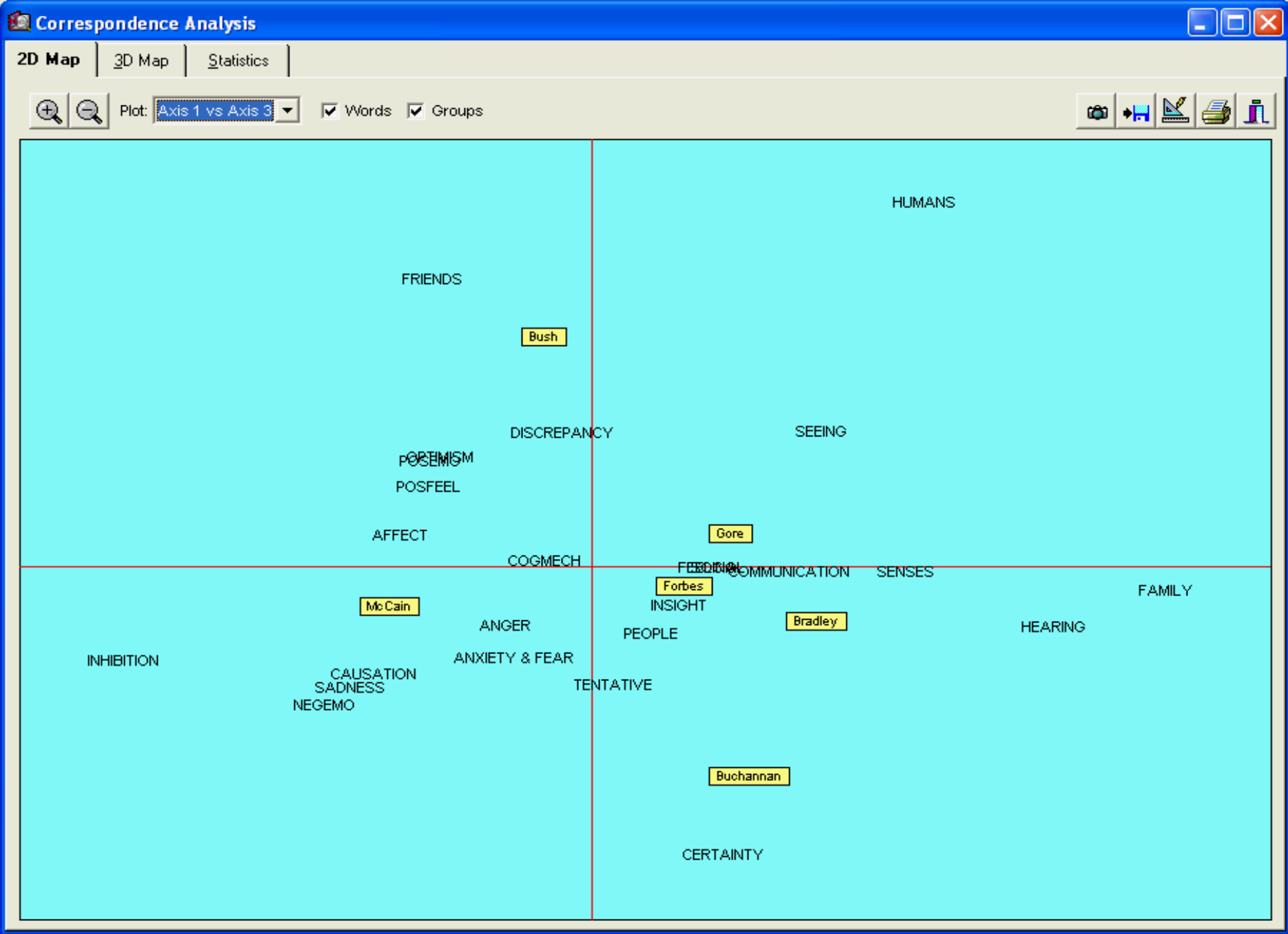
At the bottom of the window, there are buttons for "Filter", "Edit", "OK", "Recount", "Print", "Save", "Cancel", and "Help". The status bar shows "31/31 records".

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

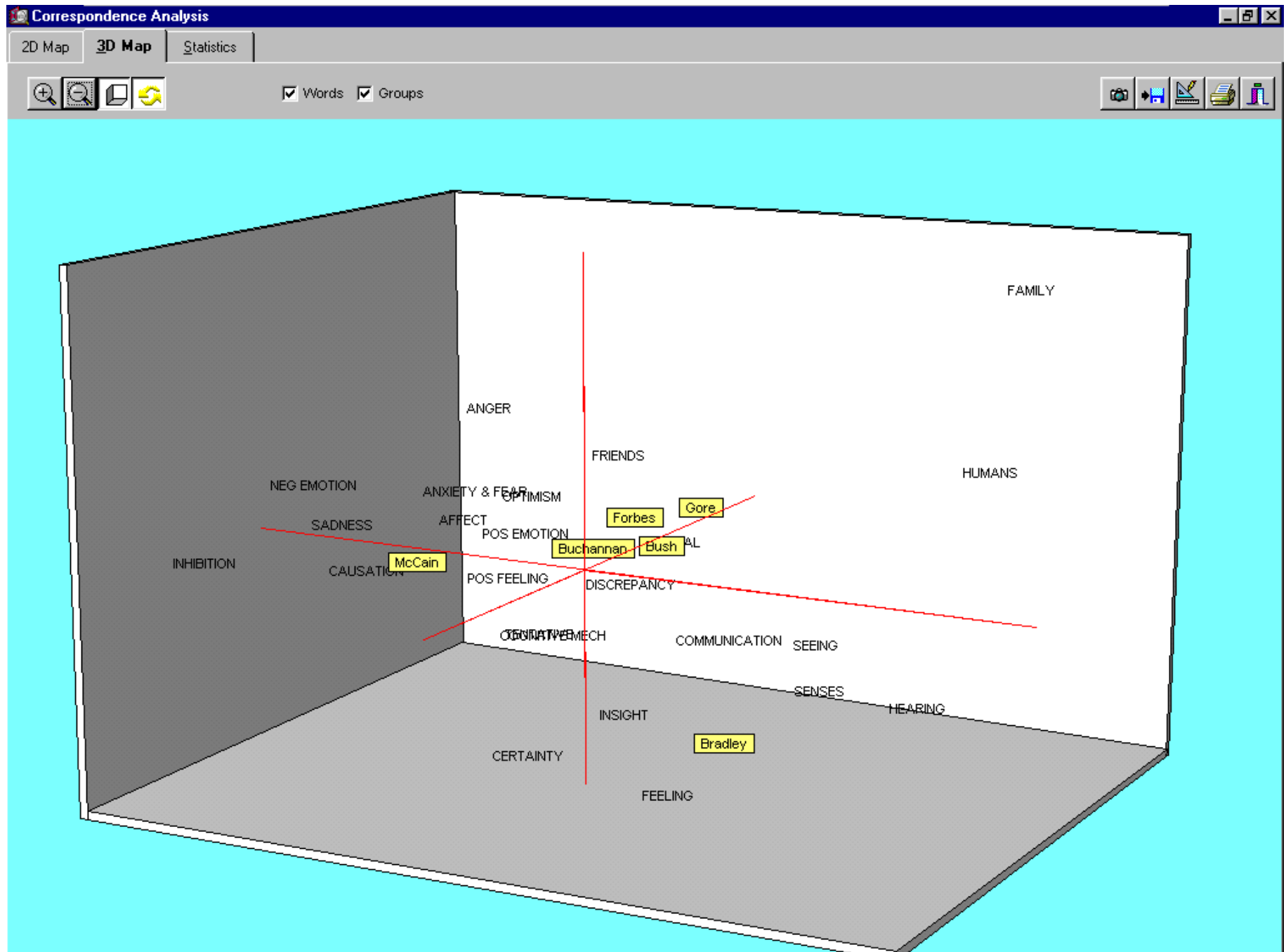
Correspondence Plot of Pennebaker's LIWC Psychological Categories and Presidential Candidates (axis 1 & 2)



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

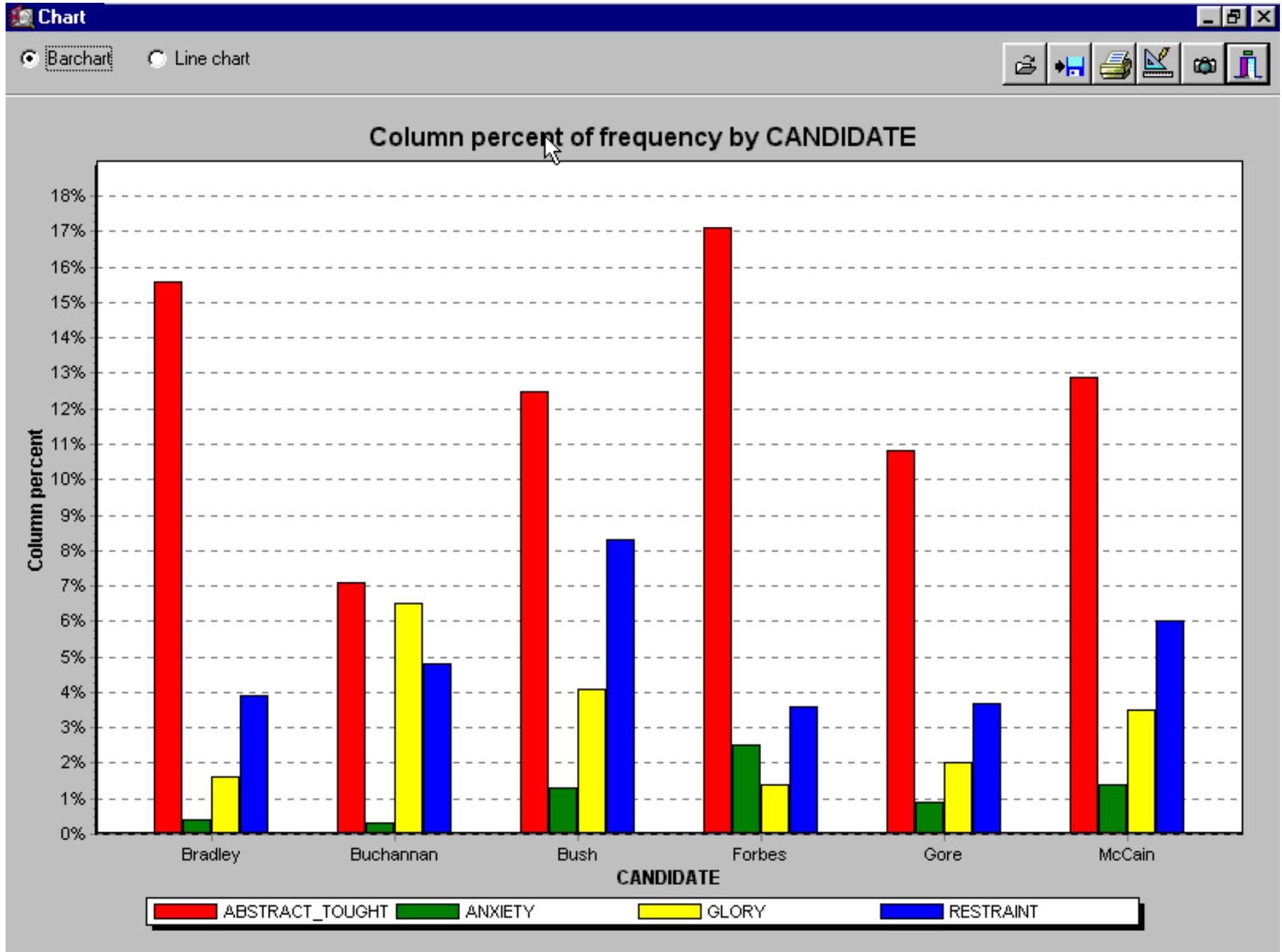
Dictionaries | Options | Word count | **Crosstab** | Key-Word-In-Context

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

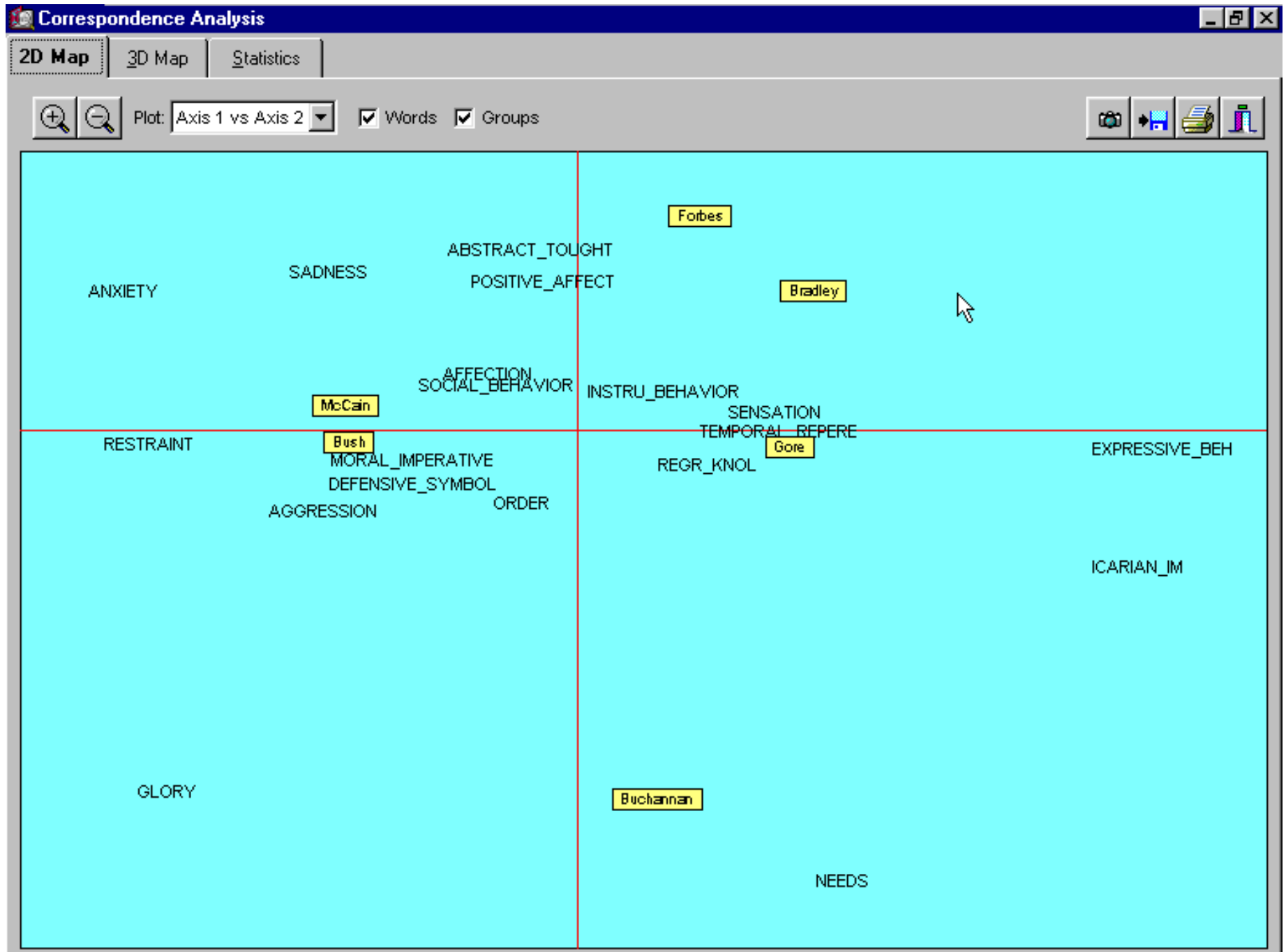
With: CANDIDATE | Sort by: Word | Frequency: Absolute Relative

	Bradley	Buchanan	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

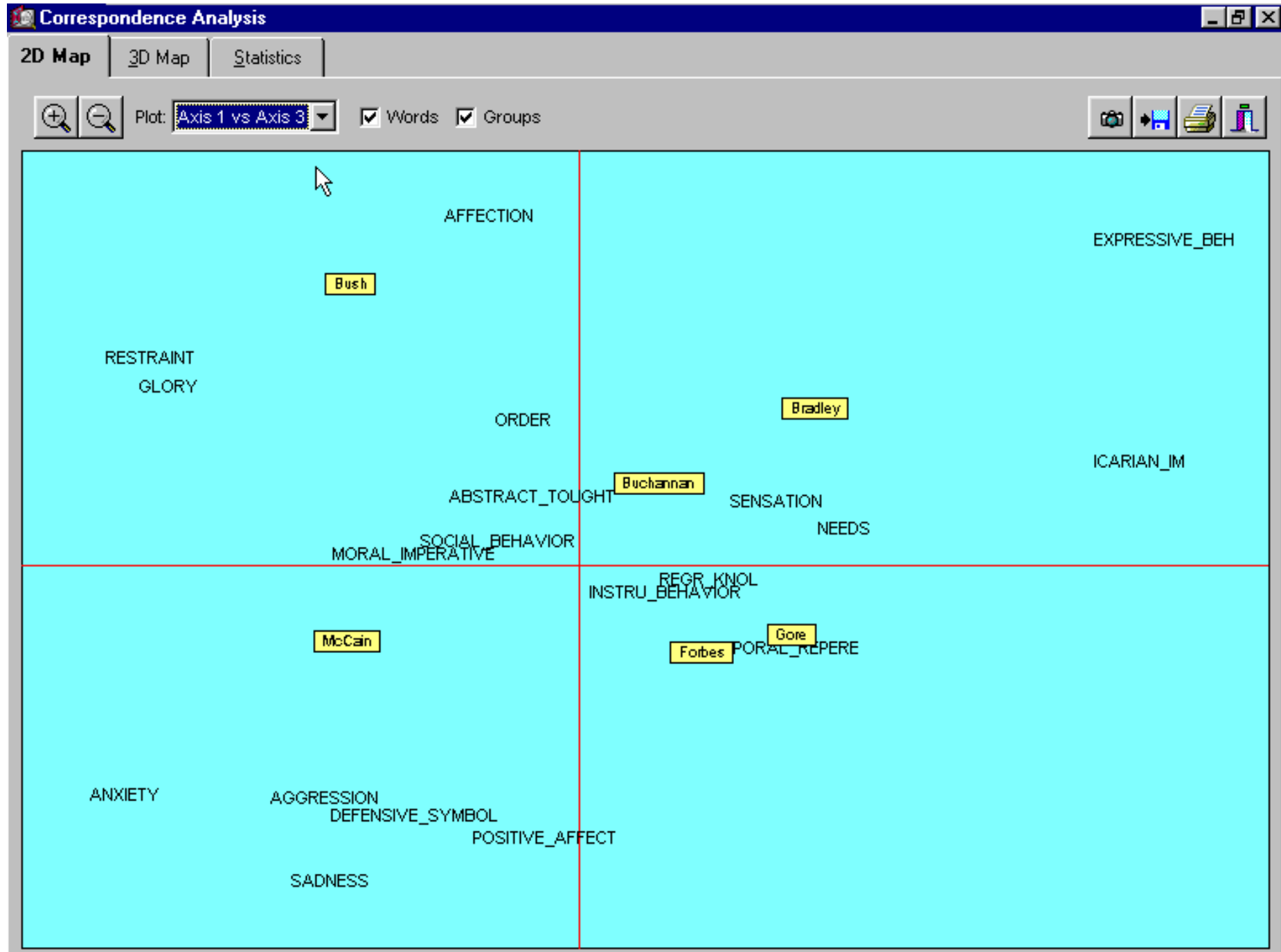
Relative Importance of Four Categories of Martindale RID among Presidential Candidates



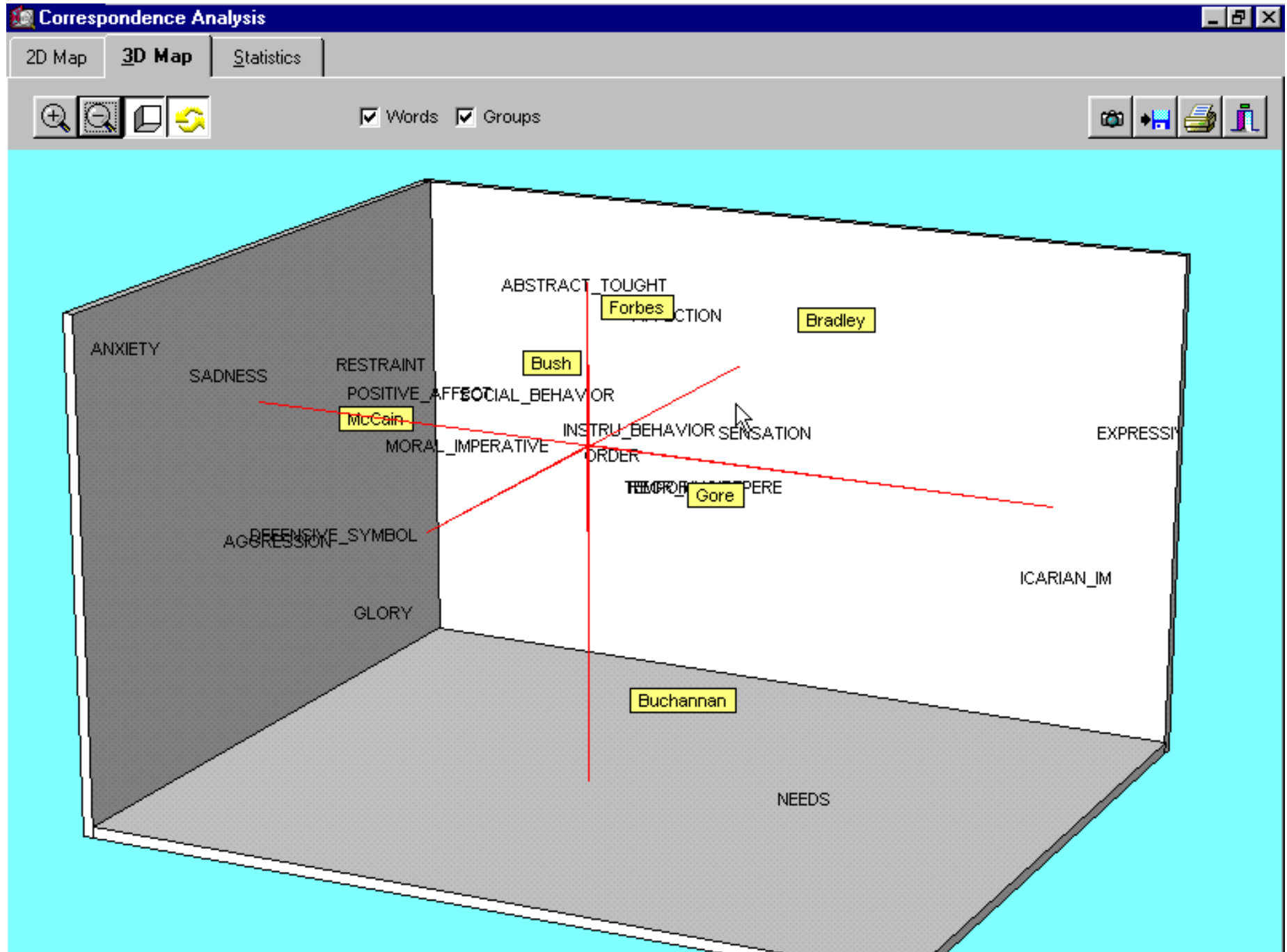
Correspondence Plot of Martindale RID Categories and Presidential Candidates (axis 1 & 2)



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

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

List: Sort by:

Word: Context delimiter:

RECORDNO	KEYWORD	CANDIDATE	TOPIC
23	knew the importance of American leadership in a dangerous world. President Reagan knew that Soviet com	McCain	Announce
24	eriment 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	owed vigilance, in the face of new and continuing dangers . As we move forward, we must honor .. this simpl	Gore	Foreign
25	eeded to combat the most serious and near term dangers . Unless we cogently prioritize threats to our nation:	McCain	Foreign
24	world. Although the next century will hold many dangers for America and our cause, it will, more than ever, '	McCain	Announce
22	ig down the Soviet threat and reducing the mortal dangers of nuclear war. Today, the Soviet Union is gone -	Gore	Foreign
20	is 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	ine 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	eful not to use the relief we feel over our apparent defeat of Mr. Milosevic to indulge an impulse to pronounc	McCain	Foreign
26	ie. 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 constitutor	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	legend scores of countries around the world on a defense budget that has fallen to the smallest share of t	Buchanan	Foreign

Number of items: 1122

31/31 records

Sample Concordance Report of Negative Emotion Words (LIWC)

Report

File Edit Search

File Edit Search

Arial 9

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

[Record #1 CANDIDATE = Bradley TOPIC = Announce]
f restaurants in the boot heel of Missouri because they wouldn't serve our black catcher. Racism **disturbed** me then, and still
angers me now. For me, the only thing deserving of hate is hate itself.

[Record #1 CANDIDATE = Bradley TOPIC = Announce]
puts the people front and center - not the president. A leadership that understands the people's **fears** as well as their hopes. A
leadership that respects the people as well as challenges them. We

[Record #1 CANDIDATE = Bradley TOPIC = Announce]
oday we would call him disabled - he suffered from calcified arthritis of the lower spine. I never **saw** him drive a car, or throw a
ball, or walk farther than a few blocks. My mother dressed him every]

[Record #1 CANDIDATE = Bradley TOPIC = Announce]
ater, when the river flooded the town's main intersection under six feet of muddy water, we also **saw** its destructive power. Crystal
City had only one stoplight then, but it had a rich array of e

[Record #1 CANDIDATE = Bradley TOPIC = Announce]
, its glacial lakes and long Atlantic shore, New Jersey is a microcosm of America. As a Senator, I **saw** my role as both
representing my state and the best interests of our country. I worked hard... f

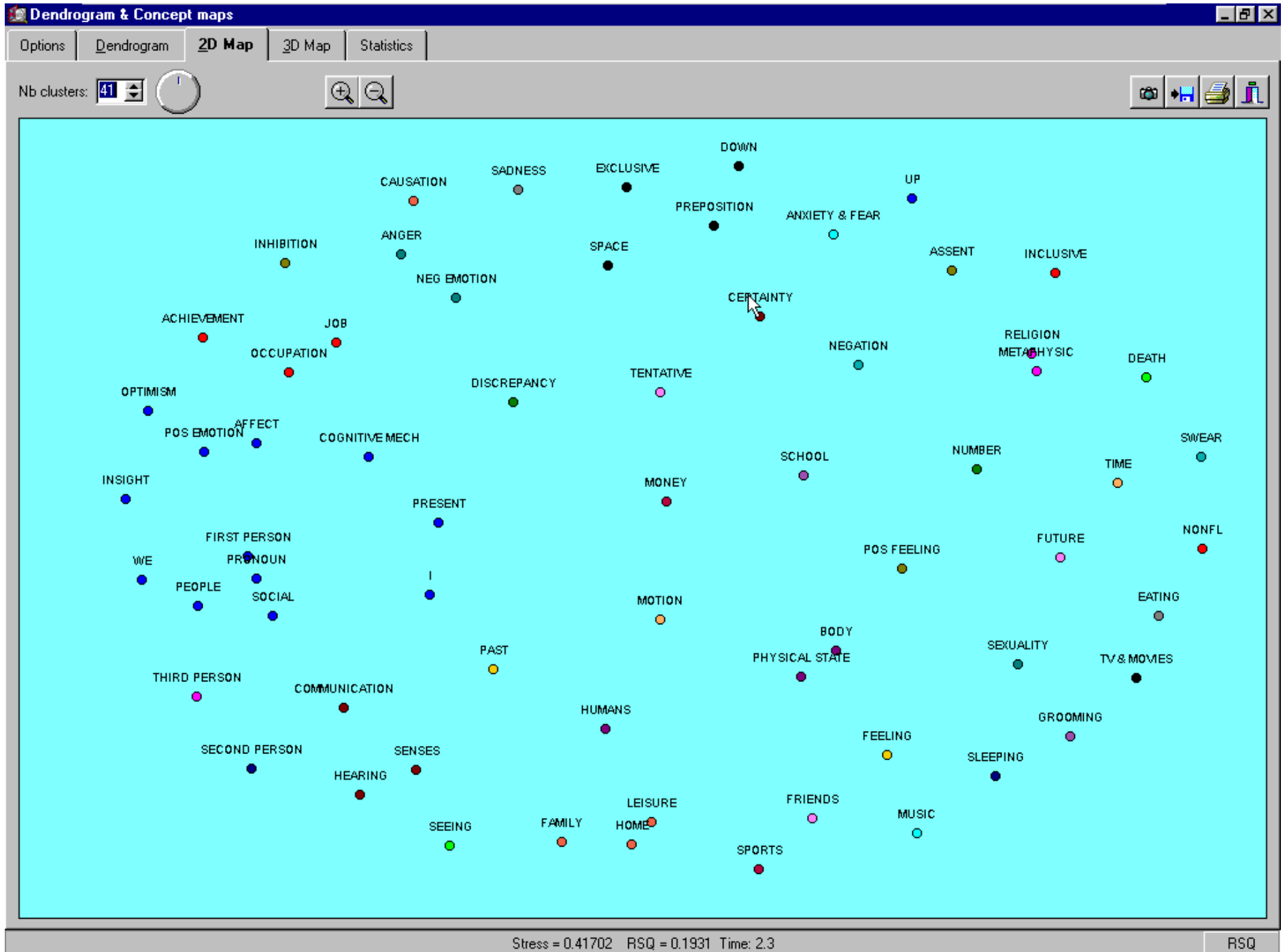
[Record #1 CANDIDATE = Bradley TOPIC = Announce]
at it brings a hot breakfast to children who go to school without it, that it brings security to **worried** seniors - but deeper in the
sense that we have a prosperity that adds up to more than the sum of

[Record #2 CANDIDATE = Bradley TOPIC = Announce]
situation. Today, many hard working parents who hold several jobs in order to make ends meet, **worry** that they don't have
enough time with their children to imbue them with a proper set of values.

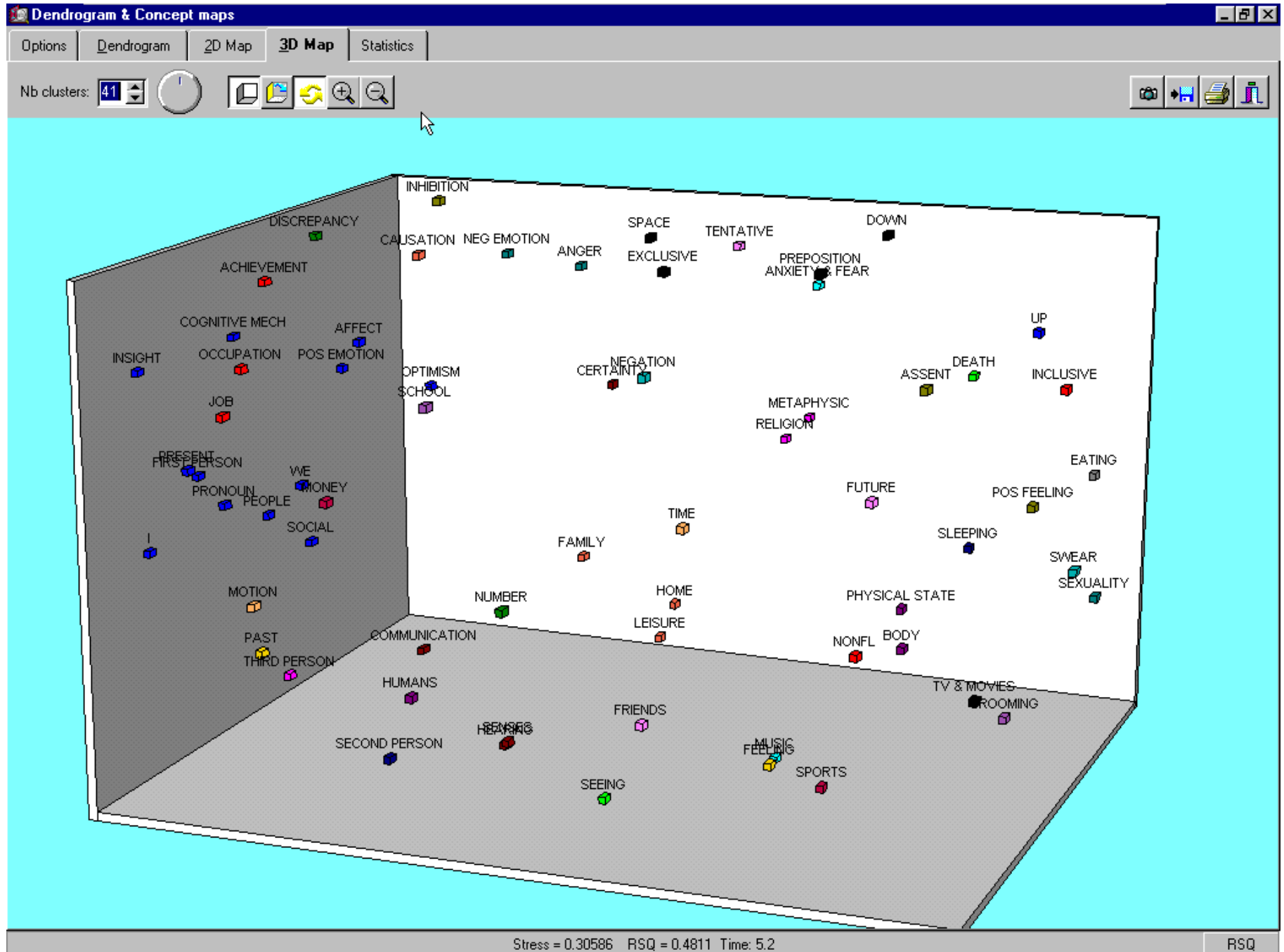
[Record #3 CANDIDATE = Bradley TOPIC = Foreign]
must do well. This new world that's taking shape is not always clear. For example, there's a **disturbing** paradox. We are more
powerful than ever before, yet we are also more vulnerable to a variety of

[Record #3 CANDIDATE = Bradley TOPIC = Foreign]
t dependence on foreign oil, Mexican oil, Venezuelan oil, it was rather dependence on oil from **insecure** sources such as in the
Persian Gulf. And the way to deal with the problem was not to build a massi

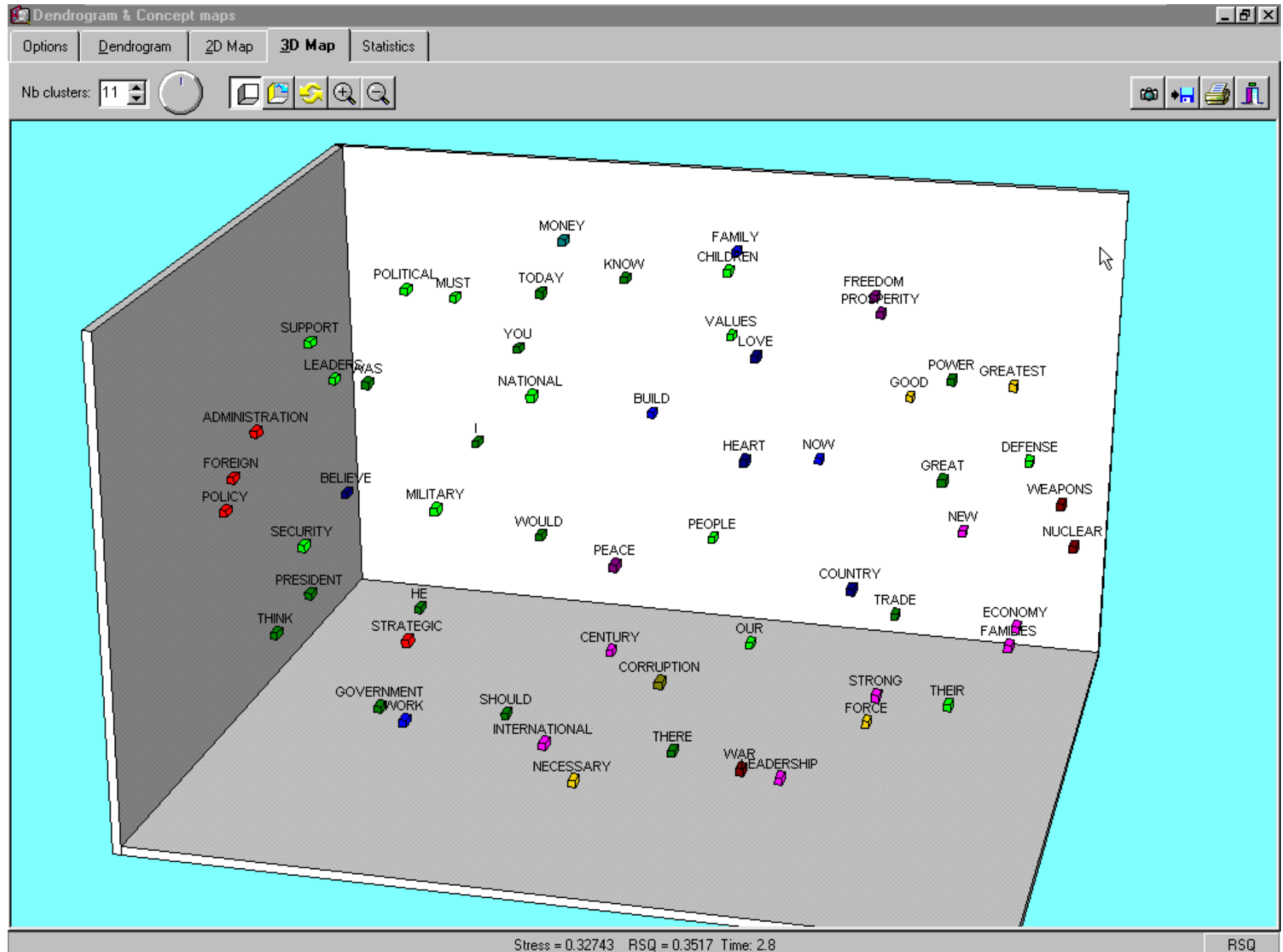
Sample 2D Multidimensional Scaling of LIWC categories



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

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

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

Words to display: Not included | Sort by: Word frequency | List | Tree/Map

	NB WORDS	% SHOWED	% PROCESSED	% TOTAL	NB RECORDS	% RECORD
WORLD	290	1.4%	0.2%	0.4%	29	93.5%
AMERICA	250	1.2%	0.2%	0.3%	27	87.1%
MORE	249	1.2%	0.2%	0.3%	30	96.8%
AMERICAN	227	1.1%	0.2%	0.3%	26	83.9%
SO	171	0.8%	0.1%	0.2%	29	93.5%
STATES	147	0.7%	0.1%	0.2%	29	93.5%
EVERY	135	0.6%	0.1%	0.2%	28	90.3%
THOSE	133	0.6%	0.1%	0.2%	26	83.9%
UNITED	121	0.6%	0.1%	0.2%	29	93.5%
CENTURY	119	0.6%	0.1%	0.2%	24	77.4%
ONLY	115	0.6%	0.1%	0.2%	28	90.3%
COUNTRY	110	0.5%	0.1%	0.2%	26	83.9%
THESE	105	0.5%	0.1%	0.1%	27	87.1%
OTHER	101	0.5%	0.1%	0.1%	24	77.4%
GREAT	97	0.5%	0.1%	0.1%	27	87.1%
MOST	95	0.4%	0.1%	0.1%	25	80.6%
NUCLEAR	95	0.4%	0.1%	0.1%	13	41.9%
CHINA	92	0.4%	0.1%	0.1%	13	41.9%
WAR	92	0.4%	0.1%	0.1%	18	58.1%
NATION	88	0.4%	0.1%	0.1%	26	83.9%
GOVERNMENT	84	0.4%	0.1%	0.1%	15	48.4%
OWN	84	0.4%	0.1%	0.1%	23	74.2%
POLICY	84	0.4%	0.1%	0.1%	22	71.0%
MANY	82	0.4%	0.1%	0.1%	26	83.9%
WHICH	81	0.4%	0.1%	0.1%	28	90.3%
LIFE	75	0.4%	0.1%	0.1%	24	77.4%
FOREIGN	72	0.3%	0.1%	0.1%	16	51.6%
RIGHT	71	0.3%	0.1%	0.1%	26	83.9%
AMERICA'S	70	0.3%	0.1%	0.1%	22	71.0%
AMERICANS	69	0.3%	0.1%	0.1%	16	51.6%
CORRUPTION	67	0.3%	0.1%	0.1%	3	9.7%
BACK	66	0.3%	0.1%	0.1%	10	31.2%

To inclusion list...
 To exclusion list
 Remove from included
 Key-Word-In-Context

Showed: 4768 Unique: 4835/72101 Time: 49s | Filter | Edit | OK | Recount

31/31 records | Print | Save | Cancel | Help

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.

{Foreign Policy Address by Senator John McCain The Landon Lecture}

No decade in this century began more auspiciously than the 1990s. That gross impediment to liberty, the Berlin Wall, was breached by the stronger forces of human yearning. The central security problem of our time - the possible clash of East and West on the plains of Germany - was resolved by the dissolution of the Warsaw Pact and the reunification of Germany. The Soviet Union imploded and ceased its militant hostility to the values of liberal democracies.

The euphoria that accompanied these remarkable events anticipated the arrival of a world of independent democracies engaged only in peaceful commercial competition with one another. But the resurrection of ancient conflicts and hideous barbarism in the Balkans; the re-appearance of other incidents of irrational nationalism and tribalism that had been sublimated by the Cold War; the near total collapse of the Russian economy and its implications for economic and political stability; the accelerated proliferation of weapons of mass destruction, and the progress to date in developing weapons made by rogue states with an implacable hostility for American interests; the dozens of wars that were waged across the globe, dimmed humanity's highest hopes.

Still, our disappointments should not obscure our progress. They should caution us, but they should not intimidate us from making the most of this moment, to continue building a safer, more humane world than the one that inaugurated this century.

I am often identified as a critic of administration policies. But I have just as publicly agreed with Administration policies when I have found them to have merit as I have criticized them when they have been feckless. I strongly supported the expansion of NATO, as well as the two great trade successes, NAFTA and the Uruguay Round. But to understand the mistakes of Clinton statecraft, we need to discuss them plainly.

At the core of the problem are two pronounced flaws: strategic incoherence and self-doubt. The first refers to the lack of a conceptual framework -- in other words, what we want the world to be and how we can help make it so. A conceptual framework establishes the relationships among our many interests in the world, provides the basis for prioritizing those interests, and obliges us to integrate policies to serve those priorities.

Early on, the Administration had a flight of fancy that posed as the means employed to support a conceptual framework - "assertive multilateralism." But it fell apart in Somalia and Bosnia, and assertive gave way to passive multilateralism, where we act in concert with other nations when they determine it necessary to safeguard their security. In Bosnia, and probably soon in Kosovo, American troops are deployed not only, but primarily, to ensure stability in Europe. Yet, only two of our NATO allies help us in any meaningful way to contain the threat from Iraq.

The second fault I find with the administration, its self-doubt, is obviously related and a primary cause of its strategic incoherence. Often evident in Administration policies is a justifying uncertainty.

31 / 31

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.

CANDIDATE: McCain

{Foreign Policy Address by Senator John McCain The Landon Lecture}

No decade in this century **began** more auspiciously than the 1990s. That **gross** impediment to liberty, the Berlin Wall, was breached by the **stronger** forces of **human** yearning. The central **security** problem of **our** time - the **possible** clash of East and West on the plains of Germany - was **resolved** by the dissolution of the Warsaw Pact and the reunification of Germany. The Soviet Union imploded and ceased its militant **hostility** to the **values** of liberal democracies.

The euphoria that accompanied these remarkable events anticipated the independent democracies engaged only in **peaceful** commercial competition; the resurrection of ancient **conflicts** and hideous barbarism in the Balkans; the incidents of **irrational** nationalism and tribalism that had been sublimated; the total collapse of the Russian economy and its **implications** for economic growth; the accelerated proliferation of weapons of mass **destruction**, and the proliferation of weapons made by rogue states with an implacable **hostility** for American interests; the dozens of wars that were waged across the globe, dimmed humanity's hopes.

Still, **our** **disappointments** **should** not obscure **our** progress. **They** **should** **but** **they** **should** not **intimidate** **us** from making the most of this moment, a more humane world than the one that inaugurated this century.

I am often identified as a critic of administration policies. **But** I have **just** a few words to say about Administration policies when I have **found** **them** to have merit as I have **found** **them** to have been feckless. I **strongly** **supported** the expansion of NATO, as well as its successes, NAFTA and the Uruguay Round. **But** to **understand** the mistakes, we **need** to **discuss** **them** plainly.

At the core of the problem are two pronounced flaws: strategic incoherence and lack of a conceptual framework -- in other words, **what** **we** **want** **we** can **help** make it so. A conceptual framework establishes the **relationship** between **our** **interests** in the world, provides the **basis** for prioritizing those **interests**, and provides policies to serve those priorities.

Early on, the Administration had a flight of fancy that posed as the **means** of a conceptual framework - "assertive multilateralism." **But** it fell apart in Somalia. **Assertive** **gave** way to passive multilateralism, where **we** act in concert with others to **determine** it necessary to safeguard **their** **security**. In Bosnia, and **probably** soon in Kosovo, American troops are deployed not only, **but** primarily, to ensure stability in Europe. Yet, only two of **our** NATO allies **help** **us** in **any** **meaningful** way to **contain** the threat from Iraq.

The second fault I **find** with the administration, its self-**doubt**, is obviously **related** and a primary **cause** of its strategic incoherence. Often evident in Administration policies is a mystifying uncertainty about **how** to act in a world where **we** are the only superpower. When the administration stands mute and **undecided** about where and **how** **they** **want** to lead the world **they** exhibit, to **friend** and foe alike, an identity crisis, an image of America an existential crisis: **who** are **we** and **why** are **we** here?

Color codes

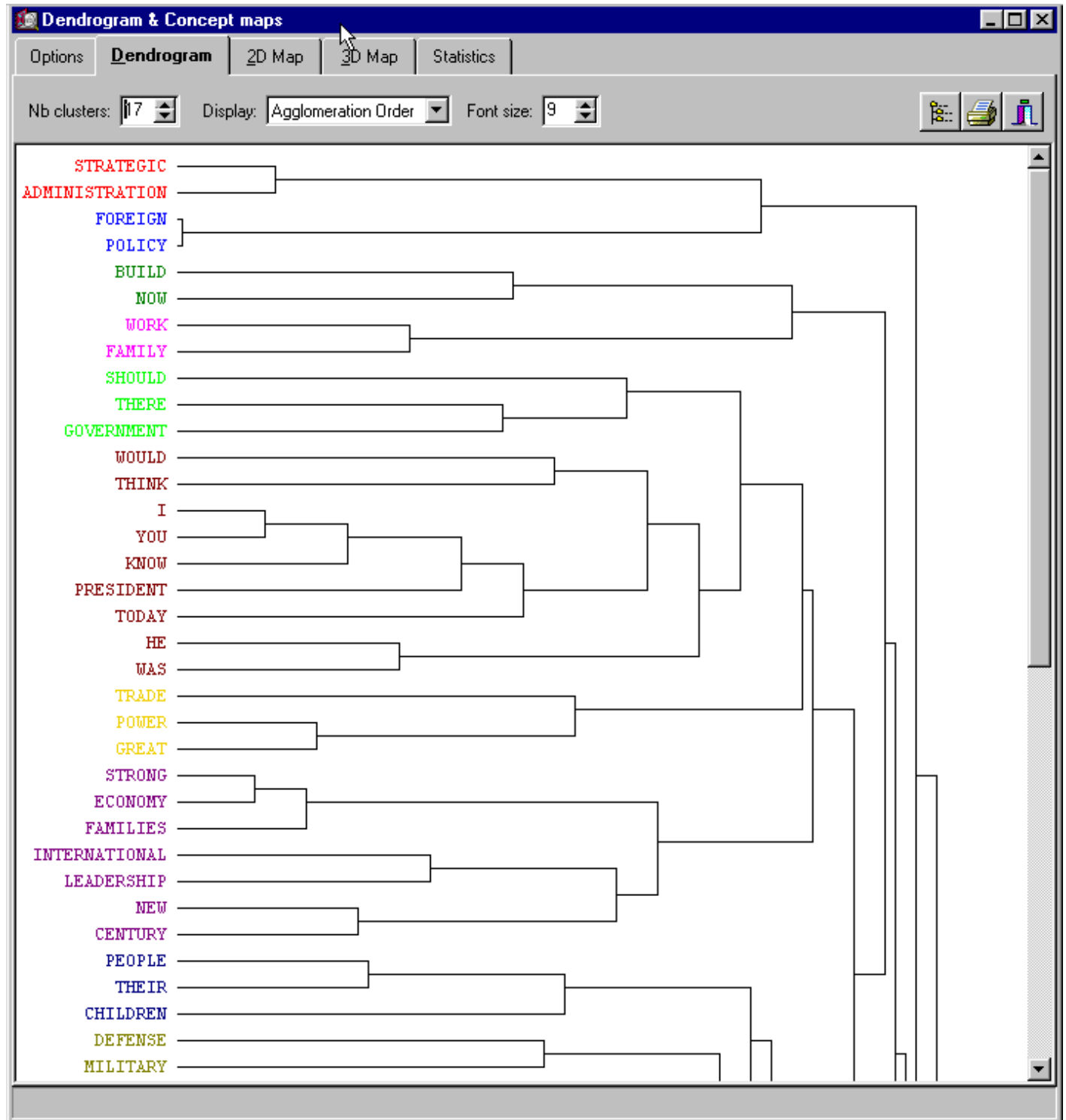
Font: [Red] Background: [Yellow]

- AFFECTIVE PROCESS
- ANGER
- ANXIETY
- CAUSATION
- CERTAINTY
- COGNITIVE_PROCESS
- COMMUNICATION
- DISCREPANCY
- FAMILY
- FEELING
- FRIENDS
- HEARING
- HUMANS
- INHIBITION
- INSIGHT
- NEGATIVE_EMOTION
- OPTIMISM

Clear Cancel OK

Show all record 31 / 31 Close

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.

