

Coming to Australia Personality in Language

Scott Nowson

Macquarie University

September 20, 2006

From A...



...to B



2001-2006

- **The Language of Weblogs:**
a study of genre and individual differences
 - Investigated how personality (and gender) are reflected in language of CMC
 - Specifically, in the language of personal online diaries, or 'blogs'

2001-2006

- The Language of Weblogs:
a study of genre and individual differences
 - Investigated how personality (and gender) are reflected in language of CMC
 - Specifically, in the language of personal online diaries, or 'blogs'

2001-2006

- The Language of Weblogs:
a study of genre and individual differences
 - Investigated how personality (and gender) are reflected in language of CMC
 - Specifically, in the language of personal online diaries, or 'blogs'

A blog by any other name...

- What is a blog?
 - Website with entries in reverse chronological order
 - Can be news reporting, opinion giving or life story telling
- Interestingly...
 - Majority of personal webloggers are younger females
 - Extraversion no leaning either way
 - Bloggers *appear* to be highly Open

A blog by any other name...

- What is a blog?
 - Website with entries in reverse chronological order
 - Can be news reporting, opinion giving or life story telling
- Interestingly...
 - Majority of personal webloggers are younger females
 - Extraversion no leaning either way
 - Bloggers *appear* to be highly Open

A blog by any other name...

- What is a blog?
 - Website with entries in reverse chronological order
 - Can be news reporting, opinion giving or life story telling
- Interestingly...
 - Majority of personal webloggers are younger females
 - Extraversion no leaning either way
 - Bloggers *appear* to be highly Open

A blog by any other name...

- What is a blog?
 - Website with entries in reverse chronological order
 - Can be news reporting, opinion giving or life story telling
- Interestingly...
 - Majority of personal webloggers are younger females
 - Extraversion no leaning either way
 - Bloggers *appear* to be highly Open

A blog by any other name...

- What is a blog?
 - Website with entries in reverse chronological order
 - Can be news reporting, opinion giving or life story telling
- Interestingly...
 - Majority of personal webloggers are younger females
 - Extraversion *no leaning either way*
 - Bloggers *appear* to be highly Open

A blog by any other name...

- What is a blog?
 - Website with entries in reverse chronological order
 - Can be news reporting, opinion giving or life story telling
- Interestingly...
 - Majority of personal webloggers are younger females
 - Extraversion no leaning either way
 - Bloggers *appear* to be highly Open

A blog by any other name...

- What is a blog?
 - Website with entries in reverse chronological order
 - Can be news reporting, opinion giving or life story telling
- Interestingly...
 - Majority of personal webloggers are younger females
 - Extraversion no leaning either way
 - Bloggers *appear* to be highly Open

A wee teaser

- Neuroticism
 - Low - many references to other people: [*NP1* <*p*>], [*that he*]
 - Similar use of references to self: [*if i*], [*that i*]
- Extraversion
 - Extraverts references to other people, Introverts more to self
 - Extraverts use more social words, talk less of their achievements
 - As found in e-mail corpus, Extraverts are [*cool* <*p*>]
- Openness to Experience
 - High scorers use more emotionally expressive language
 - They use longer words and also talk less about their achievements

A wee teaser

- Neuroticism
 - Low - many references to other people: [*NP1 <p>*], [*that he*]
 - Similar use of references to self: [*if i*], [*that i*]
- Extraversion
 - Extraverts references to other people, Introverts more to self
 - Extraverts use more social words, talk less of their achievements
 - As found in e-mail corpus, Extraverts are [*cool <p>*]
- Openness to Experience
 - High scorers use more emotionally expressive language
 - They use longer words and also talk less about their achievements

A wee teaser

- Neuroticism
 - Low - many references to other people: [*NP1 <p>*], [*that he*]
 - Similar use of references to self: [*if i*], [*that i*]
- Extraversion
 - Extraverts references to other people, Introverts more to self
 - Extraverts use more social words, talk less of their achievements
 - As found in e-mail corpus, Extraverts are [*cool <p>*]
- Openness to Experience
 - High scorers use more emotionally expressive language
 - They use longer words and also talk less about their achievements

Tease me more

- Agreeableness
 - High scorers swear less, talk less of negative emotions
 - Agreeableness correlates negatively with contextuality
 - Low - collocations contain contractions: [*they don't*], [*if i'm*]
 - High - contains non-contracted phrases: [*i am*], [*is not*]
- Conscientiousness
 - High - scorers end with punctuation: [*<p> <eop>*]
 - Low - may change topic more: [*anyway <p>*], [*actually <p>*]

Tease me more

- Agreeableness
 - High scorers swear less, talk less of negative emotions
 - Agreeableness correlates negatively with contextuality
 - Low - collocations contain contractions: [*they don't*], [*if i'm*]
 - High - contains non-contracted phrases: [*i am*], [*is not*]
- Conscientiousness
 - High - scorers end with punctuation: [*<p> <eop>*]
 - Low - may change topic more: [*anyway <p>*], [*actually <p>*]

The yellow brick road

- Computing @ Macquarie
- Text Summarisation

What can I do for you

... it is clear that human processors of information make use of a much deeper understanding of text [...] how we can improve search engine technology by integrating more knowledge about the processing and understanding of language [...] Are there insights from the cognitive sciences that can tell us how to build better tools for finding information?

- Humans understand personality from language
- Measures based on peer evaluation
- What if machines could understand personality

What can I do for you

... it is clear that human processors of information make use of a much deeper understanding of text [...] how we can improve search engine technology by integrating more knowledge about the processing and understanding of language [...] Are there insights from the cognitive sciences that can tell us how to build better tools for finding information?

- Humans understand personality from language
- Measures based on peer evaluation
- What if machines could understand personality

What can I do for you

... it is clear that human processors of information make use of a much deeper understanding of text [...] how we can improve search engine technology by integrating more knowledge about the processing and understanding of language [...] Are there insights from the cognitive sciences that can tell us how to build better tools for finding information?

- Humans understand personality from language
- Measures based on peer evaluation
- What if machines could understand personality

What can I do for you

... it is clear that human processors of information make use of a much deeper understanding of text [...] how we can improve search engine technology by integrating more knowledge about the processing and understanding of language [...] Are there insights from the cognitive sciences that can tell us how to build better tools for finding information?

- Humans understand personality from language
- Measures based on peer evaluation
- What if machines could understand personality

And I would walk 10,000 miles

- What can I do in Sydney?
 - Google gives me the official line
 - But what will I enjoy?
- Canvas by author personality
- Want opinions and recommendations of similar people

And I would walk 10,000 miles

- What can I do in Sydney?
 - Google gives me the official line
 - But what will I enjoy?
- Canvas by author personality
- Want opinions and recommendations of similar people

And I would walk 10,000 miles

- What can I do in Sydney?
 - Google gives me the official line
 - But what will I enjoy?
- Canvas by author personality
- Want opinions and recommendations of similar people

Who are you? hoo-hoo hoo-hoo

- First stage is to automate the detection of personality
- Still in the early stages
 - Stratification approach
 - Feature selection
- Results on a limited corpus very promising
- However, possibility of over-fitting
- Working on much larger corpus

Who are you? hoo-hoo hoo-hoo

- First stage is to automate the detection of personality
- Still in the early stages
 - Stratification approach
 - Feature selection
- Results on a limited corpus very promising
- However, possibility of over-fitting
- Working on much larger corpus

Who are you? hoo-hoo hoo-hoo

- First stage is to automate the detection of personality
- Still in the early stages
 - Stratification approach
 - Feature selection
- Results on a limited corpus very promising
- However, possibility of over-fitting
- Working on much larger corpus

The end

- There are many directions in which search can be extended
- Searching by author is one very interesting direction to take
- because Mr Shy doesn't want to know where the liveliest bars are and Mr Sporty doesn't care for chess clubs

The end

- There are many directions in which search can be extended
- Searching by author is one very interesting direction to take
- because Mr Shy doesn't want to know where the liveliest bars are and Mr Sporty doesn't care for chess clubs

The end

- There are many directions in which search can be extended
- Searching by author is one very interesting direction to take
- because Mr Shy doesn't want to know where the liveliest bars are and Mr Sporty doesn't care for chess clubs