Mark my words!
Linguistic style coordination in social media

Cristian Danescu-Niculescu-Mizil\textsuperscript{1}
Michael Gamon\textsuperscript{2}
Susan Dumais\textsuperscript{2}

\textsuperscript{1}Cornell University
\textsuperscript{2}Microsoft Research

01-04-2011 WWW 2011
Non-conscious coordination

When conversing, people non-consciously adapt to one another’s communicative behaviors. [Giles et al., 1991; Chartrand and Bargh, 1999]
When conversing, people non-consciously adapt to one another’s communicative behaviors. [Giles et al., 1991; Chartrand and Bargh, 1999]

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Canonical study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posture</td>
<td>Condon and Ogston, 1967</td>
</tr>
</tbody>
</table>
Non-conscious coordination

When conversing, people non-consciously adapt to one another’s communicative behaviors. [Giles et al., 1991; Chartrand and Bargh, 1999]

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Canonical study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posture</td>
<td>Condon and Ogston, 1967</td>
</tr>
<tr>
<td>Head nodding</td>
<td>Hale and Burgoon, 1984</td>
</tr>
</tbody>
</table>
When conversing, people non-consciously adapt to one another’s communicative behaviors. [Giles et al., 1991; Chartrand and Bargh, 1999]

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Canonical study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posture</td>
<td>Condon and Ogston, 1967</td>
</tr>
<tr>
<td>Head nodding</td>
<td>Hale and Burgoon, 1984</td>
</tr>
<tr>
<td>Pause length</td>
<td>Jaffe and Feldstein, 1970</td>
</tr>
<tr>
<td>Backchannels</td>
<td>White, 1984</td>
</tr>
<tr>
<td>Self-disclosure</td>
<td>Derlenga et al., 1973</td>
</tr>
<tr>
<td>Linguistic style</td>
<td>Niederhoffer and Pennebaker, 2002</td>
</tr>
</tbody>
</table>
Non-conscious coordination

When conversing, people non-consciously adapt to one another’s communicative behaviors. [Giles et al., 1991; Chartrand and Bargh, 1999]

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Canonical study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posture</td>
<td>Condon and Ogston, 1967</td>
</tr>
<tr>
<td>Head nodding</td>
<td>Hale and Burgoon, 1984</td>
</tr>
<tr>
<td>Pause length</td>
<td>Jaffe and Feldstein, 1970</td>
</tr>
<tr>
<td>Backchannels</td>
<td>White, 1984</td>
</tr>
<tr>
<td>Self-disclosure</td>
<td>Derlenga et al., 1973</td>
</tr>
<tr>
<td>Linguistic style</td>
<td>Niederhoffer and Pennebaker, 2002</td>
</tr>
</tbody>
</table>

Communicative behaviors are “patterned and coordinated, like a dance” [Niederhoffer and Pennebaker, 2002]
Non-conscious coordination

When conversing, people **non-consciously** adapt to one another’s communicative behaviors. [Giles et al., 1991; Chartrand and Bargh, 1999]

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Canonical study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posture</td>
<td>Condon and Ogston, 1967</td>
</tr>
<tr>
<td>Head nodding</td>
<td>Hale and Burgoon, 1984</td>
</tr>
<tr>
<td>Pause length</td>
<td>Jaffe and Feldstein, 1970</td>
</tr>
<tr>
<td>Backchannels</td>
<td>White, 1984</td>
</tr>
<tr>
<td>Self-disclosure</td>
<td>Derlenaga et al., 1973</td>
</tr>
<tr>
<td>Linguistic style</td>
<td>Niederhoffer and Pennebaker, 2002</td>
</tr>
</tbody>
</table>

Communicative behaviors are **“patterned and coordinated, like a dance”**  
[Niederhoffer and Pennebaker, 2002]
Non-conscious coordination

When conversing, people **non-consciously** adapt to one another’s communicative behaviors. [Giles et al., 1991; Chartrand and Bargh, 1999]

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Canonical study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posture</td>
<td>Condon and Ogston, 1967</td>
</tr>
<tr>
<td>Head nodding</td>
<td>Hale and Burgoon, 1984</td>
</tr>
<tr>
<td>Pause length</td>
<td>Jaffe and Feldstein, 1970</td>
</tr>
<tr>
<td>Backchannels</td>
<td>White, 1984</td>
</tr>
<tr>
<td>Self-disclosure</td>
<td>Derlenga et al., 1973</td>
</tr>
<tr>
<td>Linguistic style</td>
<td>Niederhoffer and Pennebaker, 2002</td>
</tr>
</tbody>
</table>

**Small scale, laboratory controlled studies.**
Linguistic style:

**How** things are said as opposed to **what** is said
Linguistic style:

**How** things are said as opposed to **what** is said

Example:

Q : “At what time does your shop close?”

A1: “At five o’clock.”
Linguistic style coordination

Linguistic style:

How things are said as opposed to what is said

Example:

Q : “At what time does your shop close?”

A1: “At five o’clock.”
A2: “Five o’clock”
Linguistic style coordination

**Linguistic style:**

*How* things are said as opposed to *what* is said

**Example:**

Q1: “At what time does your shop close?”
Q2: “What time does your shop close?”

A1: “At five o’clock.”
A2: “Five o’clock”

[Levelt & Kelter, 1982]
Linguistic style:

How things are said as opposed to what is said

Example:

Q1: “At what time does your shop close?”
Q2: “What time does your shop close?”

A1: “At five o’clock.”
A2: “Five o’clock”

[Levelt & Kelter, 1982]
Linguistic style:

*How* things are said as opposed to *what* is said

Example:

Q1: “At what time does your shop close?”
Q2: “What time does your shop close?”

A1: “At five o’clock.”
A2: “Five o’clock”

[Levelt & Kelter, 1982]
Linguistic style coordination

Linguistic coordination occurs:
- instantaneously
- non-consciously
Talk outline

Psycho-linguistics

Linguistic style coordination
Socio-linguistic problems:

- Relationship nature and communication patterns – Voluntary vs. involuntary relations (e.g., dating partners vs. in-laws) [Mongeau & Henningsten, 2008]
- Communication accommodation [Giles et al., 1991]

Large-scale socially-rich conversational data:

- Facebook, Twitter, Second Life, reality shows, etc.

Talk outline

Psycho-linguistics

Linguistic style coordination

Back to psycho-linguistics
Socio-linguistic problems:

- Relationship nature and communication patterns – Voluntary vs. involuntary relations (e.g., dating partners vs. in-laws) [Mongeau & Henningsten, 2008]
- Communication accommodation [Giles et al., 1991]

Large-scale socially-rich conversational data:
- Facebook, Twitter, Second Life, reality shows, etc.

Psycho-linguistics

Linguistic style coordination

Back to psycho-linguistics

Practical applications
Socio-linguistic problems:

- Relationship nature and communication patterns – Voluntary vs. involuntary relations (e.g., dating partners vs. in-laws) [Mongeau & Henningsten, 2008]
- Communication accommodation [Giles et al., 1991]

Large-scale socially-rich conversational data:

- Facebook, Twitter, Second life, reality shows, etc.

Linguistic style coordination

Mathematical formalism

Psycholinguistics

Practical applications

Back to psycholinguistics

Talk outline
Socio-linguistic problems:
• Relationship nature and communication patterns – Voluntary vs. involuntary relations (e.g., dating partners vs. in-laws) [Mongeau & Henningsten, 2008]
• Communication accommodation [Giles et al., 1991]

Large-scale socially-rich conversational data:
• Facebook, Twitter, Second Life, reality shows etc.

Linguistic style coordination

Mathematical formalism

Psycho-linguistics

Large scale, real life setting

Practical applications

Back to psycho-linguistics

Talk outline
Socio-linguistic problems:


• Relationship nature and communication patterns – Voluntary vs. involuntary relations (e.g., dating partners vs. in-laws) [Mongeau & Henningsten, 2008]

• Communication accommodation [Giles et al., 1991]

Large-scale socially-rich conversational data:

• Facebook, Twitter, Second life, reality shows etc.

Linguistic style coordination

Mathematical formalism

Psycho-linguistics

Large scale, real life setting

Back to psycho-linguistics

Practical applications

Talk outline
Practical applications

Why bother?

Premise: Coordination has a positive effect on communication

Previously explored applications:
- Hostage negotiations [Taylor & Thomas, 2008]
- Treatment of mental disability [Hamilton, 1991]
- Psychotherapy [Ferrara, 1991]
Practical applications

Why bother?

Premise: Coordination has a positive effect on communication

Previously explored applications:
- Hostage negotiations [Taylor & Thomas, 2008]
- Treatment of mental disability [Hamilton, 1991]
- Psychotherapy [Ferrara, 1991]
- Bigger tips! [Van Baaren et al, 2003]
  - when waitresses coordinate with the customers they get bigger tips
Practical applications

Why bother?

Premise: Coordination has a positive effect on communication

[Giles et al, 2006]

By adding:

- Robustness (beyond small-scale, real-time, face-to-face etc.)
- Framework for working with coordination at a large-scale

we could use coordination for:

- Automated dialogue systems
- Detection of forged or unnatural conversations
Socio-linguistic problems:

- Relationship nature and communication patterns – Voluntary vs. involuntary relations (e.g., dating partners vs. in-laws) [Mongeau & Henningsten, 2008]

Large-scale socially-rich conversational data:
- Facebook, Twitter, Second Life, reality shows, etc.

Linguistic style coordination

Mathematical formalism → Psycho-linguistics → Large scale, real life setting → Practical applications → Back to psycho-linguistics

Back to psycho-linguistics

Danescu-Niculescu-Mizil, Gamon, Dumais

Mark my words! 25
Coordination on Twitter

**Twitter conversations: a new opportunity**

- Large scale and naturally occurring
  - ~37% of all tweets are conversational [Kelly, 2009]
Coordination on Twitter

Twitter conversations: a new opportunity / challenge

• Large scale and naturally occurring
  ▪ ~37% of all tweets are conversational [Kelly, 2009]

• Unlike settings where coordination was observed:
  ▪ not real-time
  ▪ not face-to-face
  ▪ 140 character restriction
  ▪ wide spectrum of relation development stages
    (vs. inception-stage only in laboratory studies)
Coordination on Twitter

**Twitter conversations: a new opportunity / challenge**

- Large scale and naturally occurring
  - ~37% of all tweets are conversational [Kelly, 2009]
- Unlike settings where coordination was observed:
  - not real-time
  - not face-to-face
  - 140 character restriction
  - wide spectrum of relation development stages
    (vs. inception-stage only in laboratory studies)

*Is the phenomenon robust enough to emerge under these constraints?*
Socio-linguistic problems:

- Relationship nature and communication patterns – Voluntary vs. involuntary relations (e.g., dating partners vs. in-laws) [Mongeau & Henningsten, 2008]
- Communication accommodation [Giles et al., 1991]

Large-scale socially-rich conversational data:

- Facebook, Twitter, Second Life, reality shows, etc.

Linguistic style coordination:

- Mathematical formalism
- Large scale, real life setting
- Practical applications

Back to psycho-linguistics
Measuring linguistic style

Pleased that the Senate has confirmed Elena Kagan as our 112th Supreme Ct. justice. I am confident she will make an outstanding justice.

about 22 hours ago via HootSuite

Barack Obama

Barack Obama

congrats 2 landon donovan 4 scorin game winning goal. grt job but u ain't scoring on me in shaq vs! good luck nxt round go usa

10:16 AM Jun 23rd via Twitter for iPhone

THE_REAL_SHAQ
Measuring linguistic style

A methodology borrowed from psycho-linguistics (LIWC):

[Pennebaker et al., 2001]

<table>
<thead>
<tr>
<th>Feature families</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prepositions</td>
<td>at, to, with</td>
</tr>
<tr>
<td>Articles</td>
<td>the, an, a</td>
</tr>
<tr>
<td>Auxiliary verbs</td>
<td>maybe, perhaps</td>
</tr>
<tr>
<td>Conjunctions</td>
<td>and, whereas</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

9 feature families ~450 lexemes

Functional words, deemed by [Ireland et al., 2010] to be:
- Unrelated to topic
- Generated and processed non-consciously

(all results hold for all the 50 feature families we studied)
Pleased that the Senate has confirmed Elena Kagan as our 112th Supreme Ct. justice. I am confident she will make an outstanding justice.

about 22 hours ago via HootSuite

BarackObama
Barack Obama
Measuring linguistic style

Articles

Pleased that the Senate has confirmed Elena Kagan as our 112th Supreme Ct. justice. I am confident she will make an outstanding justice.

about 22 hours ago via HootSuite

BarackObama
Barack Obama
Measuring linguistic style

Pleased that the Senate has confirmed Elena Kagan as our 112th Supreme Ct. justice. I am confident she will make an outstanding justice.

about 22 hours ago via HootSuite

BarackObama
Barack Obama
Measuring linguistic style

Articles
Auxiliary verbs
Prepositions
Personal pronouns

Barack Obama
Barack Obama

about 22 hours ago via HootSuite
Measuring linguistic style

Articles
Auxiliary verbs
Prepositions
Personal pronouns

congrats 2 landon donovan 4 scorin game winning goal. grt job but u ain't scoring on me in shaq vs! good luck nxt round go usa
10:18 AM Jun 23rd via Twitter for iPhone

Barack Obama
Barack Obama

THE_REAL_SHAQ
Measuring linguistic style

Articles
Auxiliary verbs
Prepositions
Personal pronouns

congrats 2 landon donovan 4 scorin game winning goal. grt job but u ain't scoring on me in shaq vs! good luck nxt round go usa

about 22 hours ago via HootSuite
BarackObama
Barack Obama

10:18 AM Jun 23rd via Twitter for iPhone
THE_REAL_SHAQ

Danescu-Niculescu-Mizil, Gamon, Dumais

Mark my words!
Measuring linguistic style

Articles
Auxiliary verbs
Prepositions
Personal pronouns

Barack Obama
Barack Obama

about 22 hours ago via HootSuite

THE_REAL_SHAQ

10:18 AM Jun 23rd via Twitter for iPhone
Measuring coordination

What we want: how much A’s inclusion of an article immediately triggers the usage of articles in B’s reply?
Measuring coordination

**What we want:** how much A’s inclusion of an article immediately triggers the usage of articles in B’s reply?

**What we don’t want:** how similar A’s style is to B’s style (e.g., effect of homophily)
What we want: how much A’s inclusion of an article immediately triggers the usage of articles in B’s reply?

\[ \text{Coordination}_{(B \text{ to } A)}(\text{art.}) = \]
Measuring coordination

What we want: how much A’s inclusion of an article immediately triggers the usage of articles in B’s reply?

\[ Coordination_{(B \text{ to } A)}(\text{art.}) = P(B^{\text{art.}} | B \text{ replied to } A, A^{\text{art.}}) \]
What we want: how much A’s inclusion of an article immediately triggers the usage of articles in B’s reply?

\[
Coordination_{(B \text{ to } A)}(art.) = P(B^{art.} | B \text{ replied to } A, A^{art}) - P(B^{art.} | B \text{ replied to } A)
\]

Baseline: controls for inherent similarity

Trigger
Overall coordination: average over all pairs of users \((B,A)\)

\[
P(B^{art.} \mid B \text{ replied to } A, A^{art})
\]

\[
P(B^{art.} \mid B \text{ replied to } A)
\]
Measuring coordination

Overall coordination: average over all pairs of users \((B,A)\)

\[ P(B^{art.} \mid B \text{ replied to } A, A^{art}) \]

\[ P(B^{art.} \mid B \text{ replied to } A) \]
Measuring coordination

**Overall coordination:** average over all pairs of users \((B,A)\)

\[
P(B^{art.} | B \text{ replied to } A, A^{art}) \quad \text{and} \quad P(B^{art.} | B \text{ replied to } A)
\]
Empirical validation

Complete Twitter conversational history
for 2,200 pairs of users
215,000 conversations
Average of 100 conversations/pair
held over an average of 270 days
Empirical validation

Statistically significant (p<0.0001) effect of coordination
Empirical validation

The first time coordination is shown to occur in a large-scale, real-life setting!
Sociolinguistic problems:


- Relationship nature and communication patterns – Voluntary vs. involuntary relations (e.g., dating partners vs. in-laws) [Mongeau & Henningsten, 2008]

- Communication accommodation [Giles et al., 1991]

Large-scale socially-rich conversational data:

- Facebook, Twitter, Second life, reality shows, etc.

Linguistic style coordination

Mathematical formalism

Large scale, real life setting

Practical applications

Back to psycho-linguistics
Stylistic influence

Given two people, one can impose her style on the other more than vice-versa.
Stylistic influence

Given two people, one can impose her style on the other more than vice-versa.

\[ \text{Influence}_{(A \text{ on } B)}(\text{art.}) = \text{Coordination}_{(B \text{ to } A)}(\text{art.}) - \text{Coordination}_{(A \text{ to } B)}(\text{art.}) \]
Stylistic influence

Given two people, one can impose her style on the other more than vice-versa.

\[ \text{Influence}_{(A \text{ on } B)}(\text{art.}) = \text{Coordination}_{(B \text{ to } A)}(\text{art.}) - \text{Coordination}_{(A \text{ to } B)}(\text{art.}) \]

A statistically significant imbalance between users is prevalent in Twitter.

- revealing a complexity of the phenomenon never observed before
Stylistic influence

Is coordination symmetric?
Stylistic influence

- **Symmetric:**
  \[ \text{Coordination}_{(B \text{ to } A)}(F) > 0 \ \& \ \text{Coordination}_{(A \text{ to } B)}(F) > 0 \]

- **Asymmetric:**
  - **Default:**
    \[ \text{Coordination}_{(B \text{ to } A)}(F) > 0 \ \& \ \text{Coordination}_{(A \text{ to } B)}(F) = 0 \]
  - **Diverging:**
    \[ \text{Coordination}_{(B \text{ to } A)}(F) > 0 \ \& \ \text{Coordination}_{(A \text{ to } B)}(F) < 0 \]
Stylistic influence

- **Symmetric:**
  \[\text{Coordination}_{(B \text{ to } A)}(F) > 0 \text{ } \& \text{ } \text{Coordination}_{(A \text{ to } B)}(F) > 0\]

- **Asymmetric:**
  - **Default:**
    \[\text{Coordination}_{(B \text{ to } A)}(F) > 0 \text{ } \& \text{ } \text{Coordination}_{(A \text{ to } B)}(F) = 0\]
  - **Diverging:**
    \[\text{Coordination}_{(B \text{ to } A)}(F) > 0 \text{ } \& \text{ } \text{Coordination}_{(A \text{ to } B)}(F) < 0\]
Stylistic influence

What type of people have more stylistic influence?
Stylistic influence

What type of people have more stylistic influence?

case study: gender (of movie characters)
Stylistic influence and gender

Movie scripts conversation dataset
220,000 conversations
617 movies
between 9000 characters
with known gender!

[with Lillian Lee, forthcoming]
Movie characters coordinate!
Movie characters coordinate!

[with Lillian Lee, forthcoming]
Movie characters coordinate!

➢ Coordination is so embedded in our language generation process, that it is exhibited even in our imagined conversations.

[with Lillian Lee, forthcoming]
Stylistic influence and gender

[with Lillian Lee, forthcoming]
Stylistic influence and gender

[with Lillian Lee, forthcoming]
Stylistic influence and gender

[with Lillian Lee, forthcoming]
Stylistic influence and gender

[with Lillian Lee, forthcoming]
Contributions

Mathematical formalism: probabilistic framework

Psycho-linguistics

Linguistic style coordination

Back to psycho-linguistics
robustness
stylistic influence
relation to gender
(read the paper for more results)

Large scale environment:
Twitter conversations & movie conversations

Practical applications

Large-scale socially-rich conversational data:
Facebook, Twitter, Second life, reality shows, etc.

Psycholinguistics

Practical applications

Mathematical formalism: probabilistic framework

Psycho-linguistics

Linguistic style coordination

Back to psycho-linguistics
robustness
stylistic influence
relation to gender
(read the paper for more results)
Thank you!
Thank you!

Good karma!