Measuring Variations of Mimicry by Means of Prosodic Cues in Task-Based Scenarios and Conversational Speech

Brian Vaughan

Dublin Institute of Technology, brian.vaughan@dit.ie

Celine De Looze

Trinity College Dublin

Follow this and additional works at: http://arrow.dit.ie/dmcoth

Part of the Computer Sciences Commons, Phonetics and Phonology Commons, and the Signal Processing Commons

Recommended Citation

Measuring variations of mimicry by means of prosodic cues

Céline De Looze & Brian Vaughan
TCD, Dublin, Ireland
Mimicry in speech

Speakers imitate each other’s speech mannerisms in terms of sounds, syntax, lexicon, prosody

… accommodation, alignment, convergence, entrainment, synchrony…

Non-Mimicry: random

Mimicry: parallel patterns

Anti-mimicry: mirror patterns

Convergence: converge towards a common point

Divergence: move apart towards different points
Mimicry in speech

Speakers imitate each other’s speech mannerisms in terms of sounds, syntax, lexicon, prosody

… accommodation, alignment, convergence, entrainment, synchrony…

Non-Mimicry: random

Mimicry: parallel patterns

Anti-mimicry: mirror patterns

TRANSITION PHASES

Convergence: converge towards a common point

Divergence: move apart towards different points
Mimicry in speech

The situation where the observed behaviours of two interactants although dissimilar at the start of the interaction are moving towards behavioral matching (Burgoon et al 1995)

Speakers tend to imitate over the course of the interaction?

Phases of mimicry and non-mimicry
Mimicry measurements

Metrics developed may not capture the temporal dynamics of mimicry (except Jaffe et al, 2001; Edlund et al 2009)

Mimicry strength measured
- on the whole interaction
- on parts of the interaction
Data

Task-based scenarios: Co-operation between 2 participants to complete an imaginary shipwreck scenario. Time, score and functional constraints.
8 dialogues, 10 minutes. Male & Female, but not mixed.

Spontaneous speech:
D64 corpus (Oertel et al, 2010)
Two interactions (S1/S2 & S1/S3)
2M & 1F
30 min each
Mimicry measurements

Prosodic cues:
- Pitch level and span: f0-average + f0-max-min
- Voice Intensity: rms-Int + sd-Int
- Duration: number and mean pause duration
Mimicry measurements

Prosodic cues:
- Pitch level and span: f0-median + f0-max-min
- Voice Intensity: rms-Int + sd-Int
- Duration: number and mean pause duration

Task-based dialogues

Spontaneous speech
Mimicry measurements

Methods:
- Prosodic cues extraction: a series of overlapping windows (length = 20 sec; time step = 10 sec) (Kousidis et al, 2008; Edlund et al, 2009)

*Figure 1: Conversation Chart*
Mimicry measurements

Methods:
- Mimicry strength measurement: Pearson’s correlation coefficient of the two speakers’ time series (use of moving windows for temporal variations)

*Figure 2*
Mimicry measurements

Methods:
- Mimicry strength measurement: Pearson’s correlation coefficient of the two speakers’ time series (use of moving windows for temporal variations)

Figure 2
Mimicry measurements

Methods:
- Mimicry strength measurement: Pearson’s correlation coefficient of the two speakers’ time series (use of moving windows for temporal variations)

*Figure 2*
Mimicry measurements

Methods:
- Mimicry strength measurement: Pearson’s correlation coefficient of the two speakers’ time series (use of moving windows for temporal variations)

Figure 2
Mimicry - functions

Mimicry plays an important role in social interaction
- express deference, speakers seek each other’s approval
- its absence: maintain social distance with each other?
- signal agreement?

Level of agreement (DAMSL: 5-point scale) in task-based dialogues.

Degree of involvement (scale 0-10) in spontaneous speech.
Results

Task-based dialogues

Whole interaction
• D1: Weak for mean pitch, pitch range, intensity
• D2: Stronger for pitch, pitch range, max pitch, mean intensity

Windowed correlation
D1 & D2: Change in some of the values in either direction.
Results

Task-based dialogues

D1 Overall:
• Mean Pitch: 
r=0.377
• Pitch Range (Semi-Tones) 
r=0.11
• Min pitch: 
r=0.044
• Max pitch: 
r=0.073
• Mean intensity: 
r=-0.83
• Intensity range: 
r=0.273
• Agreement: 😊

D2 Overall:
• Mean Pitch: 
r= 0.711
• Pitch Range (Semi-Tones) 
r= 0.433
• Min pitch: 
r=0.254
• Max pitch: 
r=0.712
• Mean intensity: 
r=0.795
• Intensity range: 
r=0.752
• Agreement: 😇?
Results

Task-based dialogues: Change of values with windowed correlation (10 points: 200 seconds).

D1 windowed correlation.
1-10: Pitch range, $r=0.444$
10-20: Max pitch, $r=0.814$
20-30: Mean pitch, $r=0.657$
30-40: Mean pitch, $r=0.448$
40-50: Mean Pitch, $r=0.575$
Min pitch, $r=0.666$
50-60: Mean intensity, $r=-0.602$
60-66: Mean pitch, $r=0.690$
Min pitch, $r=-0.741$
Intensity, $r=-0.786$

D2 windowed correlation
1-10: Max pitch, $r=0.531$
10-20: Mean intensity, $r=0.905$
Intensity range, $r=0.899$
20-30: Mean intensity, $r=0.899$
Intensity range, $r=0.875$
30-40: Max pitch, $r=0.753$
Pitch range ST, $r=0.758$
40-50: Pitch range ST, $r=0.868$
50-61: Intensity range, $r=0.837$
Results

Mimicry and levels of agreement

In task-based dialogues

Mimicry when agreement AND disagreement. E.g. Mean pitch

Movement is in either direction: increase AND decrease. E.g Mean pitch.
Results

Task-based dialogues

D1 Agreement.
Results

Task-based dialogues

D2 Agreement.
Results

Spontaneous speech

On the whole interaction
- S1/S2: No mimicry
- S1/S3: Mimicry in voice intensity level, variation in pitch range ceiling, mean pauses duration

Temporal variations of mimicry
- S1/S2: Mimicry in voice intensity level
- S1/S3: Mimicry in pitch range ceiling, mean pauses duration
Results

Mimicry and degrees of involvement

*In spontaneous speech (S1/S2)*

- Mean(I) = calculated from the set of 8 prosodic cues
Results

Mimicry and degrees of involvement

In spontaneous speech (S1/S2)

- Strong correlation:
The higher the degree of involvement, the stronger the mimicry
Results

Mimicry and degrees of involvement

In spontaneous speech (S1/S2)

- Strong correlation:
  The higher the degree of involvement, the stronger the mimicry

In terms of
- Rms_Intensity (r=0.89)
- Mean_pause_dur (r=0.89)
- Number_pauses (r=0.59)
- F0-min (r=0.55)
- F0-span (r=0.51)
- F0-median (r=0.42)
Conclusion

- Mimicry can be measured by means of prosodic cues
- A non-linear phenomenon
- Temporal dynamics of mimicry as strong cues for predicting involvement
- Mimicry at points of agreement and disagreement.
Discussion

Mimicry: task-dependent?
Not necessarily a linear phenomenon
In spontaneous speech: dynamics of mimicry

Mimicry or use of the same prosodic parameters to convey the same functions (e.g. discourse, attitudinal)?

Z-score transformations for detecting mimicry but not convergence?
Future work

Methodology:
- Capture smaller variations of mimicry
- Measure anti-mimicry, convergence and divergence phases
- Measure who mimics whom?
- Improve/increase annotation of agreement

Correlation between temporal variations of mimicry and discourse structure (e.g. topic changes)

....
Thanks!