

Entrainment on Backchannel- Preceding Cues

Rivka Levitan

Agustín Gravano

Julia Hirschberg



Entrainment

'Are their heads off?' shouted the Queen.

'Their heads are gone, if it please your Majesty!'
the soldiers shouted in reply.

'That's right!' shouted the Queen. 'Can you play
croquet?'

'Yes!' shouted Alice.

'Come on, then!' roared the Queen, and Alice
joined the procession, wondering very much
what would happen next.



--Alice's Adventures in Wonderland

Entrainment

'Are their heads off?' **shouted** the Queen.

'Their heads are gone, if it please your Majesty!'
the soldiers **shouted** in reply.

'That's right!' **shouted** the Queen. 'Can you play
croquet?'

'Yes!' **shouted** Alice.

'Come on, then!' **roared** the Queen, and Alice
joined the procession, wondering very much
what would happen next.



--Alice's Adventures in Wonderland

Backchannels

Speaker A: All right so I have a- a a nail on top

Speaker B: *okay*

Speaker A: with an owl in the lower left.

Backchannel: a short segment of speech uttered by a listener to convey that she is paying attention.

A. Gravano & J. Hirschberg, 2009

Backchannel-inviting cues in task-oriented dialogue

T1: speech <pause> speech <pause> **speech**

T2: <backchannel>

Cue	Feature
Intonation	pitch slope over the IPU-final 200 and 300 ms
Pitch	mean pitch over the final 500 and 1000 ms
Intensity	mean intensity over the final 500 and 1000 ms
Duration	IPU duration in seconds and word count
Voice quality	NHR over the final 500 and 1000 ms

A. Gravano & J. Hirschberg, 2009

Backchannel-inviting cues in task-oriented dialogue

T1: IPU <pause> IPU <pause> IPU

T2: <backchannel>

Cue	Feature
Intonation	pitch slope over the IPU-final 200 and 300 ms
Pitch	mean pitch over the final 500 and 1000 ms
Intensity	mean intensity over the final 500 and 1000 ms
Duration	IPU duration in seconds and word count
Voice quality	NHR over the final 500 and 1000 ms

Entrainment on backchannel- preceding cues

- Different speakers display different combinations of cues.
- Different speakers may realize a cue in different ways.

Entrainment on backchannel- preceding cues

- Different speakers display different combinations of cues.
- Different speakers may realize a cue in different ways.
- Do speaker pairs entrain on backchannel-
preceding cues?

Entrainment on backchannel- preceding cues

- Three measures of backchannel cue entrainment:
 - Measure 1: Cue sets
 - Measure 2: Cue realization
 - Measure 3: Local cue realization

The Columbia Games Corpus

- Twelve spontaneous task-oriented conversations
- In each conversation, two people, separated by a curtain, played computer games.
- Thirteen people participated; eleven participated in two sessions each.
- All features extracted automatically in Praat.



Measure 1: Similar cue sets

- Do speakers use similar sets of cues?

Speaker 1

Rising intonation

Higher pitch

Longer duration

Speaker 2

Rising intonation

Higher intensity

Longer duration

Change in voice quality

2 cues in common

Measure 1: Similar cue sets

- Do speakers use similar sets of cues?

<u>Speaker 1</u>	<u>Speaker 2</u>
Rising intonation	Rising intonation
Higher pitch	Higher intensity
Longer duration	Longer duration
	Change in voice quality

2 cues in common

Measure 1: Similar cue sets

- Do speakers use similar sets of cues?

<u>Speaker 1</u>	<u>Speaker 2</u>
Rising intonation	Rising intonation
Higher pitch	Higher intensity
Longer duration	Longer duration
	Change in voice quality

2 cues in common

Yes (paired t -test, $p < 0.05$)

Measure 2: Cue realization

- Do speakers realize their cues in similar ways?

For any feature f modeling cue c :

$$d_{1,2}^f < \min(d_{1,*}^f, d_{*,2}^f)$$



this session exhibits entrainment on c .

Measure 2: Cue realization

<i>Session</i>	<i>Intensity</i>	<i>Pitch</i>	<i>Voice</i>	<i>Intonation</i>	<i>Duration</i>
1	✓	✓	✓	✓	
2	✓	✓	✓		
3	✓	✓		✓	✓
4	✓	✓	✓	✓	✓
5	✓	✓		✓	
6	✓	✓		✓	
7	✓	✓	✓		
8	✓	✓		✓	✓
9	✓	✓	✓	✓	✓
10	✓	✓	✓		✓
11	✓	✓	✓	✓	✓
12			✓		✓

Measure 2: Cue realization

<i>Session</i>	<i>Intensity</i>	<i>Pitch</i>	<i>Voice</i>	<i>Intonation</i>	<i>Duration</i>
1	✓	✓	✓	✓	
2	✓	✓	✓		
3	✓	✓		✓	✓
4	✓	✓	✓	✓	✓
5	✓	✓		✓	
6	✓	✓		✓	
7	✓	✓	✓		
8	✓	✓		✓	✓
9	✓	✓	✓	✓	✓
10	✓	✓	✓		✓
11	✓	✓	✓	✓	✓
12			✓		✓

Measure 2: Cue realization

<i>Session</i>	<i>Intensity</i>	<i>Pitch</i>	<i>Voice</i>	<i>Intonation</i>	<i>Duration</i>
1	✓	✓	✓	✓	
2	✓	✓	✓		
3	✓	✓		✓	✓
4	✓	✓	✓	✓	✓
5	✓	✓		✓	
6	✓	✓		✓	
7	✓	✓	✓		
8	✓	✓		✓	✓
9	✓	✓	✓	✓	✓
10	✓	✓	✓		✓
11	✓	✓	✓	✓	✓
12		✓	✓		✓

Measure 2: Cue realization

<i>Session</i>	<i>Intensity</i>	<i>Pitch</i>	<i>Voice</i>	<i>Intonation</i>	<i>Duration</i>
1	✓	✓	✓	✓	
2	✓	✓	✓		
3	✓	✓		✓	✓
4	✓	✓	✓	✓	✓
5	✓	✓		✓	
6	✓	✓		✓	
7	✓	✓	✓		
8	✓	✓		✓	✓
9	✓	✓	✓	✓	✓
10	✓	✓	✓		✓
11	✓	✓	✓	✓	✓
12			✓		✓

Measure 2: Cue realization

<i>Session</i>	<i>Intensity</i>	<i>Pitch</i>	<i>Voice</i>	<i>Intonation</i>	<i>Duration</i>
1	✓	✓	✓	✓	
2	✓	✓	✓		
3	✓	✓		✓	✓
4	✓	✓	✓	✓	✓
5	✓	✓		✓	
6	✓	✓		✓	
7	✓	✓	✓		
8	✓	✓		✓	✓
9	✓	✓	✓	✓	✓
10	✓	✓	✓		✓
11	✓	✓	✓	✓	✓
12			✓		✓

Measure 2: Cue realization

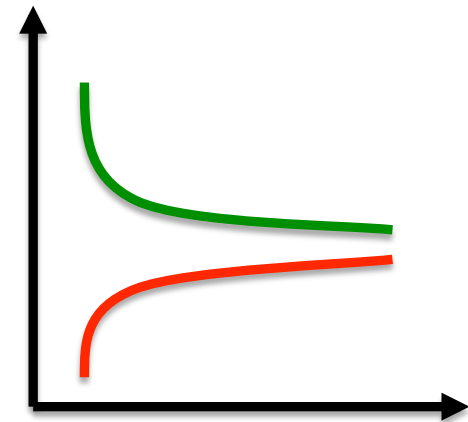
<i>Session</i>	<i>Intensity</i>	<i>Pitch</i>	<i>Voice</i>	<i>Intonation</i>	<i>Duration</i>
1	✓	✓	✓	✓	
2	✓	✓	✓		
3	✓	✓		✓	✓
4	✓	✓	✓	✓	✓
5	✓	✓		✓	
6	✓	✓		✓	
7	✓	✓	✓		
8	✓	✓		✓	✓
9	✓	✓	✓	✓	✓
10	✓	✓	✓		✓
11	✓	✓	✓	✓	✓
12			✓		✓

Measure 2: Cue realization

- Pitch, intensity and voice quality show significantly smaller differences between partners than between non-partnered pairs.

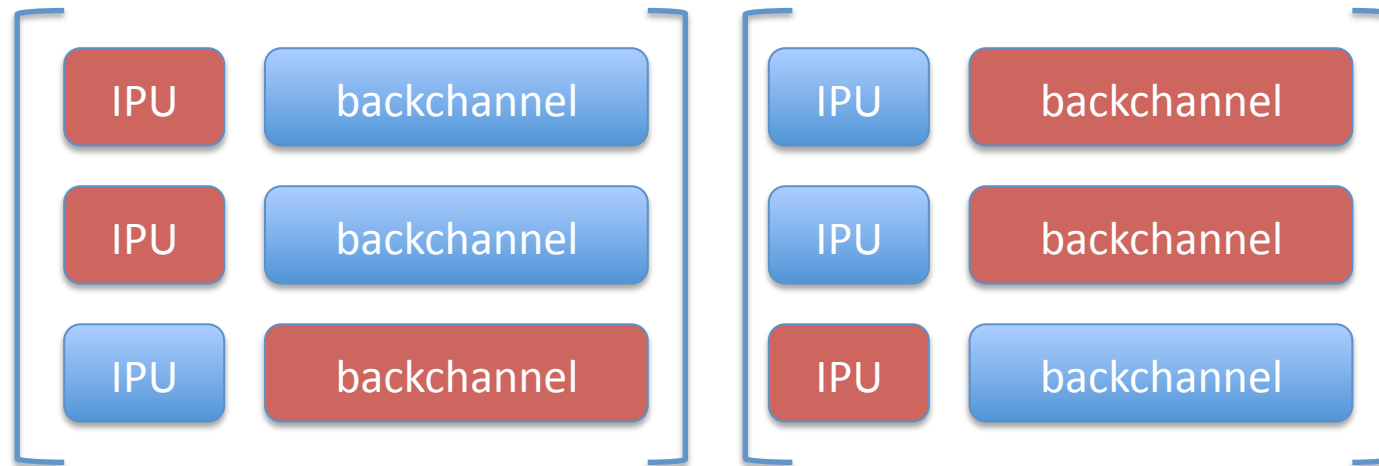
Measure 2: Cue realization

- Pitch, intensity and voice quality show significantly smaller differences between partners than between non-partnered pairs.
- Convergence: The differences between partners in intensity and pitch are smaller in the last half of a conversation than the first.



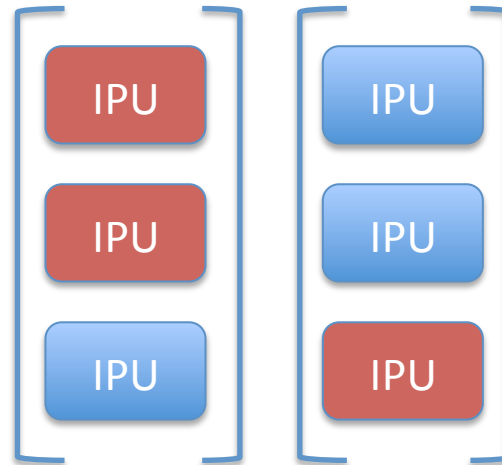
Measure 3: Local entrainment

- Does the way one speaker invites a backchannel affect how her partner invites his next backchannel?



Measure 3: Local entrainment

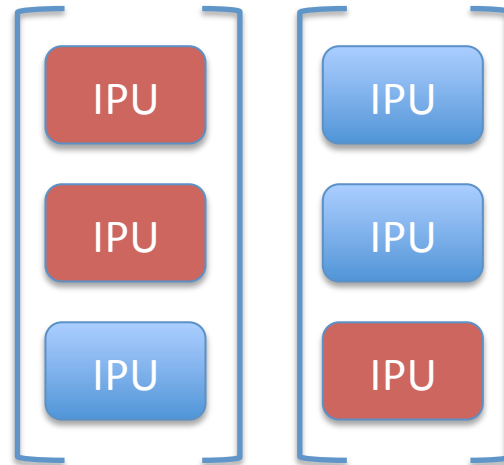
- Does the way one speaker invites a backchannel affect how her partner invites his next backchannel?



Pearson's correlation

Measure 3: Local entrainment

- Does the way one speaker invites a backchannel affect how her partner invites his next backchannel?



Pearson's correlation

⇒ Significant correlations for mean pitch and intensity ($r=0.3$).

Entrainment on backchannel- preceding cues

- Three measures of backchannel cue entrainment:
 - Measure 1: Cue sets
 - Measure 2: Cue realization
 - Measure 3: Local cue realization

Entrainment on backchannel- preceding cues

- Three measures of backchannel cue entrainment:
 - Measure 1: Cue sets ✓
 - Measure 2: Cue realization
 - Measure 3: Local cue realization

Entrainment on backchannel- preceding cues

- Three measures of backchannel cue entrainment:
 - Measure 1: Cue sets ✓
 - Measure 2: Cue realization ✓
 - Measure 3: Local cue realization

Entrainment on backchannel- preceding cues

- Three measures of backchannel cue entrainment:
 - Measure 1: Cue sets ✓
 - Measure 2: Cue realization ✓
 - Measure 3: Local cue realization ✓

Association with dialogue success

- Three measures of dialogue success:

- Score

- Proportion of interruptions

S1: <unfinished speech>

S2: <speech>

- Mean backchannel latency

S1: <speech>

S2: <pause> <backchannel>



Association with dialogue success

Entrainment measure	Success measure	<i>r</i>	<i>p</i>-value
Common cues	Latency	-0.33	0.06
	Interruptions	-0.50	0.01
	Score	0.22	N.S.
Similar means	Latency	-0.61	0.002
	Interruptions	-0.22	N.S.
	Score	0.72	6.9e-05

Association with dialogue success

Entrainment measure	Success measure	<i>r</i>	<i>p</i>-value
Common cues	Latency	-0.33	0.06
	Interruptions	-0.50	0.01
	Score	0.22	N.S.
Similar means	Latency	-0.61	0.002
	Interruptions	-0.22	N.S.
	Score	0.72	6.9e-05

Association with dialogue success

Entrainment measure	Success measure	<i>r</i>	<i>p</i>-value
Common cues	Latency	-0.33	0.06
	Interruptions	-0.50	0.01
	Score	0.22	N.S.
Similar means	Latency	-0.61	0.002
	Interruptions	-0.22	N.S.
	Score	0.72	6.9e-05

Conclusions

- ✓ Speakers use similar sets of cues.
- ✓ They realize the cues in similar ways.
- ✓ This entrainment is local as well as global.
- ✓ This entrainment is positively associated with task success and negatively associated with poor coordination.

Future Work

- Who entrains to whom?
- What is the relationship between different types of entrainment?
- Is this entrainment associated with more perceptual measures of task success?

Thank you!