Entrainment

Rivka Levitan, PhD
Guest Lecture: Advanced Topics in Spoken Language Processing
Spring 2026

What is 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*

What is entrainment?

'Jeeves,' I said, 'you're talking rot.'

'Very good, sir.'

'Absolute drivel.'

'Very good, sir.'

'Pure mashed potatoes.'

'Very good, sir.'

'Very good, sir – I mean, very good, Jeeves, that will be all,' I said.

And I drank a modicum of tea, with a good deal of hauteur.

–*Very Good, Jeeves*

Evidence of entrainment

- Lexical
  - Referring expressions: Brennan & Clark, 1992
  - High frequency words: Nenkova et al., 2008
  - Syntax: Branigan et al., 2000; Reitter et al., 2010
  - Linguistic Style Matching: Niederhoffer & Pennebaker, 2002; Danescu-Niculescu-Mizil et al., 2011
  - To a computer: Brennan, 1996; Stoyanchev & Stent, 2009

Evidence of entrainment

- Acoustic-prosodic:
  - Response time: Matarazzo & Wiens, 1967; Street, 1984
  - Intensity, pitch: Natale, 1975; Gregory et al., 2003; Ward & Litman, 2007
  - To a computer: Bell et al., 2003; Coulston et al., 2002
  - Intensity, pitch, speaking rate, voice quality, backchannel-inviting cues, pitch contours: Levitan et al. 2011, 2012, 2014, 2015, 2016

Dialogue quality

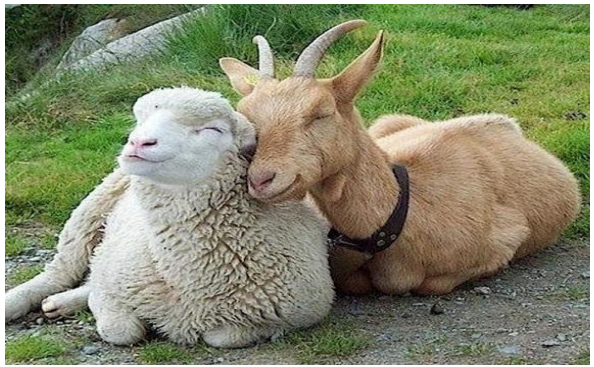
- Positive interactions in married couples (Lee et al., 2010)
- Score on the Map Task (Reitter and Moore, 2007)
- Liking, smoother interaction (Chartrand & Bargh, 1999)
- Social desirability (Natale, 1975)
- Power (Danescu-Niculescu-Mizil et al., 2012)
- Smoother interaction, task success (Nenkova et al., 2008)
- Romantic interest (Ireland et al., 2014)
- Turn taking, encouraging, trying to be liked (Levitan et al., 2012)

Dialogue quality

- Subjects who entrain
  - Perceived as more socially attractive (Putnam & Street '84, Bourhis et al '75)
  - Perceived as more competent (Street '84)
  - Conversation perceived as more intimate (Buller & Aune '88)
- Entrainment leads subjects to like their conversational partners (and their computers) more and to perceive interactions as more successful (Nass et al ‘95, Chartrand & Bargh ‘99)
- Long-term syntactic entrainment is a good predictor of actual task success in Map Task (Reitter et al ’07)

Entrainment theory

- Communication Accommodation Theory (Giles et al., 1991)
- Communication model (Natale, 1975)
- Perception-behavior link (Chartrand & Bargh, 1999)
- Interactive Alignment Theory (Pickering & Garrod, 2004)



Social

Automatic

Measuring entrainment

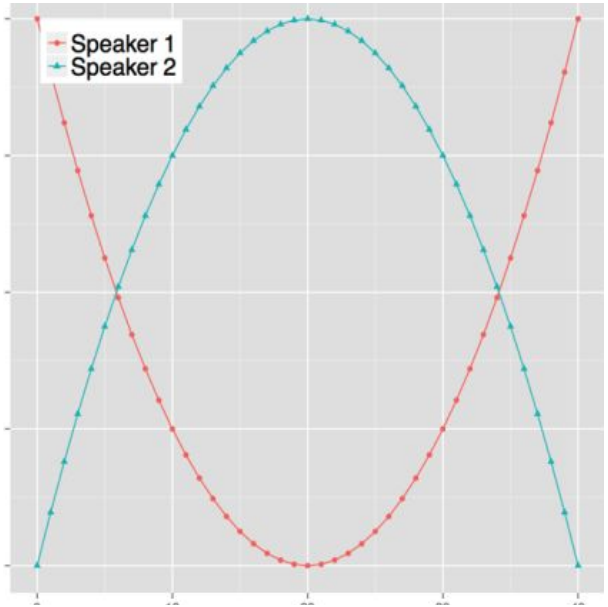
- **Global vs. local**

  Global: compare average to baseline

  - other speakers
  - self in other conversation

  Local: compare difference at turn exchanges to baseline

  - non-adjacent turns



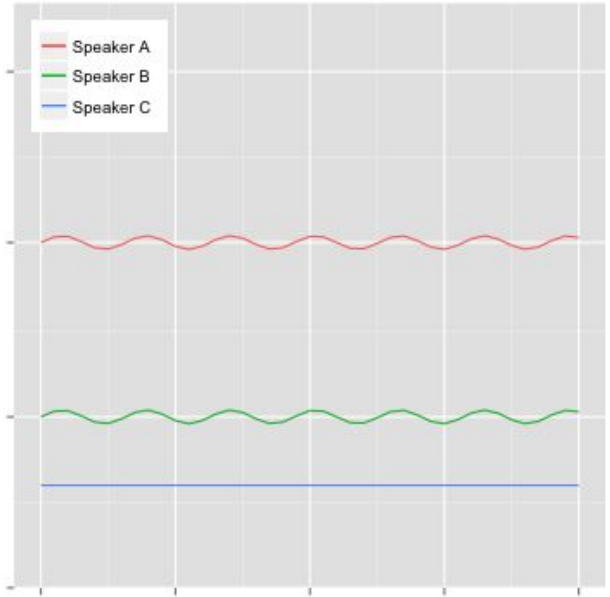
Levitan & Hirschberg, 2011

Measuring entrainment

- Global vs. local
- **Exact vs. relative**

  Exact: compare difference between adjacent feature values to baseline

  Relative: correlation of adjacent feature values



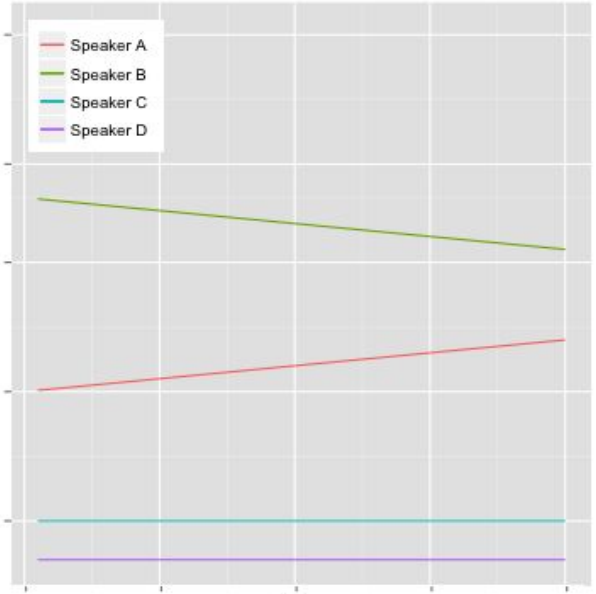
Levitan & Hirschberg, 2011

Measuring entrainment

- Global vs. local
- Exact vs. relative
- **Converging vs. constant**

  Global: compare difference in averages over time

  Local: correlate adjacent differences with time



Evidence of entrainment in the Columbia Games Corpus

- Global: intensity, speaking rate
  - Convergence: Pitch max, NHR, speaking rate (reset effect)
- Local: intensity, NHR
  - Convergence: all except jitter and speaking rate; weak
- Synchrony: moderate for intensity, none for speaking rate, others weak

Levitan & Hirschberg, 2011

Variation across speakers in other corpora

	Deception (EE)		Deception (ER)		Fisher	
	conv.	synch.	conv.	synch.	conv.	synch.
Entraining speakers	47%	53%	42%	47%	39%	46%
Valence						
positive	42%	68%	40%	65%	37%	69%
negative	52%	18%	51%	26%	52%	22%
mixed	6%	14%	9%	9%	11%	9%
#Features						
1	68%	55%	73%	64%	74%	69%
2	25%	30%	19%	28%	23%	25%
3+	7%	15%	8%	8%	3%	6%
max.	4	5	4	4	4	4

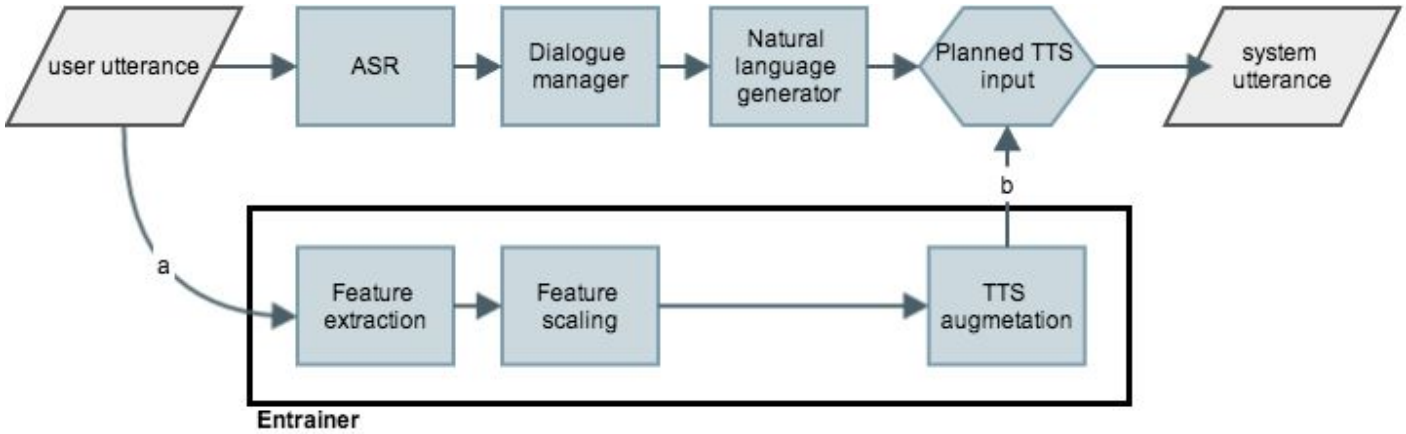
Weise et al, 2019

Variations across speakers

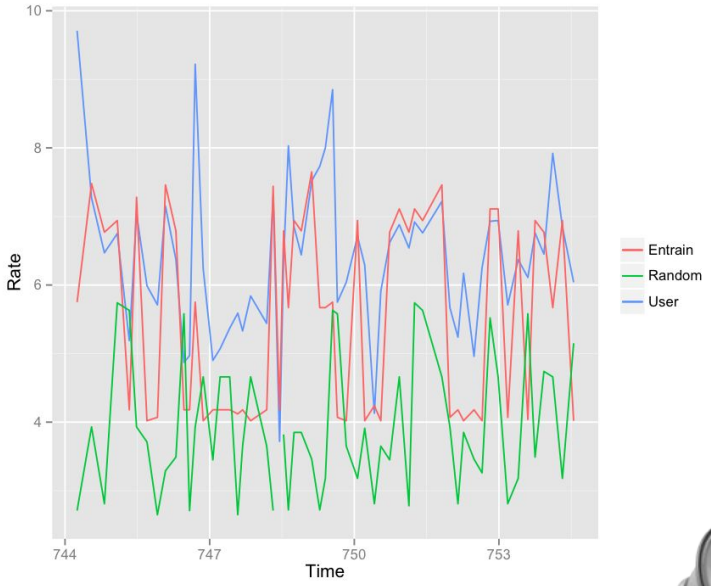
- Some speakers don’t entrain at all
- Some entrain only positively
- Some entrain only negatively
- Some entrain positively for some features, negatively for others
- This variation is not explained by gender, native language, or conversational role

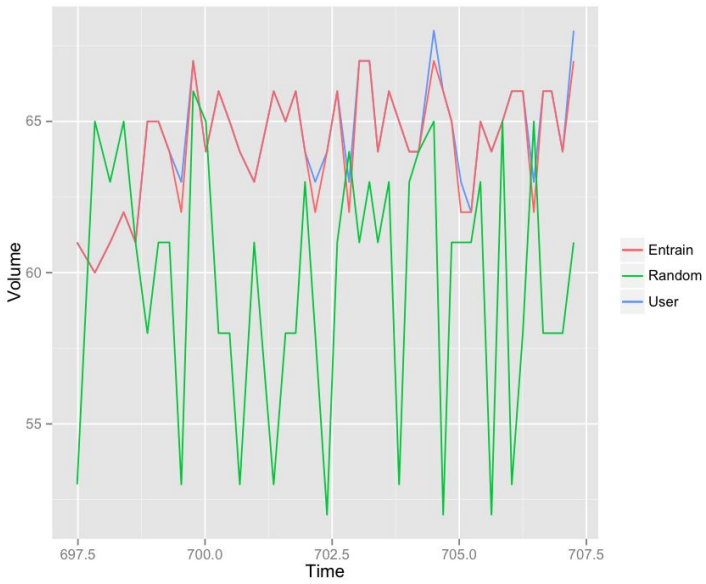
Weise et al, 2019

Implementing entrainment

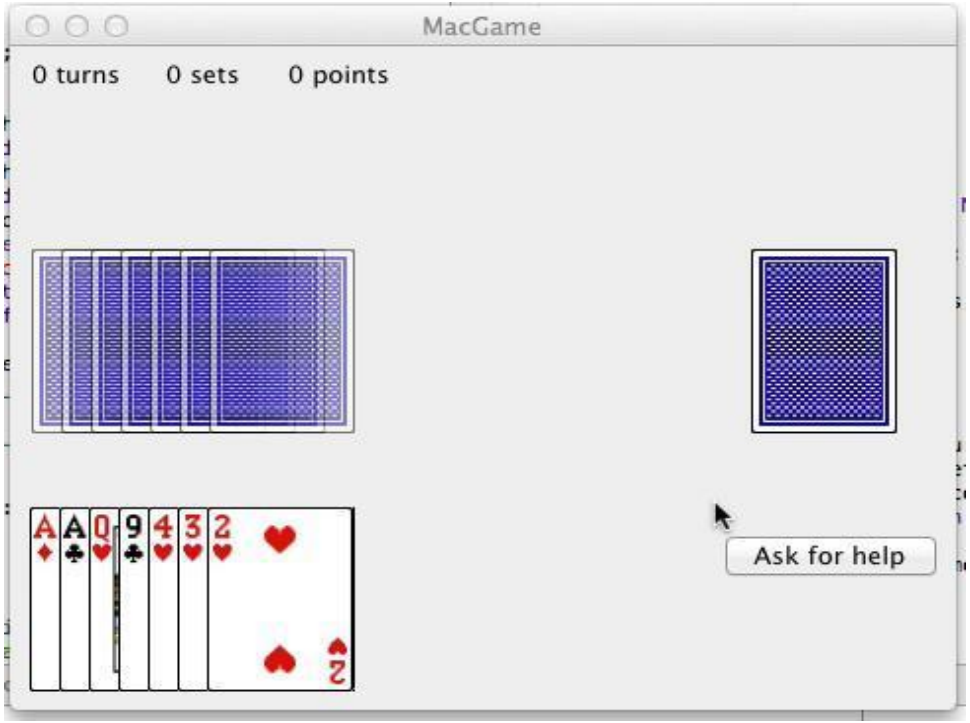


Performance





Do users prefer an entraining system?



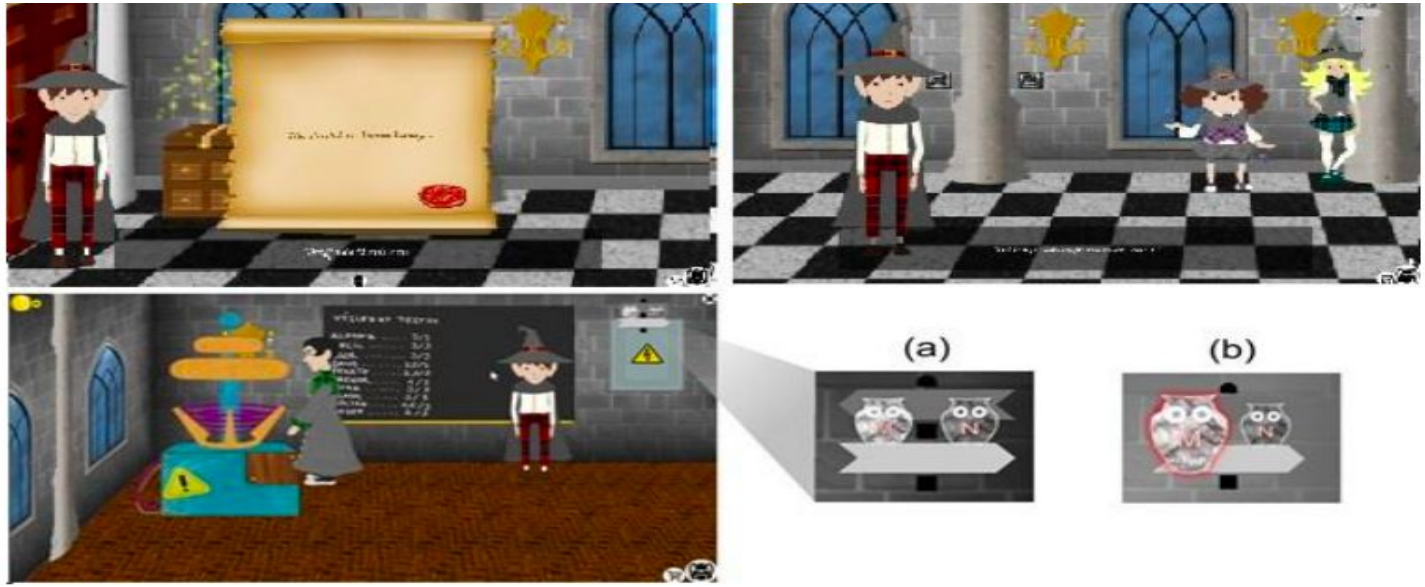
Do users prefer an entraining system?

"Bobby and Freddy can see the system's hand and will tell you which rank to ask for. They will usually give you good advice, but sometimes they will give you bad advice. One will give you bad advice more often than the other. Your role in this game is to decide who to trust."

Do users prefer an entraining system?

Language	Avatar gender	Entrainment		Baseline	Entrainment × Advice score
		Features	Method		
English	Male	Intensity Speech rate	Absolute	Random	+ ($p < 0.001$)
Spanish	Female	Speech rate	Relative	Disentrain	- ($p < 0.1$)
				Constant	no effect
Spanish	Female	Speech rate	Relative	Disentrain	- ($p < 0.05$)
				Constant	no effect

Wizard game



Wizard trust results

- Subjects trust **complementary entraining** avatar
  - Especially females
  - High-openness subjects
- Preference flips towards end of game?
  - High-stakes decision

Leveraging entrainment

Prosodic Entrainment in an Information-Driven Dialog System (Fandrianto & Eskenazi, 2012)

type	shouting	hyperarticulation
explicit	prompt to speak more quietly	prompt to speak "normally"
implicit	lower volume	increase synthesis speaking rate
backoff	switch slot objective	switch slot objective

Table 3: *Strategies by speaking style*

Leveraging entrainment

Prosodic Entrainment in an Information-Driven Dialog System (Fandrianto & Eskenazi, 2012)

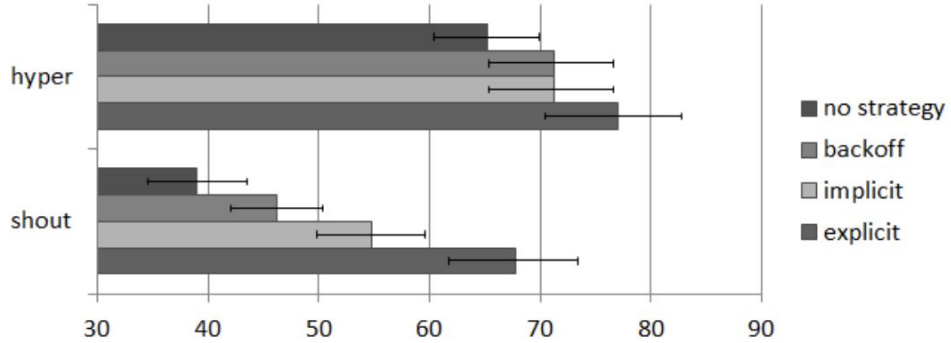


Figure 1: *Styles vs percent rate style revert to "unmarked"; error bars indicate ≈ 95% of a binomial distribution's probability mass over observations*

Leveraging entrainment

Towards choosing better primes for spoken dialog systems (Lopes et al, 2011)

Prime	PER (%)	Frequency
próximo (next)	19.0	21
próximo autocarro (next bus)	26.3	19
anterior (previous)	18.2	22
autocarro anterior (previous bus)	18.8	16
outro percurso (another route)	50.0	6
nova pesquisa (new request)	36.4	11
agora (now)	60.6	33

Leveraging entrainment

Towards choosing better primes for spoken dialog systems (Lopes et al, 2011)

Concept	Old prime	New prime
Now	*agora*	*imediatamente neste momento o mais brevemente possível o mais rápido possível*
New query	*nova pesquisa outro percurso*	*procurar novamente nova procura outra procura nova busca*
Next bus	*próximo*	*seguinte*

Leveraging entrainment

Towards choosing better primes for spoken dialog systems (Lopes et al, 2011)

next bus	W1: *próximo*	62 (83.8)	3 (4.8)
	W2: *seguinte*	12 (16.2)	59 (95.2)

Leveraging entrainment

Towards choosing better primes for spoken dialog systems (Lopes et al, 2011)

	Week 1 (%)	Week 2 (%)
Correct	32.3	41.1
ASR	52.8	44.2

Disentrainment

- Andreas Weise, Sarah Ita Levitan, Julia Hirschberg, and Rivka Levitan. Individual differences in acoustic-prosodic entrainment in spoken dialogue. Speech Communication, 115:78–87, 2019.
- Céline De Looze, Catharine Oertel, Stéphane Rauzy, and Nick Campbell. Measuring dynamics of mimicry by means of prosodic cues in conversational speech. Proceedings of ICPhS. Springer, pages 1294–1297, 2011.
- Patrick GT Healey, Matthew Purver, and Christine Howes. Divergence in dialogue. PloS one, 9(6):e98598, 2014.
- Juan Manuel Pérez, Ramiro H Gálvez, and Agustín Gravano. Disentrainment may be a positive thing: A novel measure of unsigned acoustic-prosodic synchrony, and its relation to speaker engagement. In Interspeech, pages 1270–1274, 2016.

Nonlinear entrainment

Masahiro Mizukami, Koichiro Yoshino, Graham Neubig, David Traum, and Satoshi Nakamura. Analyzing the effect of entrainment on dialogue acts. In Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pages 310–318, 2016.

lexical entrainment is stronger given some dialogue acts :-> speakers were influenced by their interlocutor’s lexical realization of a particular dialogue act, rather than the most recent turn

Non-similarity-based entrainment measures

Lee et al (2010) quantify entrainment over stylized pitch and energy contours using squared correlation coefficient, mutual information, and spectral coherence

Lee et al (2011) quantify vocal entrainment as the amount of variance of a speaker's vocal feature that is preserved when projected onto the PCA-derived space of their interlocutor's vocal features

Neural methods

Nasir et al, 2018

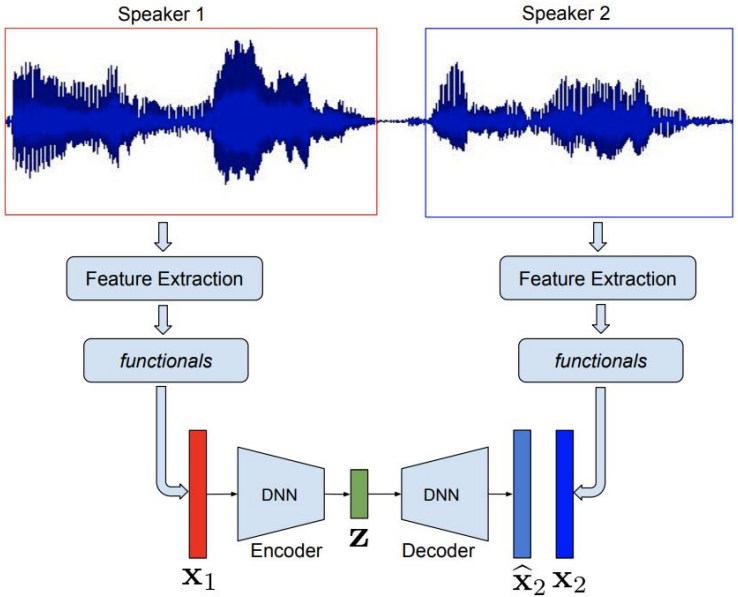
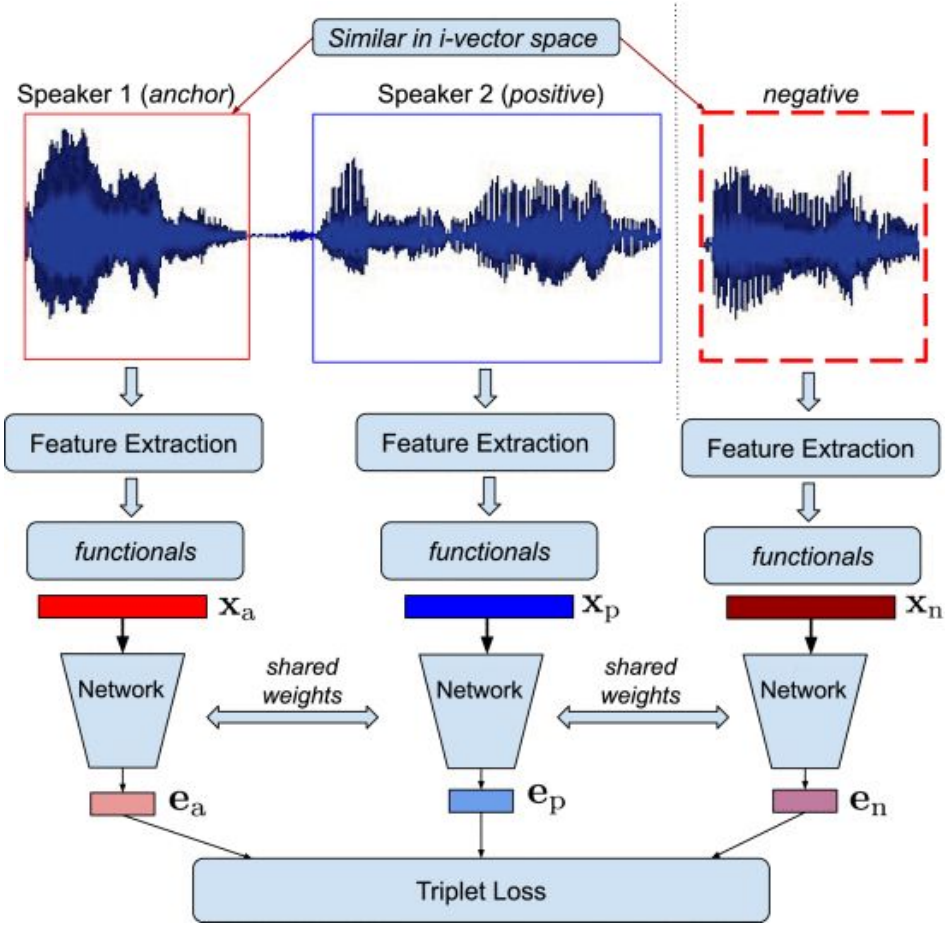


Figure 1: *An overview of unsupervised training of the model*

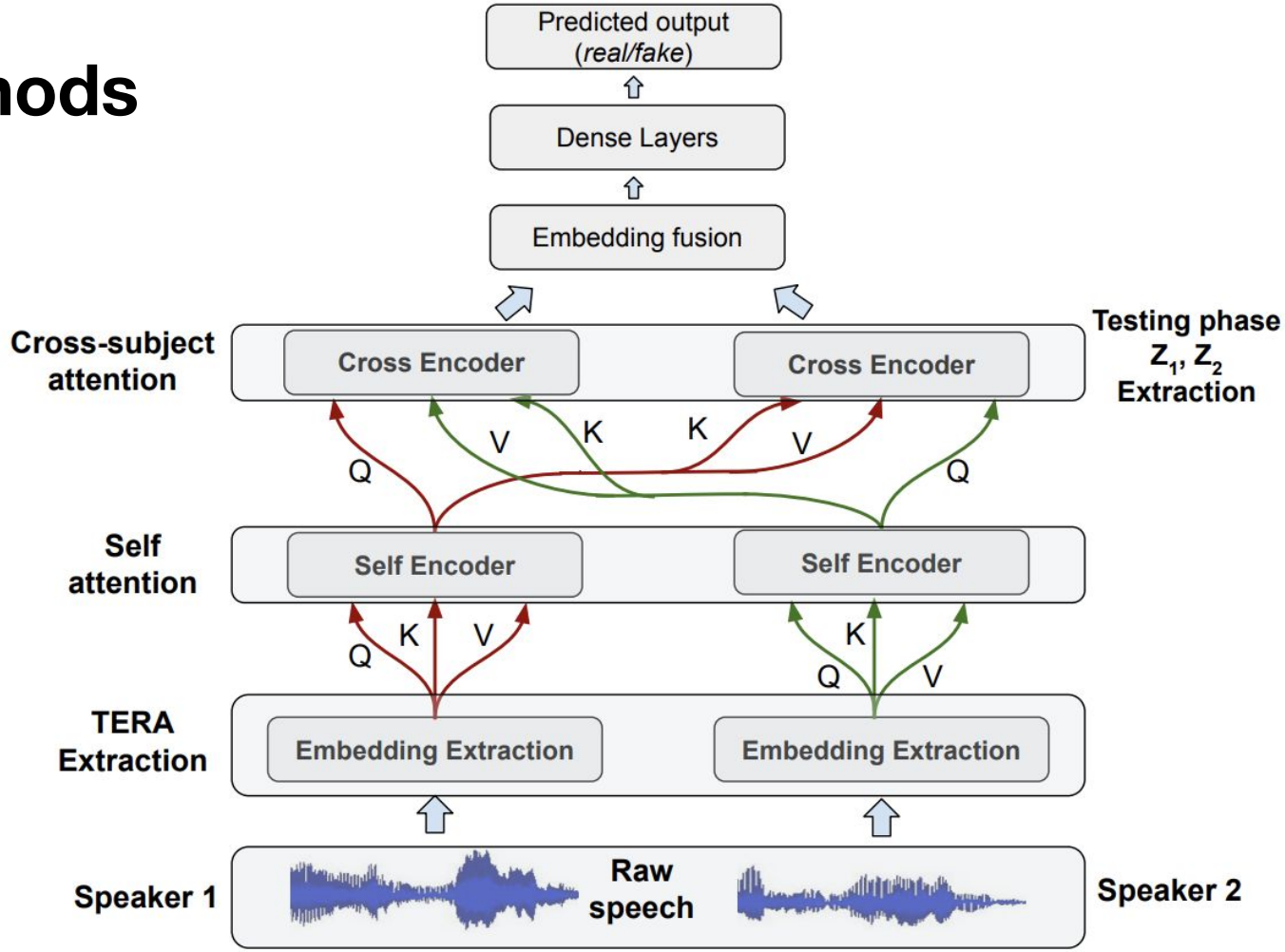
Neural methods

[Nasir et al, 2020]

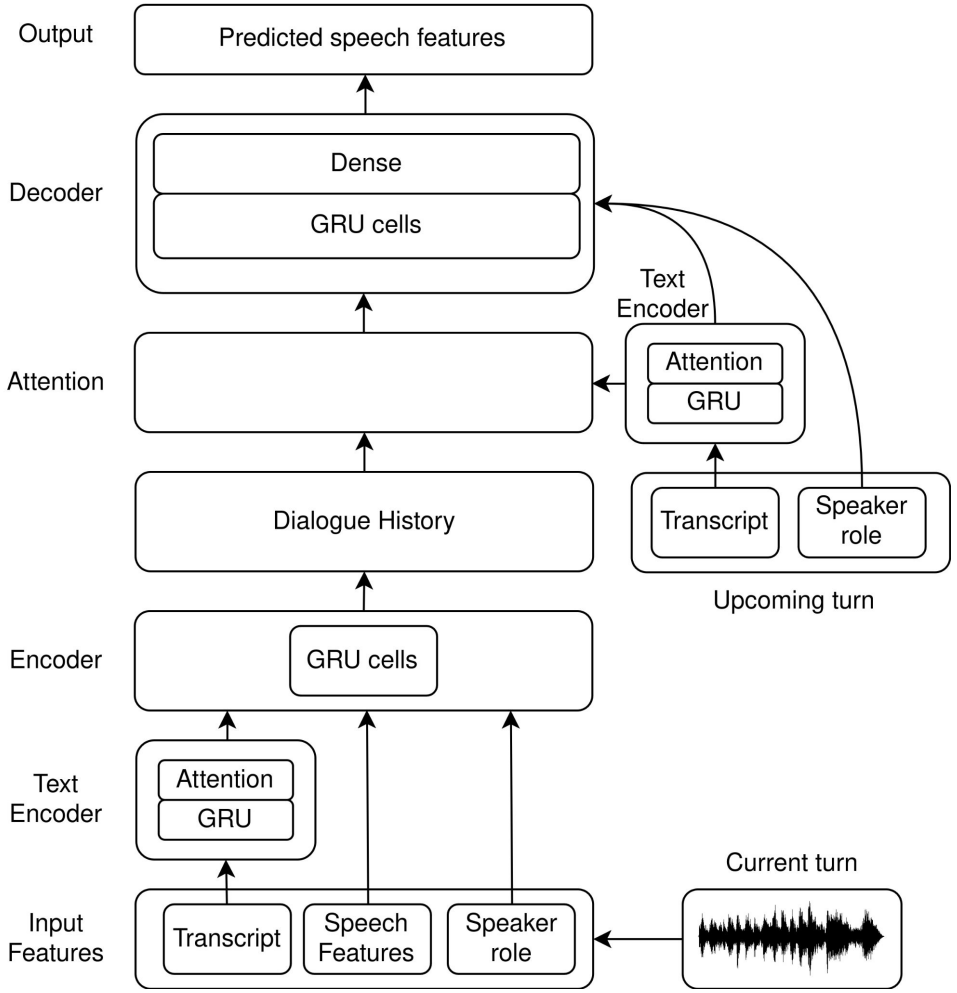


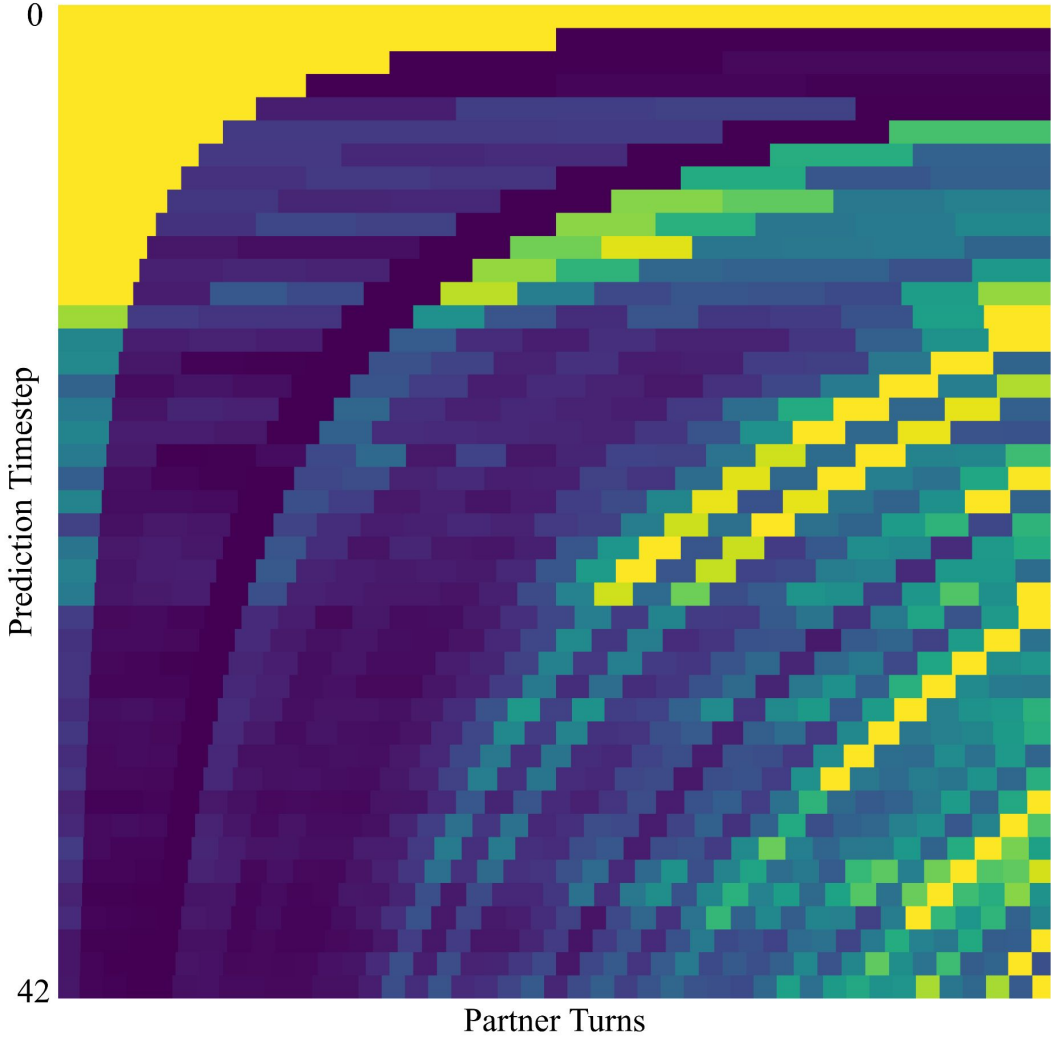
Neural methods

[Lahiri et al (2023)]



Conversational Dynamics Model




0
Prediction Timestep
42
Partner Turns

0
42
Prediction Timestep
Partner Turns
Having a Credit Card
Credit Cards & Cash
Marketing & Preapproval
Money Problems

Turn	Side	Transcript	Anchor
0	**P**	**well um with credit cards is me i i try to keep maybe just one or two i don't i don't like having credit cards for every store i i uh i just don't like them**	
1	A	what are you afraid of with them	0
2	P	what am i afraid of	0
3	A	yes	0
4	P	um i don't know if i'm really afraid of spending too much i just uh don't think that i need them you know i uh they are tempting at times but i i just you know sometimes i just don't like everybody knowing everything about me you know so	0
5	A	well that's very can be very very true	0
6	P	so and there you know everybody just gives you a credit card just so you'll spend money so	0
7	A	well that's the idea they	0
8	P	yeah	0
9	A	figure that if they give you a card at no charge for a whole year that you will use it	0
10	P	uh-huh that's right see they make money off of it whether you use it or not	0
11	A	and absolutely well the other thing of course is the fact that they hope you will not pay your bill at the end of the month so that you will be paying interest	0
12	P	that's right uh-huh	0
13	A	i use mine a great deal um for groceries for everything that i can and then just write one check at	0
14	P	uh-huh	0
15	A	the end of the month	0
16	P	that's what i do	0
17	A	for the entire thing right	0
18	P	that's what i do just a whitney i will i'm on the phone honey sorry	0
19	A	that's all right sounds like you have a little one there	0
20	P	i have two little ones yeah	0
21	A	oh you have	0
22	P	yeah so but	0
23	A	great how old are they	0
24	P	always want some um four and two and a half	0
25	A	oh boy those are two very active	0
26	P	yes uh yeah i'm i'm like you i i use my only use my credit card for um you know when i you know i just use it whenever i feel like i don't want to write a check and then but i don't charge anything that i can't payoff at the end of the month	0
27	A	uh-huh well uh do you uh do you ever use the atm machines	0
28	**P**	**so no i never do my husband does at work just to get cash out but uh i take the checkbook so i you know i just if i need cash i just tell him and he gets it out and i don't even think i know my number**	
29	A	well i find it a great use from the standpoint that you don't have to continue to write checks in order	28
30	P	uh-huh	28
31	A	to get cash	28
32	P	that's right see we he couldn't do without it he but you know since he can just do it right there at work for nothing	28
33	A	oh that helps right well	28
34	P	yeah	28

Conversational Dynamics Model

	Actual		Expected	
	%	Freq	%	Freq
Sta+Opi	0.770	4210	0.550	3009.572
Question	0.100	546	0.068	372.094
Other	0.069	376	0.107	584.904
Bac	0.043	233	0.227	1242.102
Ans+Agr	0.019	104	0.048	260.327

Turn id	Speaker	Transcript	Anchor
106	A	well that's uh that's the way i feel i i feel like i i work uh hard enough and make enough money i i can pay somebody to take care of my cars	
107	B	it it it'll save your fingers trust me	98
108	A	so uh that's right that's right well hey i appreciate the conversation	
109	B	well i enjoyed it jay	108
110	A	alrighty	
111	B	thank you	108
112	A	bye-bye	
113	B	bye-bye	108

An autoregressive conversational dynamics model for dialogue systems

Autoregressive cross-interlocutor attention scores meaningfully capture conversational dynamics

Conclusion

- You must always model the interlocutor!