End-to-end conversational agents: what's missing?

Marco Baroni

Facebook Artificial Intelligence Research

Let's Discuss Workshop at NIPS 2016

With many thanks to…

- Angeliki Lazaridou, Alex Peysakhovich
- Gemma Boleda
- Jason Weston, Marc'Aurelio Ranzato, Sumit Chopra, Antoine Bordes
- Tomas Mikolov, Armand Joulin, Germán Kruszewski
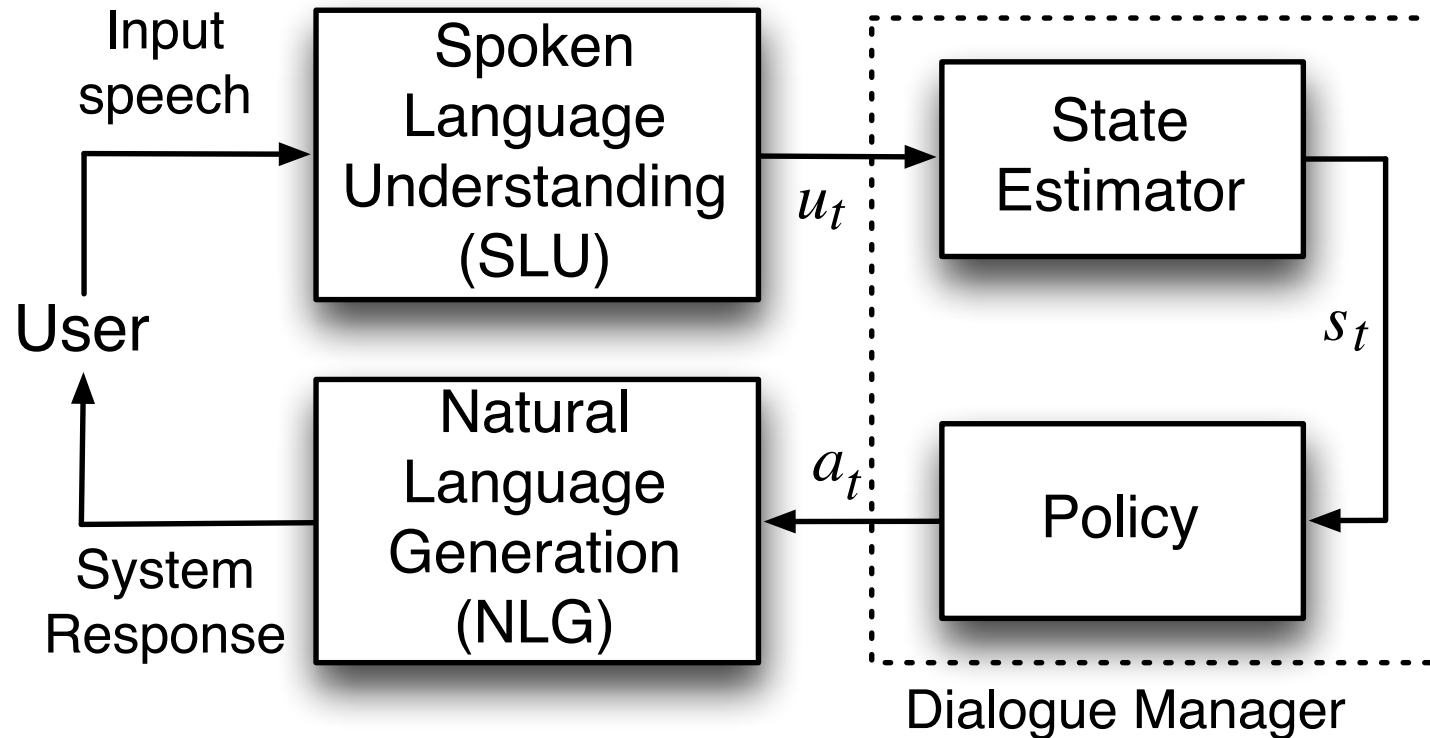- Raquel Fernandez
- Aurelie Herbelot
- …

Why is conversation so easy?
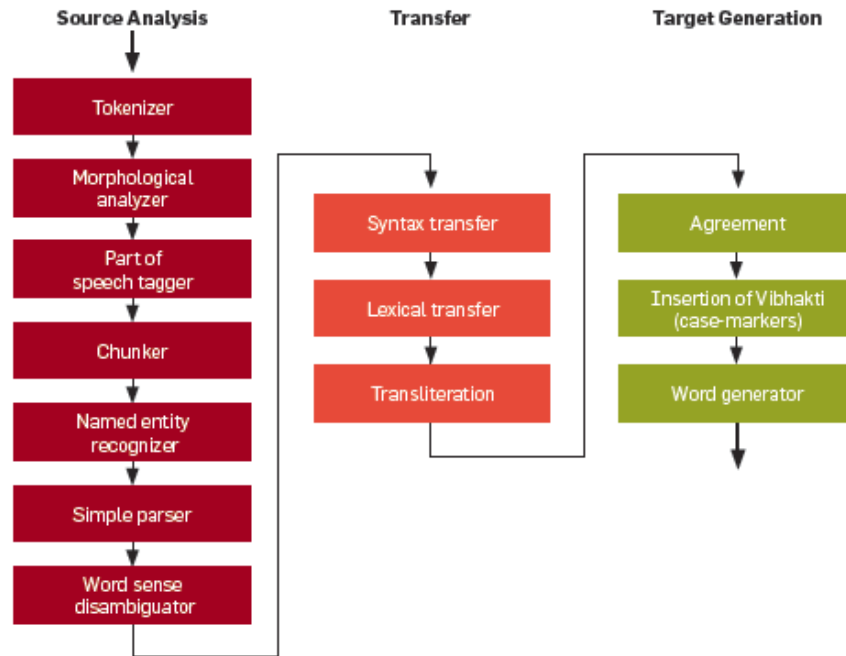
Simon Garrod[1] and Martin J. Pickering[2]

[1]University of Glasgow, Department of Psychology, Glasgow, UK, G12 8QT
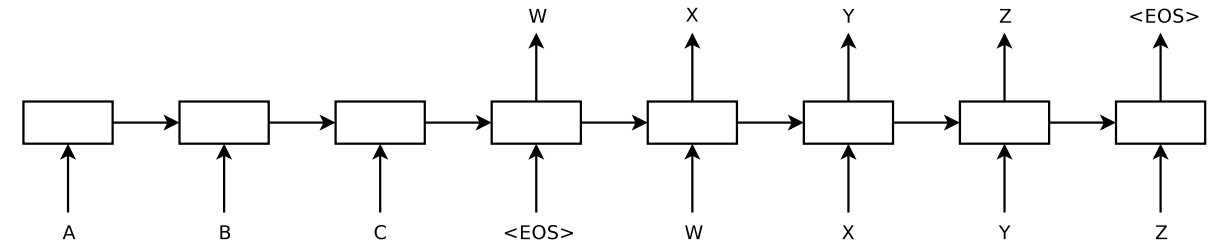[2]University of Edinburgh, Department of Psychology, Edinburgh, UK, EH8 9JZ

The conversational agent pipeline

Young et al. 2013

From pipelines to end-to-end machine translation

Anthes 2010

Sutskever et al. 2014
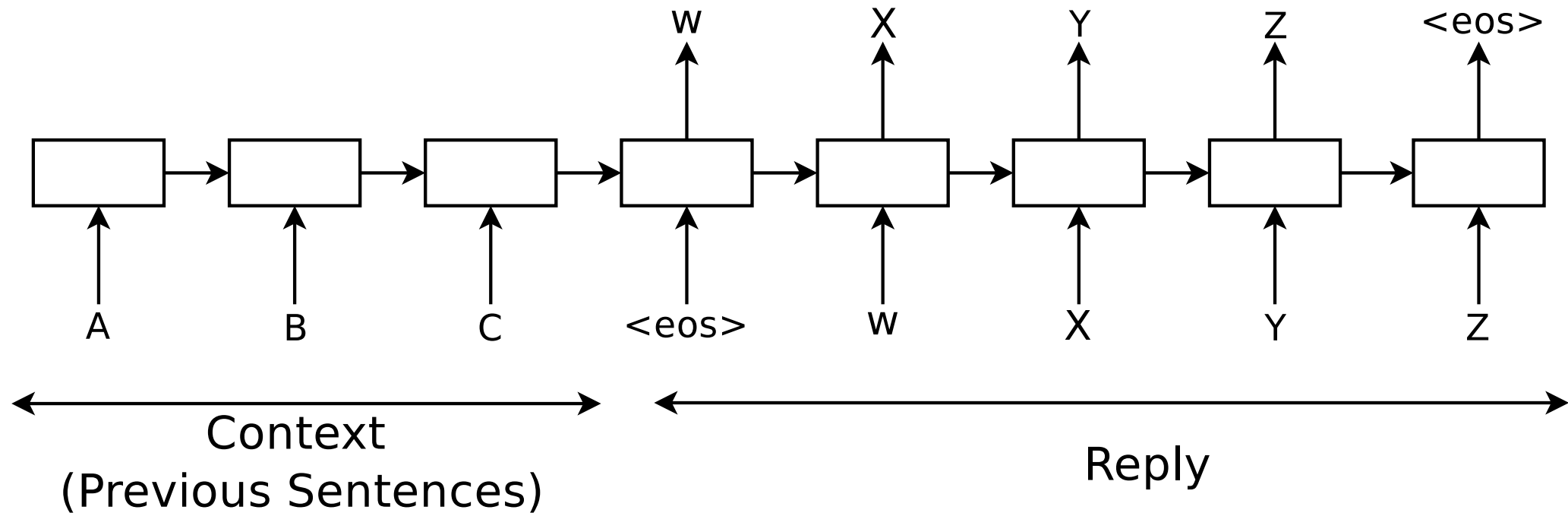
Training end-to-end machine translation

Io sono il dottore di cui in questa novella si parla talvolta con parole poco lusinghiere.

predict

I am the doctor occasionally mentioned in this story, in unflattering terms.

End-to-end conversational agents

Vinyals and Le 2015

Training end-to-end conversational agents

A: Shall we call Alan Turner?

B: Alan Turner?

A: Yeah.

B: I haven't got his number have you got it?

http://bncweb.lancs.ac.uk/

Training end-to-end conversational agents

A: Shall we call Alan Turner?

B: Alan Turner?

A: Yeah.

predict →

B: I haven't got his number have you got it?

End-to-end conversational agents

H: i am seeing an error related to vpn

M: what is the error message that you are getting when connecting to vpn using network connect?

H: connection refused or something like that

M: try this, goto <URL> page and enter your name and password

H: done

M: try logging in with <NAME> and check if it's working fine now

H: yes, now it works!

Vinyals and Le 2015

End-to-end conversational agents

H: : what do you think about messi?

M: he's a great player.

H: what do you think about cleopatra?

M: oh, she's very regal.

H: what do you think about england during the reign of elizabeth?

M: it was a great place.

Vinyals and Le 2015

End-to-end conversational agents

H: what is your job?

M: i'm a lawyer.

H: what do you do?

M: i'm a doctor.

Vinyals and Le 2015

End-to-end conversational agents

H: what are you doing?
M: nothing.

H: what is your name?
M: i don't know, sir.

H: how old are you?
M: i'm not sure.

Li et al. 2016a

Conversation as a joint project
Clark 1996

- **(Extra-linguistic) context:** a conversation will take place in a certain state of the world

- **Common ground:** the participants must coordinate on a (dynamic) set of "mutual knowledge, mutual beliefs, and mutual assumptions"

- **Purpose:** we engage in a conversation in order to achieve something

Context

A: Shall we call Alan Turner?

B: Alan Turner?

A: Yeah.

[B knows Alan Turner number]

B: I haven't got his number have you got it?

Context

A: Shall we call Alan Turner?

B: Alan Turner?

A: Yeah.

[Alan Turner's number is 333 454443]

B: His number is 911

Accessing extra-linguistic context

- Teach machine to query a data-base (Bordes and Weston 2016):

  A: Shall we call Alan Turner?
  B: Alan Turner?
  A: Yeah.
  B: *db_search(Alan_Turner, Has_Number, *])*
  DB: *333 454443*
  B: Alan's number is 333 454443

Accessing extra-linguistic context

Das et al. 2016

Accessing extra-linguistic context: problems and solutions

- Subject-elicited data: small
- Scripted interactions: limited, artificial

- Mix large-scale purely-text-based training and small/artificial real-world-based training (how?)
- Clever ways to collect grounded data (games?)
- Grounded data in the wild (movies?)

How to use contextual information

Contextual vs generic

- What's the color of cows?
- What's the color of this cow?

Precise vs vague

- Where does Alan Turner live?

- 34 Washington Street / 35 Lincoln Avenue / In the US

- In the US / In the United States / In Winesburg, Ohio / 34 Washington Street

Conversation as a joint project
Clark 1996

- **(Extra-linguistic) context:** a conversation will take place in a certain state of the world

- **Common ground:** the participants must coordinate on a (dynamic) set of "mutual knowledge, mutual beliefs, and mutual assumptions"

- **Purpose:** we engage in a conversation in order to achieve something

Common ground
Clark and Brennan 1991

A: Shall we call Alan Turner?

B: Alan Turner?

A: Yeah.

B: I haven't got his number have you got it?

Common ground
Clark and Brennan 1991

A: Shall we call Alan Turner?

B: Alan Turner?

A: Yeah.

B: I haven't got his number have you got it?

Common ground

Leslie: Gotta see them next time.

Leslie : See, I was a bit shocked, you know

Leslie: Mike said he went to Chessington, was it?

Rob: Yeah.

Leslie: Ha ha!

Leslie: Why didn't he do that when we were here?

http://bncweb.lancs.ac.uk/

Common ground in end2end systems

H: i am seeing an error related to vpn

M: what is the error message that you are getting when connecting to vpn using network connect?

H: connection refused or something like that

M: try this, goto <URL> page and enter your name and password

H: done

M: try logging in with <NAME> and check if it's working fine now

H: yes, now it works!

Vinyals and Le 2015

Common ground in end2end systems

H: what is your job?

M: i'm a lawyer.

H: what do you do?

M: i'm a doctor.

Vinyals and Le 2015

Common ground and conversation partners

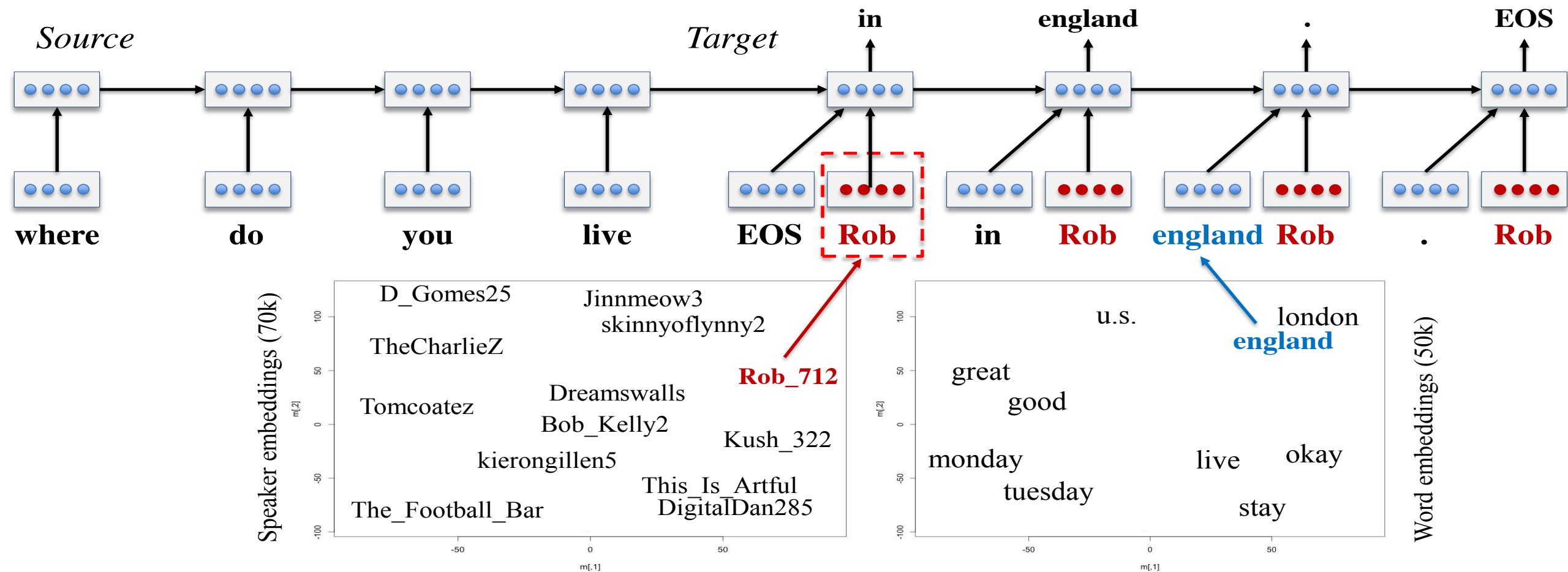Rob: Let me introduce you to my friend Mike...

Machine: Nice to meet you!

...

Machine: Mike said he went to Chessington, was it?

Rob: Yeah.

Machine: I know this person, Mike, who went to Chessington.

Alice: Yeah

Personas
(Li et al. 2016b)

Negotiating common ground

Alan: In fact it was the Society's policy now that we won't put animals to sleep unless there is an extreme cause for that...

James: Why is it that in the north of England, particularly in this region of the north of England, we seem to be worse than anywhere else?

Alan: I don't think you are. You know, since I've been down in Horsham, I have found that the northeast is not alone. There are many other areas in the country where animals are seriously abused. The northeast isn't the only place.

http://bncweb.lancs.ac.uk/

Updating common ground

A: Shall we call Alan Turner?

B: Alan Turner?

A: Yeah.

B: I haven't got his number have you got it?

...

A: Shall we call our... his name isn't Alan Turner.

B: Well what's his name?

A: His name is... Richard... Thorpe, which is why am I looking under T in my book...

Conversation as a joint project
Clark 1996

- **(Extra-linguistic) context:** a conversation will take place in a certain state of the world

- **Common ground:** the participants must coordinate on a (dynamic) set of "mutual knowledge, mutual beliefs, and mutual assumptions"

- **Purpose:** we engage in a conversation in order to achieve something

The purpose of conversations

Task-oriented conversations

A: I'd like to book a restaurant for tonight.
B: What kind of food?
A: Indian, perhaps?
B: How about Shahi Tandoor on George Street?
A: That sounds great, thank you!

Chitchat?

A: Shall we call Alan Turner?
B: Alan Turner?
A: Yeah.
B: I haven't got his number have you got it?

The purpose of conversations

A: Shall we call Alan Turner?
B: Alan Turner?
A: Yeah.
B: I haven't got his number have you got it?

The purpose of conversations

B: Before the end of the show, we've got ten minutes left...

A: Shall we call Alan Turner?
B: Alan Turner?
A: Yeah.
B: I haven't got his number have you got it?

A: No... hey, have you seen they're filming in Harewood this week?

Conversational success is inherently rewarding

Cell
PRESS

Brain-to-brain coupling: a mechanism for creating and sharing a social world

Uri Hasson[1,2], Asif A. Ghazanfar[1,2], Bruno Galantucci[3,4], Simon Garrod[5,6] and Christian Keysers[7,8]

Learning and transferring linguistic cues of success

See also Li et al. 2016c

Explicit reward

A: I'd like to book a restaurant for tonight.
B: What kind of food?
A: Indian, perhaps?
B: How about Shahi Tandoor on George Street?
A: That sounds great, thank you!
+1

Implicit reward

A: Shall we call Alan Turner?
B: Alan Turner?
A: Yeah.
B: Here's his number: 333 454443
A: Great, thank you!

Conversation as a joint project
Clark 1996

- **(Extra-linguistic) context:** a conversation will take place in a certain state of the world

- **Common ground:** the participants must coordinate on a (dynamic) set of "mutual knowledge, mutual beliefs, and mutual assumptions"

- **Purpose:** we engage in a conversation in order to achieve something

A: Shall we call Alan Turner?

B: Well, this is the last slide, we ran out of time again!

- Anthes. 2010. Automated translation of Indian languages. Comm ACM.
- Border and Weston. 2016. Learning end-to-end goal-oriented dialog. arXiv.
- Clark. 1996. Using language. CUP.
- Clark and Brennan. 1991. Grounding in communication. In Perspectives on socially shared cognition.
- Das et al. 2016. Visual dialog. arXiv.
- Garrod and Pickering. 2004. Why is conversation so easy. TICS.
- Hasson et al. 2012. Brain-to-brain coupling: A mechanism for creating and sharing a social world. Trends Cog Sci.
- Li et al 2016a. A diversity-promoting objective function for neural conversation models. NAACL 2016.
- Li et al 2016b. A persona-based neural conversation model. ACL 2016.
- Li et al 2016c. Deep reinforcement learning for dialogue generation. EMNLP 2016.
- Sutskever et al. 2014. Sequence to sequence learning with neural networks. NIPS.
- Vinyals and Le. 2015. A neural conversational model. ICML DLW.
- Young et al. 2013. POMDP-based statistical spoken dialogue systems: A review. Proc IEEE.