

Formalising backchannel relevance spaces

Christine Howes

University of Gothenburg
christine.howes@gu.se

Arash Eshghi

Heriot-Watt University
a.eshghi@hw.ac.uk

- 1 Background
- 2 Modelling feedback in DS
- 3 Conclusions

- 1 Background
- 2 Modelling feedback in DS
- 3 Conclusions

Dialogue...

- A 5143 He did mashed potatoes
J 5144 Mm.
A 5145 cabbage, savoy cabbage, carrots <pause> and he'd cu- cut them like I always cut them cos they were only them little baby carrots so, what I do I slice them down
J 5146 Yeah.
A 5147 you know, down middle like
J 5148 Yeah.
A 5149 into quarters so I do them longer
J 5150 Yeah.
A 5151 and he'd done them like that in microwave for eight minutes <pause> and er, done sprouts <pause> then he'd put this meat pie in oven
J 5152 Crikey!
A 5153 and er, done onion gravy!
J 5154 Mm mm!
A 5155 I says, ooh this gravy's lovely!
J 5156 Yeah!
A 5157 He says er, yeah he said I did some onion, and then, I got some of them, you know
J 5158 Granules?
A 5159 yeah, put some of that in
J 5160 Mm.

Dialogue...

- is incremental
- and co-constructed (Clark, 1996; Goodwin, 1981)
- even if one person does most of the talking (Bavelas et al., 2000)
- listener feedback:
 - backchannels (*mmm, uh-huh*)
 - repair (*what?, huh?*)
 - (also non-verbal, such as nodding)

Backchannels...

- can occur sub-sententially
- evidence suggests that there are specific places where they are salient
- *backchannel relevance spaces* (Heldner et al., 2013)
- analogous to transition relevance places (TRPs; Sacks et al., 1974)
- but more common
- feedback is optional at BRSs

Randomly placed backchannels...

- disrupt the flow of dialogue
- are rated as less natural
- decrease rapport
- make a robot listener seem less attentive

(Poppe et al., 2011; Kawahara et al., 2016;
Park et al., 2017)

Using backchannels...

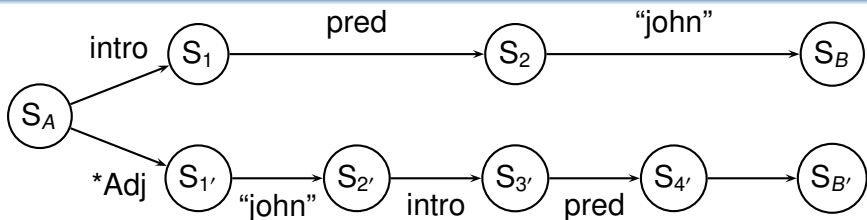
- is crucial for dialogue models
- which may use low-level features (e.g. intonation)
 - these (Cathcart et al., 2003; Gravano and Hirschberg, 2009) sound 'more human' but provide no insight into why feedback occurs where it does
- or incorporate reasoning about the interlocutor's intentions or goals
 - these (Visser et al., 2014; Buschmeier and Kopp, 2013; Wang et al., 2011) presuppose a level of complexity that is unnecessary in natural conversation Gregoromichelaki et al. (2011)

- 1 Background
- 2 Modelling feedback in DS**
- 3 Conclusions

Modelling feedback in DS

- Dynamic Syntax (DS: Kempson et al., 2001; Cann et al., 2005) can provide a formal model of where feedback should be salient
- backchannels are taken to signal (when produced), or trigger (when parsed) the execution of COMPLETION

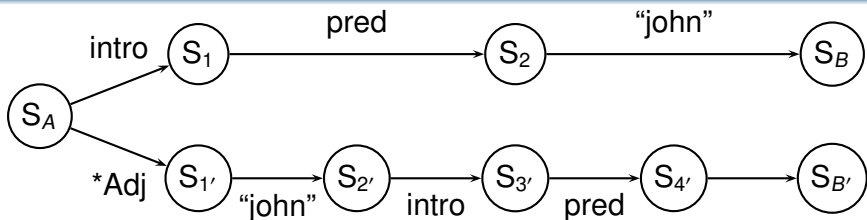
Graph-based parsing and generation



Nodes = Semantic Trees

Edges = Lexical or Computational actions

Graph-based parsing and generation



Nodes = Semantic Trees

Edges = Lexical or Computational actions

- Parsing = incremental search/construction of this Directed Acyclic Graph (DAG) (Sato, 2011)
- Probabilistic best-first parsing definable over the same structure
- Context in DS is this DAG: record of trees and actions so far (Eshghi et al., 2013; Purver et al., 2011)

Modelling feedback in DS

- To model grounding states, we augment the context DAG with two *Coordination Pointers*

Modelling feedback in DS

- To model grounding states, we augment the context DAG with two *Coordination Pointers*
 - The *self-pointer*, \blacklozenge , and
 - The *other-pointer*, \blacklozenge

Modelling feedback in DS

- To model grounding states, we augment the context DAG with two *Coordination Pointers*
 - The *self-pointer*, \blacklozenge , and
 - The *other-pointer*, \blacklozenge
- The intersection of self-pointer-to-root and other-pointer-to-root path is *grounded*

Modelling feedback in DS

- To model grounding states, we augment the context DAG with two *Coordination Pointers*
 - The *self-pointer*, \blacklozenge , and
 - The *other-pointer*, \blacklozenge
- The intersection of self-pointer-to-root and other-pointer-to-root path is *grounded*
- Discursive potential (Ginzburg, 2012) or discourse obligations (Matheson et al., 2000) as pointer divergence

Modelling feedback in DS-TTR

- Enables modelling of contextual updates arising from backchannels, CRs, short answers, or any use of context dependency.

Modelling feedback in DS-TTR

- Enables modelling of contextual updates arising from backchannels, CRs, short answers, or any use of context dependency.
- purely in terms of processing: No recourse to dialogue acts, intentions, or higher order reasoning

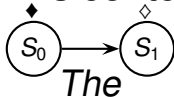
A simple model of backchannels

Dialogue
A: The

Context-final semantics

$$\left[\begin{array}{l} r \quad : \left[\begin{array}{l} x \quad : e \end{array} \right] \\ x_{=t(r,x,r)} \quad : e \end{array} \right]$$

A's context after dialogue



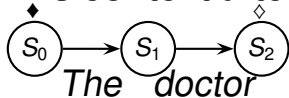
A simple model of backchannels

Dialogue
A: The doctor

Context-final semantics

$$\left[\begin{array}{l} r \quad : \left[\begin{array}{l} x \quad : e \\ p_{=doctor(x)} : t \end{array} \right] \\ x_{=t(r,x,r)} : e \end{array} \right]$$

A's context after dialogue



A simple model of backchannels

Dialogue

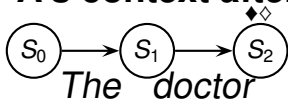
A: The doctor

B: mhm

Context-final semantics

$$\left[\begin{array}{l} r \quad : \left[\begin{array}{l} x \quad : e \\ p_{=doctor(x)} : t \end{array} \right] \\ x_{=t(r,x,r)} : e \end{array} \right]$$

A's context after dialogue



A simple model of backchannels

Dialogue

A: The doctor

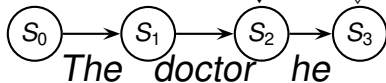
B: mhm

A: he

Context-final semantics

$$\left[\begin{array}{l} r \\ x_{=t(r,x,r)} \end{array} : \left[\begin{array}{l} x : e \\ p_{=doctor(x)} : t \end{array} \right] \right]$$

A's context after dialogue



A simple model of backchannels

Dialogue

A: The doctor

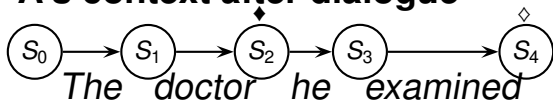
B: mhm

A: he examined

Context-final semantics

$$\left[\begin{array}{l} r \quad : \quad \left[\begin{array}{l} x \quad : e \\ p_{=doctor(x)} : t \end{array} \right] \\ x_{=t(r,x,r)} \quad : e \\ ev_{=examine} \quad : es \\ p_{=subj(ev,x)} \quad : t \end{array} \right]$$

A's context after dialogue



A simple model of backchannels

Dialogue

A: The doctor

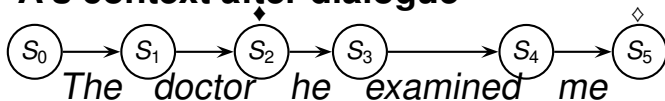
B: mhm

A: he examined
me

Context-final semantics

$$\left[\begin{array}{l} r \quad : \quad \left[\begin{array}{l} x \quad : e \\ p_{=doctor(x)} : t \end{array} \right] \\ x_{=t(r,x,r)} \quad : e \\ ev_{=examine} \quad : es \\ p_{=subj(ev,x)} \quad : t \\ x1_{=spkr} \quad : e \\ p1_{=obj(ev,x1)} \quad : t \end{array} \right]$$

A's context after dialogue



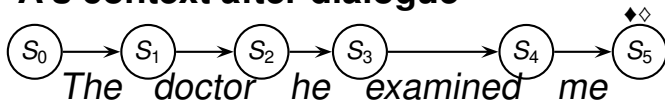
A simple model of backchannels

Dialogue
A: The doctor
B: mhm
A: he examined
me
B: okay

Context-final semantics

$$\left[\begin{array}{l} r \quad : \quad \left[\begin{array}{l} x \quad : \quad e \\ p_{=doctor(x)} : t \end{array} \right] \\ x_{=t(r,x,r)} \quad : \quad e \\ ev_{=examine} \quad : \quad es \\ p_{=subj(ev,x)} \quad : \quad t \\ x1_{=spkr} \quad : \quad e \\ p1_{=obj(ev,x1)} \quad : \quad t \end{array} \right]$$

A's context after dialogue



- But why are backchannels distributed the way they are?

Backchannelling as DS Completion

Lexical Entry for a backchannel:

mhm	IF	$?Ty(X)$	
	THEN	abort	
	ELSE	IF	$\langle \uparrow_0 \downarrow_1 \rangle \exists x. Tn(x)$
			$\langle \uparrow_0 \downarrow_1 \rangle \neg \exists x. ?x$
			$\neg \exists x. ?x$
		THEN	abort
	ELSE	do-nothing	

Precludes “**A:** John arrived with ... **B:** mhm”

Processing Clarification Requests

- Local and non-local CRs
- Extend a semantic tree in context

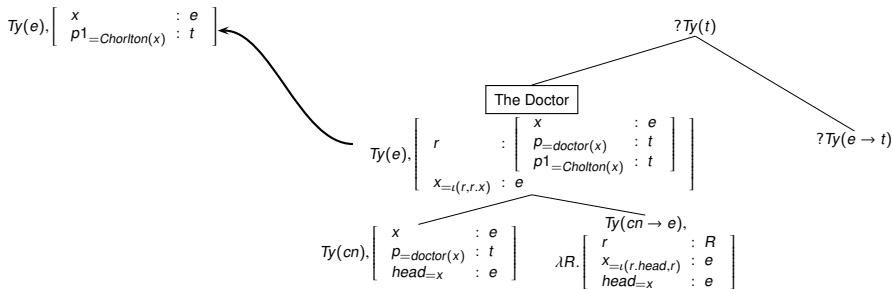
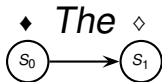


Figure: Processing *Chorlton?* in “A: the doctor B: Chorlton?”

Clarification Interaction in DS

Dialogue

A: The



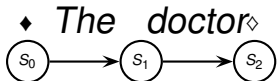
Context-Final Semantics

$$\left[\begin{array}{l} r \quad : \left[\begin{array}{l} x \quad : e \end{array} \right] \\ x = t(r, x, r) \quad : e \end{array} \right]$$

Clarification Interaction in DS

Dialogue

A: The doctor



Context-Final Semantics

$$\left[\begin{array}{l} r \quad : \quad \left[\begin{array}{l} x \quad : e \\ p=\text{doctor}(x) \quad : t \end{array} \right] \\ x=t(r,x,r) \quad : e \end{array} \right]$$

Clarification Interaction in DS

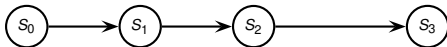
Dialogue

A: The doctor
examined

Context-Final Semantics

$$\left[\begin{array}{l} r \\ x_{=t}(r,x,r) \\ ev_{=examine} \\ p_{=subj}(ev,x) \end{array} : \left[\begin{array}{l} x : e \\ p_{=doctor}(x) : t \end{array} \right] \right]$$

◆ *The doctor examined* ◆



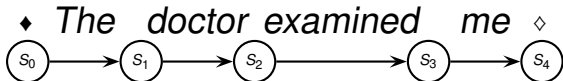
Clarification Interaction in DS

Dialogue

A: The doctor
examined me

Context-Final Semantics

$$\left[\begin{array}{l} r \quad \quad \quad : \left[\begin{array}{l} x \quad \quad \quad : e \\ p=\text{doctor}(x) : t \end{array} \right] \\ x=\iota(r.x,r) \quad : e \\ \text{ev}=\text{examine} \quad : \text{es} \\ p=\text{subj}(\text{ev},x) \quad : t \\ x^1=\text{spkr} \quad \quad : e \\ p^1=\text{obj}(\text{ev},x^1) : t \end{array} \right]$$



Clarification Interaction in DS

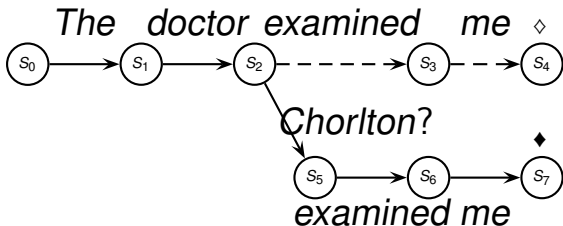
Dialogue

A: The doctor
examined me

B: Chorlton?

Context-Final Semantics

$$\left[\begin{array}{l} r \quad \quad \quad : \left[\begin{array}{l} x \quad \quad \quad : e \\ p = \text{doctor}(x) : t \\ p1 = \text{Chorl}(x) : t \end{array} \right] \\ x =_t(r,x,r) \quad : e \\ ev = \text{examine} \quad : eS \\ p = \text{subj}(ev,x) \quad : t \\ x1 = \text{spkr} \quad \quad : e \\ p1 = \text{obj}(ev,x1) \quad : t \end{array} \right]$$



Clarification Interaction in DS

Dialogue

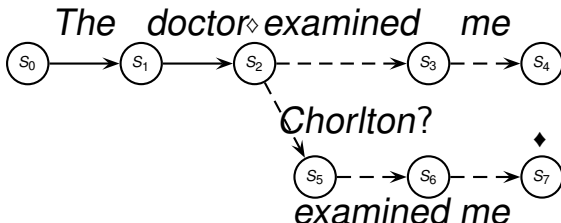
A: The doctor
examined me

B: Chorlton?

A: no,

Context-Final Semantics

$$\left[\begin{array}{l} r \quad \quad \quad : \left[\begin{array}{l} x \quad \quad \quad : e \\ p = \text{doctor}(x) : t \\ p1 \quad \quad \quad : t \end{array} \right] \\ x = t(r,x,r) \quad : e \\ ev = \text{examine} \quad : es \\ p = \text{subj}(ev,x) \quad : t \\ x1 = \text{spkr} \quad \quad : e \\ p1 = \text{obj}(ev,x1) : t \end{array} \right]$$



Clarification Interaction in DS

Dialogue

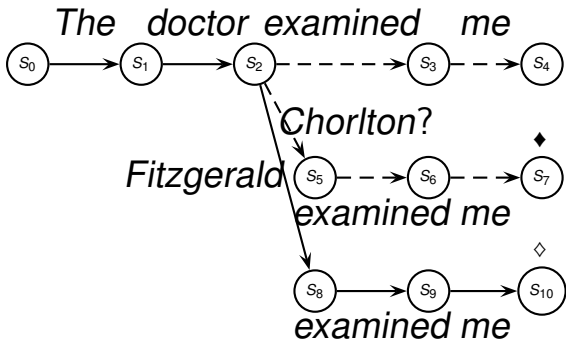
A: The doctor examined me

B: Chorlton?

A: no, Fitzgerald

Context-Final Semantics

$$\left[\begin{array}{l} r \quad \quad \quad : \quad \left[\begin{array}{l} x \quad \quad \quad : e \\ p = \text{doctor}(x) : t \\ p^1 = \text{Fitz}(x) : t \end{array} \right] \\ x =_t(r,x,r) \quad : e \\ \text{ev} = \text{examine} \quad : \text{es} \\ p = \text{subj}(\text{ev},x) \quad : t \\ x^1 = \text{spkr} \quad \quad : e \\ p^1 = \text{obj}(\text{ev},x^1) : t \end{array} \right]$$



Clarification Interaction in DS

Dialogue

A: The doctor
examined me

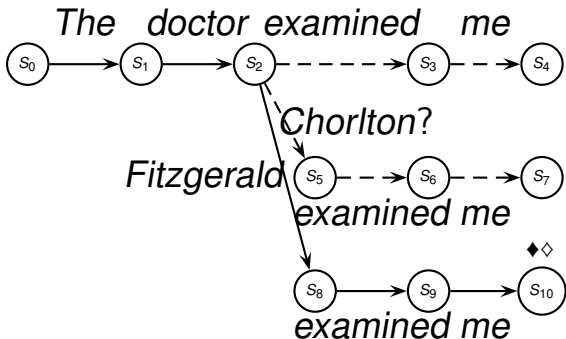
B: Chorlton?

A: no, Fitzgerald

B: uh-huh

Context-Final Semantics

$$\left[\begin{array}{l} r \quad : \quad \left[\begin{array}{l} x \quad : \quad e \\ p = \text{doctor}(x) \quad : \quad t \\ p^1 = \text{Fitz}(x) \quad : \quad t \end{array} \right] \\ x = \iota(r.x.r) \quad : \quad e \\ \text{ev} = \text{examine} \quad : \quad \text{es} \\ p = \text{subj}(\text{ev}.x) \quad : \quad t \\ x^1 = \text{spkr} \quad : \quad e \\ p^1 = \text{obj}(\text{ev}.x^1) \quad : \quad t \end{array} \right]$$



Backchannels and semantic units

- Feedback should come after (and ground) a semantic unit of information
- Parse paths leading to further qualification should be less likely after a backchannel
- And no feedback at such points is interactionally relevant

A: Matt,
B: mmm
A: lives on a boat

A: Matt,
B: ...
A: my brother, ...

Self-backchannels

- If backchannels are purely communicative then self-backchannels make no sense.
- But on our account, backchannels
 - are almost at the level of a reflex
 - indicate that the preceding talk has been integrated
 - even if performed for myself will help my interlocutor to keep track of our potential divergent parses

A: I think he's gonna get it

J: He's servicing it?

A: ready for tonight I think [mm]

J: [Oh] you've serv- oh ready for M O T like?

'Late' feedback

- the listener may be lagging slightly behind the speaker
 - may simply reflect the time taken for the listener to integrate the information into their interpretation so far
 - 'correct' placement of feedback requires prediction
 - analogously to turn-taking timing – people predict upcoming TRPs (de Ruiter et al., 2006)
 - 'Late' feedback can be interpreted as grounding the most recent increment
- i.e. move *other-pointer* to most recent DAG position at which COMPLETION was possible

'Early' feedback

- some feedback seemingly precedes the completion of a semantic unit

J: her mum really she's got a lot on, she'll have a lot on cos she's got to prepare for that wedding, you know what you're like when you, [you've got]

A: [Mm]

J: you know if you want, want to be doing things [don't you get out of house and that]

A: [Yeah, pre- preparing for a wedding, yeah] <pause>
aye

'Early' feedback

- ought to be impossible. . .
- but is expected where the completion is predictable, given the machinery of DS
 - 1 inherent predictability from lexical and computational actions that induce more tree structure with requirements for fixed decorations as well as reuse of actions and
 - 2 the parity between parsing and production

'Early' feedback

- Same thing as completions (strong/weak forms)
- This analysis is supported by the results of an online chat experiment (Howes et al., 2012)
 - artificially truncated turns
 - producing candidate completions as clarifications is a fairly common strategy
 - –when the POS is predictable, and the context is sufficiently constrained

'Early' feedback

- Clarification request completion example:
 - N: i think susie because she is t . . .
 - B: **a woman?**
 - N: ehe least important out of the three if you think about it . . .

'Early' feedback

- Clarification request completion example:
 - N: i think susie because she is t . . .
 - B: **a woman?**
 - N: ehe least important out of the three if you think about it . . .

- but if the continuation is predictable enough:
 - T: its not that fair on the girl doing th . . .
 - H: **exactly, you need to think of others and not be so selfish :P**
 - T: study

- 1 Background
- 2 Modelling feedback in DS
- 3 Conclusions**

Discussion

- Model of feedback at FRPs and non-FRPs
- Backchannels serve to align processing time lines

Discussion

- Model of feedback at FRPs and non-FRPs
- Backchannels serve to align processing time lines
- Other context-dependent forms align interpretation pathways
- Context-dependency at the centre of participant coordination and feedback in dialogue
 - clipping alternative interpretation paths

Discussion

- Model of feedback at FRPs and non-FRPs
- Backchannels serve to align processing time lines
- Other context-dependent forms align interpretation pathways
- Context-dependency at the centre of participant coordination and feedback in dialogue
 - clipping alternative interpretation paths
 - pervasiveness not coincidental. . .
 - . . . and not just about least effort

Future directions

- Integrated model of feedback in dialogue
- Put the horse before the cart!
- Corpus and experimental studies to see if any of this holds weight
- ...including non-verbal feedback

Thanks!

Questions?

References I

- Bavelas, J. B., Coates, L., Johnson, T., et al. (2000). Listeners as co-narrators. Journal of personality and social psychology, 79(6):941–952.
- Buschmeier, H. and Kopp, S. (2013). Co-constructing grounded symbols–feedback and incremental adaptation in human-agent dialogue. KI-Künstliche Intelligenz, 27(2):137–143.
- Cann, R., Kempson, R., and Marten, L. (2005). The Dynamics of Language. Elsevier, Oxford.
- Cathcart, N., Carletta, J., and Klein, E. (2003). A shallow model of backchannel continuers in spoken dialogue. In Proceedings of the tenth EACL conference, pages 51–58. Association for Computational Linguistics.
- Clark, H. H. (1996). Using Language. Cambridge Univ Press.
- de Ruiter, J., Mitterer, H., and Enfield, N. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. Language, 82(3):515–535.
- Eshghi, A., Purver, M., and Hough, J. (2013). Probabilistic induction for an incremental semantic grammar. In Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Long Papers, pages 107–118, Potsdam, Germany. Association for Computational Linguistics.
- Ginzburg, J. (2012). The Interactive Stance: Meaning for Conversation. Oxford University Press.

References II

- Goodwin, C. (1981). Conversational organization: Interaction between speakers and hearers. Academic Press, New York.
- Gravano, A. and Hirschberg, J. (2009). Backchannel-inviting cues in task-oriented dialogue. In INTERSPEECH, pages 1019–22.
- Gregoromichelaki, E., Kempson, R., Purver, M., Mills, G. J., Cann, R., Meyer-Viol, W., and Healey, P. G. T. (2011). Incrementality and intention-recognition in utterance processing. Dialogue and Discourse, 2(1):199–233.
- Heldner, M., Hjalmarsson, A., and Edlund, J. (2013). Backchannel relevance spaces. In Nordic Prosody: Proceedings of XIth Conference, Tartu 2012, pages 137–146.
- Howes, C., Healey, P. G. T., Purver, M., and Eshghi, A. (2012). Finishing each other's ... responding to incomplete contributions in dialogue. In Proceedings of the 34th Annual Meeting of the Cognitive Science Society (CogSci 2012), pages 479–484.
- Kawahara, T., Yamaguchi, T., Inoue, K., Takanashi, K., and Ward, N. (2016). Prediction and generation of backchannel form for attentive listening systems. In Proc. INTERSPEECH, volume 2016.
- Kempson, R., Meyer-Viol, W., and Gabbay, D. (2001). Dynamic Syntax: The Flow of Language Understanding. Blackwell.
- Park, H. W., Gelsomini, M., Lee, J. J., and Breazeal, C. (2017). Telling stories to robots: The effect of backchanneling on a child's storytelling. In Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction, pages 100–108. ACM.

References III

- Poppe, R., Truong, K. P., and Heylen, D. (2011). Backchannels: Quantity, type and timing matters. In International Workshop on Intelligent Virtual Agents, pages 228–239. Springer.
- Purver, M., Eshghi, A., and Hough, J. (2011). Incremental semantic construction in a dialogue system. In Bos, J. and Pulman, S., editors, Proceedings of the 9th International Conference on Computational Semantics, pages 365–369, Oxford, UK.
- Sacks, H., Schegloff, E., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. Language, 50(4):696–735.
- Sato, Y. (2011). Local ambiguity, search strategies and parsing in Dynamic Syntax. In Gregoromichelaki, E., Kempson, R., and Howes, C., editors, The Dynamics of Lexical Interfaces. CSLI Publications.
- Visser, T., Traum, D., DeVault, D., and op den Akker, R. (2014). A model for incremental grounding in spoken dialogue systems. Journal on Multimodal User Interfaces, 8(1):61–73.
- Wang, Z., Lee, J., and Marsella, S. (2011). Towards more comprehensive listening behavior: beyond the bobble head. In Intelligent Virtual Agents, pages 216–227. Springer.