A Bottom-Up Exploration of the Dimensions of Dialog State in Spoken Interaction

SIGdial 2012

Nigel G. Ward and Alejandro Vega

Models of dialog state are important, both scientifically and practically, but today's best build strongly on tradition. This paper presents a new way to identify the important dimensions of dialog state, more bottom-up and empirical than previous approaches. Specifically, we applied Principal Component Analysis to a large number of low-level prosodic features to find the most important dimensions of variation. The top 20 out of 76 dimensions accounted for 81% of the variance, and each of these dimensions clearly related to dialog states and activities, including turn taking, topic structure, grounding, empathy, cognitive processes, attitude and rhetorical structure.

The table summarizes interpretations of some of the top dimensions, with the variance explained by each, combining information from this paper and two others: "Towards Empirical Dialog-State Modeling and its Use in Language Modeling", Nigel G. Ward and Alejandro Vega, Interspeech 2012; and "Where in Dialog Space does *Uh-huh* Occur?" Nigel G. Ward, David G. Novick, Alejandro Vega, Interdisciplinary Workshop on Feedback Behaviors in Dialog, 2012, submitted; and unpublished findings.

#	Dimension	Variance
1	this speaker talking *vs.* other speaker talking	32%
2	neither speaking *vs.* both speaking	9%
3	topic closing *vs.* topic continuation	8%
4	grounding *vs.* grounded	6%
5	turn grab *vs.* turn yield	3%
6	seeking empathy *vs.* expressing empathy	3%
7	floor conflict *vs.* floor sharing	3%
8	dragging out a turn *vs.* ending confidently and crisply	3%
9	topic exhaustion *vs.* topic interest	2%
10	lexical access or memory retrieval *vs.* disengaging	2%
11	low content and low confidence *vs.* quickness	1%
12	claiming the floor *vs.* releasing the floor	1%
13	starting a contrasting statement *vs.* starting a restatement	1%
14	rambling *vs.* placing emphasis	1%
15	speaking before ready *vs.* presenting held-back information	1%
16	humorous *vs.* regrettable	1%
17	new perspective *vs.* elaborating current feeling	1%
18	seeking sympathy *vs.* expressing sympathy	1%
19	solicitous *vs.* controlling	1%
20	calm emphasis *vs.* provocativeness	1%
21	mitigating a potential face threat *vs.* agreeing, with humor	< 1%
22	personal stories/opinions *vs.* impersonal explanatory talk	< 1%
23	closing out a topic *vs.* starting or renewing a topic	< 1%
24	agreeing and preparing to move on *vs.* jointly focusing	< 1%
25	personal experience *vs.* second-hand opinion	< 1%
26	signalling interestingness *vs.* downplaying the current information	< 1%
29	no emphasis *vs.* lexical stress	< 1%
30	saying something predictable *vs.* pre-starting a new tack	< 1%
37	mid-utterance words *vs.* sing-song adjacency-pair start	< 1%
62	explaining/excusing oneself *vs.* blaming someone/something	< 1%
72	speaking awkwardly *vs.* speaking with a nicely cadenced delivery	< 1%