From Reaction to Prediction
Experiments with Computational Models of Turn-Taking

David Schlangen

Potsdam / Berlin Conversation Studies Lab

Department of Linguistics, University of Potsdam

das@ling.uni-potsdam.de

Background

- ♣ Question: What distinguishes places in a conversation where a transition is possible from those where it isn't?
- ♣ More specifically: What distinguishes them in terms of syntax and prosody?
- ♣ Method: quantitative analysis; ML; how well can we predict turn-taking decisions?
- ♣ Context: P/B-CSL devises and implements models of conversational competence.

Experiments

"is this *x* turn-final or not?"

- ‣ E1: utterances; human subj
- ‣ E2: utterances; ML
- ‣ E3: utterances, real data; ML
- ‣ E4: inter-pausal units; ML
- ‣ E5: all words; ML

Experiments

"is this *x* turn-final or not?"

- ‣ E1: utterances; human subj
- ‣ E2: utterances; ML
- ‣ E3: utterances, real data; ML
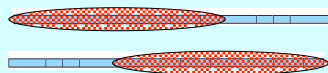- ‣ E4: inter-pausal units; ML
- ‣ E5: all words; ML

Experiments

"is this *x* turn-final or not?"

- ▾ E1: utterances; human subj

    data: created conversational situation where same (declarative) sentence would appear both turn-medial and turn-final. 8 sents; 3 spkrs. In German.

    (cf. (Cutler & Pearson 1986))

- ‣ E2: utterances; ML
- ‣ E3: utterances, real data; ML

Experiments

"is this *x* turn-final or not?"

- ▾ E1: utterances; human subj

    data: 8 sentences, all recorded in turn-medial and turn-final position.

    classifier: 24 university students

- ‣ E2: utterances; ML
- ‣ E3: utterances, real data; ML

Slide 7

Experiments

"is this *x* turn-final or not?"

- ▾ E1: utterances; human subj

  data: 8 sentences, all recorded in turn-medial and turn-final position.

  classifier: 24 university students

  question: will the speaker continue after this sentence or not? (continue y/n = wait/take)

- ▸ E2: utterances; ML
- ▸ E3: utterances, real data; ML

Slide 8

Experiments

"is this *x* turn-final or not?"

- ▾ E1: utterances; human subj

  data: 8 sentences, all recorded in turn-medial and turn-final position.

  classifier: 24 university students

  question: will the speaker continue after this sentence or not? (continue y/n = wait/take)

  results:

class	cor.	incor.	f-m
overall	56%	44%	–
wait	78%	22%	0.63
take	34%	66%	0.38

- ▸ E2: utterances; ML
- ▸ E3: utterances, real data; ML

Slide 9

Experiments

"is this *x* turn-final or not?"

- ▾ E1: utterances; human subj

  results:

class	cor.	incor.	f-m
overall	56%	44%	–
wait	78%	22%	0.63
take	34%	66%	0.38

  (Cutler & Pearson 1986) < us < (Barkhuysen *et al.* 2006)

- ▸ E2: utterances; ML
- ▸ E3: utterances, real data; ML

Slide 10

Experiments

"is this *x* turn-final or not?"

- ▸ E1: utterances; human subj
- ▸ E2: utterances; ML
- ▸ E3: utterances, real data; ML
- ▸ E4: inter-pausal units; ML
- ▸ E5: all words; ML

Slide 11

Experiments

"is this *x* turn-final or not?"

- ▸ E1: utterances; human subj
- ▸ E2: utterances; ML
- ▸ E3: utterances, real data; ML
- ▸ E4: inter-pausal units; ML
- ▸ E5: all words; ML

Slide 12

Experiments

"is this *x* turn-final or not?"

- ▾ E2: utterances; ML

  data: same

  classifier: machine learners

- ▸ E3: utterances, real data; ML
- ▸ E4: inter-pausal units; ML
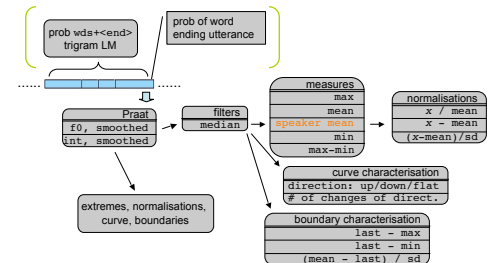
- Machine Classifiers

All ML experiments conducted with WEKA toolkit (Witten & Frank 2005)

Input represented by set of syntactic & prosodic features, computed for each word.

Evaluated using 10-fold cross-validation.

13

- Extracted Features

(cf. (Shriberg *et al.* 2000))

14

Experiments

"is this *x* turn-final or not?"

▾ E2: utterances; ML

data: pairs of recordings of sentence uttered in turn-medial and turn-final position, 8 sents.
classifier: machine learners

▸ E3: utterances, real data; ML
▸ E4: inter-pausal units; ML

15

Experiments

"is this *x* turn-final or not?"

▾ E2: utterances; ML

results:

Clsf.	Fl.	FSet	C%	IC%	Cls	F
JRip	–	fcdr,fml1,ipxm	**68.0** (w-f: 0.68)	**31.9**	wt tk	**0.72** **0.63**
JRip	–	f0 only	65.2	34.8	wt tk	0.69 0.59
JRip	–	int only	60.8	39.2	wt tk	0.62 0.59

▸ E3: utterances, real data; ML
▸ E4: inter-pausal units; ML

16

Experiments

"is this *x* turn-final or not?"

▸ E1: utterances; human subj
▸ E2: utterances; ML
▸ E3: utterances, real data; ML
▸ E4: inter-pausal units; ML
▸ E5: all words; ML

17

Experiments

"is this *x* turn-final or not?"

▸ E1: utterances; human subj
▸ E2: utterances; ML
▸ E3: utterances, real data; ML
▸ E4: inter-pausal units; ML
▸ E5: all words; ML

18

Slide 19

Experiments

"is this *x* turn-final or not?"

▾ E3: utterances, real data; ML

data: 100 dialogues from switchboard corpus; utterances (slash units)

classifier: machine learners

▸ E4: inter-pausal units; ML

▸ E5: all words; ML

19

Slide 20

Experiments

"is this *x* turn-final or not?"

▾ E3: utterances, real data; ML

results:

Clsf.	Fl.	FSet	C%	IC%	Cls	P	R	F
Maj.	–	full	69.3	30.7	wt	0.69	1	0.82
			(w-f: 0.57)		tk	0	0	0
BayN	pki	full	64.6	35.4	wt	0.76	0.72	
			(w-f: **0.65**)		tk	0.43	0.48	
BayN	pki	int only	61.2	38.8	wt	0.76	0.64	
			(w-f: 0.62)		tk	0.40	0.55	
BayN	pki	f0 only	61.4	38.6	wt	0.72	0.72	
			(w-f: 0.61)		tk	0.37	0.37	

▸ E4: inter-pausal units; ML

▸ E5: all words; ML

20

Slide 21

Experiments

"is this *x* turn-final or not?"

▸ E1: utterances; human subj

▸ E2: utterances; ML

▸ E3: utterances, real data; ML

▸ E4: inter-pausal units; ML

▸ E5: all words; ML

21

Slide 22

Experiments

"is this *x* turn-final or not?"

▸ E1: utterances; human subj

▸ E2: utterances; ML

▸ E3: utterances, real data; ML

▸ E4: inter-pausal units; ML

▸ E5: all words; ML

22

Slide 23

Experiments

"is this *x* turn-final or not?"

▸ E3: utterances, real data; ML

▾ E4: inter-pausal units; ML

data: swbd; units delineated by pauses (varying the minimal length)

200ms

▸ E5: all words; ML

23

Slide 24

Experiments

"is this *x* turn-final or not?"

▸ E3: utterances, real data; ML

▾ E4: inter-pausal units; ML

data: swbd; units delineated by pauses (varying the minimal length)

200ms

task: `take` v `wait` (≈ turn-end sil. v hesitation)

classifier: machine learners

(cf. e.g. (Ferrer *et al.* 2002, 2003))

▸ E5: all words; ML

24

Slide 25

Experiments

"is this *x* turn-final or not?"

- E3: utterances, real data; ML
- E4: inter-pausal units; ML

results:

PTR	clsf	f-w	f-t	FAR	
.500s	bsln	0	62.8	54.2	
	J48	71.8	65.3	33.2	$\Delta = 38.7\%$
.250s	blsn	0	50.3	66.3	
	J48	76.7	50.5	46.9	$\Delta = 29.3\%$
.100s	bsln	0	41.7	73.6	
	J48	82.3	44.1	51.3	$\Delta = 30.3\%$
0s	bsln	n/a	n/a	n/a	
	J48	97.6	35.5	41.6	(see Tab. 6)

- E5: all words; ML

Slide 26

Experiments

"is this *x* turn-final or not?"

- E3: utterances, real data; ML
- E4: inter-pausal units; ML
- E5: all words; ML

Slide 27

Experiments

"is this *x* turn-final or not?"

- E3: utterances, real data; ML
- E4: inter-pausal units; ML
- E5: all words; ML

Slide 28

Experiments

"is this *x* turn-final or not?"

- E3: utterances, real data; ML
- E4: inter-pausal units; ML
- E5: all words; ML

data: all words in switchboard
task: is (the utterance up to) this *word* turn-final or not?

Slide 29

Experiments

"is this *x* turn-final or not?"

- E3: utterances, real data; ML
- E4: inter-pausal units; ML
- E5: all words; ML

results:

Clsf.	Fl.	FSet	C%	IC%	Cls	F
Maj.	–	full	95.6	4.37	wt	0.98
(κ: 0; TT_ε: undef.)			(w-f: 0.93)		tk	0
J48	–	full	95.4	4.6	wt	0.98
(κ: 33; TT_ε: 4.81)			(w-f: 0.95)		tk	0.36
J48	–	syn.	95.8	4.2	wt	0.98
(κ: 24; TT_ε: 6.53)			(w-f: 0.95)		tk	0.26
J48	–	ac.	95.4	4.6	wt	0.98
(κ: 6; TT_ε: 31)			(w-f: 0.94)		tk	0.07

Slide 30

Summary / Conclusions

"is this *x* turn-final or not?"

- E1: utterances; human subj
  - better than chance; better at `wait`
- E2: utterances; ML
  - better at `wait`; int&f0 same for `take`; f0 for `wait`
- E3: utterances, real data; ML
  - perf. worse; same int/f0 pattern
- E4: inter-pausal units; ML
  - prosodic info improves over baseline
- E5: all words; ML
  - hard task; main predictive power: syntax

Conclusions

♣ Question was: *what* distinguishes.

♣ only some way towards answering this, more: are they distinguished?

♣ E5 task is of course a poor approximation of real task of *predicting* TRPs: this is not done on a word-by-word basis.

31

Future Work

♣ incremental parsing for prediction; realisation of filter model; prosody has veto

♣ move further backwards: predict end in x ms

♣ fixed window rather than word; using ASR results, etc.

32

Thank you
for your attention!

Acknowledgements:
thanks to the
Potsdam Dialogue Group for discussion!

Diagrams produced with Zeitwort:
http://www.sourceforge.net/projects/zeitwort

<-- this is the Berlin TV tower, in case you were wondering. Shown for no particular reason.

33

Corpus for E1 & E2

Instructions to speakers: "There are two sets of cards, one with situations described in sentences, the other with pictures of these situations. Your task is to read out the descriptions so that B [a confederate] can identify the cards."

Instruction to B: give backchannels, some task related chat.

34