

SPECIALIZED TOPIC PRESENTATION: SENTIMENT AND SUBJECTIVITY

Xiaosu Xue

The research question



- identify when something subjective is being said
- recognize the type of subjective content



Annotation schemes

looking closely at the problem

MPQA annotation scheme



- Key concept: private state
 - ▣ any internal or emotional state
 - ▣ described based on its functional components

- Annotation scheme
 - ▣ represented as frames
 - ▣ frames have slots for attributes and properties

Examples of frames



(4) Democrats also have doubts about Miers' suitability for the high court.

(5) Miers' nomination was criticized from people all over the political spectrum.

(6) "She [Miers] will be a breath of fresh air for the Supreme Court," LaBoon said.

(7) This the nomination of Miers is a missed opportunity of historic proportions.

(8) White House spokesman Jim Dyke said Miers' confirmation hearings are set to begin Nov. 7.

Adaptation of the MPQA scheme



- identify subjective questions
- no need to represent nested sources
- annotate at utterance level

Subjective utterances

- “a span of words (or possibly sounds) where a *private state* is being expressed, either through choice of words or prosody”

(11) Um ⟨**POS-SUBJ** it’s very easy to use⟩. Um ⟨**NEG-SUBJ** but unfortunately it does lack the advanced functions⟩ ⟨**POS-SUBJ** which I I quite like having on the controls⟩.

(13) Um ⟨**UNCERT** I’m not entirely sure what the corporate colour is⟩.

(14) ⟨**OTHER-SUBJ** I think one factor would be production cost⟩.

Objective polar utterances



- positive or negative factual information without expressing a private state

Subjective questions



- elicit the private state of the person being asked
- three types: positive, negative, general

(16) Do you like the large buttons?

(17) What do you think about the large buttons?

Sources and targets

- marked only on the subjective utterances and the objective polar utterances

(18) ⟨**NEG-SUBJ SOURCE=SPEAKER** Finding them is really a pain, you know⟩.

(22) Shall we sh well ⟨**POS-SUBJ SOURCE=SPEAKER TARGET=MEETING** we'll stick to kind of your area for now⟩.

Overlapping annotations

- the speaker expresses a private state about someone else's private state

(25) ⟨**OTHER-SUBJ SOURCE=SPEAKER TARGET=REMOTE DESIGN** I think a recurring theme here is ⟨**POS-SUBJ SOURCE=SPECIFIC EXTERNAL TARGET=REMOTE DESIGN** the company wants it to be [disfmarker] wants us to make something that's fashionable and sleek and trendy ⟩ ⟩

Evaluation

	Kappa	% Agreement
Subjective Utterances (excluding fragments)	0.56	79
Positive Subjective	0.58	84
Negative Subjective	0.62	92
Positive Subjective + Positive Objective	0.58	83
Negative Subjective + Negative Objective	0.68	93
Subjective Question	0.56	95

Table 4: Interannotator agreement for the AMIDA subjectivity annotations

Subjectivity and Polarity Classification

work with the data

Goal



- recognize subjectivity in general and distinguish between positive and negative subjective utterances

Data



- dialogue act segments of AMI corpus
- for subjectivity classification: segments overlapping with subjective utterances or subjective questions
- for pos/neg classification: segments overlapping with positive or negative subjective utterances

Features



- prosody
- word n-grams
- character n-grams
- phoneme n-grams

- individual and combined

Results

Table 4: Results Task 1: Subjective vs. Non-Subjective.

	PROS	WORDS	CHARS	PHONES	F ₁	PREC	REC	ACC
BASE-SUBJ	always chooses subjective class				60.3	43.4	100	43.4
BASE-RAND	randomly chooses a class based on priors				41.8	42.9	41.3	50.6
single	•				54.6	55.3	54.5	63.1
		•			60.5	68.5	54.5	71.0
			•		61.7	67.5	57.2	71.1
				•	60.3	66.4	55.5	70.2
double	•	•			63.9	72.1	57.6	73.4
	•		•		65.6	71.9	60.3	74.0
	•			•	64.6	72.3	58.4	73.7
		•	•		66.2	73.8	60.1	74.9
		•		•	65.2	73.2	58.8	74.3
			•	•	66.1	72.8	60.7	74.5
triple	•	•	•		66.5	74.3	60.3	75.1
	•	•		•	65.5	73.5	59.0	74.5
	•		•	•	66.5	73.3	60.8	74.8
		•	•	•	66.9	74.3	60.9	75.3
quartet	•	•	•	•	67.1	74.5	61.2	75.4

Results 2

Table 5: Results Task 2: Positive Subjective vs. Negative Subjective.

	PROS	WORDS	CHARS	PHONES	F ₁	PREC	REC	ACC
BASE-POS-SUBJ	always chooses positive subjective class				85.6	75.0	100	75.0
BASE-RAND	randomly chooses a class based on priors				75.1	74.4	76.1	62.4
single	•				84.8	74.8	98.1	73.9
		•			85.6	79.6	93.1	76.8
			•		85.9	81.9	90.5	78.0
				•	85.5	80.5	91.3	77.0
double	•	•			88.7	83.0	95.4	81.9
	•		•		88.7	83.1	95.1	81.8
	•			•	88.5	83.3	94.4	81.6
		•	•		89.5	84.2	95.7	83.3
		•		•	89.2	83.7	95.5	82.8
			•	•	89.0	84.2	94.6	82.6
triple	•	•	•		89.6	84.0	96.1	83.4
	•	•		•	89.3	83.6	95.8	82.8
	•		•	•	89.2	83.7	95.5	82.7
		•	•	•	89.8	84.4	96.0	83.8
quartet	•	•	•	•	89.9	84.4	96.2	83.8

Conclusion



- Combined features yield the best results
- Prosody seems to be the least informative
- Character n-grams seem to perform the best



Sentiment Analysis

with prosodic features

Data



- elicited short spoken reviews from 84 participants
 - ▣ nine questions asked, but only the final one, the short review, is included in the dataset
- 52 positive and 32 negative
 - ▣ mixed reviews -> negative
 - ▣ overall ranking of 4 or 5 out of 5 -> positive
 - ▣ overall ranking below 4 -> negative

Data 2



- for text-based classification:
 - ▣ subjects read a review online, write down a short summary, and indicate the overall sentiment; only reviews originally rated under 2 or above 4 were presented
 - ▣ 3268 textual review summaries: 1055 negative, 1600 positive, 613 mixed

Text-based classification baseline



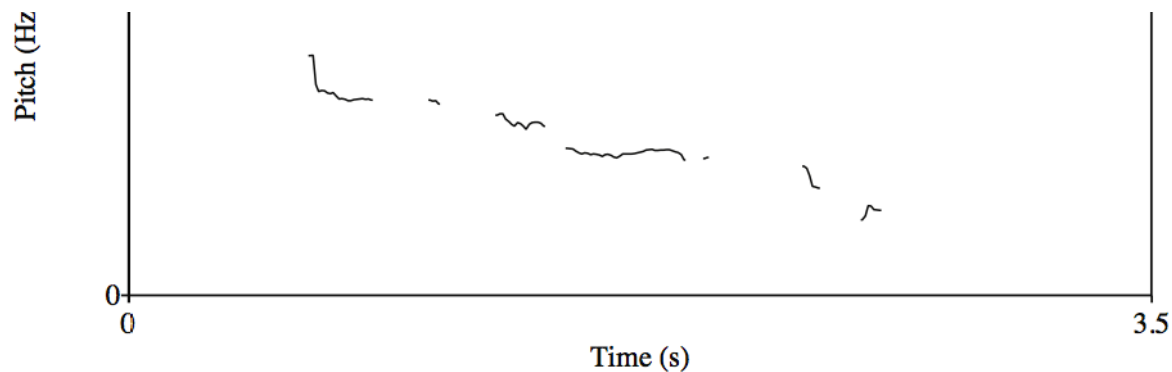
- trained an SVM classifier on the full corpus of 3268 textual review summaries
- feature: n-grams ($n=1,2,3$)

Speech recognition

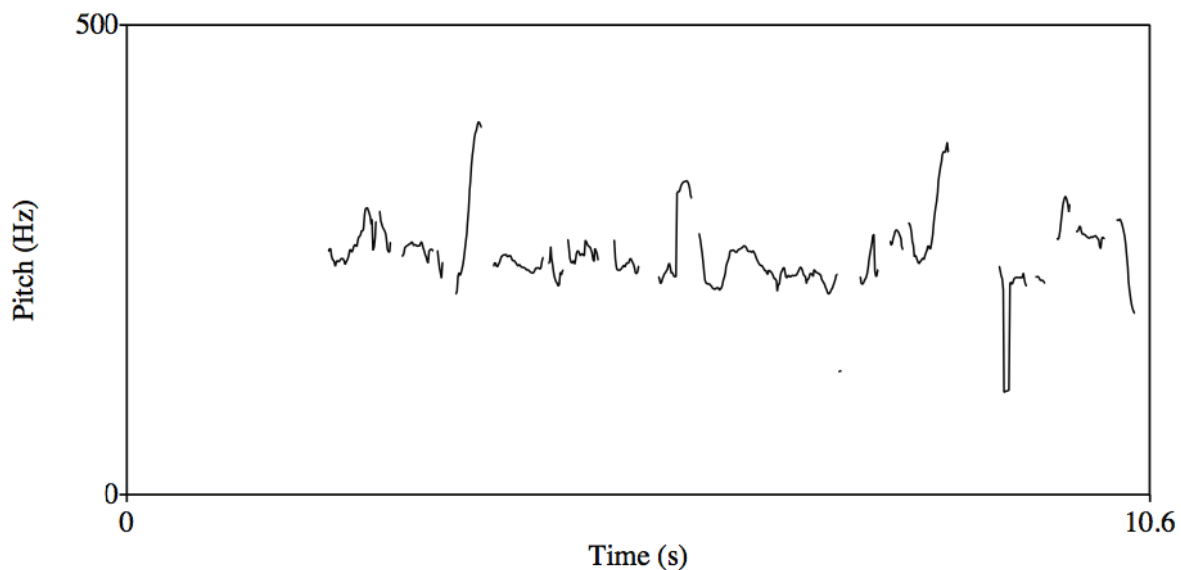


- ASR language model trained on data mined from review websites
- word accuracy: 56.8%
 - ▣ most mistakes are due to out of vocabulary proper names

Acoustic features



(a) *'It's a nice restaurant but a little disappointing.'*



(b) *'Had a great time at Tapeo trying their authentic Spanish tapas, I really enjoyed the goat cheese entrees and had a great time with friends!'*



Results



Feature combination	ASR	Trans.
1. Majority class baseline	61.9	61.9
2. Text prediction only/no acoustic features	75.0	84.4
3. Automatically selected acoustic features	68.9	77.8
4. F0 features only	72.6	81.0
5. Automatically selected F0 features only	82.5	81.0

Conclusion



- Features characterizing F0 are informative enough to significantly outperform a majority class baseline without using any textual information
- If the utterance's text is known, prosodic features confuse the classifier
- If only ASR hypothesis is known, prosody improves performance over a solely text-based model



Finally...

What I have learned



- Possible features for subjectivity and polarity classification of spoken language data
- The motivation for research on sentiment and subjectivity in spoken language data
- Study of annotation schemes helps dissect a problem and facilitates inter-research comparison
- Different ways of collecting and selecting data and the possible effect on the results

Questions for discussion



- Difference between multi-party conversations and short spoken reviews: is prosody more informative in a spoken review?
- From text to speech: what are the challenges/ advantages in the task of subjectivity detection or sentiment analysis?