

EMOTIVE OR NON-EMOTIVE: THAT IS THE QUESTION

Michał Ptaszynski

Fumito Masui

Rafał Rzepka

Kenji Araki

Kitami Institute of Technology

Hokkaido University

PRESENTATION OUTLINE

1. Problem definition
2. Language Combinatorics
3. Experiment setup
4. Results and discussion
5. Conclusions and Future work

PROBLEM DEFINITION

MAINSTREAM:

POSITIVE VS. NEGATIVE

EMOTION TYPES

DISREGARDED OR AS SUBTASK:

IS THE SENTENCE EMOTIONAL / NEUTRAL (PRESELECTION)

PROBLEM DEFINITION

emotional

neutral

“Was the speaker in an emotional state?”

PROBLEM DEFINITION

“Non-emotional” would be probably better.

emotional

neutral

“Was the speaker in an emotional state?”

Easy to ask laypeople because everyone thinks they are specialists in their own emotions.

ambiguous

PROBLEM DEFINITION

emotional

neutral

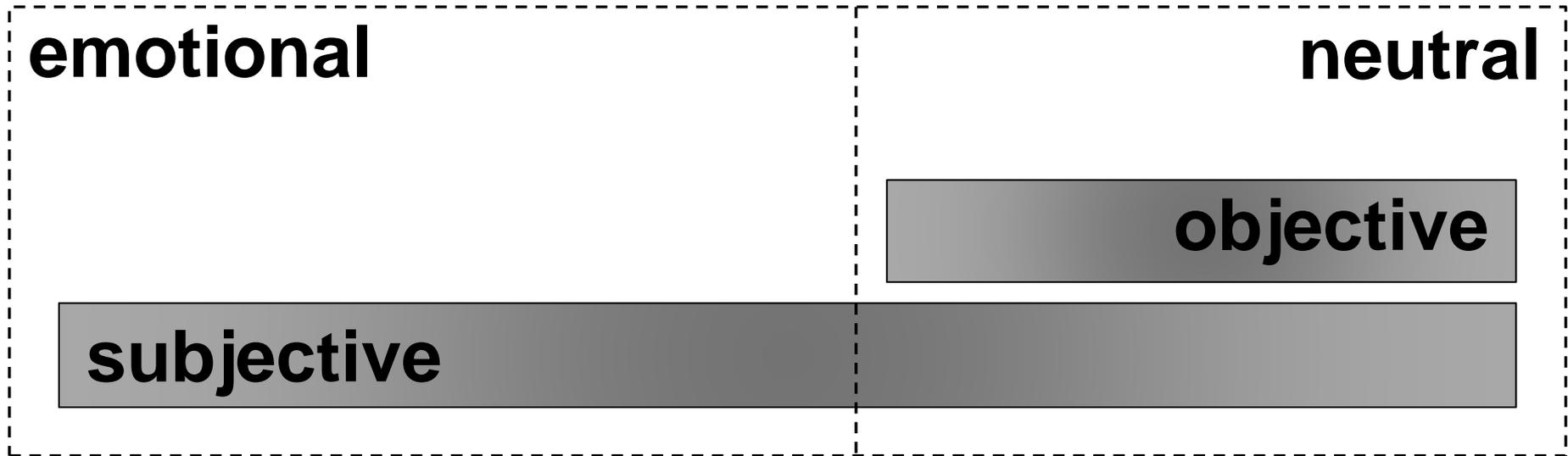
“Was the speaker in an emotional state?”

Easy to ask laypeople
because everyone thinks
they are specialists in their
own emotions.

ambiguous

1. Junko Minato, David B. Bracewell, Fuji Ren and Shingo Kuroiwa. 2006. Statistical Analysis of a Japanese Emotion Corpus for Natural Language Processing. LNCS 4114, pp. 924-929
2. Saima Aman and Stan Szpakowicz. 2007. Identifying expressions of emotion in text. In Proceedings of the 10th International Conference on Text, Speech, and Dialogue (TSD-2007), Lecture Notes in Computer Science (LNCS), Springer-Verlag.
3. Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka. 2011. Affect analysis model: novel rule-based approach to affect sensing from text. Natural Language Engineering, Vol. 17, No. 1 (2011), pp. 95-135.

PROBLEM DEFINITION



“Was the speaker in an emotional state?”

“Did the speaker present their the contents from first-person centric perspective or no specified perspective?”

PROBLEM DEFINITION

emotional

neutral

objective

subjective

“Was the speaker in an emotional state?”

“Did the speaker present their the contents from first-person centric perspective or no specified perspective?”

Doesn't have much to do with emotions
→ Only expressions of emotions tend to be first person centric as well.

PROBLEM DEFINITION

emotional

neutral

objective

subjective

“Was the speaker in an emotional state?”

“Did the speaker present their the conte
perspective or no specified perspective?”

Doesn't have much to do with emotions
→ Only expressions of emotions tend
to be first person centric as well.

Opinion could be subjective but neutral

- “I think it will rain tomorrow.”
- “In my opinion the government should have applied a different policy.”

**** Not talking about positive/negative.****

PROBLEM DEFINITION

emotional

neutral

subjective

objective

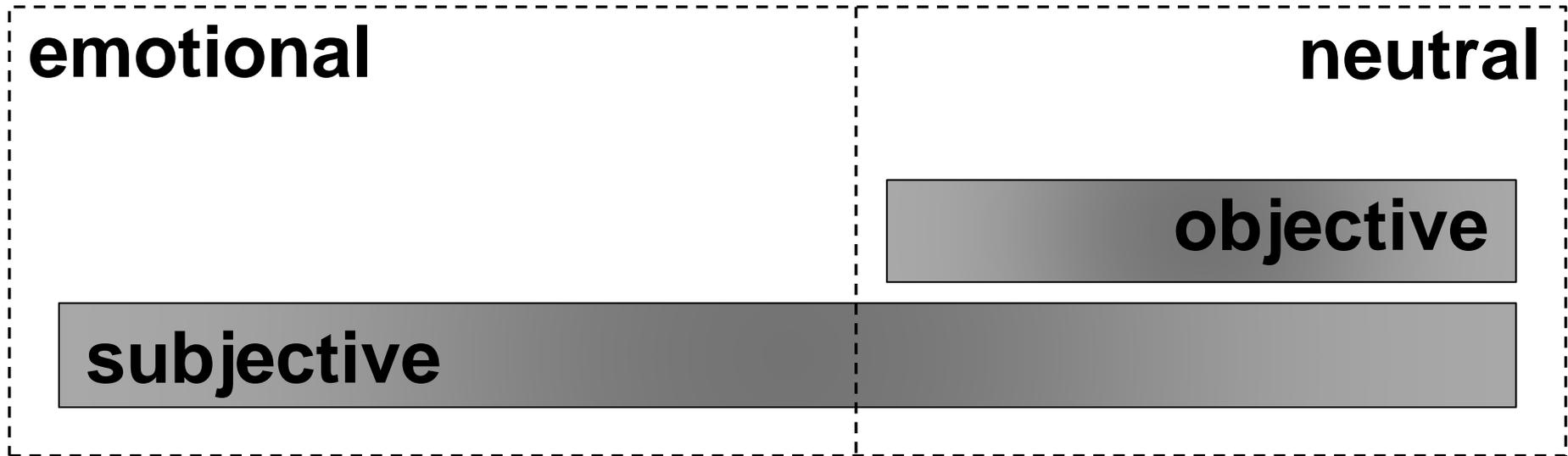
1. Janyce M. Wiebe, Rebecca F. Bruce and Thomas P. O'Hara. 1999. Development and use of a gold-standard data set for subjectivity classifications. In Proceedings of the Association for Computational Linguistics (ACL-1999), pp. 246-253.
2. Theresa Wilson and Janyce Wiebe. 2005. Annotating Attributions and Private States. Proceedings of the ACL Workshop on Frontiers in Corpus Annotation II, pp. 53-60.
3. Hong Yu and Vasileios Hatzivassiloglou. 2003. Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences. In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2003), pp. 129-136.
4. Vasileios Hatzivassiloglou and Janice Wiebe. 2000. Effects of adjective orientation and gradability on sentence subjectivity. In Proceedings of International Conference on Computational Linguistics (COLING-2000), pp. 299-305.

Opinion could be subjective but neutral

- "I think it will rain tomorrow."
- "In my opinion the government should have applied a different policy."

**** Not talking about positive/negative.****

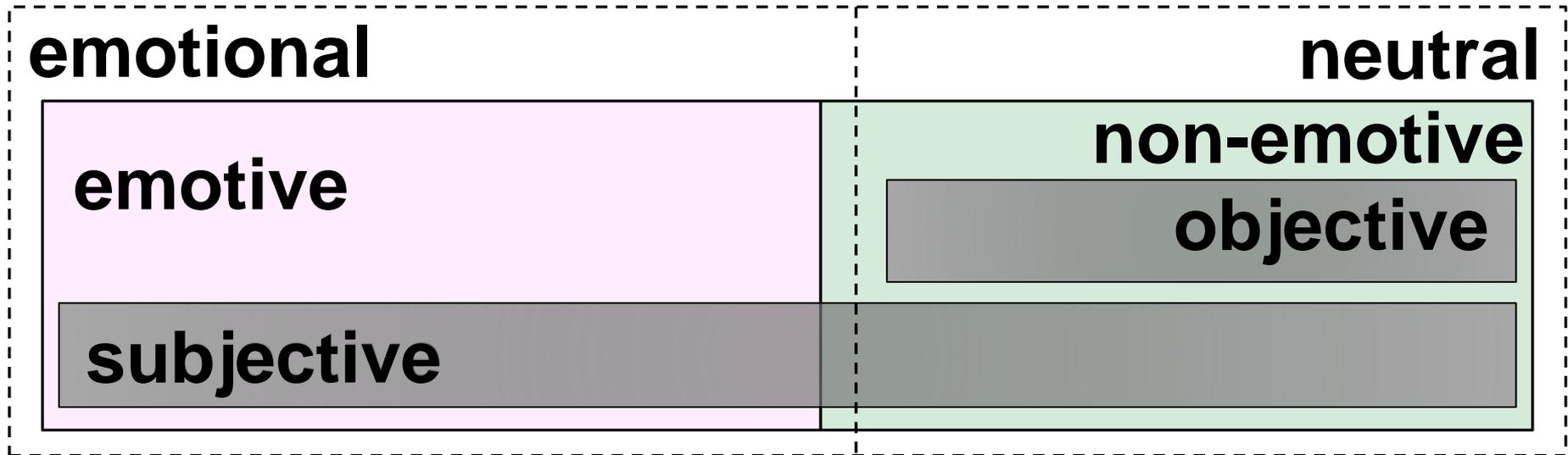
PROBLEM DEFINITION



“Was the speaker in an emotional state?”

“Did the speaker present their the contents from first-person centric perspective or no specified perspective?”

PROBLEM DEFINITION

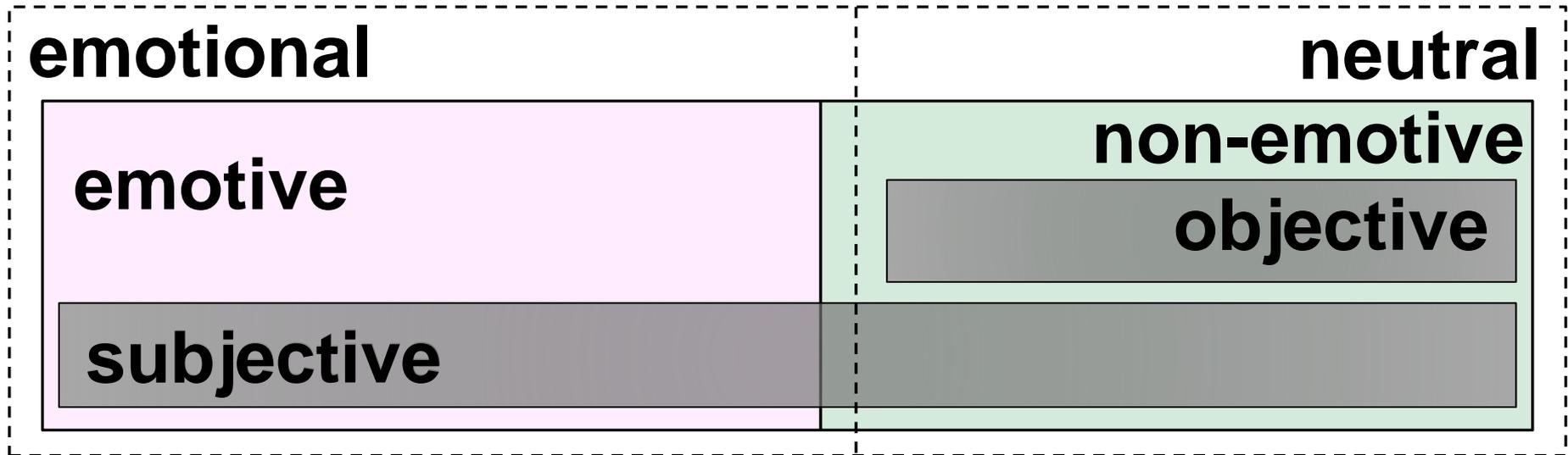


“Was the speaker in an emotional state?”

“Did the speaker present their the contents from first-person centric perspective or no specified perspective?”

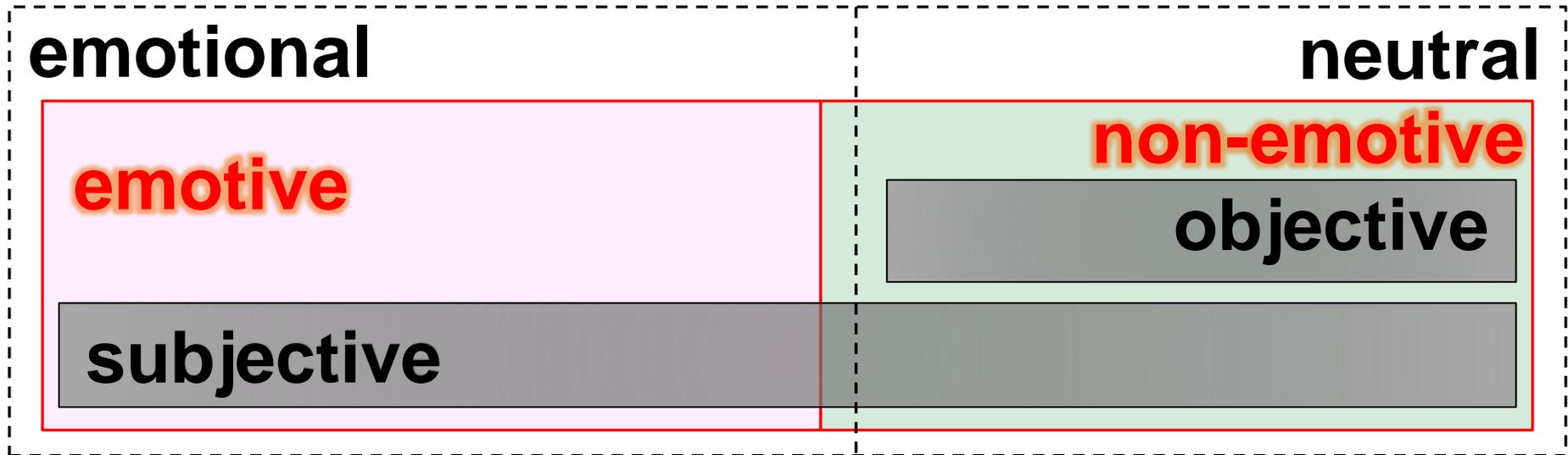
“Was the sentence expressed with emphasis (distinguishable linguistic emotive features)?”

PROBLEM DEFINITION



- All emotive sentences are emotional.
- All neutral and objective sentences are non-emotive.
- Some emotional sentences could be non-emotive.
- Subjective sentences could be emotive/non-emotive as well as emotional/neutral.

PROBLEM DEFINITION



Emotive/non-emotive could help distinguishing between

emotional/neutral

subjective/objective

PROBLEM DEFINITION

In linguistics:

Karl Buhler in 1934: 3 functions of language:
descriptive, impressive, emotive

Stevenson in 1937: emotiveness
(with regards to morality as a concept influenced by emotions)

Roman Jakobson in 1960: 6 functions of language:
+poetic, +phatic, +metalingual

- Karl Buhler. 1990. *Theory of Language. Representational Function of Language*. John Benjamins Publ. (reprint from **Karl Buhler. 1934. *Sprachtheorie***. Die Darstellungsfunktion der Sprache, Ullstein, Frankfurt a. Main, Berlin, Wien,)
- **Stevenson, C. L. 1937. *The Emotive Meaning of Ethical Terms***. In Stevenson, C. L. *Facts and Values*. Yale University Press (published 1963). ISBN 0-8371-8212-3.
- **Roman Jakobson. 1960. *Closing Statement: Linguistics and Poetics***. *Style in Language*, pp.350-377, The MIT Press.

PROBLEM DEFINITION

In linguistics:

- ◆ Emotive elements: **interjections/exclamations** (aah, ooh, whoa, great!), **hypocoristics** (endearments, dog → doggy), **emotive punctuation** (!, ??, ..., ~), **emoticons** (:-), ^o^), **onomatopoeia/mimetic expressions** (*gitaigo* in Japanese)

PROBLEM DEFINITION

In linguistics:

- ◆ Emotive elements: **interjections/exclamations** (aah, ooh, whoa, great!), **hypocoristics** (endearments, dog → doggy), **emotive punctuation** (!, ??, ..., ~), **emoticons** (:-), ^o^), **onomatopoeia/mimetic expressions** (*gitaigo* in Japanese)
- ◆ Plus - combinations of elements

ああ、今日はなんて気持ちいい日なんだ! (^o^)/
Oh, what a pleasant day today, isn't it? (^o^)/

Interjection

Exclamative
phrase

Exclamative
grammar

Exclamation
mark

Emoticon

PROBLEM DEFINITION

In linguistics:

- ◆ Emotive elements: **interjections/exclamations** (aah, ooh, whoa, great!), **hypocoristics** (endearments, dog → doggy), **emotive punctuation** (!, ??, ..., ~), **emoticons** (:-), ^o^), **onomatopoeia/mimetic expressions** (*gitaigo* in Japanese)
- ◆ Plus - combinations of elements

ああ、今日はなんて気持ちいい日なんだ! (^o^)/
Oh, what a pleasant day today, isn't it? (^o^)/

Interjection

Exclamative
phrase

Exclamative
grammar

Exclamation
mark

Emoticon

- ◆ Are there non-emotive elements ??

PROBLEM DEFINITION

In linguistics:

- ◆ Emotive elements: **interjections/exclamations** (aah, ooh, whoa, great!), **hypocoristics** (endearments, dog → doggy), **emotive punctuation** (!, ??, ..., ~), **emoticons** (:-), ^o^), **onomatopoeia/mimetic expressions** (*gitaigo* in Japanese)
- ◆ Plus - combinations of elements

ああ、今日はなんて気持ちいい日なんだ! (^o^)/
Oh, what a pleasant day today, isn't it? (^o^)/

Interjection

Exclamative
phrase

Exclamative
grammar

Exclamation
mark

Emoticon

- ◆ Are there non-emotive elements ??
- ◆ How to extract them?

PROBLEM DEFINITION

ああ、今日はなんて気持ちいい日なんだ！
(Oh, what a pleasant day today, isn't it?)

This sentence contains the pattern:

ああ * なんて * なんだ！ (Oh, what a * isn't it?)

1. This pattern cannot be discovered with n-gram approach.
2. This pattern cannot be discovered if one doesn't know what to look for.

Need to find a way to extract such frequent sophisticated patterns from corpora.

*) pattern = something that frequently appears in a corpus (more than once).

LANGUAGE COMBINATORICS

SPEC – Sentence Pattern Extraction arChitecture

Sentence pattern = ordered non-repeated combinations of sentence elements.

For $1 \leq k \leq n$, there is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ all possible k -long patterns, and

$$\sum_{k=1}^n \binom{n}{k} = \frac{n!}{1!(n-1)!} + \frac{n!}{2!(n-2)!} + \dots + \frac{n!}{n!(n-n)!} = 2^n - 1$$

LANGUAGE COMBINATORICS

SPEC – Sentence Pattern Extraction arChitecture

Sentence pattern = ordered non-repeated combinations of sentence elements.

For $1 \leq k \leq n$, there is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ all possible k -long patterns, and

$$\sum_{k=1}^n \binom{n}{k} = \frac{n!}{1!(n-1)!} + \frac{n!}{2!(n-2)!} + \dots + \frac{n!}{n!(n-n)!} = 2^n - 1$$

Extract patterns from
all sentences and
calculate occurrence.

LANGUAGE COMBINATORICS

SPEC – Sentence Pattern Extraction arChitecture

Sentence pattern = ordered non-repeated combinations of sentence elements.

For $1 \leq k \leq n$, there is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ all possible k -long patterns, and

$$\sum_{k=1}^n \binom{n}{k} = \frac{n!}{1!(n-1)!} + \frac{n!}{2!(n-2)!} + \dots + \frac{n!}{n!(n-n)!} = 2^n - 1$$

And then
classify/
compare
emotive
sentences
with non-
emotive

Normalized pattern weight

$$w_j = \left(\frac{O_{pos}}{O_{pos} + O_{neg}} - 0.5 \right) * 2$$

Score for one sentence

$$score = \sum w_j, (1 \geq w_j \geq -1)$$

LANGUAGE COMBINATORICS

SPEC – Sentence Pattern Extraction arChitecture

Sentence pattern = ordered non-repeated combinations of sentence elements.

For $1 \leq k \leq n$, there is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ all possible k -long patterns, and

$$\sum_{k=1}^n \binom{n}{k} = \frac{n!}{1!(n-1)!} + \frac{n!}{2!(n-2)!} + \dots + \frac{n!}{n!(n-n)!} = 2^n - 1$$

Normalized pattern weight

$$w_j = \left(\frac{O_{pos}}{O_{pos} + O_{neg}} - 0.5 \right) * 2$$

Score for one sentence

$$score = \sum w_j, (1 \geq w_j \geq -1)$$

EXPERIMENT SETUP

DATASET

91 sentences close in meaning, but different emotional load (50 emotive, 41 non-emotive) gathered in an anonymous survey on 30 people of different background (students, businessmen, housewives).

Examples:

Emotive

高すぎるからね

Takasugiru kara ne

'Cause its just too expensive

すごくきれいな海だなあ

Sugoku kirei na umi da naa

Oh, what a beautiful sea!

なんとあの人、結婚するらしいよ

Nanto ano hito, kekkon suru rashii yo

Have you heard? She's getting married!

Non-emotive

高額なためです。

Kougaku na tame desu.

Due to high cost.

きれいな海です

Kirei na umi desu

This is a beautiful sea

あの日と結婚するらしいです

Ano hito kekkon suru rashii desu

They say she is getting married.

EXPERIMENT SETUP

10-fold Cross Validation

DATASET

91 sentences close in meaning, but different emotional load (50 emotive, 41 non-emotive) gathered in an anonymous survey on 30 people of different background (students, businessmen, housewives).

Examples:

Emotive

高すぎるからね

Takasugiru kara ne

'Cause its just too expensive

すごくきれいな海だなあ

Sugoku kirei na umi da naa

Oh, what a beautiful sea!

なんとあの人、結婚するらしいよ

Nanto ano hito, kekkon suru rashii yo

Have you heard? She's getting married!

Non-emotive

高額なためです。

Kougaku na tame desu.

Due to high cost.

きれいな海です

Kirei na umi desu

This is a beautiful sea

あの日と結婚するらしいです

Ano hito kekkon suru rashii desu

They say she is getting married.

EXPERIMENT SETUP

Preprocessing

Sentence:	今日はなんて気持ちいい日なんだ！
Transliteration:	<i>Kyōwanantekimochiihinanda!</i>
Translation:	What a pleasant day it is today!
	Preprocessing examples
1. Tokens:	<i>Kyō wa nante kimochi ii hi nanda !</i>
2. POS:	N TOP ADV N ADJ N COP EXCL
3. Tokens+POS:	<i>Kyō[N] wa[TOP] nante[ADV] kimochi[N] ii[ADJ] hi[N] nanda[COP] ![EXCL]</i>

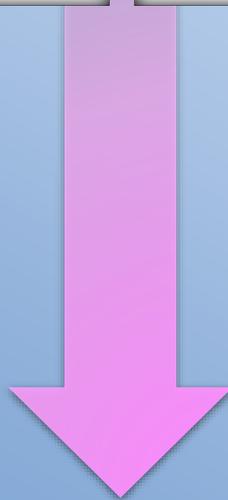
EXPERIMENT SETUP

Pattern List Modification

1. All patterns
2. Zero-patterns deleted
3. Ambiguous patterns deleted

Weight Calculation Modifications

1. Normalized
2. Award length
3. Award length and occurrence



EXPERIMENT SETUP

Pattern List Modification

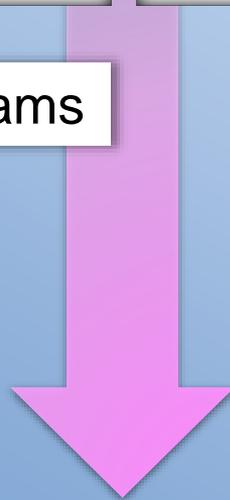
1. All patterns
2. Zero-patterns deleted
3. Ambiguous patterns deleted

Weight Calculation Modifications

1. Normalized
2. Award length
3. Award length and occurrence

All patterns vs. only n-grams

**Is it worth
the time?**



EXPERIMENT SETUP

Pattern List Modification

1. All patterns
2. Zero-patterns deleted
3. Ambiguous patterns deleted

Weight Calculation Modifications

1. Normalized
2. Award length
3. Award length and occurrence

All patterns vs. only n-grams

Automatic threshold setting

Is it worth
the time?

Data is never
perfectly
balanced.

EXPERIMENT SETUP

Pattern List Modification

1. All patterns
2. Zero-patterns deleted
3. Ambiguous patterns deleted

Weight Calculation Modifications

1. Normalized
2. Award length
3. Award length and occurrence

All patterns vs. only n-grams

Automatic threshold setting

10-fold Cross Validation

**Is it worth
the time?**

One experiment
= 280 runs

**Data is never
perfectly
balanced.**

EXPERIMENT SETUP

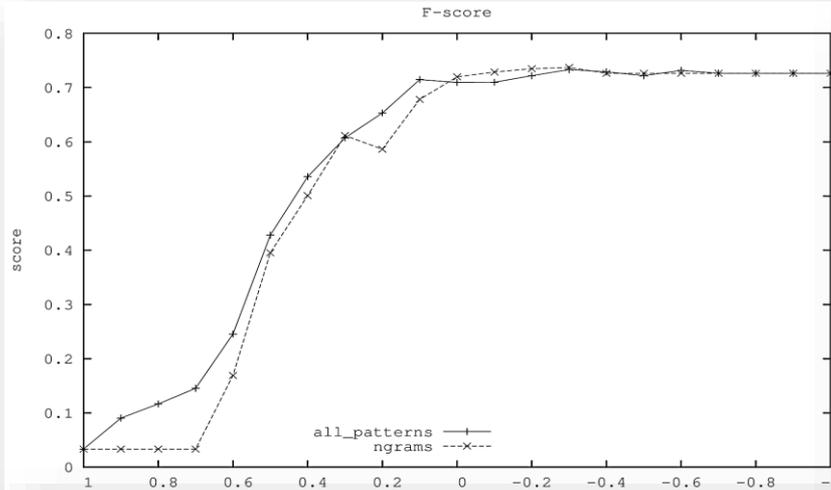
Score calculated in:

- Precision
- Recall
- Balanced F-score
- Accuracy
- Specificity
- Phi-coefficient

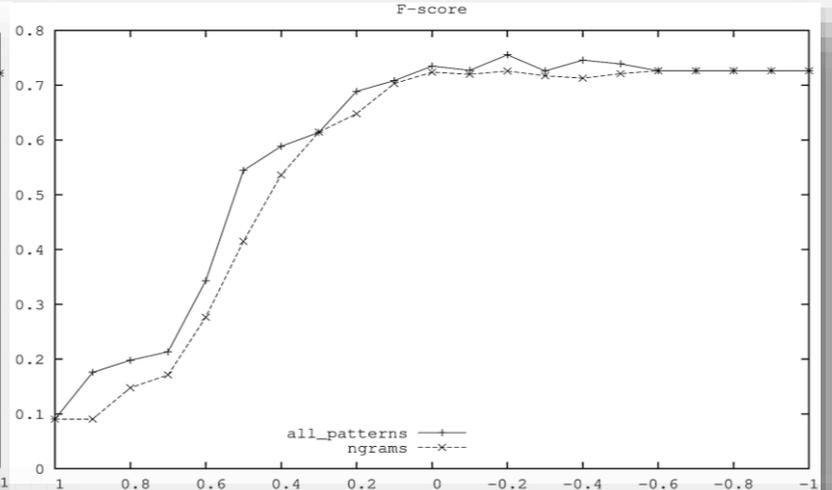
RESULTS AND DISCUSSION

Weight normalized

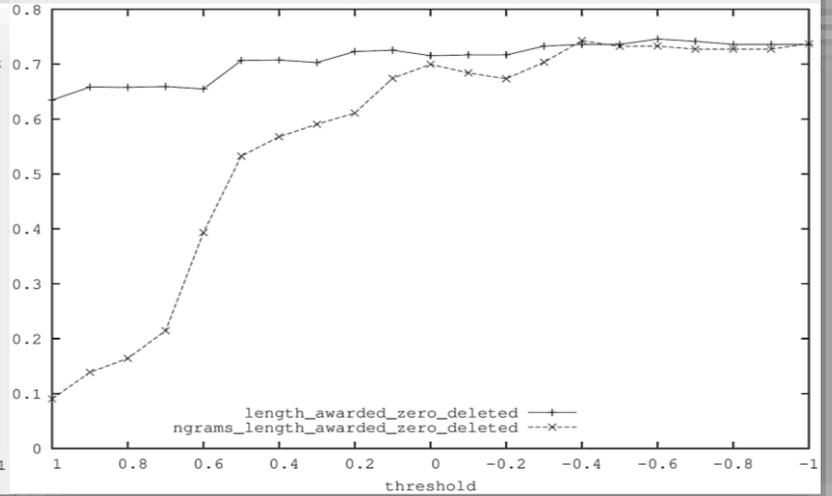
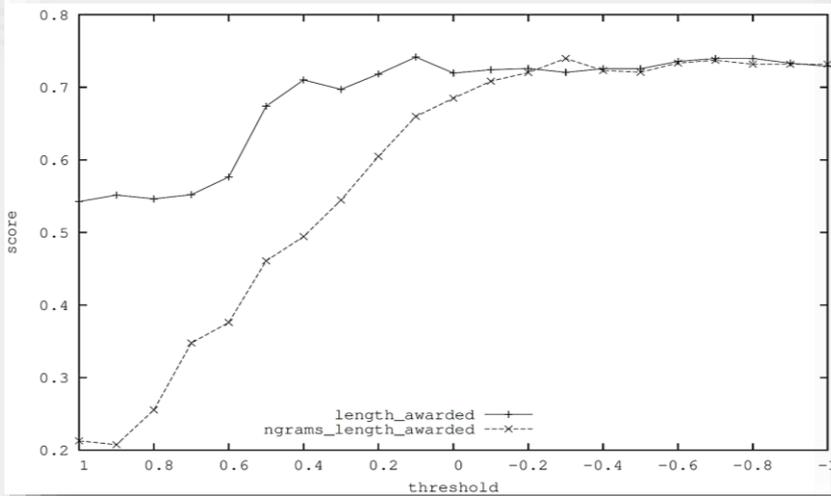
Tokenized



Tokens + POS



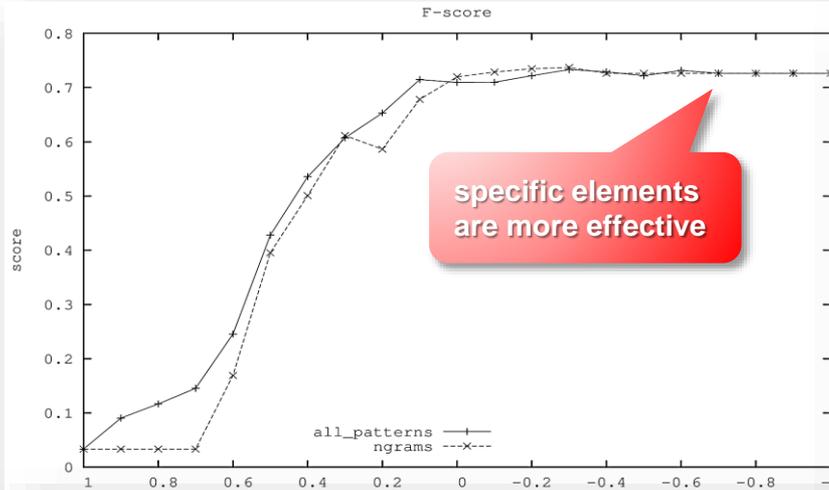
Length awarded



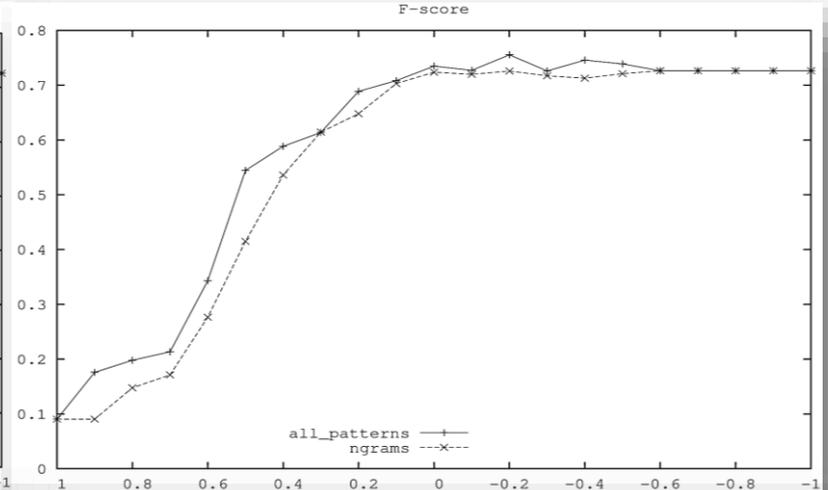
RESULTS AND DISCUSSION

Weight normalized

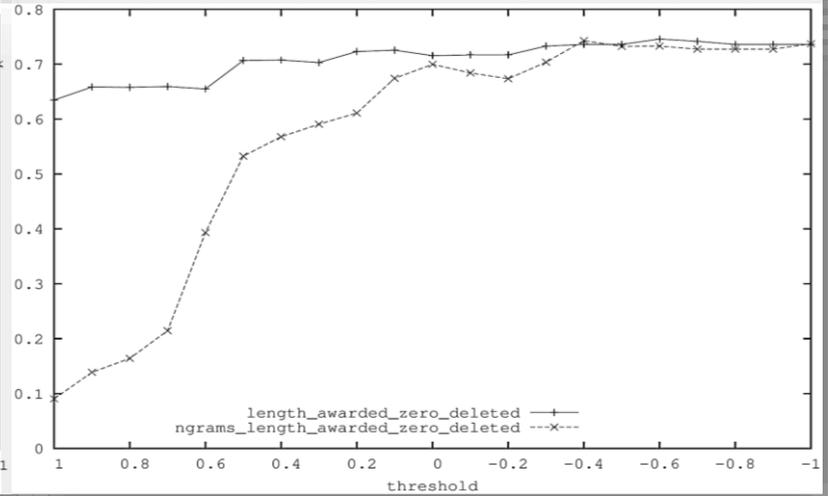
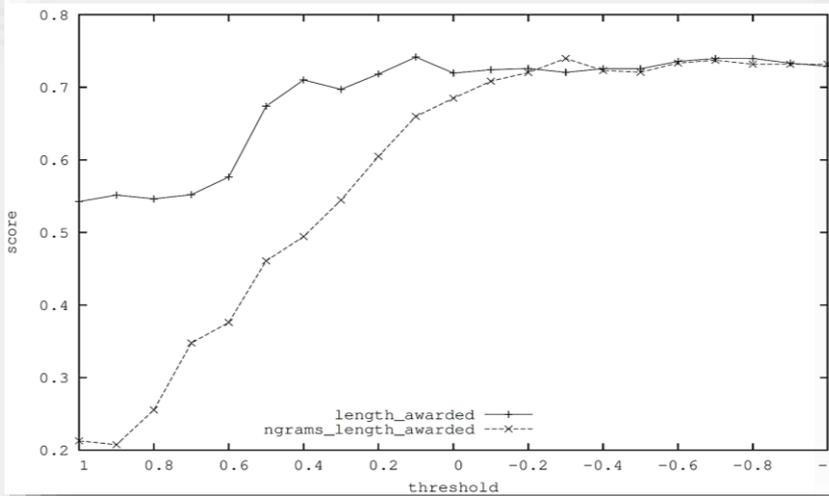
Tokenized



Tokens + POS



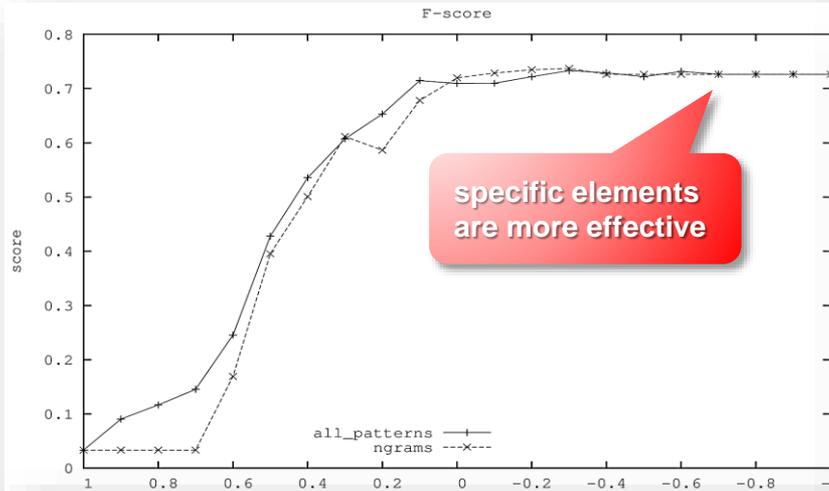
Length awarded



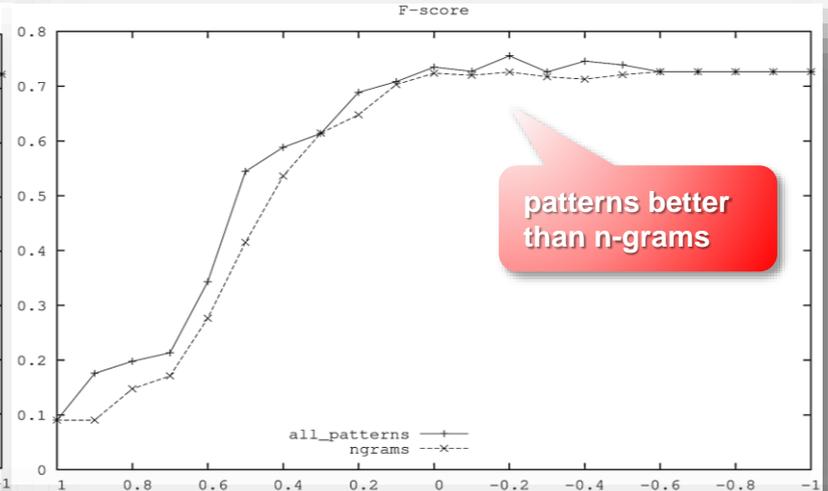
RESULTS AND DISCUSSION

Weight normalized

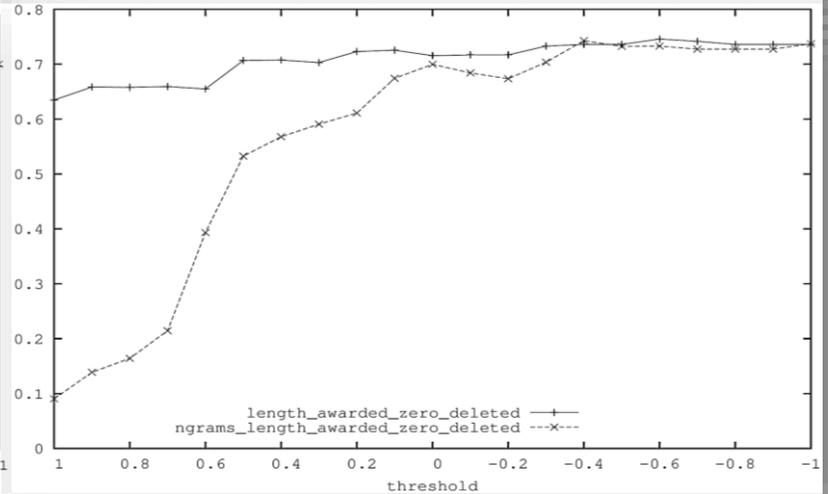
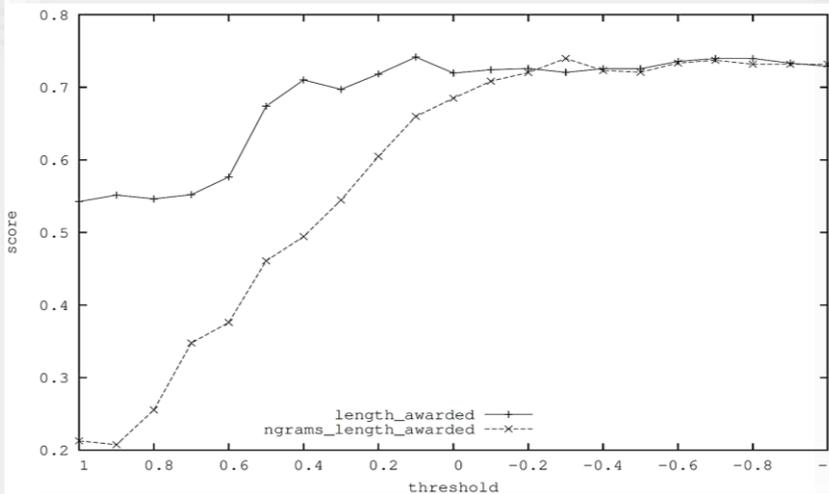
Tokenized



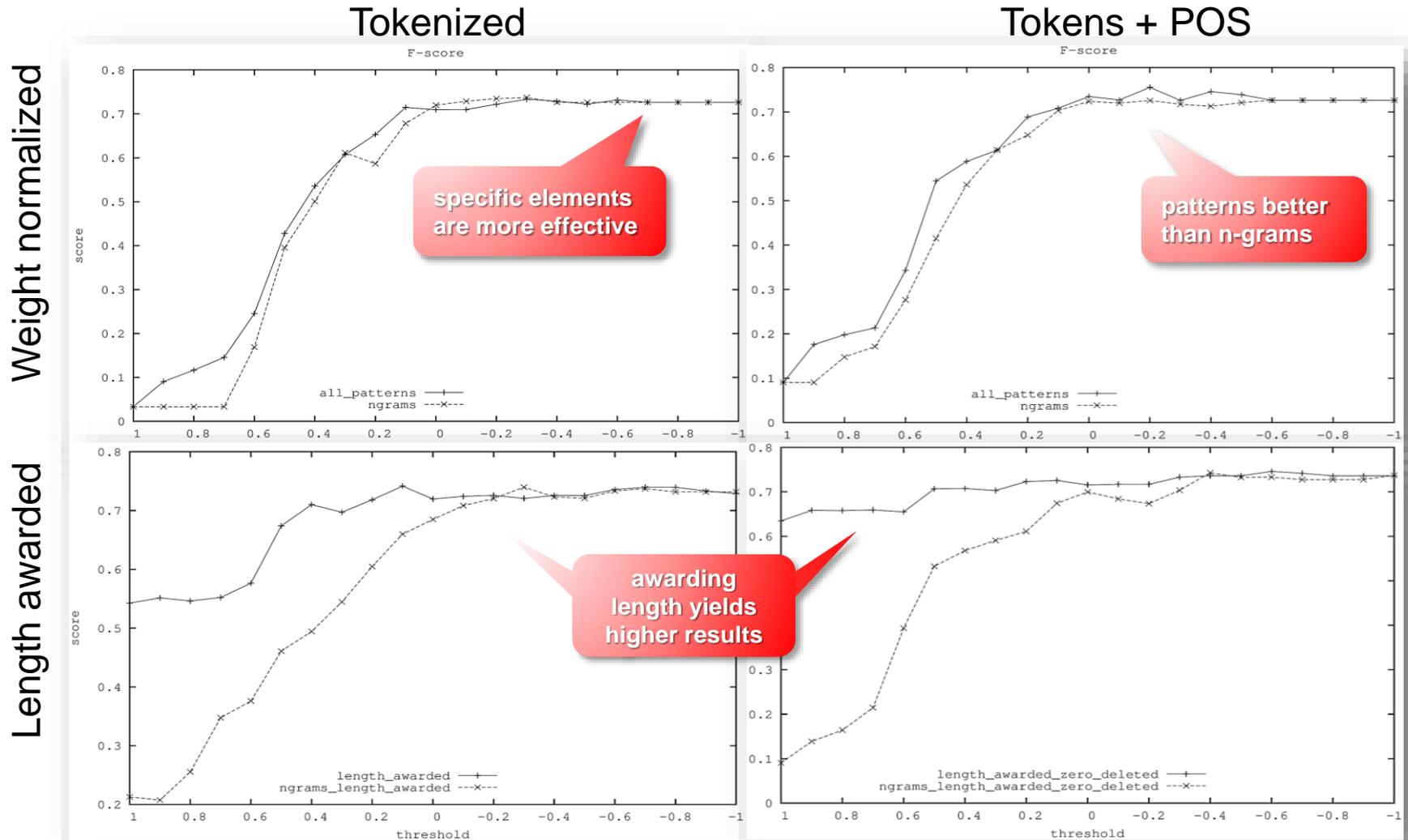
Tokens + POS



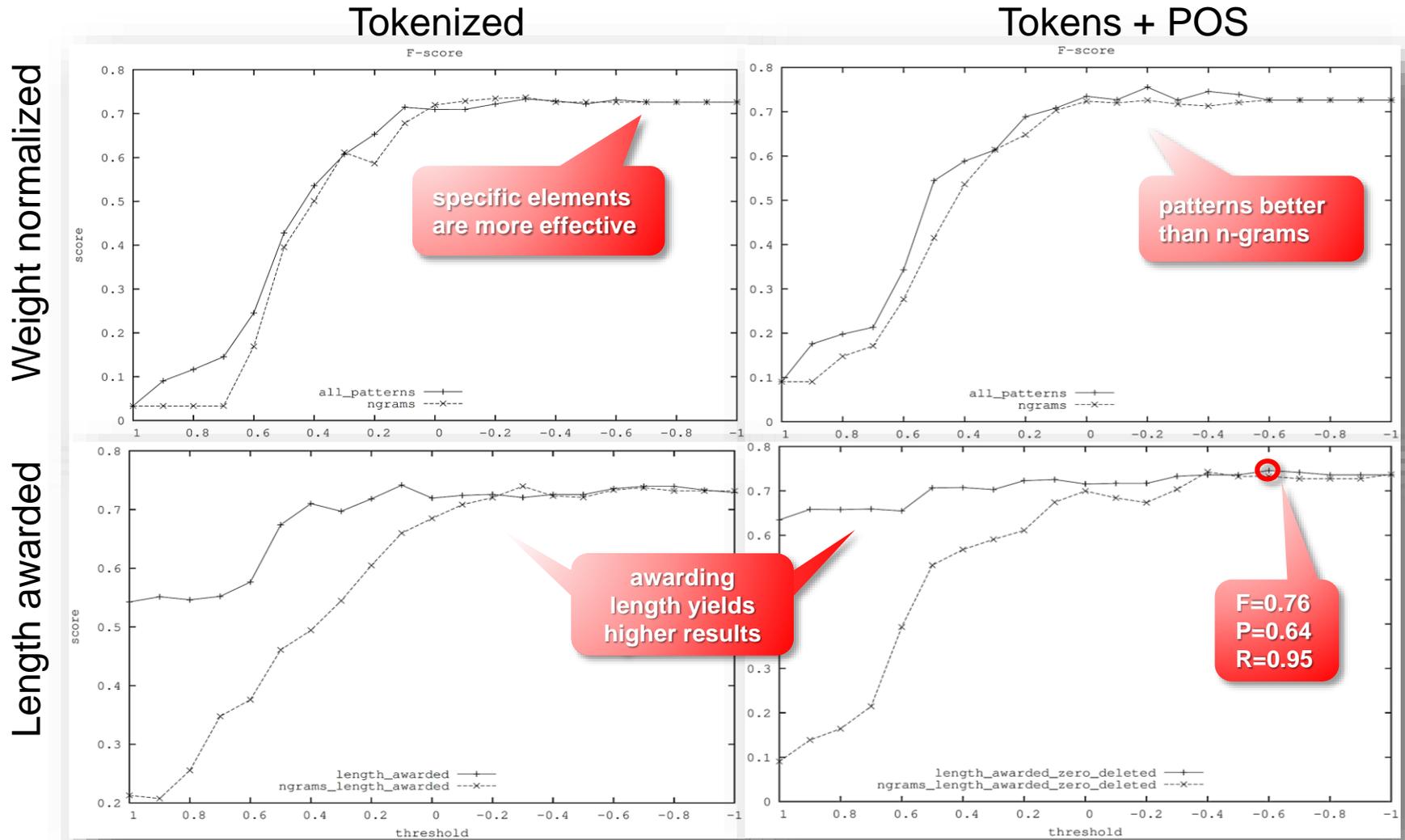
Length awarded



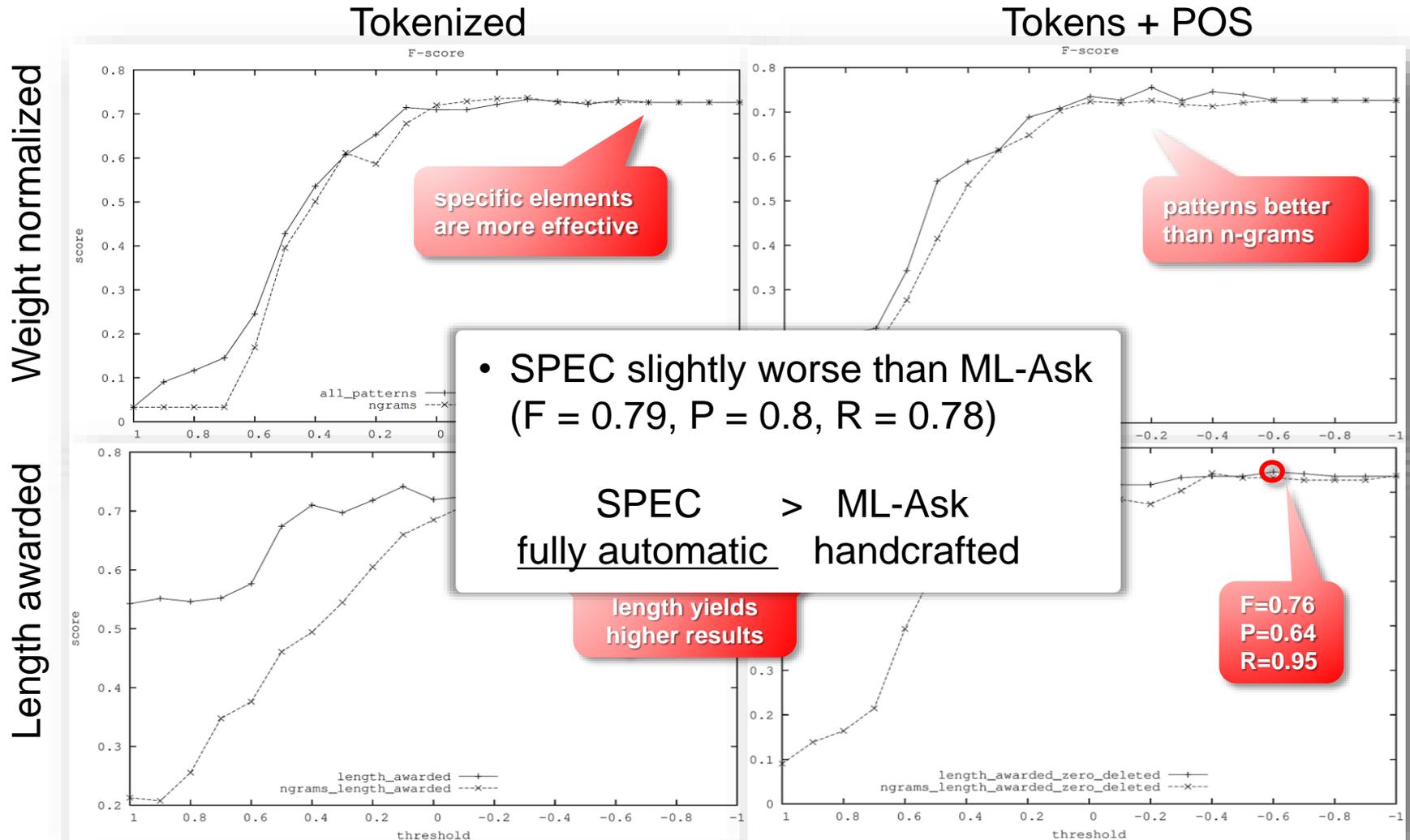
RESULTS AND DISCUSSION



RESULTS AND DISCUSSION



RESULTS AND DISCUSSION

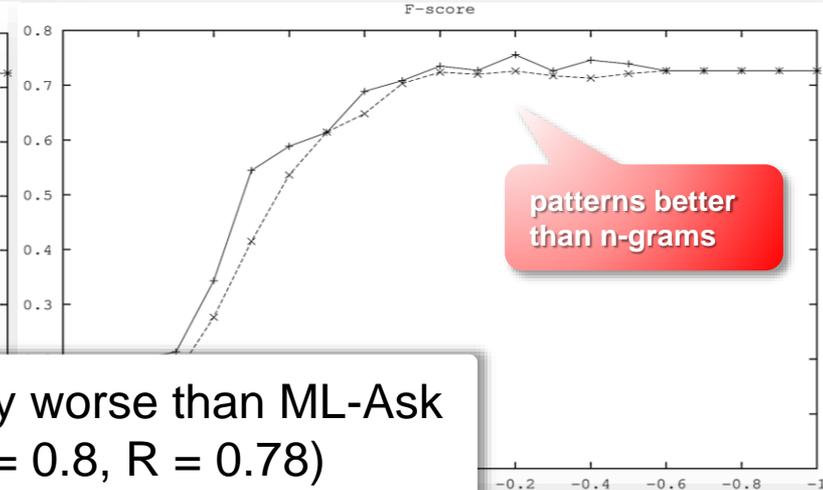
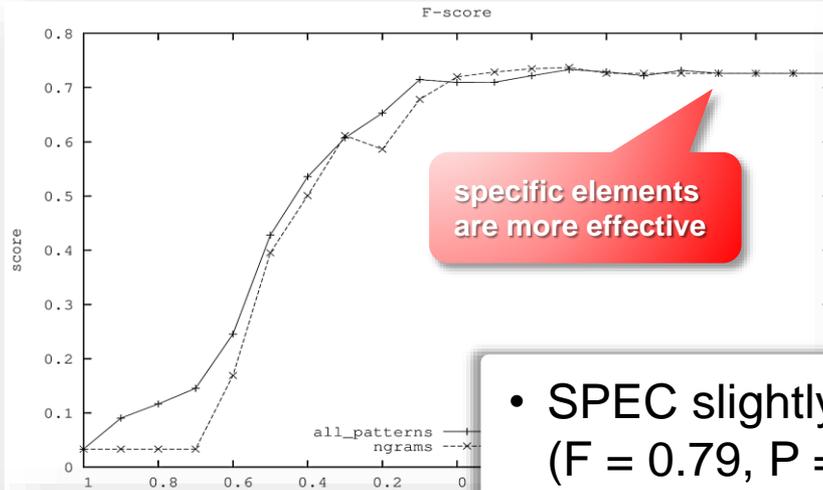


RESULTS AND DISCUSSION

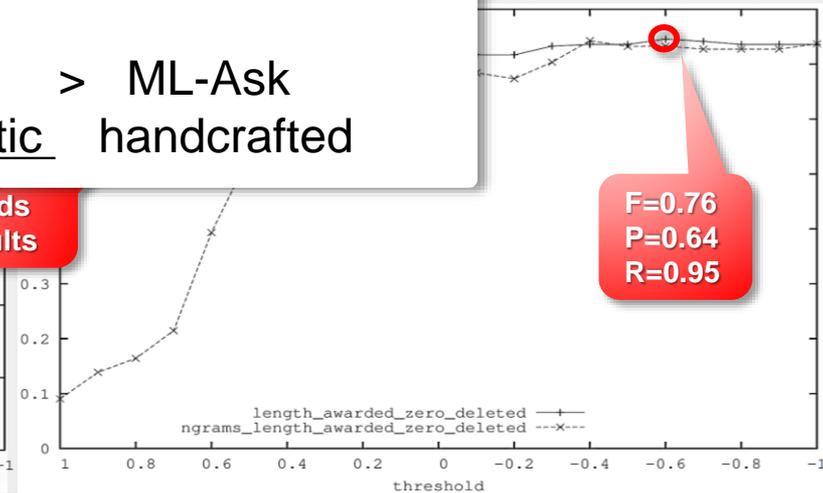
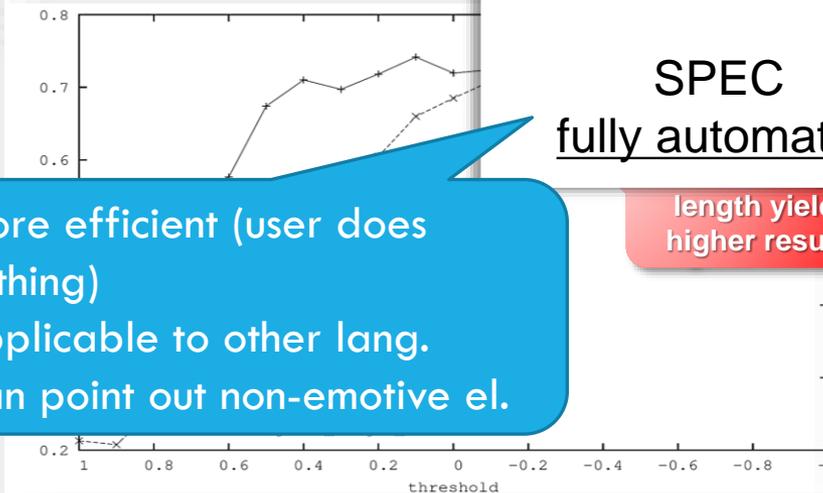
Tokenized

Tokens + POS

Weight normalized



warded



- SPEC slightly worse than ML-Ask (F = 0.79, P = 0.8, R = 0.78)

SPEC > ML-Ask
fully automatic handcrafted

- More efficient (user does nothing)
- Applicable to other lang.
- Can point out non-emotive el.

RESULTS AND DISCUSSION

Examples of extracted
Patterns (Tokenized)

Emotive

- Casual wording
- Prolongation marks
- SFPs
- Subject particles

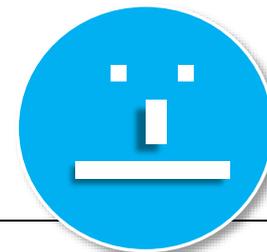
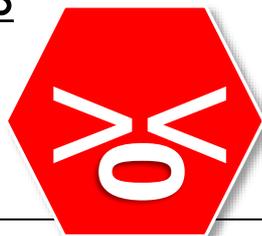
Emotive		Non-emotive	
freq.	example pattern	freq.	example pattern
14	、*た	11	い*。
12	で	8	し*。
11	ん*。	7	です。
11	と	6	は*です
11	—	6	まし*。
10	、*た*。	5	ました。
9	、*よ	5	ます
9	、*ん	5	い
8	し	4	です*。
7	ない	3	この*は*。
7	!	3	は*です。
6	ん*よ	3	て*ます
6	、*だ	3	が*た。
6	ちゃ	3	美味しい
6	よ。	3	た。
5	だ*。	2	た*、*。
5	に*よ	2	せ
5	が*よ	2	か
5	ん	2	さ

Non-emotive

- Official forms (*desu-masu*)
- More “periods”
- No exclamation marks, etc.

RESULTS AND DISCUSSION

EXAMPLE SENTENCES



Example 1.

メガネ、そこにあつたんだよ。
Megane, soko ni atta n da yo .
(The glasses were over there!)

Example 2.

ううん、舞台が見えないよ。
Uun, butai ga mienai yo .
(Ooh, I cannot see the stage!)

Example 3.

ああ、おなかがすいたよ。
Aa, onaka ga suita yo .
(Ohh, I'm so hungry)

Example 4.

高額なためです。
Kougaku na tame desu .
Due to high cost.

Example 5.

きれいな海です
Kirei na umi desu
This is a beautiful sea

Example 6.

今日は雪が降っています。
Kyou wa yuki ga futte imasu .
It is snowing today.

CONCLUSIONS & FUTURE WORK

Presented research on extracting emotive patterns.

Used SPEC - a method for automatic extraction of patterns from sentences.

Extracted the patterns from a set of emotive and non-emotive sentences.

Classified sentences (test data) with those patterns.

Compared different preprocessing techniques (tokenization, POS, token-POS).

The best results obtained patterns with both tokens and POS
(F-score = 76%, Precision = 64%, Recall 95%).

Results for only POS were the lowest. This means the algorithm works better on less abstracted data.

The results of SPEC were compared to ML-Ask affect analysis system. ML-Ask achieved better Precision, but lower Recall. However, SPEC is fully automatic and thus more efficient and language independent.

Many of the automatically extracted patterns appear in handcrafted databases of ML-Ask, which suggests it could be possible to improve ML-Ask performance by extracting additional patterns with SPEC.

In the future we'll try to quantify the correlation between emotive-subjective-emotional

THANK YOU FOR YOUR ATTENTION!

Michal Ptaszynski Kitami Institute of Technology

ptaszynski@ieee.org

<http://orion.cs.kitami-it.ac.jp/tipwiki/michal>