



Title	Contextual affect analysis : a system for verification of emotion appropriateness supported with Contextual Valence Shifters
Author(s)	Ptaszynski, Michal; Dybala, Pawel; Shi, Wenhan; Rzepka, Rafal; Araki, Kenji
Citation	International Journal of Biometrics, 2(2), 134-154 <a href="https://doi.org/10.1504/IJBM.2010.031793">https://doi.org/10.1504/IJBM.2010.031793</a>
Issue Date	2010
Doc URL	<a href="http://hdl.handle.net/2115/47138">http://hdl.handle.net/2115/47138</a>
Type	article
File Information	IJB2-2_134-154.pdf



[Instructions for use](#)

---

## **Contextual affect analysis: a system for verification of emotion appropriateness supported with Contextual Valence Shifters**

---

Michal Ptaszynski\*, Pawel Dybala,  
Wenhan Shi, Rafal Rzepka and Kenji Araki

Graduate School of Information Science and Technology,

Hokkaido University,

Kita-ku, Kita 14 Nishi 9, Sapporo 060-0814, Japan

Fax: +81-11-709-6277

E-mail: ptaszynski@media.eng.hokudai.ac.jp

E-mail: paweldybala@media.eng.hokudai.ac.jp

E-mail: shibuka@media.eng.hokudai.ac.jp

E-mail: kabura@media.eng.hokudai.ac.jp

E-mail: araki@media.eng.hokudai.ac.jp

\*Corresponding author

**Abstract:** This paper presents a novel method for estimating speaker's affective states based on two contextual features: valence shifters and appropriateness. Firstly, a system for affect analysis is used to recognise specific types of emotions. We improve the baseline system with the analysis of Contextual Valence Shifters (CVS), which determine the semantic orientation of emotive expressions. Secondly, a web mining technique is used to verify the appropriateness of the recognised emotions for the particular context. Verification of contextual appropriateness of emotions is the next step towards implementation of Emotional Intelligence Framework in machines. The proposed method is evaluated using two conversational agents.

**Keywords:** affect analysis; contextual appropriateness; web-mining; CVS; contextual valence shifters.

**Reference** to this paper should be made as follows: Ptaszynski, M., Dybala, P., Shi, W., Rzepka, R. and Araki, K. (2010) 'Contextual affect analysis: a system for verification of emotion appropriateness supported with Contextual Valence Shifters', *Int. J. Biometrics*, Vol. 2, No. 2, pp.134-154.

**Biographical notes:** Michal Ptaszynski received his MA from the University of Adam Mickiewicz in Poznan, Poland in 2006. He was a research student at Otaru University of Commerce, and since 2007, he is studying towards his PhD Degree at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include natural language processing, dialogue processing, affect analysis, sentiment analysis, HCI and information retrieval. He is a member of the IEEE, SOFT, JSAI and NLP.

Pawel Dybala received his MA from the Jagiellonian University in Krakow, Poland in 2006. He was a research student at Hokkaido University, and since 2007, he is studying towards his PhD Degree at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include natural language processing, dialogue processing, humour processing, HCI and information retrieval.

Wenhan Shi received his BE from Hokkaido University, Japan. Since 2008, he is studying towards his MA Degree at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include natural language processing, web mining, affect analysis, and information retrieval.

Rafal Rzepka received his MA from the University of Adam Mickiewicz in Poznan, Poland in 1999 and PhD from Hokkaido University, Japan in 2004. Now, he is an Assistant Professor at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include natural language processing, web mining, commonsense retrieval, dialogue processing, language acquisition, affect and sentiment analysis. He is a member of the AAAI, ACL, JSAI, IPSJ, IEICE, JCSS and NLP.

Kenji Araki received BE, ME and PhD Degrees in Electronics Engineering from Hokkaido University, Sapporo, Japan in 1982, 1985 and 1988, respectively. In April 1988, he joined Hokkai Gakuen University, Sapporo, Japan. He was a Professor of Hokkai Gakuen University. He joined Hokkaido University in 1998 as an Associate Professor of the Division of Electronics and Information Engineering. He was a Professor of the Division of Electronics and Information Engineering of Hokkaido University from 2002. Now, he is a Professor of the Division of Media and Network Technologies of Hokkaido University. His research interests include natural language processing, spoken dialogue processing, machine translation and language acquisition. He is a member of the AAAI, IEEE, JSAI, IPSJ, IEICE and JCSS.

---

## **1 Introduction**

In recent years there has been a rising tendency in Computer Science and Artificial Intelligence research to enhance Human-Computer Interaction by humanising machines to make them more user-friendly (Treur, 2007). The humanising process can consist of making robots look like humans (MacDorman and Ishiguro, 2006), but if a robot capable to act and talk with a user on the human level is to be created, it also needs to be equipped with procedures allowing it to understand human cognitive behaviours. Such robots could be very usefull, playing roles of intelligent mates for humans, for example helping children in their development (Rzepka et al., 2008) or helping drivers not to fall asleep during a long ride home (Takahashi et al., 2003).

One of the most important cognitive human behaviours present in everyday communication is expressing and understanding emotions. Emotional states influence the decision making process in humans (Schwarz, 2000) and are a vital part of human intelligence (Salovey and Mayer, 1990). Therefore one of the current issues in Artificial Intelligence is to produce methods for efficient processing of users' emotional states.

The field embracing this subject, called Affective Computing, although being a rather young discipline of study has been gathering popularity of researchers since being initiated only a little over ten years ago (Picard, 1997). The interest in such research is usually focused on recognising the emotions of users in human-computer interaction. In the most popular methods the emotions are recognised from: facial expressions (Hager et al., 2002), voice (Nwe et al., 2003) or biometric data (Teixeira et al., 2008). However these methods, usually based on behavioural approaches, ignore the semantic context of emotions. Therefore, although achieving good results in the laboratory, such methods become useless in real life. A system for recognition of emotions from facial expressions, assigning 'sadness' when a user is crying would be critically mistaken if the user was e.g., cutting an onion in the kitchen. This leads to the need of applying contextual analysis to emotion processing. Furthermore, although recent discoveries prove that affective states should be analysed as emotion-specific rather than divided simply into positive or negative valence (Lerner and Kelter, 2000), most of the behavioural approach methods are incapable to distinguish emotions in more subtle manner than the two-part classification to pairs like joy-anger, or happiness-sadness (see Teixeira et al., 2008). Furthermore, biometric methods, like functional Magnetic Resonance Imaging (fMRI) or Electroencephalography (EEG), are usually laborious, time consuming and expensive. A problem is also how much the performing of an fMRI experiment in itself would influence the participants' emotional states. For example, a subject might feel nervous because of being plugged into the fMRI apparatus, which would obviously influence the results of an emotion detection experiment.

This led to the formation of Affect Analysis – a field focused on developing natural language processing techniques for estimating the emotive aspect of text to provide a support for biometric methods or a substitution when the use of such methods becomes difficult. There have been several attempts to achieve this goal for the Japanese language. For example, Tsuchiya et al. (2007) tried to estimate emotive aspect of utterances with a use of an association mechanism. On the other hand, Tokuhisa et al. (2008) used a large number of examples from the web. This positive tendency seen in natural language processing and text mining approaches lead however to a new set of problems in emotion estimation. The lack of standardisation often causes inconsistencies in emotion classification (compare (Tsuchiya et al., 2007; Tokuhisa et al., 2008)). Moreover, none of the present methods is capable to perform a deep contextual analysis.

In this paper, we present a method capable of specifying users' emotional states in a more sophisticated way than a simple valence classification. The method contributes to standardisation of classification of emotions expressed in Japanese since we do not propose our own classification of emotions, but use the most reliable one available today. Finally, our method does not only specify what type of emotion was expressed, but also determines whether the expressed emotion is appropriate

for the context it appears in. The emotions conveyed by the speaker are estimated with an Affect Analysis system and their contextual appropriateness is verified with a web mining technique.

Previous version of this method, presented first by Ptaszynski et al. (2009), was using, as the Affect Analysis system, the baseline of the system proposed by Ptaszynski et al. (2008a). It is based on the idea of two-step classification of emotive content – general emotiveness and specific emotion types. However, one of the problems with this method was confusing the valence polarity of emotive expressions in the last step of the analysis. To solve this problem and to push Ptaszynski's method one more step towards a deeper contextual analysis of emotive content we decided to apply the analysis of CVS to the baseline system to enhance the specific emotion types determination.

## 2 Definitions

### 2.1 Definition and classification of emotions

Our working definition of emotions is based on the work of Nakamura (1993), who defines them as every temporary state of mind, feeling, or affective state evoked by experiencing different sensations. This definition is complemented by Solomon's, who argues that people are not passive participants in their emotions, but rather the emotions are strategies by which people engage with the world. Since we operate on language, the above is further complemented by Beijer's definition of emotive utterances, which he describes as every utterance in which the speaker is emotionally involved, and this involvement, expressed linguistically, is informative for the listener (Beijer, 2002).

Nakamura (1993), after a thorough study of emotions in Japanese, proposed a classification of emotions into ten types – said to be the most appropriate for the Japanese language. That is: *ki*, *yorokobi* (joy, delight), *do*, *ikari* (anger), *ai*, *aware* (sorrow, sadness, gloom), *fu*, *kowagari* (fear), *chi*, *haji* (shame, shyness, bashfulness), *kou*, *suki* (liking, fondness), *en*, *iya* (dislike, detestation), *kou*, *takaburi* (excitement), *an*, *yasuragi* (relief) and *kyou*, *odoroki* (surprise, amazement).

### 2.2 Emotional Intelligence

The idea of Emotional Intelligence (EI) was first officially proposed by Salovey and Mayer (1990) who defined it as a part of human intelligence including the abilities to

- I perceive emotions
- II integrate emotions to facilitate thoughts
- III understand emotions and to
- IV regulate emotions to promote personal growth.

In their EI Framework (Mayer and Salovey, 1997) they generally divide the first step, perception of emotions, into

- a the ability to identify emotions
- b discriminate between accurate (or appropriate) and inaccurate (or inappropriate) expression of emotions, which is a key ability for the third step in EI Framework – understanding and analysis of emotional states.

Salovey and Mayer argue that recognising emotions is only the first basic step to acquire full scope of Emotional Intelligence and does not tell us anything about whether it is appropriate for a given situation, and what actions should be undertaken as a reaction (Salovey and Mayer, 1990). Solomon (1993) argues that the valence of emotions is determined by the context they are expressed in. For example, anger can be warranted (a reaction to a direct offence) or unwarranted (scolding one's own children unjustly) and the reactions should be different for the two different contexts of anger. Not taking this fact into consideration in the grand majority of emotion processing research can drive us to the conclusion that such research assumes that positive emotions are always appropriate and negative ones inappropriate.

The attempts to implement the EI Framework usually do not go beyond theory (Andre et al., 2004), and the few practical attempts eventually still do not go beyond the first basic recognition step (Picard et al., 2001). This paper presents an attempt to go one step further. Following affective state recognition, the appropriateness of those states is checked against their contexts. This represents a step forward in the implementation of EI framework.

### 2.3 *Linguistic approach to emotions*

The semantic and pragmatic diversity of emotions is best conveyed in language (Solomon, 1993). Therefore we designed our method to be language-based. There are different linguistic means used to inform other interlocutors of emotional states. The elements of speech used to convey emotive meaning in sentence make up the feature of language called emotiveness (Stevenson, 1963), which is described by the emotive function of language (Buhler, 1934; Jakobson, 1960).

The emotive function of language is realised lexically in Japanese through such parts of speech as exclamations (Beijer, 2002), hypocoristics (endearments), vulgar language (Crystal, 1989; Potts and Kawahara, 2004) and mimetic expressions (*gitaigo*) (Baba, 2003). A key role in expressing emotions is also played by the lexicon of words describing emotional states (Nakamura, 1993). The analysis of elements of language such as intonation or tone of voice, in the communication channel limited to transmission of lexical symbols, must focus on its textual manifestations, like exclamation marks or ellipsis.

### 2.4 *Contextual Valence Shifters*

The idea of Contextual Valence Shifters (CVS) application in Sentiment Analysis was first proposed by Polanyi and Zaenen (2004). They distinguish two kinds of CVS: negations and intensifiers. The group of negations contains words and phrases like 'not', 'never', and 'not quite', which change the valence polarity of semantic orientation of an evaluative word they are attached to. The group of intensifiers contains words like 'very', 'very much', and 'deeply', which intensify the semantic

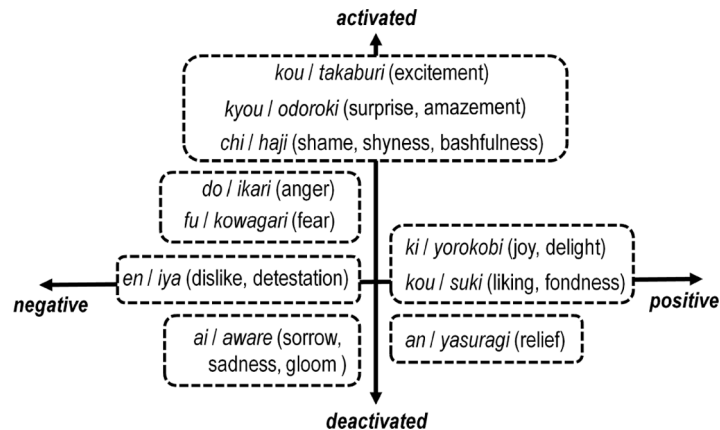
orientation of an evaluative word. So far the idea of CVS analysis was successfully applied to the field of Sentiment Analysis of texts in English (Kennedy and Inkpen, 2005). A few attempts on Japanese ground (Miyoshi and Nakagami, 2007) show that it is also applicable for the Japanese language.

Examples of CVS negations in the Japanese language are grammatical structures such as: *-nai*<sup>1</sup> (not-), *amari -nai* (not quite-), *-to wa ienai* (cannot say it is-), *matteku -nai* (not at all-), or *sukoshi mo -nai* (not even a bit-). Intensifiers are represented by such grammatical structures as: *totemo-* (very much-), *sugoku-* (-a lot), or *kiwamete-* (extremely). However, till now there were no attempts to apply CVS in the field of Affect Analysis neither in English nor in Japanese. This paper presents the first pioneering attempt of that kind.

### 2.5 Two-dimensional model of affect

The idea of a two-dimensional model of affect was first proposed by Schlosberg (1952) and developed further by Russell (1980). Its main assumption is that all emotions can be described in a space of two-dimensions: the emotion's valence polarity (positive/negative) and activation (activated/deactivated). An example of positive-activated emotion would be an 'excitement'; a positive-deactivated emotion is, for example, a 'relief'; negative-activated and negative-deactivated emotions would be 'anger' and 'gloom' respectively. This way four areas of emotions are distinguished: activated-positive, activated-negative, deactivated-positive and deactivated-negative (see Figure 1).

**Figure 1** Grouping Nakamura's classification of emotions on Russell's space



Nakamura's emotion types were mapped on two-dimensional model of affect and their affiliation to one of the spaces is determined. However for some emotion types the affiliation is not obvious (e.g., surprise can be both positive as well as negative; dislike can be either activated or deactivated, etc.) and therefore they were mapped on all of the groups they could belong to. However no emotion type was mapped on more than two adjacent fields. This grouping is then used in our system for two reasons. Firstly, in the CVS analysis procedure to verify which emotion corresponds to the one negated by a negation CVS phrase. Secondly, in emotion appropriateness

varification procedure, for estimating whether the emotion types tagged by ML-Ask belong to the same Russell's space, even if not perfectly matching with the emotive associations gathered from the web.

### 3 Methods

#### 3.1 *ML-Ask: affect analysis system*

Based on the linguistic approach and assumptions described above Ptaszynski et al. (2008a) constructed ML-Ask (Emotive Elements/Emotive Expressions Analysis System) for analysing the emotive contents of utterances. The system uses a two-step procedure:

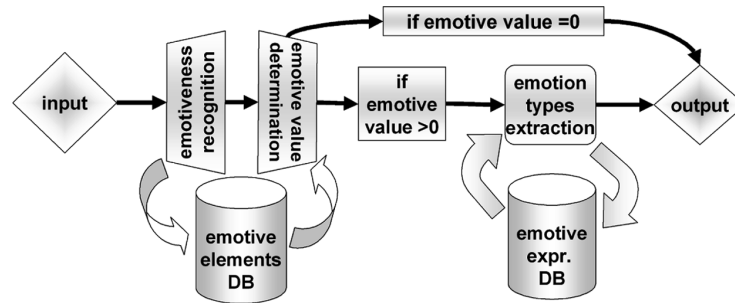
- 1 Analysing the general emotiveness of an utterance by calculating the emotive value representing the strength of the conveyed emotions.
- 2 Recognising the particular emotion types. The system is based on Ptaszynski's (2006) idea of two-part analysis of realisations of emotions in language into:
  - 1 *Emotive elements/emotemes*. Indicating that emotions have been conveyed, but not detailing what specific emotions there are. The same emotive element can express different emotions depending on context. This group is linguistically realised by such subgroups as interjections, exclamations, mimetic expressions, or vulgar language. Examples are: *sugee* (great!), *wakuwaku* (heart pounding), *-yagaru* (a vulgarisation of a verb).
  - 2 *Emotive expressions*. Parts of speech, that in emotive sentences describe emotional states. This group is realised by various parts of speech, like nouns, verbs, adjectives, etc. Examples are: *aijou* (love), *kanashimu* (feel sad), *ureshii* (happy), respectively.

The emotive element database was built using data from different research (Oshima-Takane et al., 1995–1998; Tsuchiya, 1999; Baba, 2003; Sjöbergh and Araki, 2008) and divided into interjections, mimetic expressions, endearments, vulgarities, and representations of non-verbal emotive elements, such as exclamation marks or ellipsis. An algorithm detecting emoticons was also added, as they are symbols commonly used in everyday text-based communication tools. The database of emotive expressions contains Nakamura's collection (Nakamura, 1993).

##### 3.1.1 *Affect analysis procedure*

For a textual input provided by the user, two features are calculated in order: emotiveness of an utterance and specific type of emotion. Firstly, the system searches for emotive elements in the utterance to determine the emotiveness (emotive/non-emotive). Secondly, in utterances described as emotive the system searches for emotive expressions to determine the specific type of emotion (or emotions) conveyed in the utterance. An example of analysis performed by ML-Ask is shown in Table 1.



**Figure 2** Flow chart of the ML-Ask system**Table 1** An example of analysis performed by ML-Ask

<b>Utterance</b>	<i>Kono hon saa, sugee kowakatta yo. Maji kowasugi.</i> (That book, ya know, 'twas a killer. It was just too scary.)
<b>Emotive elements</b>	<i>saa, sugee, -yo, maji, -sugi</i> (Emotive value = 5; emotiveness = yes)
<b>Emotive expressions</b>	<i>kowai</i> (fear)

### 3.1.2 Applying CVS to ML-Ask

One of the problems in the procedure described above was confusing the valence polarity of emotive expressions. The cause of this problem was extracting from the utterance only the emotive expression keywords without its grammatical context. One utterance showing such a case is presented in Table 2. In this sentence the emotive expression is the verb *akirameru* (give up [verb]) and a CVS phrase, *-cha ikenai* (Don't- [particle+verb]) is present, suggesting that the speaker is in fact negating and forbidding the emotion expressed literally. To solve this problem we apply the analysis of Contextual Valence Shifters to change the valence polarity of emotive expressions in utterances containing CVS structures. However, using only the CVS analysis we would be able to find out about an appropriate valence of emotions conveyed in the utterance, but we would not know the exact emotion type. Therefore, to specify the emotion types in such utterances we apply the idea of the two-dimensional model of affect.

**Table 2** An example of failure of emotion description in ML-Ask and improvement with CVS analysis procedure

	<i>Sentence author's tagging</i>	<i>ML-Ask baseline</i>	<i>ML-Ask + CVS</i>
<i>Akirame cha ikenai yo!</i> (Don't give up!)	[joy, delight], [excitement]	[dislike, detestation]	[joy, delight], [liking, fondness]

### *Applying the two-dimensional model of affect to the CVS procedure*

The need of changing the valences in emotion estimation research is a common problem. However, it is not uncommon that researchers use valence changing

patterns constructed by themselves without any scientific grounds. For example, Tsuchiya et al. (2007) used their own list of contrasting emotions. However, they do not notice that, as is argued by Solomon (1993), the fact that two emotions are in a contrast is not a matter of a stiff division, but is more profound and context dependent. We assumed this profundity could be specified with the help of the two-dimensional model of affect.

### *Description of the CVS procedure*

Analysis of Contextual Valence Shifters is a supplementary procedure for ML-Ask and is described as follows. When a CVS structure is discovered, ML-Ask changes the valence polarity of the emotion conveyed in the sentence. According to Russell's model, every emotion is placed in one of the four spaces. However, some emotions have a tendency to appear in two of the quarters. The appropriate emotion is determined as the one with valence polarity and activation parameters different to the contrasted emotion. If an emotion was located in only one quarter, e.g., positive-activated, the contrasting emotions would be determined as negative-deactivated. A change in the output is shown in Table 2. Originally ML-Ask selected [dislike, detestation] (see Table 2). This emotion is located in both quarters of the negative valence space. Therefore ML-Ask determines the new emotion types as positive and belonging to both of the positive quarters. The new proposed emotion types are: [joy, delight] and [liking, fondness] belonging to both positive-activated and positive-deactivated quarters.

### *3.1.3 Evaluation of ML-Ask*

To verify the system's accuracy we performed an evaluation experiment. In the experiment we evaluated the ML-Ask system's performance in the two steps of the procedure described in Section 3.1.

The evaluation is based on a corpus of natural utterances gathered through an anonymous survey. In the survey 30 people from different age and social groups participated. Each of them was to imagine or remember a conversation (or conversations) with any person (or persons) they know and write three sentences from that conversation: one free, one emotive, and one non-emotive. After that the sentences' authors tagged the utterances written by themselves in the same way, as the system's procedure – first it was determined whether an utterance was emotive. If it was, they set the emotive value (0–5) and described the specific emotion types conveyed in the emotive utterances.

We used ML-Ask to analyse this corpus of 90 sentences, and compared the results of the system with the results provided by the authors of the utterances. Although the ideal would be a perfect agreement, it is highly difficult to achieve such a result, since determining emotive features of speech differs greatly between people. Therefore, we also added an emotive tagging of the corpus by third party human evaluators to determine a general human level in recognising emotions from text in ordinary people. The third party human evaluators (10 people per sentence on average) tagged the corpus emotively in the same way as the system and for each of them we calculated the recognition accuracy by comparing the results to the results provided by the authors of the utterances. The scope of results achieved by the third

party evaluators was considered to be the human level of accuracy in recognising emotions from text.

#### *Emotiveness determination procedure*

The system's accuracy in determining the emotiveness is calculated as the harmonic mean of the accuracies on emotive and non-emotive utterances. The total accuracy of ML-Ask in this matter evaluated on 90 items was 0.83. This result is very promising, since the same value calculated for the third party human evaluators gave a wide range of results from 0.4 to 0.86. The result of ML-Ask was placed close to the top ranking like this, and therefore we can say that the system recognises emotiveness on a very high human level.

As an interesting fact, we should add that for sentences with a perfect match, where both authors and all evaluators agreed on the emotiveness (18 sentences), the system's result was 100.0%.

#### *Emotion types extraction procedure: baseline*

The system can potentially extract many emotion types (up to 10) for one utterance. However some of them can be extracted wrongly, and there is a possibility that there would still be some emotion types not extracted. Therefore we decided to calculate the system's results as the balanced F score with the emotive tags added by the authors of the utterances as a gold standard. Correctly specified emotion type is understood as recognising any and at least one emotion type from a set per utterance, including non-emotive.

The system's accuracy in estimating the specific types of emotions reached balanced F-score of 0.45. As for human evaluators, the average accuracy was 0.72. Therefore the system's accuracy reached approximately 62.5% (0.45/0.72) of the human level.

#### *Emotion types extraction procedure supported with CVS*

In the evaluation of the system supported with the CVS analysis procedure ML-Ask was able to determine correctly all of the sentences containing contextual valence shifters. The accuracy of the system rose from 62.5% to 65.3% (0.47/0.72) of the human level. The improvement is not great, because there were only a few sentences containing CVS structures in the corpus used for evaluation.

### *3.2 Web mining technique*

As a verifier of the appropriateness of the emotion types recognised by ML-Ask as conveyed by the user, we apply Shi's web mining technique for extracting emotive associations from the web based on causality (Shi et al., 2008). The technique is made up of three stages:

- a phrase extraction from an utterance
- b morpheme modification
- c extraction of emotion associations.

### 3.2.1 Phrase extraction procedure

An utterance is first processed by MeCab, a tool for part-of-speech analysis of Japanese (Kudo, 2001). Every element separated by MeCab is treated as a unigram. All the unigrams are grouped into larger  $n$ -gram groups preserving their word order in the utterance. The groups are arranged from the longest  $n$ -gram (the whole sentence) down to all groups of trigrams.  $N$ -grams ending with particles are excluded, since they gave too many ambiguous results in our pre-test phase.

### 3.2.2 Morpheme modification procedure

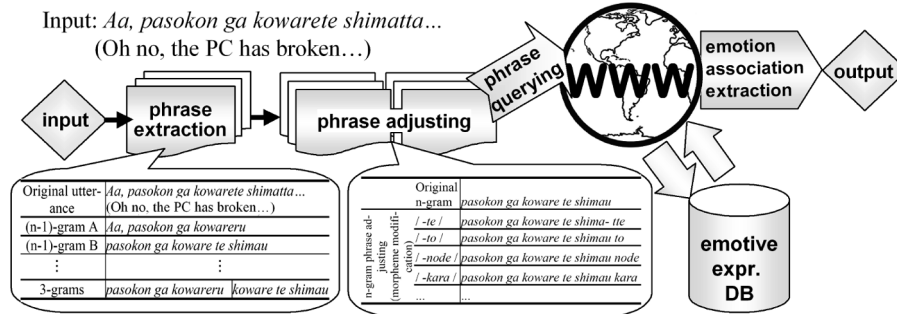
After extracting a list of phrases from the utterance all  $n$ -grams ending with a verb or an adjective are grammatically adjusted for semantically deeper web mining. They are modified in line with Yamashita's research on causality morphemes after which people tend to convey emotive meaning in Japanese (Yamashita, 1999). This research was independently confirmed experimentally by Shi et al. (2008), who distinguished 11 emotively stigmatised morphemes for the Japanese language using statistical analysis of web contents. They performed a cross reference of the appearance of the 11 morphemes with the emotive expression database (see Section 3.1) using the Google search engine. This provided results showing which of the 11 morphemes were most frequently used to express emotions. After sorting the results gradually they came to the conclusion that with the use of only the most frequently used five of them, the coverage of the web mining procedure will still exceed 90.0% of all the content. Therefore in the web mining they decided to use only the first five of the most emotively stigmatised morphemes of causality, namely: *-te*, *-to*, *-node*, *-kara* and *-tara* (see Table 3).

**Table 3** Hit-rate results for each of the 11 morphemes with the ones used in the web mining technique in bold font

<b>morpheme</b>	<b><i>-te</i></b>	<b><i>-node</i></b>	<b><i>-tara</i></b>	<i>-nara</i>	<i>-kotoga</i>	<i>-nowa</i>
<b>result</b>	<b>41.97%</b>	<b>7.20%</b>	<b>5.94%</b>	1.17%	0.35%	2.30%
<b>morpheme</b>	<b><i>-to</i></b>	<b><i>-kara</i></b>	<i>-ba</i>	<i>-noga</i>	<i>-kotowa</i>	×
<b>result</b>	<b>31.97%</b>	<b>6.32%</b>	3.19%	2.15%	0.30%	×

### 3.2.3 Emotion type extraction procedure

The morpheme modification stage provides query phrases for the web mining. All the modified phrases are used as a queries to the Google search engine with 100 snippets per query phrase. This way a maximum number of 500 snippets for each originally extracted phrase is extracted from the web and cross-referenced with the database of emotional expressions described in Section 3.1 (see Figure 3). The emotive expressions extracted from the snippets are summed up, and the results for every emotion type are listed in descending order. This way a list of emotions commonsensically associated with the queried sentence is obtained (an example is shown in Table 4). However, Ptaszynski et al. (2008b) showed that, since emotive associations extracted from the web contain some amount of noise, Shi's technique is most efficient when only the three highest results are kept and the rest is considered to be noise.

**Figure 3** Flow chart of the web mining technique**Table 4** Example of emotion association extraction from the web (3 highest results)

Sentence: <i>Aa, pasokon ga kowarete shimatta...</i> (Oh no, the PC has broken...)		
Extracted emotion type	Type extracted/all extracted types	Ratio
Fear	28/133	0.211
Sorrow, sadness	26/133	0.195
Dislike, detestation	16/133	0.120

### 3.2.4 Evaluation of the web mining technique

To evaluate the web mining technique we used the same collection of utterances as used in the evaluation of ML-Ask. However, since the web mining technique is meant not to recognise the emotions of a particular user, but rather to find a general common sense about what emotion type should be expressed in the utterance, we used the emotions tagged by the third party evaluators as gold standard. Also, since in the morpheme modification procedure the original input utterance is dismantled and additional query phrases are created, the web mining technique is capable of finding emotive associations even when an utterance is not emotive. Therefore, we used 30 fully tagged emotive utterances from the collection used in evaluation of ML-Ask. The conditions for the result to be correct were as follows:

- at least one of the extracted emotive associations belonged to the group of emotion types tagged by the third party evaluators
- the extracted emotive associations agreed with the majority of the human taggings.

Under these conditions, Shi's web mining technique achieved 72.0% accuracy in extracting emotive associations from the web.

## 4 Contextual appropriateness of emotion verification procedure

The two systems described above are combined in the Contextual Appropriateness of Emotion Verification (CAEV) Procedure described as follows. When ML-Ask

recognises an utterance as emotive and successfully discovers a specific type of emotion, the web mining technique begins the process of verifying whether the expressed emotion type is appropriate to use in such a context. It extracts from the web all emotion types commonsensically associated with the context of the utterance. The emotions discovered by ML-Ask in user utterances are compared to the web mining procedure results and if the two match, the emotions expressed by the user are determined as appropriate for the context. In such situations, a conversational agent equipped with this method can choose a dialogue strategy to sympathise with the user. Two examples of possible conversations with emotions of different valence are presented below.

**Positive-appropriate emotion:**

*User:* I'm so happy I passed the exam!

ML-Ask : joy; web mining : joy, excitement

*Agent:* Yeah! That's great!

**Negative-appropriate emotion:**

*User:* I hate him for making a fool of me in front of everyone.

ML-Ask : dislike; web mining : anger, dislike

*Agent:* You have a reason to be angry.

However if the verification procedure indicates that the expressed emotion is inappropriate for the context, the agent could undertake different measures e.g., helping the user to manage their emotions. Two possible examples are shown below.

**Positive-inappropriate emotion:**

*User:* I'd be happy if that bastard was hit by a car!

ML-Ask : joy; web mining : fear, sadness

*Agent:* Are you sure that is what you are really feeling?

**Negative-inappropriate emotion:**

*User:* I won the prize, but I feel so bored.

ML-Ask : dislike, depression; web mining : excitement, joy

*Agent:* You should be happy!

*Two-dimensional model of affect in the CAEV procedure*

As is argued by Solomon (1993) although people sometimes misattribute the specific emotion types of other people, they rarely misattribute their valence. One could for

example confuse such emotions as anger, irritation or spite, but it is rather difficult to confuse admiration with detestation, or love with hate. Therefore we decided to check if, in cases where the specific emotion types conveyed in an utterance do not perfectly match the emotive associations, at least their general features matched. By general features we mean the two ones proposed by Russell (1980) in the theory of the two-dimensional model of affect – valence and activation.

## 5 Evaluation experiment

To evaluate the method we performed an experiment. In the experiment we used two different conversational agents.

- *Modalin* is a non-task oriented keyword-based conversational agent, which uses modality to enhance web-based propositions for dialogue. The agent was developed by Higuchi et al. (2008).
- *Pundalin* is a non-task oriented conversational agent created by combining Modalin with Dybala's Pun generating system PUNDA (Dybala et al., 2008). Pundalin therefore is a humour-equipped conversational agent using puns to enhance the communication with a user.

Thirteen users (undergraduate students from different departments) performed a ten turn conversation with both agents. All conversations were then analysed by ML-Ask. 6 out of all 26 conversations contained no specified emotional states and were excluded from the further evaluation process. For the conversations containing sentences which ML-Ask described as emotive and specified the emotion types, the web mining procedure was carried out to determine whether the emotions expressed by the user were contextually appropriate. The results provided a straightforward verification of appropriateness of the expressed emotion types. All the results were stored and a questionnaire was designed to evaluate how close they were to human thinking. One questionnaire set consisted of one conversation record and three questions inquiring:

- 1 What was the valence of emotions conveyed in emotive utterances
- 2 What were the specific emotion types conveyed in the conversation
- 3 Whether they were contextually appropriate.

Every questionnaire set was filled out by another 10 people (undergraduate students different from the users who performed the conversations with the agents). For every conversation set evaluated by 10 people we calculated how many of them agreed with the system's results.

The evaluated items were: **A)** general valence determination and **B)** specific emotion types determination accuracies in ML-Ask; and the accuracy of the system as a whole to determine the contextual appropriateness of **C)** specific emotion types and **D)** valence. We also checked how applying CVS influenced the Appropriateness Verification Procedure. The overall results of the survey are shown in Tables 5 and 6 for the baseline system and in Tables 7 and 8 for system supported with CVS.

**Table 5** The number of people that agreed with the system when analysing one agent at a time

<i>No. of people</i>	<i>Modalin</i>				<i>Pundalin</i>			
	<i>10-7 ppl.</i>	<i>6-4 ppl.</i>	<i>3-1 ppl.</i>	<i>0 ppl.</i>	<i>10-7 ppl.</i>	<i>6-4 ppl.</i>	<i>3-1 ppl.</i>	<i>0 ppl.</i>
<b>A</b>	<b>5</b>	3	2	0	<b>5</b>	2	2	1
<b>B</b>	4	<b>5</b>	1	0	<b>5</b>	1	2	2
<b>C</b>	2	<b>4</b>	3	1	2	1	<b>5</b>	2
<b>D</b>	<b>4</b>	1	3	2	2	2	<b>4</b>	2

**Table 6** The number of people that agreed with the system on all texts, including two summaries

<i>No. of people</i>	<i>Overall results</i>				<i>Summary</i>	
	<i>10-7 ppl.</i>	<i>6-4 ppl.</i>	<i>3-1 ppl.</i>	<i>0 ppl.</i>	<i>Rigorous (10-4 ppl.)</i>	<i>Optimistic (10-1 ppl.)</i>
<b>A</b>	<b>10</b>	5	4	1	75.0%	95.0%
<b>B</b>	<b>9</b>	6	3	2	75.0%	90.0%
<b>C</b>	4	5	<b>8</b>	3	45.0%	85.0%
<b>D</b>	6	3	<b>7</b>	4	45.0%	80.0%

**Table 7** Same as Table 5, but now adding CVS support

<i>No. of people</i>	<i>Modalin</i>				<i>Pundalin</i>			
	<i>10-7 ppl.</i>	<i>6-4 ppl.</i>	<i>3-1 ppl.</i>	<i>0 ppl.</i>	<i>10-7 ppl.</i>	<i>6-4 ppl.</i>	<i>3-1 ppl.</i>	<i>0 ppl.</i>
<b>A</b>	<b>5</b>	3	2	0	<b>6</b>	2	2	0
<b>B</b>	4	<b>5</b>	1	0	<b>5</b>	1	2	2
<b>C</b>	2	<b>4</b>	3	1	2	1	<b>5</b>	2
<b>D</b>	<b>4</b>	1	3	2	3	2	<b>4</b>	1

**Table 8** Same as Table 6, but now adding CVS support

<i>No. of people</i>	<i>Overall results</i>				<i>Summary</i>	
	<i>10-7 ppl.</i>	<i>6-4 ppl.</i>	<i>3-1 ppl.</i>	<i>0 ppl.</i>	<i>Rigorous (10-4 ppl.)</i>	<i>Optimistic (10-1 ppl.)</i>
<b>A</b>	<b>11</b>	5	4	0	80.0%	100.0%
<b>B</b>	<b>9</b>	6	3	2	75.0%	90.0%
<b>C</b>	4	5	<b>8</b>	3	45.0%	85.0%
<b>D</b>	<b>7</b>	3	<b>7</b>	3	50.0%	85.0%



## 6 Results and discussion

In the majority of the research on emotion processing an Achilles heel is the process of evaluation. It is usually performed on a small number of participants, sometimes even one (Tsuchiya et al., 2007; Tokuhisa et al., 2008). In our research aside from the 13 people who took part in conversations with the agents, we looked to evaluate every questionnaire set ten times.

In the majority of the research on emotion processing an Achilles heel is the process of evaluation. It is usually performed on a small number of participants, sometimes even one (Tsuchiya et al., 2007; Tokuhisa et al., 2008). In our research aside from the 13 people who took part in conversations with the agents, we looked to evaluate every questionnaire set ten times.

However, since one of our goal was to contribute to the standardisation in emotion research we allowed for two interpretations of the results. In the primary, more rigorous one, we assume that the results are correct when at least 4 participants of 10 agree with the system's results. This is a fair ratio when we consider that it means that at least four people of ten provided exactly the same results as the system, which is a difficult thing to achieve in emotion research, because of the many ambiguities that go along with expressing and perceiving emotions. Therefore, we also allow for the traditional approach, which assumes that if at least one human provides the same results as the machine, then this means that the machine and the human performed the task on a comparable (human) level.

Regardless of the approach to evaluation, the survey provided many positive results. Firstly, in most cases ML-Ask's results on affect analysis were confirmed by humans. The primary evaluation of this system by Ptaszynski et al. (2008a) was performed on a small collection of 90 utterances (see Section 3.1). The evaluation 'in the field' defined ML-Ask as fully implementational in practice. ML-Ask supported with the CVS procedure acquired 80.0% of accuracy in emotion valence recognition. Interestingly, the system's result in this matter using the optimistic perspective was 100.0%.

As an interesting remark we should add that the survey participants sometimes determined the valence and the specific types of emotions in a non-standard way, e.g., for some contexts 'fear' was determined as positive, 'joy' as negative, etc. As we assume, it is another proof that emotions are not only constituted of valence, but rather the valence of an emotion is made up by a certain context.

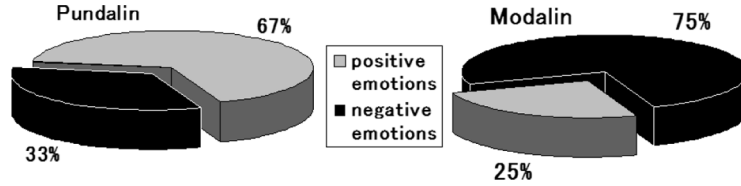
Since one of the agents was using humourous responses we also checked whether the jokes influenced the human-computer interaction. Most of the emotions expressed in the conversations with Pundalin were positive whereas for Modalin most of the emotions were negative (see Figure 4), which confirms that users tend to be positively influenced by the use of jokes in conversational agents. Some of the users even engaged in a kind of joking game and expressed their will to continue the conversation after the experiment.

As for the results of the survey, it appeared that the context appropriateness was more difficult to determine in the conversations containing puns, which is reasonable, since humour is said to be one of the most creative and therefore difficult tasks in Artificial Intelligence (Boden, 1998).

Applying the CVS analysis procedure improved the performance of the appropriateness verification procedure on the level of valence. In the rigorous

evaluation the performance was improved from 45.0% to 50.0%. In the optimistic evaluation the performance increased from 80.0% to 85.0%. Some of the successful examples as well as the ones showing the improvement after applying CVS are shown in Table 9.

**Figure 4** The total ratio of emotions positive to negative conveyed in the utterances of users with Modalin and Pundalin



**Table 9** Three examples of the results provided by the emotion appropriateness verification procedure (CAVP) with a separate display of the examples showing the improvement of the procedure after applying CVS

Part of conversation in Japanese (English translation)	ML-Ask output	web mining	CAEV	
<b>USER:</b> <i>Konpyuuta wa omoshiroi desu ne.</i> (Computers are so interesting!)	<b>Positive</b> [joy]	<b>Positive</b> [joy]	<b>Appropriate</b>	
<b>SYSTEM:</b> <i>Sore wa oishii desu ka.</i> (Is it tasty?) [about instant noodles]	×	×	×	
<b>USER:</b> <i>Oishii kedo, ore wa akita kana.</i> (Its tasty, but I've grown tired of it.)	<b>Negative</b> [dislike]	<b>Negative</b> [dislike]	<b>Appropriate</b>	
Part of conversation in Japanese (English translation)	ML-Ask baseline	CVS impr.	Web mining	CAEV
<b>SYSTEM:</b> <i>Sore wa omoshiroi tte</i> (Its so interesting!) [about conversation]	×	×	×	×
<b>USER:</b> <i>Sore hodo omoshiroku mo nakatta desu yo.</i> (Well, it wasn't that interesting.)	Positive [joy]	<b>Negative</b> [dislike]	<b>Negative</b> [fear], [sad]	<b>Appropriate</b>

With the CVS procedure applied there were as much as 8 cases of a perfect agreement with all 10 human evaluators for one questionnaire set. In conversations with Modalin ML-Ask two times reached perfect agreement in both – valence and emotion-specific determination. As for Pundalin, the perfect agreement cases appeared twice in valence determination and, which is the most important, in both kinds of determining about appropriateness of emotions.

## 7 Conclusions and future work

In this paper we introduced a method for verifying contextual appropriateness of emotions conveyed in conversations. The method, taking advantage of recent

discoveries in the field of Emotional Intelligence, presents a novel approach towards user emotion estimation in human-computer interaction. Most of the similar methods are focused on discovering certain emotion types, or more often, only the general valence of the emotions, which ignores the complexity of and the strong context dependence of emotions. However, to create a machine capable to communicate with a user on a human level, there is a need to equip it with an Emotional Intelligence Framework (Mayer and Salovey, 1997). The method described in this paper takes a step towards the practical implementation of such a framework by providing machine computable means for verifying whether an emotion conveyed in a conversation is contextually appropriate. In the proposed method a system for affect analysis is used to recognise user emotions and a web mining technique to verify their appropriateness in a particular context.

We enhanced the emotion types extraction procedure of the baseline system for affect analysis with grammatical analysis using Contextual Valence Shifters – grammatical structures including phrases such as ‘not’, ‘very much’ and ‘not much’, which determine the semantic orientation of the valence of emotive expressions. Furthermore, we applied a two-dimensional model of affect to determine which types of emotions are most probable to appear instead of the contrasted ones. ML-Ask was able to correctly handle all of the sentences containing CVS.

The contextual appropriateness of the emotions recognised by the CVS enhanced system for affect analysis was then verified by comparing them to emotive associations extracted from the web.

With a rigorous evaluation the affect analysis method was evaluated at 75.0% accuracy in determining the specific emotion types and 80.0% accuracy in determining valence of the emotions. The accuracy in determining the contextual appropriateness was 45.0% for particular emotion types and 50.0% for valence. The results, although not perfect, are very encouraging, since the same evaluation performed with a more lenient condition used popularly in the field gave results of 85.0%. Moreover there were some cases of perfect agreement with all 10 human evaluators.

An agent equipped with our system can be provided with hints about what communication strategy would be the most desirable at a certain moment. For example, a conversational agent can choose to either sympathise with the user or take precautions and help them manage their emotional states.

The theory of Emotional Intelligence, to which we refer in this paper, is a quickly developing field of research. It frequently delivers new discoveries about the structures and functions of emotions and therefore should be in focus of researchers attempting to develop means for computing human intelligence.

As our main goal, we were able to prove that computing emotions in a more sophisticated manner than simple division into positive and negative is a feasible task. Although the system as a whole is still not perfect and its components (ML-Ask and the web mining technique) need improvement, it defines a new set of goals for Affective Computing. The computing of the contextual appropriateness of emotional states is a key task on the way to full implementation of Emotional Intelligence in machines and as such is valuable to the research of Artificial Intelligence in general.

## Acknowledgements

This research was partially supported by a Research Grant from the Nissan Science Foundation and The Global Centers of Excellence Program founded by Japan's Ministry of Education, Culture, Sports, Science and Technology.

## References

- Andre, E., Rehm M., Minker, W. and Buhler, D. (2004) 'Endowing spoken language dialogue systems with emotional intelligence', *LNCS*, Vol. 3068, Heidelberg, Berlin, pp.178–187.
- Baba, J. (2003) 'Pragmatic function of Japanese mimetics in the spoken discourse of varying emotive intensity levels', *Journal of Pragmatics*, Vol. 35, No. 12, pp.1861–1889.
- Beijer, F. (2002) 'The syntax and pragmatics of exclamations and other expressive/emotional utterances', *Working Papers in Linguistics 2*, The Department of English in Lund, Lund, Sweden.
- Boden, M.A. (1998) 'Creativity and artificial intelligence', *Artificial Intelligence*, Vol. 103, Nos. 1–2, pp.347–356.
- Buhler, K. (1934) *Sprachtheorie. Die Darstellungsfunktion der Sprache*, (Theory of Language. Representative Function of Language), Ullstein, Frankfurt a. M., Wien, Berlin.
- Crystal, D. (1989) *The Cambridge Encyclopedia of Language*, Cambridge University Press, Cambridge, UK.
- Dybala, P., Ptaszynski, M., Higuchi, S., Rzepka, R. and Araki, K. (2008) 'Humor prevails! – implementing a joke generator into a conversational system', *LNAI*, Vol. 5360, pp.214–225.
- Hager, J.C., Ekman, P. and Friesen, W.V. (2002) *Facial Action Coding System*, A Human Face, Salt Lake City, UT, USA.
- Higuchi, S., Rzepka, R. and Araki, K. (2008) 'A casual conversation system using modality and word associations retrieved from the web', *Proceedings of the EMNLP 2008*, Honolulu, Hawaii, USA, pp.382–390.
- Jakobson, R. (1960) 'Closing statement: linguistics and poetics', *Style in Language*, The MIT Press, Cambridge, Massachusetts, USA, pp.350–377.
- Kennedy, A. and Inkpen, D. (2005) 'Sentiment classification of movie and product reviews using contextual valence shifters', *Workshop on the Analysis of Informal and Formal Information Exchange during Negotiations (FINEXIN-2005)*, Ottawa, Ontario, Canada.
- Kudo, T. (2001) *MeCab: Yet Another Part-of-Speech and Morphological Analyzer*, <http://mecab.sourceforge.net/> (2009.01.21).
- Lerner, J.S. and Keltner, D. (2000) 'Beyond valence: toward a model of emotion-specific influences on judgement and choice', *Cognition and Emotion*, Psychology Press, East Sussex, UK.
- MacDorman, K.F. and Ishiguro, H. (2006) 'The uncanny advantage of using androids in social and cognitive science research', *Interaction Studies*, Vol. 7, No. 3, pp.297–337.
- Mayer, J.D. and Salovey, P. (1997) 'What is emotional intelligence?', in Salovey, P. and Sluyter, D. (Eds.): *Emotional Development and Emotional Intelligence*, Basic Books, New York, pp.3–31.
- Miyoshi, T. and Nakagami, Y. (2007) 'Sentiment classification of customer reviews on electric products', *IEEE International Conference on Systems, Man and Cybernetics*, Montréal, Canada.
- Nakamura, A. (1993) 'Kanjo hyogen jiten' *Dictionary of Emotive Expressions*, Tokyodo Publishing, Tokyo (in Japanese).

- Nwe, T.L., Wei, F.S. and De Silva, L.C. (2003) 'Speech emotion recognition using hidden Markov models', *Elsevier Speech Communications Journal*, Vol. 41, No. 4, November, pp.603–623.
- Ono, H. (2002) *An Emphatic Particle DA and Exclamatory Sentences in Japanese*, University of California, Irvine.
- Oshima-Takane, Y., MacWhinney, B. (Ed.), Shirai, H., Miyata, S. and Naka, N. (Rev.) (1995–1998) *CHILDES Manual for Japanese*, The JCHAT Project, McGill University, Montreal, P.Q. Canada.
- Picard, R.W., Vyzas, E. and Healey, J. (2001) 'Toward machine emotional intelligence: analysis of affective physiological state', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 10, pp.1175–1191.
- Picard, R.W. (1997) *Affective Computing*, The MIT Press, Cambridge.
- Polanyi, L. and Zaenen, A. (2004) 'Contextual valence shifters', *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*, Stanford, USA.
- Potts, C. and Kawahara, S. (2004) 'Japanese honorifics as emotive definite descriptions', *Proceedings of Semantics and Linguistic Theory 14*, pp.235–254.
- Ptaszynski, M. (2006) *Boisterous Language. Analysis of Structures and Semiotic Functions of Emotive Expressions in Conversation on Japanese Internet Bulletin Board Forum 2 Channel*, MA, Dissertation, UAM, Poznan (in Japanese).
- Ptaszynski, M., Dybala, P., Rzepka, R. and Araki, K. (2008a) 'Effective analysis of emotiveness in utterances based on features of lexical and non-lexical layers of speech', *Proceedings of NLP2008*, Tokyo, Japan, pp.171–174.
- Ptaszynski, M., Dybala, P., Shi, W., Rzepka, R. and Araki, K. (2008b) 'Disentangling emotions from the web. Internet in the service of affect analysis', *Proceedings of KEAS'08*, Nagaoka, Japan, pp.51–56.
- Ptaszynski, M., Dybala, P., Shi, W., Rzepka, R. and Araki, K. (2009) 'Towards context aware emotional intelligence in machines: computing contextual appropriateness of affective states', *Proceedings of Twenty-First International Joint Conference on Artificial Intelligence (IJCAI-09)*, Pasadena, California, USA, pp.1469–1474.
- Russell, J.A. (1980) 'A circumplex model of affect', *Journal of Personality and Social Psychology*, Vol. 39, No. 6, pp.1161–1178.
- Rzepka, R., Higuchi, S., Ptaszynski, M. and Araki, K. (2008) 'Straight thinking straight from the net – on the web-based intelligent talking toy development', *Proceedings of SMC 2008*, Singapore.
- Salovey, P. and Mayer, J.D. (1990) 'Emotional intelligence', *Imagination, Cognition, and Personality*, Vol. 9, pp.185–211.
- Schlosberg, H. (1952) 'The description of facial expressions in terms of two dimensions', *Journal of Experimental Psychology*, Vol. 44, pp.229–237.
- Schwarz, N. (2000) 'Emotion, cognition, and decision making', *Cognition and Emotion*, Vol. 14, No. 4, pp.433–440.
- Shi, W., Rzepka, R. and Araki, K. (2008) 'Emotive information discovery from user textual input using causal associations from the internet', *FIT2008*, Fujisawa, Kanagawa, Japan, pp.267–268.
- Sjöbergh, J. and Araki, K. (2008) 'A multi-lingual dictionary of dirty words', *LREC*.
- Solomon, R.C. (1993) *The Passions: Emotions and the Meaning of Life*, Hackett Publishing, Indianapolis, Indiana, USA.
- Stevenson, C.L. (1963) *Facts and Values-Studies in Ethical Analysis*, Yale University Press, New Haven, Connecticut, USA.
- Takahashi, T., Watanabe, H., Sunda, T., Inoue, H., Tanaka, K. and Sakata, M. (2003) 'Technologies for enhancement of operation efficiency in 2003i IT Cockpit', *Nissan Technical Review*, Vol. 53, pp.61–64.

- Teixeira, J., Vinhas, V., Oliveira, E. and Reis, L. (2008) 'A new approach to emotion assessment based on biometric data', *2nd International Workshop on Human Aspects in Ambient Intelligence (HAI'08), Proceedings of the WI-IAT'08*, Sydney, Australia, pp.459–500.
- Tokuhisa, R., Inui, K. and Matsumoto, Y. (2008) 'Emotion classification using massive examples extracted from the web', *Proceedings of Coling 2008*, Manchester, UK, pp.881–888.
- Treur, J. (2007) 'On human aspects in ambient intelligence', *Proceedings of The First International Workshop on Human Aspects in Ambient Intelligence*, Darmstadt, Germany, pp.5–10.
- Tsuchiya, S., Yoshimura, E., Watabe, H. and Kawaoka, T. (2007) 'The method of the emotion judgement based on an association mechanism', *Journal of Natural Language Processing*, Vol. 14, No. 3, pp.219–238.
- Tsuchiya, N. (1999) 'Taiwa ni okeru kandoshi, iiyodomi no togoteki seishitsu ni tsuite no kosatsu', *Statistical Observations of Interjections and Faltering in Discourse* (in Japanese), SIG-SLUD-9903-11.
- Yamashita, Y. (1999) 'Kara, node, te-conjunctions which express cause or reason in Japanese', *Journal of the International Student Center*, Hokkaido University, Vol. 3 (in Japanese).

## Note

<sup>1</sup>In this paper we use italics for Japanese expressions.

## Appendix: Japanese vocabulary used in this article

### *Emotion Types*

*ai*, *aware* (sorrow, sadness, gloom)  
*an*, *yasuragi* (relief)  
*chi*, *haji* (shame, shyness, bashfulness)  
*do*, *ikari* (anger)  
*en*, *iya* (dislike, detestation)  
*fu*, *kowagari* (fear)  
*ki*, *yorokobi* (joy, delight)  
*kou*, *suki* (liking, fondness)  
*kou*, *takaburi* (excitement)  
*kyou*, *odoroki* (surprise, amazement)

### *Contextual Valence Shifters*

#### *Negations*

*amari -nai* (not quite-)  
*-cha ikenai* (Don't- [particle+verb])  
*mattaku -nai* (not at all-)  
*sukoshi mo -nai* (not even a bit-)  
*-to wa ienai* (cannot say it is-)

#### *Intensifiers*

*kiwamete-* (extremely)  
*sugoku-* (-a lot)  
*totemo-* (very much-)

### *Morphemes of causality*

*-ba*  
*-kara*  
*-kotoga*  
*-kotowa*  
*-nara*  
*-node*  
*-noga*  
*-nowa*  
*-tara*  
*-te*  
*-to*

### *Other words*

*akirameru* (to give up)  
*gitaigo* (mimetic expressions)  
*kanashimu* (to feel sad)  
*kowai* (scary)  
*maji* (no way! [interjection])  
*saa* (oh, well [interjection])  
*sugee* (great! [interjection])  
*-sugi* (over-, too much)  
*ureshii* (happy)  
*wakuwaku* (heart pounding)  
*-yagaru* (a vulgarisation of a verb)  
*-yo* (emphatic sentence ending particle)