KEER2010, PARIS | MARCH 2-4 2010
INTERNATIONAL CONFERENCE ON KANSEI ENGINEERING AND EMOTION RESEARCH 2010

EMBEDDING EMOTIONS WITHIN AUTOMATICALLY GENERATED BRAND NAMES

Denis BROUILLET[a], Anne LAURENT[b], Bénédicte LAURENT[c], Sophie MARTIN[a] and Mathieu ROCHE[b]

[a] *University Paul Valéry - Montpellier 3, France*
[b] *LIRMM – CNRS UMR 5506 - University Montpellier 2, France*
[c] *Namae Concept Corporation, France*

ABSTRACT

Brand names are everywhere and are more and more important due to the rate which new products are issued at. For instance, the French institute *INPI* (Institut National de la Propriété Industrielle) in charge of name registration collects more than two millions of new names every year. Names have been recognized to embed many characteristics of the products or company they are related to. For instance, a name like "discipline" can invoke feelings like "unhappiness". However, they have not been extensively studied, neither from the point of view of cognitive science/linguistics nor from the computer science point of view. For this reason, the automatic generation of brand names is a very difficult task, which has not been given much attention yet. In this paper, we thus propose a new method to automatically embed emotions within brand names.

Keywords: *Brand Names, Cognitive Science, Linguistics, Text Mining*

1. INTRODUCTION

It is well recognized that brand names embed many characteristics of the products or company they are related to. For instance, names like "discipline" can invoke feelings like "unhappiness".

However, although they are more and more present in our everyday life, brand names have been hardly studied from the linguistics point of view, although they are a very rich piece of language [1]. They are indeed everywhere and are of a large set of various forms: they can be taken from a dictionary, or on the contrary be a neologism, which can be very

far or very close to existing words. They can convey a meaning in one or several languages, or on the contrary be as original (even unpronounceable) as they cannot take any meaning shared by all the people/consumers.

For this reason, the automatic generation of brand names is a very difficult task, which has not been given much attention yet [2].

In this paper, we describe our method and how emotions [3] can be embedded within a name. Our method is based on a three-step process integrated in the Namae Concept platform delivered in a SaaS (Software as a Service) manner [4]:

- selection of a set of emotions and concepts,
- creation of the names,
- evaluation of the names.

In the first step, the user specifies the emotions and connotations (s)he would like to embed in the brand names (s)he is currently creating.

In the second step, we put together the concepts and emotions, and we add linguistic properties to choose letters that convey emotions and meanings. These letters have been intensively studied in many languages. For instance, in French, letters such as "k" convey the meaning of exoticism in French (as this letter is very rare in French), which is highly related to the emotion that the user will feel when seeing the brand name on a product.

The third step is performed using both computer science (distance measures) and cognitive psychology paradigms. The paradigms used allow an implicit evaluation and avoid desirability bias. Experiments are led on people to confirm/infirm unconsciously the feelings and emotions that were to be embedded in the name our method created.

This paper is organized as follows: Section 2 recalls the main studies of emotion from the cognitive point of view and of brand names from the point of view of linguistics. Section 3 introduces the automatic method used for summarizing concepts. Section 4 reports the method for creating names, while Section 5 shows how names can be studied regarding the concepts they embed. Section 6 reports experiments led to assess our method. Finally Section 7 concludes and gives the perspectives associated with our work.

2. EMOTIONS AND BRANDNAMES

In this section, we recall basic notions on emotions from the cognition point of view, before introducing how brand names have been studied in the literature of linguistics.

2.1. Emotions

As described in [3], emotion has a considerable impact on cognition whether by helping particular memories to emerge or by denying them any access to conscience. Whatever its effect might be, it is commonly agreed upon that emotion underlies and models memories; in the same way, there is a straight connection between the emotional dimension of a stimulus and the emotional dimension of the answer that is given to it.

Emotional experience expresses itself in two ways: through valence and through arousal. The valence gives the emotional value of an event on a scale that goes from “very positive” to “very negative”. The arousal is referred to as the effect of emotion that results into “arousal or stimulation” until “calming or inhibition” .

Previous research was shared out between the study of the link between the valence of events to remember and the ability to actually recover these memories, and the study of the “arousal” dimension as an explanatory factor concerning the influence of emotion on memory. The links between memory and the valence dimension on the one hand, and between memory and the arousal aspect on the other hand seem to differ from one another as one appears to help conscious encoding strategies while the other seems to affect attention in a non-conscious way.

At the time of encoding, valence would make the working out process easier, (*i.e.* finding connections between the event to remember and one's own knowledge of the world), which process reinforces memories. For instance, memorizing an event on the basis of an autobiographical processing will lead to better recall than a semantical processing of that same event. The word “spade” will be more easily memorized if it is associated with an autobiographical processing such as “leaving shelter and digging a way through the snow to go and fetch some wood”, than if it is associated with a word that is semantically connected to it, such as “pickax”.

Arousal, on the other hand, would lead to the focusing of attention on an event which would from then on be distinguished from other events, and therefore be easier to memorize. As a consequence of emotion, the attentional field of the individual is restricted, possibly focusing on particular aspects to the detriment of others.

If emotion does increase the level of recognition, this superiority appears to be relative. Indeed, just as the emotional dimension of a word increases the probability of answering “yes” to it in comparison with a non emotional word, (considering that both words have equally been shown during the learning phase), an increase of the probability of answering “yes” to a word that hasn't been shown during the learning phase can also be noticed, in a concomitant way, as long as the word has an emotional aspect. Therefore, what seems to characterize the recognition of emotional words is the reduced ability to distinguish between them: the level of wrong recognitions for emotional words is significantly higher than for neutral words.

These results lead us both to ask ourselves which are the processes that are really involved in the recovery of emotional words, and to reconsider the moment that explains the role of emotion as being the time of the recovery rather than the time of encoding or storage.

The easiness with which emotional information is recalled, in comparison with neutral information, results in an increase of answers based on the recollection of information (remember) and in the same time in an increase of answers based on familiarity (know).

The fact that familiarity can explain the results for emotional words has great heuristic value, because it means the situation in which the recollection takes place plays a role that is anything but neutral in the occurrence of the effects observed. Indeed, it is not the

emotion associated with the word that explains its recovery, but rather the fact that there is emotion at the time of the recovery created by the emotion associated to the word: it is easier to recognize a neutral word when it's shown in an emotional phrasal context than when it is shown in a neutral phrasal context.

2.2. Brand Names

2.2.1. A particular object in linguistics

Brand names have been hardly studied in linguistics, although they are an interesting kind of nouns (and expressions). Much research is indeed being done in economics and law regarding trade names used in advertising (which we would refer to as *advertising names*). However few works are published on the subject in French linguistics. These names can be either taken from a dictionary, of be neologisms, and have been often associated with proper names. However, they must be studied as a very particular piece of language, as reported in [1].

2.2.2. Linguistics and Emotions

In every language, letters, sounds ad combinations of such elements convey emotions and feelings. For instance, in French, the letter "k" conveys the meaning of exoticism because this letter is very rare. We report below the symbolic of vowels and consonants in French.

symbolic of vowels. Some vowels (both letters and sounds) and combinations can evocke:

- movement:

  speediness, quickness, lightness (dynamic concepts) versus slowness, heaviness

- shape:

  sharp / round form

- luminosity:

  light, clarity versus night, obscurity

symbolic of consonants. Some consonants (both letters and sounds) and combinations can evocke:

- luminosity: Figure 1 depicts the degree of luminosity of consonants.

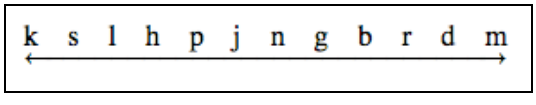

Figure 1: Luminosity and consonants: from clarity to obscurity

- hardness, shortness
- smoothness, contiguousness

In our work, we thus use some letters or other ones in order to help for creating names that convey particular feelings (see Figure 2).

3. SUMMARIZING CONCEPTS AND EMOTIONS

In this section, we show how engineering techniques can be used to try and capture emotions embedded withing brand names.

3.1. Text Mining

In [2], we have shown how sequential patterns can help to describe the frequent patterns occurring in brand names. Sequence patterns have been studied for the last 15 years. As a data mining method, they allow to automatically and efficiently extract patterns of the form <(A,B)(C)>, x% meaning that x% of a given set of objects contain A together with B and then C. For instance, they can be used to extract patterns from a given set of texts such as "in 23% of the texts, there is a sentence containing the words followed some sentences *marketing* and *Paris* further by a sentence containing the word *brand*".

Applied to brand names, such patterns are for instance of the form <(k)(y)(u)> meaning that trade names contain the letter *k* followed some letters further by the letter *y* and then letter *u*. Note that these letters are not mandatorily consecutive.

In the framework of emotions and feelings, they can be used to describe the patterns (sequence of symbols) that characterize a given set of names, for instance the set of names that have been pointed as "evoking fear" by the users.

3.2. n-grams

n-grams refer to sequences of letters and symbols [7]. They are used to extract n consecutive characters from words. n often ranges from 1 to 6, depending on the application, but is often set to 3 for NLP (Natural Language Processing) purposes [8,9]. An n-gram of size 1 is often referred to as a *unigram,* while *bigrams* refer to n-grams of size 2 and *trigrams* n-grams of size 3. They are used for genetic sequence analysis, or for language recognition systems, etc. For instance, the trigrams extracted from "house" are {hou, ous, use}.

In the framework of this work, n-grams are generated in relation with the concepts that are to be evoked. For instance, if *Love* must be evoked, we collect, either automatically (e.g. using the Internet and NLP pre-processing tools) or manually, all words related to this concept, and we decompose these words into the associated n-grams (n=1 to 6). Note that this principle is used for several kinds of concepts (e.g. user concepts, languages).

4. CREATING NAMES

In this section, we study how to embed emotions and concepts within names. This is achieved using the techniques described above:

1. embedding n-grams that are related to the emotions that must be conveyed,
2. inserting letters that are symbols of some emotion/feeling,
3. using patterns that were shown to embed particular emotions/feelings.

In order to embed n-grams, we mix them in different manners, using algorithms that were developped in collaboration with linguists and computer scientists. Roughly speaking, two methods are applied: randomized mix and guided mix. These methods cannot be described here in details as they are part of a commercial tool [4].

Note that n-grams can be replaced by syllables. We are currently studying to which extend n-gram and syllables convey meanings and emotions.

Inserting letters is achieved by embedding these letters either, and by removing the letters conveying the opposite feelings.

Figure 2 shows the tool proposed by [4] for creating such names. Here the feelings amour (love) and peur (fear) have been chosen together with some letters to evoke luminosity-clarity.

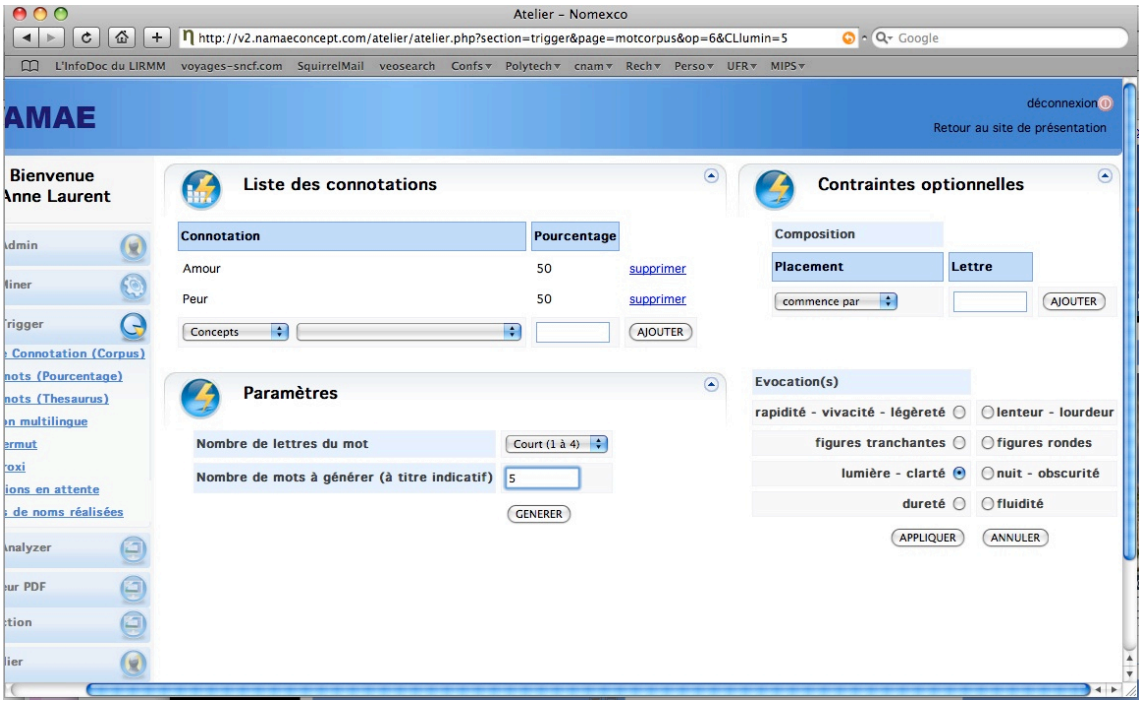

Figure 2: Namae Concept [4]: Generating a name evoking Love, Fear and luminosity-clarity

Finally, patterns extracted using data mining techniques can be used. For instance, if we discovered that brand names that were recognized to convey exoticism are of the form <*k*…*y*…*u*> then names will be sharped in order to contain a *k* then some other letters, then a *y* then some other letters and then a *u*.

5. EVALUATING NAMES

In this section, we study how to evaluate to which extent a concept is embedded in a name. For this purpose, we evaluate to which extent a given name embeds the n-grams related to the emotions-concepts that must be studied. This computation is performed using classical similarity distances [8,9,10].

We report below two distances. The interested reader can refer to [9] for a complete review of the existing distances.

5.1. Tri-gram based Measure

This measure aims at comparing two words based on the trigrams they share. For instance the words "chat" and "chaton" contain the trigrams tr(chat)= {cha, hat} (|tr(chat)| = 2) and tr(chaton)= {cha, hat, ato, ton} (|tr(chaton)|=4). In this example, two trigrams are shared: tr(chat) ∩ tr(chaton) = {cha,hat} and |tr(chat) ∩ tr(chaton)|=2.

In order to calculate the similarity degree, the following formula is used [8]:

$$Tri(word1,word2)=$$

$$1/(1+|tr(ch1)|+|tr(ch2)|-2)*|tr(ch1)\cap tr(ch2)|$$

Referrring back to our example, we have:

$$Tri\ (chat,chaton)= 1/(1+2+4-2)*2 = 0.33$$

5.2. String Matching

String matching [10] relies on the so-called *Edit Distance* (denoted by E). The Edit Distance is the minimal sum of the cost of the operations that are required to transform one string to the other one. The operations being considered are: suppression, insertion and substitution of characters. For example, two insertions are required (characters "o" and "n") to transform "chat" to "chaton". We have thus

$$E\ (chat,\ chaton) = 2$$

String Matching (denoted by Str) considers the Edit Distance:

$$Str(ch1,ch2) =$$

$$max\ (\ 0,\ (min(|ch1|,|ch2|)-E(ch1,ch2))\ /\ min(|ch1|,|ch2|)\ )$$

We thus have:

```
Str (chat, chaton)= max(0, (4-2)/4)= 0.5
```

6. EXPERIMENTS

In this section, we report experiments led to assess our method. For these experiments, we created 135 names using the method described above:

- 45 names were created by combining *neutral* names (*e.g.,* table),
- 45 names were created by combining *positive* names (*e.g.*, beauty),
- 45 names were created by combining *negative* names (*e.g.,* snake).

For instance, names like *avate* (neutral), *joliesse* (positive) and *∂auchema* (negative) were created. We then presented these names to users who were asked . This presentation was done three times (evaluation 1, evaluation 2, evaluation 3).

Figure 3 reports the percentage of matching between the category of the name (neutral/positive/negative) and the category the user felt. Note that the higher the number of evaluations, the higher the matching and the faster the response (see Table 1).

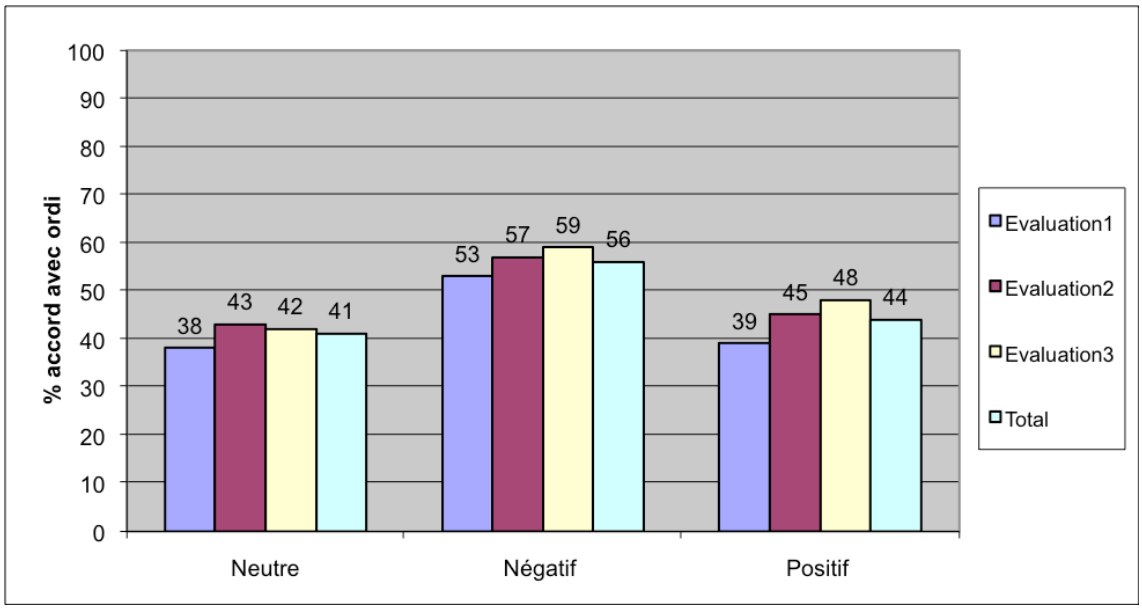

Figure 3: Percentage of matching between the user feeling and class predicted by the computer

From the results presented in Figure 3, it can be deduced that:

- The method is relevant because the percentages of matching are always greater than random (*i.e.* 33%, which corresponds to a random choice of one answer out of three possible classes).

- The negative words seem easier to detect (the matching scores are better).

Valence	Evaluation 1	Evaluation 2	Evaluation 3
Neutral	1093.61	923.98	833.96
Negative	1093.55	900.72	826.40
Positive	1038.20	952.55	867.54

Table 1: Response Time

7. CONCLUSION AND PERSPECTIVES

In this paper, we study how to embed concepts, feelings and emotions within brand names by combining methods and tools from cognitive science/psychology, linguistics, and computer science. This topic is not only of great interest for commercial reasons, but also to help people to memorize and use products. For instance, we aim at using these techniques to create efficient names for products intended for elderly people.

REFERENCES

1. Laurent B., *Nom de marque, nom de produit. Production de sens en discours publicitaire.* PhD Thesis. 2006.
2. Fiot C., Laurent A., Teisseire M. and Laurent B., Why Fuzzy Sequential Patterns can Help Data Summarization: An Application to the INPI Trade Mark Database, *In the proceedings of the Fuzz'IEEE Conference 2006*.
3. Blanc N., Syssau A. and Brouillet D. *Emotion et cognition quand l'émotion parle à la cognition*. collection Concept-psy. 2006.
4. Namae Concept. *http://www.namaeconcept.com*
5. Kronlund A. and Whittlesea W.A., Remembering After a Perception of Discrepancy: Out With the Old, in With the Two. *Journal of Experimental Learning, Memory, and Cognition*, Vol. 32, No. 5, 1174 –1184. 2006.
6. Kronlund A. *Remembering words and brand names after a perception of discrepancy*. PhD Thesis. Simon Fraser University. 2006.
7. Manning C., Schütze H., *Foundations of Statistical Natural Language Processing*, MIT Press: 1999.
8. Lin D., An Information-Theoretic Definition of Similarity. *In Proc. 15th International Conf. on Machine Learning*. Morgan Kaufmann, San Francisco, CA, pp296-304, 1998.
9. Kefi H., ontologies et aide à l'utilisateur pour l'interrogation de sources multiples et hétérogènes, université de paris 11. *PhD Thesis*, Université de paris 11, 2006
10. Maedche A. and Staab S., Measuring similarity between ontologies. *In proceedings of the European Conference on Knowledge Acquisition and Management - EKAW*. pp251-263. 2002.