

A Semantic Lexicon for Emotions and Feelings

Yvette Yannick Mathieu

Laboratoire de linguistique formelle – CNRS
UMR 7110, Université Paris 7
CP.7031, 2 place Jussieu
75251 Paris cedex 05 – France
ymathieu@linguist.jussieu.fr

Abstract

In this paper, the *FEELING* system for extracting and building semantic representations of emotions, feelings and psychological states expressed in texts is presented, as well as the linguistic knowledge used by the system. Semantic representations are described by means of a set of feature structures. Letters to the Editor is taken as a domain for the evaluation of this work.

Introduction

A semantic lexicon in the field of feeling, emotion and psychological states is presented here, as well as *FEELING*, a software using this lexicon to provide an interpretation and to generate paraphrases. A prototype-based organization of this lexicon with heritage mechanism, intensity and antonymy graphs, and a linguistic knowledge database was already proposed in Mathieu (2000).

This approach differs from Wordnet (Fellbaum 1998) on two points. First, the organization principles are different. For instance, there are no heritage or intensity relationships in Wordnet. Secondly, Wordnet is a large database resource, with no associated knowledge.

The only similar work on French is Mel'čuk's dictionary (1992), where each word description is made up of its linguistics properties and lexical rules. My perspective is rather different, I focused on a very detailed description in a semantic field with a software associated, while Mel'čuk study is about general language and no system using the dictionary knowledge.

1. Semantic Lexicon Description

I studied 950 French words for emotions and psychological states. Among them, 600 are verbs, like *aimer* (to love), *effrayer* (to frighten), *irriter* (to irritate), and 350 are nouns, like *amour* (love), *colère* (anger), etc. I propose a semantic classification, in which verbs and nouns are split into 38 semantic classes, according to their meaning. My hypothesis is that the language reflects the way one conceptualizes the world, and that words with close meanings have similar formal behaviors. My classification is partially similar to that of Johnson-Laird

and Oatley's (1989) and Wierzbicka's (1996) works on feelings. In this paper, I present my work on verbs.

1. Verbs

French verbs for feelings, and psychological states occur in two kinds of structures, as illustrated in (1) and (2), respectively.

- (1) *Paul irrite Mary*
Paul irritates Mary
- (2) *Mary hait Paul*
Mary hates Paul

These structures differ by the syntactic position of the person (*Mary*), called the “experiencer”, who has the feeling or the emotion. In (1), the experiencer is the complement, and the subject (*Paul*) is the cause of her feeling (*irritation*). In (2), the experiencer is the subject and the complement *Paul* is the object of her feeling (*hate*). About 500 verbs belong to the first category and 100 to the second one. Since the linguistic behaviors of these verbs are very different, two separate analyses have been made. Here, I describe verbs of the first category, such as *irriter* (to irritate). All these verbs were divided into 33 semantic classes (cf. Table 1), each class including verbs with the same meaning. From a prototypical point of view, one verb represents and can replace all the verbs in a class, and it names this class. For instance, the class IRRITER contains verbs like to aggravate, to irritate, to exasperate, etc.

There are three categories of verbs :

a -) Verbs which mean the experience or the causation of a rather unpleasant feeling, such as irritation, fear or disappointment. They are divided into eighteen classes, like IRRITER, EFFRAYER (to frighten), ATTRISTER (to sadden), etc.

b -) Verbs which mean the experience or the causation of a rather pleasant feeling, such as interest or fascination. They are divided into thirteen classes like INTERESSER (to interest), FASCINER (to fascinate), etc.

c -) Verbs which mean the experience or the causation of a feeling that is neither pleasant nor unpleasant, like astonishment or indifference. They belong to two classes :

ÉTONNER (to astonish) and INDIFFERER (to be indifferent to).

AIGRIR	ATTRISTER	DÉCEVOIR
DÉCONCERTER	DÉGOÛTER	DÉMORALISER
DÉRANGER	EFFARER	EFFRAYER
ENDURCIR	IRRITER	FROISSER
INHIBER	LASSER	MEURTRIR
OBSÉDER	RÉVOLTER	TRACASSER
CALMER	DÉSARMER	DISTRAIRE
ÉMOUSTILLER	ÉMOUVOIR	ÉPATER
FLATTER	INTÉRESSER	PASSIONNER
RASSURER	SATISFAIRE	FASCINER
VIVIFIER		
INDIFFÉRER	ÉTONNER	

Table 1. Semantic classes of French psychological verbs

2. Relationships between semantic classes

Semantic classes are linked by meaning, intensity and antonymy relationships, represented with simple graphs.

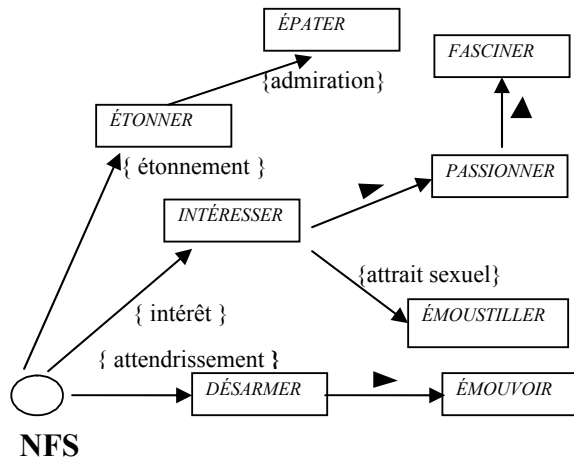


Figure 1. Meaning and Intensity relationships between "pleasant" classes

Intensity graphs, one for pleasant feelings (Figure 1), one for unpleasant ones (Figure 2), are connected graphs, oriented according to intensity of experienced feeling. A connected graph contains meaning relationships (shown partially on Figure 1 and 2). A graph contains antonymy links between classes (a subset is given in Figure 3). There is a "no feeling state" represented by a white circle noted NFS. Labeled arcs join this state to semantic classes. For example, the arc labeled "intérêt" (interest) joins the no feeling state to the INTERESSER class. These arcs can be labeled by more specific features like "admiration", or by intensity degree represented by the symbol \blacktriangleright .

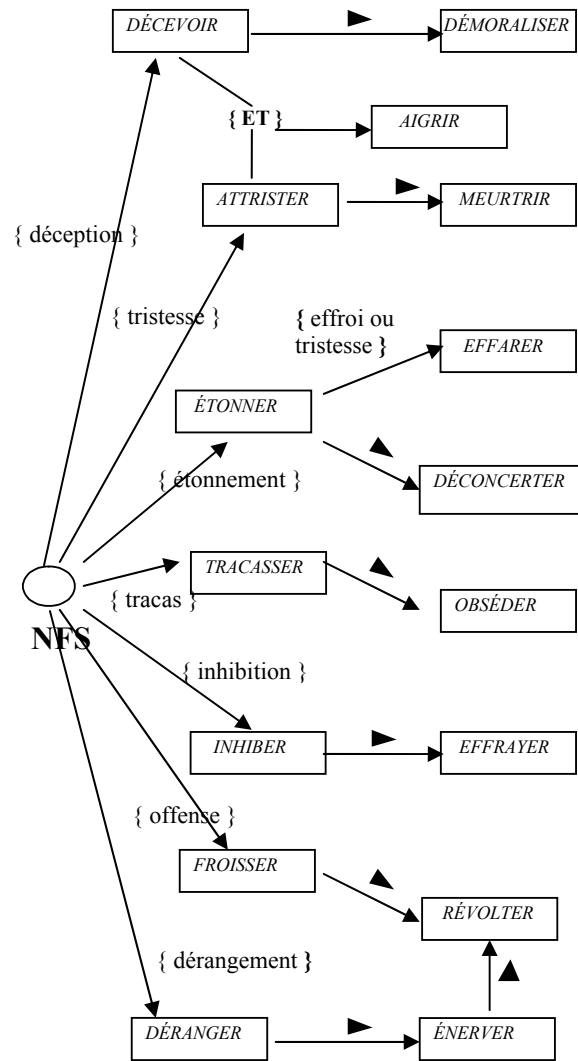


Figure 2. Meaning and Intensity relationships between "unpleasant" classes

For example, an interest increase on INTERESSER class verbs (to interest, to attract, to tempt, to entice, etc.), is described by PASSIONNER (to inspire passion) class verbs (to devour, to excite, to overexcite, to electrify, to fire, to carry away, to enthuse, etc.).

A stronger emotion of passion is reflected by FASCINER class verbs (to fascinate, to intoxicate, to hypnotize, to mesmerize, etc.).

As is shown in these two graphs, the meaning of ÉTONNER class verbs can be pleasant, unpleasant or neutral, depending on the context.

The antonymy between classes is represented with an arc and the symbol \leftrightarrow . As an example, (Figure 3), the IRRITER class is antonymous with the CALMER (to calm) class, meaning that each verb in the first class is antonymous with at least one verb in the second class (to mellow, to calm down, to relax, etc.) and vice versa.

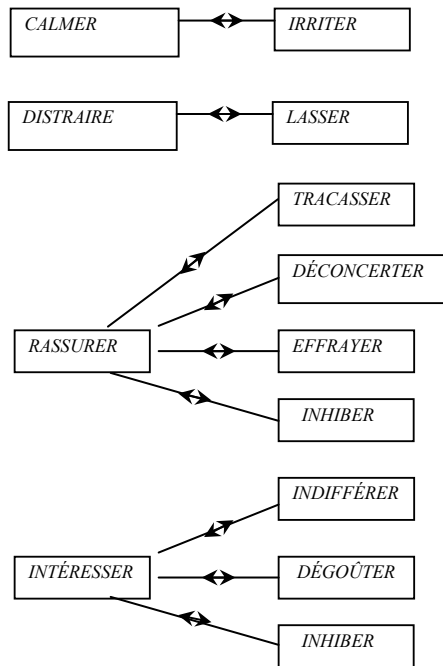


Figure 3. Subset of Antonymy relationships between semantic classes

3. Linguistic Properties of Verbs

I studied a set of linguistic properties of verbs and classes. Among them, 15 are binary and easy to formalize, while others are more complex such as arguments selection or arguments structure. I describe here a subset of these two types of properties, for a full description see Mathieu (2000).

3.1. Simple properties. A subset of these properties for IRRITER class is shown in Table 2a and 2b. The first row represents the whole class (e.g. *IRRITER*), then each following row is a verb (*agacer*, *asticoter*, etc.), and each column is a property where *V* is a Verb, *N₀* is the subject, *N₁* is the object, *V-inf* is an infinitive and *Vpp* is *V* past participle.

A plus sign indicates that a verb accepts the property. The property [the subject is agentive/ non agentive] indicates that *Luc effraye Marie* (Luc frightens Marie) has two possible meanings : that Luc frightens Marie because he wants it, or he frightens her unintentionally, by his behavior or his appearance, or something else. This property is called *N₀ agentif* (*N₀ agentive*).

Some verbs have a psychological meaning only, like *aimer*, while others, such as *irriter*, have two meanings: one which is "basic" (3), and one psychological by metaphor (4) :

- (3) *Le soleil irrite Mary* (sa peau)
The sun irritates Mary (her skin)
- (4) *Paul irrite Mary* (par son comportement)
Paul irritates Mary (by his behavior)

This property is called "Emploi non psychologique". For each verb (noted *V*), a nominalization (or deverbal, noted *V-n*) is associated, if it exists, as in *irriter/irritation*. This noun denotes the feeling or emotion felt by the experiencer.

		<i>V-n éprouvé/ressenti par N₁ = déverbal</i>	<i>V-n exercé par N₀</i>	<i>N₀ agentif</i>	<i>N₁ V</i>	<i>N₁ se V de ce que P</i>	<i>N₁ est Vpp de ce que P</i>
IRRITATION	IRRITER	-					
	agacer	<i>agacement</i>	-	+	-	+	-
	asticoter		-	+	-	-	+
	courroucer	<i>courroux</i>	-	+	-	+	+
	crisper	<i>crispation</i>	-	+	-	+	+
	énervé	<i>énervement</i>	-	+	-	+	+
	enquiquiner		-	+	-	-	-
	enrager	<i>rage</i>	-	+	+	+	-
	exaspérer	<i>exaspération</i>	-	+	-	+	+
	excéder		-	+	-	+	-
	fâcher	<i>fâcherie</i>	-	+	-	+	+
	hérisser		-	-	-	+	+
	horripiler	<i>horripilation</i>	-	-	-	-	-
	impatier	<i>impatience</i>	-	-	-	+	+
	irriter	<i>irritation</i>	-	+	-	+	+
offusquer		-	-	-	-	-	
stresser	<i>stress</i>	-	+	+	-	-	
ulcérer		-	-	-	-	-	

Table 2a. Simple properties

Intensity relationships link verbs inside each class. Thus, *exaspérer* (to exasperate) and *irriter* are in the same IRRITER class, but *exaspérer* is stronger than *irriter*. These internal intensity relationships between verbs are described by the property "Intensifieur" with the value "neutre" (neutral) for *irriter*, and the value "beaucoup" (high) for *exaspérer*.

3.2. Complex properties. Where simple properties are attribute-value pairs, others are more complex such as arguments selection and arguments structure. The experiencer is always a person, like Mary in sentence (1) *Paul irrite Mary*, or some metonymical expressions referring to a person. The form of these expressions can be "Det N de Nhum" (Determinant Noun of Human noun). There are three main categories of noun N, according to how distant the metonymy is from the person (Nhum) ; it can be (1) a body part (or "soul part") like *coeur* (heart), (2) a feeling or quality name (anger, vanity, etc.), or (3) a noun such as *hopes*, *convictions*, etc.

		N_1 est Vpp de N_0	N_1 est un Vpp	Complément en par obligatoire	Intensifieur	Aspect duratif	Emploi non psychologique
IRRITATION	IRRITER	+	-	-	-	-	-
	agacer	+	-	-	neutre	-	-
	asticoter	+	-	-	neutre	-	-
	courroucer	+	-	-	beaucoup	-	-
	crisper	+	+	-	neutre	-	+
	énervé	+	+	-	neutre	-	+
	enquiquiner	+	-	-	neutre	-	-
	enrager	+	-	-	beaucoup	-	-
	exaspérer	+	-	-	beaucoup	-	-
	excéder	+	-	-	beaucoup	-	-
	fâcher	+	-	-	neutre	-	-
	hérissier	+	-	-	beaucoup	-	+
	horripiler	+	-	-	beaucoup	-	-
	impatienter	+	-	-	beaucoup	-	-
	irriter	+	-	-	neutre	-	+
	offusquer	+	-	-	neutre	-	-
	stresser	+	-	-	neutre	-	-
	ulcérer	+	-	-	beaucoup	-	+

Table 2b. Simple properties

Some verbs select nouns of the first category only, like *briser* (to break) in *Mary a brisé le coeur de Luc* (Mary broke Luc's heart). Some verbs select nouns of the second category also, like *apaiser* ou *calmer* in *La chanson a apaisé/ calmé la colère de Mary* (The song calmed down Mary's anger), while others accept all nouns denoting psychological states, such as *satisfaire* (to satisfy) in *Les paroles de Mary ont satisfait les espoirs/ la curiosité de Luc* (Marie's words satisfied Luc's hopes / curiosity). Moreover, 110 verbs like *irriter* or *déconcerter* (to disconcert) accept non strictly human complement, although with difficulty.

The semantics of the verb and its arguments constrains the possibility of paraphrases. Thus, when the subject of the active form is a human name, as *Paul* in *Paul déçoit Luc* (Paul disappoints Luc) or *Paul énerve Luc* (Paul irritates Luc), the passive sentence with "de" is possible only for some verbs. One can say *Luc est déçu de Paul* (Luc is disappointed of Paul) but not **Luc est énervé de Paul*¹ (Luc is irritated of Paul). These properties are described by means of 40 rules represented in declarative format [Si (if) condition **Alors** (then) conclusion].

II. FEELING System

The *FEELING* system has been built in order to demonstrate the validity of my formalization. It contains four main components: (1) a communication interface, (2) a semantic knowledge database, (3) an inheritance

¹ "*" before a sentence means the sentence is ungrammatical.

mechanism, and (4) a knowledge treatment module.

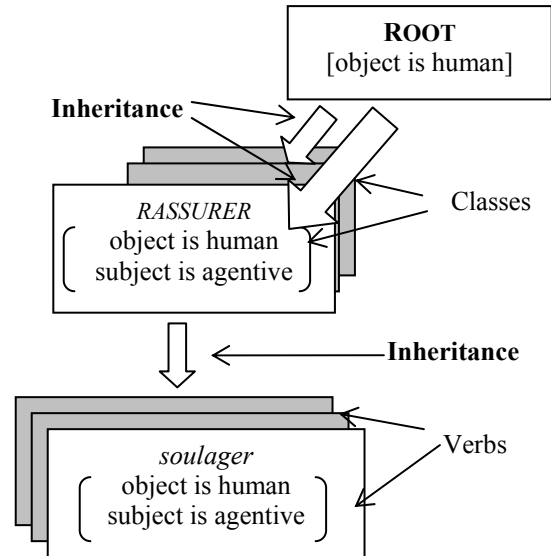


Figure 4. Inheritance

Here, I describe the knowledge representation and the inheritance mechanism. A complete description of *FEELING* may be found in Mathieu (2000).

The semantic database contains the class and verb knowledge, represented in prototype-based format (Rosch 1975, Kleiber 1990) and the graphs describing the relationships between classes. A root prototype contains the knowledge shared by all verb classes, like [the object is human]. An analogy is drawn between semantic classes and prototypes. Each verb is derived and inherits properties and rules from one prototype.

An example is given in Figure 4. The *RASSURER* (to reassure) prototype (or semantic class), which has the [the subject is agentive] property, inherits [the object is human] property from the root prototype. Each *RASSURER* class verb, like *calmer* (to calm), *rassurer*, *soulager* (to relieve), etc. inherits these two properties.

In order to process a sentence, each token has to be described with its grammatical category (noun, verb) and function (subject, object).

For example, *Luc exaspère Marie* is analyzed in:

{ Subject : *Luc*, human proper noun ;
Verb *exaspérer* ; Object : *Marie*, human proper noun }.

We rely on Fabre & Bourigault (2001) to get such syntactic information.

Given this sentence, *FEELING* provides an interpretation which describes what kind of feeling is expressed (I) :

I. INTERPRÉTATION (Meaning)
Le sentiment est éprouvé par *Marie*
(The feeling is experienced by *Marie*)

La cause est *Luc* (The cause is *Luc*)

Le sentiment éprouvé est *plutôt désagréable*
(The experienced feeling is **rather unpleasant**)

Le sentiment éprouvé est *l'exaspération*
(The experienced feeling is **exasperation**)

Then, *FEELING* provides synonymous and antonymous sentences (II et III):

II. PHRASES SYNONYMES : (Synonymous sentences)

Luc énerve beaucoup Marie
Luc (courrouce / enrage / excède / hérisse /
horripile / impatiente / ulcère) Marie
Marie est exaspérée par Luc

III. PHRASES ANTONYMES (Antonymous sentences)

Luc apaise Marie
Luc calme Marie

and generates several well formed paraphrases (IV):

IV. PARAPHRASES

Marie éprouve de l'exaspération (envers / contre) Luc
Marie éprouve beaucoup d'énervement) (envers /
contre) Luc
Marie est (exaspérée / très énervée)
Marie est (exaspérée / très énervée) par Luc
Marie éprouve de l'exaspération
Marie éprouve beaucoup d'énervement
Marie s'exaspère

Finally, *FEELING* provides sentences expressing increased and/or decreased feeling (V):

V. VARIATION DE L'INTENSITÉ du sentiment exprimé (Feeling intensity variation)

Diminution continue de l'intensité du verbe *exaspérer*
(<<) :
(Continuous decreasing intensity of *exaspérer* verb)

Luc (agace / crispe / énerve / fâche / irrite) Marie

<< *Luc (dérange / désoblige / emmerde) Marie*

<< *Luc (ennuie / ennuie / gêne / importune*
/ incommode / indispose) Marie

Augmentation continue de l'intensité du verbe *exaspérer*
(>>) :
(Continuous increasing Intensity of *exaspérer* verb)

>> *Luc (braque / choque / écœure / scandalise) Marie*
>> *Luc (/ indigne / rebelle / révolte) Marie*

III. Evaluation

An evaluation of this work is being conducted in two ways. First, a comparison between usual dictionaries and the *FEELING* lexicon was made. A large part of the lexicon, more than 10 per cent, contains words which are not described by usual dictionaries as psychological nouns or verbs, like *bétonner* (to concrete) which means *endurcir* (to harden) whereas these words have been found in texts with a psychological meaning.

Secondly, newspaper "letters to the Editor" will be analyzed by both *FEELING* and human readers. A first protocol is in progress: a corpus (50 letters) is going to be classified by *FEELING* and by three human evaluators. Those evaluators are instructed to classify sentences as reflecting a "pleasant", "unpleasant" or "neutral" feeling, and to give an intensity level of this feeling, on a scale from 1 to 3, with 1 being weak feeling, 2 being medium and 3 being strong. A comparison with *FEELING* results will be made with standard metrics, like recall and precision. Evaluation results will be available at the beginning of 2004.

Conclusion

The present work demonstrates the interest of a systematic linguistic study of verbs, and of a modeling allowing its exploitation. It shows the necessity to take intensity and antonymous relationships into account in order to analyze and interpret feeling in text.

References

- Fabre, C. and Bourigault, D. 2001. "Linguistic clues for corpus-based acquisition of lexical dependencies". Proceedings of the Corpus Linguistics 2001 Conference, UCREL Technical Papers, vol 13. Lancaster University, pp. 176-184.
- Fellbaum, C. ed. 1998. *Wordnet. An electronic Lexical Database*. Cambridge, Mass.: MIT Press.
- Johnson-Laird, P.N., and Oatley, K. 1989. "The Language of Emotions : An Analysis of a Semantic Field". *Cognition and Emotion* 3, pp. 81-123.
- Kleiber, G. 1990. *La sémantique du prototype*. Paris: puf, collection Linguistique nouvelle.
- Mathieu, Y. Y. 2000. *Les verbes de sentiment*. Paris: CNRS Edition.
- Mathieu, Y. Y. 1999. "Les prédicats de sentiment". *Langages* 136, pp. 41-52.
- Mel'čuk, I. 1992. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques III*. Montréal: Les Presses de l'Université de Montréal.
- Rosch, E. 1975. "Cognitive Representations of Semantic Categories". *Journal of Experimental Psychology* 104/3.
- Wierzbicka, A. 1996. *Semantics, Primes and Universals*. Oxford: Oxford University Press.