Collecting and Evaluating Lexical Polarity with a Game With A Purpose

Mathieu Lafourcade
LIRMM
Univ. Montpellier, France
lafourca@lirmm.fr

Nathalie Le Brun
Imagin@t
34400 Lunel, France
imaginat@imaginat.name

Alain Joubert
LIRMM
Univ. Montpellier, France
aj@lirmm.fr

Abstract

Sentiment analysis from a text requires amongst others having a polarity lexical resource. We designed LikeIt, a GWAP (Game With A Purpose) that allows to attribute a positive, negative or neutral value to a term, and thus obtain a resulting polarity for most of the terms of the freely available lexical network of the JeuxDeMots project. We present a quantitative analysis of data obtained through our approach, together with the comparison method we developed to validate them qualitatively.

1 Introduction

Being able to evaluate feelings is essential in natural language processing, whether to analyze political speeches or opinion of the general public on the provision of services, tourist, cultural, or about consumer goods. Whatever type of approach, statistics supervised or more linguistic one (Brun, 2011), such ability requires referring to a polarity lexical resource, in wich terms are endowed with positive, negative and neutral values. The polarity can be expressed using a single numerical value (Taboada *et al.*, 2011), or more: two values (positive/negative) are used in Emolex (Saif and Turney, 2013), a lexical polarity / feelings resource in English, produced by crowdsourcing (using Amazon Mechanical Turk, which can be problematic, see Fort *et al.* (2014)). SentiWordNet (Esuli and Sebastiani, 2006), as well as WordNet Affect (Strapparava and Valitutti, 2004) are extensions of WorldNet in which terms are polarized along three values (positive, negative, objective) of which the last one is opposed to the first two. Approaches through propagation by starting from a manual core (see Gala and Brun (2012) and Lafourcade and Fort (2014)) have also been performed, but such approaches may not exactly reflect the views of speakers. Learning algorithms and compositional approaches can use such polarity data (Kim and Hovy, 2004) and (Turney, 2002). The GWAP (Game With A Purpose) JeuxDeMots (JDM) (Lafourcade, 2007) resulted in a lexical network in constant expansion, in which terms are linked by lexical-semantic relations. Contributive approaches with non-experts have been analysed in (Snow *et al.*, 2013) and proven quite efficient. Within the JDM project, some alternative type of games allow to validate and verify lexical relations produced through the main game (Lafourcade *et al.*, 2015). The JDM project thus provides suitable context to test various methods of polarity information acquisition.

It may be relevant to assign to words some information in the form of finite sets of values. Thus, the polarity can be defined by three values: positive, negative and neutral. It may be noticed that many semantic features can be characterized in this way, i.e. associated with such variable-sized sets of values: feelings/emotions (anger, fear, joy, love, sadness ...), or colors (red, blue, yellow, green, orange, violet, black, white ...). Since this type of association cannot be obtained through the main game of the JDM project, we designed several other games for characterizing the words according to various criteria (I like/I don't like, associated feeling, associated color...). Applications of these data are numerous, either in discourse analysis or disambiguation. But such an annotation is complex because it is subjective and heavily influenced by the context: for example, the same remark can be considered a trait of humor, an advice, a criticism or a reprimand ... according to the enunciator, the interlocutor and context.

In this article, we firstly introduce LikeIt, a GWAP designed to collect polarization data, and how the polarization of the terms spreads within the lexical network. Then, we present the results obtained through a quantitative and qualitative analysis. Our method of qualitative assessment, based on a comparison between the polarity data and the feelings data (i.e. feelings that people spontaneously associate with a given term) is described in detail. Finally, we discuss the prospects that this work allows to consider.

2 LikeIt, a Polarity Game

Similarly to social networks, the game is to assign the assessment *I like*, *I do not like* or *I don't care* to a displayed term. Of course, this assessment is not only very subjective, but closely linked to the context. However, we hypothetise that many words have intrinsic polarity that it is possible to gather, by asking enough people. A majority polarity may emerge from answers, and if so, we can verify which one.

2.1 How Does it Work?

The player has to answer *yes*, *no* or *I do not care* to the question *do you like the idea of* followed by a term. This framework seems to be the most flexible and most comprehensive way to enrich the lexical network with polarity information. This allows in particular to distinguish between the terms for which people are mostly indifferent (majority of *I do not care, neutral* polarity) and those that raise sharply divided opinions (roughly equal amounts of *yes* and *no*, polarity equally divided between *positive* and *negative*). Within the context of word sense disambiguation, preliminary results show that polarity is sufficient for selecting the correct meaning of a term in about 50% of cases. Polarity data may also be used in opinions analysis, by combining the polarities of highly polarized terms (i.e. those whose highest polarity is greater than 50% of the cumulative values of the three possible polarities). Figure 1 shows screenshots of LikeIt game. Among the qualities that make this vote game by consensus an effective GWAP (Lafourcade *et al.*, 2015), it can be emphasized:

- **simplicity**: although the response procedure (*yes*, *no*, *I don't care*) is identical to that of surveys, diversity of vocabulary and topics is such that people do not feel they complete a survey. In addition, the response by a simple click makes possible to play from a smartphone or tablet, to pass the time during relatively short waiting situations (waiting room, queue, transport, ...). Quantitatively, very short games and immediate rerun make likeIt a very effective game to collect data.
- **diversity of vocabulary and topics and variability of response**: a number of words elicit mixed feelings, even opposing (e.g., the term *operating room*, theoretically seen as positive, but negative if we are personally concerned), and the feelings of a player can evolve over time and according to circumstances. Thus the word *bachelor* or *school exam* creates a negative feeling among high school students, but significantly positive for graduates. The choice to provide some very general vocabulary makes it interesting and varied game.
- **reactivity**: as soon he answered, the player can see the percentage of people who share his opinion, which may induce some emotions about the fact of being or not like everyone. Direct feedback is thus given, while the game is immediately rerun with a new question.

2.2 The Collected Data

For each word, the responses of players generate a triplet of values representing the number of votes for each of the three possible polarities. Their percentage distribution represents what we call the polarization of the term, similar to a three-component vector, whose norm can be calculated. The higher are the intensity (i.e. the number of votes for the word) and the vector norm, the more reliable the polarization is. The minimum intensity from which the polarization can be considered as reliable is difficult to define because various factors are involved, but we can estimate the minimum number of votes as 20 times the number of poles (i.e. at least 20 votes for a monopolarity, 40 votes for bipolarity, and 60 votes for a tripolarity). A word is highly polarized when one of the three values is greater than 50%. Table 1 shows some examples of different polarizations, with corresponding intensities and norms.

Although this is out of the scope of this article, we should note that considering polarization as a vector of the three polarities (with possibly null polarities) lead us to very interesting manipulation, comparison and combination possibilities pertaining to vectors and to norms. Basically, the Manhattan norm is the count of votes, and the p-norm with $p = 2$ is the Euclidian norm (the one mentioned here). Roughly speaking, in the context of sentiment analysis of a given text, combining polarizations of contained words means adding such normed vectors.

Figure 1: Two consecutive screenshots of LikeIt. Further to the answer given in the left screen (*barrier*), the player immediately can see at the top of the next screen (right image and bottom zoom), the percentage of players who share his view: the game thus provides a feedback to the player while being immediately rerun with a new question.

Term	Distribution of polarities (%)	Intensity
gift	**POS: 82** NEUT: 14 NEG: 4	Nb votes : 280 norm : 232.73
retirement	POS: 48 NEUT: 18 NEG: 34	Nb votes : 303 norm : 190.60
policemen	POS: 29 NEUT: 15 **NEG: 56**	Nb votes : 274 norm : 177.45
autumn	POS: 37 NEUT: 44 NEG: 18	Nb votes : 277 norm : 168.34

Table 1: Examples of polarizations obtained with LikeIt: the term *gift* is strongly positively monopolarized, while others show a more heterogeneous distribution. The norm value is the norm of the vector composed of the values of positive, neutral and negative polarities. The higher the norm, the more confident we can be in the the representativeness of the polarity distribution.

2.3 The Term Selection Algorithm

A very large proportion of words having a neutral overall polarity, if we randomly select the terms within the network, the game may be monotonous and boring for the player. In addition, the network includes highly specialized terms, which is interesting if you know the term, but discouraging otherwise. For these reasons, the terms are selected within the network via a propagation algorithm whose principle is:

- A term T whose neutral polarity represents less than 50% in the distribution of the three polarities is selected randomly;
- The proposed word is either T, with a probability *p* of 0.5, either a neighbor N that is randomly selected among neighbors of T, with probability *1-p*;
- In order to accelerate the propagation, the probability *p* is changed under various conditions (empirically determined). If the total number of votes is under 30 (resp. over 300, over 1000) for N, then $p = 0.25$ (resp. 0.75, 0. 9);
- The propagation algorithm was initiated by manually assigning a positive polarity to the term *good* (1 *positive* vote) and a negative polarity to the word *bad* (1 *negative* vote).

This simple algorithm performs within the network a propagation between the words for which polarity information is relevant, i.e. those that are not strongly neutral, and this while partially avoiding the terms that have already a lot of polarity votes. Thus, a neutral term will be mostly selected through its neighbors. A highly linked term (to other terms in the network) will be polarized more quickly than others as it will be more often reached through neighboring (at least as long as the number of votes remains below the set threshold).

2.4 Observed Experimental Biases

A first bias observed is due to polysemy: for a polysemous term, it is possible that the player's response is influenced by an anecdotal sense, but strongly negative or positive. Thus, polarization of *vache* (cow), whose the dominant sense, the animal, is broadly neutral (or even slightly positive) can be influenced by the meaning *vache (méchant)* (nasty), which is strongly negative. Indeed, the players, who are *de facto* in a polarity context, think at first to the most polarized sense and thus they assign it a negative polarity. It is the same for *fumier* (dominant sense: manure, substituted sense: insult), *cellule* (cell) (dominant sense: biology, substituted sense: jail cell), etc. However, the refinements of these terms show a polarization consistent with that expected. Some terms are bipolarized because the polarity can vary depending on the diegetic perspective (the player identifies with the character and is involved in the context) or extradiegetic (he adopts an external perspective). Thus *dragon, orc, vampire, witch,* etc. are both negative (diegetic perspective) and positive (extradiegetic perspective). It is a second bias. A third bias, which tends to favour the positive polarization, is explained in the next section.

3 Evaluation of Polarity Data

3.1 Quantitative Evaluation

During the first three months, more than 25,000 terms were polarized (i.e. characterized with information on the polarity) with a total of over 150,000 votes. Within 3 years, more than 385,000 words were polarized with more than 100 million votes. The network containing about 490,000 words, we see that about 75% were reached by the propagation algorithm [1] .

283 votes per term on average
178 votes per positive polarity on average
88 votes per neutral polarity on average
83 votes per negative polarity on average
120 votes per polarity on average

Table 2: Quantitative data of polarity obtained with LikeIt: the average number of votes for terms and polarities.

We can clearly see on table 2 that the average number of positive votes is higher than neutral and negative votes altogether. Hence players seem more reluctant to vote neutral or negative than positive. This could be interpreted as an analogy to what happens on social networks, where people are invited to click *like* to show their approval, but where there is no way to indicate that one do not like, disapprove, or even just is indifferent. Thus, it is possible that many people unconsciously behave in a "socially correct" way, i.e. giving only positive opinions and passing over terms that would generate a negative one. We should note however than the mean number might not always be a good indicator of the distribution of the votes, especially when the distribution roughly follows a power law. The median value is certainly more meaningful.

As regards the global distribution of polarities (table 3), there is a slight predominance of neutral polarity, which is not surprising. Although the algorithm is designed not to offer too many neutral terms, current vocabulary still remains predominantly neutral. On the other hand, the positive polarities are almost twice as high as the negative ones, which may be explained in different ways, in addition to the above assumption.

Data in tables 3 and 5 thus appear to be biased towards the positive polarity that represents 55% of votes. Indeed, interviewing the players, it turns out that many terms rather perceived as neutral (e.g. *Odonata*) are often labeled positively. The bias seems to be the result of the adage that "I love what I do not hate". It is difficult to assess the impact of such a bias because the terms that would be positive or neutral are not known *a priori*, but this effect would be in addition to the one mentioned above (reluctance to express a negative opinion) to explain the strong predominance of positive votes.

The positive bias can also be explained as an effect of the term selection algorithm : the proposed terms are mostly named entities or words in fields which usually arouse approval : thus the vast majority of famous people are perceived rather positively, especially actors and actresses; in the same way, named entities of works (films, paintings, novels. . .) mainly generate positive feelings, as well as most of the culinary vocabulary , especially names of culinary specialty, of drinks . . .

The distribution of polarities according to number of votes in table 4 has a median value around 80 (it means that there are so many polarities with a number of votes lower than 80, as of polarities with a number of votes higher than 80). It is quite enough votes for being statistically meaningful. We could consider that at least 20 votes are

[1] all data are freely available at the following url in real-time: url anonymized

336,461 positive polarities	(37.1 %)
373,403 neutral polarities	(41.1 %)
198,103 negative polarities	(21.8 %)
907,967 polarities	(100 %)

59,943,107 positive votes	(54.9 %)
32,879,876 neutral votes	(30.1 %)
16,332,188 negative votes	(15 %)
109,155,171 votes	(100 %)

Table 3: Quantitative data obtained with LikeIt: distribution of polarities (left) and votes (right). We can see the distribution of polarities does not stick exactly to the distribution of votes. There is a majority of neutral polarities but a majority of positive votes.

needed to define a representative polarity; 811,666 polarities are above this threshold (89% of all polarities).

The average (120 see table 2) is shifted to the right due to a number of relations with a very high number of votes. These are the "hub" terms of the network, *i.e.* the very general terms, which are connected to several tens of thousands of words. For instance, the term *animal* has more than 26,052 outcoming relations. Such terms are more often proposed than less connected words, and thus rapidly collect a large number of votes.

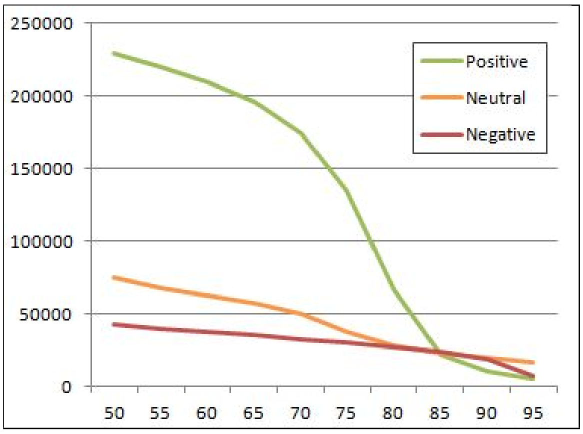

Figure 2: This figure shows the distribution of the majority polarities ($> 50\%$). Such polarities are largely positive and are twice the number of neutral and negative polarities altogether. However for higher distributions (from 60% to 80%) number of positive drops sharply.

In table 5 we see that the dominating polarities combinations are: positive/neutral bipolarity (43%, not necessarily with the same weight) and "positive/neutral/negative" tripolarity (42%, not necessarily evenly distributed). For the first, it confirms that people tend to vote either neutral or positive, or more precisely to vote positively even if they are rather indifferent. Conversely, a negative vote would truly reflect a marked opinion. For the second, it indicates that many words arouse an opinion shared, although in these polarities distributions there may be a strong dominance of one among the three. This distribution also shows that unanimity is rare: only 6.4% shows a single polarity. Figure 3 is cumulative and shows that there are many (in proportion) negative and neutral polarities with a low number of votes, and significantly more positive polarities over 200 votes. This is consistent with the hypothesis mentioned above : people seem more likely to vote positively than negatively or neutrally. In figure 4, the distribution of polarities according to their weight (linear and log) shows that over approximately 400 votes, negative polarities are more numerous than others. This is due to the presence in the network of very negative "hubs" : hightly connected words for which the vote is almost always negative, as *death, illness, accident, cancer* ...Ups and downs are a consequence of the structure of the network, the algorithm and the fact that players can pass over, all combined.

6,835 terms with positive polarity only	(1.8 %)
13,006 terms with neutral polarity only	(3.4 %)
4,388 terms with negative polarity only	(1.1 %)
167,396 terms with positive/neutral polarity only	(43.4 %)
627 terms with positive/negative polarity only	(0.2 %)
31,563 terms with neutral/negative polarity only	(8.2 %)
161,752 terms with positive/neutral/negative polarity	(42 %)
385567 terms with at least one polarity	(100 %)

Table 5: Quantitative data of polarity obtained with LikeIt: the distribution of terms according to mono, bi or tripolarization.

3.2 Qualitative Evaluation Method

The problem of the qualitative evaluation of our data is complex insofar as there is no lexical resource of polarity to which the polarity data from LikeIt could be compared. A manual assessment which would be to check the relevance of the polarity assigned to a number of terms is unthink-

53,881 polarities < 10 votes	854,086 polarities ≥ 10 votes
96,301 polarities < 20 votes	811,666 polarities ≥ 20 votes
222,988 polarities < 40 votes	684,979 polarities ≥ 40 votes
➤ 468,072 polarities < 80 votes	439,895 polarities ≥ 80 votes ◀
639,269 polarities < 160 votes	268,698 polarities ≥ 160 votes
863,043 polarities < 320 votes	44,924 polarities ≥ 320 votes
907,261 polarities < 640 votes	706 polarities ≥ 640 votes
907,926 polarities < 1280 votes	41 polarities ≥ 1280 votes

Table 4: Distributions of polarities depending upon the number of votes which median value is around 80, the median for each polarity being given in table 3.

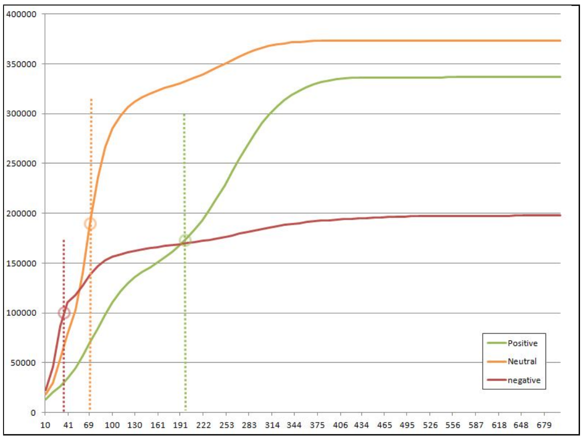

Figure 3: Cumulative number of polarities according to the number of votes (weight). The median values concerning the negative and neutral polarities (resp. 36 and 70) are significantly lower than the median value for positive polarities (about 200).

able due to the data size. In addition, how would we select the terms to be checked? Within the project JDM, two games allow associations between terms and the feelings they evoke: terms relative to feelings can be proposed openly via a text field in the main game, and in a semi-open way (chosen by clic or given through free answer in advanced mode) in Emot game (Lafourcade *et al.*, 2015) (url anonymized) .

So, for each term, we get a list of weighted associated feelings as follows:

- **gift**: joy (1712)(+); surprise (1142)(+); happiness (980)(+); love (780)(+); pleasure (741)(+); friendship (660)(+); gratitude (310)(+); disappointment (260)(-); amazement (222)(+); gratefulness (210)(+); generosity (200)(+); satisfaction (160)(+); contentment (140)(+); enjoyment (120)(+); desire (100)(+); embar-rassment (90)(-); emotion (81)(+); delight (80)(+); impatience (70)(-); jealousy (70)(-); happy (60)(+); party (50)(+); liking (50)(+); frustration (50)(-); awkwardness (50)(-);
- **policeman**: security (1027)(+); fear (1007)(-); violence (817)(-); hatred (357)(-); apprehension (297)(-); anger (186)(-); strength (137)(*); protection (127)(+); repression (127)(-); insecurity (127)(-); anxiety (117)(-); revolt (117)(-); insecurity (127)(-); injustice (97)(-); brutality (97)(-); panic (97)(-); respect (97)(+); terror (87)(-); aggressiveness (117)(-); fury (87)(-); distrust (87)(-); worry (77)(-); pain (77)(-); reject (77)(-); blue funk (67)(-); blindness (66)(-); mistrust (65)(-); shame (63)(-); incomprehension (57)(-); distress (57)(-); relief (57)(+); fright (32)(-); disquietude (32)(-);
- **arm**: strength (110)(*); protection (100)(+); support (80)(+); union (5)(-); indifference (4)().

The terms concerning feelings were the first to be reached by the propagation algorithm, so they are polarized. In the list above, for each feeling term, following the weight of the relation, a symbol in brackets indicates the majority polarity (which accounts for over 50% of votes) or the absence of a dominant polarity. The (+) corresponds to a positive dominant polarity, (-) indicates a negative dominant polarity, () a predominantly neutral polarity, and (*) indicates the absence of a majority polarity. We so notice that the term *strength*, associated with *arm* and with *policeman* does not present any majority polarity.

A polarization can thus be calculated for a term, by making the sum of polarity vectors of every feeling term associated, and it can be compared to that stemming from the LikeIt game. We compare then a polarity inferred to a polarity directly

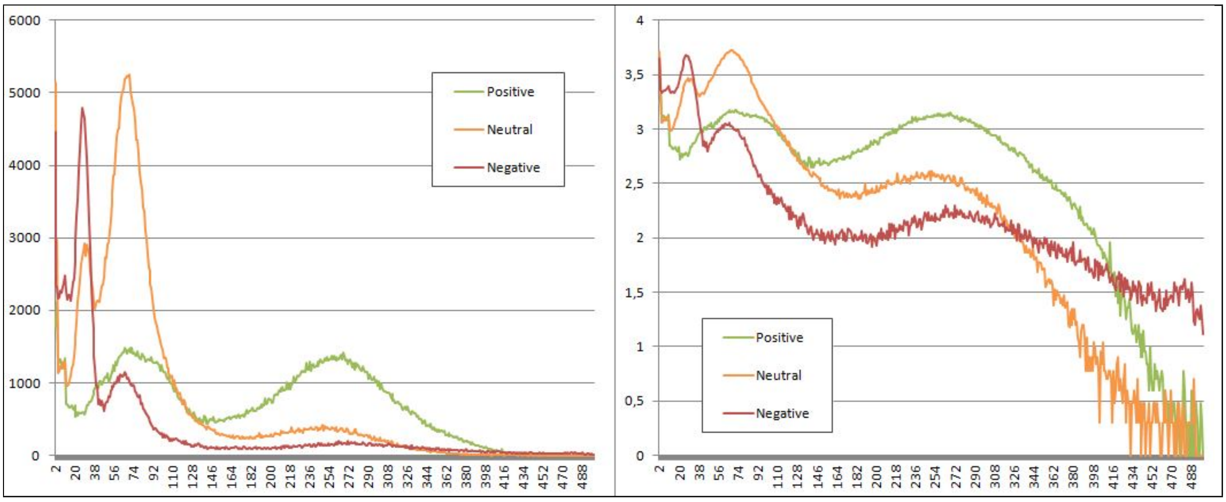

Figure 4: Distribution of polarities (on the left) and of the log of polarities (on the right) according to their number of votes (weight). The number of polarities above 500 votes is low (see table 4) - they are not shown on these figures.

established by the players. This is done via a *cosine* measure and a measure of the *max* (*max*=1 if both dominant polarities coincide). The advantage of such an approach is that it can be automated, so we can reserve the effort of manual inspection for divergent cases. We calculated the *cos* and *max* values, and ordered the first 5,000 terms by decreasing weights for the feelings relation (thus the most often played for this relation at the first).

The average of the maximal polarities from the game (*mpa*) can be seen as the maximum rate of agreement reached on average by the general opinion, for the *n* most played words. Between 1,000 and 5,000 first most played words, the difference between *mpa* and unanimity (100-*mpa*) varies between 15 and 12 %: it seems logic that the number of divergent opinions increases with the number of votes. The manual review of cases of divergence (*max* = 0) shows that they mainly concern the terms that can be perceived from a diegetic or extradiegetic perspective, such as:

thesis, earwig, analysis, moray, micropenis, woman [agent-of] express something, dragon, custard pie attack…

To associate feelings with a given term, the player seems to get a diegetic perspective, while he adopts an external one (extradiegetic perception) to assign one polarity to a given word with LikeIt. Indeed, all the cases of difference concern words polarized negatively via the associated feelings, and positively via LikeIt. Note that the highly polarized words are not concerned by the perspective diegetic / extradiegetic. Moreover, we emphasize that the terms that elicit the most subjectivity of opinion display a heterogeneous polarity, but its distribution into *positive/negative/neutral* is consistent in both modes of assessment.

4 Conclusion and Future Work

Our results and the method we developed to characterize the polarity through various GWAP allow to consider a number of perspectives. First, it is to continue the double approach (polarity inferred from associated feelings, and polarity directly assigned through the game LikeIt) to further expand the already abundant lexical resource of polarity (385,000 words with a polarity information as a freely available resource).

Then, our approach can be extrapolated: indeed, all types of characteristics (size, temperature, weight / balance, temporality, location ...) may be characterized and quantified using crowdsourcing through GWAP. But a preliminary study to identify the most useful and informative has necessarily to be undertaken, to avoid boring and thus demotivating the players by multiplying this type of games. Note that the data generated through these games, that require only knowledge and a good command of language, are of good quality, which justifies this approach.

It is also necessary to keep in mind that the polarities data are not static but potentially fluctuating, especially in time, and depending on the circumstances. For example, the term *volcano* rather arouses curiosity or indifference, but when an im-

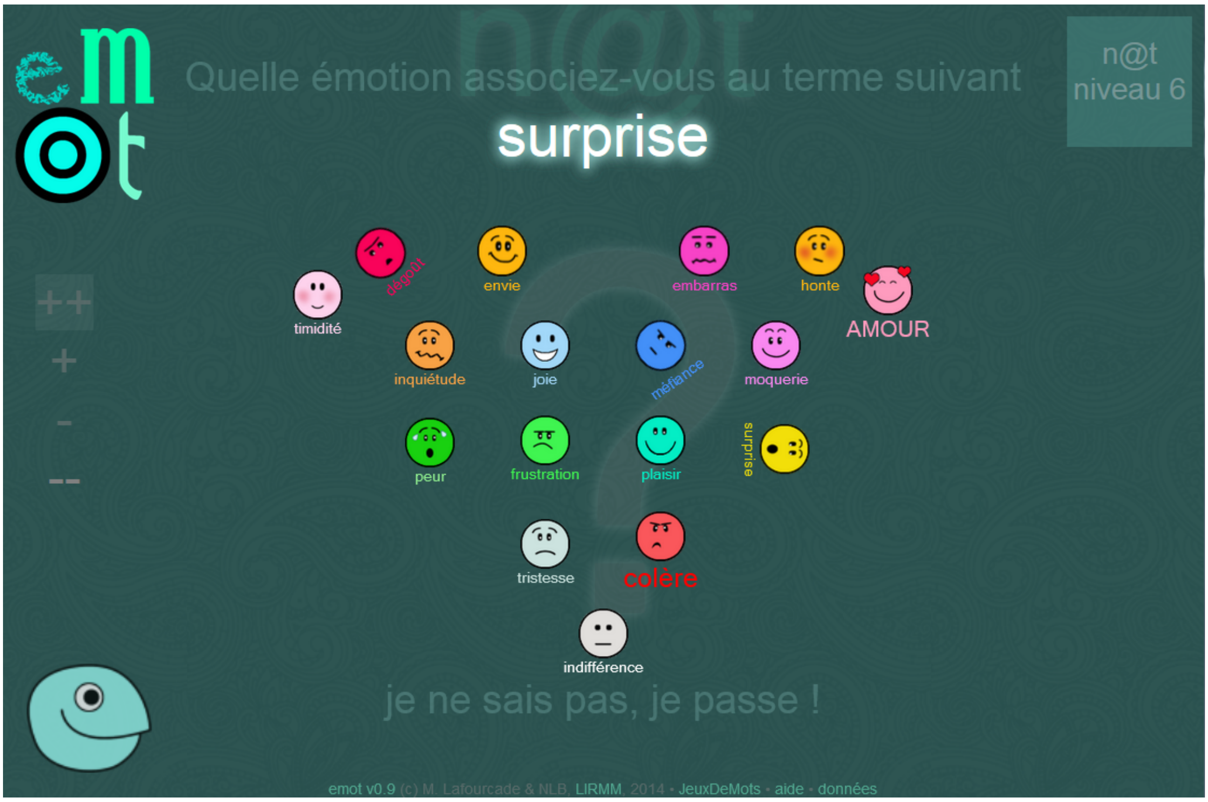

Figure 5: A screenshot of the Emot game. The player is invited to choose one associated feeling aroused by the word *surprise*. The data obtained with Emot allow us to cross-evaluate those obtained with LikeIt.

n first terms	**Cos average**	**Max average**	**Maximal polarities average**
1 000	0.80	0.76	85.65 %
2 000	0.83	0.79	86.55 %
3 000	0.80	0.75	87.40 %
4 000	0.82	0.77	87.49 %
5 000	0.83	0.79	87.63 %

Table 6: Qualitative assessment of polarization data from LikeIt compared with those calculated from the associated feelings. There is a significant correlation between the polarization defined by LikeIt and that induced by the associated feelings.

minent eruption threatens populations or air traffic, anxiety and fear become the majority among the feelings expressed. Similarly, feelings about a celebrity, or a work (named entities) can be very fluctuating over time, and if contradictory feelings appear for the same word in the network, introduce a notion of context may be interesting, for example *DSK [context] IMF*, and *DSK[context] Sofitel*.

We could polarize the words automatically, based on their relations within the network: for example, the relation *characteristic* is very polarizing; *widow [characteristic] sad* allows to assign a negative polarity to *widow*. However, the crowdsourcing approach is generally more reliable and faster, both for highly monopolarized words and those whose polarity is more heterogeneous.

The approach and tools presented in this article are relatively new, and the number of polarized terms represents a significant proportion (70%) of the entire network. It can be assumed that the most interesting common words are those which are the most played in JeuxDeMots, hence the most appropriately linked to other words, as claimed in (Chamberlain *et al.*, 2006). As our propagation algorithm selects the vast majority of such terms, we may conclude that our approach allows to effectively polarize them. Given the results, we reckon we have demonstrated the feasibility, the interest and the perspective of our project, and broadly undertook to build the corresponding resource.

References

Agarwa C. and Bhattachary P. 2006. Augmenting Wordnet with Polarity Information on Adjectives. *3rd International Wordnet Conference*, Jeju Island, Korea, South Jeju (Seogwipo), 2006, 7 p.

Brun C. 2011. Detecting opinions using Deep Syntactic Analysis. *Proceedings of Recent Advances in Natural Language Processing*, (RANLP 2011), Hissar, Bulgaria, pp. 392-398.

Chamberlain J., Fort K., Kruschwitz U., Lafourcade M. and Poesio M. 2013. Using Games to Create Language Resources: Successes and Limitations of the Approach. *Theory and Applications of Natural Language Processing.* Gurevych, Iryna; Kim, Jungi (Eds.), Springer, ISBN 978-3-642-35084-9, 2013, 42 p.

Esuli A. and Sebastiani F. 2006. SentiWordNet: a publicly available lexical resource for opinion mining. *Proceedings of LREC-06*, Gêne, Italie, 6 p.

Fort K., Adda G., Sagot B., Mariani J. and Couillaut A. 2014. Crowdsourcing for Language Resource Development: Criticisms about Amazon Mechanical Turk Overpowering Use. *Lecture Notes in Artificial Intelligence*,Springer, pp. 303-314, 2014, 978-3-319-08957-7.

Gala N. and Brun C. 2012. Propagation de polarités dans des familles de mots : impact de la morphologie dans la construction d'un lexique pour l'analyse d'opinions. *Actes de Traitement Automatique des Langues Naturelles*, (TALN 12), Grenoble, juin 2012, pp. 495-502.

Kim S. and Hovy E. 2004. Determining the sentiment of opinions. *Proceedings of COLING- 04*, (TALN 12), Barcelone, Espagne, pp. 1367-1373.

Lafourcade M. 2007. Making people play for Lexical Acquisition. *Proceedings of 7th Symposium on Natural Language Processing*, Thailand, 13-15 December 2007, 8p.

Lafourcade M., Le Brun N. and Joubert A. 2015. Jeux et intelligence collective – résolution de problèmes et acquisition de données sur le Web. *Collection Science cognitive et management des connaissances (sous la direction de Joseph Mariani et Patrick Paroubek)*, ISTE éditions, 2015, 156 p.

Lafourcade M. and Fort K. 2014. Propa-L: a semantic filtering service from a lexical network created using Games With A Purpose. *Proceedings of the Ninth International Conference on Language Resources and Evaluation*, Reykjavik, Islande, 26-31 May 2014, pp. 1676-1681.

Saif M. and Turney P. 2013. Crowdsourcing a Word-Emotion Association Lexicon. *Computational Intelligence*, 29 (3), pp. 436-465.

Snow R., OConnor B., Jurafsky D., and Y. Ng A. 2008 Cheap and Fast But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, Honolulu, Association for Computational Linguistics, pp. 254263.

Strapparava C. and Valitutti A. 2004. WordNet Affect: an affective extension of WordNet. *Proceedings 4th International conference on Language Resources and Evaluation*, (LREC-04), Lisbon, Portugal, pp. 1083-1086.

Taboada M., Brooke J., Tofiloski M., Voll K. and Stede M. 2011. Lexicon-based methods for sentiment analysis. *Computational Linguistics,* Volume 37 (2), pp. 267-307.

Turney P. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Proceedings of ACL-02*, Philadelphia, USA, pp. 417-424.

Vegnaduzzo V. 2004. Acquisition of subjective adjectives with limited resources. *AAAI spring symposium on exploring attitude and affect in text: Theories and Applications*, Stanford, USA.