Toward a cognitive organization for electronic dictionaries, the case for semantic proxemy

Bruno Gaume, Karine Duvignau, Laurent Prévot, Yann Desalle
Université de Toulouse, CNRS
{gaume,duvignau,prevot,desalle}@univ-tlse2.fr

Abstract

We compare a psycholinguistic approach of mental lexicon organization with a computational approach of implicit lexical organization as found in dictionaries. In this work, we associate dictionaries with 'small world' graphs. This multidisciplinary approach aims at showing that implicit structure of dictionaries, mathematically identified, fits the way young children categorize. These dictionary graphs might therefore be considered as 'cognitive artifacts'. This shows the importance of semantic proximity both in cognitive and computational organization of verbs lexicon.

1 Introduction

According to (Dik, 1991) a linguistic theory should be compatible with psycholinguistic research on language acquisition, treatment, production, interpretation and memorization of linguistic expressions. We agree with this view and postulate that elaborating electronic dictionaries on the ground of a linguistic theory, satisfying Dik's principle, will confer them good ergonomics that will increase their usability. Our approach is to some extent comparable to WordNet initiative (Fellbaum, 1998), in the sense that we are trying to characterize speakers' mental lexicon.

In this paper, we focus on verb lexical organization through the examination of verbal pivot metaphorical utterances (VPMU). Such utterances involve an understudied structural aspect of the lexicon: *interdomain co-hyponymy* (Duvignau, 2002; Duvignau and Gaume, 2008). In this context, we take semantic proximity as a central principle for cognitive ergonomics influencing dynamic lexical acquisition and adult lexical organization. VPMU generally consists in substituting elements from different semantic domains. They are usually considered as deviants while they might constitute a linguistic illustration of the categorial flexibility advocated in (Piaget, 1945; Ny, 1979; Hofstadter, 1995). They might therefore reveal an early lexical structuring mode that may form a ground for improving electronic dictionaries.

This paper presents a mathematical method able to discover the areas in which this structuring mode appears in dictionaries. Our approach is to take advantage of the mathematical structure of the network generated by verb definitions. This structure has been mentioned in (Watts and Strogatz, 1998), studied for WordNet by (Sigman and Cecchi, 2002), refined in (Gaume et al., 2002) and exploited in the current proposal.

The paper is organized as follows. The next section brings evidence of categorization by semantic proximity from early lexicon acquisition experiments. Section 3 presents the computational model, hereafter *'proxemy'*. Section 4 details our work on lexical graphs while section 5 compares the results of experimental studies with those of the computational model.

2 Toward a categorization by semantic proximity: evidences from early lexicon acquisition

In order to show the importance of semantic approximation, we have chosen to support our claim with productions observed at the crucial period of lexical construction (between 2 and 4 years-of-

Coling 2008: Proceedings of the workshop on Cognitive Aspects of the Lexicon (COGALEX 2008), pages 86–93
Manchester, August 2008

age) and to compare these with adult speakers that have a stabilized lexicon.

2.1 Inter-domains vs. intra-domain semantic approximations

Studies in this field are almost exclusively limited to nominal utterances. (Duvignau et al., 2005) established the existence of the production of metaphor-like utterances with a verbal pivot in 2-4 years-old children and proposed to consider them, at this stage of language development, as semantic approximations and not as mistakes or true metaphors. Duvignau distinguished two kinds of semantic approximations: Inter-domains proximity and intra domain proximity between verbs (Duvignau, 2002).

- Inter-domains proximity / co-hyponymy between verbs : a 'linguistic approximation'

(1) Elle déshabille l'orange (She undresses the orange) [Age: 3 years] [movie: a lady peels an orange]

In this category of approximation, the verb used by the speaker constitutes a reference to a semantic domain different from the one of element it is combined to *('undress' / 'orange')*. For this reason, the approximate character of the verb is understandable independently of the context of the utterance: detecting the approximation occurs at the linguistic level. We call this type of production 'semantic approximation'. They might constitute a metaphor or an 'analogic surextention'.

When someone has a conventional verb in the mental lexicon *('to peel')* and use a non conventional but relevant verb like *'to undress the orange'* for the action *[to peel the orange* his verbal semantic approximation constitutes a metaphor. On the contrary when someone does not have a conventional verb in the mental lexicon but manages to use a non conventional but relevant verb in saying *'to undress'* for this action, his verbal semantic approximation constitutes a 'surextension' but not an error because of the lexical relation that links these verbs. In fact, according to (Duvignau and Gaume, 2008) *'to undress'* and *'to peel'* are related by an inter-domains synonymic relation.

- Intra-domain proximity / co-hyponymy between verbs: a 'pragmatic approximation' In this category, illustrated by (2) the approximate character of the verb comes only from a non-correspondence between the verb used and the reality it designates. This happens with utterances in which the use of the verbal form does not create any semantic tension within the utterance but designates a way of carrying out an activity that does not correspond precisely to the action undertaken.

(2) Elle coupe l'orange (She cuts the orange)[age: 3 years][movie: a lady peels an orange]

We propose an experimental study of the production of verbal semantic approximations like (2) or (1) by way of a naming task of 17 action-movies with young children (from 2 to 4 years old). We compare their performances with adult's ones.

2.2 Experimental Design

In order to elicit the production of semantic approximations we proposed to all our participants an action-video naming task. The population sample consisted of:

- 54 non-disturbed children (2-4 years old), monolingual in French
- 77 non-disturbed young adults (18-40 years old), monolingual in French

The action movies sequences are coming from the *Approx* protocol (Duvignau et al., 2005). The material consists in 17 action-movies sequences described in table 1.

The 17 action movies are presented in random order to each participant. Instructions were given at the time the action in the movie was completed and its results were visible (e.g when the glass is broken). At that moment a question was asked to the participant:*'What did the woman do? (just now)'*

2.3 Results

Each of the children produced between 2 and 5 approximations: *'Elle casse la tomate' -She breaks a tomato '* [action = to squash], *'Elle épluche le bois' - She peels the wood'* [action = to strip the bark off a log]. Globally, children produced semantic approximations for 34 % of the naming tasks, which were distributed as follows: 24 % intra-domain semantic approximations, 10 % inter-domains semantic approximations. They produced them significantly more frequent than adults : 5 % with 4% intra-domain semantic approximations and 1 % inter-domains semantic approximations.

/DAMAGE/	/TAKE OF /	/SEPARATE/
1- burst a balloon 2- screw up a piece of paper 3- break a glass with a hammer 4- squash a tomato with the hand 5- tear up a newspaper	6- peel a carrot with a peeler. 7- peel an orange with one's hands. 8- strip the bark off a log 9- undress a doll 10- take apart a lego structure. 11- peel a banana	12- make bread-crumbs by hand. 13- slice bread with a knife. 14- break up bread with one's hands 15- shred parsley with a knife. 16- saw a wooden plank 17- tear up a shirt

Table 1: *Approx* 17 action movies

The *student Test* shows the difference between children and adults in terms of production of semantic approximation is very significant: here $p < 0,01$ while $p < 0,05$ is enough.

These results signal the importance of semantic approximations and of semantic proximity between verbs in the cognitive organization of verbs lexicon.

In the rest of the paper we present a computational model of semantic proximity and then compare this model with the experimental data obtained from the children.

3 *Proxemy*: a computational approach

A theory of language useful for computational work must account for language statistical regularities. Zipf law (Zipf, 1949) satisfy this observation but provides little insight on lexical structural organization. More recent graph theory studies (Ferrer-i-Cancho and Sole, 2001; Sigman and Cecchi, 2002), capitalizing on results in other scientific domains, provided interesting contributions to the establishment of such a theory of language. All structures discovered in this field research satisfy the “hierarchical small word” (HSW) definition (see section 3.1). Our approach takes place in this general framework. Our specificities are:

- a new linguistic and psycholinguistic insight that guides us and help us on our results validation;
- the kind of objects studied (dictionaries);
- our analysis of graph structure resulting in a computational model of *semantic proximity* among vertices (here vertices are French verbs).

The study by (Resnik and Diab, 2000) signaled that although existing models for verb similarity performed reasonably well against human judgments, none managed to handle certain types of metaphorical pairs such as *to undress / to peel off* that are nonetheless declared to be rather similar by speakers. We aim to develop a model addressing this issue.

3.1 Small World Networks

Networks corresponding to structures found in real world (henceforth *real world networks*) are sparse: in a graph with n nodes, the maximum number of possible edges is $O(n^2)$ while the number of edges in real networks is generally inferior to $O(nlog(n))$. Watts and Strogatz (Watts and Strogatz, 1998) proposed two indicators to characterize a large sparse graph G:

- **L** : *the characteristic path length*, i.e the mean of the shortest path between two nodes of G
- **C** : *the clustering coefficient*, $C \in [0, 1]$, it measures the graph tendency to host zones very dense in edges. (The more clustered the graph is, the more the graph's **C** approaches 1, whereas in random graphs **C** is very close to 0).

In applying these criteria to different types of graphs, Watts and Strogatz found that:

- real world networks have a tendency to have a small **L**: generally there is at least one short path between any two nodes ;
- real world networks have a tendency to have a large **C**: this reflects a relative tendency for two neighbors on the same node to be interconnected;
- random graphs have a small **L**: If someone builds a graph randomly with a density of edges comparable to real world networks, it will obtain graphs with a small *L*;
- random graphs have a small **C**: They are not composed of aggregates. In a random graph

there is no reason why neighbors of a same node are more likely to be connected than any two other nodes, hence their poor tendency to form aggregates.

Watts and Strogatz proposed to call the graphs having these two characteristics (a small **L** and a large **C**) *small worlds* (SW). They recognized these SW in all the real world networks they observed, and therefore postulated for being a SW was an universal property of real world networks. A complete presentation of *Small Words* can be found, for example, in (Newman, 2003).

More recent research has shown that most SW also have a hierarchical structure (*hereafter hierarchical small worlds, HSW*). The distribution of the vertices incidence degrees follows a power law. The probability $P(k)$ that a given node has k neighbors decreases as a power law, $P(k) \approx k^{-\lambda}$, where λ is a constant characteristic of the graph (Barabási and Albert, 1999), while random graphs conforms to a Poisson Law.

In the next section, we present *'proxemy'*, a semantic proximity measure based on a distance we define. A interesting particularity of this distance is to calculate the distance between two vertices on the ground of the complete graph, and not only on their direct neighbors.

3.2 The mathematical model

PROX (PROXemy) is a stochastic method designed for studying "Hierarchical Small Worlds".[1] This method takes graph as input and transform them in a Markov chain whose states are graph vertices. Metaphorically, energy particles wander randomly from vertex to vertex through the edges of the graph. It is their trajectory dynamics that give us the structural properties of the graph.

PROX takes a graph in input and output a similarity measure between the vertices of the graph. Our problem is therefore the opposite than the one of Pathfinder networks (PFNETs see (Schvaneveldt et al., 1988)). PFNETs take a full proximity matrix in input and output a sparse graph. Their goal is to minimize the number of edges required in the sparse graph to be able to approximate the full distance matrix corresponding to the initial full proximity matrix.

[1]In this paper we will use the term *'proxemy'* to refer to the obtained by PROX algorithm. It corresponds to some kind of semantic proximity.

PROX build a similarity measure between the vertices. The hypothesis is that areas having a high density in edges (hereafter, these areas will be called *aggregates*) correspond to closely related verb meanings (in a graph of verbs).

Given a graph with n vertices, $G = (V, E)$, we will note $[G]$ the matrix $n \times n$ such that $\forall r, s \in V$, $[G]_{r,s} = 0$ if $\{r, s\} \notin E$ and 1 otherwise. $[G]$ is called the adjacency matrix of G.

Given $G = (V, E)$ a reflexive graph with n vertices. $[\hat{G}]$ is a $n \times n$ matrix defined by $\forall r, s \in V$, $[\hat{G}]_{r,s} = \frac{[G]_{r,s}}{\sum_{x \in V}\{[G]_{r,x}\}}$. $[\hat{G}]$ is the Markovian matrix of G.

$[\hat{G}]$ is the $n \times n$ matrix is a transition matrix of the homogeneous Markov chain whose states are the vertices of the graph such that the probability of going from one vertex $r \in V$ at an instant t onto another $s \in V$ at the instant $t + 1$ is equal to:

- 0 if $\{r, s\} \notin E$ (s is not neighbor of r)
- $1/D$ if $\{r, s\} \in E$ and r has D neighbors (s is a neighbor of r)[2]

Given $G = (V, E)$ a reflexive graph with n vertices and $[\hat{G}]$ its Markovian matrix, $\forall r, s \in V, \forall t \in \mathbb{N}^*$, $PROX(G, t, r, s) = [\hat{G}^t]_{r,s}$

$PROX(G, t, r, s)$ is therefore the probability for a particle departing from r at the instant zero to be on s at the instant t.

Therefore when, $PROX(G, t, r, s) > PROX(G, t, r, u)$, the particle has more probability to be, at instant t on s than on u and it is graph structure that determine these probabilities.

For the rest of this paper we will set the value of t to 4 since **L** is less than 4 in the kind of graph we are concerned with. Therefore, we take into account the global graph simply by calculating $PROX(G; 4; r; s)$.

Now we have defined our model we will present lexical graphs on which we apply it.

4 Lexical graphs

Several types of lexical graphs can be built according to the type of the semantic relation used for defining the graph's edges. The two principal types of relations used are:

[2]In the context of this presentation of the model we do not consider weighted graphs. However when building the graphs we do consider information, such as the position of the word in the definition, for giving weight to the edges.s

- Syntagmatic relationships, like co-occurrence relationships: they define edges between nodes corresponding to words found near to each other in a corpus.
- Paradigmatic relationships, like synonymy: they define, on the ground of lexical databases such as WordNet (Fellbaum, 1998), edges between nodes of words being in a synonymy relationship in such resource.

Moreover, we are interested into less specific relations, called *semantic proximity relations* or *semantic relatedness*, and which covers both paradigmatic and syntagmatic dimensions.

4.1 Dictionary graphs

Meaning in dictionary definition is at least partially brought by the relations they create between the words constituting the entries. Our approach consists in exploiting the small word properties of the graphs corresponding to dictionaries. More precisely, we are taking advantage of our hypothesis that *aggregates* correspond to areas of closely related senses. We illustrate our approach on two kinds of dictionary, two traditional dictionaries, *Le Grand Robert*[3] and *TLFi*[4], and an synonym dictionary *(Dicosyn)* made of compilation of synonym relations extracted from seven other dictionaries (Bailly, Benac, Du Chazaud, Guizot, Lafaye, Larousse et Robert).[5]

We create a graph from a dictionary in the following way. The entries constituted the vertices. Edges between two vertices A and B were added if and only if B appears in A's lemmatized definition[6] as illustrated in Figure 4.1

We proceed in this way for each entry and obtained a graph of the dictionary. By extracting the subgraph composed only of verbs, the 'neighborhood' we get for the verb 'écorcer' is illustrated by Figure 4.1. Then we render the graph symmetric and reflexive. These modifications on the graphs are allowed thanks to its paradigmatic nature. Graphs created in this way are typical *small world network*. For example, DicoSyn-Verb has 9043 vertices and 50948 edges, its **L** is 4,1694 and its *C* 0,3186.

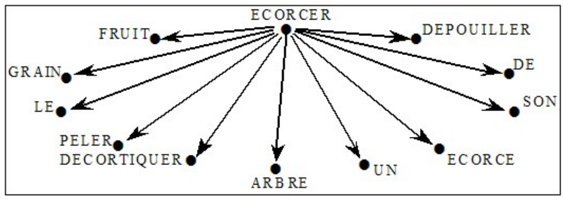

Figure 1: Sub-graph near 'écorcer (to bark – a tree–)' from *Le Grand Robert*

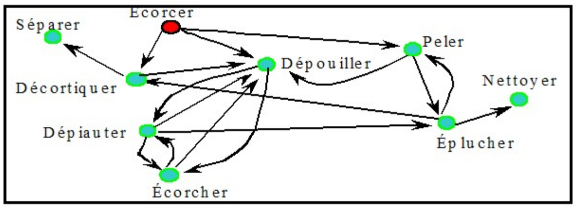

Figure 2: Sub-graph of the verbs near 'écorcer' from Robert

(Duvignau, 2002) has shown that co-hyponymy verb lexical organization according fits with a power law distribution of incidence degrees. In our opinion, (i) the hierarchical organization of dictionaries is a consequence of the special role of the hypernymy relation together with the polysemy of some specific vertices; (ii) the strong **C** reflects the role of interdomain co-hyponyny (Duvignau, 2002; Duvignau and Gaume, 2003). For example, in French language, *'casser (to break)'* appears in many definitions: *'émietter (to crumble)', 'fragmenter (to fragment)', 'détériorer (to damage)', 'révoquer (to dismiss)', 'abroger (to abrogate)'*. This results in a very high incidence for the vertex *'casser (to break)'*. Moreover, many triangles exist *({casser, émietter, fragmenter}, {casser, révoquer, abroger}...,)* and they help to create aggregates. These areas that are bringing co-hyponyms closer in the resulting graph.

4.2 Disambiguization for creation dictionary graphs

Word Sense Disambiguation is a general issue for natural language processing that we need to address when we build our graphs. We need to disambiguate the verbs we found in the definition facing a similar problem as (Harabagiu et

[3]A significant amount of work has been done to encode *'Le Grand Robert* in a graph.

[4]We would like to thank ATILF for making the TLFi resource available to us.

[5]*Dicosyn* has been first realized at ATILF (Analyse et Traitement Informatique de la Langue Française), before being corrected at CRISCO laboratory *(http://elsap1.unicaen.fr/dicosyn.html)*.

[6]Lemmatization has been realized with TreeTagger *(http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/)*.

al., 1999). For example, in French dictionary *Le Grand Robert*, there are two distinct entries for the verb 'causer': *to cause* (3) and *to chat* (4).

(3) CAUSER-1: être la cause de. *(to be the cause of)*

(4) CAUSER-2: S'entretenir familièrement avec qqn. *to chat with*

Of course, the word 'causer' may appear in other definitions like 'bavarder' *(to chat)* . Although a French speaker knows that the 'causer' in (5) refers to the definition (4) our system for building the graph cannot disambiguate. The solution we propose is to (i) first create a fictive vertex which is not a dictionary entry and then (ii) adds two edges {CAUSER, CAUSER-1} and {CAUSER, CAUSER-2 }. When 'causer' is found in another definition like (5), we add the edge { BAVARDER, CAUSER } as illustrated in Figure (5).

(5) BAVARDER "Parler beaucoup, longtemps ou parler ensemble de choses superficielles. - Parler; babiller, bavasser (fam.), cailleter, caqueter, **causer**, discourir, discuter, jaboter, jacasser, jaser, jaspiner (argot), lantiponner (vx), papoter, potiner. Bavarder avec qqn ... "

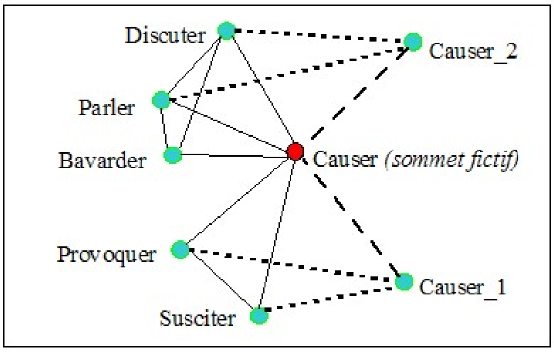

Figure 3: Disambiguation: 'Causer', fictive vertice

In Figure (5), many edges are hidden for clarity reasons. Dashed edges ({Discuter, Causer2}) result from the fact 'Discuter' and 'Parler' are in the definition of 'Causer-2'.

At this stage, we apply PROX to such graph as the one Figure (5) in order to get a matrix $[\hat{G}^4]$ as defined in section 3.2. $[\hat{G}^4]_{bavarder,causer-1} < [\hat{G}^4]_{bavarder,causer-2}$. This comparison allows us to disambiguate.

More generally, let suppose we found a word with k entries in a definition, we will then have $S_1, \ldots, S_k$ vertices corresponding to the entries a fictive vertex S. In case there is an edge $\{A, S\}$ it is replaced by $\{A, S_i\}$ where S_i is such that $[\hat{G}^4]_{A,S_i} = MAX_{0<i\leq k}\{[\hat{G}^4]_{A,S_i}\}$. Then we remove all fictive vertices from the graph to get a disambiguated graph.

We can then apply PROX a last time on the disambiguated graph in order to get the closest word of a word according to our proxemy measure. For example, the PROX-closest words of *écorcer (to bark –a tree–)*, calculated with $t = 6$ are: *1 ECORCER (to bark), 2 DÉPOUILLER (strip), 3 PELER (peel), 4 TONDRE (mow, shear), 5 ÔTER (remove), 6 ÉPLUCHER (peel, pare), 7 RASER (shave), 8 DÉMUNIR (divest), 9 DÉCORTIQUER (decorticate), 10 ÉGORGER (slit the throat of), 11 ÉCORCHER (skin), 12 ÉCALER (husk), 13 VOLER (steal), 14 TAILLER (prune), 15 RÂPER (grate), 16 PLUMER (pluck), 17 GRATTER (scrape), 18 ENLEVER (remove), 19 DÉSOSSER (bone), 20 DÉPOSSÉDER (dispossess), 21 COUPER (cut), 22 BRETAUDER (shear sloppily), 23 INCISER (incise), 24 GEMMER (tap), 25 DÉMASCLER (remove first layer of cork)*[7]

5 Proxemy and Experimental studies

Prox is a robust method: changing randomly a few edges does not change significantly the results. The repartition of aggregates is not strongly affected by a random redistribution of some edges. However the relevance of our proxemy approach of lexical networks is tied to the linguistic representativity of the networks we use. Therefore, we tested the PROX model of four different dictionary graphs and we compared them to the psycholinguistic experimental results presented in section 2. The graph we compared were:

1. Graph.TLFI.Verb, a graph built as explained in 4.1 from TLFi[8] dictionary,

2. Graph.Robert.Verb, a graph built as explained in 4.1 from *Le Grand Robert* dictionary,

3. Graph.DicoSyn.Verb, in which there is a edge between two verbs if there are given as syn-

[7] Proposing a translation for such fine grained and sometimes polysemous words is impossible since proposing the translation include a certain form of disambiguisation as it is suggested by the work of (Gale et al., 1992).

[8] http://atilf.atilf.fr/tlf.htm

onyms by one of the synonym dictionary composing DicoSyn

4. Graph.DicoSyn_20 built from Graph.DicoSyn but in which 20% of the edges are randomly removed and re-added.

For each of these graphs we looked at two variables to be related with the psycho-linguistics experiments: the answers incidence and the proximity of answers to a *'reference verb'*

Answers incidence We compare in the graph the average incidence degree between adult (ID_{adult}) and children answers ($ID_{children}$).

Average incidence	Children	Adults
Graph.TLFi.Verb	61	29
Graph.Robert.Verb	236	126
Graph.DicoSyn.Verb	102	58
Graph.DicoSyn.Verb _20	66	40

Table 2: Results for 'Answers incidence'

The proximity of answers to a *'reference verb'* Three linguist judges determined together for each movie which was the most appropriate verb to describe the action performed in the movie (hereafter R_i is the reference verb for the movie M_i). For a given movie M_i, an answer may therefore be ranked according to its proxemy according to R_i.

For a lexical graph $G = (V, E)$ composed of n words, and for a reference verb $R_i \in V$, one can define $rank_{Ri}$ for ranking all the vertices of V in decreasing order resulting from a PROX iteration $PROX(G, t, Ri, \bullet)$ on V (see section 3.2).

Average rank relatively to reference verbs	Children	Adults
Graph.TLFi.Verb	270	173
Graph.Robert.Verb	121	76
Graph.DicoSyn.Verb	105	44
Graph.DicoSyn.Verb _20	185	94

Table 3: Proximity between answers and reference verb

Our first hypothesis was that $ID_{adult} < ID_{children}$. According to the hypothesis children would learn first words corresponding to high incidence vertices. Then they would use them for talking about an large lexical area (e.g 'casser' (to break) is used by children while adults use a more precise verb like 'déchirer' (to tear) which has a lower incidence in dictionary graphs).

Our second hypothesis was that the mean of the rank of the children answers according to the reference verb is higher that the adult ones. When a child is attempting to communicate an event *(e.g déchirer un livre, to tear a book)* for which he does not have an already constituted verbal category, he would do an analogy with a past event *(e.g to break a glass)* and use this verb for describing the current event *(e.g casser un livre, to break a book).* The adult could use a number of more accurate verbs but their proxemic rank, with regard to the reference verb, is generally lower than the children ones.

The table 2 shows the results concerning answers incidence. Although some variability is observed across the graphs, our first hypothesis is validated for the 4 graphs. On the three first graphs the average incidence of answers is roughly twice as the adults one.

The table 3 illustrates the results concerning proxemic rank of answers according to the reference verb. Again, in spite of some variability across the graphs our second hypothesis is validated as well. Moreover, having in mind that the graph has about 10 000 vertices, we observe that although less close that adults answers, the children answers remain relatively close to the reference verb according to our proxemic measure.

6 Conclusion

Our psycholinguistic approach allows us to establish that semantic proximity between verbs play a fundamental role during the period of early lexical acquisition. We signaled the existence in the organization of the lexicon of a relation of co-hyponymy between verbs. Based on these first observations we consider that productions based on semantic proximity are particularly interesting: they manifest the existence, at the surface level of discourse, of a lexical relation of inter-domain 'semantic proximity' between verbs not yet considered in linguistics.

Moreover we have seen that semantic approximations for verbs appear to fit the proximity values calculated by PROX. On the ground of these first results, we postulate that constructing electronic dictionaries on the ground of linguistic theory of lexical semantic organization that fits with early lexicon acquisition as well with adult lexical organization will provide them interesting ergonomics properties. This should increase their usability and might be taken into account for normalizing electronic dictionaries.

For example, we are developing a 'proxemic electronic dictionary' from TLFi. Such dictionaries enable to find an uncommon but precise verb like *'to bark'* by using (i) a common verb like *'to undress'* which is related to *'to bark'* by semantic proximity and (ii) a word *(e.g 'tree')* bringing a relevant semantic domain. Moreover, in the definition of *'to bark'* one can find: *'tree', 'grain' 'fruit'* which are close from each other according to PROX ran on nouns. Finally, when we look for verbs that are close from both *'to undress'* and *'tree'*, PROX provides the verbs: *'to cut, to ring, to peel, to notch, to bark, to incise,...'* which constitute relevant verbs. Such a dictionary can be useful for didactic studies where it can complements approaches like and NLP for word sense desambiguization (Gaume et al., 2004) or de-metaphorization.

References

Barabási, Albert-László and Réka Albert. 1999. Emergence of scaling in random networks. *Science*, 286:509–512, October.

Dik, S. 1991. Functional grammar. In Droste, F. and J. Joseph., editors, *Linguistic theory and grammatical description*. Amsterdam : Benjamins.

Duvignau, K. and B. Gaume. 2003. Linguistic, psycholinguistic and computational approaches to the lexicon: Contributions to early verb-learning. *Journal of the European Society for the Study of Cognitive Systems*, 6(1).

Duvignau, Karine and Bruno Gaume. 2008. Between words and world: Verbal 'metaphor' as semantic or pragmatic approximation? In *Proceedings of International Conference 'Language, Communication and Cognition'*.

Duvignau, K., B. Gaume, and S. Kern. 2005. Semantic approximations intraconceptăvs. interconcepts in early verbal lexicon: flexibility against error. In *Proceedings of ELA 2005, Emergence of language abilities: ontogeny and phylogeny*.

Duvignau, K. 2002. *La métaphore berceau et enfant de la langue*. Ph.D. thesis, Université Toulouse Ű Le mirail.

Fellbaum, C., editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.

Ferrer-i-Cancho, Ramon and Ricard V. Sole. 2001. The small world of human language. *Proceedings of The Royal Society of London. Series B, Biological Sciences*, 268(1482):2261–2265, November.

Gale, W., K. Church, and D. Yarowsky. 1992. A method for disambiguating word senses in a large corpus. *Computers and the humanities*, 26(2):415–439.

Gaume, B., K. Duvignau K., O. Gasquet O., and M-D. Gineste. 2002. Forms of meaning, meaning of forms. *Journal of Experimental and Theoretical Artificial Intelligence*, 14:61–74.

Gaume, B., N. Hathout, and P. Muller. 2004. Désambiguisation par proximité structurelle. In *Proceedings of TALN 2004*.

Harabagiu, Sanda M., George A. Miller, and Dan I. Moldovan. 1999. Wordnet 2 - a morphologically and semantically enhanced resource. In *SIGLEX 1999*.

Hofstadter, D. 1995. *Fluid concepts and creative analogies*. New York : Basic Books.

Newman, M. 2003. The structure and function of complex networks.

Ny, J-F. Le. 1979. *La sémantique psychologique*. PUF.

Piaget, J. 1945. *La formation du symbole chez lŠenfant,*. Delachaux et Niestlé.

Resnik, P. and M. Diab. 2000. Measuring verb similarity. In *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society*.

Schvaneveldt, R. W., D. W D.W Dearholt, and F.T Durso. 1988. Graph theoric foundations of pathfinder networks. *Computers and Mathematics with Applications*, 15:337–445.

Sigman, M. and G.A. Cecchi. 2002. Global organization of the wordnet lexicon. *Proc. Natl. Acad. Sci.*, 99(3):1741–1747.

Watts, D.J. and S.H. Strogatz. 1998. Collective dynamics of small-world networks. *Nature*, 393:440–442.

Zipf, G. K. 1949. *Human behavior and the principle of least effort*. Addison-Wesley.