Using Extra-Linguistic Material for Mandarin-French Verbal Constructions Comparison *

Pierre Magistry[a,b], Laurent Prévot[c], Hintat Cheung[a], Chien-yun Shiao[a], Yann Desalle[d], and Bruno Gaume[d]

[a] Graduate Institute of Linguistics, National Taiwan University, Taiwan
{d97142005, hintat, r96142005}@ntu.edu.tw
[b] CLCLP - Institute of Linguistics, Academia Sinica, Taiwan
[c] Laboratoire Parole et Langage, CNRS — Université de Provence, France
laurent.prevot@lpl-aix.fr
[d] CLLE-ERSS, CNRS — Université de Toulouse, France
{desalle, gaume}@univ-tlse2.fr

Abstract. Systematic cross-linguistic studies of verbs syntactic-semantic behaviors for typologically distant languages such as Mandarin Chinese and French are difficult to conduct. Such studies are nevertheless necessary due to the crucial role that verbal constructions play in the mental lexicon. This paper addresses the problem by combining psycho-linguistics and computational methods. Psycho-linguistics provides us with a bilingual corpus that features verbal construction associated with carefully built extra-linguistic material (short video clips). Computational approaches bring us distributional semantic models (DSM) to measure the distance between linguistic elements in the extra-linguistic space. These models allows for cross-linguistic measures that we evaluate against manually annotated data. In this paper, we discuss the results, potential shortcomings involving cultural variability and how to measure such bias.

Keywords: Cross-linguistic study, Distributional Semantic Models, Psycholinguistics, Extralinguistic context

1 Introduction

The study reported in this paper is part of a broader project (M3: Model and Measurement of Meaning) which investigates lexical organization of French and Mandarin. The project combine psycho-linguistic and computational approaches for investigating the verb lexical organization for both languages. We therefore deal with two kinds of data: On the one hand the productions obtained through psycho-linguistic experiments realized both with French and Mandarin Chinese speakers; On the other hand we systematically exploit existing electronic resources. However, this paper does not address on the later aspect of this work. We focus here on crossing the results of both methodologies: psycho-linguistics experiments and analyses with computational and statistical methods. As suggested by (Biber, 1988), particular attention is given to extra-linguistic material that is used here as a context for applying a distributional approach. Distributional approaches are usually applied to linguistic material. However, in this study we attempt to use the

* We want to thank Karine Duvignau and Marion Fossard for collecting the French data as well as Jialin Wu, Yufeng Yao, Miaoru Huang, Tehsin Wu and Sumen Chen for collecting the Mandarin data. The research presented in this paper is supported by the ANR-NSC (France-Taiwan) bilateral project M3 (Modeling and Measurement of Meaning). More information about M3 project on these websites: http://erss.irit.fr/flexsem (France) http://140.112.147.149:81/m3/ (Taiwan). We are also grateful to the anonymous reviewers.

videos themselves for unraveling some aspects of the mental lexicon. The experiments proposed in the following sections demonstrate the validity of such an approach. Our work can be situated in a growing body of semantic studies that approximate semantic objects with distributional information. The core idea is that similar semantic objects have similar distributional properties. The meaning of a word is modelled as a vector in a multidimensional space. Although our approach is rooted into empirical evidence, we do not discard more theoretical insights of lexical organisation. We rather argue for a combination of distributional semantics for content units such as nouns or verbs and of more traditional formal semantics for more functional units like closed-class words. Distributional semantics are well equipped for studying the paradigmatic organisation of open-class words. Once our empirical investigation achieved, we will relate it to existing theoretical frameworks dealing with verb semantics such as (Ahrens *et al.*, 2003).

The paper is structured as follows. Section 2 introduces the protocol used for eliciting the linguistic material of this study. Section 3 details how extra-linguistic material can be used as a space for projecting linguistic data from both languages. Then we introduce our bilingual alignment in section 4 and we use it for evaluating and analyzing our statistical results (in section 5). The approach raises some issues about the bias that the extra-linguistic material may introduce. We will try to address this concern in section 6 before concluding and discussing the impact of this study on our longer term objectives.

2 Collecting data

In order to compare the lexical structure of verbs in French and Mandarin, we used the Approx protocol initially built to study the production of verbal approximates and the hierarchical aspects of the lexicon in normal and pathological development (Duvignau *et al.*, 2005; Gaume *et al.*, 2008).

2.1 Material

The material of this protocol consists of 17 speechless action-video clips divided as in Figure 1 into 3 action concepts: TO DETERIORATE, TO REMOVE, TO SEPARATE.

In each video clip, a woman alters an object from an initial state to a final state with her hands or with an instrument.

TO DETERIORATE	TO REMOVE	TO SEPARATE
Burst a balloon	Peel a carrot	Make bread-crumbs
Crumple a piece of paper	Peel an orange	Cut a bread
Break a glass	Pull the bark of a log off	Break a bread off
Crush a tomato	Undress a doll	Chop parsley
Tear off a newspaper	Take down legos	Saw a plank of wood
	Peel a banana	Unsew a shirt

Figure 1: The 17 action video clips of the Approx protocol

These video clips were shown to several participants. For the M3 project, we focused on two populations in France and in Taiwan : healthy children and adults.

2.2 Participants

The French participants were 170 children between the ages of 2 and 11 years old. The French baseline population was constituted by 75 young adults between 18 and 38 years old. Children were recruited from kindergartens and elementary schools in France.

The Taiwanese participants were 80 children between the ages of 3 and 9 years old. Children were divided into in four different age groups: 3-year-olds (n=20; mean age= 3.05 years;

SD=1.82), 5-year-olds (n=20, mean age=5.02 years; SD=2.27), 7-year-olds (n=20, mean age=7 years; SD=2.68) and 9-year-olds (n=19, mean age=9.01 years; SD=2.06). Twenty children were recruited from kindergartens and elementary schools located Taiwan. Sixty university students were also recruited so that their responses could serve as a baseline for the children's response.

2.3 Procedure

Before showing the video, participant were told that "a lady" (or a "a big sister" in Mandarin with children participants) was going to perform. After showing the video, the experimenter asks the children what the woman has done. After the first answer, a rewording of the action was asked to the participant. In the French study, the video clips were shown in random order. In the Mandarin study they were beforehand counterbalanced, thus creating 6 different sets of video clips. Between each video clip event, a distractor (i.e. popular cartoon characters in Mandarin Team) was shown to the children to avoid perseveration effects.

For each answer, verbal forms were extracted from both answers and then lemmatized. In Mandarin, compound verbs V1V2 were recorded and analyzed as V1V2, V1 alone and V2 alone (see section 4.1). In French, only the more meaningful verbal form was recorded. In this way, we built two databases (one for each language), listing all the verbs produced by the participants for each video clip. Our quantitative study will focus on young adult productions while a manual alignment will be done for the children's productions.

3 Movies as dimension spaces for linguistic comparison

For each verb produced, we count how many times it has been used for describing each movie. We therefore study the distribution of the verbs in the 17-dimension space generated by the video clips. The coordinates of a verb in this space corresponds to the number of times a verb has been used for the corresponding video clip, as illustrated in Figure 2. From this distributional data, we will analyse the semantic structures underlying both languages and their differences. Distributional similarity as an approximation for semantic similarity is discussed in details by (Curran, 2003), see also (Evert and Lenci, 2009) for a complete overview and more recent state of the art.

Verb	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
briser	24	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
couper	1	52	1	17	9	0	0	75	0	0	2	0	5	0	5	0	0
...	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
v-切	0	0	0	43	0	0	0	24	0	0	0	0	2	0	0	0	0
v-剝	0	4	0	0	19	0	0	0	47	0	46	26	0	0	0	0	0
...	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
v1-弄	0	0	14	0	4	1	0	0	0	0	0	1	0	0	0	0	2
v2-成兩半	0	13	0	0	0	0	0	11	0	0	0	0	6	0	21	0	0
...	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Figure 2: Examples of verb coordinates in the video clips space

The projection can be done for both French and Mandarin verbs. The resulting coordinates are in the same semantic space and therefore comparable. Moreover, in the case of Mandarin, we can go further and measure components V1 and V2. Various measures have been proposed for comparing vector in multidimensional spaces. See (Evert and Lenci, 2009) for a introductory overview and (Bullinaria and Levy, 2007) for a deeper systematic investigation of the effect of mathematical parameters in DSM. However, in our case we simply measure distance in the 17-dimension space with Euclidean distance applied onto normalized vectors.

$$dist(va, vb) = \sqrt{\sum_{i=1}^{n}(va_i - vb_i)^2} \quad (1)$$

Although dependent on protocol set-up for eliciting the data (in particular the video clips) this data constitute a crucial starting point for our study. Sections 6 and 7 investigate the shortcomings and solutions of such an approach.

4 Aligning Chinese and French verb data

For each video clip we gathered the Taiwanese and French answers and we generated the full list of potential translation pairs. We are currently performing an evaluation of the pairs. The evaluation is carried on by two persons: a Chinese native speaker and a French native speaker. They worked together for assessing the validity of each translation pair. At the time of writing, about 10% of the 11000 pairs have been evaluated. The protocol is straightforward. The annotators are presented with a table listing the pairs and they have to decide whether or not (binary choice) the two verbs constitute an acceptable translation of each other. The corresponding video clip to the verbs was shown beforehand in order to reduce polysemy issues for the translation.

To understand the main translation issues, it is important to detail the kind of answer obtained through the protocol.

4.1 Mandarin Data Classification

The role of resultative construction and related issues has been heavily studied in Mandarin (Thompson, 1973; Lu, 1977; Gu, 1992; Cheng and Huang, 1994; Gao, 1997). Because of the compositional nature of verbal forms, most verbs can be viewed as made of two sub-events. In the case of resultative construction, the first verbal component represents the action part while the second indicates the result. Various proposals have been presented in accounting for the thematic relations between the two verbal components (Li, 1990; Chang, 1998). Accordingly, the responses from children and adults can be classified into the following categories:

- Action verb: responses involving only an action verb. For example, for the event "cut parsley", children's answer might be something like "切 香菜" *(cut parsley)*
- Resultative verb: responses involving only an resultative verb. For the same event, children's answer might be something like "香菜 斷-了" *(parsley snapped-PFV)*
- Resultative verb compound: responses consisted of an action verb plus a resultative verb, forming a resultative verb compound. Responses from children could be something like "切-開 香菜" *(cut-open parsley)*.
- resultative verb construction: responses like action verb-INTO-resultative verb (切-成-碎碎的) or action verb-DE-resultative verb (切-的-碎碎的 *cut-INTO/DE-pieces*) are considered to be resultative verb constructions.

Some answers, although easily interpretable, clearly belong to childish language, e.g 打打打 *(hit-hit-hit)*. Their alignment is somewhat difficult. Some childish language production might be available in the other language but it is rather tricky to decide about translation relations between these forms.

4.2 Diathesis alternation

In French data, some other phenomena are raising relating issues. Some verbs may be used in different ways for describing our video clips as illustrated in example (1).

(1) a. La verre a explosé.
The glass exploded.
b. La dame a explosé le verre.
The woman exploded the glass.
c. La dame a fait exploser le verre.
The woman caused the glass to explode.

While in standard French verbs like *éclater, exploser, craquer,...* are intransitives (example (1-a)) and have only an unaccusative subject as argument, colloquial French allows sometimes for transitive uses like example (1-b) which is derived from the causative construction FAIRE + V (TO MAKE + V) that can be translated into *"Agent make Patient V"*. This ambiguity could be an issue when evaluating Mandarin-French. Mandarin productions include clearly full V1V2 construction but also stative verbs alone (like 破掉, *broken-away*) or more complicated stative construction like 碎的很可怕 *(frighteningly into pieces)* (as stated above). However, since the eliciting question of the protocol was *"What did the woman do?"*, the French answers mostly correspond to the colloquial transitive use of these verbs. A more detailed analysis of French answers should nevertheless be carried on in the future for distinguishing the different alternations.

4.3 Productivity of Mandarin V1V2

Mandarin is well known for its verbal compound productivity (Li, 1990; Chang, 1998; Chen, 2008). One of the widely examined verb compounds is resultative verb compound. As mentioned above, V1 of RVC encodes the action and V2 the result. The limitations on the combination between V1 and V2 are largely semantic. Such productivity can be observed in young children. (Erbaugh, 1992) reported that a two-year-old boy produced a novel resultative construction 壞人打死掉 *(the bad guy hit-to death-away)*. Also, the degree of precision of the construction proposed by some adults has no equivalent in French language (without entering into lengthy paraphrases). It is specially the case for some video clips, like the "newspaper" one: 撕成對半, 撕成兩半 *(torn to form two halves)*. Such remarks emphasize the need for including V1 and V2 alone into our study.

5 Results

After removing single occurrence words, we are left with 191 verbs (43 French verb, 85 Mandarin verb, 63 V1 or V2). A first comparison has been realized by measuring the average distance in the video clips space between translation-pairs and the other pairs. Notice that all the pairs are real translation candidates since both verbs have been used for the same video clip.

The average distance, in the video clip space, for translation pairs is about 0.8 and about 1.25 (on a $[0 - \sqrt{2}]$ scale) for other pairs. With a threshold of 0.95 the agreement between the manual alignment and our distance is about 87%. A more precise evaluation of our ability to detect translation pairs presents however more a complex picture (Precision: 0.55, Recall: 0.46, F-score 0.5). This figure was confirmed by a manual annotation of the positive pairs which provided similar figure for recall, precision and f-score.

Another measure consist in evaluating the ranks of translation equivalents according to the distance in our video clip space. We tried to make the decision on neighborhood ranking rather than distance threshold but it yielded roughly the same results (with an earlier but slower drop down). To improve the evaluation process given the lack of gold standard, we manually evaluated the output of a system that would consider every pairs with a distance less than 0.95 to be a proper translation. Again we obtain similar results. However, we also noticed that often (around 25% of false-positive) the French verb would still be an acceptable translation of either the V1 or V2 part (but lacks some precision). V1V2 appear as a crucial element of the analysis, section 5.1 will provide some details about it.

An intrinsic shortcoming of our approach has to do with "dimension sparseness". DSM are supposed to be applied in high-dimensional spaces so they can take advantage of PCA (Principal Component Analysis) and other dimension reduction techniques. Some verbs intuitively very dissimilar are nevertheless "synonyms" in our video clip space because the space carved by the video clips is not discriminating enough for sorting out these verbs. For example *'crever (to burst)'* and *'taper (to hit)'* in French appear only into the *balloon* video clip. They are therefore "near-synonyms" in our 17-video clips space. There is no easy way to work around this issue. To have

a richer extra-linguistic context, we must have more video clips and those are costly to shoot. On a more positive note, it is worthy to underline that similar materials have been developed and continue to be used in other teams for studying other domains of the lexical organisation. See for example video clips for motion, put/take, cut/break and a few other kind of verbs developed at the Max-Planck Institute (Majid *et al.*, 2007).

5.1 Mandarin verb compounds

As we mention in section 4.1 Mandarin Chinese has Verb-Resultative constructions and compounding that are highly productive and compositional. Many difficulties encountered for alignment of French and Chinese data are related to them. For data exploration purpose we decided to project in the same semantic space all the levels of the Chinese data: verbs (compound or construction V1+V2), activity part of the verb (V1) and resultative or directional part of each verb (V2).

In the cases in which data sparseness is not involved we find four different types of relevant objects in the close neighborhood of a V2 :

- other V2 (near-synonyms)
- V1 that are often used with the V2 (collocation)
- V1V2 compounds (including the same V2 or near-synonyms)
- French verb (translation)

For example, the nearest neighbors of the V2 爆 *(explode)* are the verbs 打爆 *(to hit-explode)*, the V1 拍 *(to clap)* and the French "éclater" *(to explode/burst)*, those of V2 碎 *(into pieces)* are French "briser" *(to smash)*, V2 的粉碎 *(smashed to pieces)*, French "casser" *(to break)*, V1 敲 *(to knock)* and verb 打碎 *(to hit-into pieces)*.

It is not always easy to decide whether a verb in French focus more on the result or on the activity without looking at tense or other linguistic context, but French verb like "éclater" *(to burst)*, "briser" *(to smash)*, "casser" *(to break)* typically denote a result (even though as mentioned before, unlike Chinese V2, they can have a transitive use involving an agent).

Unfortunately, our space does not account for all the V2, as well as it does for 爆 *(explode)* and 碎 *(into pieces)*. The main factor here is data sparseness, both in term of linguistic data and extra linguistic data. Some lexical items have been uttered for only one video and this dimension sparseness leads to false translation relation such as V2 破 *(smashed)* and French "crever" *(to burst)* (which is more specific than 破) for the *burst_balloon* video clip.

Another issue is that the semantic similarities we obtain from this method concern only the aspects of the meaning related to our video clips. The closest neighbors of the directional V2 下來 *(to remove/come-down)* are V1 扒 *(to strip off)*, V2 下 *(down)*, V2 掉 *(fall/away)*, French "retirer" *(to pull out)* and "enlever" *(to remove)*. This seems correct but we have to keep in mind that it is only one possible meaning for 下來 *(to remove)* and the other meaning *(downward path)* is not pointed out by the French data.

6 Measuring cultural variation

A first shortcoming of the study presented above concerns the potential cultural dependence of the video clips. Although the material has been carefully controlled for not being cultural dependent some artifacts might still be at work. From a more general viewpoint, any attempt for finding lexical equivalence has to address cultural variations between the languages involved. The work in this section is a proposal for handling this general issue.

One way to detect cultural discrepancy related to the video clips or at least to spot less relevant video clips for our purpose of alignment is to perform a classification of the video clips based on linguist data. The most straightforward way is to project the video clips in the space defined by

verb use (with each verb as one dimension). However this method would have to face the issue that data points would be in two different spaces: Chinese verbs on one hand and French verbs on the other. To enable us to compare video clips we used a random walk in a weighted graph approach (See below). Using this method, we are able to project video clips in the same space (defined by video clips as dimensions), while we still take in account the influence of one language or another.

Our projected data points are video clips typed according to a language (fr or ch). We therefore have 34 data points into the 17-dimensions video clips space. Then we compare the distance between French and Chinese interpretations of each video clip and perform a cluster analysis.

6.1 Formal definition and computation

To conduct the computation of the coordinates, we followed the method stated in (Gaume *et al.*, 2004) with a number of steps $i = 2$. If $G = (V, E)$ is our graph with $|V| = n$ vertices, we consider the adjacency matrix of G that we note $[G]_{i,j}$. $[G]_{i,j} = 1$ iff there is an edge between the vertices i and j, 0 otherwise.

The next step is to create the Markovian matrix of G, such that

$$[\hat{G}]_{r,s} = \frac{[G]_{r,s}}{\sum_{x \in V}([G]_{r,x})}$$

Following this definition, $[\hat{G}]^2$ will be a matrix that contains all the coordinates we need.See (Gaume *et al.*, 2004) for more details about this method.

6.2 Random walks in experimental data

To obtain the coordinates of the video clips in the video clips space, we first built two weighted bigraphs (one for each language). The nodes are either video clips or verbs. There is a vertex between a given verb and a given video clip if and only if a verb has been uttered to describe a video clip. The vertices are weighted as estimation of conditional probabilities using occurrence counts. The result is a markovized bipartite and symmetric graph that can be used to compute the coordinates. This is done by looking at how probabilities are spreading into the graph from each video clip as a starting node.

Our graphs being bigraphs, 2-step random walks starting from a video clip will yield non-null probability only for video clips. Those probabilities on 17 video clips are then used as dimensions in a space. (See Figure 3)

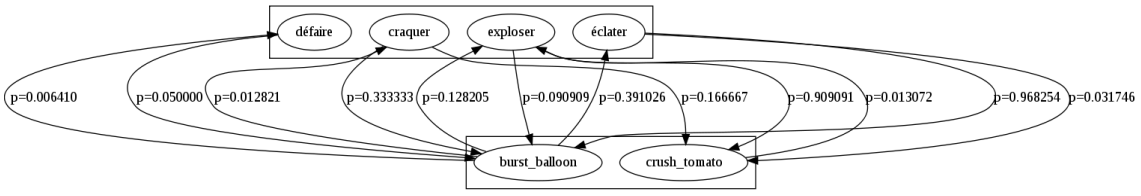

Figure 3: Movie-Verb bigraphs and 2-step random walks

Figure 3 shows a small part of the French graph (two video clips and five verbs). On the complete graph, the sum of the probabilities of all outgoing vertices from a node is equal to 1. To compute the coordinates of a given video clip, simply sum up the probabilities of all the paths going from this video clip to each video clip, in two steps.

This method projects extra-linguistic situations in the same space. Their position still depend on the language used to describe each situation. Distances between video clips is computed in a similarly than in classic DSM approaches.

We first applied PCA and hierarchical clustering to observe whether some video clips are likely to be culturally (or linguistically) dependent. This is interesting *per se* for our longer term objectives[1] but also here for evaluating the validity of the verbal DSM dimensions. If in this space a video clip described in one language is far from the same video clip described in the other language, it is questionable to use it as an axis for our bilingual lexical space.

The PCA has been mainly used for data exploration (49% of cumulative proportion of variance remains with only 3 dimensions, and 37% in a 2-dimensional plane). The hierarchical clustering however appear more informative. We also use computed distance between video clips (value or rank). As we can see on the dendrogram (Figure 4), video clips such as *break_up_bread, crumb_bread* and *peel_carrot* are likely to be problematic (distances with their cross-linguistic counterpart of 0.94,0.79 and 0.74 respectively while "good couples" range between 0.2 and 0.5). Interestingly those video clips use either a typical French bread or a French carrot whose size is significantly different from a Taiwanese carrot.

Figure 4: Movies clustering results obtained by random walk in the bigraph

7 Conclusion and Future Work

This paper has explored the possibilities opened by distributional and stochastic studies of lexical semantics grounded on extra-linguistic material. We have presented a full experimental set up. The coherence of our interdisciplinary approach is supported by compatible results with manual annotations. We then have proposed a way to measure linguistic dependence of the extra-linguistics material. These steps are important when developing a fully interdisciplinary approach as the one advocated here. The work presented is however only a preliminary step for the more specific studies mentioned below.

From a cognitive semantic view, the video clips actually represent the physical world. The elicited answers are conceptual encoding processes from two groups of speakers. Each participant has its own perspective in viewing these events (for example, *peeling a carrot* or *sharpening something*). We are currently working on an evaluation using video clip ratings for trying to

[1] The evolution of this classification during first language acquisition can be studied but we keep this for future work.

capture this level of conceptualisation and constitute a better representation of the extra linguistic context.

Moreover, we are having a closer look at the syntactic dependencies of the verbs involved in the answers, in particular in terms of subcategorisation frames. Later on this will enable us to use more linguistic contexts.

The correlation between our distance and the semantic contribution of the verbal elements demonstrate the overall validity of distributional approaches of semantics. An important aspect of our future work will be to determine how such models can be mapped to more theoretical approaches. For example, investigating the different possibilities vectors combination of various elements of the verbal construction such as V1V2 will be one of our concern for future study.

The 17 video clips are all about state-changing events, including an agent, an action, and an object. So they are within a well-defined semantic space. It will be however interesting to scale up the study on verbs with different event structures. Motion events, such as those in used in Slobin's Frog Stories (Slobin, 2004), have very different semantic characteristics. We will be extending our study to different verb types to examine the cross-lingual dependency of semantic distance and event type.

References

Ahrens, K., C.-R. Huang, and Y.-H. Chuang. 2003. Sense and meaning facets in verbal semantics: A MARVS perspective. *Language and Linguistics*, 4(3):469–484.

Biber, D. 1988. *Variation across speech and writing*. CUP.

Bullinaria, J. A. and J. P. Levy. 2007. Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, 39:510–526.

Chang, C. H.-H. 1998. V-V compounds in Mandarin Chinese: Argument structure and semantics. In J. L. Packard, editor, *New approaches to Chinese word formation: Morphology, phonology and the lexicon in modern and ancient Chinese*, pages 77–101. New York: Mouton de Gruyter.

Chen, J. 2008. *The acquisition of verb compounding in Mandarin Chinese*. PhD thesis, de Vrije Universiteit Amsterdam. http://hdl.handle.net/1871/12679.

Cheng, L. L.-S. and C.-T. J. Huang. 1994. On the argument structure of resultative compounds. In M. Y. Chen and O. J. Tzeng. , editors, *In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*, pages 187–221. Pyramid Press.

Curran, J. R. 2003. *From Distributional to Semantic Similarity*. PhD thesis, University of Edinburgh.

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

Erbaugh, M. 1992. The acquisition of Mandarin. In D. I. Slobin, editor, *The crosslinguistic study of language acquisition*, pages 373–455. Mahwah, NJ: Lawrence Erlbaum.

Evert, S. and A. Lenci. 2009. Distributional semantic models. Advanced course at ESSLLI 2009, Bordeaux, July 27-31, 2009. http://wordspace.collocations.de/doku.php/course:start.

Gao, Q. 1997. Resultative verb compounds and ba-construction in Chinese. *Journal of Chinese Linguistics*, 25:84–130.

Gaume, B., K. Duvignau, L. Prévot, and Y. Desalle. 2008. Toward a cognitive organization for electronic dictionaries, the case for semantic proxemy. In *Coling 2008: Proceedings of the Workshop on Cognitive Aspects of the Lexicon (COGALEX 2008)*, pages 86–93, Manchester.

Gaume, B., P. Muller, and N. Hathout. 2004. Word sense disambiguation using a dictionary for sense similarity measure. In *Proceedings of Coling*.

Gu, Y. 1992. *The syntax of resultative and causative compounds in Chinese*. PhD thesis, Cornell University, Ithaca, NY. unpublished.

Li, Y. 1990. On V-V compounds in Chinese. *Natural Language and Linguistic Theory*, 8:177–207.

Lu, J. H.-T. 1977. Resultative verb compounds vs. directional verb compounds in Mandarin. *Journal of Chinese Linguistics*, 5:276–313.

Majid, A., M. Bowerman, M. Staden, and J. Boster. 2007. The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics*, 18(2):133–152.

Slobin, D. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Stromqvist and L. Verhoeven. , editors, *Relating Events in Narratives: Typological and Conceptual Perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.

Thompson, S. A. 1973. Resultative verb compounds in Mandarin Chinese: A case for lexical rules. *Language*, 49:361–379.