Distributed Representations of Lexical Sets and Prototypes in Causal Alternation Verbs

Edoardo Maria Ponti*
University of Cambridge

Elisabetta Jezek**
Università degli Studi di Pavia

Bernardo Magnini†
Fondazione Bruno Kessler

Lexical sets are sets of words filling a specific argument position of a verb in one of its senses. They can be extracted from corpora automatically. The purpose of this paper is demonstrating that the properties of lexical sets are mirrored by their distributed vector representations. This provides insights onto many linguistic phenomena, such as the causative-inchoative alternation in the verbal domain. A first experiment aims at investigating the internal structure of the sets, whose meanings are known to be radial and continuous categories cognitively. In particular, we establish an equivalence between the prototype of a set of words and the centroid of a set of vectors. A second experiment shows that the distance between the intransitive subject set and transitive object set is correlated with the spontaneity of the event expressed by the verb, defined in terms of morphological coding and frequency.

1. Introduction

Lexicographic attempts to cope with verb sense disambiguation often rely on "lexical sets" (Hanks 1996), which represent the lists of corpus-derived words that appear as arguments for each distinct verb sense. The arguments are the "slots" that have to be filled to satisfy the valency of a verb (subject, object, etc.). In previous works (Montemagni, Ruimy, and Pirrelli 1995, *inter alia*), lexical sets were collected manually and were compared through set analysis. The measure of similarity between two sets was proportional to the cardinality of their intersection. We believe that possible improvements may stem from deriving the lexical sets automatically and from exploiting the semantic information of the fillers fully.

In this work, we devise a method to extract lexical sets from a huge corpus and use a distributional semantics approach to perform our analyses. More specifically, we represent fillers as word vectors and compare them through spatial distance measures. In order to test the relevance for linguistic theory of this approach, we focus on a case study, namely verbs undergoing the causative-inchoative alternation. Based on two experiments, we show how our methodology clarifies the some controversies: the relationship among the members of a same set of words, and the position of each of these verbs along the so-called spontaneity scale (see § 1.2) (Haspelmath 1993; Samardzic and Merlo 2012; Haspelmath et al. 2014).

* ep490@cam.ac.uk
** jezek@unipv.it
† magnini@fbk.eu

The structure of the paper is as follows. In § 1, we define the core notions of this study, including lexical sets, causative-inchoative verbs, and word embedding. § 2 presents the method and the data, whereas § 3 reports the results of the experiments. Finally, § 4 draws some conclusions from the them and proposes possible future lines of research.

1.1 Lexical Sets

A lexical set can be defined as the set of words that occupy a specific argument position within a single verb sense, such as {*gun*, *bullet*, *shot*, *projectile*, *rifle*...} for the sense 'to shoot' of *to fire*, or {*employer*, *teacher*, *attorney*, *manager*...} for its sense 'to dismiss'. The notion of lexical set was firstly introduced by Hanks (1996). Its purpose is explaining how the semantics of verbs is affected by the patterns of complements they are found with. Hanks' approach is justified by the pervasiveness of patterns in corpora: these patterns are instantiated by specific lexical items typically occurring in the argument positions. These items, called fillers, form sets belonging to different patterns of meaning. Hanks and Pustejovsky (2005) and Hanks and Jezek (2008) propose an ontology where fillers are clustered into semantic types, i.e. categories such as [[Location]], [[Event]], [[Vehicle]], [[Emotion]]. These form a hierarchy that branches into more specific types.

The authors, however, warn that these categories are problematic, as lexical sets tend to "shimmer" (Jezek and Hanks 2010): their membership tends to change according to the verb they associate with. For example, two verbs, *wash* and *amputate*, both typically select [[Body Part]] as their direct object, but they select different prototypical members of the set, as can be seen in the examples below from Jezek and Hanks (2010):

(1) [[Human]] *wash* [[Body Part]]
where [[Body Part]]: {<u>hand</u>, hair, face, foot, mouth...}

(2) [[Human]] *amputate* [[Body Part]]
where [[Body Part]]: {leg, limb, arm, <u>hand</u>, finger...}

The "shimmering nature" of lexical sets is not an accidental phenomenon. Rather, it stems from the fact that verb selectional restrictions may cut across conceptual categories due to the specific predication introduced by the verb. In this perspective, this property provides an ideal entry point for studying the interplay between the lexicon and cognition. In our study, we aim at contributing to clarify the notion of prototype, which is inherent in the shimmering nature of lexical sets, by grounding it on empirical evidence exploiting the tools and methodologies offered by recent developments in distributional semantics.

1.2 Causative-Inchoative Verbs

In principle, lexical sets can be constructed for every verb. In this work, however, we limit our inquiry to a specific subset of verbs, namely causative-inchoative verbs in Italian. The choice of this specific subset of verbs is due to the fact that they were investigated thoroughly in the literature, and nonetheless many uncertainties still surround their properties. This provides a testbed for our method, which can be easily extended to other domains.

Causative-inchoative verbs appear either as transitive or intransitive. In the first case, an agent brings about a change of state; in the second, the change of a patient is presented as spontaneous (e.g. *to break*, as in "Mary broke the key" vs. "the key broke"). The two alternative forms of these verbs can be morphologically symmetrical or asymmetrical: if so, one has a derivative affix and the other does not. The so-defined set of alternating verbs varies cross-linguistically (Montemagni, Ruimy, and Pirrelli 1995). Alternations regarding physical change-of-state and manner-of-motion are found in English, whereas they are limited to psychological and physical changes-of-state in Italian. In Japanese and Salish languages, also verbs like *to arrive* and *to appear* do alternate (Alexiadou 2010). From a semantic point of view, Italian causative-inchoative verbs are required to be telic and have an inanimate patient (Cennamo 1995). Morpho-syntactically, they are generated from an asymmetrical derivation, called "anti-causativisation." The intransitive form is marked since it is sometimes preceded by the clitic *si*: its presence is mandatory, optional or forbidden according to verb-specific rules (Cennamo and Jezek 2011). Because of this, many different categorisations of Italian causative-inchoative verbs were attempted (Folli 2002; Jezek 2003).

Causative-inchoative verbs in general are endowed with peculiar properties. Haspelmath (1993) showed that verbs with a cross-linguistic preference for a marked causative form denote a more "spontaneous" situation. Spontaneity is intended by the author as the likelihood of the occurrence of the event without the intervention of an agent. In this way, a correlation between the form and the meaning of these verbs was borne out. Moreover, Samardzic and Merlo (2012) and Haspelmath et al. (2014) demonstrated that verbs appearing more frequently (intra- and cross-linguistically) in the inchoative form tend to derive the causative form morphologically. Here, the correlation holds between form and frequency. Vice versa, situations entailing an agentive participation prefer to mark the inchoative form and occur more frequently in the causative form. These two results can arguably establish important generalisations regarding the meaning, form, and frequency of causative-inchoative verb forms. However, the role of spontaneity remains unclear: this work is non-committal with respect to the role it plays in semantics. Rather, it is considered a notion useful for labelling the observed variations in morphology and frequency.

1.3 Word Embedding

Once established the domain, we need to provide a reliable method of inquiry. The automatic classification of verbs into classes, such as causative-inchoative verbs, was performed successfully based on sub-categorisation frames and selectional preferences, concepts partly overlapping with lexical sets (Joanis, Stevenson, and James 2008). Most notably, the lexical items were compared via distributional semantics (McCarthy 2000) and induced from automatic parses of heterogeneous and wide corpora (Schulte Im Walde 2000). These results validated empirically the claim that verbs in the same class share similar argument patterns. Moreover, they proved that lexical sets can be characterised through distributional semantics fruitfully.

Our work is inspired by these attempts, but its direction of research is reversed. Indeed, rather than classifying verbs given the information about their arguments, it analyses the arguments given a specific verb class, in order to shed light on their properties from the perspective of linguistic theory. In general, we aim at grounding the notion of lexical sets on empiric evidence and clarifying their connection with the notion of prototypicality.

The full exploitation of the semantic information inherent to argument fillers for verbs can take advantage from some recent developments in distributional semantics. Recently, efficient algorithms have been devised to map each word of a vocabulary into a corresponding vector of n real numbers, which can be thought as a sequence of coordinates in a n-dimensional space (Mikolov et al. 2013). This mapping is yielded by unsupervised machine learning, based on the assumption that the meaning of a word can be inferred by its context, i.e. its neighbouring words in texts. This is known as Distributional Hypothesis (Harris 1954). Distributed models have some relevant properties: the geometric closeness of two vectors corresponds to the similarity in meaning of the corresponding words. Moreover, its dimensions possibly retain a semantic interpretation such that non-trivial analogies can be established among words.

2. Data and Method

The data are sourced from a sample of ItWac, a wide corpus gathered by crawling texts from the Italian domain in the web using medium frequency vocabulary as seeds (Baroni et al. 2009). This sample was further enriched with morpho-syntactic information through the MATE-tools parser (Bohnet 2010)[1] and filtered by sentence length (< 100). Eventually, sentences in the sample amounted to 2,029,454 items. A target group of 20 causative-inchoative verbs was taken from Haspelmath et al. (2014): they are listed in Table 1, together with the count of the extracted lexical sets for the relevant semantic macro-roles (see below).

lemma	**translation**	S	O
aprir(si)	to open	195	1337
scuoter(si)	to rock	10	69
affondare	to sink	18	74
(far) bollire	to boil	2	2
girare	to spin	155	243
raccoglier(si)	to gather	85	505
uscire/portare fuori	to go/put out	325	638
svegliar(si)	to awake/wake	68	89
romper(si)	to break	83	419
connetter(si)	to connect	39	134
divider(si)	to split	129	246
chiuder(si)	to close	289	606
seccare	to dry	15	14
congelare	to freeze	10	30
alzar(si)	to arise/raise	75	304
finire	to stop	1092	721
riempir(si)	to fill	58	166
aumentare	to improve	534	998
scioglier(si)	to melt	94	143
bruciare	to burn	75	174

Table 1
List of 20 causative-inchoative verbs and the count of their fillers for each argument position.

1 LAS scores for the relevant dependency relations: 0.751 with dobj (direct object), 0.719 with nsubj (subject), 0.691 with nsubjpass (subject of a passive verb).

The procedure for the automatic extraction of the fillers consisted in the following steps. Once the sentences had been parsed, the target predicates were identified inside the dependency trees. The lemmas of these verbs and the forms of their arguments were stored in a database. Argument fillers were grouped not according to their grammatical relation, but rather to their semantic macro-roles according to (Dixon 1994): subjects of transitive verbs (A), subjects of intransitive verbs (S) and objects (O). In particular, the subjects of verb forms accompanied by the *si*-clitic were treated as intransitive subjects.[2] These operations resulted in a database structured as a list: in each row a verb is followed by the fillers it is associated with in a specific text occurrence. For example, compare an original sentence and its corresponding entry:

(3) *Plinio il* ***Vecchio*** *non* ***cita*** *più il* ***Po*** *di Adria*
Pliny the Elder doesn't mention anymore the Po of Adria
perché l' ***Adige*** *aveva* ***subito*** *una* ***rotta*** *ed era*
because the Adige had undergone an overflowing and had
confluito nella Filistina.
merged with the Filistina.

V	A	S	O
citare	Vecchio	_	Po
subire	Adige	_	rotta

The database was later collapsed by verb lemma so that each became associated to three sets of fillers (one per macro-role). Each of these sets is a corpus-based lexical set. Compared to manually picked lexical sets, they are more noisy but less sparse: the vastness of the data mitigates the errors in the parsing step. Moreover, the automation in lexical set extraction allows to access the fillers of virtually every verb: resources based on manual selection like T-PAS (Jezek et al. 2014), on the other hand, are limited to a small amount of verbs.

Afterwards, each of the argument fillers was mapped to a vector relying on a space model pre-trained through Word2Vec (Dinu, Lazaridou, and Baroni 2015). It was generated by a CBOW algorithm with negative sampling, 300 dimensions, a context window of 10 tokens, pruning of infrequent words and sub-sampling. In order to compare the elements in this space models, many different metrics are available, including pure geometrical (Euclidean) distance. In this work, we rely on the popular metrics of cosine distance. Assume that **a** and **b** are vectors, and that $\mathbf{a}_i$ and $\mathbf{b}_i$ are their i^{th} components, respectively. The cosine similarity between **a** and **b** is then defined as follows:

$$\cos(\theta) = \frac{\sum_{i=1}^{n} \mathbf{a}_i \mathbf{b}_i}{\sqrt{\sum_{i=1}^{n} \mathbf{a}_i^2} \sqrt{\sum_{i=1}^{n} \mathbf{b}_i^2}} \tag{1}$$

The opposite, namely cosine distance d, is simply defined as $-\cos(\theta)$. As for the values that $d(\mathbf{a},\mathbf{b})$ can assume, the minimum is at 0 (angles completely overlap) and the maximum is at 1 (orthogonal vectors).

2 Subjects of verbs inflected in the passive voice were treated as objects, instead.

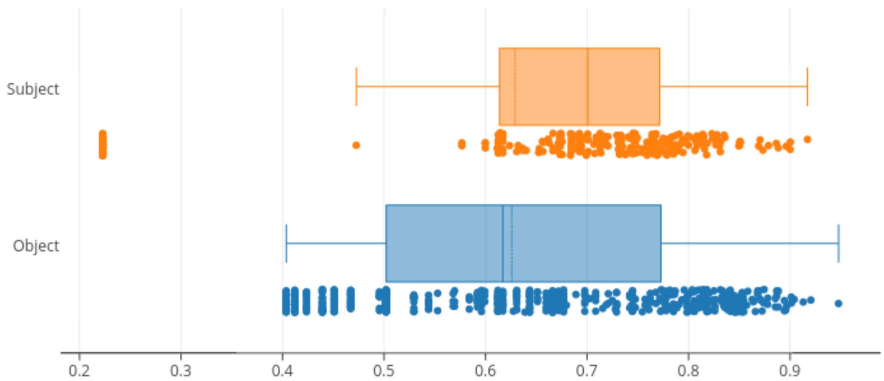

Figure 1
Cosine distance of vectors corresponding to the fillers of a verb from their centroid.

3. Experiments

In order to bring to light the linguistic information concealed in the automatically extracted lexical sets, we devised two experiments. One investigates the internal structure of lexical sets. In fact, previous works based on set theory treated them as categoric sets, of which a filler is either a member or not. Research in psychology, however, has long since demonstrated that the members of a linguistic set are found in a radial continuum where the most central element is the prototype for its category, and those at the periphery are less representative (Rosch 1973; Lakoff 1987).[3] Word vectors allow to capture this spatial continuum. The second experiment, moreover, is aimed at studying how the centroids of different argument slots are related. In particular, we compute the cosine distance between the centroids of the fillers for the roles of intransitive subject and transitive object.

3.1 Distance of Set Members from Centroid

Once the fillers have been mapped to their corresponding vectors, a lexical set appears as a group of points in a multi-dimensional model. The centre of this group is the Euclidean mean among the vectors, which is a vector itself and is called centroid. In the first experiment, we evaluate the co-ordinates of the centroid of the lexical sets of the instransitive subject (S) and the object (O) for each verb[4]. Then we measure the cosine

3 For previous work on lexical sets considering prototypicality in the context of the notion of shimmering, see Jezek and Hanks (2010).
4 Every filler was weighted proportionally to its absolute frequency.

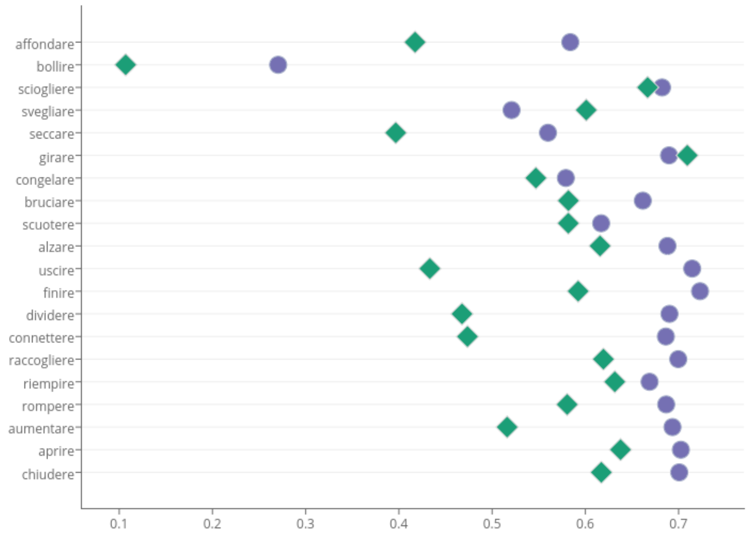

Figure 2
Median value of cosine distances of intransitive subject set (blue circles) and the object set (green diamonds) for verbs in Haspelmath's scale.

distance of every vector member of the sets from its centroid. In semantic terms, this should correspond to assessing how far a filler is from its prototype.

We obtained two sets of cosine distance values for each verb: these can be plotted as boxes and whiskers, like in Figure 1. The example represents the distances of each filler from the two centroids of *dividere* 'to split'. The boxes contain the values in the second and third quartiles, whereas the vertical lines stand for the median and the extremes[5]. From all these distance values, we picked the median value of each lexical set. The plot of these medians for the S set and the O set of each verb ordered according to Haspelmath's ranking is shown in Figure 2.

The questions at the heart of this experiment were: how are lexical set structured? Do their elements distribute uniformly in the space, or rather gather near or far the prototype? An important result can be observed from these plots: the S lexical set lies in a more compact range of distances, whereas O is more scattered. On the other hand, the vectors of S tend to be farther from the centroid, as demonstrated by the ranges where their distance values fall. This implies that O behaves more similarly to a radial category, whereas S just populates the periphery. In fact, the medians of the S sets are quite always higher than the median of O for the same verb.

5 The median is the value separating the higher half of the ordered values from the lower half.

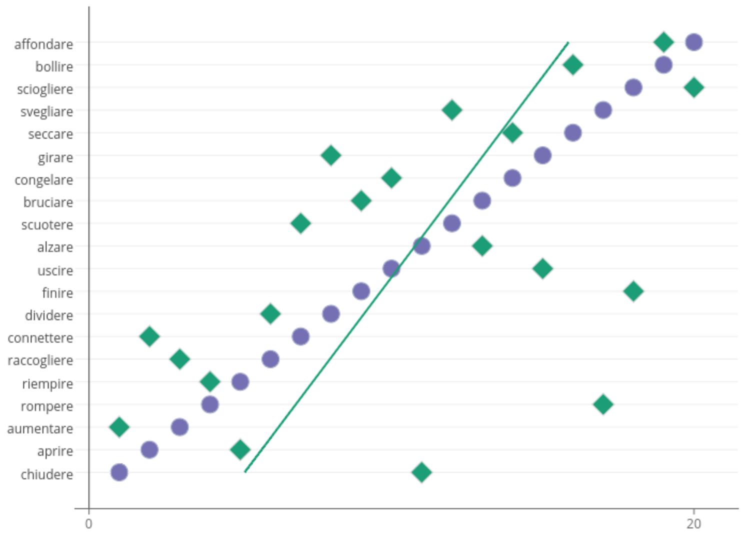

Figure 3
Ranking based on cross-linguistic form frequencies (blue circles) against ranking based on distance between the centroids of S and O in Italian (green diamonds).

3.2 Distance of Centroid Pairs

Thus far, lexical sets of the same verb have been considered independently from each other. However, a comparison is needed to assess whether any relation holds between them. The second experiment consisted in estimating the cosine distance between the centroid of S and the centroid of O for each verb. This operation was aimed at finding to which extent the lexical sets of S and O overlap. In fact, Montemagni, Ruimy, and Pirrelli (1995) and McCarthy (2000) assessed some asymmetries between these lexical sets, which in principle should share all their members.

Inspecting our results, the distance between S and O seems to behave as a measure of spontaneity, intended as cross-linguistic frequency of the forms of a verb: the more the centroids tend to be set apart, the more the verb tends to have a more frequent intransitive form. In fact, we compared the ranking of 20 alternating verbs according to the ratio of the cross-linguistic frequency of their transitive and intransitive forms (Haspelmath et al. 2014) and a ranking based on the centroid distances of the same verbs. Both these rankings are plotted in Figure 3: every verb is associated with its position in the two scales.

Both scales display a common tendency. In particular a Spearman's ranking test was performed over them, yielding a mild positive correlation of $\rho = 0.56391$ with a quite strong confidence, i.e. with $p < 0.01$. In order to compare our method based on distributional semantics with plain set analysis, we considered an alternative measure for the ranking: the cardinality of the intersection $S \cap O$ weighted by the union $S \cup O$.

In this case, Spearman correlation was $\rho = 0.42255$, but it was not significant because of $p \approx 0.06$. It should be also noted that our ranking sometimes diverges dramatically from the frequency-based ranking: for instance, *romper(si)* 'to break' is highly agentive according to its behaviour in frequency, but it is spontaneous according to our criterion.

4. Conclusions

Our work provided evidence that lexical sets of Italian causative-inchoative verbs are non-uniform categories, whose distribution around the prototype varies to a great extent. This distribution is sensitive to the argument slot: intransitive objects display a more homogeneous one, whereas all the fillers of intransitive subjects lie on the outskirts of the category. This difference might be due to different selectional restrictions applied to the subjects. In alternative, it is possibly an artifact due to the fact that the sample of objects is usually wider and hence more representative.

Moreover, a correlation was discovered between the cosine distance of lexical sets of a given verb and the cross-linguistic behaviour of its forms, i.e. the tendency to appear more frequently as intransitive or as transitive. Possibly, the correlation is not stronger for a series of reasons. On the one hand, our lexical sets might be noisy, or not rich enough to guarantee full representativeness. Moreover, the choice of inter-linguistic equivalent translations may be faulty, because of some possible difference in meaning. Finally, Haspelmath's scales may approximate the cross-linguistic tendencies of morphology and frequency to an unsatisfactory extent, since the samples of languages they rely upon are limited to 21 for morphology and 7 for frequency. Figure 4 shows a synopsis of our result in the context of the correlations established in previous works.

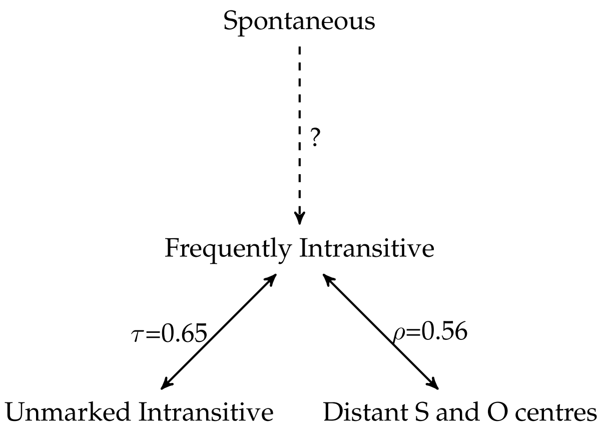

Figure 4
Synopsis of the correlations among features of causative-inchoative verbs. The measures are based on Kendall Tau test (τ) and Spearman's ranking test (ρ).

In Figure 4, solid lines stand for correlations proven based on cross-linguistic evidence (frequency-form) and evidence from the Italian language (frequency-lexical sets). The dotted line, on the other hand, suggests the existence of and underlying motivation for the correlations (i.e. spontaneity), which nonetheless remains unproven and undetermined in its nature. Its possible validation is left to future research, but remains tricky due to its purely semantic nature.

To amend the above-mentioned limitations, further research should resort to an enhanced database with a wider sample, try to reduce the parsing error (and hence filler identification error), and add sense disambiguation for polysemous word forms (Grave, Obozinski, and Bach 2013). Also, it should choose different pre-trained vector models, in order to try and replicate these results. In particular, the new vector models could be optimized for similarity through semantic lexica (Faruqui et al. 2015) or based on syntactic dependencies (Séaghdha 2010). The experiments in this work may be extended to other languages, either individually or through a multi-lingual word embedding model (Faruqui and Dyer 2014).

Finally, the results in this work may turn out to be relevant for other areas, such as applications in natural language processing and neuro-linguistics. In particular, our method provides insights about the relation between fillers and causal alternations: the role of event participants is also crucial in identifying causal relations in discourse (Ponti and Korhonen 2017). Moreover, our method can provide a quantitative measure to compare linguistic prototypes to cognitive prototypes in behavioural experiments, similarly to what has been attempted through the analysis of linguistic networks (Duch, Matykiewicz, and Pestian 2008).

References

Alexiadou, Artemis. 2010. On the morpho-syntax of (anti–) causative verbs. *Syntax, lexical semantics and event structure*, pages 177–203.

Baroni, Marco, Silvia Bernardini, Adriano Ferraresi, and Eros Zanchetta. 2009. The wacky wide web: a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation*, 43(3):209–226.

Bohnet, Bernd. 2010. Very high accuracy and fast dependency parsing is not a contradiction. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 89–97. Association for Computational Linguistics.

Cennamo, Michela. 1995. Transitivity and vs order in italian reflexives. *Sprachtypologie und Universalienforschung*, 48(1-2):84–105.

Cennamo, Michela and Elisabetta Jezek. 2011. The anticausative alternation in italian. *I luoghi della traduzione*, pages 809–823.

Dinu, Georgiana, Angeliki Lazaridou, and Marco Baroni. 2015. Improving zero-shot learning by mitigating the hubness problem. *workshop contribution at ICLR 2015*.

Dixon, Robert MW. 1994. *Ergativity*. Cambridge University Press.

Duch, Włodzisław, Paweł Matykiewicz, and John Pestian. 2008. Neurolinguistic approach to natural language processing with applications to medical text analysis. *Neural Networks*, 21(10):1500–1510.

Faruqui, Manaal, Jesse Dodge, Sujay K. Jauhar, Chris Dyer, Eduard Hovy, and Noah A. Smith. 2015. Retrofitting word vectors to semantic lexicons. In *Proceedings of NAACL*.

Faruqui, Manaal and Chris Dyer. 2014. Improving vector space word representations using multilingual correlation.

Folli, Raffaella. 2002. *Constructing Telicity in English and Italian*. Ph.D. thesis, Oxford University.

Grave, Edouard, Guillaume Obozinski, and Francis Bach. 2013. Hidden markov tree models for semantic class induction. In *CoNLL-Seventeenth Conference on Computational Natural Language Learning*.

Hanks, Patrick. 1996. Contextual dependency and lexical sets. *International Journal of Corpus Linguistics*, 1(1):75–98.

Hanks, Patrick and Elisabetta Jezek. 2008. Shimmering lexical sets. In *Proceedings of the XIII EURALEX International Congress (Barcelona, 15-19 July 2008)*, pages 391–402.

Hanks, Patrick and James Pustejovsky. 2005. A pattern dictionary for natural language processing. *Revue Française de linguistique appliquée*, 10(2):63–82.

Harris, Zelig. 1954. Distributional structure. *Word*, 10(23):146–162.

Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. *Causatives and transitivity*, 23:87.

Haspelmath, Martin, Andreea Calude, Michael Spagnol, Heiko Narrog, and Elif Bamyaci. 2014. Coding causal–noncausal verb alternations: A form–frequency correspondence explanation. *Journal of Linguistics*, 50(03):587–625.

Jezek, Elisabetta. 2003. *Classi di verbi tra semantica e sintassi*. Ed. ETS.

Jezek, Elisabetta and Patrick Hanks. 2010. What lexical sets tell us about conceptual categories. *Lexis*, 4(7):22.

Jezek, Elisabetta, Bernardo Magnini, Anna Feltracco, Alessia Bianchini, and Octavian Popescu. 2014. T-pas: A resource of corpus-derived types predicate-argument structures for linguistic analysis and semantic processing. In *Proceedings of LREC*.

Joanis, Eric, Suzanne Stevenson, and David James. 2008. A general feature space for automatic verb classification. *Natural Language Engineering*, 14(03):337–367.

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Cambridge University Press.

McCarthy, Diana. 2000. Using semantic preferences to identify verbal participation in role switching alternations. In *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference*, pages 256–263.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. In *Workshop at ICLR*.

Montemagni, Simonetta, Nilda Ruimy, and Vito Pirrelli. 1995. Ringing things which nobody can ring. a corpus-based study of the causative-inchoative alternation in italian. *Textus online only. 8 (1995), N. 2, 1995*, 8(2):1000–1020.

Ponti, Edoardo Maria and Anna Korhonen. 2017. Event-related features in feedforward neural networks contribute to identifying causal relations in discourse. In *Proceedings of EACL, Workshop on Linking Models of Lexical, Sentential and Discourse-level Semantics*, pages 25–30.

Rosch, Eleanor H. 1973. Natural categories. *Cognitive psychology*, 4(3):328–350.

Samardzic, Tanja and Paola Merlo. 2012. The meaning of lexical causatives in cross-linguistic variation. *Linguistic Issues in Language Technology*, 7(12):1–14.

Schulte Im Walde, Sabine. 2000. Clustering verbs semantically according to their alternation behaviour. In *Proceedings of the 18th conference on Computational linguistics-Volume 2*, pages 747–753.

Séaghdha, Diarmuid O. 2010. Latent variable models of selectional preference. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 435–444.