Tracing metaphors in time through self-distance in vector spaces

Marco Del Tredici
ILLC, Univ. of Amsterdam
Amsterdam, The Netherlands
marcodeltredici@gmail.com

Malvina Nissim
CLCG, Univ. of Groningen
Groningen, The Netherlands
m.nissim@rug.nl

Andrea Zaninello
Zanichelli editore
Bologna, Italy
azaninello@zanichelli.it

Abstract

English. From a diachronic corpus of Italian, we build consecutive vector spaces in time and use them to compare a term's cosine similarity to itself in different time spans. We assume that a drop in similarity might be related to the emergence of a metaphorical sense at a given time. Similarity-based observations are matched to the actual year when a figurative meaning was documented in a reference dictionary and through manual inspection of corpus occurrences.

Italiano. *Nel presente esperimento costruiamo spazi vettoriali progressivi nel tempo su un corpus diacronico dell'italiano e calcoliamo la distanza di alcuni termini rispetto a loro stessi in differenti periodi. L'ipotesi è che un calo di similitudine possa essere indicativo dell'acquisizione di un significato metaforico. Tale ipotesi è valutata attraverso una risorsa lessicografica esterna e l'annotazione manuale dei contesti dei termini nel corpus.*

1 Introduction

It is widely acknowledged that metaphors are pervasive in language use, and that their detection and interpretation are crucial to language processing (Group, 2007; Turney et al., 2011; Shutova, 2015).

One tricky aspect related to metaphors is their dynamic nature: new metaphors are created all the time. For example, in recent years the Italian term "talebano" ('Taliban'), previously only used to refer to the Islamic fundamentalist political movement founded in the Nineties in Afghanistan (Example 1), has come to define more generally someone who is extreme in his or her positions, for example regarding food, use of medicines, and the like (Example 2).[1]

(1)   (*lit.*) l'operazione [...] ha permesso di arrestare un **talebano** esperto in esplosivi

(2)   (*fig.*) [...] senza l'atteso top player, e di un allenatore **talebano** della tattica

If the metaphorical meaning becomes commonly used, it might get recorded in reference dictionaries, too. Indeed, for the case of "talebano" the Italian dictionary Zingarelli (Zingarelli, 1993–2017) has recorded the metaphorical extension ("che (o chi) è dogmatico, integralista") in the year 2009, while until then only the literal meaning was included.

Most of the computational work on metaphors has focused on their identification and interpretation using a variety of techniques and models, such as clustering (Shutova and Sun, 2013), LDA topic modeling (Heintz et al., 2013), tree kernels (Hovy et al., 2013), but all from a purely synchronic perspective.[2] The way metaphors develop across time, instead, and whether the shift of a word's literal meaning to a figurative one can be automatically detected and modelled is as of now a little investigated aspect.

As a contribution in this sense, we build on the basic observation that if a metaphorical meaning is acquired by a term at a certain point in time, the context of use of that term will, at least partially, change. In this paper we offer a proof of concept of this assumption, based on a selection of terms. (Dis)similarity of contexts is measured relying on the distributional semantics approach, and thus on the terms' vector representations, and the existence of a metaphoric shift is derived from the Zingarelli dictionary of Italian.

---

[1]All of the examples in this paper are from the newspaper *la Repubblica*, see Section 4.2.

[2]For a detailed survey on current NLP systems for metaphor modeling see (Shutova, 2015).

2 Approach

According to the principle of distributional semantics, the meaning of a word is represented by vectors that encode the contextual information of that word in a corpus (Turney et al., 2010). All vectors representing words are included in a distributional semantic space in which similar words are represented by vectors that are close in that space, while different words are distant.

We rely on the intuition that if a term develops a metaphoric sense, its contexts of occurrence will start to differ, at least partially, from those observed for the very same term at the time the metaphorical meaning had not emerged yet. This implies that detecting a distance in space across time could be indicative of a meaning shift. Hence, instead of comparing different terms synchronically, we focus on their *self-distance* across time, thus tracing their diachronic evolution of meaning.

Practically, we train vector representations of words in consecutive time spans, and compare such representations to one another, for a set of pilot terms. As a default, a term is expected to exhibit a vector representation roughly similar to itself across time. If we observe a drop in similarity between vectors in consecutive spaces, we can hypothesise the emergence of a new sense for this term, potentially metaphoric.

By using the information recorded for the selected terms in a reference dictionary for the Italian language, we observe whether there is some correspondence between the observed similarity drop, if present, and the time of inclusion of a figurative sense. Finally, for each year cluster, we manually inspect the occurrences of our target terms in order to see if changes of use can be observed.

We are aware of the fact that changes in distance of a word to itself across time might be triggered by phenomena other than the rise of a metaphoric shift. Indeed, especially for polysemous words, extra-linguistic factors could cause the dominance of one sense over the others at a given time. In a larger-scale, bottom-up approach to detect metaphorical shifts, this would need to be properly accounted for. In the context of this proof-of-concept, we control for this factor by choosing words that are not or are minimally polysemous (see Section 4.1).

3 Related Work

The automatic modelling of diachronic shift of meaning has been investigated employing several different techniques. Among these, most recently, Latent Semantic Analysis (Sagi et al., 2011; Jatowt and Duh, 2014), topic clustering (Wijaya and Yeniterzi, 2011) and dynamic topic modeling (Frermann and Lapata, 2016). Vector representations for diachronic shift of meaning have been used by Gulordava and Baroni (2011), with a simple co-occurence matrix of target words and context terms. Jatowt and Duh (2014) and Xu and Kemp (2015) experimented both with a bag-of-words approach and a more linguistically motivated representation that also captures the relative position of lexical items in relation to the target word.

Recently, Word Embeddings (Mikolov and Dean (2013), see also Section 4.3) have been used to investigate diachronic meaning shifts: vectors are usually created independently for each time span and then mapped from one year to another via a transformation matrix, thus leveraging the stability of the relative positions of vectors in different spaces (Kulkarni et al., 2015; Zhang et al., 2015; Hamilton et al., 2016).

An alternative approach, which we also adopt – with a slight change – in our work, is introduced by Kim et al. (2014), who propose a simple but effective methodology to make vectors trained on different corpora directly comparable: embeddings created for year y are used to initialise the vectors for year $y + 1$. The process is progressively applied to all time spans.

4 Experiment

Following the approach described in Section 2, we selected a small set of pilot terms from a lexicographic reference, and observed their space development across time, on a diachronic corpus for Italian that we collected for this purpose. Due to the absence of datasets in which words are annotated for meaning change, a qualitative analysis of a set of hand-selected words like the one we propose has established itself as a common evaluation method in previous work on diachronic meaning change (Frermann and Lapata, 2016).

4.1 Lexicographic reference and term selection

The Zingarelli dictionary is a reference dictionary for the Italian language, updated and published every year, both in digital and paper version. The dictionary is traditionally dated one year ahead of the year it is published, hence the Zingarelli 2017

Table 1: Selected terms. **a-date** = first attested; **d-date** = decision date for extended meaning to be included in dictionary; **i-date** = actual inclusion date in Zingarelli for extended meaning.

term	literal	figurative	a-date	d-date	i-date
implosione	implosion	cedimento, tracollo improvviso (collapse)	1932	2013	2015
kamikaze	kamikaze	chi compie un'impresa rischiosa o destinata al fallimento (daredevil, reckless)	1944	2007	2009
rottamatore	dismantler	nel linguaggio giornalistico e della politica, chi si propone di allontanare e sostituire un gruppo dirigente considerato antiquato (new broom)	1990	2012	2014
talebano	Taliban	che (o chi) è dogmatico, integralista (hard-liner, extremist)	1995	2007	2009
tsunami	tsunami	evento che determina lo sconvolgimento di un assetto costituito (devastation, havoc)	1907	2008	2010

is published in June 2016, and it refers to decisions about new words and new meanings (including metaphorical ones) made up until December 2015.

We analysed the behaviour of a small set of terms extracted from the dictionary. We searched the 2017 edition to extract nouns that record a figurative meaning, limiting our search to words whose first occurrence is recorded in the 20th or 21st century. Newly born words (including borrowings) are more likely to show a meaning shift in the time span considered in our search (1984-2015) than older words (especially if derived directly form Latin, where the figurative meaning was also originally highly available, so probably arisen earlier). Out of a total of 447 hits, five target words were chosen for this pilot study. They are reported in Table 1 together with relevant information.

In order to minimise (at least in the context of this experiment) the influence of polysemy in the observable similarity distance across years, we verified that the selected terms are not polysemous, or minimally so. For the words "rottamatore", "talebano", and "tsunami", the Zingarelli records one sense only. For the word "implosione" three senses in total are recorded, two of which are however technical language, in the fields of linguistics (phonology) and psychology, and we assume will not be used much in newswire. For "kamikaze" the Zingarelli records one meaning only (Japanese pilot) to which is associated the extended sense of someone who kills himself in a terrorist attack; in our corpus the extended meaning is clearly the primary one, and the figurative sense that we consider is derived from it (see also Section 4.4).

4.2 Corpus

We created a diachronic corpus of approximately 60 millions tokens by collecting articles from the Italian newspaper *la Repubblica* from 1984 (the first year for which data is available digitally) to 2015. All texts were tokenised and lowercased. Because we are interested in how a term's context changes over time, we had to determine time-spans for our corpus, and we settled on two-year blocks, for a total of 16 time spans, the first one being 1984-1985 and the last 2014-2015. These subcorpora are used to train consecutive vector space models.

4.3 Model

We implemented vector representations using the skip-gram architecture introduced by Mikolov and Dean (2013). Such representations (Word Embeddings) are low dimensional, dense and real-valued vectors that have been proved to preserve syntactic and semantic information in several NLP tasks (Baroni et al., 2014).

Vectors created on different corpora cannot be directly compared, since every semantic space implements arbitrary orthogonal transformations and hence there is no direct correspondence between word vectors in different semantic spaces (Zhang et al., 2015). This would hold true also for our data, since we create a different corpus for each time span. Therefore, in order to create comparable vector representations for each word in any time span, we adopt the methodology introduced by Kim et al. (2014) (see Section 3), slightly modifying it. While Kim et al. (2014) use vectors of span y to initialise the vectors for year $y + 1$, we do the opposite, i.e.

we start with 2014-15, and use those vectors to initialise the 2012-13 time span, and thus backwards until 1984-85.

This methodological choice is due to the fact that the majority of the words in the set we considered for this experiment (included the selected target words, see 4.1) have few or no occurrences in the first time spans of the corpus: for example, "rottamatore" and "talebano" occur for the first time in 96/97. Indeed, using Kim et al. (2014)'s original approach, which we implemented in a preliminary experiment, the vectors for these words were correctly initialised, but were basically random vectors with no meaningful information. Conversely, our *reverse* setting, while still offering the same opportunity to trace shifts of meaning across time, allows to initialise all target words on a time span (14/15) in which they occur a number of times sufficient to create a more stable, meaningful representation.

Using the `gensim` library (Řehůřek and Sojka, 2010), we trained the models with the following parameters: window size of 5, learning rate of 0.01 and dimensionality of 200. We filtered out words with frequency lower than 5 occurrences. The vocabulary was initialised over the whole dataset.

4.4 Results and discussion

Figure 1 shows the similarity values for one time span to the next (dotted line), together with the average shift of meaning of a subset of 5000 nouns randomly selected (solid line). While we cannot draw any statistically significant conclusions from such little data, we aim at potentially observing patterns of shift of meaning through change of vector representations that could be used for developing predictive metrics of metaphorical shifts in time.

We interpret the results of our models according to (i) information in the Zingarelli dictionary and (ii) a manual inspection of the context of use of our target words in the corpus.

For (i), we verify if, for a given term, an observable correlation exists between changes in its vector representations and the insertion of a figurative sense in the dictionary. Results show that such a correlation exists for "talebano", "rottamatore", and "tsunami". For these words a drop in cosine similarity can be observed between three and five years before the insertion of the figurative meaning in the dictionary. This fits well with the timing for new meanings to be recorded in lexicographic resources (see Section 4.1). The nouns "kamikaze"

and "implosione", instead, show a more stable evolution of meaning in time, with no clear drop in cosine similarity, and thus no evident correlation between changes in vector representations and insertion of a figurative meaning in dictionary.

For (ii), we manually inspected the contexts in which target terms occur in the the corpus as literal or metaphoric, in order to check if some relevant change in words usage could be observed in correspondence to drops in cosine similarity between time spans.

"Tsunami" occurs 27 times between 84/85 and 02/03: in 88.9% of the cases the word is used literally, with only 3 metaphorical uses in 98/99 (mirrored in a slight drop in cosine similarity). Of the 930 occurrences from 04/05 to 14/15, only 59.1% are literal. In Figure 1 we can observe a major drop in cosine similarity exactly between 04/05 and 06/06.

"Rottamatore" occurs 4 times between 84/85 and 08/09, always used literally. From 10/11 on, there are 156 occurrences, all metaphorical. Thus, the drop corresponds to change in usage here too.

"Talebano" occurs 12 times between 84/85 and 02/03, with 83.3% of literal usage. Once again, the drop in cosine coincides with the time span in which the term started to be used metaphorically: between 02/03 and 08/09 40% of the occurrences of "talebano" are metaphorical. Then, another relevant drop is observed between 08/09 and 10/11, and this is due to the sudden return of the literal usage of this word (86.1%), which continues also in the following years.

As already noticed, "kamikaze" and "implosione" do not seem to undergo a clear shift. As for the former, the analysis of its contexts of use reveals that indeed it is not possible to clearly identify, in our corpus, when exactly the term started to be used metaphorically: of the 25 occurrences of "kamikaze" in 84/85, 32% are metaphorical. This trend is fairly constant, and it explains why the vector representation of "kamikaze", which from the very beginning conflates literal and metaphorical usages, is stable in time. There is only a relevant change starting from 10/11: from this period onwards, the metaphorical use decreases, and almost all the occurrences are literal.[3] Accordingly, this

---

[3]Interestingly, this increase of literal usage is observed in the same years also for "talebano", a term that is semantically related to "kamikaze". This observation would require further investigation in connection with the socio-political events of those time spans.
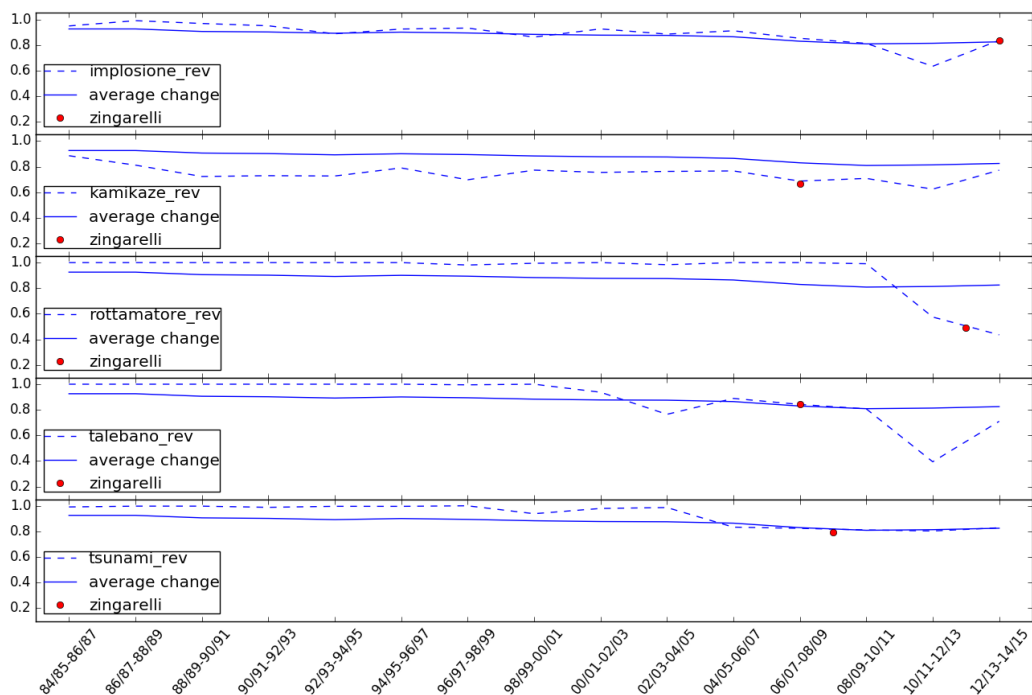
Figure 1: Cosine similarity values across time spans for target words (dotted line), average similarity of nouns (solid line) and date of insertion of metaphorical meaning in the Zingarelli dictionary (red dot).

almost exclusively return to the literal meaning corresponds to a slight increase in cosine similarity between the two last time spans.

"Implosione" occurs 433 times overall and in 92.4% of them is used metaphorically, but in few and specific contexts. A metaphorical, quite specific, sense of "implosione" is thus the main sense for this term in our corpus, and this is why we observe, on average, a high similarity across time spans. There is only a small drop between 10/11 and 12/13, when the word started to be used in the context of the economical crisis ("l'implosione dell'euro").

To sum up, both "kamikaze" and "implosione" show a similar stable behaviour in time, with only small drops. However, while for "kamikaze" such stability is due to a relatively constant ratio between literal and metaphorical meanings, in the case of "implosione" the observed stability is given by the constant predominance of the metaphorical sense across all the time spans.

5 Conclusion and future work

This work was meant as an exploration of the assumption that the emergence of the metaphorical use of a term might be mirrored in changes in cosine similarity of the term to itself across time. Such assumption has been partially confirmed by the comparison to the Zingarelli dictionary, while we found a more robust evidence when inspecting the terms' contexts of use manually.

Future work will stem from methodology and observations discussed here. Specifically, we plan to investigate further several aspects of this initial work, including the relation between changes in cosine similarity and frequency of use of a word: to which extent a change of the former relates to an increase of the latter? Mostly though, we plan to run experiments on larger sets of words with the aim to consolidate and then further exploit the mainly qualitative observations reported here towards the development of reliable predictive metrics which can serve to detect the emergence of shifts automatically, in a completely bottom-up fashion.

Acknowledgments

References

Marco Baroni, Georgiana Dinu, and Germán Kruszewski. 2014. Don't count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 238–247, Baltimore, Maryland, June. Association for Computational Linguistics.

Lea Frermann and Mirella Lapata. 2016. A bayesian model of diachronic meaning change. *Transactions of the Association for Computational Linguistics*, 4:31–45.

Pragglejaz Group. 2007. MIP: A method for identifying metaphorically used words in discourse. *Metaphor and symbol*, 22(1):1–39.

Kristina Gulordava and Marco Baroni. 2011. A distributional similarity approach to the detection of semantic change in the Google books ngram corpus. In *Proceedings of the GEMS 2011 Workshop on GEometrical Models of Natural Language Semantics*, pages 67–71. Association for Computational Linguistics.

William L Hamilton, Jure Leskovec, and Dan Jurafsky. 2016. Diachronic word embeddings reveal statistical laws of semantic change. *arXiv preprint arXiv:1605.09096*.

Ilana Heintz, Ryan Gabbard, Mahesh Srinivasan, David Barner, Donald S Black, Marjorie Freedman, and Ralph Weischedel. 2013. Automatic extraction of linguistic metaphor with LDA topic modeling. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 58–66.

Dirk Hovy, Shashank Srivastava, Sujay Kumar Jauhar, Mrinmaya Sachan, Kartik Goyal, Huiying Li, Whitney Sanders, and Eduard Hovy. 2013. Identifying metaphorical word use with tree kernels. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 52–57.

Adam Jatowt and Kevin Duh. 2014. A framework for analyzing semantic change of words across time. In *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries*, pages 229–238. IEEE Press.

Yoon Kim, Yi-I Chiu, Kentaro Hanaki, Darshan Hegde, and Slav Petrov. 2014. Temporal analysis of language through neural language models. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, pages 61–65. Association for Computational Linguistics.

Vivek Kulkarni, Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. 2015. Statistically significant detection of linguistic change. In *Proceedings of the 24th International Conference on World Wide Web*, pages 625–635. ACM.

T Mikolov and J Dean. 2013. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*.

Radim Řehůřek and Petr Sojka. 2010. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta, May. ELRA.

Eyal Sagi, Stefan Kaufmann, and Brady Clark. 2011. Tracing semantic change with latent semantic analysis. *Current methods in historical semantics*, pages 161–183.

Ekaterina Shutova and Lin Sun. 2013. Unsupervised metaphor identification using hierarchical graph factorization clustering. In *HLT-NAACL*, pages 978–988.

Ekaterina Shutova. 2015. Design and evaluation of metaphor processing systems. *Computational Linguistics*, 41(4):579–623.

Peter D Turney, Patrick Pantel, et al. 2010. From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research*, 37(1):141–188.

Peter D Turney, Yair Neuman, Dan Assaf, and Yohai Cohen. 2011. Literal and metaphorical sense identification through concrete and abstract context. In *Proceedings of the 2011 Conference on the Empirical Methods in Natural Language Processing*, pages 680–690.

Derry Tanti Wijaya and Reyyan Yeniterzi. 2011. Understanding semantic change of words over centuries. In *Proceedings of the 2011 international workshop on DETecting and Exploiting Cultural diversiTy on the social web*, pages 35–40. ACM.

Y. Xu and C. Kemp. 2015. A computational evaluation of two laws of semantic change. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society*.

Yating Zhang, Adam Jatowt, Sourav S Bhowmick, and Katsumi Tanaka. 2015. Omnia mutantur, nihil interit: Connecting past with present by finding corresponding terms across time. In *Proc. of ACL*, pages 645–655.

N. Zingarelli. 1993–2017. *Lo Zingarelli - Vocabolario della lingua italiana*. Zanichelli editore, Bologna.