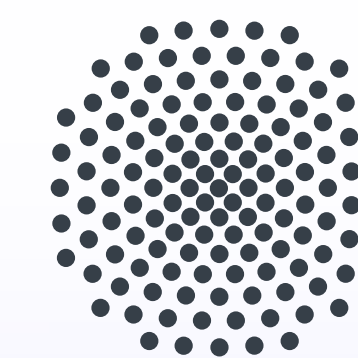


An Annotation of Semantic Change based on Usage Relatedness

Dominik Schlechtweg*, Sabine Schulte im Walde*, Stefanie Eckmann†

*Institute for Natural Language Processing, University of Stuttgart

†Historical and Indo-European Linguistics, LMU Munich



University of Stuttgart
Germany

Introduction

- **aim:**
 - build a computational model detecting semantic change in corpora
- **benefits:**
 - test theories empirically on a large scale
 - inspire research by detecting new changes
- **problem:**
 - models should be evaluated on a sufficient number of semantic developments
 - there is no reliable test set of semantic change for any language
- **solution:**
 - collect a set of words displaying semantic change in a specific corpus in an annotation study

Related Work

- various polysemy annotation studies of use pairs on proximity scales
- Soares da Silva (1992), Brown (2008): semantic relatedness
- Erk, McCarthy, and Gaylord (2009, 2013): semantic similarity

Annotation Scale

- 4: Identical
3: Closely Related
2: Distantly Related
1: Unrelated
- 0: Cannot decide

Table 1: Four-point Scale of Relatedness derived from Brown (2008, p. 250).

References

- Blank, A. (1997). *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen*. Tübingen: Niemeyer.
- Brown, S. W. (2008). Choosing Sense Distinctions for WSD: Psycholinguistic Evidence. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers* (pp. 249–252). Stroudsburg, PA, USA: Association for Computational Linguistics.
- Erk, K., McCarthy, D., & Gaylord, N. (2009). Investigations on word senses and word usages. In *Proceedings of ACL-09*.
- Erk, K., McCarthy, D., & Gaylord, N. (2013). Measuring word meaning in context. *Computational Linguistics*, 39(3), 511–554.
- Fritz, G. (2006). *Historische Semantik*. Stuttgart/Weimar: Metzler.
- Soares da Silva, A. (1992). Homonímia e polissemia: Análise sémica e teoria do campoléxico. In *Actas do xix congresso internacional de lingüística e filología románicas* (Vol. 2, pp. 257–287). La Coruña: Fundación Pedro Barrié de la Maza.

Acknowledgements

The research was supported by DFG Collaborative Research Centre SFB 732.

Semantic Change

- semantic innovation creates **polysemy** (cf. Fritz, 2006, p. 57)
- following Blank (1997, p. 113) we distinguish two types of semantic change:
 - i) **innovative meaning change:** the emergence of polysemy

German **packen**, ‘to grab > to grab; to hit figuratively’

earlier: *Ich wollte reden, er packte mich aber mit starker Faust und warf mich zur Thüre hinaus...*
‘I wanted to talk, but he grabbed me with a strong fist and kicked me out of the door...’

later: *Da erst packte mich die Verzweiflung und mit der Verzweiflung zugleich unnennbare Wuth.*
‘Only then I was hit by desperation and with the desperation also by undescribable rage.’

later: *Der alte Schäferle packte ihn am Arme und rief:*
‘The old Schäferle grabbed his arm and shouted:’

- ii) **reductive meaning change:** loss of a fully-established meaning

German **schinden**, ‘to torture; to skin > to torture’

earlier: *Jhr schindet jhnen die Haut ab vnnnd das Fleisch von Jhren Beinen...*
‘You skin them and remove the flesh from their legs...’

earlier: *Dann es funden sich mehr als fünffhundert Zeugen / welche hernacher öffentlichen wider jhn zeugeten und gnugsam bewiesen / wie er das arme Volck hatte bedranget und geschunden.*
‘Then there were more than five hundred witnesses / who later publicly gave testimony against him and sufficiently proved / how he had pressured and tortured the poor folk.’

later: *Wer sich stundenlang schindet und schwitzt, etwa bei einem Marathon, verliert Mineralstoffe im Übermaß.*
‘Who tortures himself for hours and sweats, as e.g. in a marathon, loses minerals in abundance.’

Semantic Proximity

- Blank (cf. 1997, p. 406–419) distinguishes four degrees of **semantic proximity** for pairs of word uses relying on **prototype theory**:

- 4: **identity:** sem. identical, both uses have same meaning
- 3: **context variance:** sem. very near, referents belong to same category
- 2: **polysemy:** sem. distant, ref. from diff. cat. but have sem. relation (similarity, contiguity...)
- 1: **homonymy:** sem. very distant, ref. from diff. cat. and no sem. rel.

- **basic idea:** we measure the mean semantic proximity of uses of a word over time
 - **increase** suggests innovative meaning change (polysemization)
 - **decrease** suggests reductive meaning change

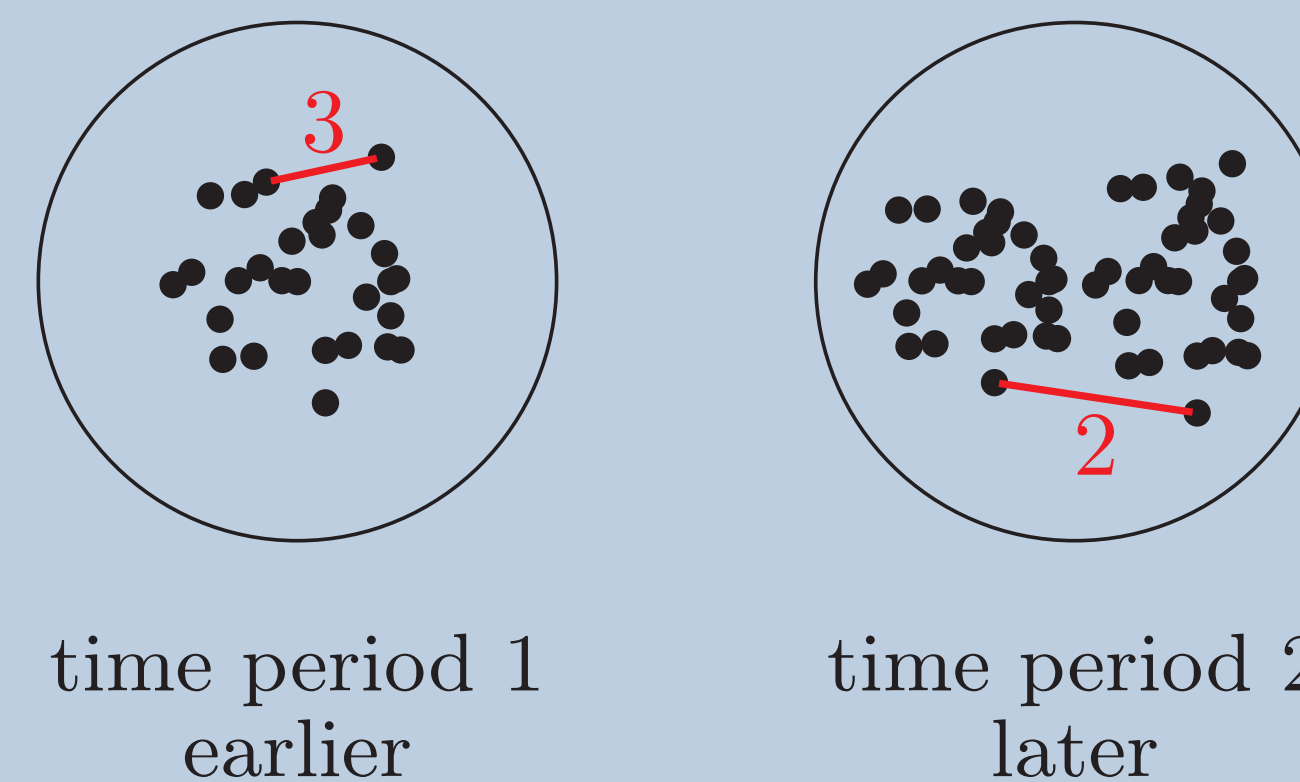


Figure 1: 2-dimensional use spaces in two time periods of a target word w undergoing polysemization. Dots represent uses of w and lines represent the **semantic proximity** of two such uses (measured by the number written next to it).

Advantages

- use pairs:
 - * **propagation:** captures degrees of prevalence
- scale:
 - * **graduality:** captures degrees of change
- relatedness:
 - * **intersubjectivity:** previous studies high annotators’ agreement on semantic relatedness
 - * **feasibility:** non-experts can annotate

Problems

- sample size?
- hardly distinguishable semantic constellations:
 - * homonymy vs. polysemy with 3 meanings
 - * polysemy vs. semantic generality
- use variance here?
- tendency of modern interpretation
- intelligibility of historic language
- ...

Annotation Examples

	A	B	C	D
1	target sentence 1 Ihr treibt das Geben noch weiter als ich das Naschen. Schon der vorige Kuchen, den ich allein esse, wiegt 10 Geburtstagen auf. Ich nehme diesen von deiner Liebe gar an; aber der Ultimakuchen aus Hof — dieser ist Penultima- und der erste Antepenultima-Kuchen — kann, soll etwas davon für mich abgeschnitten werden, auf mein h. Wort, nur für gleiche Bezahlung angenommen werden.	judgment	comment	target sentence 2 Der Benediktiner hatte eine Schlüssel der schönsten und größten Pfirsichen zum Geschenk erhalten, die er zum Nachtschlaf auftragen ließ, und dem Prinzen die köstlichste mit einer lächelnden und frommen Miene hinreichte. Der Prinz theilte sie mit seiner Geliebten, und sie aßen beyde die Pfirsiche ohne Verdacht. Man stund auf.
Rating 4 (Identical).				
1	target sentence 1 Hier schlug er seine Bude auf, und rief den Pöbel zusammen, seine schöne Raritäten zu schauen. Das Volk drang hinzu, Mägde und Knechte, Jungfrauen und Wittwen, Kinder und Greise. Der Teufel gaukelte ihnen allerley Histörchen vor, die er mit frommen Erläuterungen und moralischen Sprüchen begleitete.	judgment	comment	target sentence 2 Zum Namenstag. Die sechs nächsten Wesen sind ihr Mann und die fünf Kinder. 260.
Rating 3 (Closely Related).				
1	target sentence 1 Ferner: solange der einzelne Geschäftsmann alle Zahlungen unmittelbar aus seiner Kasse leistet, bezw. in dieselbe einnimmt, muss er zu den Zeiten, wo regelmäßig grössere Summen fällig werden, einen erheblichen Barbestand beschaffen und andererseits in den Zeiten überwiegender Eingänge dieselben sogleich zweckmässig unterzubringen wissen. Die Konzentrierung des Geldverkehrs in den grossen Banken enthebt ihn dieses periodischen Zwanges zur Aufhäufung und Drainierung; Erhöht ein letztes Beispiel.	judgment	comment	target sentence 2 Ich fand im 43 ^{ten} 44 ^{ten} lang, wahrscheinlich 51 ^{ten} 52 ^{ten} . Auf der Bank haben wir keinen Seetang gesehen, auch nicht westlich seit dem 8ten Juli. Wenige Tage, ehe wir die Bank von Neufundland berührten, waren, was in dieser Jahreszeit (Mitte Juli's) sehr ungewöhnlich ist, grosse Eismassen gesehen worden, die sich nach Südwest bewegten, während der gewöhnliche Strom auf dem südlichen Theil der Bank nach Norden gerichtet ist.
Rating 2 (Distantly Related).				
Rating 1 (Unrelated).				