



**University of Stuttgart**  
Germany

# An Annotation of Semantic Change based on Usage Relatedness

October 28, 2017

Dominik Schlechtweg, Sabine Schulte im Walde  
Institute for Natural Language Processing, University of Stuttgart

# Introduction

- ▶ our aim:
  - ▶ build a computational model detecting semantic change in corpora
- ▶ the problem:
  - ▶ there is no reliable test set of semantic change for any language
- ▶ the solution:
  - ▶ collect a set of words displaying different types of semantic change in a specific corpus

## A Definition

- ▶ “Semantische Neuerung **erzeugt Polysemie.**” (Fritz, 2006, p. 57, our emphasis)
- ▶ “Wir unterscheiden also zwei Arten von Bedeutungswandel:
  - **innovativer Bedeutungswandel:** Dies ist die Entstehung einer neuen Bedeutung mit vollausgeprägten Bedeutungsebenen bei einem Wort neben der alten. [...] Ausgangsbedeutung und neue Bedeutung stehen [...] in Polysemie zueinander. Die Polysemie der beiden Bedeutungen ist Ergebnis [...] des Bedeutungswandels.
  - **Reduktiver Bedeutungswandel:** Dies ist der Wegfall einer voll-ausgeprägten Bedeutung.” (Blank, 1997, p. 113, our emphasis)

# Polysemy

- ▶ “Ein problematischer Begriff” (Blank, 1997, p. 406)
- ▶ hard to distinguish between:
  - (i) context variance
  - (ii) polysemy
  - (iii) homonymy

## Polysemy vs. Context Variance

- ▶ extensional: “Kontextvarianz einer Bedeutung [...] [liegt] dann vor, wenn es sich bei den konkreten Referenten der Verwendung eines Wortes zwar eindeutig um nicht-prototypische Vertreter handelt [...], wenn man diese Referenten aber dennoch ein und derselben Kategorie [...] zuordnen würde [...]. In die Semantik übertragen heißt dies: Die spezifische Verwendung eines Wortes bewegt sich semantisch im Rahmen einer bestimmten Bedeutung” (Blank, 1997, p. 414)
- ▶ semantisch: “Ist die semantische Nähe zwischen den beiden Verwendungen so eng, daß keine der üblichen semantischen Relationen zwischen ihnen hergestellt werden kann, daß man also die eine Verwendung weder als Metapher, noch als Metonymie, noch als Oberbegriff etc. der anderen verstehen kann, dann liegen wahrscheinlich Kontextvarianten vor.” (Blank, 1997, p. 416)

# Polysemy vs. Context Variance

- ▶ Example of Context Variance (Blank, 1997, p. 412):

*My arm hurts.*

vs.

*Look at the arm of the statue.*

- ▶ Example of Polysemy (Blank, 1997, p. 413):

*My arm hurts.*

vs.

*An arm of the sea.*

## Polysemy vs. Homonymy

- ▶ “Homonymie liegt dann vor, wenn zwei semantisch verschiedene Verwendungen eines Signifiant in keiner semantischen Relation (Metapher, Metonymie, Ober-/Unterbegriffsrelation etc.) stehen, wenn die Relation der Bedeutungen nicht «konzeptuell motiviert» [...] ist.” (Blank, 1997, p. 417)

# Polysemy vs. Homonymy

- ▶ Example of Homonymy (Blank, 1997, p. 417):

*My arm hurts.*

vs.

*The number of men under arms is no longer the decisive factor in warfare.*

- ▶ Example of Polysemy (Blank, 1997, p. 413):

*My arm hurts.*

vs.

*An arm of the sea.*



## Semantic Proximity

- ▶ “[Es] wurden Kriterien entwickelt, mit denen der Status und die semantische Zuordnung von individuellen Verwendungen oder Typen von Verwendungen eines Wortes besser erfaßt werden können. Dies geschah auf dem Hintergrund einer klareren Definition von «**semantischer Nähe**»” (Blank, 1997, p. 418, our emphasis)
- ▶ **semantic proximity** as defined by Blank relies on prototype theory
- ▶ from this he deduces the main criterion for identification of polysemy (and hence of semantic change):
  - semantic proximity expressed by the existence of a semantic relation between two uses of a word**
- ▶ (i) context variance = semantically very near, no semantic relation (or synonymy)
- ▶ (ii) polysemy = semantically more distant, semantic relation
- ▶ (iii) homonymy = semantically very distant, no semantic relation

# Semantic Proximity in Practice

- ▶ Blank mentions a study of Soares da Silva (1992) where sentence pairs were rated for their semantic proximity ('0' = no semantic relation, '4' = high semantic relatedness)
  - in line with Blank we can see semantic relatedness as a continuum with homonymy on one end, meaning identity on the other and polysemy in the middle
- ▶ similar studies in computational linguistics for word sense disambiguation:
  - ▶ Brown (2008) for determining the degree of relatedness between a word's meaning in different contexts
  - ▶ Erk, McCarthy, and Gaylord (2009, 2013) for determining the degree of similarity between a word's meaning in different contexts

## Related Concepts – Semantic Variation

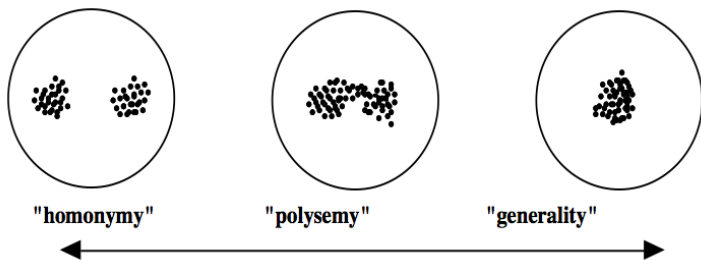


Figure 1 : Zlatev (2003, p. 482, Figure 12)'s continuum from generality to homonymy. Each dot is a particular contextual interpretation of a given expression in a multidimensional use-space (inadequately presented as a 2-dimensional circle). Generality is similar to context variance.

# Semantic Proximity as a Proxy for Semantic Change

- ▶ we know: semantic proximity distinguishes semantic change from non-change (polysemy from contextual variance and identity)
- ▶ we also know: there are already successful synchronic studies annotating similar notions (similarity/relatedness)
  - we use semantic proximity as a proxy for semantic change
- ▶ basic idea: we measure the semantic proximity of uses of a word over time
  - ▶ increase means innovative meaning change (polysemization)
  - ▶ decrease means reductive meaning change

## Example – Polysemization

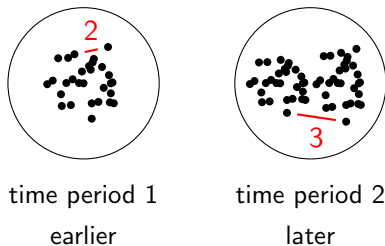


Figure 2 : 2-dimensional use spaces in two time periods of a target word  $w$  undergoing polysemization. Dots represent uses of  $w$  and lines represent the **semantic proximity** of two such uses (measured by the number written next to it).

# Advantages?

- ▶ graduality
- ▶ prevalence
- ▶ fuzziness of senses
- ▶ fuzziness of definitions

# Problems?

- ▶ sample size
- ▶ polysemy of source
- ▶ ...

# Annotation Guidelines (Draft) – 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).



# Annotation Examples – Identity/Synonymy

	A	B	C	D
1	<b>target sentence 1</b>	<b>judgment</b>	<b>comment</b>	<b>target sentence 2</b>
2	Ihr treibt das Geben noch weiter als ich das Naschen. <b>Schon der vorige Kuchen, den ich allein esse, wiegt 10 Geburtstagsorten auf.</b> Ich nehme diesen von deiner Liebe gar an; aber der Ultima-Kuchen 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.	4		Der Benediktiner hatte eine Schüssel der schönsten und größten Pfirsichen zum Geschenk erhalten, die er zum Nachtsch auftragen ließ, und dem Prinzen die köstlichste mit einer lächelnden und frommen Miene hinreichte. <b>Der Prinz theilte sie mit seiner Geliebten, und sie aßen beyde die Pfirsiche ohne Verdacht.</b> Man stund auf.

Table 2 : rating 4 (Identical).

# Annotation Examples – Contextual Variance

	A	B	C	D
1	<b>target sentence 1</b>	judgment	comment	<b>target sentence 2</b>
3	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, <b>Kinder</b> und Greise. Der Teufel gaukelte ihnen allerley Histörchen vor, die er mit frommen Erläuterungen und moralischen Sprüchen begleitete.	3		Zum Namenstag. Die sechs nächsten Wesen sind ihr Mann und die fünf <b>Kinder</b> . 260.

Table 3 : rating 3 (Closely Related).

# Annotation Examples – Polysemy

	A	B	C	D
1	<b>target sentence 1</b>	<b>judgment</b>	<b>comment</b>	<b>target sentence 2</b>
4	Die Fraw aber / gleichsamb ob sie nichts von der Fabel wiste / fragte fleissig nach allem / wer das Kind were / warumb man es verstossen / vnd den jungen Knaben / von solcher Schönheit daß die Natur nichts an jhm vergessen / also weggeworffen hette. Er gab für / jhm were weiter nichts wissendt / als dass er jhn auff einem <b>Scheidwege</b> am Wald / da niemandt als Hirten vnd Jäger hin kämen / gefunden; dahin er entweder vnbarhertziger weise / oder auß höchstdringender Noth müßte gelegt worden seyn. Also nam Sicambre das weinende Kindt / vnd schweigte es / in dem sie jhm zu trincken gab.	2		Auf der Erde werden Kinder Gottes geboren, die von den Menschen Genies genannt werden. Ein jedes von ihnen wird in früher Jugend an einen <b>Scheidweg</b> gestellt und muß seine Wahl treffen. Seine Gabe ist ihm verliehen für andere sowohl wie für sich selbst.

Table 4 : rating 2 (Distantly Related).

# Annotation Examples – Homonymy

	A	B	C	D
1	<b>target sentence 1</b>	<b>judgment</b>	<b>comment</b>	<b>target sentence 2</b>
5	Ferner: solange der einzelne Geschäftsmann alle Zahlungen unmittelbar aus seiner Kasse leistet, bzw. 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. <b>Die Konzentrierung des Geldverkehrs in den grossen Banken</b> enthebt ihn dieses periodischen Zwanges zur <b>Aufhäufung und Drainierung</b> ; Endlich ein letztes Beispiel.	1		Ich fand lat. 43 ° 14 ", long. wahrscheinlich 51°¼. <b>Auf der Bank</b> haben wir <b>keinen Seetang gesehen, auch nicht westlich seit dem 8ten Juli</b> . 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.

Table 5 : rating 1 (Unrelated).

# Bibliography

- 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 filoloxía románicas* (Vol. 2, pp. 257–287). La Coruña: Fundación Pedro Barrié de la Maza.
- Zlatev, J. (2003). Polysemy or generality? Mu. In H. Cuyckens, R. Dirven, & J. R. Taylor (Eds.), (Vol. 23, p. 447 - 494). Mouton de Gruyter.