

Mehrworteinheiten

Zusammenfassung & Seminarkritik

Timm Lichte

HHU Düsseldorf

Letzte Sitzung, 30.01.2018



Mehrworteinheiten: Definition

■ Definition

“Pain the neck for NLP” (Sag et al. 2002)

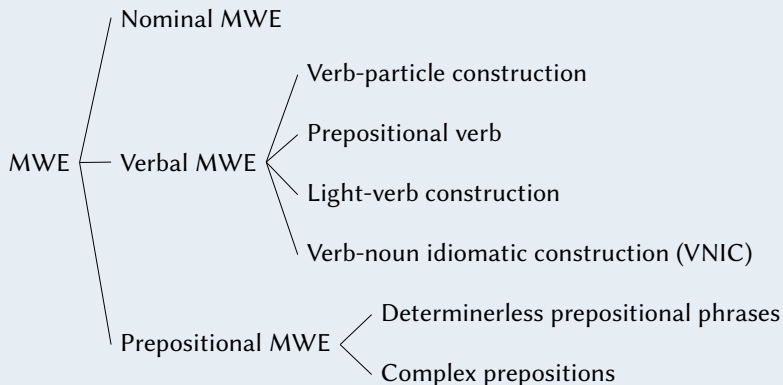
Complex expressions which have surprising properties not predictable from their component words. (Baldwin & Kim 2010)

Multiword expressions (MWEs) are lexical items that: (a) can be decomposed into multiple lexemes; and (b) display lexical, syntactic, semantic, pragmatic and/or statistical idiomaticity. (Baldwin & Kim 2010)

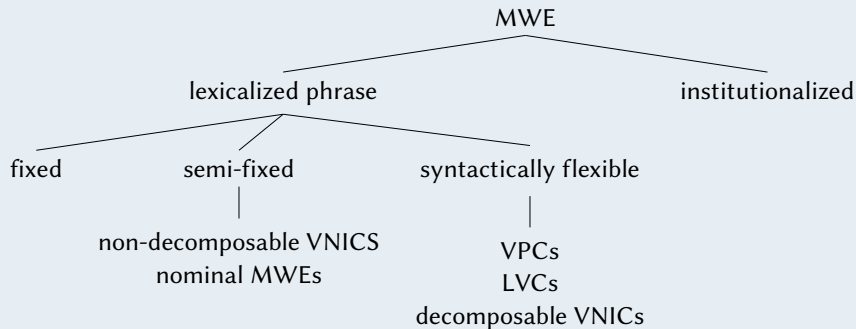
Multiword expressions (MWEs) are units of more than one lexem stored in the mental lexicon (\leadsto “familiarity”).

Mehrworteinheiten: Typen

■ Typen (Baldwin & Kim 2010)



- Klassen (Sag et al. 2002, Baldwin & Kim 2010)



- Präzisionsgrammatiken
- Lexikographie
- Identifikation und Extraktion
- Parsing
- Statistisches Maschinelles Übersetzen
- Psycholinguistik

Wissen:

- Kennenlernen der verschiedenen (computer-)linguistische Aspekte von Mehrworteinheiten

Fähigkeit:

- wissenschaftliche Texte lesen
- über wissenschaftliche Texte referieren
- über wissenschaftliche Texte diskutieren
- über wissenschaftliche Texte schreiben

X tun lernt man nur, indem man X tut.

Kritikpunkt: Textauswahl

- 1 Diskussionsbeiträge (Anwesenheit + vorbereitende Lektüre)
- 2 Referat

Was bleibt hängen? Viel mehr, wenn man den Text vorher gelesen hat!

Zweck eines Referats?

- den Text filtern (wichtig/unwichtig)
- den Text gliedern (Hypothesen, Ansatz, Ergebnisse, Diskussion, Ausblick)
- den Text vervollständigen (Beispiele, Definitionen)
- den Text diskutieren (Fragen, Kritik)

Kritikpunkt: zu lange/kurze Referate pro Sitzung?

Wie sehen gute Folien aus?

Form (Schriftarten, Farben, Schaubilder, Gliederung, Tafelinsatz)
&
inhaltliche Darstellung (Text-Bild-Verhältnis, Gedankenführung)

sind **nicht** im Weg!

⇒ obligatorische Powerpoint- und \LaTeX -Vorlagen

Wie umgehen mit Texten und Fachterminologie auf Englisch?

Folien auf komplett auf Englisch (außer klar markierte Kommentare)!

Wie hält man ein gutes Referat?

- 1 Respect your audience! (Kenntnisstand, Aufnahmefähigkeit, ...)
- 2 Know your slides!
- 3 Show rather than talk!
- 4 Say what your're doing!
- 5 Don't say sorry!
- 6 Be honest!
- 7 Try to have fun!

Was bleibt dann hängen?

Viel mehr, wenn man den Text vorher gelesen hat!

Inhalt:

- Fragestellung → Beantwortung durch Literatur → eigener Standpunkt
- Aufgaben
 - Eingrenzung, Abgrenzung des Themas
 - Darstellung von Frage und Antwort
 - Bewertung, eigener Standpunkt, eigene Ideen?
- imaginiertes Publikum: Kommilitonen, die den Kurs vielleicht nicht besucht haben (Dozent)

Form:

- Bestandteile: Titelseite, Inhaltsverzeichnis, Einleitung, Hauptteil, Schluss, Literaturverzeichnis
- Umfang: ca. 15 Seiten
- richtiges Zitieren
- Textprozessor: MS-Word, OO-Writer, \LaTeX
- Abgabe: PDF

Freischuss: spätestens drei Wochen vor Abgabe

- [1] Baldwin, Timothy & Su Nam Kim. 2010. Multiword expressions. In Nitin Indurkha & Fred J. Damera (eds.), *Handbook of natural language processing*, 2nd, 267–292. Boca Raton, FL: CRC Press.
- [2] Sag, Ivan A, Timothy Baldwin, Francis Bond, Ann Copestake & Dan Flickinger. 2002. Multiword expressions: A pain in the neck for NLP. In Alexander Gelbukh (ed.), *Computational linguistics and intelligent text processing* (Lecture Notes in Computer Science 2276), 1–15. Berlin: Springer.