

CoRA Company Relation Analysis

Das CoRA-Projekt: Ziele

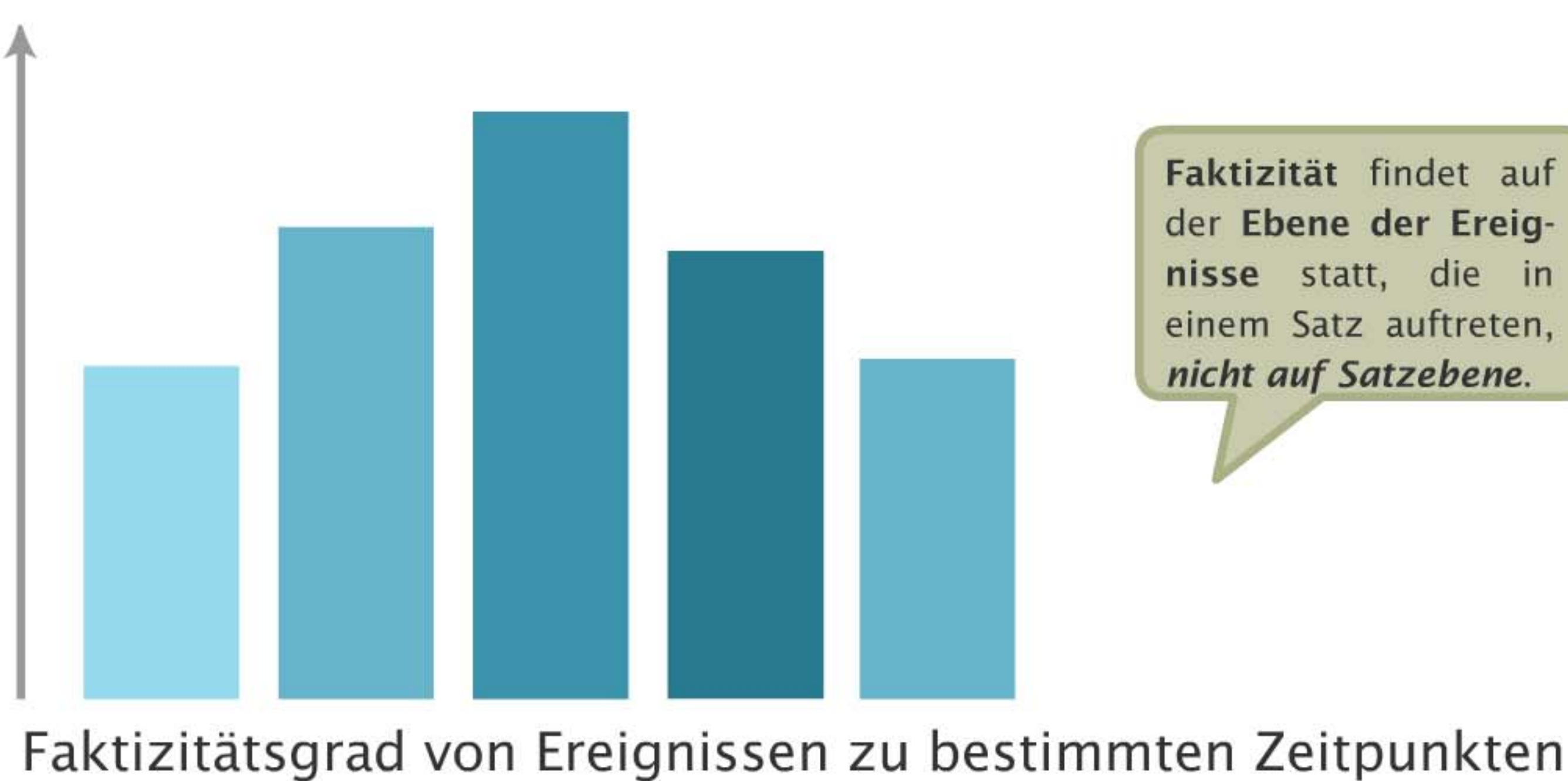
- **Linguistische Analyse** (Genauigkeit und Ganzheitlichkeit bei der Deutung) von **Firmenrelationen in englischsprachigen Finanz- und Wirtschaftsnachrichten** (Financial Times, Wall Street Journal, etc.)
- Erstellung von sprachspezifischen Modulen im Bereich der **Informationsextraktion** zur Erkennung von **Personalentscheidungen, Firmenfusionen** und weiterer interessanter **Relationen zwischen Unternehmen** und ihren **Konkurrenten**, oder ihren **Angestellten**, oder ihren **Kunden**.

Anwendungen

Die **Ergebnisse** dieses Forschungsvorhabens sind für diverse **Anwendungen im Informationsmanagement** von besonderer Bedeutung: Hierzu gehören z.B.

- **Suchmaschinen** (Finden von Unternehmensdaten in der aktuellen Presse),
- **Frage-Antwort-Systeme** (Analyse der zeitl. Abfolge und Stadien von Ereignissen)
- **Klassifikationssysteme** (Kategorisierung von Firmen nach Beurteilungen),
- **Datenbankerstellung** (Datenbanken mit Firmendossiers)

Die Faktizität eines Ereignisses



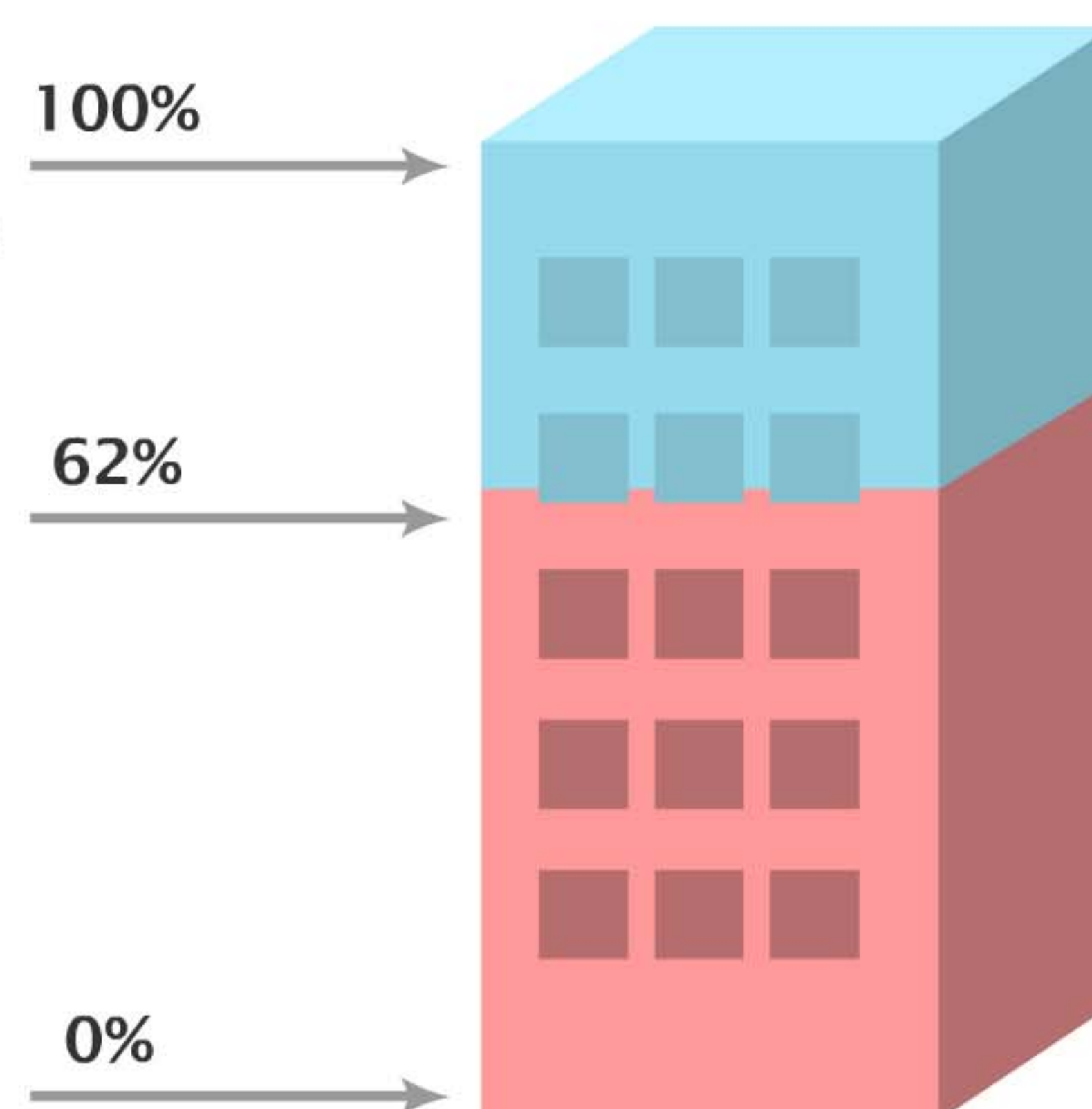
Zwischen Sicherheitsgrad und Polarität

- **Mithilfe sprachlicher Mittel** drückt man den **Sicherheitsgrad über den faktischen Status des Ereignisses** aus.
- Der Sicherheitsgrad betrifft die Art, wie das **Ereignis aus der Sicht des Autors** dargestellt wird:
Ist es sicher, möglich oder unmöglich?
- **Interaktion** zwischen dem **Grad der Sicherheit** (möglich, wahrscheinlich, etc.) und der **Polarität (positiv vs. negativ)**.

Firmenfusionen im Fokus der Linguistik

Faktizitätsabstufungen

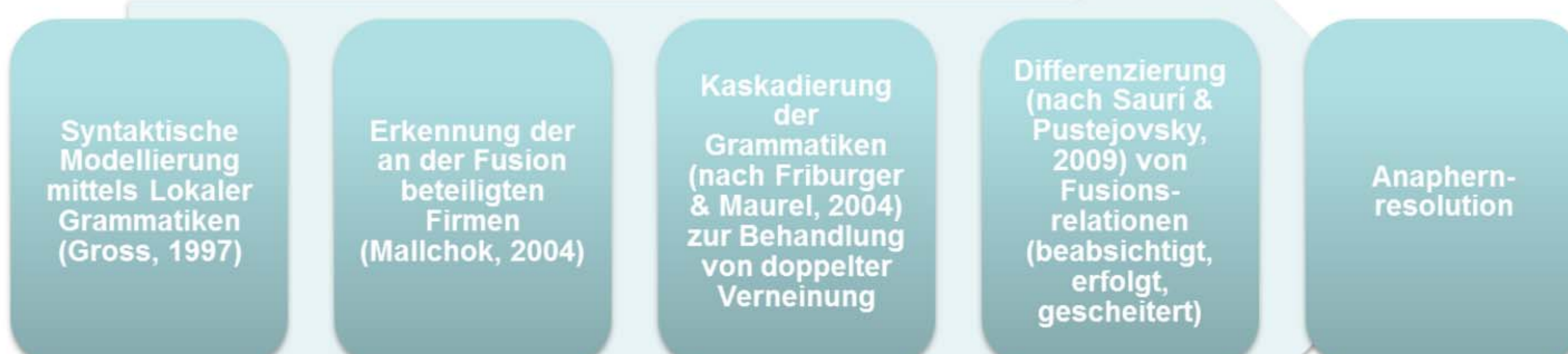
- erfolgte Fusion
- Beginn einer Fusion
- (fast) sicher stattgefundenene Fusion
- Ankündigung einer Fusion
- Wunsch einer Fusion,
- Planung einer Fusion
- Abstimmung einer Fusion
- Einigung über eine Fusion
- Verhandlungen um eine Fusion
- Vorschlag einer Fusion
- gescheiterte Fusion



Sprachliche Mittel für Faktizität

- **Polaritätspartikel** (Aussage in ihr Gegenteil verkehren)
*It is not the case that [Yahoo **never** merged with Google].*
 - **Epistemische Modalitätspartikel**
 - **Hilfsverben** wie „**may**“ (PS+) oder „**must**“ (Uu)
 - **Adverbien** wie „**certainly**“ (CT+) oder „**apparently**“ (PR+)
 - **Adjektive** wie „**unsure**“ (PS-).
 - **„Event-selecting predicates“**
Yahoo **denied** [that it **may merge**] with Google. (CT-)
It is **not** [**impossible** (CT-)] to **merge**. (PS+)
- (PS± = possibly positive/negative; Uu = fully underspecified; CT+ = certainly positive; PR+ = probably positive;)

Linguistische Modellierung



Referenzen

- Friburger, Nathalie; Maurel, Denis: *Finite-state transducer cascades to extract named entities in texts*. *Theor. Comput. Sci.* 313(1): 93-104 (2004)
- Gross, Maurice: *The construction of local grammars*. In: Roche; Schabes (Hrsg.): *Finite-state language processing* 329-354. 1997
- Mallchok, Friederike: *Automatic Recognition of Organization Names in English Business News*, LMU München, Doktorarbeit, 2004
- Sauri Roser; Pustejovsky, James: *FactBank: A Corpus Annotated with Event Factuality*. In: *Language Resources and Evaluation* 43 (2009), S. 227-268
- Sauri, Roser; Pustejovsky, James: *Are You Sure That This Happened? Assessing the Factuality Degree of Events in Text*.