

1. Fragen

- (a) Nennen Sie drei Aspekte, in denen sich Computerlinguistik und Sprachprodukttechnologie unterscheiden.
- (b) Nennen Sie die drei Einsatzgebiete von Sprachprodukten.
- (c) Nennen Sie drei Sprachprodukte und erläutern Sie deren Funktionalität.
- (d) Was versteht man unter einem Wort im Unterschied zu einer Wortform?

2. Fachsprachenkonzept

Gegeben sei der folgende Text:

Z1 Mit Tief „Fritz“ über der Biskaya gelangt feuchte, aber allmählich mildere Luft nach
Z2 Mitteleuropa. In Spanien und Frankreich bringt das Tief kräftige Regengüsse. Über
Z3 dem Baltikum und in weiten Teilen von Skandinavien sorgt Hoch „Birgit“ für
Z4 trockenes, aber kaltes Winterwetter. Im östlichen Mittelmeerraum liegt ein weiteres
Z5 Tiefdruckgebiet mit Regen.

- (a) **Wie viele *types*** enthält der Text und **wie viele *tokens***?
- (b) Betrachten Sie den Unterschied zwischen *types* und *tokens*. Was fällt Ihnen auf? Welche Folgerungen können wir daraus ziehen?
- (c) **Analysieren** Sie den Text und
 - markieren Sie **fünf** für die Anwendungsdomäne wichtige Begriffe (keine Eigennamen)
 - nennen Sie wenigstens **zwei** für den Text auffällige syntaktische Konstruktionen
- (d) **Erläutern** Sie den Begriff der **Fachsprache** und nennen Sie die drei wesentlichen **Abgrenzungskriterien** einer Fachsprache gegenüber der Allgemeinsprache.

3. Linguistische Ebenen

- (a) Nennen Sie ausgehend von der untersten Ebene die sog. linguistischen Ebenen und verdeutlichen Sie die linguistischen Ebenen an folgendem Beispiel:

Im östlichen Mittelmeerraum liegt ein weiteres Tiefdruckgebiet mit Regen.

- (b) Was ist der Unterschied zwischen einem Phonem und einem Morphem?
- (c) Erläutern Sie die Prinzipien *Konkatenation* und *Abstraktion*. Wie lassen sich die linguistischen Ebenen mit Hilfe dieser Prinzipien rekonstruieren?
- (d) Präzisieren Sie am Beispiel der **syntaktischen Analyse** die Aufgaben, die von einem Programm für die automatische Verarbeitung natürlicher Sprache zu lösen sind.

4. Satzbaumuster und ComputerTalk

- (a) Bestimmen Sie die Satzbaumuster für die folgende Frage:

Wieviel kostet eine Fahrkarte von Regensburg nach Kiel an einem Sonntag ?

- (b) Was besagt die These des ComputerTalk?
(c) Welche praktischen Konsequenzen können wir daraus für die Entwicklung von Sprachprodukten ziehen?

5. Trigramme und Editierdistanz

- (a) Bestimmen Sie die Trigramme der nachfolgenden Wortformen

wirk, work, workaholic, wirkaholic

- (b) Bestimmen Sie mit dem Levensthein-Algorithmus die Editierdistanz zwischen *wirk* und *work*

- (c) Bestimmen Sie die mit dem Dice-Maß die Ähnlichkeit zwischen den Wortformen *wirk, work, workaholic* und *wirkaholic*. Wie wirkt sich die Wortlänge auf die Ähnlichkeit aus?