

## ÜBERLEGUNGEN ZU UMFANG UND AUFBAU DES WORTSCHATZES

### 1. Reduktion des lateinischen Wortschatzes

Entwicklung der Wortkunden seit den 60er Jahren:

- Leitschuh/Hofmann (1965): 3700 Wörter
- Klaus/Klingelhöfer (1966): 4000 Wörter
- Konrad/Raab (1979): zur Bewältigung der Lektüre müssen ca. 2500 Wörter genügen
- Neuere Wortkunden basieren auf Frequenzstatistiken und umfassen 2500-3000 Wörter.

Welche Kriterien und welche Lektüre sollen Grundlage für eine Reduktion sein?

C.Utz: 'Werkstattbericht' Bamberger Wortschatz:

Verlässliches **Fundamentum** + autorenspezifischer **Augmenta**.

### 2. Textcorpus und Datenbestand

Ziel der Wortschatzsammlung soll die Fähigkeit sein lat. Originallektüre lesen zu können. Der Lektürecorpus wird mit Hilfe einer Häufigkeitsstatistik untersucht.

Bisherige Wortkunden geben wenig Auskunft über die Auswahlkriterien.

Gesamtkorpus Bamberger Wortschatz:

- Gesamtumfang: 140482 Wortformen (incl. Namen) → 133326 Wortformen

Ermittlung des Ausgangskorpus:

- aktuelle Lehrpläne u. administrative Richtlinien von acht Bundesländern mit den meisten Lateinschülern
- einschlägige Lektüreausgaben

Datenbestand:

- Zuordnung der Wortformen zu lexikalisch nachschlagbaren Stichwörtern → Lemmatisierung
- Query-Corpus: kodierter Textcorpus → zu jeder Wortform: Satzkontext, zugehöriges Lemma und morphosyntaktischer Status
- Beispiel: l [ADEO] w [ADIBIT] a [CA] s [08:016] fq [56]

### 3. Untersuchungen zu Fundamentum und Augmenta

- Unterscheidung von Fundamentum und Augmenta → Ökonomie des Lernens und Behaltens
- Vergleich Bamberger Wortschatz und frühere Untersuchungen  
ca. 1000 Wörter decken 80% der Textmenge
- Diagramm (S.158) zeigt: 1248 Vokabeln erreichen 83% Textabdeckung, für die restlichen 17% wären 6000 Vokabeln zusätzlich notwendig.
- Ermittlung der Vokabeln: - sortieren nach Frequenz  
- Berücksichtigung ab Frequenzwert 16 → 1320 Lemmata