

Hinweise zur Nutzung der Listen transkribierter Einheiten¹

Die angegebenen Häufigkeiten können aus folgenden Gründen variieren bzw. real von den in der Liste angegebenen Anzahlen abweichen:

- In der gesprochenen Sprache sind Abbrüche, Pausen zwischen Wortteilen, Wiederaufnahme, Stottern, Versprechen usw. charakteristisch. Diese tauchen auch als solche in der Liste auf, da die Transkription nach bestimmten Konventionen (GAT 2) nah an der Aussprache erfolgt.

Beispiele:

Abbruch:

SV_DE_068: würde sich auch sehr sch äh stark **anschie** ((lacht)) ((lacht)) ((unverständlich)) ja

Pausen:

EV_PL_020: mit **eindeuti** (.) **ger** äh äh kausaler semantik

Wiederaufnahme:

PG_UK_020: von seiner **autobiogra** (0.3) **biografie**

PG_DE_055: noch weitere punkte die sie **empfe** als empfehlenswert (.) betrachten

Stottern:

SV_PL_006: das erste **argu arg agument** ist hm

Versprechen:

PG_DE_136: ham zu dieser zeit (.) **eingefa angefangen** °h

- In der gesprochenen Sprache können Wörter auf unterschiedliche Weise realisiert werden und wurden dementsprechend auch in verschiedener Form transkribiert. Somit kann sich das gleiche Wort mit unterschiedlichen Belegen an unterschiedlichen Stellen in der Liste transkribierter Einheiten finden. Dazu zählen von der Standardsprache abweichende Realisierungen wie regionalsprachliche Phänomene, Reduktionen, Klitisierungen und Tilgungen.

Beispiele:

abweichende Artikulation:

PG_PL_212: also um **achtzehnhundert** nicht wahr und dann

SV_UK_010: im **achsehhundert** °h vierundreißig

Regionalsprachliche Phänomene:

PG_UK_028: für sie da interessant **isch** (0.6) s_is jetzt ne offene frage

EV_DE_097: ja klar da war_s **jeschrieben**

Klitisierungen:

EV_UK_001: im perfekt **isses** (.) isses auch noch problematisch

Reduktionen/Assimilationen/Tilgungen:

EV_UK_005: aspekte verschiedener kulturen **miteinander** vermengen (0.5)

vs. EV_UK_006: also systeme **mitnander** vergleicht

PG_UK_010: seine familie und die **deutschsprachi gemeinschaft**

PG_DE_114: is klammernöffnendes **elemen** dann isses

EV_DE_097: aber das **sin**

¹ Es handelt es sich nicht um Wortformen im lexikographischen Sinne, sondern um die Aufstellung aller transkribierten Einheiten.

- Aufgrund bestimmter Transkriptionskonventionen werden Wortteile in der Liste fälschlicherweise als eigenständige Wörter erfasst.

Beispiele:

EV_DE_094: in der **lehraus** und **fortbildung**

(weder „lehrausbildung“ noch „lehrerfortbildung“ werden als eigenständige Wörter in der Liste abgebildet, „lehraus“ und „fortbildung“ hingegen schon)

PG_UK_020: von seiner **autobiogra** (0.3) **biografie**

(„autobiographie“ erscheint nicht als komplett artikuliertes Wort, auch bei Wiederaufnahme wird nicht das komplette Wort verbalisiert)

- Auch Vorschläge und Alternativen zu schwer- oder unverständlichen Wörtern, die auf der Interpretation der Transkribenten beruhen, sind Teil der Liste (hier ist also von einigen - wenn auch wenigen - fehlerhaften Belegen auszugehen).

Beispiele:

PG_UK_023: auf die ähm (0.6) **(auf)** die ähm

EV_DE_097: und zu **(deinem/dem)** ersten mit dem

EV_DE_104: sie mit verfassungs äh **(gericht/gerüst)** zu tun haben

- Bei der automatischen Erstellung der Liste werden Sonderzeichen als Wortgrenzen gedeutet und dementsprechend Wörter ggf. nicht in kompletter Form aufgeführt.

Beispiel:

PG_PL_110: **(gegenüber)stellen**

(„gegenüberstellen“ wird nicht als Wort gezählt, dagegen erscheinen ein Treffer für „gegenüber“ und einer für „stellen“)

Dies ist auch der Fall bei Klitisierungserscheinungen mit Unterstrich. Die Häufigkeiten dieser Vorkommen verteilen sich entsprechend auf die Einzelwörter.

Beispiele:

PG_UK_002: benutzt °h äh **gab_s** da irgendwelche probleme

EV_UK_008: zurück und dann **wird_s** °h **wird_s** gespielt

(Statt der Treffer „gab_s“ und „wird_s“ werden in der Liste jeweils die Einzelwörter „gab“ bzw. „wird“ und „s“ angezeigt)

- Es muss auch von gelegentlichen Transkriptionsfehlern ausgegangen werden (z.B. versehentliches Setzen von Leerzeichen und damit Trennung von Wörtern).
- Die Listen enthalten auch Einträge, die aus Kommentaren in der Transkription zu sprachbegleitenden Handlungen stammen, wie z.B. *schnalzt*, *schmatzt*, *lachend* usw.

LR_0200 [v]: was ich überhaupt nicht nachvollziehen <<**lachend**> kann> °h