



Hintergrund

Wortmaterial, das in der Diagnostik und Therapie von erworbenen Sprechstörungen verwendet wird, sollte unter Berücksichtigung unterschiedlicher (psycho-)linguistischer und phonetischer Einflussfaktoren sorgfältig ausgewählt und strukturiert werden.

Aktuell verfügbare Wortdatenbanken sind aus mehreren Gründen nicht zufriedenstellend:

- Die in einigen größeren Datenbanken gelisteten Wortfrequenzen basieren ausschließlich oder überwiegend auf geschriebener Sprache (z.B. CELEX, Baayen et al., 1995; *dlxDB*, Heister et al., 2011).
- Die Korpora, die über die Wortfrequenz hinaus weitere lexikalische Parameter bereitstellen (z.B. Erwerbalters, Vertrautheit mit dem Konzept, semantische Typikalität), haben häufig einen begrenzten Umfang [1].
- Einige Korpora sind zwar nach phonologischen Aspekten strukturiert (z.B. *ArtikuList*, Betke et al., 2010), enthalten jedoch keine verwertbaren Parameter über die artikulatorische Komplexität der Wörter.

Vorgehen

SUBTLEX-DE Wortfrequenzdatenbank
N = 190.000 Einträge (token)

Wortfrequenzen basieren auf gesprochener Sprache (Untertitel von Filmen und Fernsehserien) (Download unter: <http://crr.ugent.be/SUBTLEX-DE/>)

Automatisierte Selektion mithilfe von *Balloon* (Textverarbeitungstoolkit; [2]) nach:

- **Wortart:** unflektierte Inhaltswörter, d.h. Nomen (NN; keine Eigennamen), Adjektive (ADJ), Verben (VVINF)
- **Wortlänge:** 1 bis 3 Silben

Manuelle Korrektur

- **inhaltlich** problematische Begriffe wurden ausgeschlossen (z.B. Maledicta)
- **formal** wurden Nichtwörter, orthographisch fehlerhafte Wörter, Wörter ohne Eintrag im DUDEN ausgeschlossen

N=12.829

Phonologische Nachbarschaftsdichte & Minimalpaare (für 56% von SUBTLEX-np)



CLEARPOND-Korpus [4]

PTAN: Anzahl aller phonologischen Nachbarn des Wortes (durch Substitution, Elision oder Addition eines Phonems)

PTAF: mittlere Frequenz aller phonologischen Nachbarn des Wortes

PTAW: Wortliste aller phonologischen Nachbarn des Wortes

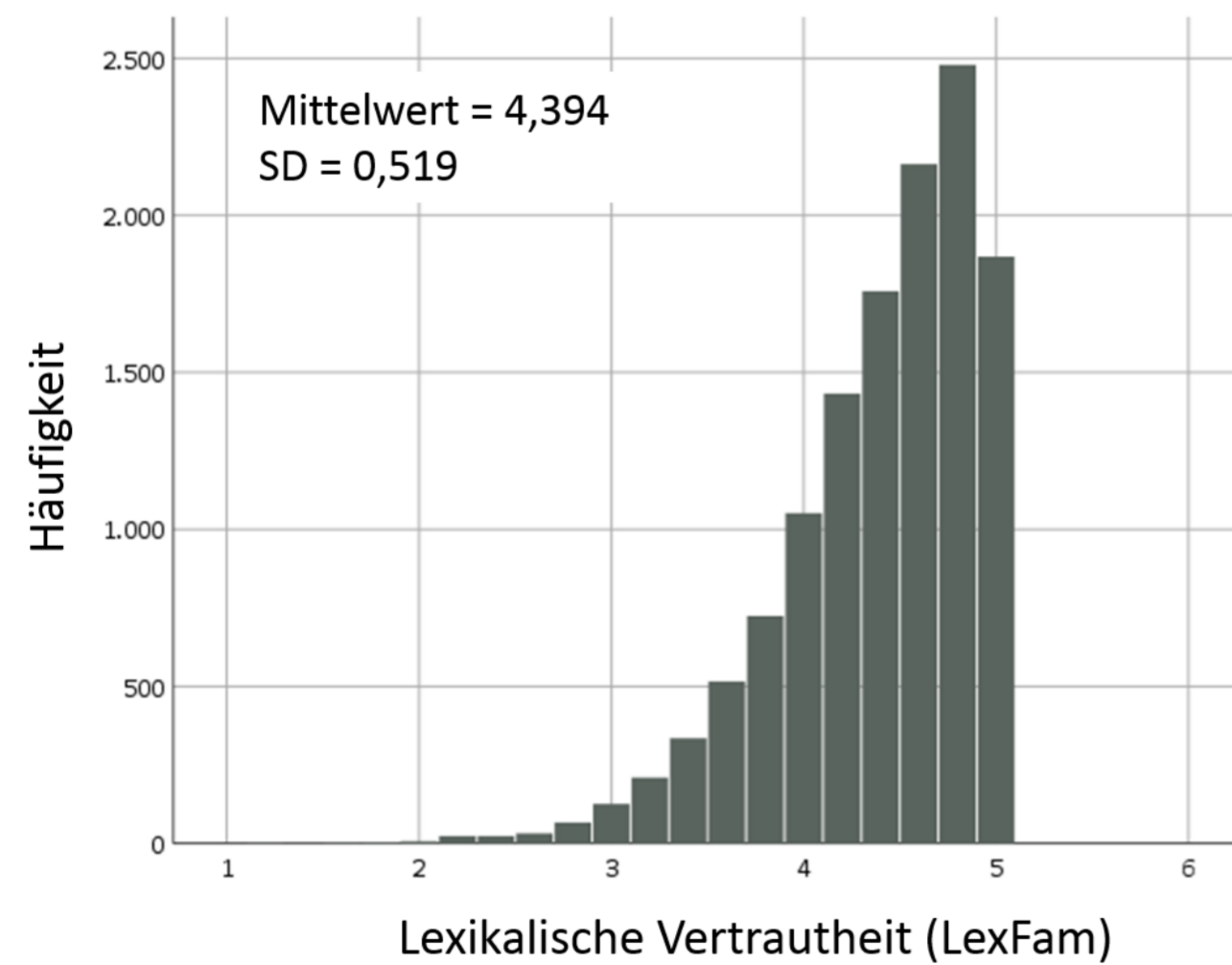
Lexikalische Vertrautheit (Lexical Familiarity)



Online-Experiment mit *Pavlovia* (<https://pavlovia.org/>) mit sprachgesunden Probanden (N=59; Alter, Geschlecht und Bildung kontrolliert)

Bitte schätzen Sie auf einer Skala von 1 bis 5 ein, wie vertraut Ihnen die folgenden Wörter sind. Sehr vertraute Wörter sind Wörter, die Sie oft lesen, hören und/oder selbst verwenden. Unvertraute Wörter sind Wörter, die Sie entweder nicht kennen oder die Sie sehr ungewöhnlich finden.

LexFam: Mittelwert über jeweils 5 Urteile



Sampa-Transkript



automatische Generierung über die Software *Balloon* [2]; manuell korrigiert

Sampa: Speech Assessment Methods Phonetic Alphabet (IPA vereinfacht und angepasst an Standardtastaturen bzw. maschinenlesbar)

Artikulatorische Komplexität:



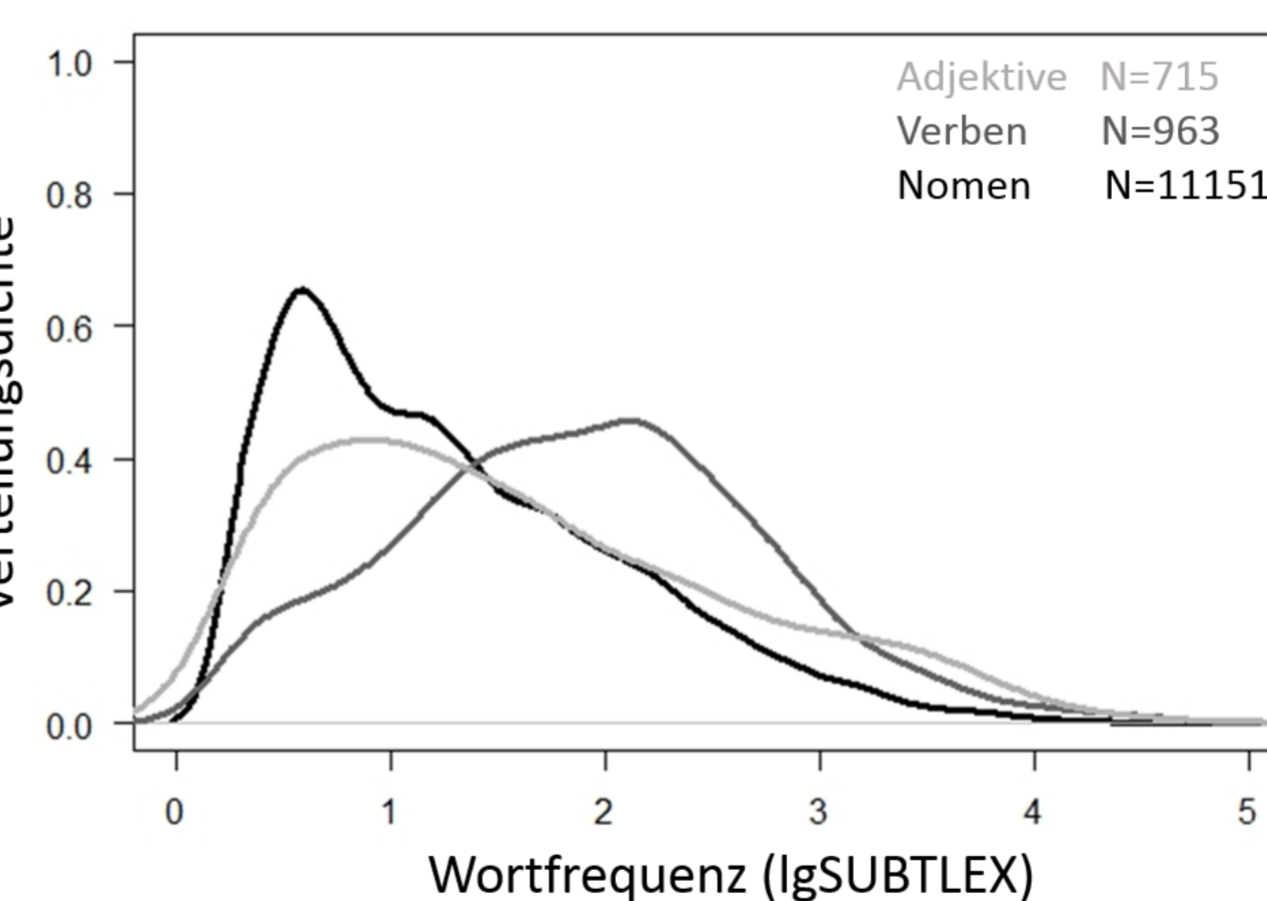
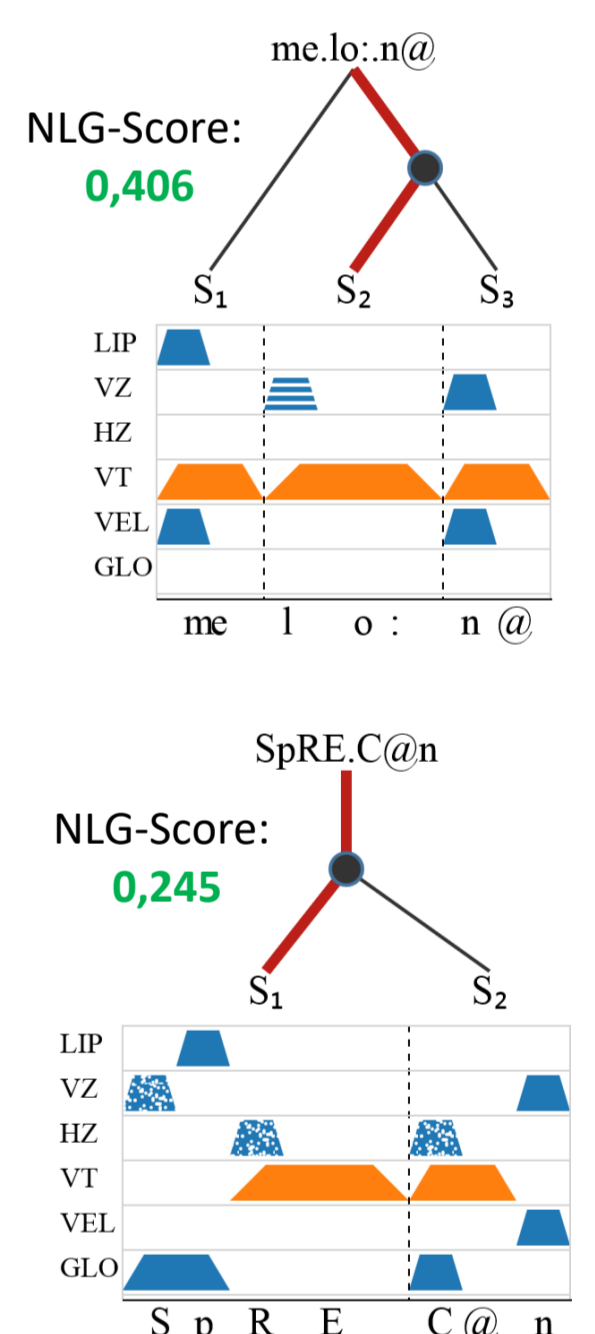
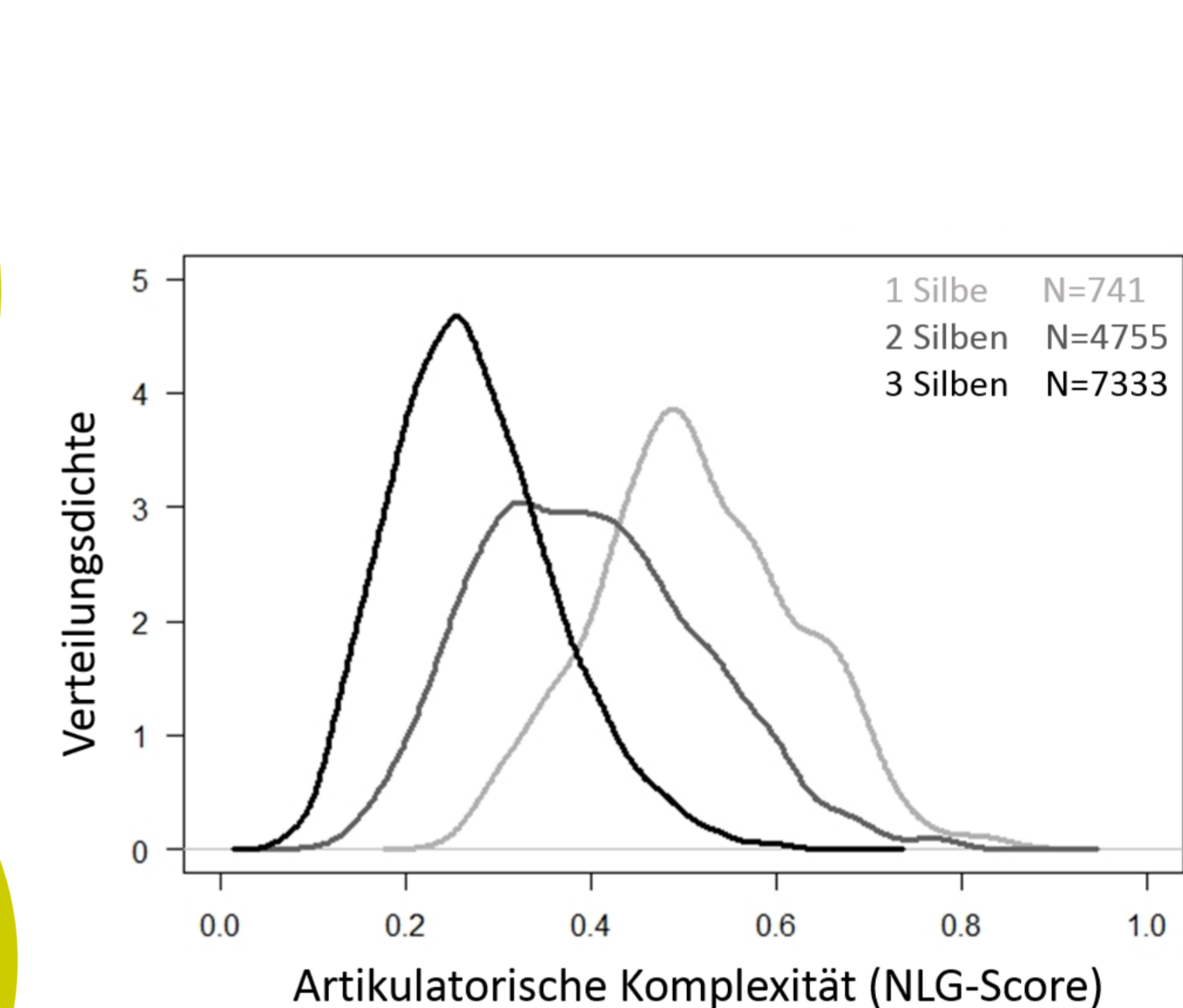
Gestenscore Rechner

Grundlage: Nicht-lineares Gestenmodell (NLG-Modell) [3]

(<https://neurophonetik.de/gesten-koeffizientenrechner>)

NLG-Score: Schätzwert der artikulatorischen Komplexität

- berechnet über eine nicht-lineare Funktion der gestischen, silbischen und metrischen Struktur
- basiert auf 120 bzw. 66 Realisierungen von 232 Wörtern durch Patienten mit Sprechapraxie



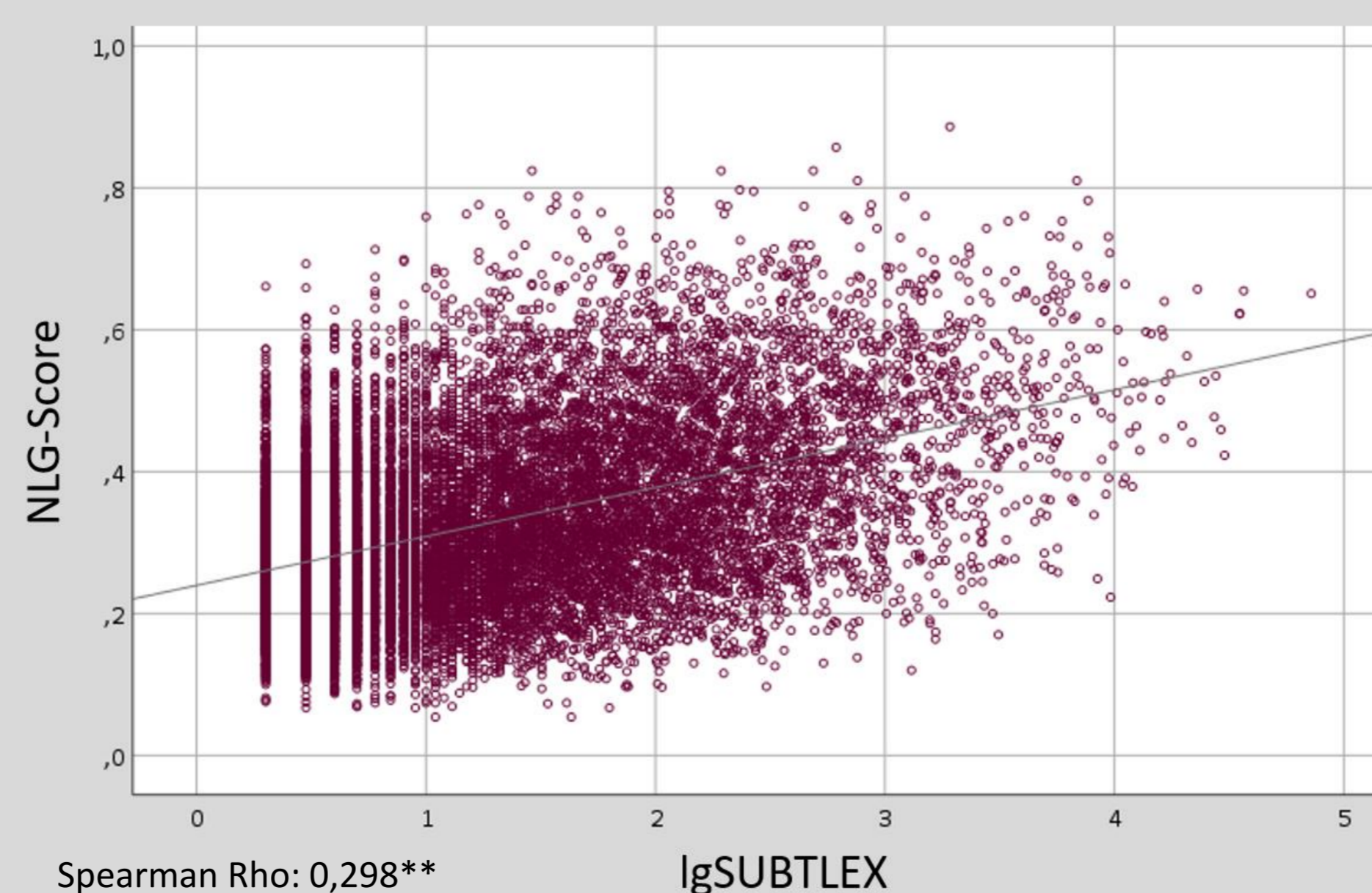
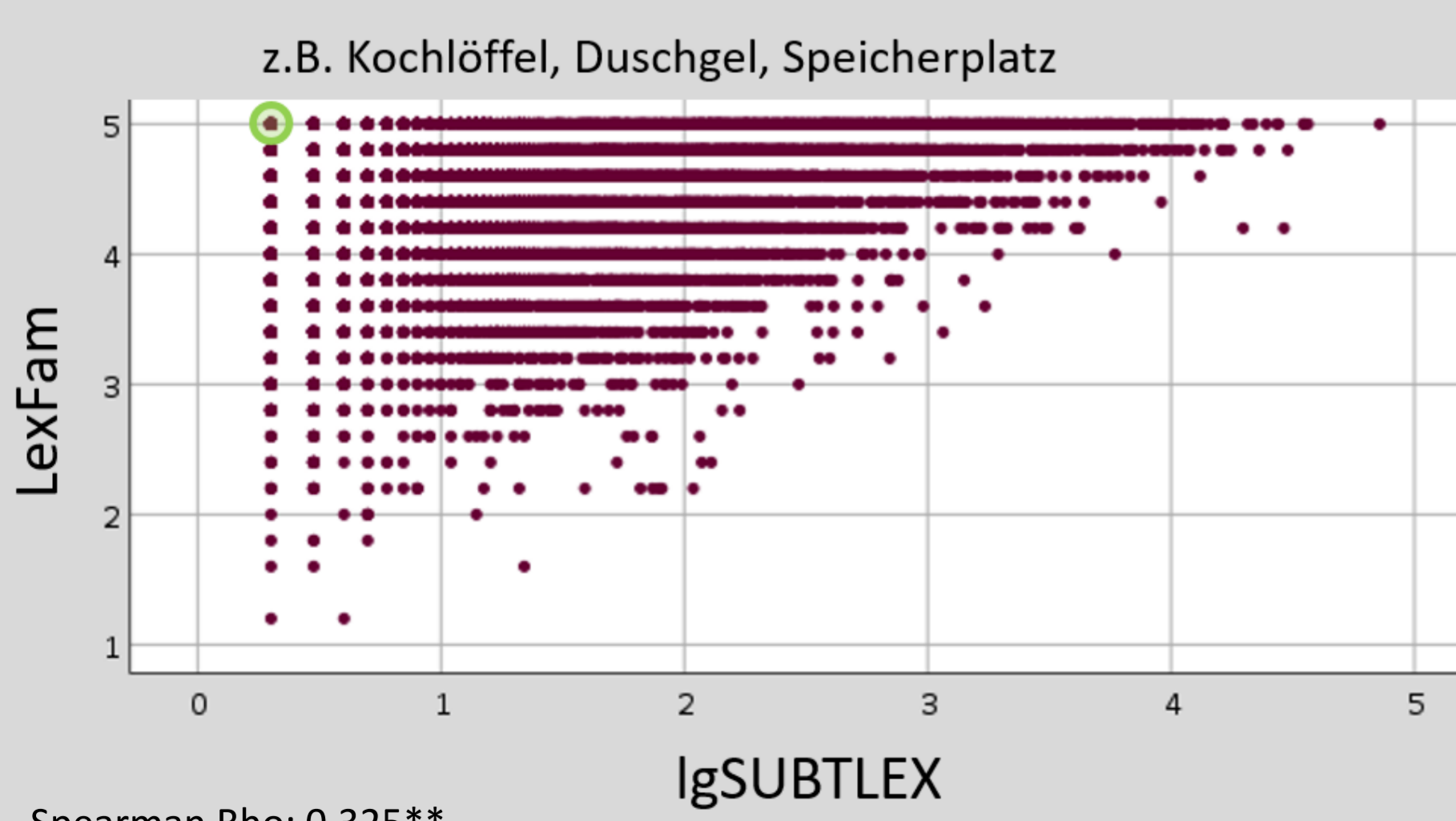
Wortfrequenz



SUBTLEX-DE (s.o.)

SUBTLEX-Werte erklären deutlich mehr Varianz in den Reaktionszeiten von lexikalischen Entscheidungsaufgaben als andere Frequenzmaße (z.B. CELEX) [5]

Zusammenhang einzelner Parameter



Anwendungsbeispiel 1: KommPaS

(KommunikationsParameter für Sprechstörungen) [6]

→ Systematische Auswahl von Wörtern zur tele-diagnostischen Verständlichkeitsmessung bei Dysarthrie
 → Paralleltests mit vergleichbarer Schwierigkeit

z.B. Haben Sie Zitrone verstanden?

(3 Silben, hochfrequent, NLG=0,204, LexFam=5,0)

Haben Sie Zauberer verstanden?

(3 Silben, hochfrequent, NLG=0,278, LexFam=4,8)

Anwendungsbeispiel 2: HWL

(Hierarchische Wortlisten – Nachsprechtest für die Sprechapraxiediagnostik [7])

Literatur

- [1] Schröder, A., Gemballa, T., Ruppig, S., & Wartenburger, I. (2012). German norms for semantic typicality, age of acquisition, and concept familiarity. *Behav Res Methods*, 44(2), 380-394.
- [2] Reichel, U. (2012). Perma and Balloon: Tools for string alignment and text processing. Paper presented at the 13th Annual Conference of the International Speech Communication Association, Portland.
- [3] Ziegler, W., & Aichert, I. (2015). How much is a word? Predicting ease of articulation planning from apraxic speech error patterns. *Cortex*, 69, 24-39.
- [4] Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLOS ONE*, 7(8).
- [5] Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, 58(5), 412-424.
- [6] Lehner, K. & Ziegler, W. (im Druck). Crowdbasierte Methoden in der Diagnostik neurologischer Sprechstörungen. *Aphasie und verwandte Gebiete*.
- [7] Liepold, M., Ziegler, W., & Brendel, B. (2003). Hierarchische Wortlisten. Ein Nachsprechtest für die Sprechapraxiediagnostik. In W. Ziegler, G. Kerkhoff, N. Troppmann & G. Goldenberg (Hrsg.), *EKN - Materialien für die Rehabilitation*. Dortmund: Borgmann.

Gefördert durch:

