

## Erstellung repräsentativer Textdatenbanken englischer Varietäten

### The International Corpus of English: Malta & Puerto Rico

Seit 1990 arbeiten über 20 Forscherteams weltweit an dem Projekt „International Corpus of English“ (ICE) mit dem Ziel der Erstellung vergleichbarer Korpora (Textdatenbanken) des Englischen. Für die meisten Varietäten stellt dies die erste systematische Erfassung der nationalen Varietät dar. Es existieren bereits mehr als 10 solcher ICE Korpora (u.a. Canada, Jamaica, Great Britain, Hong Kong, India, Ireland, Singapore, East Africa, New Zealand, Sri Lanka und Nigeria).

Jede der Komponenten beinhaltet 1 Mio Wörter (60% gesprochene, 40% geschriebene Daten) und folgt einem einheitlichen Korpusdesign. Um die Varietät möglichst umfassend abzubilden, werden insgesamt Daten aus 32 Genres erhoben (z.B. *face-to-face conversations*,

*legal cross-examinations, broadcast news, student essays, administrative writing, press editorials*).

Der Lehrstuhl für Englische Sprachwissenschaft arbeitet hierbei intensiv an der Erstellung der ICE Komponenten für Malta und Puerto Rico, in denen das Englische jeweils starken Einflüssen seiner Kontaktsprachen (Maltesisch, Italienisch bzw. Spanisch) unterliegt. Da bei der Zusammenstellung des Korpus größtenteils gedruckte Texte und Audioaufnahmen zum Einsatz kommen, ist die Datenaufbereitung ein aufwändiger Schritt. Sie umfasst nicht nur die Digitalisierung (Transkription bei Audio-dateien, Scan und Texterkennung bei Texten), sondern auch das manuelle Hinzufügen von Markup und eine Annotation der Wortarten (*parts-of-speech tagging*).

Aufgrund der bisher kompilierten Daten wurden bereits systematische Erkenntnisse zu ausgewählten Bereichen der Lexik, Syntax und Pragmatik gewonnen.



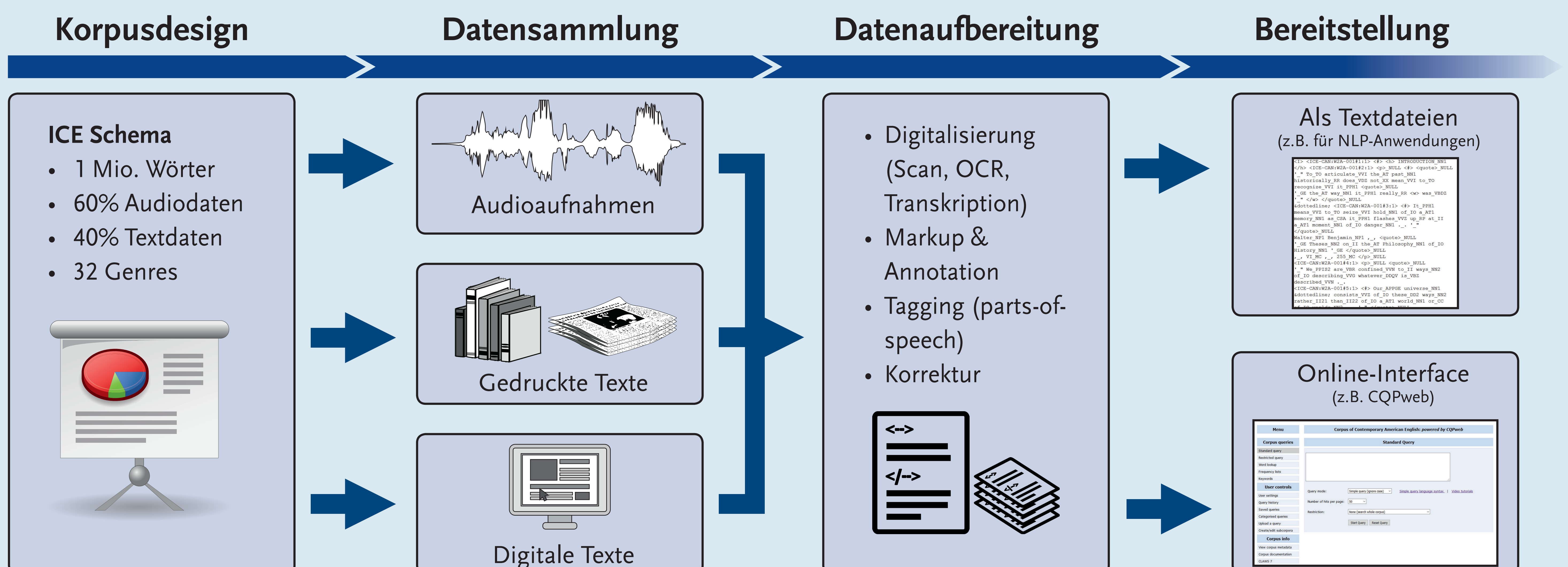
#### Publikationen (Auszug):

Bonnici, Lisa, Michaela Hilbert & Manfred Krug. 2011. Maltese English. In: Bernd Kortmann & Kerstin Lunkenheimer (eds.) *The electronic World Atlas of Varieties of English [eWAVE]*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://www.ewave-atlas.org/>.

Michaela Hilbert & Manfred Krug. 2010. The compilation of ICE Malta: State of the Art and Challenges Along the Way; *ICAME Journal* 34: 54-63.

Krug (2015) "Maltese English." In: Jeffrey P. Williams, Edgar Schneider, Daniel Schreier, Peter Trudgill & (eds.) *Lesser-known Varieties of English* (Vol. 2). Cambridge: CUP, 8-50.

Krug, Manfred; Anna Rosen. 2012. Standards of English in Malta and the Channel Islands. In: Raymond Hickey (ed.) *Standards of English: Codified Varieties around the World*. Cambridge: Cambridge University Press, 117-138.



Kontakt: Otto-Friedrich-Universität Bamberg  
Lehrstuhl für Englische Sprachwissenschaft einschl. Sprachgeschichte,  
Prof. Dr. Manfred Krug, Prof. Dr. Julia Schlüter, Michaela Hilbert, Anja Linne, Ann-Charlott Stegbauer, Fabian Vetter

