Redefining psycholinguistic cognates: Linguistic and historical considerations

Isabella Fritz, Joshua Booth & Aditi Lahiri, University of Oxford

Much of what we know about the bilingual mental lexicon is based on experiments comparing cognates with non-cognates. In psycholinguistics, the term *cognate* refers to translation equivalents that don't only share meaning but also have a phonological and orthographical (for same script-languages) overlap [1]. As highlighted in a recent review paper [2], the current psycholinguistic cognate definition raises the question of a cut-off point between cognates and non-cognates which is operationalised differently across studies. Concretely, this means that various approaches (algorithms/ratings) exist to calculate phonological and orthographical similarity, often based on the assumption that all differences are either weighted equally (e.g., [3]), or the weighting is not based on empirical evidence (e.g., ALINE [4]). Importantly, current approaches neglect the fact that languages are structured systems and that phonological variation across languages is systematic.

Based on our recent line of research, we propose an alternative approach to investigate the role of the L1 phonology in L2 word recognition that is based on theoretical and historical linguistics. We advocate for a linguistically accurate use of the term *cognate*, a concept that in linguistics refers exclusively to inherited words (English: *daughter* [ˈdɔːtə], German: *Tochter* [ˈtɔxtɐ̯]) whilst all borrowed words are considered loans. Loans may come from the same source (from French: E *reptile* [ˈɹɛpˌtʌɪ̯l], G *Reptil* [ˈʁɛpˈtiːl]) which we refer to as *shared loans*. Alternatively, the first language has borrowed the relevant item from the second (English) as in *chocolate* [tʃɔklət] – Mandarin 巧克力 [t͡ɕʰjaʊ̯³kʰɤ⁴li⁴] which was then adapted to the L1 phonological grammar. Cognates only exist in related languages and they do undergo regular sound changes. For example, following the Second Sound Shift (approximately dated to the sixth century), all initial interdental fricatives [θ] in English became voiced stops in German [d] (e.g., *thick* – *dick*, see Table 1).

Why is this relevant for psycholinguistic studies? For cognates, sound change is regular, systematic, and sound correspondences between classes of words remain. This systematicity holds for shared loans, i.e., the ways in which non-native phonemes and metrical patterns are adapted from one language into another is entirely systematic. However, phonological grammars differ between shared loans and cognates leading to varying degrees of phonological and orthographic overlap and on different phonological levels (see Table 2).

In regard to models of bilingual word recognition, recent accounts supporting an integrated bilingual lexicon [5] assume orthographically-driven co-activation based on visual word recognition studies. If, however, we want to advance our understanding of phonology-based co-activation [cf. 6], we need to test step-by-step which types of phonological mismatch between L1 and L2 are permitted so that phonological co-activation still occurs [7]. However, previous studies on L2 processing leave open the question of how this "mismatch" is to be defined or quantified. And this is exactly what our research agenda set out to achieve.

We will draw from a set of ERP cross-modal priming studies with a Lexical Decision Task and with a variety of L1-L2 combinations (i.e., Bengali, Dutch & German L1, English L2). Across multiple studies, we show that bilingual speakers are sensitive to systematic cross-linguistic correspondences and differences in both shared loans and cognates. In our research so far, we investigated systematic cross-linguistic differences/similarities on a featural and prosodic level. However, we will show that such linguistically informed study designs can be used to test for a wide range of phonological phenomena.

/s/–/z/ [OBS], [CONT], [COR], [STRID]				/w/–/v/ [CONT], [LAB], [VOICE]				/θ/–/d/ [OBS], [COR]			
***s**ack*	/**s**æk/	***S**ack*	/**z**ak/	***w**ind*	/**w**ɪnd/	***W**ind*	/**v**ɪnd/	***th**ick*	/**θ**ɪk/	***d**ick*	/**d**ɪk/
***s**and*	/**s**ænd/	***S**and*	/**z**and/	***w**ine*	/**w**ʌɪ̯n/	***W**ein*	/**v**aɪ̯n/	***th**ing*	/**θ**ɪŋ/	***D**ing*	/**d**ɪŋ/
***s**ea*	/**s**iː/	***S**ee*	/**z**eː/	***w**olf*	/**w**ʊlf/	***W**olf*	/**v**ɔlf/	***th**orn*	/**θ**ɔːn/	***D**orn*	/**d**ɔʁn/
/s/ has no voice feature /z/ = [VOICE]				/w/ = [SONORANT] /v/ = [OBSTRUENT]				/θ/ = [CONTINUANT] /d/ = [PLOSIVE] (underspecified)			

Table 1 Classes of corresponding initial consonants in English–German cognates, with the relevant differentiating phonological feature provided below. Shared features are provided in the first row of the table (OBS = OBSTRUENT, CONT = CONTINUANT, LAB = LABIAL, COR = CORONAL, STRID = STRIDENT).

Status	Related E—G words	Orthography	Phonology
Cognates	nest–Nest [nɛst]	✓	✓
	mild–mild [mɪlt]	✓	×
	mouse–Maus [maʊ̯s]	×	✓
	night–Nacht [naxt]	×	×
Recent loans	chef–Chef [ʃɛf]	✓	✓
	garage–Garage [gaˈʁaːʒə]	✓	×
	Cognac–Kognak [ˈkɔnjak]	×	✓
	bureau–Büro [byˈʁoː]	×	×

Table 2: The ways in which related words can differ across English and German in terms of orthographic and phonological overlap.

References

[1] Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, *41*, 496-518.

[2] Roembke, T. C., Koch, I., & Philipp, A. M. (2024). What makes a cognate? Implications for research on bilingualism. *Bilingualism: Language and Cognition*, 1-6.

[3] Mulder, K., Wloch, L., Boves, L., ten Bosch, L., & Ernestus, M. (2021). Cognate status modulates the comprehension of isolated reduced forms. *Language, Cognition and Neuroscience, 37*(5),

[4] Kondrak, G. (2003). Phonetic Alignment and Similarity. *Computers and the Humanities, 37*(3), 273-291.

[5] Dijkstra T, Wahl A, Buyuenhuijs F., et al. Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*. 2019;22(4):657-679.

[6] Jiang, N. (2019). Phonology-based bilingual activation among different-script bilinguals?. *Bilingualism: Language and Cognition*, *22*(4), 693-694.

[7] Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175-197. doi:10.1017/S1366728902003012