Journal of Memory and Language 95 (2017) 49–67

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Word frequency predicts translation asymmetry

Anna Ibrahim [a,*], Patricia E. Cowell [a], Rosemary A. Varley [b]

[a] Department of Human Communication Sciences, University of Sheffield, 362 Mushroom Lane, Sheffield S10 2TS, United Kingdom
[b] Division of Psychology and Language Sciences, University College London, 332 Chandler House, 2 Wakefield Street, London WC1N 1PF, United Kingdom

ARTICLE INFO

Article history:
Received 5 April 2016
revision received 1 February 2017
Available online 21 February 2017

Keywords:
Bilingualism
Frequency effects
Translation asymmetry
Language dominance
RHM

ABSTRACT

Bilingualism studies report asymmetries in word processing across languages. Access to L2 words is slower and sensitive to semantic blocking. These observations inform influential models of bilingual processing, which propose autonomous lexicons with different processing routes. In a series of experiments, we explored an alternative hypothesis that the asymmetries are due to frequency of use. Using a within-language 'translation' task, involving high/low frequency (HF/LF) synonyms, we obtained parallel results to bilingual studies. Experiment 1 revealed that HF synonyms were accessed faster than LF ones. Experiment 2 showed that semantic blocking slowed retrieval only of LF synonyms, while form blocking produced powerful interference of both HF and LF words. Experiment 3 examined translation speed and sensitivity to blocking in two groups of Russian-English bilinguals who differed in frequency of use of their languages. Translation asymmetries were modulated by frequency of use. The results support an integrated lexicon model of bilingual processing.

Introduction

Frequency effects are widely documented in language processing. High frequency (HF) words are processed faster and are more robust to various types of interference. Effects of frequency are found in picture naming (Oldfield & Wingfield, 1965), visual word recognition (Howes & Solomon, 1951) and lexical decision (Paap, McDonald, Schvaneveldt, & Noel, 1987). In aphasia, low frequency (LF) words are more susceptible to disruption than HF ones (Caramazza & Hillis, 1990). In healthy lexical processing, stimulus degradation impacts more on LF forms (Bangert, Abrams, & Balota, 2012). Furthermore, Michael and Gollan (2005) report that LF words are more vulnerable to Tip-Of-the-Tongue states than HF ones. Bilingual studies reveal faster access to words of a frequently-used language (usually the L1) as opposed to the one used less often (Kroll & Stewart, 1994). However, with more frequent use of L2, the frequency bias can change, with faster access to L2 words (Basnight-Brown & Altarriba, 2007; Heredia, 1996; Heredia & Altarriba, 2001; Sunderman & Priya, 2012). Overall, there is extensive evidence of processing advantage for HF over LF forms in monolingual and bilingual speakers.

In bilingual translation, greater susceptibility of L2 than L1 forms to interference (e.g., presentation of stimulus words blocked into semantic categories) has also been reported (Kroll & Stewart, 1994). In the Revised Hierarchical Model (RHM), Kroll and Stewart developed an influential account of processing asymmetries between languages. The basic architecture of the RHM is that words from each language are stored in separate lexicons, and their semantic representations in a single shared module. In its early formulation, Kroll and Stewart proposed that the connections between the lexicons and meanings were asymmetrical. L1 words had direct connections to their meanings, while L2 words accessed them via their L1 equivalents. During translation, presentation of L1 forms resulted in automatic semantic activation prior to access to the L2 forms. As a consequence of this semantic mediation, L1 → L2 translation is slower than in the reverse direction. Furthermore, semantic manipulation, such as blocking stimuli into semantic categories, results in activation of multiple overlapping conceptual representations and further slows processing from L1 to L2 due to the need to inhibit non-target items. By contrast, L2 → L1 translation proceeds by direct lexical connections and is therefore faster and immune to semantic manipulation.

In support of the RHM, Kroll and Stewart (1994) conducted an experiment investigating bidirectional translation between L1 and L2. The study had two components: first, examining translation speeds from L2 → L1 and vice versa; second, determining the impact of semantic interference (created by blocking stimulus lists into semantically-related items) on translations in both directions. They recruited Dutch (native and dominant) – English bilinguals and presented them with sets of nouns that were organised

* Corresponding author.
E-mail addresses: anna.ibrahim@sheffield.ac.uk (A. Ibrahim), p.e.cowell@sheffield.ac.uk (P.E. Cowell), rosemary.varley@ucl.ac.uk (R.A. Varley).

http://dx.doi.org/10.1016/j.jml.2017.02.001
0749-596X/

into random lists or blocked into semantic categories. Words were translated in both directions. The results were consistent with the predictions of the RHM: participants were slower translating L1 → L2 than in reverse. Furthermore, L1 → L2 translations were slowed under semantic blocking conditions, while L2 → L1 translations were immune to semantic blocking.

A core aspect of the RHM is that it is a developmental model. It describes processing in states of non-balanced bilingualism where the L1 remains the dominant language, such as in early L2 acquisition. Kroll and De Groot (1997) propose that with increased proficiency, links between the lexicons and the semantic, conceptual levels become more similar and, therefore, asymmetries in lexical access become less pronounced. Furthermore, the model has been modified in response to evidence that speakers even in the early stages of L2 acquisition can show direct semantic mediation of both L1 and L2 words (Poarch, Van Hell, & Kroll, 2015). Experiments employing tasks such as semantic priming have shown that response times in lexical decision can be facilitated by brief presentation of a semantic prime in either language (Brysbaert & Duyck, 2010; Dimitropoulou, Duñabeitia, & Carreiras, 2011; Duyck & Warlop, 2009; Perea, Duñabeitia, & Carreiras, 2008; Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2009). Consistent with the RHM developmental aspect, patterns of semantic modulation vary with exposure and competence in the L2. Typically, semantic priming is stronger in the dominant language but may shift to the L2 in instances where that is the commonly used language.

Although aspects of the RHM have been modified in the face of evidence of early semantic mediation and to address evolving patterns of asymmetry related to language use, the model still holds to core assumptions of independent lexicons and different structural and sequential processing routes in L1 and L2 access, particularly in the case of imbalanced bilingual states. However, Brysbaert and Duyck (2010) question the assumption of functional and structural autonomy between L1 and L2 lexicons. Evidence of automatic and simultaneous access to words from both languages has accrued from paradigms such as eye-tracking (Spivey & Marian, 1999) and lexical decision (Van Heuven, Dijkstra, & Grainger, 1998). These studies report interference from words of one language when accessing words of another language in monolingual tasks performed by bilinguals. In response, Kroll, Van Hell, Tokowicz, and Green (2010) suggest that perceptual level orthographic and phonological similarity between L1 and L2 forms might be the source of parallel activation of forms in separate lexicons. However, a number of studies report parallel activation of L1 and L2 items in instances where there was little form similarity between words of different languages (e.g., Von Holzen & Mani, 2012), and in the case of stimuli involving different scripts such as English and Mandarin Chinese (e.g., Moon & Jiang, 2012; Thierry & Wu, 2007). An alternative to the RHM architecture of independent lexicons is the proposal of a single integrated lexicon, where both L1 and L2 items are stored and accessed based on common processing mechanisms.

In the face of conflicting accounts of bilingual lexical processing, we attempt to explain the source of the asymmetries that Kroll and Stewart observed. We present a series of experiments that focus on word production and develop an alternative account of asymmetries in processing speed and susceptibility to interference. We propose that the asymmetries may result from differences in the frequency of word use across languages. In most bilingual states, there are imbalances of language use, with the L1 often being the dominant language. As a result, L1 (higher frequency) words are easier to access and more robust to interference than their L2 equivalents. The frequency account does not involve structural or functional differences in processing between lexicons. The same asymmetry might be observed within a language when 'translating' between two frequency-contrasted synonyms, as well as across L1 → L2 translation equivalents. Furthermore, LF words will also be more sensitive to interference resulting from blocked presentation (e.g., into semantic categories) than their HF equivalents within or across languages.

We explored this hypothesis in three experiments. First, we developed a within-language 'translation' task involving synonyms where one member of the pair was of higher frequency than its twin. This allowed us to model 'translation' effects in a paradigm where explanation could not rest upon multiple distinctly-processed lexicons. Further, the frequency imbalance allowed us to model the developmental perspective of the RHM as the experiments involved retrieval of more/less entrenched forms. Thus, Experiment 1 compared access to a HF versus LF synonym, determining if there were asymmetries in processing speed, dependent upon the direction of 'translation'. Using the same task, Experiment 2 explored the impact of blocking stimuli on retrieval of LF (Experiment 2A) and HF (Experiment 2B) members of a synonym pair. We examined if asymmetry in the effect of stimulus blocking was unique to semantics, or whether grouping by form produced similar asymmetry.

Interference related to blocking of stimuli into form-related categories is predicted by connectionist models, such as BIA+ (Dijkstra & Van Heuven, 2002) and stems from the necessity to inhibit competing overlapping word forms in recognition and production. Such form-similarity interference effects are reported in word recognition (Davis & Lupker, 2006; Grainger & Van Heuven, 2003) and production (in overt articulation and inner speech access to tongue twisters) (Dell, 1986, 1988; Dell & Reich, 1981; Oppenheim & Dell, 2008). In bilingualism studies, Sunderman and Kroll (2006) and Sunderman and Priya (2012) also observed form-similarity interference in translation recognition tasks.

In Experiment 3, we explored whether bilingual speakers also demonstrated frequency-modulated behaviour in a traditional translation task. We recruited two groups of Russian(native)-English bilinguals who differed in their frequency of use of L2. We predicted that frequent (dominant) users of L1 would translate faster and show greater resilience to blocking conditions in the L2 → L1 direction. Dominant users of L2, however, were predicted to show a reversed effect with faster translations and resilience to blocking conditions in the L1 → L2 direction.

Methodology

All experiments were granted ethical approval by an institutional ethics panel. Volunteers in experiments gave informed consent to participation.

Across experiments, stimuli included abstract and concrete nouns (see Appendix A for the stimulus lists). Polysemous nouns were avoided, although this was problematic for abstract nouns, some of which had more diffuse semantic representations (Kroll & Tokowicz, 2007). Experiments 1 and 2 employed English synonym pairs selected from the Longman Synonym Dictionary (Urdang, 1986). Members of each pair contrasted in frequency of use (British National Corpus)[1] with one being of markedly higher frequency than its twin, e.g., Enemy (HF, 49 instances per million (ipm)) – Foe (LF, 3.8 ipm). Bilingual stimuli (Experiment 3) were Russian-English translation equivalents. They were matched on frequency of use (I-RU (Russian) and I-EN (English) corpora (Sharoff,

[1] Word frequency matching in monolingual experiments (1 and 2) was also performed on BNC Zipf-scale values (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014) available subsequent to the data collection for this study. The results were consistent with those reported in this paper for both experiments. It was not possible to obtain Zipf-scale values for the bilingual Experiment 3, as no comparison corpus was available.