Iconicity effects in translation direction: Bimodal bilingual lexical processing

Benjamin Anible

Follow this and additional works at: http://digitalrepository.unm.edu/ling_etds

Recommended Citation

http://digitalrepository.unm.edu/ling_etds/41
Benjamin Anible
Candidate
Linguistics
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Jill P. Morford, Chairperson
Sherman Wilcox
Phyllis P. Wilcox
Barbara Shaffer
Judith F. Kroll
Iconicity effects in translation direction:
Bimodal bilingual lexical processing

by

Benjamin Anible

2008 Bachelor of Science, Rochester Institute of Technology, ASL-English Interpretation
2010 Master of Arts, University of Rochester, Linguistics

DISSERTATION
Submitted in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
in
Linguistics

The University of New Mexico
Albuquerque, New Mexico

July, 2016
Dedicated to my father, Christian. I miss your encouragement and insight every day.
# Table of Contents

Approval Page ........................................................................................................... i  
Table of Contents ..................................................................................................... iv  
List of Tables ........................................................................................................... vi  
List of Figures .......................................................................................................... vii  
Acknowledgements ................................................................................................. viii  
Abstract of the Dissertation ................................................................................... x  
Introduction ............................................................................................................. 1  
Iconicity in Translation Recognition ....................................................................... 8  
   Methods ................................................................................................................. 15  
      Participants ........................................................................................................ 15  
      Materials .......................................................................................................... 16  
      Procedure .......................................................................................................... 20  
Data Analysis .......................................................................................................... 21  
   Model Construction ............................................................................................... 22  
   Results ................................................................................................................ 23  
Discussion ................................................................................................................ 28  
   Bimodal Lexical Access ......................................................................................... 29  
   Iconicity Effects ................................................................................................... 30  
   Imageability Effects .............................................................................................. 32  
   Theoretical Considerations ................................................................................. 34  
Conclusion .............................................................................................................. 38  
Iconicity in Translation Production ....................................................................... 41  
   Study 1: Picture Naming ....................................................................................... 45  
      Participants ...................................................................................................... 46  
      Procedure ........................................................................................................ 47  
      Results ............................................................................................................ 49  
      Discussion ....................................................................................................... 54  
   Study 2: Forward Translation ............................................................................... 56  
      Accuracy .......................................................................................................... 57  
      Reaction Time ................................................................................................. 60  
   Study 3: Backward Translation .......................................................................... 63
List of Tables

Table 1 – Translation set phonological overlap rankings. ........................................................... 17
Table 2 – Control measures for experimental stimuli ........................................................................ 20
Table 3 – Main effects and interactions with SRT-score and imageability on RT ............................. 23
Table 4 – Control measures for experimental stimuli ........................................................................ 48
Table 5 – Main effects and interactions: Picture naming accuracy .................................................... 51
Table 6 – Main effects: Picture naming RT ......................................................................................... 54
Table 7 – Main effects and interactions: Forward accuracy .............................................................. 58
Table 8 – Main effects and interactions: Forward RT ......................................................................... 61
Table 9 – Main effects and interactions: Backward accuracy .............................................................. 63
Table 10 – Main effects and interactions: Backward RT ................................................................. 66
List of Figures
Figure 1 – ASL signs for SIT, PROBLEM, TABLE, and COUCH. ......................................................... 17
Figure 2 – (a) Model estimates. Main effects of (b) condition, (c) imageability, and (d) SRT-score. ......................................................................................................................................................... 24
Figure 3 – Interaction of fixed effects on condition by (a) SRT-score and (b) imageability. .... 24
Figure 4 – Three-way interaction of iconicity by SRT-score by imageability. ....................... 25
Figure 5 – ASL signs for KISS, FIX, CARRY, and CONVINCE. ....................................................... 34
Figure 6 – Patterns in high imageability iconic networks: (a) experts & (b) novices. .......... 37
Figure 7 – Patterns in low imageability iconic networks: (a) experts & (b) novices.............. 38
Figure 8 – Sign for “hide” in profiled vs. non-profiled image conditions................................................. 47
Figure 9 – Correlation of: (a) optical flow, and (b) spacebar with by-hand measures. ........ 50
Figure 10 – Nonsigner vs. signer accuracy for profiled vs. non-profiled pictures. ............... 52
Figure 11 – (a) Main effects of picture type RT, (b) imageability, and (c) number of translation equivalents. ........................................................................................................................................................................... 54
Figure 12 – Main effects of (a) iconicity, and (b) SRT-score on forward accuracy.................. 59
Figure 13 – Interaction of iconicity and SRT-score on forward accuracy................................. 59
Figure 14 – Main effects of (a) iconicity, (b) SRT-score, and (c) frequency on forward RT. ..... 62
Figure 15 – Interaction of iconicity and SRT-score on forward RT............................................. 62
Figure 16 – Interaction of iconicity and SRT-score on accuracy. Error! Bookmark not defined.
Figure 17 – Main effects of (a) iconicity, and (b) SRT-score on backward accuracy.............. 64
Figure 18 – Interaction of iconicity and SRT-score on backward accuracy............................. 64
Figure 19 – Main effects (a) frequency & (b) number translation equivalents on backward RT. 67
Figure 20 – Interaction of iconicity and SRT-score on backward RT............................................. 67
Figure 21 – Imagistic and diagrammatic interactions in novice bimodal bilinguals in RHM. ..... 83
Acknowledgements

I would like to express my gratitude to my supervisor Dr. Jill Morford for her support during my PhD work at the University of New Mexico. Without her experience and expertise, this research would never have been successfully completed. In particular, I want to thank her for her support during my frustrating mistakes (e.g. when I accidentally reversed my iconicity scale), for all the impromptu discussions at the lab white board, and the many email exchanges where results finally started to make sense.

I wish to acknowledge the community of Deaf and hearing signers who have made it possible for me to do this work. Special thanks in particular go to Brian Burns, Paul Twitchell, Debbie Wager, Sara Siyavoshi, Teri Jaquez, Martin Watkins, Adam Stone, Jessica Contreras, Erica Israel, Erin Wilkinson, Clifton Langdon, Casey Thornton, Deb Cates, Amy Letteri, and to many others I’ve accidentally overlooked.

I’m thankful to the committee of experts who refereed and substantially improved my work. Dr. Judith Kroll, Dr. Sherman Wilcox, Dr. Phyllis P. Wilcox, and Dr. Barbara Shaffer have been open and available to discuss my research and instrumental in opening up new ways of thinking about linguistics, bilingualism, and signed languages. Working with each of them has been greatly beneficial.

Thanks are also due to the wonderful community of scholars at the University of New Mexico. Among many others, Keiko Beers, Laura Hirrel, Clay Beckner, Logan Sutton, Evan Ashworth, Iphigenia Kerfoot, Jeannine Kammann-Soon, Shelece Easterday, Susan Brumbaugh, Motomi Kajitani, Amy Lindstrom, Aubrey Healey, Ricardo Napoleão, and Jackelyn Van Buren have provided excellent advice and company along the way. Extra thanks goes out to the recording talent. Brian Burns for the ASL stimuli and Mitchell Sances for the English stimuli.
A special thanks goes to my friend and colleague Corrine Occhino for her encouragement and camaraderie. Starting our doctoral studies at the same time is one of the most fortuitous accidents of my career. My gratitude also goes to Dr. Brenda Nicodemus and Dr. Peter Hauser, for their graciousness and warmth when I visited their labs to collect data for this thesis. The VL2 Student Research Committee and UNM Graduate and Professional Student Association generously provided funding for my data collection. Thanks also go out to the students in Barbara Shaffer’s interpreter practicum course for all their hard work helping with accuracy coding: Katherine Wagy, Lyle Jeff, Sarah Walton, Rebekah Miller, Alex Luce, Adrianna Lucero, Emily Mize, and Olivia Ross.

I owe a great debt to my family who have loved me each step of the way through my academic career and provided me the opportunities to pursue the things that excite and fascinate me. You helped more than you know, Mom, Katia and Marie. Gemma, you are my truest ally and closest confidant. Without your support, love, and help I could have never completed this dissertation. Thank you for holding up my arms.
Second language learners of American Sign Language (ASL) have an unusual preference for translating “forward” into their second language (Nicodemus & Emmorey 2013), in contrast to spoken language bilinguals’ preference for translating “backward” into their first language (Seleskovich 1976). The typical bilingual preference for backward lexical translation is modeled by Kroll and Stewart (1994) who showed that forward translation uses slower conceptual mediation and backward translation uses faster lexical mediation, but this asymmetry alone fails
to explain signed language bilinguals’ forward preference. Prior attempts to explain these conflicting patterns for signing vs. non-signing bilinguals based on the prevalence of iconicity in signed languages are inconclusive (Bosworth & Emmorey 2010, Thompson et al. 2009, 2010). However, these failed to distinguish between language-external (transparency and imageability) and language-internal (iconicity) motivated form-meaning mappings. This dissertation investigates whether language-external and language-internal motivated mappings between form and meaning differentially impact conceptual and lexical activation during translation.

Study 1 revisits translation recognition experiments (Talmas et al. 1999, Sunderman & Kroll, 2006) in which participants judge lexical items from two languages and decide if they are equivalent. Results show incorrect translation distractors that are related in both form and meaning to the correct translation inhibit recognition in both novices and experts; however, they parallel semantic distractors for novices, but phonological distractors for experts. Further, language-external conceptual imageability affects novices and experts in a similar manner for unrelated distractors, but not for iconic distractors. Study 2 investigates the interaction of proficiency, imageability, and iconicity during production in forward and backward translation. While both imageability and iconicity facilitate forward translation for novices, they do not facilitate backward translation. Experts, on the other hand, show no facilitation for either measure in forward translation, but do benefit from iconicity in backward translation.

In sum, this dissertation provides evidence that language external motivations facilitate conceptual access for novice learners of signed languages and language internal motivations grow in importance with proficiency. This dynamic explains the unusual preference of spoken/signed bilinguals for forward translation at lower levels of proficiency and provides support for a cognitive usage-based model of the bimodal bilingual lexicon.
Introduction

In physics, the atom is the smallest component of an element that still contains its chemical properties; if an atom is split the components no longer keep the properties of that element. An atom is arguably the smallest unit of matter even though it can be split into protons and neutrons which can in turn be split into elementary particles. These subatomic particles no longer carry the essence of the element. Linguistics has transplanted this understanding of the building blocks of matter as an analog the building blocks of meaning. We learn in introductory linguistics courses that a morpheme is the smallest unit of meaning. It is composed of phonemes which are meaningless. Phonemes are composed of equally meaningless features. Morpheme as atom, phoneme as sub-atomic particle and finally feature as quark – the metaphor is very seductive. Perhaps in an effort to gain some of the legitimacy and glamour of the physical sciences, linguists were overzealous in adopting the metaphor of morphemes are atoms without fully accounting for all the evidence.

Since Saussure’s (1916) discussion of the primarily arbitrary nature of the sign, certain linguistic phenomena that do not map to the morphemes are atoms metaphor have been relegated to the periphery of linguistic inquiry. Ideophones, are one example. They have been documented in Japanese (Itô & Mester, 1995), Aboriginal Australian languages (Courtenay, 1974), Mayan languages (Durbin, 1973), and many pidgins and creoles (DeCamp, 1974). What these systems all have in common is that the smallest unit of meaning is something smaller and less easily definable than a morpheme. “Meaning” in these systems exists subsymbolically (Hay & Baayen, 2005). According to a subsymbolic view neither the word, nor the morpheme are viable candidates for the basic units of the lexicon. Instead, finer grained units exist via paradigmatically supported matches to construct meanings that are not discrete. Support for this
view has been found by studying the processing of bound stems, affixes, and phonaesthemes (Bergen, 2004; Burani & Thornton, 2003; Wurm, 1997). Thus, even speakers in languages without established ideophone systems access highly schematic sound-meaning mappings in addition to more conventional/arbitrary word forms (Monaghan, Shillcock, Christiansen, & Kirby, 2014; Pena, Mehler, & Nespor, 2011).

For example, Reilly and Kean (2007), in an English corpus-based analysis, detected patterned phonological characteristics for nouns that varied based on their imageability. Imageability is a measure of how often adults report that a strong mental image is evoked by an item. Etymology, syllable structure, phonological complexity, word length, prosody, and neighborhood density were included in the analysis. The strongest correlations between factors and imageability were from word length and etymology. Specifically, longer words were highly correlated with low imageability nouns and shorter words with high imageability nouns; Germanic word-origin was correlated with high-imageability and Latinate with low imageability.

In the MORPHMES ARE ATOMS metaphor, correlation between properties of meaning and properties of form is contradictory – elementary particles are not elements in and of themselves and so features of form should not have meanings in and of themselves, either.

In another study, Monaghan, et al. (2011) propose that arbitrariness and systematicity are competing forces in language where highly arbitrary language parts allow for ease of comprehension since lack of patterning causes low neighborhood densities, but systematicity allows for ease of learning. They hypothesize that a mix of arbitrariness and systematicity is ideal for language learning. Using machine and human learning of artificially constructed languages as well as corpora of English and French they tested this hypothesis. Machine and human learning studies found that arbitrariness in words was advantageous for language
learning. For the corpus studies they found that word onsets were more often arbitrary and the coda was more often systematic – suggesting this is the way the languages studied handle the need for both arbitrary and systematic information in the lexicon. These findings suggest that phonological-semantic arbitrariness within the lexicon of English, and presumably all spoken languages, is not inviolate. While duality of patterning need not be completely dismissed, theories that rely on discreteness at various levels (such as word processing, reading, and language acquisition) are recommended for reexamination.

Much psycholinguistic literature is not unamenable to this perspective. Consider models of how the meaning of a word like flash is activated. The distributed COHORT model proposes there is simultaneous activation of both phonological and semantic units (Gaskell & Marslen-Wilson, 1997, 2001, 2002). In bottom-up, phonological processing, language users successively narrow the list of candidates from all the words they know that begin with [f] to those that begin with [fl] and so on. In top-down processing the candidates activated in the initial moments of phonological access profligate related concepts in semantic space, aiding the comprehension of phonological input. The result is that in the early time-course of lexical processing, spreading activation has little resemblance to familiar word meanings where a specific form is paired with a specific meaning.

This model of lexical access supports the perspective of usage-based cognitive theories in which a word is a level of symbolic structure that emerges from successive usage events over time in the mind of a speaker (Bybee, 2001). These units emerge from the abstracted sum of language experiences. In this way, related sounds are grouped together and related meanings are grouped together along with the context in which they occurred. When instantiations of sounds are connected to instantiations of concepts, then words are perceived and understood. For
example, the word *flash* is built up from the meanings of related words like *flimmer, flicker, flame, flare* all having to do with the emission of light as well as words like *bash, brash, clash, crash, dash* all having to do with violence and/or speed (A. Fischer, 1999). The combined meaning of *fl-* joined to the combined meaning of *-ash* then results in a meaning of *flash* that is largely inevitable and non-arbitrary, or has some probability associated with it.

Pairing of form and meaning is a critical step in the acquisition of language and results from this literature are also situated comfortably within a usage-based cognitive framework. In both first and second language acquisition, related sounds can be mistaken for each other or mispronounced and related meanings can be confounded. This process occurs when a child has received too few tokens to solidify the exact form and meaning of a given symbolic unit. When talking about the crescent shape in the sky at night a child might produce a related form like “Look at the *boon*, daddy!” – demonstrating that the exact quality of the first segment of *moon* is still flexible in their representation. A child might also overextend the scope of a unit, mistaking a subordinate for a hyperordinate and say “Look at your *moon*, daddy!” in reference to a fingernail, because both the moon and a fingernail can have a crescent shape. Other common lexical errors include mispronouncing words with difficult/uncommon articulations (Ferguson, Menn, & Stoel-Gammon, 1992; Maye, Werker, & Gerken, 2002) or that violate prosodic patterns in the language (Saffran, Aslin, & Newport, 1996). Common errors at the conceptual level include mistaking the scope of super- and sub-ordinate categories or the fine-grained distinctions between other related concepts like *more/less* and *buy/sell* (Brown, 1973). These separate types of language errors are typically interpreted as support for the fundamental distinction between semantic and phonological levels of language processing, but in a usage-based framework they
also provide evidence for scalar connections between form and meaning that stretch more conventional understandings of the nature of a word and/or morpheme.

The majority of language acquisition literature (not to mention the majority of psycholinguistic literature) deals primarily with English. Signed languages are unique in part due to the prevalence of visually motivated mappings between form and meaning. The sign for “ball” in American Sign Language (ASL) looks like holding a ball, the sign for “cat” looks like whiskers, etc. The frequency of these “iconic” mappings are likely due to the greater ease of codifying visual aspects of the world with our hands compared to codifying auditory ones with our voices. Early explorations in child language acquisition of ASL found little effect of iconicity (Meier, 1987; Orlansky & Bonvillian, 1984; Petitto, 1987), but more recently studies have found evidence that contradicts these initial conclusions (Thompson, 2011). Parts of the ASL lexicon, due to their iconicity, are easier to remember and learn (Thompson, Vinson, Woll, & Vigliocco, 2012). In this study parents reported on the comprehension and production of 84 BSL signs in their children age 30 months to 8 years-old. Iconicity of the signs was assessed on a seven point scale as rated by deaf adult BSL users (Vinson, Cormier, Denmark, Schembri, & Vigliocco, 2008). Results showed that increasing iconicity resulted in better comprehension and production of signs and that this effect increased with age. This indicates that in child language acquisition of a signed language, core assumptions about word learning are challenged.

This contention makes it clear that iconicity is a nuanced issue and while the complete picture is not yet clear – it is probable that lexical representations in signed languages pose challenges for our understanding of the nature of the connection between form and meaning in similar ways as evidence from ideophones, phonaesthemes and other subsymbolic meaning structures. For signed languages, in particular, duality of patterning is unlikely to be a rule so
much as the end result of a process at various stages throughout the lexicon. Proficiency in a signed language involves more than knowing a collection of lexical forms and their respective meanings. It involves knowing both the meanings of lexical items and also their formational and conceptual similarities that are systematic within the networks of the language. It seems likely that the behaviors of lexical and conceptual storage are not clearly separable for some levels of ASL. As Stokoe once framed this issue: “The metaphor (…) that jumps to mind is the Möbius strip: the input is the output—with a twist!” (2001, p. 439). For signed languages, categorical separation of phonological and semantic levels of processing may be disingenuous.

This dissertation argues that while subsymbolic analogical relationships are an intrinsic part of any lexicon (Krott, Baayen, & Schreuder, 2001), spoken or signed, they are likely uniquely influenced by imagistic motivations in signed languages. The challenge, then, is to isolate the effects of imagistic motivations in a signed language lexicon at the subsymbolic level from those that are built up from analogical interactions. English-ASL bimodal bilinguals are uniquely positioned to answer this question. The stark imbalance of imagistic motivations between their two languages means iconicity effects are unlikely to be due to qualities of English, only ASL. The other side of this coin is that a view of lexical access that is nuanced enough to disambiguate language external effects like imagistic motivations from language internal effects like subsymbolic analogical relationships has the potential to inform outstanding questions about the structure of a spoken and signed language residing in the mind of a bilingual.

In particular, Kroll and Stewart’s (1994) Revised Hierarchical Model (RHM) posits weak lexical links connecting second language (L2) word forms to conceptual storage and strong L2 word to first language (L1) word connections that circumvent direct L2-conceptual access. This helps to explain spoken language bilinguals’ preference for backward translation from their L2
into their L1. However, recent studies have found that English-ASL interpreters prefer forward translation over backward translation so long as ASL is their L2 (Nicodemus & Emmorey, 2013). The very robust finding across multiple studies that hearing bilinguals show a preference for L2 to L1 translation is a pattern that is captured in the RHM. Recent evidence indicates that this well-documented pattern does not hold for bimodal bilinguals. This dissertation investigates whether these different patterns across unimodal spoken language bilinguals, and bimodal bilinguals are related to the prevalence of subsymbolic analogical relationships. This question will be examined first in a study of recognition of translation equivalents between English and ASL, and second in a study of translation production. Both studies are prepared as individual manuscripts, ready for independent review. Implications for modeling bimodal bilingual access within a hierarchical model will then be outlined based on the results of these studies in the concluding remarks.
Iconicity in Translation Recognition

Abstract

RT and accuracy measures are reported for a translation recognition study in which ASL-English bilinguals at various levels of ASL proficiency rejected English translation distractors for ASL signs phonologically, semantically and iconically related to the correct translation. Imageability ratings of concepts impacted performance in all conditions such that when imageability ratings were high, participants showed inhibition for phonologically related distractors and when imageability was low participants showed inhibition for semantically related distractors regardless of proficiency. For iconically related distractors, imageability impacted performance differently as proficiency increased; high imageability caused inhibition in experts, but low imageability caused inhibition in novices. These patterns suggest that imageability and iconicity interact with proficiency – experts process iconically related distractors phonologically, but novices process iconically related distractors semantically. Implications for measuring and modeling form/meaning connections in signed languages are discussed within a usage-based cognitive framework.

Keywords: bilingualism, sign language, iconicity, translation recognition, usage-based, cognitive linguistics
Iconicity in American Sign Language-English Translation Recognition

Signed languages are unique in part due to the prevalence of visually motivated mappings between form and meaning. The sign for “ball” in American Sign Language (ASL) looks like holding a ball, the sign for “cat” looks like whiskers, etc. The frequency of these “iconic” mappings are likely due to the greater ease of codifying visual aspects of the world with our hands compared to codifying auditory ones with our voices. Since Saussure’s (1916) discussion of the primarily arbitrary nature of the sign, iconicity has been viewed as the red-headed stepchild of symbolic relationships. Recently, a growing literature has begun to show that iconicity is less of an exception than originally thought, appearing across modalities and languages. Rather than being antithetical to normal language use, it exists harmoniously alongside arbitrary structures. Spoken language linguists document iconicity at multiple levels of linguistic structure; the discourse level (Haiman, 1980, 2003; Kaiser, 1999), the morphosyntactic level (Bybee, 1985; Givón, 1985), and the morpho-phonemic level (A. Fischer, 1999; Reilly & Kean, 2007). Researchers are increasingly arguing that iconicity as a systematic property should be viewed as a quality of all languages that operates in a balance with arbitrariness (Monaghan et al., 2011, 2014). Iconicity in signed languages, should not be seen as a special feature primarily relegated to the visual/manual modality, but as a manifestation of a general human tendency (Perniss, Thompson, & Vigliocco, 2010; Perniss & Vigliocco, 2014).

The goal of this study is to replicate the findings from spoken language translation recognition experiments, and to extend research in this domain to the visual modality in order to explore the effects of visually motivated language internal systematicity in processing as proficiency in a signed language increases. Translation recognition is an experimental paradigm that is well suited to resolving questions about the interaction of form and meaning.
Developmental hypotheses have also been successfully vetted for spoken language bilinguals using this paradigm. Participants view lexical items from two languages and decide if they are equivalent. The way in which two items are unequivalent reveals the degree to which conceptual and lexical levels of memory are activated using reaction time and accuracy measurements. Prior research has found that novice spoken language (unimodal) bilinguals are more sensitive to incorrect (*distractor*) second language (L2) translations of first language (L1) words that are phonologically related to the correct L2 translation, whereas expert unimodal bilinguals are more inhibited by distractors that are semantically related to the target (Talamas, Kroll, & Dufour, 1999). In the Revised Hierarchical Model (Kroll & Stewart, 1994), this is attributed to novices activating L1 translation equivalents as an intermediate step in accessing semantic representations (*lexical mediation*), whereas experts activate semantic representations directly (*conceptual mediation*). More recent research (Misra, Guo, Bobb, & Kroll, 2012; Sunderman & Kroll, 2006; Wu & Thierry, 2010) with spoken language bilinguals suggests that behavioral data may not always be sensitive to proficiency differences. It is probable that any replicable effects in spoken/signed language bilinguals will occur at later (i.e. conceptual) levels of processing.

There is reason to expect differences in translation recognition between unimodal bilinguals and spoken/signed (bimodal) bilinguals. Baus et al. (2013) study of English-ASL bilinguals found that novice bimodal bilinguals were faster to recognize correct transparently iconic sign translation pairs (e.g. STIR/*stir*) presented simultaneously as picture and text, than non-transparently iconic sign translation pairs (e.g. SING/*sing*). Expert bimodal bilinguals were slower to recognize correct transparent translations than non-transparent translations. These results are difficult to interpret relative to a standard spoken language translation recognition paradigm because participants were responding to simultaneously presented stimuli in two
modalities, because responses were to correct rather than incorrect translations, and because there were no purely phonologically or semantically related translation pairs to which transparently related pairs could be compared. There is, however, a compelling case to be made that novices and experts do process translations differently depending on the level of transparency of the ASL sign. This may be at least partially attributable to imageability; the extent to which a word’s referent evokes a visual or auditory image, which has been positively correlated with non-signer ratings of transparency (Baus et al., 2013; Thompson, Vinson, & Vigliocco, 2010).

Unimodal studies that have examined high and low imageability concepts in bilingual memory have found similar effects for high imageability concepts for both expert and novice bilinguals, indicating there is always conceptual mediation for words with high imageability (de Groot & Poot, 1997; Van Hell & De Groot, 1998). For example “free” is a word with low imageability, but “kiss” has high imageability. These studies suggest that words with high imageability are translated primarily using conceptual mediation. Words with low imageability, on the other hand, are translated primarily using lexical mediation for novices, but conceptual mediation for experts. While there are issues of a familiarity confound, even if frequency is matched for high and low imageability items, this pattern can be explained by proposing that high imageability concepts activate more shared conceptual nodes than low imageability concepts, regardless of levels of L2 language experience, but that low imageability concepts activate only language specific conceptual nodes for novices, favoring lexical connections in translation (Dong, Gui, & Macwhinney, 2005).

It is likely overly optimistic to assume that imageability effects would behave the same in an oral/auditory language as in a manual/visual one. However, the observation that high
imageability concepts facilitate conceptual processing in second language learners without regard to language experience does suggest that differences between novice and expert learners of ASL attributed to transparency require careful interpretation since they may instead stem from language external properties including both imageability and transparency. Two important findings make a compelling case that transparent motivation creates a link between form and meaning that affects lexical processing in signed languages. The first finding is one also found in spoken language literature on simulation (Barsalou, 2008) where phrases such as “The ranger saw the eagle in the sky,” evoke a visualization of extended wings which facilitates naming a picture of an eagle with wings extended over naming a picture of an eagle with wings folded (Zwaan & Madden, 2005). In signed languages, this semantic facilitation effect is found for individual lexical items that are transparent (Grote & Linz, 2003; Ormel, Knoors, Hermans, & Verhoeven, 2009; Thompson, Vinson, & Vigliocco, 2009). Another related effect, so far unique to sign language research, is the finding that transparency inhibits RTs in form-based decision tasks, such as phoneme monitoring (Anible, Occhino-Kehoe, & Kammann, 2013; Thompson et al., 2010). When signers are asked to indicate whether a particular phonological parameter is present in a sign, decision latencies for transparent signs are longer than for non-transparent signs. These two effects suggest that transparency impacts processing whether or not semantics are explicitly evoked.

It is unfortunate that most psycholinguistic treatments typically define iconicity operationally only as a transparent mapping due to their reliance on non-signer iconicity judgments. Transparency is likely related to language internal motivation in ways that have yet to be completely understood, but recent explorations of this topic (Lepic, 2015; Occhino-Kehoe, 2016) have suggested that the majority of “iconic” signs are diagrammatic (Haiman, 1980;
Ungerer, 1999), rather than solely *imagistic*, indicating that conceptual associations motivate lexical structure and vice versa.

In this perspective, language external motivations are conceptualizations based on general human experiences. The concept of throwing a ball inherently encodes mental representations of experiencing ball-throwing, experiencing watching ball-throwing, etc. These representations are often highly imagistic and are capable of being encoded to linguistic behaviors more directly in signed languages than in spoken languages. Language internal motivations, on the other hand, do not necessarily overtly encode a visual experience or manual behavior, but are instead part of a collection of linguistic forms that share conceptual similarities. Language external motivations are more likely to be accessible without experience in a given signed language, but language internal motivations are only accessible once the morphological and diagrammatic patterns of a signed language have been mastered.

What this means from an experimental stand-point is that distractors that are related through diagrammatic iconicity share similarity both phonologically and semantically with the correct translation. Treating iconicity as a diagrammatic and language internally systematic property gives important explanatory benefits. For example, why else would signers perceive their own signs to be more iconic than the signs of other signed languages (Occhino-Kehoe, Anible, Wilkinson, & Morford, forthcoming)? A cognitive/usage-based approach to iconicity (Occhino-Kehoe, 2016; P. Wilcox, 2000; S. Wilcox, 2004) also implicitly accounts for degree of language exposure, since the learning of these language internal patterns is accumulated over the course of an individual’s experience with a signed language.

This study hypothesizes that the effect of language external motivation (operationalized as imageability) in a translation recognition task will not vary based on level of language
experience, but that the effect of language internal motivation (operationalized as diagrammatic iconicity) will change as proficiency in a signed language grows. If effects of high imageability are similar to reported effects of high transparency in signers, then participants should show more interference (i.e. the difference compared to the unrelated condition) when rejecting phonological distractors for concepts with high imageability, but show less interference when rejecting semantic distractors for high imageability concepts. On the other hand, if imageability affects bimodal bilinguals the same way as unimodal bilinguals (as in de Groot & Poot, 1997), both experts and novices should show interference when rejecting semantic distractors, but not show interference when rejecting phonological distractors for high imageability concepts. Translation recognition for low imageability concepts should cause novice participants to show interference when rejecting phonologically related distractors, but not semantically related distractors and experts should show increased interference when rejecting semantically related distractors, but not phonologically related distractors (as in Talamas et al., 1999). From these interference measures we can infer the presence of an inhibitory mechanism.

There is no a priori evidence for the behavior of diagrammatic iconicity in signed language processing, so iconically related distractors may not be processed any differently from unrelated distractors (but see spoken language work on morphological relatedness Feldman, 2000; Qasem & Foote, 2010). If there is inhibition from iconicity, this study predicts there should be differences based on language experience; (1) inhibition from iconically related distractors could be greater than the inhibition from phonologically related distractors, but less than the inhibition from semantically related distractors indicating an additive effect of phonology and semantics; (2) inhibition from iconically related distractors could be less than inhibition from both semantic and phonological distractors indicating a subtractive effect of
language internal processes. Another possibility is (3) that the interaction of iconicity and imageability might be non-linear, meaning that novices and experts could differentially favor phonological or semantic activation for iconically related distractors as experience with language increases. Each of these possibilities has interesting implications for entrenched beliefs about the separation of phonological and semantic levels in language processing for monolinguals, bilinguals and bimodals.

**Methods**

One difficulty in examining iconicity is task selection. In the past, studies of iconicity have suffered from criticism that participants exhibited effects of iconicity primarily because the task explicitly identified the features of iconic representation being profiled by individual stimuli. Grote and Linz (2003) found that when German signers were presented with the DGS (Deutsche Gebärdensprache) sign COW (iconic of a cow’s horn), followed by either a picture of a cow’s head (iconic condition), a spotted cowhide (non-iconic condition), or a suitcase (unrelated condition) in a sign-picture identification task, they showed faster RTs when presented with a cow’s head than a standing cow. No difference was observed for non-signers. However, because the sign could be used to refer to both COW and for HORN, the faster RTs in the iconic condition could be the result of the sign’s polysemy, rather than iconicity alone. The present study examines the effects of iconicity without the risk of explicitly priming the identity of the visual aspects of iconic items.

**Participants**

Fifty-seven English-ASL bilinguals (44 female, 13 male) participated in the experiment and were compensated either monetarily or with course credit. Data from two participants were excluded due to equipment failure. Participants were native speakers of English 18 to 58 years old ($M =$
24.6, $SD = 8.9$). Participants’ experience with ASL ranged from as little as three semesters of classroom instruction to experienced sign language interpreters ($M = 10.3$ years of experience, $SD = 8.6$). ASL proficiency was assessed using the ASL Sentence Repetition Task (SRT) (Hauser, Paludnevičienė, Supalla, & Bavelier, 2008), a test developed to assess levels of proficiency for native users of ASL. Participants viewed 20 sentences of increasing complexity and were asked to repeat the sentence back exactly after it was signed to them. Exact repetitions were scored as 1; repetitions with replacements or errors were scored as 0. A native deaf signer evaluated the participants’ responses. Participants’ SRT-score ranged from 0 to 10 ($M = 2.98$, $SD = 2.56$). Often, current ASL proficiency assessments rely on participants self-rating their production and comprehension skills (c.f. B. Anible et al., 2015). Although low average scores indicate that the test is not an ideal measure for second language learners, the SRT measure should still be considered superior to self-ratings.

**Materials**

There were a total of 48 translation sets. Each set included four incorrect translation distractors; unrelated, phonologically related, semantically related, and iconically related in both form and meaning to the correct translation, which was never presented. Only filler trials were presented as correct translations. For example the English word “sit” was the correct English translation for the distractor set, PROBLEM (phonologically related), TABLE (semantically related) and COUCH (iconically related).

Phonologically related distractors included signs that overlapped in at least one of three sign parameters; handshape, location or movement. A weighted scoring system was used to assess visual phonological relatedness in a psychologically motivated paradigm. Overlap in movement and location was given a *high* rating, overlap in movement and handshape or
handshape and location was given a medium rating and overlap in just one parameter was given a low rating (See Table 2). These levels of overlap are based on Corina and Hildebrandt’s (2002) finding that participants asked to judge the phonological similarity between a target-sign and surrounding flanker-signs rated some combinations of sign parameters higher than others. High overlap accounted for 32% of the items, mid overlap for 36%, and low overlap for the remaining 32%.

<table>
<thead>
<tr>
<th>Phonological relatedness</th>
<th>Overlap type</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>movement/location</td>
</tr>
<tr>
<td>moderate</td>
<td>movement/handshape</td>
</tr>
<tr>
<td>low</td>
<td>handshape</td>
</tr>
<tr>
<td></td>
<td>location</td>
</tr>
<tr>
<td></td>
<td>movement</td>
</tr>
</tbody>
</table>

Table 1 – Translation set phonological overlap rankings.

Figure 1 – ASL signs for SIT, PROBLEM, TABLE, and COUCH.

Semantically related distractors were semantic associates of the correct translations. To retrieve relatedness measures that have demonstrable effects on processing, the top three associates for each item were drawn from the Edinburgh Associative Thesaurus (Kiss,
Armstrong, Milroy, & Piper, 1973) and the highest ranked associate that was not part of an English collocation was selected. For example, for “kiss”, “me” is the highest ranked associate because of its presence in the collocation “kiss me”, but “love” is the next highest ranked associate and was selected for the item in the semantically related condition of the set. Additionally, semantic associates in the top three occasionally had no single ASL equivalent and/or are translated using the same lexical item. For example, for “camp”, the highest ranked associate was “tent” – these two terms are near homophones in ASL. In these cases the next highest item with a single translation in ASL was selected (HOLIDAY). The average ratio in the Edinburgh Associate Thesaurus of responses for the semantically related distractor given the correct translation was about one in ten ($M = 0.09, SD = 0.1$).

Iconically related distractors were both phonologically and semantically related to the correct translation. In cognitive iconicity (S. Wilcox, 2004), signs highlight individual properties of meaning as well as the individual properties of form. The V-handshape can potentially represent the number two, or pairs of things such as eyes or legs, etc. Semantic properties of concepts can have multiple instantiations as well. The concept of “sit” includes many features that are interconnected like the position of the legs or the object that is sat upon. Signs are capable of selecting any of these semantic characteristics to profile. A pair of signs that are related iconically share properties of both form and meaning as in SIT and COUCH. Phonologically, both use the bent V-handshape. Semantically, both select the way legs behave while sitting as the property that is profiled. In SIT the sign profiles one person’s legs, and in COUCH the sign profiles a set of legs in a row. Other signs, like DESK, are semantically related to “sitting,” but profile a different property like “flatness.” The sign PROBLEM is phonologically related to SIT, but the same properties of form profile unrelated semantic features.
of different conceptual domains.

After completing the experimental task, participants were asked to rate the number of English translation equivalents for the correct translations of the items. For example, the English word “run” has at least four lexical items in ASL that express the concepts of “jogging”, “machinery functioning”, “liquid flowing,”, and “a run in fabric.” Some studies have shown that a higher number of translation equivalents can slow lexical access in tasks involving translation (Baus et al., 2013; Laxén & Lavaur, 2010; Tokowicz & Kroll, 2007). The average score of this test was 1.83 (SD = .72). Four items were removed because the average number of translation equivalents was more than 2.5 standard deviations above the mean. Items in other conditions were not controlled for number of translation equivalents due to the rigorousness of collecting these values from participants.

Items in all conditions were matched as closely as possible for word length, frequency, concreteness and imageability. These measures have all been shown to affect lexical access in online processing (Hudson & Bergman, 1985; Jescheniak & Levelt, 1994; Poarch, Van Hell, & Kroll, 2015). Scores were drawn from the MRC Psycholinguistic Database (Coltheart, 1981). In cases where an item was not available in this database, the average of the condition was used. These cases accounted for 5% (13/278) of the control values used. A series of one-way ANOVAs was performed to evaluate differences between the experimental conditions (excluding fillers) on the control measures. There were no significant effects of condition on length ($F(3,177) = 0.91$, $p = \text{n.s.}$, $MSE = 2.63$), log frequency ($F(3,177) = 1.36$, $p = \text{n.s.}$, $MSE = 2.22$), concreteness ($F(3,177) = 1.19$, $p = \text{n.s.}$, $MSE = 10477.49$), or imageability ($F(3,177) = 1.75$, $p = \text{n.s.}$, $MSE = 7058.99$) across conditions.

---

1 A dictionary of 150837 words with up to 26 linguistic and psycholinguistic attributes for each entry.
Procedure

Four lists of 96 trials were constructed so that no item in a translation set would be repeated in a list for any individual participant. Each list contained 48 fillers and 48 critical “incorrect” distractors, with twelve items per condition. The 48 fillers were the same across the lists. The presentation order of the trials was randomized so that items in the same conditions would never occur sequentially. A practice block of five trials was shown at the start of the experiment to familiarize participants with the protocol. Participants were tested individually and received instructions in spoken English as well as written English instructions that appeared on the computer screen. The task was to decide whether two words were translation equivalents of each other. The translations sets were always presented so that the ASL sign was displayed first,
followed by the English word. All stimuli were presented as either a short video clip of a sign or an audio recording of a word. Auditory presentation was selected over visual written presentation since ASL-English interpreters typically use spoken, not written English, in interpreting contexts and because there is evidence that code blends of speech and sign together have advantages for the comprehension system (Emmorey, Petrich, & Gollan, 2012) that would pose additional challenges in comparing results to spoken language translation recognition studies. Before the two items were presented a fixation point was displayed in the center of the screen. The participant started each trial by pressing the “enter” key on the keyboard. The fixation timer was displayed for 200 ms, then the video and the timer for that trial began simultaneously. 500 ms after the video ended, the audio stimulus played. The participant pressed either the “yes” key (on the left hand) or the “no” key (on the right hand), after which the timer stopped. Participants were instructed to keep their hands over the response keys and to make their responses as quickly and accurately as possible. The procedure was designed using Expyriment, a Python library for cognitive and neuroscientific experiments (Krause & Lindemann, 2014). The code is available on-line (Anible & Anible, 2015).

**Data Analysis**

Only correct responses on critical trials were included in the RT analysis. Correct responses were “no” key-presses after non-filler trials. Accuracy was automatically coded by Expyriment. RTs that were faster than 200 ms or slower than 5000 ms were removed from the analysis as outliers. Twelve observations, less than .3% of the data was trimmed in this way. Typically, very fast RTs are indicative of anticipatory responses and very slow times are indicative of attention lapses, not reactions to stimuli (Ratcliff, 1994). Additional trimming by subject and item means was not performed since analyses use mixed-effects modeling rather than by-subject and by-item ANOVAs – where the need to optimize for central values is more critical (Baayen & Milin,
Mixed-effects models with crossed random effects for participants and items (Baayen, 2008; Baayen, Davidson, & Bates, 2008) were estimated using the lme4 package (Bates, Maechler, Bolker, & Walker, 2013). P-values were calculated using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2013).

Model Construction
Mean accuracy was .94; accuracy per item ranged between .75 and 1.00 (SD = .18); accuracy per subject ranged between .80 and 1.00 (SD = .21). An omnibus model for RT was estimated first, using only correct trials. Fixed effects included linear effects and interaction between experimental variable, participant variables (age, and SRT-score) and stimulus variables (word length, frequency, concreteness, imageability and number of translation equivalents, phonological relatedness). Random effects were random intercepts for participants and items. Models were gradually trimmed to remove nonsignificant fixed effects. Word length, frequency and concreteness and phonological relatedness had no significant effects or interactions, but there were main effects and interactions for imageability. An interaction for the subject variable SRT-score was found, but no effect was found for age.

The model was subjected to criticism focusing on the residuals (Baayen & Milin, 2010). Data points (n = 46; 2% of the data) with absolute standardized residuals exceeding 2.5 standard deviations were removed from the data set and the model was re-estimated, resulting in improved fit. Correlation of fitted values with trimmed RTs was $R^2 = .53$, up from $R^2 = .43$; BIC = 31167.22, down from 33364.25. All subsequent analyses employ the trimmed data set.

Experimental conditions, SRT-score, and item imageability interacted significantly with the intercept, indicating that language experience and the ease with which concepts can be visualized affected the way items were processed. The final model included the modulation of
the effects of the experimental variable (phonologically related, semantically related, and iconically related) in comparison with the base-line (unrelated), crossed by the imageability of the translation targets (Imageability), and by the language experience of the participants (SRT).

In sum, the model included: (a) the set of fixed effects; (b) the fixed linear effect of experimental variable interacting with SRT-score, imageability and their interactions; and (c) a random structure including intercepts for participants (subject) and stimuli (item). The model (in R notation) was specified as (RT ~ Condition * Imageability * SRT + (1|item) + (1|subject).

Results

Table 4 shows the results of fitting this model for the experimental variables and interactions. To help interpret these results, Figure 2 and Figure 3 display the corresponding main effects and interaction plots, created with the effects package (Fox, 2003) and the ggplot2 package (Wickham, 2009). Figure 2 shows (a) the model estimates, and main effects of (b) the experimental condition, (c) imageability, and (d) SRT-score on the unrelated translation condition mean RT intercept. Figure 3 shows the interaction of the experimental conditions on RT with SRT-score (a) and with imageability (b). Figure 4 shows the three-way interaction of the experimental condition and imageability and SRT-score on RT, where the six panels show the incremental effects of increasing SRT-score.

<table>
<thead>
<tr>
<th>Interactions</th>
<th>Main effect × SRT-score</th>
<th>× Imageability</th>
<th>× SRT-score × Imageability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>t</td>
<td>p</td>
</tr>
<tr>
<td>Conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonologically related</td>
<td>-281</td>
<td>-1.76</td>
<td>.078</td>
</tr>
<tr>
<td>Semantically related</td>
<td>583</td>
<td>3.63</td>
<td>.000</td>
</tr>
<tr>
<td>Iconically related</td>
<td>425</td>
<td>2.45</td>
<td>.014</td>
</tr>
<tr>
<td>Imageability</td>
<td>-0.4</td>
<td>-1.69</td>
<td>.093</td>
</tr>
<tr>
<td>SRT</td>
<td>-36.5</td>
<td>-1.18</td>
<td>.240</td>
</tr>
</tbody>
</table>

Table 3 – Main effects and interactions with SRT-score and imageability on RT.
Figure 2 – (a) Model estimates. Main effects of (b): condition, (c) imageability, and (d) SRT-score.

Figure 3 – Interaction of fixed effects on condition by (a) SRT-score and (b) imageability.
How fast are participants able to reject phonologically related distractors, compared to unrelated distractors? For example, after they are presented with the target sign (SIT) how long does it take to reject the phonologically related distractor (problem) in comparison to the unrelated distractor (yellow)? Participants are quicker to reject the phonologically related distractor as the translation of SIT than the unrelated distractor at a level approaching significance at p = 0.1 (β = -281.7, t = -1.76, p = .078). There is a significant positive interaction between phonology and imageability, indicating slower RTs for rejecting the phonologically related distractors as the imageability of items increases (β = .68, t = 2.18, p = .02). When the imageability of the target is high (e.g. KISS), participants are slower to reject the phonological

---

2 Windows do not represent equal number of participants; participant pool was larger at the novice end of the scale.
distractor (fix) than the unrelated distractor (emergency), but when the imageability of the target is low (CARRY) there is no difference for the phonologically related distractor (convince) than the unrelated distractor (rose). See Figure 3(b). There is no interaction between the phonologically related distractor and SRT-score and no three-way interaction between phonology, SRT-score and imageability.

How fast are participants to reject semantically related distractors? For example, after being presented with the target sign (SIT) how long does it take to reject the semantically related distractor (desk) in comparison to the unrelated distractor (yellow)? Participants were significantly slower to reject the semantic distractor than the unrelated distractor (β = 583.3, t = 3.63, p = .000). There is a negative interaction between semantics and imageability, indicating faster RTs when rejecting the semantically related distractors as the imageability of the target increases (β = -.91, t = -2.91, p = .003). When the imageability of the target is low (HIDE), participants are slower to reject the semantic distractor (disappear) than the unrelated distractor (ugly), but when the imageability of the target is high (PAY) there is no difference between the semantically related distractor (bank) and the unrelated distractor (destroy). See Figure 3(b). There is no interaction between semantics and SRT-score and no three-way interaction between semantics, SRT-score and imageability. In sum, less imageable ratings resulted in more semantic interference, but more imageable ratings resulted in more phonological interference.

How fast are participants to reject iconically related distractors? For example after being presented with the target word (SIT) how long does it take to reject the iconically related distractor (couch) in comparison to the unrelated distractor (yellow)? Participants are significantly slower to reject the iconic distractor than the unrelated distractor (β = 425.5, t = 2.45, p = .014). There is a negative interaction between iconicity and SRT-score, indicating
faster RTs when rejecting the iconic distractor as proficiency in ASL increases at a level approaching significance at \( p = 0.05 \) (\( \beta = -78.95, t = -1.85, p = .06 \)). Participants with lower SRT-scores show greater inhibition when rejecting the iconically related distractors than the unrelated distractors, but participants with higher SRT-scores show less inhibition, see Figure 3(a). The interaction of iconicity and SRT-score is modulated by a three-way interaction between iconicity, SRT-score and imageability, reaching significance at the \( p = 0.1 \) level, indicating slower RTs when rejecting items with high imageability ratings for participants at the upper end of the SRT-score scale compared to participants at the lower end of the scale (\( \beta = .14, t = 1.74, p = .08 \)). Participants with low ASL proficiency are slower to reject the iconic distractor than the unrelated distractor when the imageability of the target is low (PREACH/feedback), but when the imageability of the target is high (DIG/clean) they show no difference. Participants with high ASL proficiency show the opposite pattern – they are slower to reject the iconic distractor than the unrelated distractor when the imageability of the target is high, but when the imageability of the target is low they show no difference. See Figure 4.

An omnibus accuracy model with identical fixed effects to the RT model was also constructed using subjects and items as random effects. Only the experimental variables of semantic relatedness (\( \beta = 6.93 \times 10^{-2}, t = -5.43, p < .005 \)) and iconic relatedness (\( \beta = 1.40 \times 10^{-1}, t = -10.97, p < .005 \)), but not phonological relatedness, reached significance (Figure 5). This indicates the alignment of the accuracy measure with the main effects of the RT model, on a less sensitive scale. No other fixed effects significantly predicted variation. Critically, the lack of an interaction with SRT-score indicates there were no speed-accuracy trade-offs (Reed, 1973) for the novice participants as is sometimes found for less proficient bilinguals.
Figure 5 – (a) Model estimates and (b) main effect of condition on accuracy.

Discussion
This study predicted that language external motivations would be shared regardless of language experience, but that the effect of language internal motivation would change as proficiency increases. Language external motivation was predicted to increase inhibition when participants rejected phonological competitors, but decrease inhibition when they rejected semantic competitors, matching reported effects of transparency for signers. This study also predicted an effect of language internal motivation, such that novices and experts would demonstrate different patterns of activation, with a number of different possible outcomes given the presence or absence of an interaction with imageability.

Predictions that the effects of language external motivation (as measured by imageability) on language processing would be shared across proficiency measurements were confirmed by the results of the model. Semantic distractors were processed slower than unrelated distractors when imageability was low and processed faster as imageability increased. Phonological distractors
were processed slower than unrelated distractors when imageability was high and processed faster as imageability decreased. This pattern was the same for both novices and experts.

The predictions that language internal motivations are processed differently than language external motivations as proficiency increases were also confirmed by the results of the model. Specifically, the language internal motivation effect (as measured by iconically related distractors), was significantly modulated by the language external effect (as measured by the imageability scale) such that experts experienced inhibition when imageability was high, but novices experienced inhibition when imageability was low. When motivation was language internal, experts experienced increasing inhibition as imageability increased, but novices experienced decreasing inhibition. These findings appear to indicate that experts are primarily processing diagrammatic iconicity as phonological competitors since the pattern matches that of phonological inhibition increasing with higher imageability. Novices, on the other hand, are primarily processing diagrammatic iconicity as semantic competitors since the pattern matches that of semantic inhibition decreasing with higher imageability. These findings have important implications for studies of signed language processing in that they challenge core assumptions about the nature and behavior of iconic representations in the lexicon. Specifically, assumptions that conflate the effects of transparency with the effects of imageability, and that largely ignore language internal motivation, are challenged by the results of this study.

**Bimodal Lexical Access**

Effective lexical processing requires extensive language experience. A second language learner can only make sense of speech (or sign) in their L2 through repeated recognition and categorization of linguistic input. Models of unimodal bilingual lexical access predict that novice bilinguals are required to spend the lion’s share of their cognitive processing in perception and
recognition of novel language signals. This results in far less time and effort being allotted to higher level processes like meaning interpretation, pragmatic inference and a speaker’s own message construction (i.e. language production). A result of this competition for cognitive resources is that until proficiency in signal recognition increases, a short-cut to meaning is often employed by learners where L2 entries are linked directly to L1 entries and from there to conceptual memory in novice bilinguals’ lexicons. The evidence for larger cognate effects for learners is consistent with this idea (Costa, Caramazza, & Sebastian-Galles, 2000; Costa, Santesteban, & Caño, 2005). Over time, L2 word forms are linked more strongly to conceptual representations and lexical mediation is no longer needed. Or at least, there is evidence that this is the case for “unimodal” bilinguals who know two spoken languages (de Groot, 1992; Sunderman & Kroll, 2006; Talamas et al., 1999).

This study does not replicate these findings for bimodal English-ASL bilinguals. Imageability affects speed and accuracy of processing regardless of proficiency level. The data indicate a trend for the most novice bimodal bilinguals (SRT = 0) in the sample to pattern most like expert unimodals. For novices, the interference (the difference compared to the unrelated condition) for semantic distractors was greater ($M = 57.49$ ms) than for the phonological distractor ($M = 6.11$ ms). Comparatively, for the highest ranked expert bimodal bilinguals (SRT > 6) in the sample, semantic ($M = -13.48$ ms) and phonological ($M = -16.92$ ms) distractors both impacted performance. What other factors might be influencing experts and novices in a signed language so that they process translations so differently than unimodal bilinguals?

**Iconicity Effects**
While experts and novices do not show a significant difference between only phonological or only semantic distractors, there is an effect approaching significance at the $p = .05$ level between
groups for iconically related distractors (that are competitors in both form and meaning). The interference between unrelated and iconically related conditions for the lowest ranked novices ($M = 92.75$ ms) was greater than for the highest ranked experts ($M = 62.93$ ms). This study provides evidence that, because of signed languages’ uniqueness in how language forms are connected to their mental representations via the manual/visual modality, the way that conceptual or lexical mediation is used contrasts with how spoken language bilinguals use these mechanisms. The iconicity of signs in ASL allows bimodal bilinguals to directly access meaning from form without lexical mediation of the first language (Morford, Occhino-Kehoe, PiñAr, Wilkinson, & Kroll, 2015). Bimodal bilinguals are distracted by iconically related translation competitors – unlike unimodal bilinguals who do not have this privileged thoroughfare between phonological and semantic representations – and show a different pattern of lexical or conceptual mediation between translation equivalents than is standard for spoken language bilinguals.

In the cognitive iconicity perspective, iconic mappings are learned and built up over iterative exposure, like all language forms. In usage-based models, bilinguals who have more experience with a signed language are more sensitive to these kinds of conceptual representations than those with more limited exposure. While a novice may recognize that the sign for BANANA in ASL profiles the act of peeling a banana, it is less likely they will recognize that the sign for DANCE profiles actions of legs. The former example represents language external knowledge, and the latter represents language internal knowledge. A proficient signer of ASL recognizes that the extended 1st and 2nd fingers (“V” handshape) are part of a diagrammatic pattern in the language where many signs recruit the same phonological characteristics for related meanings (e.g. the signs STAND, DANCE, JUMP, FALL-DOWN, and RIDE all use this pattern, see Figure 6). Another example of a sign that is more diagrammatic
and less imagistic is the sign for CHASE which profiles a schematic concept of “movement of one object in relation to another”. Following Lepic’s (2015, p. 171) treatment of morphological nuclear families in ASL, a proficient signer of ASL recognizes that the dual fists with extended thumbs (“A” handshape) produced in neutral space (in front of the body) with movement left unspecified, encodes a “moveable objects” family which includes signs that are morphologically related in ASL, but not in English (e.g. the signs FOLLOW, CHASE, AVOID, and PASS). Moreover, other signs that do not exactly fit into the family are still related, though not as closely, in both form and in meaning (e.g. CHALLENGE, EQUAL, and MEET).

Results of this study indicate that sensitivity to a signed language’s internal patterns is modulated in important ways by the relative imageability (or perhaps prototypicality) of the mental representation of a meaning and its connections to mental representations of form.

**Imageability Effects**

Imageability is the measurement of how easy or difficult it is to visually or acoustically imagine a concept. All imageability ratings were collected on the English tokens of the “correct” translation since there is currently no database that has measured the imageability of the ASL lexicon. While there is no doubt some variation in this parameter for any given lexical item cross-linguistically, this study assumes that English imageability ratings capture an effect that is rooted in the conceptual system.

There was a main effect of imageability where the more imageable a concept was, the faster the RT in comparison with the average, approaching significance at the p = .05 level (β = -0.14, t = -1.69, p = .093), suggesting that explicit manipulation of this characteristic in a future study would result in robust effects. This result matches imageability findings for spoken

---

3 Fischer (2015, p. 454) notes that signs like these are lexicalized/grammaticalized from the old LSF sign ONE.
There were two interactions between imageability and the experimental condition. First, as the imageability of items in the phonologically related distractor condition increased, so did the average RT. Second, as the imageability of items in the semantically related distractor condition increased, the average RT decreased. In other words, there was a trend for participants to activate less competition at the conceptual level for highly imageable concepts when presented with an incorrect, but semantically related translation. For instance, participants were faster to reject semantically related translation distractors for the highly imageable concept “pay”, as in the semantically related pair PAY/bank than the less imageable concept “hide” as in the semantically related pair HIDE/disappear. The competitors in the domain activated by PAY made rejecting bank as a possible translation equivalent relatively easier than the competitors in the domain activated by HIDE made rejecting disappear as a correct translation.

There was also a trend for participants to need to activate less competition at the lexical level for highly imageable concepts when presented with an incorrect, but phonologically related translation. At the highest end of the imageability scale there was inhibition for the phonologically related distractors compared to unrelated distractors. For instance, participants were slower to reject phonologically related translation foils for the highly imageable concept “fix” (as in the phonologically related pair FIX/kiss) than the competitors in the domain activated by the less imageable concept “convince” (as in the phonologically related pair CARRY/convince). The competitors in the domain activated by FIX made rejecting kiss as a possible translation equivalent relatively easier than the competitors in the domain activated by CARRY made rejecting convince as a correct translation.
These results match the prediction that the visual nature of signed language causes effects of imageability to manifest differently across modalities and are consistent with effects of iconicity when operationally defined as transparency and measured via non-signer ratings of iconicity. In particular, Thompson et al., (2009, 2010) found the same pattern of semantic facilitation and phonological inhibition for highly transparent signs, similar to the effects demonstrated here for highly imageable signs.

**Theoretical Considerations**
How can we account for the effects of imageability on phonological and semantic access?

Studies that have examined iconicity in signed languages have found that alignment between sign form and referent speeds up semantic processing (Thompson et al., 2009), but slows down phonological processing (Thompson et al., 2010). By conflating iconicity with transparency these studies pick only the low hanging fruit; iconicity can be transparent, diagrammatic, or a combination of both. While the motivation of some iconic forms is accessible (transparent) even to non-signers and novice ASL learners, other kinds of (diagrammatic) iconicity are accessed differently for native and expert second language users. In this study, imageability has the same effects as transparency measurements in these other studies, in that greater imageability
decreases time required to reject competitors when the task recruits conceptual memory, but also increases time required to reject competitors when the task requires lexical memory.

Thompson, et al., (2010) did control for imageability, but did not find it to be a significant predictor of variability. That study included both nouns and verbs, so it is possible that imageability may play a more significant role for verbs than for nouns. Concreteness is typically considered closely related to imageability, but is slightly different in important ways and may be a better measure of language external motivation for nouns. Concreteness ratings are obtained by asking participants to rate the extent to which a word’s referent can be touched or felt (Reilly & Kean, 2007). Concreteness did not reach significance in the model for this study during omnibus testing, possibly because stimuli included only verbs. This design makes the results of this study different than many of the previous unimodal studies. A data set with both nouns and verbs might show effects of concreteness in addition to imageability. Further examination of the interaction between imageability, concreteness, part of speech and iconicity is needed. There is some evidence that object and process concepts interact with iconicity in different ways. Ortega and Morgan (2015) found that non-signers in a lexical priming task had their lexical decision times facilitated by both iconic signs that profile actions (such as TAKE-PICTURE-WITH-CAMERA) and objects (such as BOX), but that experienced L2-signers had their lexical decision times facilitated only by action iconicity signs. This study focused on verbs in order to facilitate elicitation of an iconicity effect from both novice and expert sign language users.

Another possible explanation for the interaction of imageability with ASL lexical access is that concepts that are more imageable are simply more likely to be iconic, indicating that the language external force of the degree to which a concept is amenable to visual representation has
a novel effect on visual-language processing compared to spoken language processing. The effects of systematic intra-language relations where groups of signs that share phonological overlap also share semantic overlap was significant for both experts and novices, indicating that both the lexical and conceptual levels are activated by these cohorts of shared form and meaning. This is a novel finding for the study of signed language processing.

Even more novel, is the observed interaction between language experience, language external imageability, and language internal systematicity. This finding has broad applications for investigations of signed language processing because it disambiguates two sometimes confounded factors at play in the comprehension of signed language lexical representations. If imageability plays the same role for signed languages as it does for spoken languages, both bimodal novices and experts should have experienced increased semantic interference for targets with high imageability. Instead, both groups showed increased phonological interference for high imageability items. A possible explanation of this pattern is that imageability affects the representation of motivation in the lexicon in a way that causes form/meaning mappings with high imageability to act as phono-semantic anchors for dense networks of lexical items that are parasitic on a single prototypical codification of an image schema (Figure 6). Activation of items within dense networks, then, would primarily inhibit phonological access because there are many form competitors vying for a single conceptual representation. This effect would grow stronger as more lexical items are connected to the network.
Figure 6 – Patterns of activation in high imageability iconic networks: (a) experts & (b) novices.

Signs with low imageability are more likely to be members of sparse networks with less rigid mapping of conceptual representation to phonological realization (though they still maintain their motivation). Activation of items within sparse networks would primarily facilitate semantic access because specific conceptual representations would preferentially activate fewer form competitors. Additionally, novices and experts would differ in the degree to which they have mapped specific signs to form/meaning networks, affecting whether they primarily experienced phonological inhibition or semantic facilitation. The greater the number of signs mapped and the density of the network to which they are mapped would modulate the inhibition signers experience when accessing forms (Figure 7).
The models proposed here have much in common with some previous discussions of concreteness and abstractness in conceptual processing. In particular, the Context Availability Model (Kieras, 1978; Schwanenflugel & Shoben, 1983) proposed that related conceptual information is only weakly associated with abstract concepts, but for concrete concepts, related information is strongly associated. Activation of concrete conceptual information is more likely to retrieve networks that act together. This idea has also been vetted with computer modeling of experimental results (McRae, de Sa, & Seidenberg, 1997) that found biological exemplars (e.g. birds, mammals, fruits, vegetables, etc) are more likely to pull up networks that act in unison than non-biological exemplars (clothing, furniture, tools, vehicles, weapons, etc). These results, and the results of this study support the notion that the conceptual system activates abstract and concrete information in unique ways that can impact language processing.

**Conclusion**
In summary, this study’s results are consistent with the hypothesis that motivated signs activate links between form and meaning. This pattern of lexical access shows that bimodal bilinguals access translations differently than unimodal bilinguals. Rather than showing increased semantic
interference when translating high-imageability concepts (de Groot & Poot, 1997), bimodal bilinguals show increased phonological interference. This is consistent with evidence of semantic facilitation in deaf signers when processing conceptual properties for motivated signs (Thompson et al., 2009), but phonological inhibition when processing lexical properties (Thompson et al., 2010). Results also indicate measurements of motivation that are primarily language external (such as non-signer transparency judgments and imageability measurements) are too simple to accurately capture how lexical representations in a signed language change as proficiency increases. Language internal motivations interact with language external motivations so that experts experience increased phonological activation when imageability is high, but novices experience increased semantic activation. In other words, experts process diagrammatic iconicity as phonological competitors since the pattern matches that of phonological inhibition increasing with higher imageability, but novices process diagrammatic iconicity as semantic competitors since the pattern matches that of semantic inhibition decreasing with higher imageability.

A usage-based approach to cognitive iconicity (Occhino-Kehoe, 2016; P. Wilcox, 2000; S. Wilcox, 2004) provides a framework that simultaneously explains both language external and internal patterns of motivation as proficiency increases. Concepts with high-imageability are more likely to have form/meaning mappings that are accessible to novices, but these mappings are still accessible to experts as well. Motivated signs are not independent of the entrenchment of language internal patterns of form and meaning. Rather than decreasing their impact on language processing as proficiency grows, they increase in tandem, building on the original heuristic aid given by codification of visual stimuli to conceptual storage. Eventually, motivation effects even extend to concepts with low-imageability where form/meaning mappings are more schematic
and instantiated as a collection of related form/meaning mappings, instead of primarily one-to-one mappings of form and meaning. Proficiency in a signed language involves more than knowing a collection of lexical forms and their respective meanings. It involves knowing both the meanings of lexical items and also their formational and conceptual similarities that are systematic within the networks of the language. In all, the results of this study lay bare that the behaviors of lexical and conceptual storage are not clearly separable for ASL. As Stokoe once framed this issue: “The metaphor (…) that jumps to mind is the Möbius strip: the input is the output—with a twist!” (2001, p. 439). For signed languages, categorical separation of phonological and semantic levels of processing may be disingenuous. These findings expand our understanding of language processing and will inform future explorations of representational memory in both signed and spoken languages.
Iconicity in Translation Production

Abstract

Does iconicity in the sign lexicon influence hearing bimodal bilinguals’ preference for L1-L2 (forward) translation? RT and accuracy measures are reported for a translation production study in which English-ASL bilinguals at novice and expert levels of ASL proficiency translated lexical items both forward (L1-L2) and backward (L2-L1). Results indicate that in conditions where conceptual mediation is typically used by spoken language bilinguals, greater iconicity in the ASL stimuli facilitates translation latencies. Iconicity was measured by the speed participants could name pictures that highlighted the visual features that motivated the ASL form. Novices were faster and more accurate to translate English words into more iconic ASL signs than less iconic signs. Experts were faster to translate more iconic ASL signs into English than less iconic signs. Implications for measuring and modeling form-meaning relationships in signed languages are discussed within a usage-based cognitive framework.

Keywords: bilingualism, sign language, iconicity, translation production, usage-based, cognitive linguistics
Iconicity in English-American Sign Language Translation Production

The current understanding of bimodal bilinguals – bilinguals who know a spoken and a signed language – lags behind what we know about unimodal bilinguals, particularly in regard to lexical processing. It is unclear what qualities of the visual nature of signed languages might influence translation between a bimodal bilingual’s first (L1) and second (L2) language. Recent studies have found that sign language interpreters prefer forward translation over backward translation so long as American Sign Language (ASL) is their L2 (Nicodemus & Emmorey, 2013), and that very early novices produce faster forward translations for signs whose meanings are easily guessed by non-signers (Baus et al., 2013).

One outstanding question is whether or not early bimodal bilinguals might alternate from lexical or conceptual mediation when translating between their two languages as is the case for bilinguals of two spoken languages. This topic was examined by Potter et al. (1984) who proposed two possible answers to this question, distinguishing between a word-association hypothesis and a conceptual mediation hypothesis. In both models strong links exist between the L1 lexicon and conceptual memory. In the first model a link also exists between the lexical representations of the L1 and the L2. In the second model a link exists between the L2 lexical representations and conceptual memory. To test which model was correct, participants were asked to translate words from their L1 (Cantonese) into their L2 (English) and also to name pictures in their L1. Results showed that pictures were named slightly (but not significantly) faster in the L1 than L1 words were translated into L2. Because the word association model predicts slower picture naming time than translation time, the results were interpreted as support for conceptual mediation. However, Potter et al. based their conclusions on a task that only tested L1-L2 translation (forward translation), not L2-L1 translation (backward translation).
Kroll and Stewart (1994) observed that the conceptual mediation model alone could not satisfactorily describe the observed asymmetries in translation performance that bilinguals report. Specifically, spoken language interpreters express a preference to interpret from their second language into their first (Seleskovich, 1976), suggesting they may benefit from the speed of lexical connections in producing L2 translations. Stimuli in Kroll and Stewart’s study were ordered in semantically related categorized lists (e.g. dress, suit, shoes) as a way to activate conceptual memory rather than by using picture naming. Kroll and Stewart found that semantically related items were translated into the L2 slower than when translated into the L1. Stimuli in categorized lists were also translated into the L2 slower than in mixed lists (e.g. orange, lion, ambulance). These results indicated that forward translation was conceptually mediated, but backward translation was lexically mediated and that the strength of the connections between L1 and L2 lexical items to conceptual memory was weaker for L2 lexical forms, particularly for novices. Kroll and Stewart modeled these differences in their Revised Hierarchical Model (RHM), combining both the word-association and conception mediation models’ ordered rank structure. The RHM makes two important claims: first, that backward will be faster than forward translation and second, that this effect lessens given the fluency of the bilingual (Kroll, Michael, Tokowicz, & Dufour, 2002; Poarch, Van Hell, & Kroll, 2015; Sholl, Sankaranarayanan, & Kroll, 1995). More fluent bilinguals have stronger lexical-conceptual links in their L2, and as a result, rely on L1 word association less often to comprehend L2 words.

In sum, prior research investigating lexical processing in unimodal bilinguals suggests that the relationship between form and meaning varies with language proficiency, with forms becoming more closely associated to their meanings as proficiency increases. Further, the strength of form-meaning association determines the extent to which concepts are activated
A striking difference between signed languages and spoken languages is the former’s capacity and propensity for “iconicity” – visually motivated mappings between phonological and semantic levels of representation. Models of bilingual access that overlook the potential effects of iconicity may fail to accurately model lexical access in bimodal bilinguals. While there is growing evidence that iconicity in signed language is relevant for on-line comprehension and production (Baus, Carreiras, & Emmorey, 2013; Emmorey, 2014; Grote & Linz, 2003; Ortega & Morgan, 2015; Thompson, 2011; Thompson, Vinson, & Vigliocco, 2009, 2010), there is currently debate about how best to measure it (Occhino-Kehoe, 2016; Occhino-Kehoe, Anible, Wilkinson, & Morford, forthcoming). Past studies have tended to rely on non-signer judgements of iconicity that are susceptible to confounding concreteness and/or imageability ratings with the language internal dynamics of iconicity (but see Anible, in prep). To avoid these pitfalls, picture naming latencies will be used as an alternative in the current project. It is possible iconicity could impact lexical processing by facilitating conceptual mediation and/or disrupting lexical mediation.

In a series of three studies, this investigation first establishes a novel method to operationalize iconicity as an independent measure via picture naming. Subsequently, this measure is used to evaluate effects of iconicity and language proficiency on forward and backward translation in hearing bimodal bilinguals. The predictions are that iconicity should affect the speed and accuracy with which bimodal bilinguals translate lexical items. In forward, conceptually mediated (L1-L2) translation, increasing iconicity should primarily facilitate translation time and increase accuracy because L2 forms that are closely related to the concept should benefit from increased conceptual activation. In backward, lexically mediated (L2-L1)
translation, increasing iconicity should inhibit translation time and decrease accuracy because faster, more accurate lexical mediation is prevented by conceptual mediation that cannot be suppressed. Moreover, these iconicity effects should interact with second language proficiency. Anible (in prep) found that language-internal iconicity leads to the activation of phonologically related lexical cohorts in expert L2 signers but semantically related lexical cohorts in novice signers using a translation recognition paradigm. Hence, a logical prediction would be that effects of iconicity for novices in forward, generally conceptually mediated translation, should be greater than in backward translation because it uses lexical mediation. For experts, effects of iconicity should also exist at the lexical level.

**Study 1: Picture Naming**
A first step in testing these predictions is a cognitively grounded measurement of iconicity. In keeping with a perspective that aligns iconicity in sign language with the growing literature on simulation (Barsalou, 2008; Vukovic & Williams, 2014), iconicity values were assessed using a picture naming task where signers’ reaction times (RTs) were compared when naming pictures either profiling or not profiling the perceptual features selected in the articulation of the ASL stimulus verbs (e.g. HIDE features an “agent” encoded by the dominant hand, under a “surface” – encoded by the non-dominant hand, see Figure 8). Profiling is best understood as the mental process of selecting a subset of a perceptual stimulus as a focus of attention (Langacker, 2008, p. 66). The mapping of articulatory features to conceptual schemas is a documented property of signed language lexicons (Taub, 2001; P. Wilcox, 2000; S. Wilcox, 2004). Participants viewed pictures of the same concept in two conditions; (a) a picture that profiled the iconic properties of a concept matching those profiled by the ASL sign (e.g. hiding under a desk), (b) a picture that profiled a different quality of the concept (e.g. hiding behind a tree). Past studies have found that signers are faster to name pictures in a profiled condition than in a non-profiled condition (Grote
& Linz, 2003; Thompson et al., 2009). The average RT for each item in the profiled condition for the participants was subsequently used as a measure of the relative iconicity of the stimuli that were translated in the production studies.

**Participants**
A control group of fifty non-signing undergraduates (41 female, 9 male) enrolled in Introductory Linguistics courses at the University of New Mexico were asked to complete the picture naming task in English. The non-signer participants were native speakers of English 18 to 46 years old ($M = 21.4$, $SD = 4.8$). Accuracy of responses ($M = .79$, $SD = .40$) was calculated after data collection by independent coders. One participant was excluded from the analysis due to low accuracy ($M = .41$).

Fifty-seven English-ASL bilinguals (44 female, 13 male) also participated in the experiment. Data from two of these participants were excluded due to equipment failure. Participants in both groups were compensated either monetarily or with course credit.

Participants were native speakers of English 18 to 58 years old ($M = 24.6$, $SD = 8.9$) who started learning ASL between 5 and 26 years old ($M = 16.1$, $SD = 4.1$). Participants’ experience with ASL ranged from three semesters of classroom instruction to experienced sign language interpreters ($M = 10.3$ years of experience, $SD = 8.6$). ASL proficiency was measured using the ASL Sentence Repetition Task (SRT) (Hauser et al., 2008), a test developed to assess levels of proficiency for native users of ASL. Participants viewed 20 sentences of increasing complexity and were asked to repeat each sentence back exactly after viewing. Exact repetitions were scored as 1; repetitions with any replacements or errors were scored as 0. A native deaf signer evaluated the participants’ responses. Participants’ SRT-scores ranged from 0 to 10 ($M = 2.98$, $SD = 2.56$). Often, current ASL proficiency assessments rely on participants self-rating their production and
comprehension skills (c.f. B. Anible et al., 2015). Although low average scores indicate that the test is not an ideal measure for second language learners, the SRT measure should still be considered superior to self-ratings.

**Procedure**

Participants were presented with 48 line drawings that were matched with the ASL verbs in the translation production tasks and were asked to name the picture in English (control group) or ASL (experimental group) as quickly as possible. All participants were tested individually and were given English instructions by the experimenter. A practice block of five pictures was given prior to the start of the experiment. Participants were presented with the randomized pictures one at a time at the center of the computer screen. A fixation point preceded each picture which then was replaced by a line drawing depicting an action. Signing participants were instructed to place their hands on the spacebar from the start of each trial until they started their response in ASL. Non-signing participants were instructed to start speaking as soon as they were ready after viewing the picture. The participants’ responses were video-recorded and later coded for accuracy and, in the case of the signing participants, reaction time. Accuracy was coded by hand. RT was measured using the computer vision algorithm *optical flow* (see explanation, below).

![Image](image.png)

**Figure 8** – Sign for “hide” in profiled vs. non-profiled image conditions.
Items were matched as closely as possible for word length, frequency, concreteness and imageability. These measures have all been shown to affect lexical access in online processing (Hudson & Bergman, 1985; Jescheniak & Levelt, 1994; Poarch et al., 2015). Scores were taken from the MRC Psycholinguistic Database (Coltheart, 1981). In cases where an item was not available in this database, the average of the condition was used. For both concreteness and imageability scores, these cases accounted for 4% (2/48) of the control values used.

<table>
<thead>
<tr>
<th>Length</th>
<th>Log frequency</th>
<th>Concreteness</th>
<th>Imageability</th>
<th># of Trans. Equiv.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>4.75</td>
<td>1.31</td>
<td>9.72</td>
<td>1.45</td>
<td>452.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>499.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.83</td>
</tr>
</tbody>
</table>

Table 4 – Control measures for experimental stimuli.

After completing the picture naming task, participants were asked to rate, in English, the number of translation equivalents of the verbs the pictures referenced. This was done in preparation for using picture naming RTs as iconicity measurements in the analysis of results from Studies 2 and 3. For example the English word “run” has multiple lexical items in ASL that express the concepts of “jogging”, “machinery functioning”, “liquid flowing,” “a run in fabric”, etc. Some studies have shown the number of translation equivalents can slow lexical access in tasks involving translation (Baus et al., 2013; Laxén & Lavaur, 2010; Tokowicz & Kroll, 2007). Four items were removed because the average number of translation equivalents was more than 2.5 standard deviations above the mean. Another four items were removed because naming accuracy was extremely low.

---

4 A dictionary of 150837 words with up to 26 linguistic and psycholinguistic attributes for each entry.
Results

Sign RTs were measured using computer vision to track optical flow (Lucas & Kanade, 1981). Optical flow is the pattern of motion of objects between consecutive frames caused by the movement of an object or camera. Sign onset was measured as the point at which the average velocity of movement of all tracked points in a video was greater than the mean plus the standard deviation. The time-stamp at this frame was used as the reaction time value for naming the picture.

To assess the accuracy of the optical flow algorithm, sign onset for approximately one item from each participant ($N = 60$) was coded by hand. Sign onset was measured as the first frame both hands were visible with the articulated handshape(s) of the target sign. RTs that were faster than 200 ms or slower than 10000ms were removed from the analysis as outliers. Since this task used production data in a visual modality – and sign production takes longer on average for lexical items than vocal production (Bellugi & Fischer, 1972) – a longer cut off than the standard 5000ms (Baayen & Milin, 2010) was selected to capture the full range of processing times participants exhibited. Five observations (8%) of the data was trimmed in this way. The optical flow onset times were more highly correlated with the onset times coded by-hand ($r = .93, p < .001$) than were the spacebar onset times ($r = .75, p = n.s.$). Since optical flow proved to be a better measure of actual participant behavior than the RTs from the spacebar release, the RTs returned by optical flow are used for this experiment. Spacebar release RTs were considerably smaller than hand-coded onset times, indicating that participants released the spacebar well before they started signing in a large number of trials.
Figure 9 – Correlation of: (a) optical flow, and (b) spacebar with by-hand measures.

Accuracy
Accurate responses were those in which participants either spoke or signed the exact target verb after viewing the picture. Responses in which a different or related lexical item was produced were considered inaccurate. To assess whether any difference between condition accuracy was attributable to qualities of the images, rather than to language experience, accuracy between groups and conditions was measured using mixed-effect modeling with crossed random effects for participants and items (Baayen, 2008; Baayen et al., 2008) and were estimated using the lme4 package (Bates et al., 2013). P-values were calculated using the lmerTest package (Kuznetsova et al., 2013).

First, an omnibus model for accuracy was estimated. Fixed effects included linear effects and interaction between the experimental variable (profiled vs. non-profiled), participant variables (age, gender, group) and stimulus variables (word length, frequency, concreteness,
imageability and number of translation equivalents). Random effects were random intercepts for participants and items. The model was trimmed to remove nonsignificant fixed effects. Word length, frequency, concreteness, and number of translation equivalents had no significant effects or interactions. There was a main effect of group ($\beta = -0.05, t = -1.92, p = .057$), indicating that signers were less accurate on average to name the pictures than non-signers. There was a main effect of imageability ($\beta = 0.00, t = 2.44, p = .020$), indicating that as the imageability of a concept increased the accuracy of naming the concept, regardless of condition, increased. Critically, there was an interaction of the condition and group ($\beta = 0.06, t = 2.55, p = .011$), indicating that signers were significantly less accurate to name non-profiled pictures than profiled pictures compared to non-signers (see Figure 10).

<table>
<thead>
<tr>
<th>Fixed Effects and Interactions</th>
<th>$\beta$</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition = Non-profiled</td>
<td>0.03</td>
<td>1.58</td>
<td>.113</td>
</tr>
<tr>
<td>Group = Signers</td>
<td>-0.05</td>
<td>-1.92</td>
<td>.057</td>
</tr>
<tr>
<td>Imageability</td>
<td>0.00</td>
<td>2.44</td>
<td>.020</td>
</tr>
<tr>
<td>Condition = Non-profiled $\times$ Group = Signers</td>
<td>0.06</td>
<td>2.55</td>
<td>.011</td>
</tr>
</tbody>
</table>

Table 5 – Main effects and interactions: Picture naming accuracy.
These results indicate that the experimental manipulation (profiled vs. non-profiled images) did not significantly affect accuracy for non-signing participants; there was no difference in non-signing participant accuracy in naming the pictures profiling the iconicity of ASL signs compared to accuracy in naming the pictures that did not profile signs’ iconicity. Any differences in condition for signers can be attributed to the influence of ASL knowledge, rather than a visual quality of the pictures or challenges in representing concepts as images. In particular, signers are more accurate for profiled pictures than non-signers so the effect of accuracy is another indication that the experimental manipulation is impacting signers, but not non-signers. One unexpected finding in this analysis was that non-signer participants had higher average accuracy than signing participants. Presumably, alternating between profiled and non-profiled pictures resulted in some amount of residual inhibition for signing participants, possibly stemming from a
perpetuation effect of relying on form/meaning connections to guide lexical decisions for the profiled pictures, but not for the non-profiled pictures.

**Reaction Time**

An omnibus model for RT was estimated, again using mixed-effect modeling by condition with crossed random effects for participants and items. RT analyses were performed only on the experimental group (signers), not the control group (non-signers). Fixed effects included linear effects and interaction between the experimental variable (profiled vs. non-profiled), participant variables (age, SRT-score) and stimulus variables (word length, frequency, concreteness, imageability and number of translation equivalents). Random effects were random intercepts for participants and items. The model was trimmed to remove nonsignificant fixed effects. Word length, frequency, and concreteness had no significant effects or interactions, but there were main effects for imageability and number of translation equivalents. No effects were found for SRT-score or age.

The model was subjected to criticism focusing on the residuals (Baayen & Milin, 2010). Data points ($n = 55; 3\%$ of the data) with absolute standardized residuals exceeding 2.5 standard deviations were removed from the data set and the model was re-estimated, resulting in improved fit. Correlation of fitted values with trimmed RTs was $R^2 = .51$, up from $R^2 = .41$; BIC = 26652.91, down from 28159.28. There was a significant difference in the speed of naming the non-profiled vs. profiled images ($\beta = -390.19, t = -6.84, p < .001$), with profiled pictures being named faster on average than non-profiled pictures. Imageability and number of translation equivalents additionally predicted a significant portion of the variability; increasing imageability globally decreased naming times ($\beta = -2.58, t = -2.75, p = .009$) and increasing average number of translation equivalents globally increased naming times ($\beta = 415, t = 2.29, p = .028$), see
Figure 11.

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>β</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition = Profiled</td>
<td>-391.15</td>
<td>-6.66</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Imageability</td>
<td>-2.58</td>
<td>-2.75</td>
<td>.009</td>
</tr>
<tr>
<td>Number of Translation Equivalents</td>
<td>415.18</td>
<td>2.29</td>
<td>.028</td>
</tr>
</tbody>
</table>

Table 6 – Main effects: Picture naming RT.

Figure 11 – (a) Main effects of picture type RT, (b) imageability, and (c) number of translation equivalents.

**Discussion**

Past studies have relied on offline measures of iconicity using either signer or non-signer iconicity judgements to operationalize iconicity. In the current study, by contrast, an online measure of iconicity was developed by eliciting naming times for the target stimuli using images that did or did not profile the conceptual basis for the signs’ articulatory form. This approach to
operationalizing iconicity acknowledges that iconicity is not a objective characteristic of a sign, but is instead a mapping of an individual’s construal of form and construal of meaning, mediated by that individual’s own language experience. As the naming speed of a profiled image decreases, this indicates that the iconicity of the item is higher on average for participants in this study. This measurement is used to explore the effect of iconicity on translation production in the following experiments.

Evidence for the psychological reality of separate storage of word-forms from their meanings comes from three types of psycholinguistic experiments measuring naming time, recall and rapid serial presentation. Participants in naming time experiments are able to name words much faster than they are able to name pictures. In recall experiments, when subjects are asked to remember a number of unrelated items, pictures are remembered better than words. When pictures or words are presented very quickly, words are remembered better than pictures. All three findings argue for separate lexical and imagistic (conceptual) memory storage. Words are named faster than pictures because participants do not have to access conceptual memory to produce the spoken form associated with print. Random pictures are remembered better than random words because they necessitate access to conceptual memory (which persists longer) whereas words need not. When random pictures and words are presented fast enough that semantic memory cannot be accessed, words are remembered because their spoken forms can be maintained in the short term phonological-loop long enough to be recited (Potter, 1979).

More recent literature from embodied cognition studies suggests a different loci of processing for imagistic memory and indicates there may be access to all associated information. For example, Richardson et al. (2003) found readers simulate horizontal and vertical paths that are intrinsic to concrete and abstract verbs (pull vs. drop, agree vs. disrespect). Matlock (2004)
found that implied fictive motion (the river runs through downtown) creates analogous motion simulation in space and Richardson & Matlock (2007) found these simulations produce corresponding eye movements. Schubert (2005) found that reading words that are connected to high or low power focuses participants’ attention upwards or downwards. While studies are not necessarily in conflict with prior explorations of imagistic meaning they frame the issue new ways that elucidate the underlying cognitive processes governing processing effects.

Because picture naming is such a reliable way to access conceptual memory, the RTs in this experiment for the profiled condition are likely to be a good measure of the degree to which participants were affected by the conceptual representation of the iconicity profiled in the articulation of a sign. Importantly, since only signers showed a difference in accuracy between profiled and non-profiled images, differences in naming speed between the two conditions are unlikely to be attributable to qualities of the images themselves such as image complexity or prototypicality.

This newly developed measure of iconicity allows for a novel method of investigating effects of iconicity and avoids some pitfalls that previous studies have had relating to non-signer judgements of iconicity that are susceptible to confounding concreteness and/or imageability ratings with the language internal dynamics of iconicity. Prior studies of translation production in bimodal bilinguals have failed to find effects of iconicity, particularly for highly proficient bilinguals. This new cognitively motivated measure of iconicity may be more sensitive to gradual changes in the effects of iconicity on translation production as proficiency increases in bimodal bilinguals.

**Study 2: Forward Translation**
Two translation production tasks, forward and backward, followed each other sequentially for
each participant. For both translation tasks, participants were seated at a Toshiba Satellite L755 laptop with their hands on the spacebar. Participants were required to depress the spacebar to initiate each trial. During forward translation trials, they then heard a spoken English word over headphones and produced the ASL translation by raising their hand from the spacebar. During backward translation trials, participants watched a video of a sign and vocally produced the English translation. Order of translation direction was counterbalanced across participants. Accuracy was coded by hand for both forward and backward translations. Voice RTs in backward translation were measured by a voice onset key from the beginning of the ASL stimulus video. Sign RTs in forward translation were measured with optical flow. Participants (signers only) and materials were the same as in Study 1. Participants completed Study 2 and Study 3 before picture naming, so training effects from repetition of stimuli were present only in Study 1.

Results were analyzed using mixed-effect models with crossed random effects for participants and items. Iconicity values were imported from the picture naming task. In that task, lower RT values equated with higher iconicity (more naming facilitation) and higher RT values with lower iconicity (less naming facilitation). These RTs were subsequently transformed for ease of interpretation by adding the absolute of the highest RT to the inverse of all values such that an increasing score denotes increasing iconicity. This sets the iconicity scale used at approximately 0 – 2000.

**Accuracy**

Mean accuracy was .89; accuracy per item ranged between .24 and 1.00 (SD = .22); accuracy per subject ranged between .64 and 1.00 (SD = .28). An omnibus model for accuracy was

---

5 Removing items with accuracy below .5 did not change the overall pattern of results.
estimated. Fixed effects included linear effects and interactions between the experimental variable (iconicity), participant variables (age and SRT-score) and stimulus variables (word length, frequency, concreteness, imageability, and number of translation equivalents). Random effects were random intercepts for participants and items. The model was trimmed to remove nonsignificant fixed effects. SRT-score was the only factor in either participant or stimulus variables to reach significance. The final model included the modulation of the effects of the experimental variable (iconicity) by the language experience of the participants (SRT-score) and a random structure including intercepts for participants (subject) and stimuli (item). The model (in R notation) was specified as (Accuracy ~ Iconicity * SRT + (1|item) + (1|subject). Table 7 shows the results of fitting this model for the experimental variables and interactions. To help interpret these results, Figure 12 and Figure 13 show the corresponding main effects and interaction plots, created with the *effects* package (Fox, 2003) and the *ggplot2* package (Wickham, 2009).

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed Effects and Interactions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iconicity</td>
<td>0.00</td>
<td>2.56</td>
<td>.013</td>
</tr>
<tr>
<td>SRT-score</td>
<td>0.11</td>
<td>5.39</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Iconicity × SRT-score</td>
<td>-0.00</td>
<td>-4.34</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Table 7 – Main effects and interactions: Forward accuracy.
Figure 12 – Main effects of (a) iconicity, and (b) SRT-score on forward accuracy.

Figure 13 – Interaction of iconicity and SRT-score on forward accuracy.\(^6\)

\(^6\) Windows do not represent equal number of participants; participant pool was larger at the novice end of the scale.
How accurate are participants when translating from their L1 (English) into their L2 (ASL)? For example after hearing an English word (problem) how accurate are they to translate into an ASL equivalent (PROBLEM)? Participants have higher accuracy when translating from L1 into L2 as iconicity increases ($\beta = 0.00, t = 2.56, p = .013$). Participants have higher accuracy translating L1-L2 as SRT-score increases ($\beta = 0.11, t = 5.39, p < .001$). However, a significant interaction between iconicity and SRT-score ($\beta = -0.00, t = -4.34, p < .001$) indicates that iconicity increases forward translation accuracy for less proficient signers, but decreases accuracy for more proficient signers. Novice participants were more accurate to translate the English word drink to the high iconicity sign DRINK than the low iconicity measure to MEASURE, but proficient signers show the reverse pattern.

**Reaction Time**
An omnibus model for RT was estimated next, using only correct trials. Fixed effects included linear effects and interactions between experimental variable, participant variables (age and SRT-score) and stimulus variables (word length, frequency, concreteness, imageability, and number of translation equivalents). Random effects were random intercepts for participants and items. Models were trimmed to remove nonsignificant fixed effects. Word length, concreteness, imageability, and number of translation equivalents had no significant effects or interactions, but there were main effects for frequency. An interaction for the subject variable SRT-score was found, but no effect was found for age. The model was subjected to criticism focusing on the residuals. Data points ($n = 27; 3\%$ of the data) with absolute standardized residuals exceeding 2.5 standard deviations were removed from the data set and the model was re-estimated, resulting in improved fit. Correlation of fitted values with trimmed RTs was $R^2 = .40$, up from $R^2 = .30$; BIC = 15426.93, down from 16222.75. All subsequent analyses employ the trimmed data set. The
final model included the modulation of the effects of the experimental variable (iconicity) by the language experience of the participants (SRT-score) and fixed effect of frequency (log frequency) and a random structure including intercepts for participants (subject) and stimuli (item). The model (in R notation) was specified as (RT ~ Iconicity * SRT + Log Frequency + (1|item) + (1|subject). Table 8 shows the results of fitting this model for the experimental variables and interactions. To help interpret these results, Figure 14 shows main effects of (a) iconicity, (b) SRT-score, and (c) frequency on RT. Figure 15 shows the interaction of iconicity and SRT-score on RT.

<table>
<thead>
<tr>
<th>Fixed Effects and Interactions</th>
<th>β</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iconicity</td>
<td>-0.42</td>
<td>-3.24</td>
<td>.002</td>
</tr>
<tr>
<td>SRT-score</td>
<td>-158.80</td>
<td>-3.03</td>
<td>.003</td>
</tr>
<tr>
<td>Log Frequency</td>
<td>-93.35</td>
<td>-2.66</td>
<td>.012</td>
</tr>
<tr>
<td>Iconicity × SRT-score</td>
<td>0.05</td>
<td>1.89</td>
<td>.059</td>
</tr>
</tbody>
</table>

Table 8 – Main effects and interactions: Forward RT.

How fast are participants when translating from their L1 (English) into their L2 (ASL)? For example after hearing an English word (problem) how fast do they translate it into the ASL equivalent (PROBLEM)? Participants are faster to translate from L1 into L2 as iconicity increases ($\beta = -0.42, t = 3.24, p = .002$). Participants are also faster translating L1-L2 as the frequency of items increases ($\beta = -93.35, t = -2.66, p = .012$). There is a significant interaction approaching significance at the $p = .05$ level between iconicity and SRT-score ($\beta = 0.05, t = 1.89, p = .059$), indicating that iconicity speeds up forward translation for less proficient signers, but this effect diminishes as proficiency increases. Novice participants were faster to translate the
English word *drink* to the high iconicity ASL sign DRINK than the low iconicity *measure* to MEASURE. Effects of frequency on translation performance, by contrast, did not interact with proficiency.

Figure 14 – Main effects of (a) iconicity, (b) SRT-score, and (c) frequency on forward RT.

Figure 15 – Interaction of iconicity and SRT-score on forward RT.
Study 3: Backward Translation
Results for this study were also analyzed using mixed-effects with crossed random effects for participants and items. Iconicity values used were identical to those used in Study 2.

Accuracy
Mean accuracy was .86; accuracy per item ranged between .22 and 1.00 (SD = .23); accuracy per subject ranged between .41 and 1.00 (SD = .29). An omnibus model for accuracy was estimated first. Fixed effects included linear effects and interaction between the experimental variable (iconicity), participant variables (age and SRT-score) and stimulus variables (word length, frequency, concreteness, imageability, and number of translation equivalents). Random effects were random intercepts for participants and items. The model was trimmed to remove nonsignificant fixed effects. SRT-score was the only factor of participant and stimulus variables to reach significance. The final model included the modulation of the effects of the experimental variable (iconicity) by the language experience of the participants (SRT-score) and a random structure including intercepts for participants (subject) and stimuli (item). The model (in R notation) was specified as (Accuracy ~ Iconicity * SRT + (1|item) + (1|subject). Table 9 shows the results of fitting this model for the experimental variables and interactions.

<table>
<thead>
<tr>
<th>Fixed Effects and Interactions</th>
<th>β</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iconicity</td>
<td>0.00</td>
<td>4.44</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>SRT-score</td>
<td>0.06</td>
<td>6.06</td>
<td>.001</td>
</tr>
<tr>
<td>Iconicity × SRT-score</td>
<td>-0.00</td>
<td>-4.24</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Table 9 – Main effects and interactions: Backward accuracy.

\(^7\) As in Study 2: Forward Translation, removing items with accuracy below .5 did not change the overall pattern of results.
Figure 16 – Main effects of (a) iconicity, and (b) SRT-score on backward accuracy.

Figure 17 – Interaction of iconicity and SRT-score on backward accuracy.

How accurate are participants when translating from their L2 (ASL) into their L1 (English)? For example after seeing an ASL sign (PROBLEM) how fast do they translate it into
the English equivalent (problem)? Participants show greater accuracy translating from L2 into L1 as iconicity increases ($\beta = 0.00, t = 4.44, p < .001$). Participants show greater accuracy when translating from L2 into L1 as SRT-score increases ($\beta = 0.06, t = 6.06, p = .001$). There is a significant interaction between iconicity and SRT-score ($\beta = 0.00, t = 4.24, p < .001$), indicating that novice signers are more accurate when translating more iconic signs into English than less iconic signs. Translation accuracy improves with proficiency to the point that highly proficient signers are at ceiling for backward translation accuracy regardless of iconicity. Novice participants were more accurate to translate the high iconicity ASL sign DRINK to the English word *drink* than the low iconicity MEASURE to *measure*.

**Reaction Time**

An omnibus model for RT was estimated next, using only correct trials. Fixed effects included linear effects and interactions between the experimental variable, participant variables (age and SRT-score) and stimulus variables (word length, frequency, concreteness, imageability, and number of translation equivalents). Random effects were random intercepts for participants and items. Models were trimmed to remove nonsignificant fixed effects. Word length, concreteness, and imageability had no significant effects or interactions, but there were main effects for frequency and number of translation equivalents. A main effect and interaction for the subject variable SRT-score was found, but no effect was found for age. The model was subjected to criticism focusing on the residuals. Data points ($n = 26; 3\%$ of the data) with absolute standardized residuals exceeding 2.5 standard deviations were removed from the data set and the model was re-estimated, resulting in improved fit. Correlation of fitted values with trimmed RTs was $R^2 = .65$, up from $R^2 = .58$; BIC = 12398.96, down from 13185.32. All subsequent analyses employ the trimmed data set. The final model included the modulation of the effects of the
experimental variable (iconicity) by the language experience of the participants (SRT-score),
fixed effect of frequency (log frequency) and number of translation equivalents (translations) and
a random structure including intercepts for participants (subject) and stimuli (item). The model
(in R notation) was specified as (RT ~ Iconicity * SRT + Log frequency + Translations +
(1|item) + (1|subject). Table 10 shows the results of fitting this model for the experimental
variables and interactions. To help interpret these results, Figure 18 shows main effects of (a)
frequency, and (b) number of translation equivalents on RT. Figure 19 shows the interaction of
iconicity and SRT-score.

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed Effects and Interactions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iconicity</td>
<td>-0.06</td>
<td>-.73</td>
<td>.469</td>
</tr>
<tr>
<td>SRT-score</td>
<td>-8.76</td>
<td>-.46</td>
<td>.647</td>
</tr>
<tr>
<td>Log Frequency</td>
<td>-67.85</td>
<td>-2.75</td>
<td>.010</td>
</tr>
<tr>
<td># of Trans. Equiv.</td>
<td>187.47</td>
<td>-1.72</td>
<td>.095</td>
</tr>
<tr>
<td>Iconicity × SRT-score</td>
<td>-0.02</td>
<td>-2.06</td>
<td>.039</td>
</tr>
</tbody>
</table>

Table 10 – Main effects and interactions: Backward RT.
How fast are participants when translating from their L2 (ASL) into their L1 (English)?
For example after seeing an ASL sign (PROBLEM) how fast do they translate it into the English equivalent (problem)? Participants are faster translating L2-L1 as the frequency of items increases ($\beta = -67.85$, $t = -2.75$, $p = .010$). Participants are slower translating L2-L1 as the number of translation equivalents increases approaching significance at the $p = 0.1$ level ($\beta = 187.47$, $t = -1.72$, $p = .095$). There is a significant interaction between iconicity and SRT-score ($\beta = -0.02$, $t = -2.06$, $p = .039$). As proficiency increases, responses are faster, but greater gains are made for more iconic signs than for less iconic signs. Expert participants were much faster than novices to translate the high iconicity ASL sign DRINK to the English word drink than the low iconicity MEASURE to measure. Iconicity matters for the more proficient signers only in the backward direction.

**Discussion**

This study predicted that iconicity would affect the speed and accuracy of translation for bimodal bilinguals. In forward (L1-L2) translation, increasing iconicity was predicted to facilitate translation time and increase accuracy, because motivated mapping of conceptual representation to visual form would facilitate L2 lexical access. In backward, (L2-L1) translation, increasing iconicity was predicted to inhibit translation time and decrease accuracy because faster, more accurate lexical mediation (at least for novices) should be prevented by conceptual mediation. This study also predicted that novices and experts should not be affected by iconicity in the same way; form-meaning relationships were predicted to be processed primary conceptually in early language acquisition, but as proficiency grows they would have increasing effects at the lexical level.

Results show that in forward translation there is a main effect of group. Experts are faster to translate from English into ASL than novices. There is also a main effect and interaction of iconicity with group. Novices show facilitation when translating iconic items from English into
ASL, but experts do not, indicating conceptual mediation is facilitated in highly iconic signs for novices, but not for experts. In backward translation, there is no main effect of either iconicity or group, but there is an interaction of group and iconicity, indicating experts experience facilitation when translating iconic items from ASL into English. This shows that conceptual mediation is facilitated for highly iconic signs for experts, but not for novices.

The effect of iconicity in translation production in both novices and experts is to facilitate activation of conceptual-L2 links. This means that forward translation, which normally uses conceptual mediation, is significantly facilitated in novices, but experts show no effect since they are presumably at, or near, ceiling for production in the manual modality. In backward translation, novices use lexical mediation – though likely not for all items since there is a non-significant trend for iconic facilitation. Experts do show the iconic facilitation effect, which is aligned with findings that expert bilinguals use conceptual mediation for both forward and backward translation; both for behavioral measures (Chen & Leung, 1989; Macizo & Bajo, 2006), and more recently in an ERP study (Moldovan, Demestre, Ferré, & Sánchez-Casas, 2016).

**Mediation in Translation**
Since Kroll and Stewart’s (1994) seminal study demonstrated that forward translation uses conceptual mediation but backward translation uses lexical mediation, studies on bilingualism have theorized that these two distinct pathways arise due to how adults typically learn their second language compared to how a first language is acquired. Connections between lexical items and their conceptual representations that are built up early in life and over the course of a speaker’s interaction with their native language are automatic and strong. In fact, many monolinguals struggle to intuitively understand the notion of separate lexical and conceptual
processing since speaking can feel one and the same with thinking. After speakers pass through the critical period of language acquisition (Lenneberg, 1967), they find learning to associate lexical forms with concepts is greatly facilitated by recruiting their strong, pre-existing, L1 associations. This is the basis of lexical mediation; it is easier to connect an L2 form to an L1 form, already strongly entrenched with its conceptual representation, than it is to connect an L2 form directly to conceptual memory. Critically, because the direction of this association is from L2 to L1, forward translation cannot take advantage of lexical mediation and must rely on conceptual mediation – connecting concepts to L2 forms after processing L1 lexical items.

The nature of a bilingual’s language existing in separate modalities (visual/manual vs. aural/oral) makes direct comparisons between signs and words – and thus between unimodal and bimodal bilinguals – very difficult (Emmorey et al., 2012). Novices in these studies show asymmetry in the degree to which they recruit conceptual facilitation from iconic stimuli – leaning on iconic motivations in forward, conceptually mediated translation but not in backward lexically mediated translation. If it were possible to directly compare spoken language reaction times to signed language reaction times, comparative differences for forward translation based on the iconicity of the stimuli in ASL might be revealed for experts. Measuring the time-course of phonological activation via ERP has the potential to answer this question, but behavioral measures must remain speculative.

The most proficient bimodal bilinguals in these studies are superior to novice bimodal bilinguals in terms of accuracy and translation speed in both directions. Importantly, iconicity affects the two groups differently. A recently published survey of English-ASL interpreters reported that the majority of interpreters, regardless of number of years of experience, prefer forward to backward interpretation (Nicodemus & Emmorey, 2013). This preference does
weaken as number of years of experience increases. Interestingly, Children of Deaf Adults (CODAs) did not express the same pattern, instead reporting a preference for backward translation. Some possible explanations for forward translation preferences in second language learners of ASL include training (interpreter training programs typically emphasize English-ASL interpretation), language mastery (even the most experienced interpreters may not be at native levels of proficiency), and the majority of interpreting assignments involve rendering English source material into ASL for Deaf clients. There is evidence novices may struggle to successfully monitor visual/manual output (Nicodemus & Emmorey, 2015), which could contribute to this preference as well. The results of this study, on the other hand, show that in forward translation, bimodal bilinguals have a “cushion” for the hard cognitive work of translation in the range of approximately 2000 ms (the difference in RTs in the forward and backward translation conditions), and their output is much more consistent across a range of signs. By contrast, in backward translation, interpreters are generating a much faster response, and the more skilled they become, the more uneven their performance is with regards to iconicity (i.e. the more iconic signs are, the faster the translation time).

These results suggest that the preference for forward translation reported by English-ASL interpreters may not completely reside in socio-linguistic or proprioceptive peculiarities of using a signed language, but also in the prevalence of lexical-conceptual connections in signed languages’ lexicons that serve to set them apart from spoken languages in terms of their distribution of arbitrariness and systematicity. Evidence for this hypothesis comes from the observation that speed and accuracy is increased for novices in forward translation as sign iconicity increases, but that this preference weakens as experience increases.
Imageability Effects

Imageability is the measurement of how easy or difficult it is to visually or acoustically imagine a concept. All imageability ratings were collected on the English translations of stimuli since there is currently no database that has measured the imageability of the ASL lexicon. While there is no doubt some variation in this parameter for any given lexical item cross-linguistically, this study assumes that English imageability ratings capture an effect that is rooted in the shared conceptual system. Imageability is a particularly interesting control measure for studies of lexical access in signed languages. The ability of the visual modality to encode visual aspects of concepts in lexical forms suggests that the amenability of visualizing a concept could interact with how lexical forms emerge and are processed. Effects of imageability (and the closely related metric concreteness) are discussed in some studies that examine iconicity effects, and are often highly correlated with iconicity measures (Baus et al., 2013; Thompson et al., 2010, 2012, Anible, in prep).

An examination of imageability ratings for these studies indicates that for novices, imageability effects are indistinguishable from iconicity effects both in forward and in backward translation. When the correlation of RT and iconicity is compared to the correlation of RT and imageability for participants with low proficiency (SRT-score ≤ 2) there is no significant difference between them. Using the same trimmed data sets as in Study 2 and Study 3, the results of a comparison of two overlapping correlations (iconicity and imageability) based on dependent groups (RTs of low proficiency participants) reveals no difference between correlations for forward translation RTs \( r (\text{RT | iconicity}) = -.15 \times r (\text{RT | imageability}) = -.06; n = 379, p = .14 \) or backward RTs \( r (\text{RT | iconicity}) = -.07 \times r (\text{RT | imageability}) = .01; n = 344, p = .19 \). For example, in forward translation novices are faster to translate the more iconic sign SIT to sit than the less iconic sign MUSIC to music as well as being faster to translate the more imageable
concept “smoke,” and the less imageable concept “carry.” In backward translation they show no differences for either.

In experts (RTs of high proficiency participants with SRT-score ≥ 7), on the other hand, imageability has little influence either in forward or in backward translation, but iconicity does have an effect in backward translation. The results of an identical comparison as the one preformed on novices reveals no difference between correlations of iconicity and imageability for forward translation RTs \( r \text{(RT | iconicity)} = -.11 \times r \text{(RT | imageability)} = -.02; n = 213, p = .23 \], but does reveal a significant difference for backward RTs \( r \text{(RT | iconicity)} = -.24 \times r \text{(RT | imageability)} = -.03; n = 219, p < .001 \). For example, in backward translation experts are faster to translate *sit* to the more iconic sign SIT than *music* to the less iconic sign MUSIC. They do not, however, show any facilitation in translating the more imageable concept “smoke” than the less imageable concept “carry.” In forward translation they show no differences for either iconicity or imageability.

Baus et al. (2013) reported that in forward translation iconic signs were translated faster than non-iconic signs for novices, but not for experts and that in backward translation iconic signs were translated slower than non-iconic signs for experts, but not for novices. This study replicates the forward direction results; for forward translation iconic signs were translated faster than non-iconic signs by novices. Results for experts, by contrast, are contradicted; in backward translation iconic signs were translated faster by experts, but not by novices. Differences for experts between these two studies, and the results for novices appear to be substantially elucidated by disambiguating effects of imageability from those of iconicity. In a cognitive view, effects of iconicity should increase as participants become more knowledgeable about the

---

8 For all reported correlations the inter-factor correlation of iconicity and imageability was approximately \( r = .25 \) with some variability stemming from some items being removed from the data set by trimming.
language. This suggests that in forward translation, iconicity is functioning as transparency – the ability of non-signers to guess the meaning of signs based on their form, alone (Klima & Bellugi, 1979). Imageability for novice English-ASL bilinguals behaves functionally equivalent to transparency. Access to transparent conceptualizations are available to everyone, regardless of their experience with the signed language. Experts link visual construals to signs regardless of their transparency, but novices receive a boost that enhances their recall of the sign form. This suggests that during early stages of second language acquisition of a signed language the link between conceptual representation and a signed L2 lexicon is primarily unidirectional from meaning to form. In backward translation, novices don’t benefit from transparency since sign form does not transparently map to meaning as clearly (the concept – L2 mapping is unidirectional), but as signers become more proficient in the language, the concept – L2 form mapping strengthens bi-directionally, and we begin to see the effect of iconicity on translation time because iconic signs are more closely associated to their meanings than less iconic signs.

Unlike unimodal bilinguals who can benefit from direct lexical connections even as experts, signers do not appear to sidestep conceptual mediation. Instead, the better you know the language, the more differentiated the iconicity effects become from imageability effects.

**Conclusion**

This study found that iconicity affected the speed and accuracy of translation by bimodal bilinguals. Specifically, in forward (L1-L2) translation, increasing iconicity primarily facilitated translation time and increased accuracy for novice bilinguals, because high imageability facilitated L2 lexical activation due to the motivated mapping of conceptual representation to visual form. Experts, who were faster and more accurate than novices in forward translation showed no facilitation effects in forward translation. This result is interpreted as indicating that these participants were translating single words at ceiling when relying on visual RT
measurements. In backward, lexically mediated (L2-L1) translation, increasing iconicity did not force conceptual mediation by novice bilinguals; RTs were not inhibited by increasing iconicity values nor was accuracy decreased, as predicted. Rather, novices showed no effects of iconicity because iconic facilitation of conceptual access played no role in translation from ASL to English. Experts, however, were facilitated by increasing iconicity in backward translation (though accuracy was at ceiling). These results partially match earlier predictions based on Anible (in prep) for novices, indicating there are facilitation effects in forward translation comprehension due to prevalence of conceptual facilitation from iconicity, and inhibition effects for experts. The former was found, but the latter was not. Rather than causing inhibition, iconicity in translation was found to only result in facilitation when translation between languages is conceptually mediated and translation latencies are not already performed as fast and as accurately as possible. Whereas novices may have engaged in lexically mediated L2-L1 backward translation, the evidence presented here is consistent with the view that experts engaged in conceptual mediation.

These findings leave open questions about the extent to which iconicity may affect sentential or discursive level interpretation. Iconicity has been documented in signed language discourse (Dudis, 2007; Russo, 2004) and is likely to contribute additional challenges in rendering on-line interpretations of source text between signed and spoken languages. Particularly, highly iconic classifier constructions in ASL and other signed languages often lack concise translations in English or another spoken language in much the same way that metaphorical and idiomatic expressions resist faithful translation between source and target in two spoken languages. Comparisons between iconic motivations and metaphorical extension have already received some treatment in the literature (Meir, 2010), and the findings revealed in
this paper, and in related studies (Baus et al., 2013) contribute to our understanding of these forces in language processing.

An important implication for future research drawn from the results of this study is the difference between *language external* and *language internal* facilitation of conceptual access through the mechanism of motivated mappings between form and meaning. Language external motivations are conceptualizations based on general human experiences. The concept of throwing a ball inherently encodes mental representations of experiencing ball-throwing, experiencing watching ball-throwing, etc. These representations are often highly imagistic and are capable of being encoded to linguistic behaviors more directly in signed languages than in spoken languages. Language internal motivations, on the other hand, do not necessarily overtly encode a visual or manual behavior, but are instead part of a collection of linguistic forms that share conceptual similarities. Language external motivations are more likely to be accessible without experience in a given signed language, but language internal motivations are only accessible once the morphological and diagrammatic patterns have been mastered. The iconicity measurements in this study tap into language internal motivations because only signers and not non-signers showed a difference between pictures that profiled the language internal motivation of the sign.

Imageability was completely confounded with the effects of iconicity for novices in this study, but had little effect on experts, supporting previous findings that early on in the acquisition process the relative amenability to visualization of a concept increases the strength of concept-L2 links (Baus et al., 2013; de Groot & Poot, 1997). As proficiency in a signed language grows this heuristic appears to decrease in importance for producing translation equivalents and instead “horizontal” connections between language forms and their related conceptual
representations take over. This result is best understood within a usage-based approach to cognitive iconicity (Occhino-Kehoe, 2016; P. Wilcox, 2000; S. Wilcox, 2004) which provides a framework that explains both external and internal patterns of language motivation while simultaneously incorporating developmental hypotheses. In this perspective, learning of language internal patterns is accumulated over the course of an individual’s experience with a signed language.

Discussions of cognates in unimodal bilingual studies have striking similarities to this perspective. The cognate status of a word depends on formal similarity across translations. The *cognate effect* is the observation that bilinguals have low latencies in picture naming tasks for items whose translations are phonologically similar across languages (as in the Spanish-English pair “guitarra” – *guitar*) (Christoffels, Firk, & Schiller, 2007; Costa et al., 2000, 2005). Van Hell and De Groot (1998) argue that cognate status might influence conceptual processing since cognates should have larger conceptual overlap compared with non-cognates. Iconicity in signed languages, then, seems to behave like a “universal cognate” in the sense that it activates conceptual representations in contexts where lexical processing would otherwise occur.

Recent explorations of iconicity for signed languages (Lepic, 2015; Occhino-Kehoe, 2016) have concluded that the majority of “iconic” signs are *diagrammatic* (Haiman, 1980; Ungerer, 1999), rather than solely *imagistic*, indicating that conceptual associations motivate lexical structure and vice versa. Many studies have used non-signer judgements of iconicity to assess the effect of this property in on-line language comprehension and production, but non-signers are, by definition, insensitive to language internal patterns that have an increasing effect as language proficiency increases. Without taking into account both diagrammatic and imagistic motivations, investigations of iconicity effects in signed languages run the risk of skewing results toward
behaviors that are primarily present early in the second language acquisition process and overlook effects that surface only after more complete knowledge of a language’s patterns have been learned and internalized.
Conclusions

Together these two studies provide evidence for a cognitive, usage-based approach to sign language lexical motivation. Signed languages are not unique from spoken languages because of their subsymbolic networks, but because subsymbolic units are more likely to be imagistically motivated in comparison to spoken languages (Lepic, 2015; Occhino-Kehoe, 2016).

Current views of bilingualism assume complete independence of form and meaning, even in the case of bimodal bilinguals (Giezen, Emmorey, & Blumenfeld, 2013; Morford, Wilkinson, Villwock, Piñar, & Kroll, 2011; Shook & Marian, 2013). Advances in cognitive linguistics and the results of these studies indicate that this assumption is not tenable, either for spoken or signed languages. Investigating lexical processing of more vs. less iconic domains of the lexicon in bimodal bilinguals who vary in language dominance, allows us to improve our understanding of bilingual lexical processing. Because of its developmental hypothesis, the RHM is a good candidate for a consideration of this nature.

In comprehension, these studies showed that bimodal bilinguals exhibit increased phonological interference when processing semantic competitors that were analogically related to each other at the subsymbolic level. This is consistent with evidence of semantic facilitation in deaf signers when processing conceptual properties for motivated signs (Thompson et al., 2009), but phonological inhibition when processing lexical properties (Thompson et al., 2010). Results also indicated that measurements of motivation that are primarily language external (such as non-signer transparency judgments and imageability measurements) are too simple to accurately capture how lexical representations in a signed language change as proficiency increases. Language internal motivations interact with language external motivations so that experts experience increased phonological activation when imageability is high, but novices experience
increased semantic activation. In other words, experts process diagrammatic iconicity as phonological competitors since the pattern matches that of phonological inhibition increasing with higher imageability, but novices process diagrammatic iconicity as semantic competitors since the pattern matches that of semantic inhibition decreasing with higher imageability.

In production, diagrammatic iconicity affected the speed and accuracy of translation by bimodal bilinguals. In forward (L1-L2) translation, increasing iconicity primarily facilitated translation time and increased accuracy for novice bilinguals, because high imageability facilitated L2 lexical activation due to the motivated mapping of conceptual representation to visual form. Experts, who were faster and more accurate than novices in forward translation, showed no facilitation effects from iconicity in forward translation. This result is interpreted as indicating that these participants were translating single words at ceiling when relying on visual RT measurements. In backward, lexically mediated (L2-L1) translation, novices showed no effects of iconicity because iconic facilitation of conceptual access played no role in translation from ASL to English. Experts, however, were facilitated by increasing iconicity in backward translation.

Results of these studies suggest a modification to the RHM that includes a more nuanced view of conceptual-L2 connections based on the qualities of the interaction between form and meaning for many lexical items in signed languages. In particular, results of interactions with imageability measurements in these studies build on past evidence that lexical items with this quality are not sufficiently modeled by the RHM. Van Hell & De Groot (1998) found effects of imageability in both forward and backward translations for both fluent and less fluent bilinguals indicating there is little or no lexical mediation for highly imageable words. Other evidence that form/meaning connections are problematic for the RHM come from studies exploring number
magnitude effects; a phenomenon where smaller numbers are able to be produced and perceived faster than larger numbers (Duyck & Brysbaert, 2004). For instance, it’s easier to select the larger number in the sequence 2-3 than it is to select in the sequence 7-8. Since any semantic point in a number sequence has a non-lexical referent (the Arabic number) and a lexical referent (the linguistically specific form/meaning pairing) this effect is particularly interesting for exploring lexical translation. When testing translation time, the magnitude effect was observed in backward translation, but not in forward translation. Proficiency decreased the strength of the findings, though not significantly. To rule out that the magnitude effect was not due to frequency (since smaller numbers are more frequent), subjects were taught a novel language’s (Estonian) number words. Identical effects were found. These results also appear to be in conflict with RHM predictions. An interpretation of these results that lies within the framework of the RHM is that conceptual mediation for numbers in backward translation comes from mapping number forms onto an existing cognitive structure. In this interpretation, conceptual information is encoded onto a metaphorical number line where smaller numbers occur earlier than larger numbers. In both this case and in the case of imageability effects, pairing of a pre-existing mental conceptualization to an L2 language form might equalize normal faster L2-L1 lexically mediated translation latencies with conceptually “facilitated” L1-L2 translation.

Priming experiments may support this interpretation. Evidence of priming in a semantically driven picture-naming task supports the idea that form/meaning connections can facilitate forward translation. Sholl et al. (1995) presented line drawings and asked participants to name the objects in their L1 and L2. Primed L2 words were translated faster forwards than backwards. The pairing of a pre-existing mental conceptualization to and L2 language form affects translation speed in the same way as if it had been primed.
Finally, the studies performed in this dissertation also provide evidence that bimodal bilinguals are affected by richly detailed embodied representations in similar ways to unimodal bilinguals (Vukovic & Williams, 2014). There is growing evidence that linguistic input activates simulation mechanisms (Barsalou, 2008). Phrases with visual imagery like “The ranger saw the eagle in the sky,” evoke conceptualization of extended wings which facilitate participants to name a picture of an eagle with wings extended faster than with wings folded (Zwaan & Madden, 2005). Equivalent effects are evoked in hearing and deaf signers, but by a single lexical item (Grote & Linz, 2003; Thompson et al., 2009). Combining these insights with what we now know about the morphological structure and processing of the lexicon from psycholinguistic experimentation and computational modeling we can propose a more nuanced model for bimodal bilingual lexical access that accounts for subsymbolic and diagrammatic interactions as well as the influence of imagistic motivations.

The directional access of ASL sign SIT is modeled for novice English-ASL bilinguals in Figure 1. Forward (L1-L2) translation is aided by the imageability of the concept, facilitating access to the form, but the subsymbolic network of other signs that recruit the “legs as index and middle finger” connection (e.g. DANCE, WALK, RIDE, STAND, etc.), is not completely active. Critically, this means that the imageability links are unidirectional from conceptual memory to the L2. Backward (L2-L1) translation cannot benefit from imagistic motivations and diagrammatic connections are similarly not yet strong enough. Instead, as in unimodal bilinguals, lexical mediation is the most efficient route from L2 forms to L1 forms. As proficiency in ASL increases, the links from L2 forms to conceptual representations will grow stronger and, as the results of the second paper indicate, result in facilitated L2-L1 translation as the diagrammatic iconicity of the items is more accessible.
The data from these studies support the predictions of a model in which imagistically motivated sub-lexical forms in a signed language can take advantage of shared conceptual representations cross-linguistically that are subsequently built up and distributed analogically within the L2 system as proficiency increases. This proposal has important implications, not only for revealing aspects of translation performance, but also for exploring the role of increasing proficiency in bimodal bilingual memory and for suggesting new directions in the pursuit of understanding second language acquisition. In particular, explicit instruction in language internal motivations as in the case of morphological cohorts (e.g. DANCE, WALK, RIDE, STAND, etc.), has the potential to increase proficiency in comprehension and production in a similar way that learning sound-symbolic patterns helps people learn new words in spoken languages (Lockwood, Hagoort, & Dingemanse, 2016).
The metaphor MORPHEMES ARE ATOMS, like all good metaphors, excites our imaginations and frames our understanding of complex and intangible processes in useful ways. Even an over-extension of the metaphor where subsymbolic form/meaning connections could be understood as subatomic in some sense, leaves us satisfied with the notion that one thing we understand well can be an analog for something that is still somewhat nebulous. Ultimately, however, any metaphor can only get us so far. We should be wary of the seemingly logical extensions the metaphor suggests, because the observations that fit are equally important as the observations that don’t; what are imagistic motivations in our atomic model of language meaning, for instance? These have no mapping within our metaphor. Any attempts to use an explanatory tool, predictively are likely to suffer from circularity and obfuscation. While the architecture that underlies language processing and structure is undoubtedly fundamentally physical, chemical, and quantifiable, our ability to observe, manipulate, and understand it is currently limited and we must build new metaphors/models that sufficiently explain the results we observe.
Bibliography


http://doi.org/10.1016/j.brainres.2007.01.137

http://doi.org/10.1080/14640748108400805


Keras, D. (1978). Beyond pictures and words: Alternative information-processing models for


http://doi.org/10.1191/0267658302sr201oa


http://doi.org/10.1006/jmla.1994.1008


Retrieved from http://CRAN.R-project.org/package=lmerTest


Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the


Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations
http://doi.org/10.1207/s15516709cog2705_4


3514.89.1.1

Comprehension of Abstract and Concrete Verbal Materials. *Journal of Experimental
Psychology: Learning, Memory, and Cognition, 9*(1), 82–102.

(Ed.), *Translation: Applications and Research* (pp. 92–116). New York.

translation: A test of asymmetries in bilingual memory. *Psychological Science, 6*(1), 45–


University Press.


http://doi.org/10.1006/jmla.1997.2524
