Learning from L2 ASL Fingerspelling Comprehension

Abstract

Researchers have claimed that transitions are the most important part of fingerspelling comprehension. Other work has shown that, while transitions may be one cue involved in successful fingerspelling comprehension, they are not the only cue utilized. There also seem to be differences in cue-weighting between skilled signers and ASL learners. This study assesses the efficacy of a training program designed to help ASL students learn to re-weight cues to fingerspelling comprehension to improve their performance on this skill. Results indicate even short exposure to explicit phonetic instruction significantly improves participants' ability to understand fingerspelling. Our work suggests that ASL curricula should include this type of training to improve students' fingerspelling comprehension abilities.

1. Introduction

Fingerspelling, or spelling on the hands using manual letters, is a process used in some signed languages to represent represent written words in the manual modality. In the case of American Sign Language (ASL), one-handed manual letters represent the orthographic characters in the English alphabet. There are many compelling reasons which make fingerspelling an appealing area of study but this study focuses on the puzzling fact that students and researchers report that fingerspelling is one of, if not the hardest aspect of ASL acquisition (e.g., Wilcox, 1992; Patrie and Johnson, 2010). Despite this, there has been very little research on how to address this problem pedagogically. The present work aims to do just that through an integration of past work on fingerspelling comprehension in skilled deaf signers and hearing L2 learners, as well as work on the use of explicit phonetic instruction in L2 teaching generally. The remainder of this section summarizes relevant literature on these topics.

In a fingerspelling recognition and comprehension task, participants were often able to comprehend a word – they would produce an ASL translation equivalent – but not be able to spell it back correctly (Hanson, 1981). The researcher concluded, therefore, that deaf signers do not read fingerspelling letter-by-letter, but rather appear to be sensitive to the shape of the word (Padden, 1998, among others). What is not clear from this experiment is what cues within the fingerspelling stream signers are using. Schwarz (2000) examined this issue by isolating certain sources of information within the fingerspelling stream to see how this impacts word comprehension. Wilcox (1992) suggested the relevant cue for successful fingerspelling comprehension lies the transition segments because they are the most information rich portion of the signal. Schwarz's experiment also addresses this assertion.

Using fingerspelling video data, the researcher masked one hold segment – portion in which the signer holds a letter posture – within each word. The mask replaced hold video frames with a magnified image of the signer's palm, with the fingers edited out. The configuration of the hand was imperceptible; see Figure 1. Deaf signers were asked to write down the word they saw fingerspelled on screen.

Schwarz (2000) calculated her findings in three ways, two of which are relevant for the present work. The Strict Approach accepts only exact matches to the stimulus presented. For example, if the word H-A-P-P-Y were fingerspelled, only a written response of *happy* would be accepted as correct. The Envelope Approach was based on Akamatsu's *movement envelope* (1985); see Figure 2. If the stimulus item were L-O-S-E-R, a written response of *loner* would be counted as correct because both -S- and -N- are short letters, which means the two forms have the same movement envelope. A response of *lower*, however, would be marked incorrect. Because -

W- is a tall letter, its inclusion results in a different movement envelope than the original stimulus item. Each of these fingerspelled tokens is pictured in Figure 3.

(a) (b)

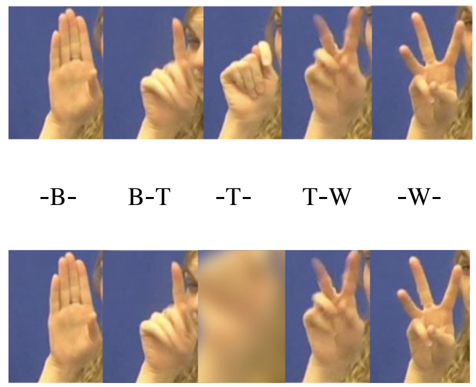


Figure 1: Recreation of stimuli according to Schwarz (2000). The (a) figure includes still images of the fingerspelled abbreviation B-T-W. The (b) figure includes a mask over the -T- hold created by zooming in on the signer's hand. The color clearly matches with the signer's hand, but the configuration of her hand is not distinguishable.

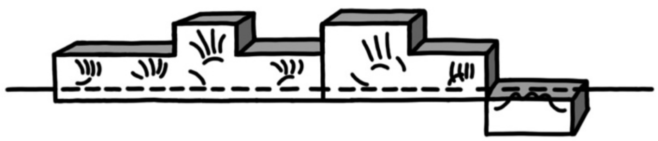
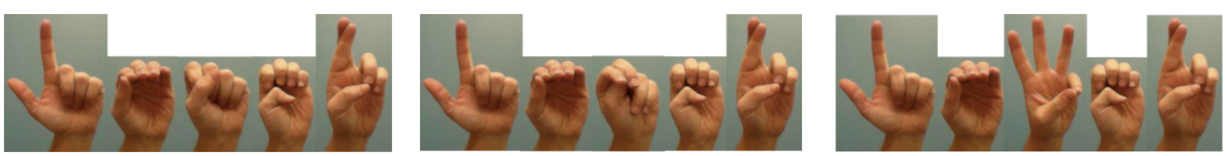


Figure 2: Re-creation of Akamatsu's visualization of the movement envelope for the fingerspelled word S-A-F-E-W-A-Y. The schematic depicts the perceived contour of the item rather than the specifics of each letter individually. This figure can be read as "two short letters followed by a tall letter, followed by short, tall, and short letters, and ending with an extra short letter."



(a) L-O-S-E-R (b) L-O-N-E-R (c) L-O-W-E-R

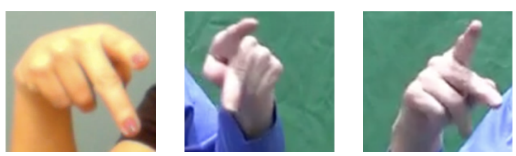
Figure 3: Three fingerspelled items. Items (a)-(b) both count as correct according to the Envelope Approach in Schwarz (2000), however (c) is incorrect because unlike -S- and -N-, the fingerspelled letter -W- is tall, thus creating a different movement envelope.

The results revealed better performance in the Envelope versus Strict Approach. This suggests that while transitions are insufficient for correctly identifying fingerspelled words, they do give some clue as to the overall shape, offering independent evidence for Hanson's conclusion regarding the use of the shape of fingerspelling for word comprehension. In closing her thesis, Schwarz (2000) suggests several logical ways to build from her work on fingerspelling comprehension. One suggestion is particularly relevant. She notes that "...fingerspelling is a linguistic system with built-in intrinsic redundancies that aid in . . . perception" (p. 80). Transitions between letters – potentially containing information about the previous and or subsequent segments – the movement envelope, and canonical handshapes, when they occur all benefit listeners to varying degrees. Thus, a logical next step is to examine which of these built-in redundancies is most useful cue for successful fingerspelling comprehension. The next pair of studies examine this very question.

Geer and Keane (2014) also used masking to isolate cues listeners attend to for fingerspelling comprehension. Instead of masking a single hold at the beginning, middle, or end of the word as Schwarz did, Geer and Keane masked each of the holds or transitions in a fingerspelled utterance, depending on the experimental condition. Unlike Schwarz (2000), who studied deaf signers, Geer and Keane (2014) examined fingerspelling comprehension among second-language learners of ASL in their third-semester of language instruction. Participants watched video clips of fingerspelled items produced by a native signer. A black screen mask was inserted for the duration of each of the transitions for the *holds-only* condition and for the duration of each of the holds to create the *transitions-only* condition. This study demonstrated that student learners perform better when only provided with the hold portion of the fingerspelling signal. Second-language learners seem to weight canonical handshapes most heavily, while Hanson and Schwarz's findings that suggest deaf signers focus more on the movement envelope.

In a larger follow-up study, Keane and Geer (2016) showed that, in addition to performing significantly better in the *holds only* condition, errors are strongly predicted by the presence of non-default palm orientation in certain manual letters. Most manual letters are produced with the wrist extended or slightly hyperextended, elbow flexed past 90 degrees, and the forearm rotated such that the palm is facing outward. Several letters are produced with a different orientation which does not conform to this default. The letters -G-, -H-, -P-, and -Q- face inward and -P- and -Q- also face downward. Students correctly identified fingerspelled words 56% of the time. Their performance on items with non-default orientation letters was worse; 32% were correctly identified.

An additional potential complication with the fingerspelled letter -P- is that some (near)native signers distinguish -K- and -P-, not with wrist flexion as is the case in citation form, but with a (sometimes slight) orientation change, an inward rotation, or *supination* of the forearm; see Figure 4. This leads to the specific error of misreading the letter -P- as -K- which Hanson (1981) also found in her study. Errors involving confusion of -K- and -P- were particularly prevalent.



(a) (b) (c)

Figure 4: Different productions of -P-. Both (a) and (b) show a more canonical production of -P- with flexion at the wrist, while (c) features very slight flexion at the wrist and inward rotation, or supination, of the forearm.

A possible explanation for the results Geer and Keane (2014); Keane and Geer (2016) found, which seem inconsistent from studies of skilled signers may be related to the manner by which adults acquire fingerspelling. Young children acquiring ASL fingerspelling first produce forms following the movement envelope because they are unable to produce adult target forms due to lack of finger dexterity, but their forms do capture the overall shape of adult forms (Akamatsu, 1985). Padden (2006) found that as children develop literacy, they go through a period in which their ability to produce fingerspelling declines as they become aware of the relationship between English letters and their manual representations in ASL. As their literacy improves, they are once again able to produce the adult(like) target. Adult-learners do not follow this two-step process. From their first exposure to fingerspelling, they understand the mapping between orthographic characters and their manual representation. This seems to impact their ability to see the holistic shape that children are so attuned to at a young age.

Additional evidence that the shape of fingerspelled words is important for comprehension comes from the concept of *local lexicalization*. Brentari (1998) details how longer fingerspelled items that are produced repeatedly in a discourse undergo local lexicalization. The term "local" refers to the context of utterance. Some lexicalizations are not widely accepted but rather are coined locally for a specific discourse, then perhaps not used subsequently. In these locally lexicalized forms, the features of the full fingerspelled word that are retained are those that involve the biggest excursions through the signing space and are therefore presumably more salient to the perceptual system. This process retains the larger movements like the those introduced by manual letters with non-default palm orientation yet this set of letters – -G-, -H-, -P-, and -Q- – is the same as those that strongly predicted comprehension errors in student learners in their second year of ASL acquisition. Testing the rules Brentari (1998) proposes to govern local lexicalization, Stone et al. (2016) asked whether signers are in fact sensitive to highly sonorous locally lexicalized forms in much the same way hearing adults are sensitive to sonority constraints in spoken languages. Participants were presented with carefully fingerspelled low-frequency words of English and then with high and low sonority lexicalized forms and asked which they'd prefer as a lexicalized variant of the carefully fingerspelled word. The 35 deaf adults tested preferred high sonority forms, those that retained large movements, 83% of the time, while sign-naïve adults performed at chance, preferring highly sonorous forms only 52% of the time. This difference between groups is statistically significant.

Research suggests that skilled signers make use of cues in fingerspelling comprehension that L2 learners do not. A question that follows from these studies is whether ASL learners can be trained to make better use of these cues to improve their fingerspelling comprehension. We next review several projects on explicit phonetic instruction in L2 learning and discuss how this relates to how this method could be implemented for ASL learners and fingerspelling.

Explicit instruction is a form of teaching in which a "rule explanation forms part of the instruction...or if learners are asked to attend to particular forms and try to find the rules themselves" (DeKeyser, 2003, p. 321). The studies which have tested the use of explicit instruction for pronunciation have enjoyed largely positive results (e.g., Couper, 2003; Derwing et al., 1998; Macdonald et al., 1994). These are detailed here with respect to how these works may inform studies on explicit instruction for learners of ASL as a second language and how the present work fits within the larger literature on second-language learning and teaching.

Speakers appropriately weight various cues in their native languages such that they are highly successful in phoneme identification/discrimination. When cue-weights in a speaker's L1 and L2 differ, errors in identification may result (Ylinen et al. 2010, among others). A good example of this is vowel identification in English speakers versus Finnish or Greek speakers (Ylinen et al., 2010; Giannakopoulou et al., 2013). In Finnish and Greek, vowel length is phonemically contrastive but this is not the case for English. This does not mean that all vowels have equal length in English, but it means that English speakers know to ignore that cue, or at least to give it very little weight.

Ylinen et al. (2010) compared L1 and L2 cue-weighting among native English and L2

Finnish speakers then asked whether that weighting system could be altered with training. Using 19 English minimal pairs distinguished by the vowels /i/ and /ɪ/, they tested how accurately native English- and Finnish-speaking participants identified the correct word. With natural stimuli, participants performed equally well, but when vowel length was made equal (/i/ is typically longer), Finnish participants scored poorly. As one might predict based on their language background, Finnish speakers (incorrectly) weighted the cue of vowel length over the most relevant cue of tenseness.

Finnish speakers completed a training program in which they listened to four speakers (two males and two females), representing various English dialects from the US, UK, and Canada, producing the same 19 minimal pairs on which they had already been tested. The training proceeded as follows: correct responses resulted in feedback and advancing to next trial; incorrect responses resulted in feedback and a repeated trial. A post-test revealed no difference between English and Finnish speakers, regardless of whether stimuli were natural or synthesized, demonstrating that cue-weighting can be altered with training.

A similar study with Greek adults and children uncovered similar results (Giannakopoulou et al., 2013). Additionally, researchers evaluated the efficacy of a training which used only natural stimuli and one which used both modified and natural stimuli. Giannakopoulou et al. also found that cue-weighting can be altered (and with greater success in children than adults), but also that the training that combined modified and natural stimuli was most effective in teaching participants to re-weight cues to English vowel identification than natural stimuli alone. This suggests that forcing participants to be attentive to cues they otherwise ignore will help them learn to attend to cues they should weight more heavily.

Using DeKeyser's definition, participants in these studies were made to infer a rule about cue-weighting for English vowels, but it was never put into words for them. Other studies have examined the effects of verbally explaining how to identify and produce phonetic segments found only in learners' L2. Researchers like Schmidt (2001) have argued that this type of conscious knowledge about features of the target language is necessary for interlanguage development. Saito (2007, 2011) assessed the efficacy of explicit phonetic instruction in Japanese students learning English. Ultimately the goal was to improve their English pronunciation, but the training also involved an identification task, as research has showed that language perception often precedes production (Flege, 1995, 2003; Kuhl, 2000).

Saito (2011) divided participants into two groups, one of which received explicit instruction on English-specific phonemes. The training consisted of identification and discrimination portions. The former taught participants about the articulatory properties of certain English-specific sounds. Next, participants were asked to produce the sounds they were taught. The latter phase involved teaching students about what Japanese segments they might mistake for English-specific segments (e.g., /æ/ vs /a/, /f/ vs. /ɸ/, /ð/ vs. /z/, etc.). Participants then practiced producing the Japanese and English segments noting the ways in which they differed based on the articulatory characteristics they had been taught. Results revealed that learners who received the explicit training performed significantly better on measures of comprehensibility compared to the control group.

These studies indicate that explicit instruction positively affects L2 outcomes, and thus might also be effective for ASL students. Students have to learn to weight the following cues: canonical handshape, transitions, the movement envelope (Schwarz, 2000). Thus the practical significance of her movement envelope finding seems to be related to cue-weighting; skilled signers place a heavier weight on the shape of an utterance, even when they can't make out all of the letters. It is quite possible then, that the reason students struggle with fingerspelling comprehension is simply because they are erroneously giving too strong a weight to the wrong cues. For example, Geer and Keane (2014) show that students are relying very heavily on the hold segments of fingerspelling while informal testing with skilled signers suggests they weight this cue less heavily. Formal testing of this question is in preparation. What we suggest and test in this article is that, like Ylinen et al. (2010) and Giannakopoulou et al. (2013), students can be taught to adjust the weights they've assigned to various cues in fingerspelling to improve their comprehension abilities.

An additional consideration is the method by which ASL learners are taught fingerspelling, as this could impact how they learn to comprehend fingerspelling, which is clearly problematic. Li and Juffs (2015) noted that one factor affecting Japanese-accented English is the manner in which students are taught to read English characters. Instead of learning what sounds a particular letter can make individually, consonants are always paired with vowels. Instead of learning something like "the letter 'r' says /ɹ/", students learn something like "the letter 'r' says /ɹa/". This, Li and Juffs argue, trains English learners to always produce vowels after consonants because they don't have a mental representation of the phoneme itself. The case of fingerspelling may be similar. Because students learn manual representations of letters in isolation, they become too focused on seeing those specific forms without regard for how they can be influenced by the context in which they appear, namely, the letters which precede or follow them. Just as Japanese English students have to un-learn this CV pattern for correct English pronunciation, ASL students have to learn to focus on aspects of fingerspelling other than just the static portions of the signal. That is what this training program aims to remedy.

The goal of this study is to answer the following research questions.

1. Will students who receive explicit training improve more than students receiving implicit fingerspelling training?
2. Will explicit training have more, less, or equal effect on fingerspelling comprehension in the two experimental conditions – holds-only versus transitions- only – and the control conditions?

The next section outlines our methodology for this experiment.

2. Methods

2.1 Participants

Eighteen students in their third-semester of language-learning participated in this experiment in exchange for course credit. The class was split into two groups; one received the explicit fingerspelling training, while the other received the implicit training. Groups were balanced for grades received in students' highest completed level of ASL at the time of study. Table 1 presents additional group characteristics. As this table demonstrates, groups were not only balanced with respect to grades received in their previous ASL course, but also with respect to age, age of acquisition, and gender.

Characteristic	**Explicit training**	**Implicit training**
Gender	f = 8, m = 1	f = 6, m = 3
ASL 2 grade	A=3, A-=2, B+ =3, B=1	A=3, A-=2, B+ =3, B=1
ASL required for major	2	1
Avg. age ASL acquisition	22	21
Avg. age at time of study	24	24

Table 1: Participant characteristics by group

2.2 Delivery

Using the survey delivery system *Qualtrics*, participants entered a unique study identification number to complete a pre- and post-test. These tests included the same content and basic design as those used in Keane and Geer (2016). Data were collected over a three-week period during a summer semester. Participants had one week to complete the pre-test, during which time they did not have access to their respective training. In the following week, they gained access to their training. In the week after the training, they completed the post-test, again, during which time they could not access the training again.

2.3 Stimuli

Video stimuli for this project consisted of fingerspelling video clips from the corpus collected and annotated by Keane and colleagues (Keane, 2014; Keane et al., 2013b, 2015, 2013a). In their corpus, Keane and colleagues identified stable hand configurations, which they termed apogees, as periods in which the velocity of the hand was zero or approached zero (see §3.2.2 of Keane, 2014, for more detailed information about segment identification in this corpus). In other words, the portion of the signal where the hand is as still as possible is the hold portion. The portions in between held postures are transition segments. Raw video footage was modified in the following ways.

- slowed to half-speed (all)
- black screen masks inserted into the transition portions of each of the clips to create the *holds-only* condition (30 items)
- black screen masks inserted into the hold portions of each of the clips to create the *transitions-only* condition (30 items)

The stimuli were all slowed to half speed for this project to avoid a floor-effect. Before data collection for the study reported in Geer and Keane (2014) began in earnest, several students volunteered to test project stimuli. At normal speed, students were unable to comprehend fingerspelled stimuli, even without masking. To combat this issue but still examine relevant cues in the fingerspelling stream, video clips were slowed to facilitate student testing. Crucially, half-speed stimuli retain their timing properties. For example, consider Figure 5. In this particular token, the hold portions are longer than the transition portions. Even when the item is slowed to half speed, this remains true. Instead of -C- lasting 149ms, however, its duration would be double that, or 298ms. This is important because previous work asserted that the reason transitions were most important for fingerspelling comprehension is because they are temporally longer (Wilcox, 1992). As Keane and colleagues have shown, this is not always the case and in fact, as this token exemplifies, sometimes the duration of the holds is greater than that of the transitions (Keane, 2014; Keane et al., 2013b, 2015).

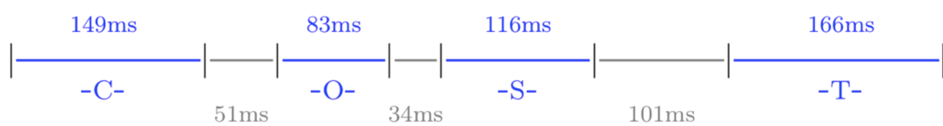


Figure 5: Timeline depicting the proportion of time spent in transition and hold segments of the fingerspelled utterance C-O-S-T. Blue segments, indicated with a letter label, represent holds and gray segments represent transitions. For this particular token, holds are appreciably longer than transition segments. The average duration of hold segments in this token is 128.5ms and the average duration for transition segments is 62ms.

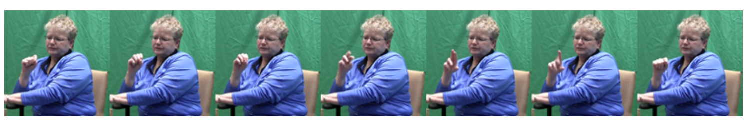
The pre- and post-tests included a total of 94 video clips. To generate this list, common four- to six-letter words were extracted from the CELEX database (Baayan et al., 1995). Of these, four served as practice items for participants to familiarize themselves with the task before it actually began. The first full block of 15 tokens had no masking. These served as control items. The second full block consisted of 30 hold only items and the third full block had 30 transition only items. The final block consisted of another 15 control items.

2.4 The training programs

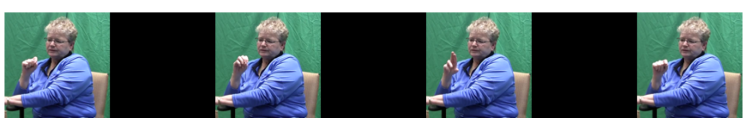
This section details the two training programs participants received. Recall that DeKeyser stipulates that training is explicit when students are made aware of a rule for some aspect of their L2, or given information which would allow them to infer the rule on their own. In the study described here, the explicit training is one in which rules are stated for students, who can then consider them consciously. The implicit training, while still focused on fingerspelling, does not allow students to even infer rules about the structure of fingerspelling or the environments which condition certain types of phonetic variation.

2.4.1 The explicit training program: Experimental group

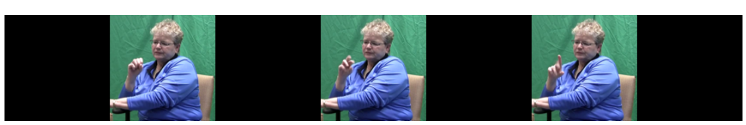
In between their pre- and post-tests, participants in this training group received an explicit fingerspelling training. There were two main sections. The first included information about the structure of fingerspelling. For example, consider the images in Figure 6. With these still images, the training drew participants' attention to the the differences between hold and transition segments. In hold portions of the signal, the video is clear and little-to-no movement, identified by blurring, is visible. Transition segments do exhibit blurring and the configuration of the hand shows features of the previous and subsequent letters. This means one could potentially use transition information to predict subsequent letters, leading to faster lexical recognition. In Figure 6b the letters -S-, -O-, -R-, and -T- can be identified, while in Figure 6c the first image is slightly blurry and looks to have features of both -S- and -O-, the second with features of -O- and -R-, and the final one with -R- and -T-. These images are described in detail in the explicit training.



(a) Clear (control



(b) Holds only

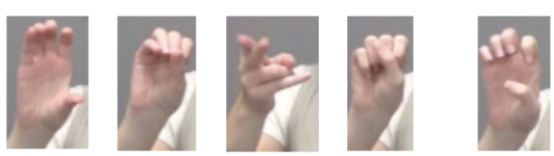


(c) Transitions only

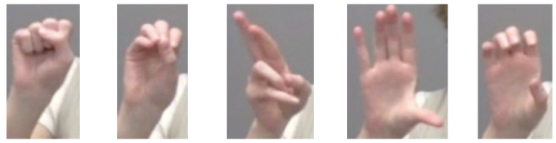
Figure 6: Example token in still images, extracted from videos, of the word S-O-R-T. The clear video (a) presents participants with unmodified stimuli – all portions of the signal are present. The holds only condition (b) presents stimuli in which the frames of transition are masked and the

transitions only condition (c) provides the opposite; the frames in which a posture is held are masked.

The second part of the training focused on describing the types of phonetic variation students might encounter when they see fingerspelling. Consider the two words in Figure 7. Both include a U-R bigram. To make this otherwise hard-to-differentiate letter combination more distinguishable, this signer adds ulnar deviation (sideways flexing of the wrist toward the pinky) and supination (upward rotation) of the forearm, which, taken together, indicate the appearance of -U- and -R-. In both the 7a and 7b examples, neither letter is realized in its canonical form, but this extra movement makes it clear both letters are intended, as it does not occur with either letter individually, nor with the combination -R- and -U-. This movement epenthesis is consistent with Brentari's proposed phonological rules governing local lexicalization. One likely explanation for the development epenthetic movement with -U- and -R- but not the reverse, -R- and -U-, is because of bigram frequency; -U- and -R- is the 50th most frequent English bigram, occurring in 0.54% of words, while -R- and -U- is far less common, occurring in only 0.128% of English words (Norvig, 2015). Other examples of gross movements to attend to in fingerspelling include epenthetic wrist flexion (bending downward) with the appearance of -Y- word internally or word finally, and a blending of features of the letters -G-, -H-, and -T-, pictured in Figure 8. The types of variation discussed in this training were selected for two reasons: previous work has noted the phenomena in question and because they were the aspects with which the first author's students reported struggling the most. This training lasted approximately 30-40 minutes.

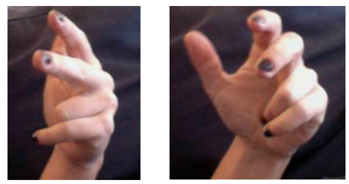


(a) C O U-R S E



(b) S O U-R C E

Figure 7: (a) C-O-U-R-S-E and (b) S-O-U-R-C-E naturally fingerspelled. (a) exhibits both ulnar deviation and supination of the forearm, while (b) exhibits the former only.



(a) (b)

Figure 8: Two images of the transition between -G-, -H-, and -T- in the fingerspelled word N-I-G-H-T. Image (a) shows the orientation shift (supination) required by the letters -G- and -H-. Image (b) shows outward rotation (pronation) of the forearm, or returning to the default, palm-out orientation. Both images show the anticipation of the letter -T-, which requires the index and middle finger to be separated, to allow for insertion of the thumb between them. Also note that in (a), the pinky remains partially extended from the previous letter -I-, while in (b) it has returned to a more flexed position. Knowledge of anticipatory and perseveratory coarticulation would help in comprehension of words with this letter combination.

2.4.2 The implicit training program: Control group

Students in the control, or implicit training group also received training in fingerspelling. Crucially, however, specific rules about fingerspelling were not explained to participants, nor could they infer one based on the training (DeKeyser, 2003).

The control group's training was meant to be a review of how fingerspelling is generally taught in ASL curricula. It included information about formation of the citation forms of each of the letters in the manual alphabet. This training included still images of each of the manual letters, except for -J- and -Z-, which were presented with short video clips since their production involves movement. In addition, participants were reminded that while most letters are produced with the palm facing away from them, there are some exceptions to this generalization. They were shown still images of two angles of letters produced with non-default palm orientation. The training concluded with reminders of how to produce double letters in fingerspelling and that fingerspelling should be produced smoothly and without jerking movements. This training, like the first, lasted approximately 30-40 minutes.

2.5 Summary of training programs

The validity of this experiment rests on the fact that these training programs are as similar as possible, except for the obvious ways in which they have to differ with respect to explicit versus implicit instruction. Both training programs had the same number of slides and took roughly the same length of time. All of the media included in the explicit training was also included in the implicit training, but participants' attention was not drawn to specific aspects of still images or videos.

3. Results

Data from all 18 participants were submitted to a mixed effects logistic regression model computed in `R`. Before detailing the findings, we first describe why this type of model was chosen, how the model is interpreted, and what the results from this model indicate (see Jaeger, 2008 for more on use of logit models in the social sciences).

Responses in this experiment were matched to the target words so the outcome could either be *correct* or *incorrect*. In the model, these values are represented by the numbers 1 and 0, respectively. The predictions from this model represent the odds of a correct answer. In a linear regression model, predictions could potentially yield values greater than 1 or less than 0. This is problematic because "greater than 100% probability of a correct answer" doesn't make sense, nor does "less than 0% probability of a correct answer". A logistic regression model constrains predictions to the 0 to 100% range.

The findings address the two research questions posed earlier in this section. The output of the mixed effects model can be found in Table 2. The text that follows will help the reader interpret this table. Visualizations are also available and will be detailed subsequently.

	coefficient (standard error)
(Intercept)	-1.26 (0.53)*
conditionClearB	1.42 (0.76)
conditionholdsOnly	0.68 (0.56)
conditiontransOnly	-0.91 (0.56)
testTypePostTest	-0.26 (0.51)
groupA	0.68 (0.50)
conditionclearB:testTypePostTest	-0.24 (0.56)
conditionholdsOnly:testTypePostTest	0.44 (0.47)
conditiontransOnly:testTypePostTest	0.18 (0.50)
conditionclearB:groupA	-0.62 (0.57)
conditionholdsOnly:groupA	-0.59 (0.43)
conditiontransOnly:groupA	-0.38 (0.47)
testTypePostTest:groupA	1.68 (0.68)*
conditionclearB:testTypePostTest:groupA	-0.58 (0.75)
conditionholdsOnly:testTypePostTest:groupA	-0.98 (0.62)
conditiontransOnly:testTypePostTest:groupA	-1.18 (0.65)
AIC	2636.29
BIC	2912.04
Log Likelihood	-1271.14
Num. obs.	2610
Num. groups: qNumber	90
Num. groups: word	90
Num. groups: group:subjCode	16
Variance: qNumber.(Intercept)	0.00
Variance: word.(Intercept)	2.01
Variance: word.conditionclearB	0.30
Variance: word.conditionholdsOnly	0.01
Variance: word.conditiontransOnly	0.10
Variance: word.testTypePostTest	0.05
Variance: group:subjCode.(Intercept)	0.52
Variance: group:subjCode.conditionclearB	0.31
Variance: group:subjCode.conditionholdsOnly	0.06
Variance: group:subjCode.conditiontransOnly	0.08
Variance: group:subjCode.testTypePostTest	0.66
Variance: Residual	1.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 2: Mixed effects logistic regression coefficient estimates and standard errors.

Again, this model yields an outcome of correct or not, which means it is evaluating whether responses match what word was actually fingerspelled. The predictors are `condition` (allclearA, holdsonly, transonly, allclearB), `testtype` (pre- or post-), and `group` (explicit-A, implicit-B) and all of their interactions. There are grouping variables that vary the intercept based on: number in the test, the word they responded to, and subject. Additionally, the effects of condition and test-type were allowed to vary by word and by subject. In Table 2, a positive value for a predictor indicates that that it is correlated with relatively more correct answers, and a negative value for a predictor correspondingly correlates with relatively fewer correct answers. The magnitude and direction correspond to the relative likelihood of a correct answer. Thus, a bigger positive number means the answers are more likely to match, and a bigger negative number means the answers are less likely to match.

The mixed effects model requires that, for each of the predictors, one level be set as the reference level. This means that the interpretation of Table 2 depends on understanding which level is set as the reference. It may also be useful to understand why each of these reference levels was selected. Please refer to Table 3.

Predictor	**Number of levels**	**Reference level**
Test-type	2	Pre-test
Group	2	Group B (implicit training)
Condition	4	Clear A

Table 3: Table of predictor variables and their reference levels

For the test-type predictor, the pre-test is the reference level because we are interested in how much change there is from pre- to post-test. This means that if participants improve from pre- to post-test, we should see a positive number, or a positive number of greater magnitude than in the pre-test. For the group predictor, the group of students who received the implicit training, called group B here, is the reference level because we were interested in learning whether group A, or the group which received the explicit training, will diverge from the assumed baseline group B represents. The condition has four levels. Clear A is the reference level, which means that all comparisons are related to the first block of the experiment. Positive values represent performance better than the reference level and negative values represent performance worse than on the reference level. We would expect then, based on Geer and Keane (2014); Keane and Geer (2016), that values would be equal or positive in the holds-only condition and negative in the transitions-only condition.

To interpret any values that meet significance, again refer to the reference level for that predictor. If an interaction is significant, it means we would expect a stronger (or weaker, depending on the sign) response than either variable on its own. Additionally, when a value is significant, it means that the effect size big enough to overcome the noise of the variability in the data. This point will be important to remember in the discussion section. The intercept represents the prediction for all reference levels, so in this model, the value is the prediction for group B in the pre-test in the clear A condition. This value is listed as significant in the table. The significance indicates that, because the value is negative, overall performance in this group was poor and correct answers were provided less than 50% of the time, a fact which is visible also in Figure 9. This has no bearing on either research question so the intercept will not be discussed further except to note that it demonstrates the extent to which students struggle with fingerspelling comprehension.

The only effect that is significant in this model is the interaction of training type and test-type. Specifically, the students with explicit training, in the post-test, did better than any other group by a statistically significant margin ($p < 0.05$). This finding speaks specifically to the first research question. Students tended to perform better in the holds-only condition, as compared to the transitions only condition, but the difference between the two did not reach significance. The trend is in the same direction that was seen by Geer and Keane (2014), and it is possible these trends could become significant with a larger number of participants. With respect to the second research question, this model suggests the training had roughly equal impact on all conditions, but performance is still weakest in the transitions only condition.

The box plot presented in Figure 9 illustrates the basic findings in this experiment which, in addition to Table 2, speak to the first research question. Here, visually, we can see that participants in both groups performed similarly on the pre-test, correctly identifying approximately 37% of fingerspelled items. The performance of participants in the implicit training group remained the same on the post-test. Those in the explicit training group improved significantly suggesting that the training was effective in improving fingerspelling comprehension scores in student learners.

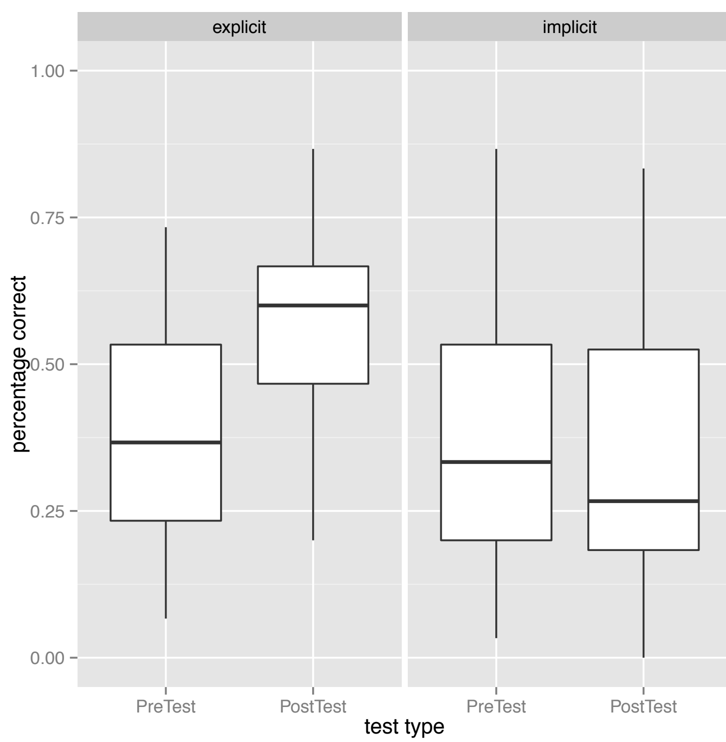


Figure 9: Box plot showing the percentage correct broken down by test-type (pre- or post-) and training group (explicit or implicit). The lines in the center of the boxes are the median, the boxes contain the inner quartile range, the whiskers contain 95% of the data observed. The difference between groups on the post-test is statistically significant.

The box plot presented in Figure 10 speaks to some of the trends mentioned above, but which were not statistically significant. One trend, which is visible in this plot, is that all participants performed better in the holds-only versus transitions-only condition. Based on previous works (Geer and Keane, 2014; Keane and Geer, 2016), we would have expected this difference to reach significance, but here it did not, perhaps due to the number of participants in each group and thus a lack of statistical power.

There are several noteworthy features of this figure. The first is consistent with what is predicted based on Figure 9, namely, that the explicit training group showed improvement from pre- to post-test, while the implicit training group did not. Examining the figure more closely, other interesting observations can be made. Notice that, in general, not only did participants in the explicit training group improve from pre- to post-test, but they performed more similarly to each other, indicated by the smaller boxes and/or shorter whiskers, than the participants in the implicit training group. Finally, notice that, while explicit training group participants improved across conditions, they still struggled the most with the transitions-only condition.

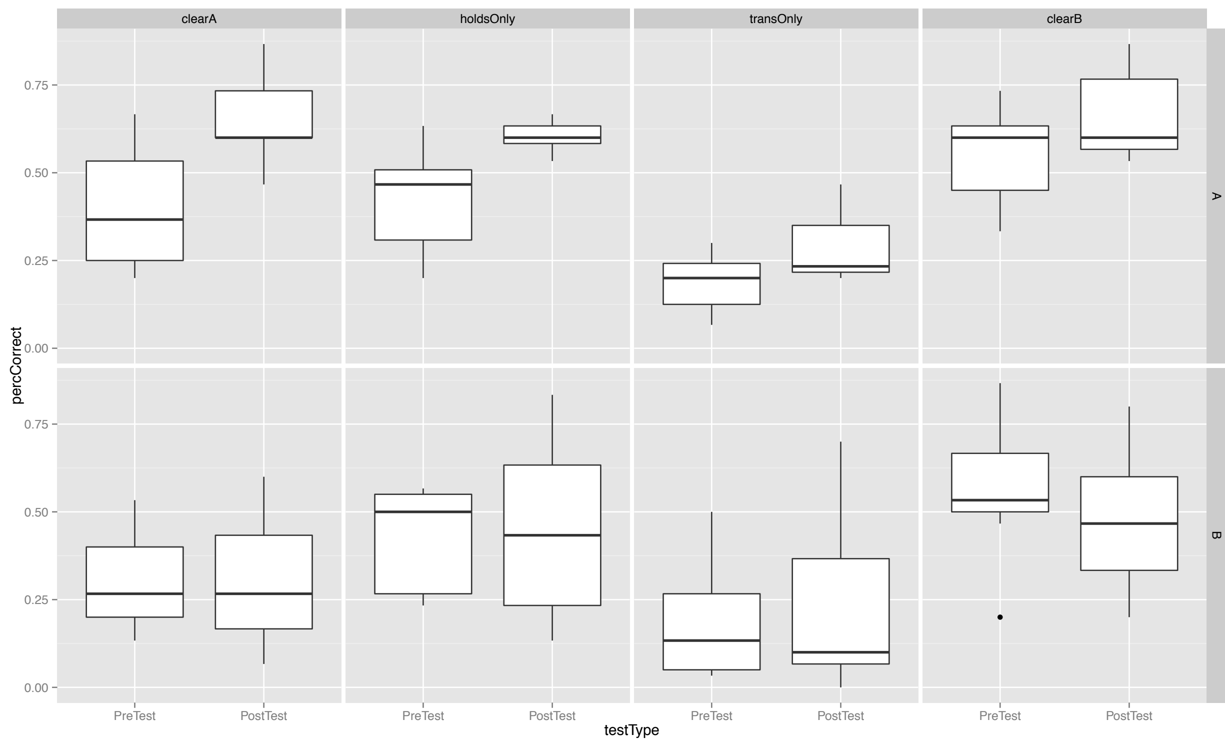


Figure 10: Box plot showing the percentage correct broken down by test-type (pre- or post-), condition (clearA, holds only, transitions only, clearB), and training group (A = explicit or B = implicit). The lines in the center of the boxes are the median, the boxes contain the inner quartile range, the whiskers contain 95% of the data observed.

4. Discussion

The goal of this study was to determine whether explicit phonetic instruction would benefit adult students acquiring ASL as a second language, with a particular focus on fingerspelling comprehension. Prior work suggests that part of their challenge with understanding fingerspelling could be related to incorrect cue-weighting (Geer and Keane, 2014; Keane and Geer, 2016). Based on these works and others which have examined explicit phonetic instruction in spoken language-learning, the study examined two research questions: (1) Can explicit phonetic instruction benefit ASL-learners on a fingerspelling comprehension task, and (2) Will it impact fingerspelling comprehension equally in various conditions?

The first question, the results indicate that explicit knowledge of fingerspelling structure and phonetic variation benefits language-learners. In addition to improved performance by a statistically significant margin, participants reported relief in having more tools to tackle fingerspelling comprehension generally. They liken it to the difference between only knowing how to read print, but being forced to read cursive. The training, they felt, helped them understand how to read "cursive" fingerspelling. They also noted the value of understanding the types of variation they might encounter in finger spelling. These results are encouraging and suggest that ASL curricula should incorporate this type of instruction.

The results that bear on the second question indicate that the training may be effective across conditions. This is indicated by a lack of a significant three-way interaction between group, test-type, and condition (refer to Table 2). In some respects, this is discouraging because we hypothesized that bringing students' attention to the information they can exploit in the transition segments would help them to improve on this condition in particular, but this was not borne out. There are several reasons which could explain this null result. It is possible there was simply too much noise that the small data set could not overcome and a significant effect went undetected. Support for this supposition comes from the fact that participants did improve in this condition, just not enough to reach significance. An alternative explanation, however, could be that the training did not provide participants enough of an opportunity to learn how to make use of cues available in the transition segments of fingerspelling. The study by Giannakopoulou et al. (2013) suggests that perhaps in addition to the training they already receive, participants would benefit more from an additional training segment which incorporated both natural and synthesized stimuli. The current training program used only natural stimuli.

The first author met with study participants after all data had been collected to discuss their experience in the experiment. They were told a follow-up study would be conducted in the near future and their feedback could be useful for preparing it. All students indicated they felt their training was beneficial to them. Students in the explicit training group said they appreciated being made aware of the "inner workings" of fingerspelling. They felt it made them better prepared to appreciate fingerspelling as a whole, rather than letter-by-letter, which is what skilled signers are doing already. This is an interesting remark because most teachers instruct students to "sound words out" rather than stating the name of each letter in their heads, because it's possible to correctly identify a long string of letters, but not understand the whole word. The goal of this instruction is to encourage students to see the whole word and not each of the letters, but it seems this specific instruction is ineffective. So, while teachers tell students to "sound out" fingerspelled words, it would seem they are not giving students the tools to learn how to do this. Worse still, as Li and Juffs (2015) suggested, how we teach students could be hindering their ability to see the forest (whole words) instead of the trees (individual letters). Students who received this training said it helped shed light on that. In addition, students expressed frustration with variation they encountered in fingerspelling prior to the study. "It's hard when one signer produces [some fingerspelled letter] like this [showing some variant of the letter], while other signers do it like that [showing some other variant of the letter]." Arming them with knowledge of the types of variation they might encounter helped them feel more secure in tackling fingerspelling comprehension tasks.

Taken together, the results of this experiment indicate that ASL teachers should begin to incorporate some type of phonetic instruction in their curricula. Students not only perform better but also feel more confident and empowered in their fingerspelling comprehension. Phonetic instruction for lexical signs may also be useful, though this has yet to be tested. The results of this experiment are encouraging that this may be the case and should be the topic of future study.

Acknowledgements

With gratitude, we acknowledge Cheyenne Matthews-Hoffman for re-creating the graphic in Figure 2 from Akamatsu (1985), the Carlota Smith Memorial Fellowship awarded to the first author for funding part of this project, and the audiences of the 9th LREC (2014) and the 12th TISLR (2016) for valuable feedback on this and closely related work. Any errors that remain are ours.

References

Akamatsu, C. T. (1985). Fingerspelling formulae: A word is more or less than the sum of its letters. In Stokoe, W. and Volterra, V., editors, SLR '83: *Sign Language Research*, pages 126–132. Linstok Press, Silver Spring.

Baayan, H. R., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX Lexical Database*. *Release 2 (CD-ROM)*. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, Pennsylvania.

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology*. MIT Press, Cambridge, MA. Couper, G. (2003). The value of explicit pronunciation syllabus in ESOL teaching. *Prospect*, 18(3):53–70.

DeKeyser, R. (2003). Implicit and explicit learning. In Long, M. and Doughty, C., editors, *Handbook of second language acquisition*, page 313-348. Blackwell, Malden, MA.

Derwing, T. M., Munro, M. J., and Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3):393-410.

Flege, J. (1995). Second language speech learning. theory, findings, and problems. In Strange, W., editor, *Speech perception and linguistic experience: Issues in cross-language research*, pages 233–277. York Press, Timonium, MD.

Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In Meyer, A. and Schiller, N., editors, *Phonetics and phonology in language comprehension and production: Differences and similarities*, pages 319 – 355. Mouton de Gruyter, Berlin.

Geer, L. and Keane, J. (2014). Exploring factors that contribute to successful fingerspelling comprehension. In Calzolari, N., Choukri, K., Declerck, T., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., and Piperidis, S., editors, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14),* pages 1905–1910, Reykjavik, Iceland. European Language Resources Association (ELRA).

Giannakopoulou, A., Uther, M., and Ylinen, S. (2013). Enhanced plasticity in spoken language acquisition for child learners: Evidence from phonetic training studies in child and adult learners of English. *Child Language Teaching and Therapy*, 29(2):201–218.

Hanson, V. L. (1981). When a word is not the sum of its letters: Fingerspelling and spelling. In *Proceedings of the 3rd National Symposium on Sign Language Research and Teaching*, Haskins.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*. 59:434-446

Keane, J. (2014). *Towards and articulatory model of handshape: What fingerspelling tells us about* *the phonetics and Phonology of Handshape in American Sign Language*. PhD thesis, Department of Linguistics, University of Chicago.

Keane, J., Brentari, D., and Riggle, J. (2013a). Coarticulation in ASL fingerspelling. In *Proceedings of the North East Linguistic Society*, number 42. North East Linguistic Society.

Keane, J., Brentari, D., and Riggle, J. (2015). *The Segment in Phonology and Phonetics*, Chapter Segmentation and pinky extension in ASL fingerspelling. Wiley.

Keane, J. and Geer, L. (2016). Understanding fingerspelling comprehension. poster. *Theoretical Issues in Sign Language Research 12*; Melbourne, Australia.

Keane, J., Riggle, J., and Brentari, D. (2013b). The timing of ASL fingerspelling. Conference presentation. *Texas Linguistics Society 14*.

Kuhl, P. (2000). A new view of language acquisition. *Proceedings of the National Academy of Science*, 97:11850–11857.

Li, N. and Juffs, A. (2015). The influence of the moraic structure on L2 English syllable-final consonants. In *Proceedings from the Annual Meeting on Phonology 2014 Conference*.

Macdonald, D., Yule, G., and Powers, M. (1994). Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning*, 44:75-100.

Norvig, P. (2015). *English Letter Frequency Counts: Mayzner Revisited or ETAOIN SRHLDCU*. `http://norvig.com/mayzner.html`, Accessed: 2015- 12-16.

Padden, C. (1998). The ASL lexicon. *Sign Language and Linguistics*, 1:33–51.

Padden, C. (2006). Learning to fingerspell twice: Young signing children's acquisition of fingerspelling. In Schick, B., Marschark, M., and Spencer, P. E., editors, *Advances in the Sign Language Development of Deaf Children*, pages 189–201. Oxford University Press, New York.

Patrie, C. J. and Johnson, R. E. (2010). *Fingerspelled Word Recognition through Rapid Sequential Visual Presentation*. DawnSign Press, San Diego, CA.

Saito, K. (2007). The influence of explicit phonetic instruction on pronunciation teaching in EFL settings: The case of English vowels and Japanese learners of English. *Linguistic Journal*, 3:17–41.

Saito, K. (2011). Examining the role of explicit phonetic instruction in native- like and comprehensible pronunciation development: an instructed SLA approach to L2 phonology. *Language Awareness*, 20(1):45–59.

Schmidt, R. (2001). Attention. In Robinson, P., editor, *Cognition and second language instruction,* page 132. Cambridge University Press, Cambridge.

Schwarz, A. L. (2000). *The perceptual relevance of transitions between segments in the fingerspelled signal*. Master's thesis, The University of Texas at Austin, Austin, TX.

Stone, A., Bosworth, R., and Petitto, L. A. (2016). Perceptual sensitivity to sonority in visual language: Native signers & naïve infants. Presentation. *Theoretical Issues in Sign Language Research 12*; Melbourne, Australia.

Wilcox, S. (1992). *The Phonetics of Fingerspelling*. John Benjamins, Philadelphia.

Ylinen, S., and Antti Latvala, M. U., Vepsäläinen, S., Iverson, P., Akahane- Yamada, R., and Näätänen, R. (2010). Training the brain to weight speech cues differently: A study of Finnish second-language users of English. *Journal of Cognitive Neuroscience*, 22:1319–1332.