Registry of Interpreters for the Deaf, Inc.

Journal of Interpretation

Volume 26 | Issue 1 Article 5

University American Sign Language Learners: Longitudinal Self- and Faculty Evaluation Ratings

Jennifer S. Beal
Valdosta State University, jsbeal@valdosta.edu

Nanci A. Scheetz
Valdosta State University, nascheet@valdosta.edu

Jessica W. Trussell
Rochester Institute of Technology, jwtnmp@rit.edu

Andrew McAllister
Maryland School for the Deaf, andrewjlmcallister@gmail.com

Jason D. Listman
National Technical Institute for the Deaf, jdlnss@ntid.rit.edu

University American Sign Language Learners: Longitudinal Self- and Faculty Evaluation Ratings

Jennifer S. Beal
Nanci A. Scheetz
Valdosta State University

Jessica W. Trussell
National Technical Institute for the Deaf

Andrew McAllister
Maryland School for the Deaf

Jason D. Listman
National Technical Institute for the Deaf

Abstract

Students who are Deaf or hard of hearing (D/HH) represent a small yet diverse population of students with individual needs who often receive educational services provided by sign language interpreters and teachers of the Deaf/hard of hearing (D/HH). Many interpreters and teachers appear unprepared to model fluent American Sign Language (ASL) skills when working with D/HH students who use sign language for communication and instruction. We investigated the ASL skills of 19 interpreting and Deaf education candidates within one university preparation program at two points in time: the end of ASL I class (Time 1) and a year later at the end of ASL IV (Time 2). We used video recordings of candidates' signed renditions of a picture book, a rubric of 12 sign language indicators with five levels of proficiency across each indicator, and ratings conducted independently by the candidates and the five authors. Four of these authors were university professors in two different Deaf education/interpreting preparation programs and the fifth was a teacher at a residential school for the Deaf. Three have typical hearing and use ASL as a secondary language; two are Deaf and use ASL as their primary language. We compared candidates' self-ratings to those of the five authors. We found that candidates tended to over-estimate their skills at T1; self-ratings and author ratings increased from T1 to T2, and candidates had higher agreement with most authors at T2 compared to T1. In addition, we found differences among ratings between the university faculty and the high school teacher. We discuss these differences in our findings and address implications for evaluating and improving university candidates' ASL skills.

Introduction

Students who are Deaf or hard of hearing (D/HH) represent a small yet diverse population of students with individual needs. Based on the most recently available data collected in Fall 2013, 1.2% of students aged 3-21 years, or 78,927 students, received services under the Individuals with Disabilities Act (IDEA) category of "hearing impairment" (Deaf/hard of hearing (D/HH); U. S. Department of Education, 2016). This frequency of incidence has remained consistent from 2004 forward. Most of these students spend the majority of their instruction in a general education setting, with a small percentage (12%) in other environments such as schools for the Deaf (Schildroth & Hotto, 1995; U. S. Department of Education, 2016). A large portion of D/HH students, estimated between 28-46%, use sign language for communication and instruction, either alone or paired with spoken language (Gallaudet Research Institute, 2009, 2013). For these students, interpreters and teachers of the Deaf/hard of hearing provide educational services either through interpretation of information provided by general educators or direct instruction. Approximately one-fourth of identified D/HH students in the U.S. utilize interpreters (6,839 interpreters; U. S. Department of Education, 2016); however, this number may be underestimated based on the titles used for interpreters (e.g., educational assistants; Storey & Jamieson, 2004). No statistics are currently available related to the number of teachers of the Deaf/HH, although Lou (1988) reported more than 10,000 in the U.S.

Interpreters and teachers often lack adequate American Sign Language (ASL) fluency to provide unlimited access to instruction and communication, despite comprehensive university preparation programs (Dodd & Scheetz, 2003; Schick, Williams, & Bolster 1999; Schick, Williams, & Kupermintz, 2006; Yarger, 2001). The professional accreditation organizations for interpreters and D/HH educators require them to maintain agreed-upon standards, provide effective and proficient communication (Easterbrooks, 2008), and represent fluent language models who can adjust their language use during communication and instruction for effective academic outcomes (i.e., the Collegiate Commission on Interpreter Education (CCIE); the Council for the Accreditation of Education Preparation (CAEP); the Council for Exceptional Children; and the Council on Education for the Deaf) (Haug, 2005).

ASL Second Language Acquisition

American Sign Language (ASL) is a visual language with its own grammatical structure that differs significantly from English (Neidle, Kegl, MacLaughlin, Bahan, & Lee, 2000). Meaning is carried in combinations of signs that are simultaneously composed of handshapes, orientations, locations, and movements, all paired with non-manual markers such as eye gaze, head tilt, body shift, and mouth movements (Neidle, Kegl, MacLaughlin, Bahan, & Lee, 2000). The majority of interpreter and teacher candidates use spoken English as their first language and learn ASL as a second language in their university preparation programs (Dodd & Scheetz, 2003; Bontempo & Napier, 2007; Bontempo, Napier, Hayes, & Brashear, 2014; Corbett & Jensema, 1981; Krause, Kegl, & Schick, 2008, Smith & Dicus, 2015; Stauffer, 2011; Storey & Jamieson, 2004; van Dijk, Boers, Christoffels, & Hermans, 2011; Wang, Napier, Goswell, & Carmichael, 2015; Woodward, Allen, & Schildroth, 1988; Yarger, 2001). Cummins' Linguistic Interdependence Hypothesis (Cummins, 1984, 2000) details the challenges in learning a second language, in this case the visual language of ASL, when one's first language is auditory-based spoken English.

While languages share some underlying cognitive and linguistic aspects, transfer of these aspects from one language to another may be limited when one language is not spoken or represented in writing, such as ASL (Mayer & Akatamasu, 2000; Mayer & Wells, 1996). To attain fluency, earlier access to ASL is optimal. Those who learn ASL at later ages, including interpreters and educators, tend to have limitations (Authors, under review, 2015; Chamberlain & Mayberry, 2000; McIntire & Reilly, 1988; R. L. McKee & D. McKee, 1992).

Several factors affect learners in their acquisition of ASL as a second language, referred to as 'second modality learners' (Chen-Pichler, 2009, 2011). First, the difference in modality between a spoken and auditory language (spoken English) and a visual language (ASL) requires learners to adjust to the different articulators in a signed language, the requirement of visual attention, and eye contact, eye gaze, facial expressions, pointing, and use of physical space and touch (i.e., tapping one's shoulder) (Chen-Pichler, 2009; Kemp, 1998; R. L. McKee & D. McKee, 1992). These skills are deemed characteristics of fluent signers by native Deaf signers (Lupton, 1998); non-native signers frequently omit these characteristics or produce them inaccurately (Chen-Pichler, 2009; Kemp, 1998; R. L. McKee & D. McKee, 1992), including eye gaze, handshapes, movement, non-manual markers, and vocabulary and classifier choices (i.e., depicting verbs) (Budding, Hoopes, Mueller, & Scarcello, 1995; Rosen, 2004; Taub, Galvin, Pinar, & Mather, 2008). Specifically, one sample of educational interpreters from Australia omitted the use of depiction (showing the actions of characters within motion events) and demonstrated English interference (signing in English word-order instead of using ASL structure) when interpreting in the classroom and tended to focus on a superficial rather than discourse-level representation of information (Bontempo & Hutchinson, 2011).

Other factors that affect learners' sign language fluency (ASL or British Sign Language, BSL) include the age of acquisition, the environment in which sign language was learned, the type and amount of sign language exposure, one's motivation for learning and practicing it, and individual personality traits (Bontempo, Napier, Hayes, & Brashear, 2014; Kemp, 1998; Lang, Foster, Gustina, Mowl, & Liu, 1996; Rosen, 2004). For example, high self-esteem was the largest predictor of BSL competence within a large sample of interpreters (Bontempo, Napier, Hayes, & Brashear, 2014). ASL learners may feel awkward when approaching Deaf adults for communication practice and are inhibited by their lack of confidence when using ASL (R. L. McKee & D. McKee, 1992). They may have fears of failure, rejection, and embarrassment (Lang, Foster, Gustina, Mowl, & Liu, 1996) and feel overwhelmed with the responsibility to "keep up" with the conversation (Kemp, 1998). This in turn may lead to "insufficient effort expended in using ASL outside of class" (R. L. McKee & D. McKee, 1992, p. 147). This combination of factors presents a challenge to those wishing to become fluent ASL users.

Interpreters' and Teachers' ASL Fluency

Professional and pre-professional interpreters' ASL fluency levels vary. Using the Sign Language Proficiency Interview (SLPI; Newell, Caccamise, Boardman, & Holcomb, 1983), Stauffer (2011) reported that the mean self-rating for interpreter candidates after their ASL IV course was *intermediate*. However, not much is known about SLPI ratings earned by recent Interpreter Education Preparation (IEP) graduates either 2- or 4-year programs. Some graduates take the National Interpreter Certification test (NIC), a certification exam developed by the

national Registry of Interpreters for the Deaf and the National Association of the Deaf. The NIC is overseen by the Registry of Interpreters for the Deaf (RID). Candidates who earn NIC Certification must demonstrate professional knowledge and skills that meet or exceed the minimum professional standards necessary to interpret in a variety of contexts. The NIC assesses candidates' interpretation skills but does not measure their ASL skills. It is often assumed that if interpreters pass the NIC exam, they are fluent in ASL and can interpret in a variety of contexts (although the RID and NIC do not state this). However, there is no formal rating on candidates' ASL skills separate from interpretation in order to evaluate their overall ASL competency. Furthermore, the NIC is not mandated by all states; therefore, some recent graduates from IEP programs are able to work in community and mainstream settings with minimal interpreting or signing experience.

Federal and state guidelines overseeing interpreters' ASL skills in educational settings are ambiguous. Interpreter qualifications are often decided by individual school districts, which are not generally knowledgeable about competencies required to effectively interpret academic content for Deaf students (Schafer & Cokely, 2016). In addition to the NIC certification test, there is an assessment specifically designed for interpreters working in K-12 settings, the Educational Interpreter Performance Assessment (EIPA). This assessment measures an interpreter's ability to interpret academic content. The EIPA, however, is not a certification in and of itself. The Registry of Interpreters for the Deaf does recognize educational interpreters with an EIPA rating of 4 or above as having full membership within the organization. The EIPA is a diagnostic tool that some states have adopted as a form of credential for interpreters working in educational settings (states accept an EIPA rating of 3 or above). Schick, Williams, and Bolster (1999) reported that fewer than half of a sample of 59 educational interpreters met the minimum required score on the EIPA. Within another sample of 46 interpreters, the mean score fell between *advanced beginner* and *intermediate* on the EIPA (Yarger, 2001). Even experienced, credentialed, or nationally certified interpreters are sometimes not able to effectively interpret academic information in ASL (Schick, Williams, & Kupermintz, 2006). Even with published data pertaining to interpretation competency, little is known specifically about interpreters' ASL competency.

Teachers of the Deaf/HH frequently must pass state-mandated content assessments, such as the Georgia Assessments for the Certification of Educators (GACE). Most Deaf students are served in local public schools with an itinerant teacher of the Deaf/HH who travels among schools to provide educational services to students or an educational interpreter in the general education classroom, and many with the provision of both (U. S. Department of Education, 2016). In one large sample of 870 U.S. secondary students served under IDEA's category of "hearing impairment," 52% of students who attended "regular secondary schools" used sign language (i.e., ASL or other manual communication systems), compared to 98% of those who attended schools for the Deaf, based on a parental report (Shaver, Marschark, Newman, & Marder, 2013, p. 211). Yet there is no established national standard for teachers' ASL fluency/proficiency level, and local school systems who serve Deaf students rarely require a specific level (Authors, 2015).

Many schools for the Deaf use a bilingual approach in which ASL is the language of instruction; however, only a portion require ASL proficiency levels of their teachers. For

instance, Beal-Alvarez and Scheetz (2015) reported limited responses to an e-mail survey that indicated sixteen states do not require a specific level of ASL proficiency for teachers through the ASLPI (ASLPI; Gallaudet University, 2014) or the SLPI (Newell, Caccamise, Boardman, & Holcomb, 1983), while eight do. These assessments are interactive conversational measures of candidates' ASL fluency that require a testing fee and provide results weeks or months after candidates complete their preparation programs. This prolonged timeline means that teaching candidates lose valuable preparation time for their classes, and that incumbent faculty in these programs are unaware of the candidate's areas of weakness. Upon preparation program exit, Curle and Jamieson (2011) reported that about half of teacher candidates were at a *beginner* level and about half were at an *intermediate* or *advanced* level based on faculty and student self-evaluations. Dodd and Scheetz (2003) provide survey results from a sample of 110 teachers of the Deaf/HH which indicate that a large majority of teachers complete ASL courses in their teacher preparation programs. Woodward and Allen (1987) reported 140 of 1,888 teachers surveyed use ASL in the classroom; however, there was no documentation of their skill levels. Other reviews (e.g., Goodman, 2006; Jones & Ewing, 2002) provide preparation program overviews but no data on pre-service teachers' ASL fluency. Only one published study specific to a teacher of the Deaf/HH found that the sole teacher in the study was aware of errors in her sign production but appeared unaware of how to match her communication to that of her preschool students (Erting, 1988).

Overall, the results of studies on professional interpreters and interpreter and teacher candidates suggest that most are not prepared to provide optimal communicative access to academic information in the classroom via ASL and need ongoing professional development to become fluent signers (Bontempo & Hutchinson, 2011; Bontempo & Napier, 2007; Dean & Pollard, 2001; Lang, Foster, Gustina, Mowl, & Liu, 1996; Schick et al., 1999; Storey & Jamieson, 2004; Yarger, 2001). While interpreter certification requires ongoing training, the responsibility for ensuring sign language proficiency prior to the provision of educational services for Deaf children appears to rest on university preparation programs.Currently, about 45 interpreter and 61 Deaf education university preparation programs exist in the U.S. with a minority requiring pre-admission ASL fluency (Carter, 2015; Goodman, 2006), unlike program requirements in other countries, such as Australia and Britain (Humphries & Allen, 2008; Jacobowitz, 2005; Napier, 2004; Swaney, 2015). Those that do require a fluency/proficiency level tend to do so near or at the end of the program and use measures developed in-house or externally through interview-format assessments such as the ASLPI (Gallaudet University, 2014) or the SLPI (Newell, Caccamise, Boardman, & Holcomb, 1983; see Authors, 2015, for a review).

Criterion-Based Assessments

Currenlty, there are no available criterion-based assessments for sign language fluency that provide feedback in a timely manner to guide candidate training and program changes. (Bontempo & Napier, 2007; Wang et al., 2015). Informal methods of sign language assessment frequently involve the use of a rubric with numerical ratings across specific components, which are efficient and give immediate feedback (Authors, 2015; Easterbrooks & Huston, 2008; Lupton, 1998; Wang, Napier, Goswell, & Carmichael, 2015), but may reveal a greater "degree of uncertainty and subjective judgment" (Wang, Napier, Goswell, & Carmichael, 2015, p. 1).

Rubrics that define ASL fluency (Lupton, 1998) may serve as a progress-monitoring tool across university preparation programs. Based on evaluations by early and native Deaf signers, Lupton (1998) identified the following characteristics as indicative of ASL fluency: facial expression, body movement, acting out, creating a picture, appropriate speed, no mouthing, clear fingerspelling, and appropriate eye contact. These characteristics are included in the Learning Assessment of ASL Proficiency Rubric created by the National Consortium of Interpreter Education Centers (Beldon, 2012), which rates candidates' abilities across indicators as *inappropriate, inconsistent, appropriate,* or *consistent and appropriate*.

Similarly, the Signed Reading Fluency Rubric (SRFR; Easterbrooks & Huston, 2008) includes the following indicators frequently found in signed narratives: speed, facial expression, body movement, sign space, sign movement, fingerspelling, use of space, role taking, eye gaze, directionality, pronominalization, and classifiers (semantic and size-and-shape-specifiers, or SASS). The complete rubric can be viewed in Easterbrooks and Huston (2008); minor adaptations to the rubric for the present study's picture book task included changing "text" to "picture book," "read-aloud" and "reading" to "narrative retell," and "reader" to "signer" (Authors, 2015). These indicators are evaluated as *not observed*, *emerging, beginning, developing*, or *mature/fluent* with expanded descriptions for each indicator at each level. The use of rubrics includes both benefits and challenges. The SRFR is the only ASL rubric with published data on its reliability and validity across raters (see Authors, 2015; Easterbrooks & Huston, 2008 for reviews). Rubrics can provide quick assessment measures and results directly related to specific tasks, such as narratives, with thorough descriptions of skills to be evaluated, and include the option for self-evaluation and triangulation across raters.

Challenges to rubrics include rater subjectivity (McNamara, 2000), including individual preferences for signing style (Lupton, 1998; Wang, Napier, Goswell, & Carmichael, 2015), individual interpretation of rating criteria (McNamara, 1996), intra- and inter-rater reliability (Bachman, 1990), validity, and time-consuming training (Wang, Napier, Goswell, & Carmichael, 2015). Wang and colleagues (2015) provide an overview of inter-rater training procedures and note that variability among raters will remain even with rater training. For instance, Wang and colleagues analyzed the evaluations of two interpreter educators and one interpreter practitioner, all of whom were typical hearing, native signers (the authors did not indicate whether the raters were children of Deaf adults, or CODAS) with national accreditation for translation and interpreting, of an interpreting team's simultaneous English to Auslan (Australian Sign Language) and Auslan to English interpretation of a conference. They reported higher ratings and higher agreement between the interpreter educators. Wang et al. concluded that the educators "were more experienced in testing an assessment against criteria, scales, and standards" (p. 11), that assessment was a regular part of their work, and that the competencies on the rubric were directly related to their teaching units. They also noted that the two educators had exposure to a broad spectrum of interpreting abilities into which they placed the interpreters they evaluated. In contrast, the third rater, who gave significantly lower scores, likely had less experience with both evaluation and breadth of signers, paired with "high personal standards" (p. 12). However, Wang et al. did not provide intra-rater reliability results.

Seeking a quick and efficient assessment of candidates' fluency at the end of their ASL IV course, but a year prior to their program completion, Authors (2015) investigated candidate

self-evaluation and faculty evaluation of candidates' narrative renditions of a picture storybook using the twelve indicators of the SRFR (Easterbrooks & Huston, 2008). Candidates' self-ratings ranged from *emergent* to *fluent*, with a mean rating of *developing*. Candidates tended to self-rate their fluency higher overall than faculty about half of the time, similar to previous findings of signers' overestimation of their fluency (Lang et al., 1996; McDermid, 2009; Schick et al., 1999; Stauffer, 2011; Yarger, 2001), although Stauffer (2011) reported a significant moderate to strong correlation between candidates' self-ratings and their instructors' ratings. Compared to faculty ratings, candidates showed more variability in their self-ratings of signing speed, movement, and use of pronominalization (i.e., establishing and referring to objects/people in space; Authors, 2015). Their ratings were similar to those of faculty for body movement and eye gaze (Authors, 2015). These findings align with those of R. L. McKee and D. McKee (1992).

We investigated the ASL fluency of candidates in a Bachelor's degree university program that requires interpreting and Deaf education candidates to complete four ASL courses that utilize specific curricula (*Signing Naturally*), supplemental activities (online or professional videos, activities and games in class, books, and articles), and instructional approaches (i.e., classes were conducted in the target language from the beginning of instruction) similar to some other university preparation programs in the U.S. (R. L. McKee & D. McKee, 1992; Rudser, 1988; Swaney, 2015). Candidates were evaluated using formative and summative assessments across all four ASL courses, including the use of video-linked professor and self-evaluation comments via online software, so that candidates could compare their comments to those of their peers and professors. ASL I and ASL II were compacted into intense one-month sessions (June and July, respectively) and ASL III and ASL IV occurred across Fall and Spring semesters, respectively. Concurrent courses beyond ASL I-IV differ between interpreter and teacher candidates due to the specific skills needed for each professional role. Interpreter candidates complete two supplementary ASL-related courses: Linguistics of American Sign Language, and Fingerspelling, Numbers, and Classifiers. Teacher candidates complete a Manual Communication course. This course introduces various grammatical features of ASL and examines conceptually accurate sign language in English word order (Conceptually Accurate Signed English and Contact Signing).

At the end of ASL IV in the present university program, all candidates are required to take the ASLPI and achieve a minimum of a 1+ (on a scale of 1 to 5) prior to the onset of their interpreting or student teaching internships. Furthermore, those obtaining the minimum 1+ are required to re-take the ASLPI with the expectation that they earn a 2 prior to the end of the semester. These scores were established in collaboration with administrators at schools for the Deaf and the Deaf Education and Interpreting Advisory Board, composed of university faculty and students as well as community stakeholders. Additionally, candidates are encouraged to participate in the university ASL Club and ASL social activities outside of the classroom with Deaf community members and their program peers, similar to other programs (Rosen, 2004). Authors (2015) provide additional curriculum details.

Rationale for this Study

Previously, researchers identified gaps in interpreters' ASL fluency at the end of their preparation programs and even after they worked multiple years within the profession; limited data are available for teachers of the Deaf/HH (Authors, 2015). Most published assessment results are from formal measures without a self-evaluation component (e.g., ASLPI, EIPA, SLPI). While longitudinal investigations of interpreters' and teachers' fluency across time have been suggested (Authors, 2015), they remain glaringly absent in the literature. Informal longitudinal assessments conducted at two points in time during a preparation program, by way of self- and faculty evaluations, may update interpreter and teacher candidates' ongoing ASL-learning goals related to both their university courses and their professional development (Bontempo & Napier, 2007). Additionally, outside of trained evaluators for formal assessments such as the ASLPI and the SLPI, Deaf stakeholder perceptions are noticeably absent from L2 learners' skill ratings.

Our aims in the present study are as follows: 1) Investigate changes in signed narrative renditions of a picture book at the end of ASL I (T1) and again at the end of ASL IV (T2) for interpreter and Deaf education candidates. 2) Investigate author inter-rater reliability of the evaluation rubric across candidates and authors. 3) Identify areas in need of change within the university preparation program.

Methods

Participants.

The University Institutional Review Board (IRB) approved the research protocol for this study. Participation in this study was voluntary, unrelated to course grades, and candidates signed a consent form to participate. A total of 19 female candidates in the interpreting (n = 11) and teacher preparation (n = 8) programs participated in this study (see Table 1). All candidates completed a background form, including questions regarding age, gender, university program (interpreting or Deaf education), childhood language used at home, preferred language, other languages used, years signing, and self-rated ASL fluency level. All candidates fell between 20-28 years of age and reported spoken English as their primary communication mode and ASL as another language used. All candidates except two reported that their childhood home language was spoken English. At the end of ASL I, candidates' length of signing time ranged from nine months to two years, with the exception of candidate 3, and most (n = 15) learned ASL as a result of their university courses. At the end of ASL I (T1), five candidates rated their ASL fluency as *basic*, twelve as *conversational*, one as *fluent*, and one *in need of remediation*. At T2, all candidates had completed ASL I, II, III, and IV; eleven (mostly interpreter candidates) completed Numbers, Fingerspelling and Classifiers; and six of the Deaf education candidates completed Manual Communication.

Table 1. Candidates' background information at T1 based on self-report.

Candidate	Age	Program	Other Languages	Years;months signing	Self-rated fluency level
1	18-20	Interpreter	ASL	0;9	basic
2	18-20	Interpreter	ASL	0;9	conversational
3	21-24	Interpreter	ASL	8	fluent
4	21-24	Interpreter	ASL	0;9	conversational
5	21-24	Interpreter	ASL	0;9	conversational
6	21-24	Interpreter	ASL	0;9	conversational
7	21-24	Interpreter	ASL, Spanish	0;9	conversational
8	21-24	Interpreter	ASL	3	conversational
9	21-24	Interpreter	ASL	0;9	conversational
10	21-24	Interpreter	ASL	0;9	basic
11	18-20	Interpreter	ASL	4	conversational
12	25-28	Deaf Ed	ASL	0;9	basic
13	21-24	Deaf Ed	ASL	0;9	conversational
14	21-24	Deaf Ed	ASL	0;9	conversational
15	21-24	Deaf Ed	ASL	0;9	conversational
16	21-24	Deaf Ed	ASL	0;9	basic
17	21-24	Deaf Ed	ASL	0;9	conversational
18	21-24	Deaf Ed	ASL, Spanish	0;9	in need of remediation
19	21-24	Deaf Ed	ASL	2	basic

DATA COLLECTION.

We replicated the procedures of Beal-Alvarez and Trussell (2015). The researchers provided an outline of the study to potential candidates near the end of ASL I and ASL IV courses. At the end of both courses candidates: 1) completed a background form (see Table 1); 2) previewed a picture book (with no printed text) and rehearsed signing it (*A Day in the Park,* Dinardo, 1988); 3) individually video-recorded their signed rendition of the picture book in a quiet location in the university library or in their home (for online learners); and 4) watched their video and conducted a self-evaluation using an adapted version of the SRFR (Easterbrooks & Huston, 2008). We selected the SRFR because it has published data related to this narrative storybook task (see Authors, 2015) and previous rater reliability (Authors, 2015; Easterbrooks & Huston, 2008), contains comprehensive indicator and level descriptions of the elements we expected to see in candidates' narrative renditions, provides an opportunity for self-evaluation and ratings by multiple assessors, and is efficient to administer and score, with an average of 20 to 30 minutes per storybook video. Using paper or digital copies, candidates circled a level of fluency across each indicator for their self-evaluation. No additional information related to specific indicators

was provided to candidates. The intent was to assess candidates' abilities to independently comprehend and recognize ASL components at the end of both ASL I and ASL IV. Candidates uploaded their videos online to an invited location for author access and turned in paper or digital copies of the background form and completed rubrics.

All authors work with Deaf individuals and use ASL on a daily basis. Beal and Trussell have fourteen and nine years of experience, respectively, teaching Deaf students at the Pre-K-12 and university levels and having Advanced Plus ratings on the SLPI (Newell, Caccamise, Boardman, & Holcomb, 1983). Scheetz has 30 years of experience teaching Deaf students and has been a nationally certified interpreter for 25 years. Beal and Trussell teach Deaf education courses at the university level. McAllister and Listman are Deaf, use ASL as their primary language, and have Master's degrees in Deaf Education. The fourth author has five years of experience teaching high school English at a school for the Deaf and the fifth author has eight years of teaching experience, ASLTA provisional certification, and teaches university ASL and interpreting courses.

Data analysis.

Because each candidate rendered the picture book at x (T1) and x2 (T2), there was a total of 38 videos. We followed Quinto-Pozos' (2007) method of blinding authors to the condition of the video (whether T1 or T2) by randomizing the videos across condition and dividing them among the authors for coding. Each video was watched and independently rated by two hearing authors and one Deaf author using the adapted SRFR rubric, following the procedures of Authors (2015). Additionally, due to variation in ratings between the hearing authors and the fourth Deaf author, a subset of videos was rated by the fifth Deaf author. For each video, we rated each of the twelve indicators across five fluency levels and calculated the total score across the indicators (total of 12 indicators x 5 levels per indicator = 60 possible points). Author 1 coded 16 T1 and 13 T2 videos (29 total), Author 2 coded 12 T1 and 18 T2 videos (30 total); Author 3 coded 15 T1 and 11 T2 videos (26 total); Author 4 coded all T1 and T2 videos (38 total); and Author 5 coded 2 T1 and 10 T2 videos.

To investigate changes in interpreter and Deaf education candidates' signed narrative renditions of a picture book across time we used candidate self-evaluations and evaluations by hearing and Deaf university faculty at T1 and T2 and calculated means and standard deviations (SDs), and compared them to self- and author-evaluation scores using correlations and Analysis of Variance (ANOVA) statistics. To investigate inter-rater reliability of the evaluation rubric across candidates and authors we compared individual evaluation scores using correlations and ANOVA statistics. The first three authors completed intra-rater agreement by re-coding three randomly selected videos approximately three months after the initial ratings and compared initial and subsequent ratings, while the fourth author re-coded six purposefully selected videos based on differences between his ratings and those of the first three authors (Wang, Napier, Goswell, & Carmichael, 2015). Finally, to identify areas in need of change within our university preparation program, we analyzed specific rubric indicators that differed among candidates and authors. Below, we present our findings.

Results

In this study, we first investigated changes in candidates' signed narrative renditions of a picture book at the end of ASL I (T1) and again at the end of ASL IV (T2). We used candidate self-evaluations and evaluations by hearing and Deaf university faculty at T1 and T2 and calculated means, SDs, and ANOVA statistics. At T1, candidates' self-evaluation scores ranged from 12 to 43 (out of a possible 60). At T2, candidates' self-evaluations ranged from 12 to 41. While candidates' self-evaluation mean score increased by 4.7 points from T1 (M = 29.7, SD = 9.42) to T2 (M = 34.4, SD = 7.15), this difference was not significant (F [1, 40] = 3.33, p = .075). However, candidates' scores at each time strongly and significantly correlated (r = .517, p = .016).

Author ratings for the candidates at T1 ranged from 8 to 42. At T2, authors' ratings ranged from 9 to 45. At T1, all author means were lower than candidate means, but only the fourth Deaf author scored candidates significantly lower than their self-evaluations (see Table 3). Similarly, at T2, all author means were lower than candidate means. Both Deaf authors rated candidates significantly lower than their self-ratings. None of the correlations between authors and candidates for T1 or T2 were significant with the exception of the first author at T2 (N = 13, r = .553, p =.050). While candidates' mean score changed by nearly five points, the fourth Deaf author's mean score between time periods changed by less than one point (he rated all videos). The hearing authors, who rated different candidates at each time point, changed by 4 to 8 points across time periods.

Table 3. Mean comparison results at T1 and T2 for individual author ratings and candidate self-evaluations.

Author	F	p	Cand. M	Cand. SD	Author M	Author SD
			T1			
JB [a]	F [1, 28] = 2.23	.146	29.0	9.35	23.7	10.17
NA [a]	F [1, 19] = 1.04	.321	29.6	9.51	24.6	12.21
JT [a]	F [1, 26] = 1.13	.298	27.9	10.80	23.9	9.08
AM [b]	F [1, 36] = 17.05	.000*	29.1	9.72	16.4	9.23
			T2			
JB	F [1, 24] = .838	.369	36.9	2.30	35.3	5.60
NS	F [1, 34] = .731	.399	34.1	7.64	31.8	8.34
JT	F [1, 20] = .480	.496	32.2	9.24	29.8	6.52
AM	F [1,40] = 68.00	.000*	34.4	7.15	17.14	6.4
JL [b, c]	F [1, 18] = 8.26	.010*	37.2	2.35	28.4	9.40

[a] Typical hearing; [b] Deaf; [c] Results only for T2; * p < .05.

We investigated intra- and inter-rater reliability using the evaluation rubric across candidates and authors. For intra-rater reliability, Author 1's difference in ratings ranged from 3 to 4 points (out of a total of 60). Author 2's difference ranged from 4 to 8 points. Author 3's difference ranged from 1 to 3 points. Author 4's differences ranged from 1 to 4 points. We investigated differences in ratings among authors. Because each author did not rate each candidate, we compared ratings for pairs of authors on the candidates they did rate. There were

no significant differences between any of the hearing authors' ratings for videos at T1 or T2 (see Table 4) or the hearing authors and the fifth Deaf author.

Table 4. Mean comparison results at T1 and T2 by author pairs.

Author 1	Author 2	F	*p*	Author 1 *M*	Author 2 *M*
T1					
JB [a]	NA [a]	F [1, 12] = .041	.843	24.29	23.00
JB	JT [a]	F [1, 18] = .600	.449	24.3	21.4
JB	AM [b]	F [1, 30] = 8.88	.006*	24.3	14.9
NA	JT	F [1, 10] = .001	.982	27.0	27.2
NA	AM	F [1, 22] = 5.31	.031*	25.8	15.6
JT	AM	F [1, 28] = 3.62	.068	24.5	18.1
T2					
JB	NS	F [1, 18] = 1.03	.324	35.7	32.5
JB	JT	F [1, 4] = .214	.668	34.0	30.3
JB	AM	F [1, 24] = 43.04	.000*	35.3	18.6
NS	JT	F [1, 14] = .127	.727	30.9	29.6
NS	AM	F [1, 34] = 31.45	.000*	31.8	17.5
JT	AM	F [1, 20] = 45.06	.000*	29.8	14.8
JB	JL [b, c]	F [1, 12] = 4.34	.059	34.2	28.9
NS	JL	F [1, 14] = 3.49	.083	37.4	28.8
JT	JL	F [1,8] = 3.37	.104	34.0	27.2
AM	JL	F [1, 18] = 13.44	.002*	15.6	28.4

[a] Typical hearing; [b] Deaf; [c] Results only for T2; * $p < .05$.

At T1, however, the hearing authors significantly differed in their ratings compared to the fourth Deaf author (the classroom teacher) in every case, with the exception of the third author (hearing university faculty). At T2, the two Deaf authors' ratings again differed significantly, with the fourth author consistently rating candidates lower than the fifth author (F [1, 18] = 13.44, p = .002). To triangulate our data, and similar to Wang et al.'s inter-rater investigation, we looked for patterns by rubric indicator within candidates' and authors' differences in ratings at each time point (see Figures 1 and 2).

Figure 1. Ratings by all four authors and self-ratings across candidates at T1.

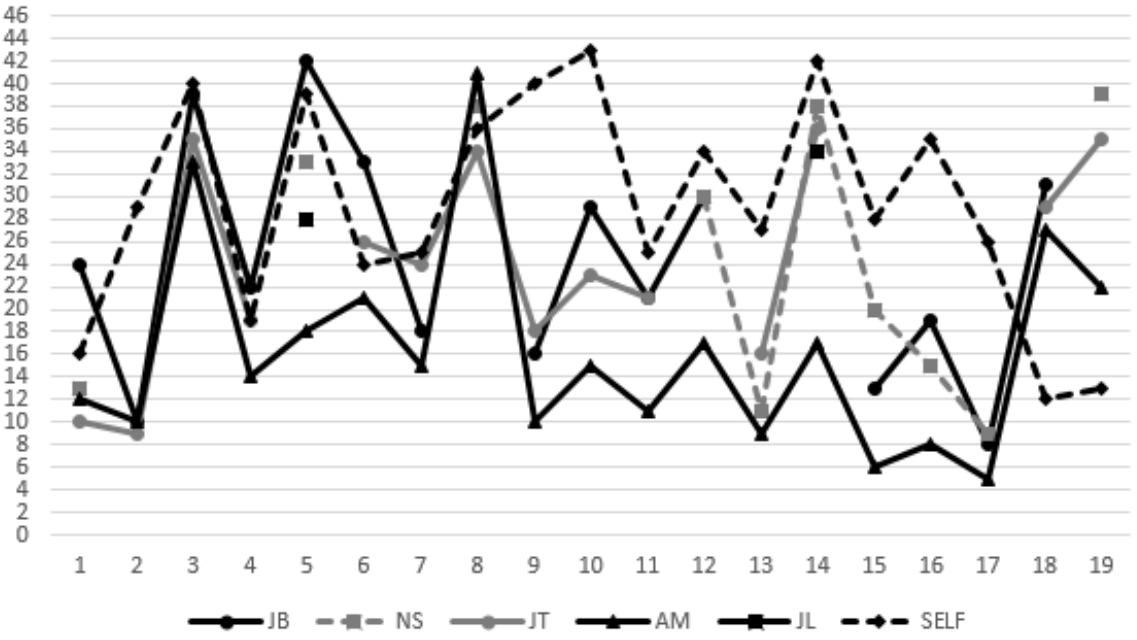
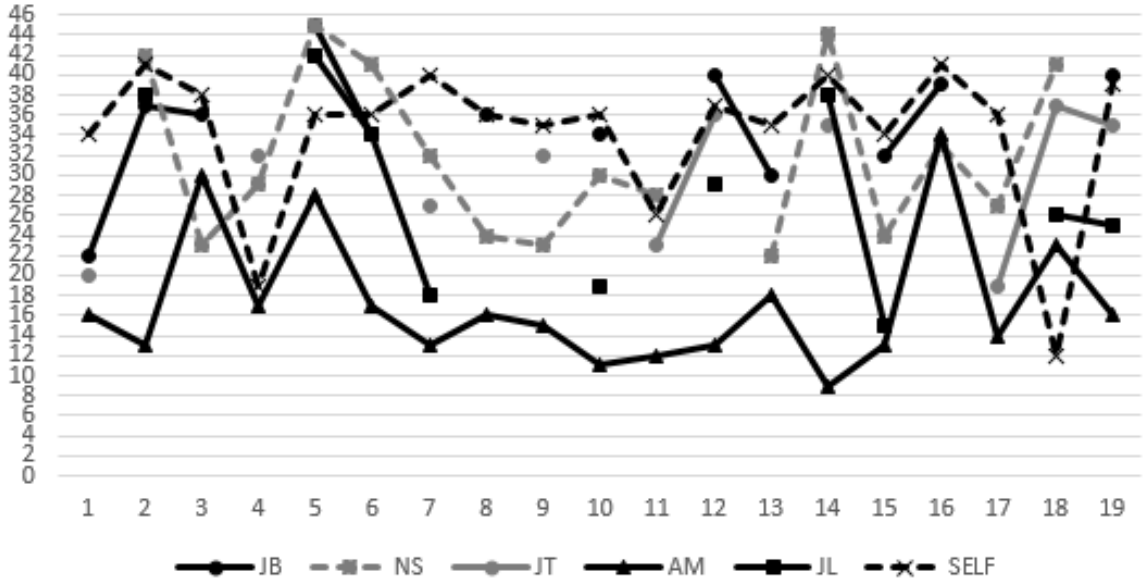


Figure 2. Ratings by all four authors and self-ratings across candidates at T2.



At T1, the authors appeared to have a similar shape and a difference in scale in their graphed ratings across candidates. For example, all authors rated candidate 2 low and candidate 3 high, followed by a low rating for candidate 4 and a high rating for candidate 8. At T2, author ratings were more diverse across candidates and graphed lines crossed more frequently, representing more variation in ratings. This correlates with the significant differences reported above.

Finally, we analyzed specific rubric indicators that differed among candidates and authors (see Table 5). At T1, candidates demonstrated inflated ratings across all indicators, defined as a score higher than all author ratings. Two or more authors disagreed by two or more levels on every indicator for at least one candidate at T1. Candidates' inflated ratings at T2 differed from T1, although role taking and pronominalization fell within the top three most frequently inflated indicators at both time points. At T2, sign space, speed, and pronominalization fell within the top three most frequent disagreements, while the authors had fewer disagreements for facial expression from T1 to T2. From T1 to T2 there was a decrease in disparity between candidate self-evaluations and author ratings for facial expression, body movement, sign movement, and classifiers, suggesting possible increase in students' fluent use of these indicators. However, there were no noticeable changes from T1 to T2 for speed, sign space, use of space, role taking, directionality, eye gaze (all components of constructed action) and there was an increase in disparity for pronominalization. These are areas for possible direct instruction across ASL courses.

Table 5. Candidate and author disagreements by SRFR indicator at T1 and T2.

SRFR Indicator	# of Candidate and Author Disagreements T1[a]	# of Candidate and Author Disagreements T2	# of Author disagreements T1[b]	# of Author disagreements T2
Speed	3	2	7	15
Facial expression	5	2	9	12
Body movement	6	0	5	12
Sign space	5	6	9	13
Sign movement	5	0	4	7
Fingerspelling	7	3	4	11
Use of space	3	2	5	14
Role taking	5	4	4	10
Eye gaze	2	2	4	12
Directionality	2	2	5	13
Pronominalization	5	9	6	14
Classifiers (i.e., depicting verbs)	5	3	4	12

[a] Candidates' self-evaluation score was higher than all author ratings; [b] Two or more authors disagreed by 2 or more levels.

Discussion

We aimed to investigate changes in interpreter and Deaf education candidates' signed narrative renditions across time. It appears some candidates over-estimate their ASL fluency, as demonstrated by differences in self- and author ratings at each Time, similar to previous findings (Authors, 2015; Lang, Foster, Gustina, Mowl, & Liu, 1996; McDermid, 2009; Schick, Williams, & Bolster, 1999; Stauffer, 2011; Yarger, 2001). In some cases, indicators for which candidates overestimated their fluency aligned with those for which authors had higher disagreement, suggesting similar areas of difficulty in ASL evaluation, as reported by R. L. McKee and D. McKee (1992). Although candidates were not directly asked about their justifications when self-rating, one might speculate that candidates became more aware of ASL components across time

and fine-tuned their initial over-estimated ratings at the end of ASL IV. At T1, candidates may not have been aware of what they should include in their renditions, which might be remedied in the future by asking them to compare and contrast their renditions to those of fluent Deaf adults (Beal-Alvarez & Trussell, 2015). Also, candidates' T1 and T2 self-evaluation scores strongly and significantly correlated, perhaps suggesting that their opinions of their ASL fluency are consistent across time, in that those whose scores were higher at T1 also were higher at T2. In contrast to Stauffer's (2011) findings, candidate and author ratings did not correlate in the present study, with the exception of the first author at T2.

We also aimed to assess inter-rater reliability using the evaluation rubric across candidates and authors. Inclusion of Deaf raters provided the opportunity to triangulate scores across multiple raters, all of whom bring different experiences and perspectives to the assessment process. Authors had high intra-rater reliability across time (Wang et al., 2015, did not provide intra-rater reliability results, which would contribute to a comparison on rubric reliability across raters). Notably, the hearing authors and the fifth Deaf author, all of whom are university faculty, tended to have high agreement, even though the third and fifth authors had never seen the candidates sign previously, which aligns with Wang et al.'s (2015) findings. Wang et al. noted that the two university educators in their study had exposure to a broad spectrum of interpreting abilities into which they placed the interpreters they evaluated. In contrast, the fourth rater, who had given significantly lower scores, likely had less experience with evaluation of university L2 signers. It is likely that the authors in the present study who were university faculty are more experienced in assessing the ASL fluency of university learners, while the fourth author, who is a classroom teacher, may have a different perspective relative to interpreter and teacher fluency based on his experience as a K-12 educator and consumer of interpreting services. Nevertheless, all of these perspectives present a triangulated view of university candidates' fluency and instructional needs during their preparation programs. Author ratings may have been affected by their experience with candidates across courses and their preferences for individual signing styles (Lupton, 1998; Wang et al., 2015). To mitigate the effect of bias in the present study, we triangulated candidate ratings by including self-evaluations, those of two university faculty who have taught the candidates one course each at the time of data collection, and those of two university faculty and one classroom teacher who have never seen the present candidates' signing.

Finally, we analyzed candidate performance across specific rubric indicators to identify areas in need of change within our university preparation program. Data collection at the end of ASL I provided an opportunity for candidates to become aware of the difficulty of some ASL components, similar to previous results (R. L. McKee & D. McKee, 1992) and presented an opportunity for them to work on these components across their subsequent ASL and university courses, as opposed to only at the end of ASL IV, when candidates in the present program complete the ASLPI. Use of the SRFR provided efficient longitudinal feedback on candidates' performance across specific narrative indicators and two time periods so that results could drive candidates' learning and authors' program alignment with candidates' needs. Candidates appear to need increased direct instruction in ASL components, especially those that are used by fluent native signers (Aarons & Morgan, 2003; Authors, 2015; Beal-Alvarez & Trussell, 2015; Cormier, Smith, & Sevcikova, 2013; Lupton, 1998; Taub & Galvan, 2001) and that are non-existent in one's L1 (Rosen, 2004), such as the use of space for establishing and referencing

characters across narrative events (pronominalization), and the use of classifier handshapes, depiction, and constructed action (i.e., facial expression, eye gaze, and role taking). Future research on candidates' narrative renditions might compare them to native or near-native signers' renditions, such as those presented in Beal-Alvarez and Trussell (2015). Evaluations might include Deaf community members as raters, who are consumers of interpreting services, and the use of paired rating sessions, as opposed to independent ratings, to discuss how each rater, including candidates themselves, arrived at her or his respective score. Finally, while university faculty may not have control over the provision of earlier ASL exposure, we can increase ASL exposure at the university level beyond required courses and Deaf community events to Living and Learning communities (i.e., dorms in which ASL is used as the language of communication; Maltby, Brooks, Horton, & Morgan, 2016; Soldner & Szelenyi, 2008) and recruit more Deaf university students as language partners to create immersion opportunities for candidates to continuously think and interact using ASL.

A prevalent limitation relative to the current study and assessment of ASL fluency in general is the lack of an established standard definition of "fluency" or "proficiency" for university candidates (and in-service interpreting and teachers of the Deaf/HH). In the present study, authors and candidates assessed twelve indicators across five fluency levels. As candidate fluency approached mature-fluent descriptions, the authors exhibited more disagreement in ratings (T2). While each level is described in detail, lack of a singular overall fluency definition likely contributed to rater variability. Another challenge of using ASL rubrics with L2 ASL learners is that feedback is typically provided in written English, as opposed to directly in ASL; changing the feedback format via time-linked video comments that provide the opportunity for modeling in ASL might be a beneficial method of evaluation. Future investigations might evaluate the effectiveness of this format. This storybook task was insufficient for rating candidates' fingerspelling fluency, as some candidates did not use fingerspelling within this task and others spelled only a few words. The small number of candidates within this study limits generalizations outside of the present preparation program; however, it adds to previous research by extending documentation of candidates' ASL fluency both at the end of one and four ASL courses.

The present results provide a snapshot of university interpreter and Deaf education candidates' ASL acquisition as a second language and suggests future directions university preparation programs may implement to address candidates' needs. Clearly, across published research and in the present study, L2 ASL learners need exposure to and acquisition of ASL at much earlier ages than university entry (Akmeşe, 2016). Most candidates in the present study had less than one year of signing experience and limited interaction with the signing Deaf community. Recently the number of high school students who take ASL as a foreign language has increased (Rosen, 2004), which begins to address earlier ages of ASL acquisition and exposure for second language learners. However, ASL as a standard offering within K-12 instruction would address ASL acquisition for both second language learners and Deaf students for whom ASL is their first language. At the university level, preparation programs need to increase learning opportunities for ASL learners on campus.

References

Aarons, D., & Morgan, R. (2003). Classifier predicates and the creation of multiple perspectives in South African Sign Language. *Sign Language Studies 3*(2), 125-56.

Akmeşe, P. P. (2016). Examination of sign language education according to the opinions of members from a basic sign language certification program. *Educational Sciences: Theory and Practice, 16,* 1189-1225, doi:10.12738/estp.2016.4.0248.

Beal, J. S., & Faniel, K. (in press). Hearing L2 Sign Language Learners: How Do They Perform on ASL Phonological Fluency? *Sign Language Studies.*

Beal-Alvarez, J. S., & Trussell, J. (2015). Depicting verbs and constructed action: Necessary components in Deaf adults' storybook renditions. *Sign Language Studies, 16*(1), 5-29.

Bachman, L. F. (1990). *Fundamental consideration sin language testing.* Oxford, England: Oxford University Press.

Beal-Alvarez, J. S., & Scheetz, N. A. (2015). Pre-service teacher and interpreter American Sign Language Abilities: Self-evaluations and evaluations of deaf students' narrative renditions. *American Annals of the Deaf, 160*(3), 316-333.

Beal-Alvarez, J. S., & Trussell, J. (2015). Depicting verbs and constructed action: Necessary components in Deaf adults' storybook renditions. *Sign Language Studies, 16*(1), 5-29.

Beldon, E. (2012). *Learning Assessment of ASL Proficiency rubric*. National Consortium of Interpreter Education Centers. Retrieved from: http://www.interpretereducation.org/aspiring-interpreter/mentorship/mentoring-toolkit/pre-and-post-learning-assessment-of-asl-proficiency/

Bontempo. K., & Hutchinson, B. (2011). Striving for an 'A' grade: A case study in performance management of interpreters. *International Journal of Interpreter Education, 3*, 56-71.

Bontempo, K., & Napier, J. (2007). Mind the gap! A skills analysis of sign language interpreters. *The Sign Language Translator and Interpreter, 1,* 275-299.

Bontempo, K., Napier, J., Hayes, L., & Brashear, V. (2014). Does personality matter? An international study of sign language interpreter disposition. *The International Journal for Translation and Interpreting Research, 6*(1), 23-46. doi: ti.106201.2014.a02.

Budding, C., Hoopes, R., Mueller, M., & Scarcello, K. (1995). Identification of foreign sign language accents by the Deaf. In L. Byers, J. Chaiken, & M. Mueller (Eds.), *Gallaudet University Communication Forum, 4,* 1-16. Washington, DC: Gallaudet University Press.

Carter, C. A. (2015). *Current practices: Pre-admission assessment of American Sign Language and English language competency in interpreter education programs*. Master's Thesis,

Western Oregon University. Retrieved from: http://digitalcommons.wou.edu/cgi/viewcontent.cgi?article=1021&context=theses

Chamberlain, C., & Mayberry, R. (2000). Theorizing about the relationship between ASL and reading. In C. Chamberlain, J. Morford, & R. Mayberry (Eds.), *Language Acquisition by Eye* (pp. 221-259. Mahwah, NJ: Erlbaum.

Chen-Pichler, D. (2009). Sign production by first-time hearing signers: A closer look at handshape accuracy. *Cadernos de Saude, 2*, 37-50. Retrieved from: https://www.gallaudet.edu/Documents/Academic/LIN/Chen%20Pichler%202009%20Cadernos.pdf

Chen-Pichler, D. (2011). Sources of handshape error in first-time signers of ASL. In G. Mathur & D. J. Napoli (Eds.), *Deaf around the world* (pp. 96-121). New York, NY: Oxford University Press Inc.

Corbett, E., & Jensema, C. (1981). *Teachers of the hearing impaired: Descriptive profiles.* Washington, DC: Gallaudet College Press.

Cormier, K., Smith, S., & Sevcikova, Z. (2013). Predicate structures, gestures, and simultaneity in the representation of action in British Sign Language: Evidence from deaf children and adults. *Journal of Deaf Studies and Deaf Education 18*(3), 370-90.

Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. San Diego, CA: College Hill Press.

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire.* Clevedon, England: Multilingual Matters Ltd.

Curle, D. M., & Jamieson, J. R. (2011). Differences in student characteristics in face-to-face and online cohorts in a teacher preparation program in education of the deaf and hard of hearing. *MERLOT Journal of Online Learning and Teaching, 7*(4), 502-514.

Dean, R. K., & Pollard, R. Q. (2001). Application of demand-control theory to sign language interpreting: Implications for stress and interpreter training. *Journal of Deaf Studies and Deaf Education, 6*(1), 1-14.

Dinardo, J. (1988). *A day in the park.* New York, NY: Modern Publishing.

Dodd, E. E., & Scheetz, N. A. (2003). Preparing today's teachers of the deaf and hard of hearing to work with tomorrow's students: A statewide needs assessment. *American Annals of the Deaf, 148*(1), 25–30.

Easterbrooks, S. R. (2008). Knowledge and skills for teachers of individuals who are deaf or hard of hearing: Initial set revalidation, *Communication Disorders Quarterly, 30*(1), 12-36. doi: 10.1177/1525740108324043.

Easterbrooks, S. R., & Huston, S. (2008). The signed reading fluency of students who are deaf/hard of hearing. *Journal of Deaf Studies and Deaf Education*, *13*, 37-54.

Erting, C. J. (1988). Acquiring linguistic and social identity: Interactions of deaf children with a hearing teacher and deaf adult. In M. Strong (Ed.), *Language learning and deafness* (pp. 192-219). New York, NY: Cambridge University Press.

Gallaudet Research Institute. (2009). Regional and national summary report of data from the 2011-2012 Annual Survey of Deaf and Hard of Hearing Children and Youth. Washington, DC: GRI, Gallaudet University.

Gallaudet Research Institute. (2013). Regional and national summary report of data from the 2011-2012 Annual Survey of Deaf and Hard of Hearing Children and Youth. Washington, DC: GRI, Gallaudet University.

Gallaudet University (2014). *American Sign Language Proficiency Interview (ASLPI)*. Retrieved from: http://www.gallaudet.edu/asldes/aslpi.html

Goodman, C. A. (2006). Teaching manual communication to preservice teachers of the deaf in an accredited comprehensive undergraduate teacher preparation program. *American Annals of the Deaf, 151*(1), 5-15.

Haug, T. (2005). Review of sign language assessment instruments. *Sign Language & Linguistics, 8*, 61-98.

Humphries, T., & Allen, B. M. (2008). Reorganizing teacher preparation in deaf education. *Sign Language Studies, 8*(2), 160-180.

Jacobowitz, E. L. (2005). American Sign Language teacher-preparation programs in the United States. *Sign Language Studies, 6*(1), 76-110.

Jones, T., & Ewing, K. (2002). An analysis of teacher preparation in deaf education programs approved by the Council on Education of the Deaf. *American Annals of the Deaf, 147*(5), 267-271.

Kemp, M. (1998). Why is learning American Sign Language a challenge? *American Annals of the Deaf, 143*(3), 255-259.

Krause, J. C., Kegl, J. A., & Schick, B. (2008). Toward extending the Educational Interpreter Performance Assessment to Cued Speech. *Journal of Deaf Studies and Deaf Education, 13*(3), 432-450. doi: 10.1093/deafed/enm059.

Lang, H. G., Foster, S., Gustina, D., Mowl, G., & Liu, Y. (1996). Motivational factors in learning American Sign Language. *Journal of Deaf Studies and Deaf Education, 1*(3), 202-212.

Lou, M. P. (1988). The history of the education of the deaf in the United States. In M. Strong (Ed.), *Language learning and deafness* (pp. 75-112). New York, NY: Cambridge University press.

Lupton, L. (1998). Fluency in American Sign Language. *Journal of Deaf Studies and Deaf Education, 3*(4), 320-328.

Maltby, J. L., Brooks, C., Horton, M., & Morgan, H. (2016). Long term benefits for women in a science, technology, engineering, and mathematics living-learning community. *Learning Communities Research and Practice, 4*(1), Article 2. Retrieved from: http://washingtoncenter.evergreen.edu/lcrpjournal/vol4/iss1/2

Mayer, C., & Akatamasu, C. T. (2000). Deaf children creating written texts: Contributions of American Sign Language and signed forms of English. *American Annals of the Deaf, 145*(5), 394-403.

Mayer, C., & Wells, G. (1996). Can the linguistic interdependence theory support a bilingual-bicultural model of literacy education for deaf students? *Journal of Deaf Studies and Deaf Education, 1*(2), 93-107.

McDermid, C. (2009). Social construction of American Sign Language-English interpreters. *Journal of Deaf Studies and Deaf Education, 14*(1), 105-130. doi: 10.1093/deafed/enn012.

McIntire, M., & Reilly, J. (1988). Nonmanual behaviors in L1 and L2 learners of American Sign Language. *Sign Language Studies, 61,* 351-375.

McKee, R. L., & McKee, D. (1992). What's so hard about learning ASL? Students' and teachers' perceptions. *Sign Language Studies, 75,* 129-158.

McNamara, T. (1996). *Measuring second language performance.* New York, NY: Longman.

McNamara, T. (2000). *Language testing.* Oxford, England: Oxford University Press.

Napier, J. (2004). Sign language interpreter training, testing, and accreditation: An international comparison. *American Annals of the Deaf, 149*(4), 350-359.

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (2000). *The syntax of American Sign Language: Functional categories and hierarchical structure.* Cambridge, MA: The MIT Press.

Newell, W., Caccamise, F., Boardman, K., & Holcomb, B. R. (1983). Adaptation of the Language Proficiency Interview (LPI) for assessing sign communicative competence. *Sign Language Studies, 41,* 311–352.

Quinto-Pozos, D. (2007). Why does constructed action seem obligatory? An analysis of classifiers and the lack of articulator-referent correspondence. *Sign Language Studies 7*(4), 458-506.

Rosen, R. S. (2004). Beginning L2 production errors in ASL lexical phonology. *Sign Language and Linguistics, 7,* 31-61.

Schafer, T., & Cokely, D. (2016). Understanding the challenges of interpreters working with children in K-12 settings: Needs assessment. National Consortium of Interpreter Education Centers. Retrieved from: http://www.interpretereducation.org/resources/need-assessments/

Schick, B., Williams, K., & Bolster, L. (1999). Skill levels of educational interpreters working in public schools. *Journal of Deaf Studies and Deaf Education, 114*(2), 144-155.

Schick, B., Williams, K., & Kupermintz, H. (2006). Look who's being left behind: Educational interpreters and access to education for deaf and hard-of-hearing students. *Journal of Deaf Studies and Deaf Education, 11*(1), 3-20. doi:10.1093/deafed/enj007.

Schildroth, A., & Hotto, S. (1995). Race and ethnic background in the annual survey for deaf and hard of hearing children and youth. *American Annals of the Deaf, 140,* 96-99.

Shaver, D. M., Marschark, M., Newman, L., & Marder, C. (2013). Who is where? Characteristics of deaf and hard-of-hearing students in regular and special schools. *Journal of Deaf Studies and Deaf Education, 19*(2), 203-219. doi: 10.1093/deafed/ent056.

Smith, C., & Dicus, D. (2015). A preliminary study on interpreting for emergent signers. *Sign Language Studies, 15*(2), 202-224.

Soldner, M., & Szelenyi, K. (2008). A national portrait of today's living-learning programs. *The Journal of College and University Student Housing, 35*(1), 14-31.

Stauffer, L. K. (2011). ASL students' ability to self-assess ASL competency. *Journal of Interpretation, 21*(1), Article 7. Retrieved from: http://digitalcommons.unf.edu/joi/vol21/iss1/7

Storey, B. C., & Jamieson, J.R. (2004). Sign language vocabulary development practices and internet use among educational interpreters. *Journal of Deaf Studies and Deaf Education, 9*(1), 53-67. doi: 10.1093/deafed/enh012.

Swaney, M. G. (2015). *American Sign Language Curriculum and Materials.* Master's Thesis, University of Tennessee, Knoxville, TN. Retrieved from http://trace.tennessee.edu/utk_gradthes/3516.

Taub, S., & Galvan, D. (2001). Patterns of conceptual encoding in ASL motion descriptions. *Sign Language Studies 1*(2), 175-200.

Taub, S., Galvin, D., Pinar, P., & Mather, S. (2008). Gesture and ASL L2 acquisition. In R.M. de Quadros (Ed.), *Sign Languages: Spinning and Unraveling the Past, Present, and Future from TISLR9* (pp. 639-651). Petropolis, Brazil: Editora Arara Azul.

U. S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs. (2016). *38th Annual Report to Congress on the implementation of the Individuals with Disabilities Education Act, 2015*, Washington, DC. Retrieved from: https://ed.gov/about/reports/annual/osep/2016/parts-b-c/38th-arc-for-idea.pdf

van Dijk, R., Boers, E., Christoffels, I., & Hermans, D. (2011). Directionality effects in simultaneous language interpreting: The case of sign language interpreters in the Netherlands. *American Annals of the Deaf, 156*(1), 47-55.

Wang, J., Napier, J., Goswell, D., & Carmichael, A. (2015). The design and application of rubrics to assess signed language interpreting performance. *The Interpreter and Translator Trainer,* 1-21. doi: 10.1080/1750399X.2015.1009261.

Woodward, J., & Allen, T. (1987). Classroom use of ASL by teachers. *Sign Language Studies, 54*, 1-10.

Woodward, J., Allen, T., & Schildroth, A. (1988). Linguistic and cultural role models for hearing-impaired children in elementary school programs. In M. Strong (Ed.), *Language learning and deafness* (pp. 184-191). New York, NY: Cambridge University Press.

Yarger, C. C. (2001). Educational interpreting: Understanding the rural experience. *American Annals of the Deaf, 146*(1), 16-30.

Appendix: Tables & Figures

Table 1. Candidates' background information at T1 based on self-report.

Candidate	Age	Program	Other Languages	Years; months signing	Self-rated fluency level
1	18-20	Interpreter	ASL	0;9	basic
2	18-20	Interpreter	ASL	0;9	conversational
3	21-24	Interpreter	ASL	8	fluent
4	21-24	Interpreter	ASL	0;9	conversational
5	21-24	Interpreter	ASL	0;9	conversational
6	21-24	Interpreter	ASL	0;9	conversational
7	21-24	Interpreter	ASL, Spanish	0;9	conversational
8	21-24	Interpreter	ASL	3	conversational
9	21-24	Interpreter	ASL	0;9	conversational
10	21-24	Interpreter	ASL	0;9	basic
11	18-20	Interpreter	ASL	4	conversational
12	25-28	Deaf Ed	ASL	0;9	basic
13	21-24	Deaf Ed	ASL	0;9	conversational
14	21-24	Deaf Ed	ASL	0;9	conversational
15	21-24	Deaf Ed	ASL	0;9	conversational
16	21-24	Deaf Ed	ASL	0;9	basic
17	21-24	Deaf Ed	ASL	0;9	conversational
18	21-24	Deaf Ed	ASL, Spanish	0;9	in need of remediation
19	21-24	Deaf Ed	ASL	2	basic

Table 2. Individual candidate self-evaluation and authors' rubric ratings across T1 and T2 out of a total possible 60 points.

Cand.	T1						T2					
	Self	JB	NS	JT	AM	JL[a]	Self	JB	NS	JT	AM	JL
1	16	24	13	10	12	-	34	22	-	20	16	-
2	29	10	-	9	10	-	41	37	42	-	13	38
3	40	39	-	35	33	-	38	36	23	-	30	-
4	19	22	-	19	14	-	19	-	29	32	17	-
5	39	42	33	-	18	28	36	45	45	-	28	42
6	24	33	-	26	21	-	36	34	41	-	17	34
7	25	18	-	24	15	-	40	-	32	27	13	18
8	36	-	38	34	41	-	36	36	24	-	16	-
9	40	16	-	18	10	-	35	-	23	32	15	-
10	43	29	-	23	15	-	36	34	30	-	11	19
11	25	21	-	21	11	-	26	-	28	23	12	-
12	34	30	30	-	17	-	37	40	-	36	13	29
13	27	-	11	16	9	-	35	30	22	-	18	-
14	42	-	38	36	17	34	40	-	44	35	9	38
15	28	13	20	-	6	-	34	32	24	-	13	15
16	35	19	15	-	8	-	41	39	33	-	34	-
17	26	8	9	-	5	-	36	-	27	19	14	-
18	12	31	-	29	27	-	12	-	41	37	23	26
19	13	-	39	35	22	-	39	40	-	35	16	25

(-) Indicates no data. [a] Deaf author.

Table 3. Mean comparison results at T1 and T2 for individual author ratings and candidate self-evaluations.

Author	F	*p*	Cand. *M*	Cand. *SD*	Author *M*	Author *SD*
			T1			
JB [a]	F [1, 28] = 2.23	.146	29.0	9.35	23.7	10.17
NA [a]	F [1, 19] = 1.04	.321	29.6	9.51	24.6	12.21
JT [a]	F [1, 26] = 1.13	.298	27.9	10.80	23.9	9.08
AM [b]	F [1, 36] = 17.05	.000*	29.1	9.72	16.4	9.23
			T2			
JB	F [1, 24] = .838	.369	36.9	2.30	35.3	5.60
NS	F [1, 34] = .731	.399	34.1	7.64	31.8	8.34
JT	F [1, 20] = .480	.496	32.2	9.24	29.8	6.52
AM	F [1,40] = 68.00	.000*	34.4	7.15	17.14	6.4
JL [b, c]	F [1, 18] = 8.26	.010*	37.2	2.35	28.4	9.40

[a] Typical hearing; [b] Deaf; [c] Results only for T2; * $p < .05$.

Table 4. Mean comparison results at T1 and T2 by author pairs.

Author 1	Author 2	F	*p*	Author 1 *M*	Author 2 *M*
		T1			
JB [a]	NA [a]	F [1, 12] = .041	.843	24.29	23.00
JB	JT [a]	F [1, 18] = .600	.449	24.3	21.4
JB	AM [b]	F [1, 30] = 8.88	.006*	24.3	14.9
NA	JT	F [1, 10] = .001	.982	27.0	27.2
NA	AM	F [1, 22] = 5.31	.031*	25.8	15.6
JT	AM	F [1, 28] = 3.62	.068	24.5	18.1
		T2			
JB	NS	F [1, 18] = 1.03	.324	35.7	32.5
JB	JT	F [1, 4] = .214	.668	34.0	30.3
JB	AM	F [1, 24] = 43.04	.000*	35.3	18.6
NS	JT	F [1, 14] = .127	.727	30.9	29.6
NS	AM	F [1, 34] = 31.45	.000*	31.8	17.5
JT	AM	F [1, 20] = 45.06	.000*	29.8	14.8
JB	JL [b, c]	F [1, 12] = 4.34	.059	34.2	28.9
NS	JL	F [1, 14] = 3.49	.083	37.4	28.8
JT	JL	F [1,8] = 3.37	.104	34.0	27.2
AM	JL	F [1, 18] = 13.44	.002*	15.6	28.4

[a] Typical hearing; [b] Deaf; [c] Results only for T2; * $p < .05$.

Table 5. Candidate and author disagreements by SRFR indicator at T1 and T2.

SRFR Indicator	# of Candidate and Author Disagreements T1[a]	# of Candidate and Author Disagreements T2	# of Author disagreements T1[b]	# of Author disagreements T2
Speed	3	2	7	15
Facial expression	5	2	9	12
Body movement	6	0	5	12
Sign space	5	6	9	13
Sign movement	5	0	4	7
Fingerspelling	7	3	4	11
Use of space	3	2	5	14
Role taking	5	4	4	10
Eye gaze	2	2	4	12
Directionality	2	2	5	13
Pronominalization	5	9	6	14
Classifiers (i.e., depicting verbs)	5	3	4	12

[a] Candidates' self-evaluation score was higher than all author ratings; [b] Two or more authors disagreed by 2 or more levels.

Figure 1. Ratings by all four authors and self-ratings across candidates at T1.

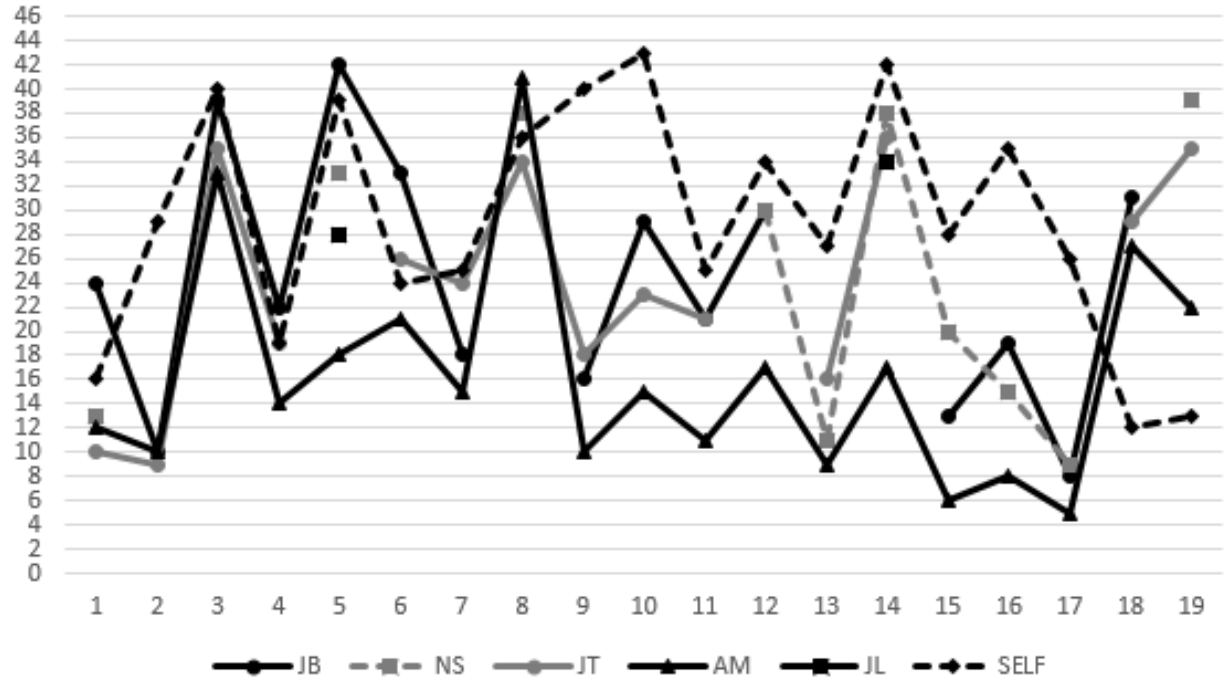


Figure 2. Ratings by all four authors and self-ratings across candidates at T2.