1. Introduction

Producing narrative is a skill that must be acquired in order to become a competent language user in a first as well as in a second language. In children, narrative competence is often acquired late (Karmiloff-Smith, 1979), in part because children only start stringing sentences together regularly, once they have acquired the linguistic skills necessary to form sentences. Cohesive discourse, however, depends on more than the ability to employ resources from phonology, morphology, and syntax. An important additional skill is the pragmatics of taking into account what the addressee knows. One place where this necessity is obvious is in the domain of reference to persons and objects. When children first begin narrating, they often produce referring expressions that are uninformative from the addressee’s point of view. This is because they do not adhere to adult norms of varying referring expressions as a function of accessibility in the discourse.

In adult second language learners, however, such pragmatic competences are in place from the beginning. Such learners have often had years of experience creating cohesive narratives in their first language. Consequently, we might expect narrative competence in second language acquisition to depend only on the ability to produce appropriate linguistic means. Studies of this topic, however, have shown that this expectation is not borne out. Beginning and intermediate second language learners produce narratives that are referentially unlike the native norm. This pattern has been attested in a variety of source-target language pairs, leading scholars to assume the existence of some universal interlanguage stage for reference. However, past studies in the field have focused on language pairs in the spoken modality. In order to claim universality, we need to see the interlanguage pattern attested in the visual modality as well.

To date, only one study has investigated patterns of reference tracking in a signed second language. Bel, Orteils, and Morgan (2014) tested proficient second language learners of Catalan Sign Language, and found some evidence of the pattern described for spoken L2 learners. However, because Bel et al.’s
subject group consisted of highly proficient learners, we do not yet know how the performance of beginning and intermediate signed L2 learners compares with these findings. The present study approaches this problem by examining reference tracking in beginning/intermediate L2 learners of American Sign Language in a controlled narrative task.

2. Creating cohesion in ASL
2.1 Referring expressions in native ASL

Like spoken languages, ASL has a range of referential expressions. This section discusses the ones relevant to the present work. First, nouns in ASL may be nominal signs (e.g. BOY, GIRL, POPSICLE, etc.) or they may be fingerspelled (e.g. P-A-R-K). Unlike English, ASL mainly uses bare nouns, and uses context to determine how given or accessible a given referent is. Signers can also modify nouns by accompanying them with points in signing space. Under some analyses of ASL, spatially modified nouns have been treated as definite noun phrases or noun phrases with locative adverbials (Neidle, Kegl, MacLaughlin, Bahan, & Lee, 2000). Across various linguistic analyses, however, bare nouns appear to be the preferred option in narratives (Swabey, 2002; Frederiksen & Mayberry, submitted). Space also plays a role in the production of ASL pronouns. An anaphoric pronoun is essentially a point to a locus in the signing space, which has been established previously and therefore is associated with a referent. Because of this, ASL anaphoric pronouns generally refer to a specific referent, rather than a class of potential referents, as is the case in spoken languages (Sandler and Lillo-Martin 2006). In addition to overt pronouns, ASL uses zero anaphora, or null pronouns extensively (Klima & Bellugi, 1979, Lillo-Martin, 1986). Because ASL has different types of verbs, there are also different types of zero anaphora that correspond to the null arguments of these verb types. Typically, studies of ASL have discussed plain verbs, agreement verbs, and locative verbs\(^1\). Plain verbs do not agree with either subject or objects, so when a plain verb is used without overt arguments, it is the context that determines the subject and/or object of the verb. Agreement verbs can take both subject and object agreement\(^2\) (Padden, 1988), although recent studies of other sign languages have suggested that verb agreement might be optional (de Beuzeville, Johnston, & Schembri, 2009). The arguments of an agreement verb may be both overt and null. With null arguments, the beginning and ending points of the verb correspond to loci associated with referents. That way, the addressee can establish the intended subject/object. Alternatively, the signer can use his or her own locus the starting or ending point of the verb. In

\(^1\) Locative verbs behave similarly to agreement verbs, and for the purposes of this study, we group them together
\(^2\) Although some scholars have begun to argue against the movement of agreeing verbs as actual grammatical/linguistic agreement (e.g. Liddell, 2000)
this case, the signer can either role-shift, or rely on the discourse context to make the verb arguments clear. Role-shifting is a common phenomenon in signed languages. This entails the signer indicating a previously established locus by the orientation of their body or gaze (Winston, 1991). This process changes the reference of the first person locus to a 3rd person referent (Friedman, 1975).

The final type of referential devices that we will discuss is classifiers and classifier predicates. Classifiers are handshapes and movements that stand in for a referent by virtue of certain of its features (Frishberg, 1975). Classifier predicates are classifier handshapes used to depict motion or location. The most commonly used grouping of classifiers is as size and shape specifiers, handle classifiers, and semantic classifiers. Size and shape specifiers (SASSes) depict the form or outline of entities. These are nominal or adjectival in nature (Johnston & Schembri, 2007). Handle classifiers show how an object is held by a human hand, and are often used to designate the object referent in agreement and locative verbs. Semantic classifiers represent groups of entities with the same handshape, based on some semantic similarity of the referents. For example, ASL has one semantic classifier for upright human being, and another for vehicles. For that reason, semantic classifiers are less clearly iconic than the other types (Supalla, 1982).

### 2.2 Using referring expressions to track referents in narratives

We have seen in the previous section that a range of nominal and predicative reference types is available to ASL signers to construct sentences. We have not yet, however, discussed the principles underlying the choice of referential expression at various points in a narrative. We believe that this question is crucial to our understanding of ASL discourse structure and coherence.

In order for a narrative to be cohesive, the sender needs to establish and uphold reference to persons and objects that figure in their story. We know from spoken language research that speakers achieve cohesion by choosing the form of their references according to a set pattern. Although this is an oversimplification, we can describe this pattern as a preference for referring with full linguistic form when the discourse entity is new or has been out of focus for a while, and a preference for referring with less linguistic material, when referring to accessible or given discourse entities (Chafe, 1976; Givón, 1983; Ariel, 1988). In English, full referential forms include nouns (e.g. *the dog*), and less full referential expressions are pronouns (e.g. *it*) or zero anaphora (*Ø*), as in the example in 1).

1) ‘the dog found a bone and Ø ate it’

This pattern has been attested in many languages across the world. However, the topic has only been investigated in a few signed languages (Wulf, Dudis,
Bayley, & Lucas, 2002; Swabey, 2002, 2011; Morgan, 2006; McKee, Schembri, McKee, & Johnston, 2011), despite the fact that signers as well as speakers have the need to create cohesive stories. Looking at Auslan and New Zealand Sign Language (McKee et al., 2011) and American Sign Language (Wulf et al., 2002, Swabey, 2002, 2011), these studies found the first evidence that signers are more likely to use null subjects when referring to highly accessible referents. Frederiksen and Mayberry (submitted) also investigated ASL reference tracking and found that native signers use different classifier types and different predicate types for different referential purposes.

However, much remains to be accounted for regarding referential cohesion in signed languages. Given the many differences between the signed and spoken modalities, reference tracking in signed language might diverge from the patterns observed in spoken languages. This also suggests that second language learners whose languages span the two modalities may face different challenges than unimodal bilinguals.

3. How learners track reference in their second language

The construction of cohesive discourse in a second language has been of interest in both generative and functional approaches to acquisition. As mentioned previously, L2 learners struggle to become native-like in this area, despite the finding that native reference tracking follows similar principles cross-linguistically. This is perhaps unexpected, since we might expect learners to transfer their L1 principles in discourse, as we know they do in other linguistic areas. However, evidence from a large number of studies has shown that reference in a second language tends to be overly explicit compared to the native norm. Although individual differences exist (e.g. Ahrenholz, 2005) this pattern has been found in L2 varieties of many languages, for example Mandarin (Polio, 1995), Swedish (Gullberg, 2003), French (Gullberg, 2003, 2006), Korean (Jung, 2004), English (Williams, 1988; Munoz, 1995), Spanish (Saunders, 1999), Italian (Chini 2005; Sorace & Filiachi, 2006), and Japanese (Yoshioka, 2008). The fact that this pattern emerges regardless of source-target language pairs (Hendriks, 2003) has led to an assumption that over-explication is a property of interlanguage in general, rather than of particular language parameters. However, a variety of explanations for why interlanguage looks this way has been proposed, including considerations of markedness and transfer strategies (Munoz, 1995; Polio, 1995), communicative strategies (Williams, 1988), process-oriented learner strategies (Carroll & Lambert, 2003, Gullberg, 2003, 2006), and vulnerability at the syntax-semantics interface (Sorace & Filiachi, 2006, Sorace, 2011).

Intertwined with the explanations for L2 over-explicitness is the question of what should be considered over-explicit reference tracking. Some studies have found that learners overuse lexical noun phrases compared to pronouns (e.g. Polio, 1995; Jung, 2004; Gullberg, 2003, 2006), where others find an over-reliance on pronouns compared to null elements (e.g. Munoz, 1995, Sorace &
Filiachi, 2006). Still others have found that learners overuse both nouns and pronouns compared to native speakers (Saunders, 1999; Yoshioka, 2008). It is clear that the native norm for reference tracking plays a role in determining the type of patterns that learners may produce. In addition, the learner patterns may change as a function of proficiency. For example, it has been suggested that over-explicit reference tracking only appears at the intermediate stage of second language acquisition (Ahrenholz, 1995; Hendriks, 2003).

Different reasons for the particular structure of interlanguage aside, it seems clear that some properties of reference tracking are shared between different learner varieties. It also seems clear that the similarities are not dependent on the type of source and target languages. However, at present the vast majority of studies of L2 reference tracking have looked at two spoken languages. Studies of signed source/signed target language pairs are non-existent, and to the best of our knowledge, only one study has investigated a spoken-signed language pairing (Bel et al., 2014). However, such extensions of the L2 reference tracking work are essential for determining the robustness of the phenomenon, and the extent to which it is modality dependent. Bel et al. investigated how the narratives of L2 signers of Catalan Sign Language compared with native signer narratives. Their study indicated that the learners generally overused pronouns compared to the native signers, but not necessarily in the context of maintained reference. However, Bel et al. studied proficient signers, rather than signers on the intermediate or beginning levels. The latter groups, however, are where most spoken language studies have found a prominence of overt reference. Clearly, further studies are needed to map out the patterns of cross-modal second language reference tracking in more detail, both in different signer groups and in different sign languages. In the present study we begin to address these questions by investigating a wider range of referential expressions in the narratives of beginning to intermediate signers of American Sign Language.

4. Method
4.1 Participants

Second Language Learners: 8 hearing second language learners of American Sign Language participated in the experiment (age range: 18-22 years). The ASL level of the participants ranged from beginning learners (around 25 weeks of instruction) to low-level intermediate (around 45 weeks of instruction). We compare the results for the L2 learners with those of a native signer control group. These data have been reported in detail elsewhere (Frederiksen and Mayberry, submitted). In the present study we discuss the natives’ data only in as far as it is relevant to interpreting the learners’ results.

Native Signers: 8 Deaf native signers (age range: 20-55 years) participated. Each had begun learning ASL from birth in the home. Because some had hearing parents there was variation in the amount of ASL input they received from birth. The participants varied in their use and knowledge of English.
4.2 Stimuli

Each signer retold four simple wordless stimulus stories. The stories were based on Karmiloff-Smith’s (1979) balloon stories. Each story consisted of six causally related events and involved a main character, a secondary character, and a featured object. We kept the story line identical across all stimuli and deliberately chose a simple narrative structure that allowed for introduction as well as maintenance and reintroduction of entities, while ensuring similar patterns across stories.

4.3 Procedure

The participants first gave informed consent and then filled out a questionnaire about their language background. After receiving instructions about the procedure, the participants first watched a story on a laptop, then retold it, until all four stories had been told. We balanced the presentation of stimulus items across participants following a Latin Squares design. Each narrative was recorded on video.

4.4 Transcription and annotation

We transcribed the signed narratives using ELAN (Crasborn & Sloetjes, 2008), and identified clause boundaries, largely based on the presence of predicates. We then identified all referring expressions in the narratives and noted which entity each referred to. Next, we determined the discourse status of each referent, following Gullberg (2006). We considered every first mention of a referent as ‘introduced’. Subsequent mentions of a previously introduced referent were coded only if they were clausal subjects. Such subjects were coded as ‘maintained’ if they had been mentioned in any position in the previous clause and as ‘reintroduced’ if they appeared following a clause where the referent in question had not been mentioned at all. All other referring expressions were not considered for analysis. Finally, we noted the reference type of each remaining referring expression as a Noun, Pronoun, Zero Anaphora or Classifier.

5. Results

The eight L2 signers’ story retellings produced a total of 504 references. As a comparison, the native signers produced 450 references retelling the same stories. The L2 signers therefore produced 52.83% of the total of 954 referring expressions. This distribution shows that the L2 signers were able to retell the stimulus stories, and it suggests that we can compare their overall performance with that of the native signers.
We first asked whether the L2 signers distribute their referential expressions the same way as the native signers. Table 1 shows the proportion of L2 and native signers’ referring expressions that were used to introduce, maintain or reintroduce characters or objects. While the two signer groups are using the same proportion of expressions to introduce characters, there is a difference between groups for the maintained and reintroduced contexts. The L2 signers use a smaller proportion of maintained reference than the native signers, as well as a larger proportion of reintroduced reference.

We then asked whether we would find additional differences in the way the L2 signers used particular types of referring expressions within each discourse status, as compared with the native signers. Recall that we expect the L2 signers to be overly explicit in their use of referring expressions, that is, we expect them to use more nouns and pronouns in maintained contexts than the native signers. Table 2 shows what proportion of native and L2 signers’ referring expressions were nouns, pronouns, zero anaphora and classifiers in each discourse status.
Looking first at the introduced discourse status only, graphed in Figure 1, we see that the preferred referring expressions for introducing a referent in ASL are nouns and, to a smaller extent, classifiers. The L2 signers use these expressions in proportions that are similar to the native signers’. This suggests that the L2 learners have acquired the typical way of character introduction in ASL.

![Figure 1. Mean % of introduced references by group and reference type.](image)

The maintained context is displayed in Figure 2. Here we see more differentiation between the two signer groups. Both groups use zero anaphora as their preferred means, and pronouns as their least preferred means of maintaining reference, but the L2 signers use a smaller proportion of classifiers than native signers for this purpose. Consequently, the L2 signers use roughly the same proportion of nouns and classifiers to maintain reference, whereas the native signers have a clear preference for classifiers over nouns in maintained contexts. This analysis suggests that L2 signers are not overusing pronouns or underusing zero anaphora, as predicted by the over-explication hypothesis. However, we see that the L2 signers are indeed using a greater proportion of nouns in maintained contexts, compared with the native signers, and this finding is in line with our hypothesis.

Last, we regard the reintroduced references, shown in Figure 3. Here again, we see some differences between the signer groups. The L2 signers use more nouns, and less zero anaphora than the native signers in this referential context.
Moreover, they use pronouns and classifiers, which the native signers appear to substantially disprefer for this function.

Recall that our analysis of the use of the three discourse statuses (introduced, maintained, reintroduced) found different patterns in the L2 vs. the native signer group. Because this finding was unexpected, we speculated about its causes. What might cause native signers to use a very small proportion of reintroductions in their narratives, which the L2 signers would not have acquired? Given that the affordances of the visual modality differ from those of the spoken modality with respect to the number and behavior of the articulators, we hypothesized this pattern might arise from the native signers’ use of the sign medium. As opposed to the spoken modality, the visual modality allows for the simultaneous use of more than one articulator, for example by holding constant a handshape on the non-dominant hand, while driving the narrative forward with the signs on the dominant hand. If we exclude gestural expressions, spoken languages are articulated in a sequential fashion where referent mentions substitute on another in time. In a signed language, however, it is possible to keep multiple referents active in the discourse at the same time. Such processes might play an important role in the creating of cohesion in signed narratives, and they may not be straightforward for learners to acquire. On these grounds, we predicted that the native signers in our data would show a greater proportion of holds across clauses than the L2 signers. We restricted our analysis to cases where the same handshape was held on either hand across a clause boundary.

**Figure 2. Mean % of maintained references by group and reference type.**
We then calculated the ratio of cross-clause holds to the total number of clauses in the narrative. We first did this individually for each signer, and then we computed a mean proportion for each group. For the L2 signers, we found a mean proportion of 3.6%, and for the native signers the mean proportion was 7.8%. This pattern would seem to support our hypothesis, and help explain why native signers reintroduce less than L2 signers.

6. Comparing L2 reference tracking in the spoken and signed modalities

The goal of this paper was to investigate how native English speakers track reference in L2 ASL. Studies of spoken L2 reference tracking have shown a general trend towards learner over-explicitness, which appears to be independent of source-target language pairs. In line with these findings, we hypothesized that L2 learners tracking reference in ASL would exhibit the same tendency. We therefore expected to see an overuse of overt elements, like pronouns and nouns in the context of maintained reference, as well as an underuse of null elements. Unlike the previous research on L2 signers of Catalan Sign, our study did not find the learners to overuse overt pronouns for referent maintenance in their L2. We also did not find an underuse of null elements in the same context. We did however find a greater use of nouns in maintained contexts in the learner narratives, compared with the native signers. We also found some unexpected

Figure 3. Mean % of reintroduced references by group and reference type.
results. First, although the proportion of classifiers used by each group is comparable when we average across discourse statuses, we find different distributions in the two groups. While the L2 learners use similar proportions of classifiers in maintained and reintroduced contexts, the native signers only use classifiers for the purpose of maintaining reference, and never when they reintroduce. This finding might suggest that although L2 signers appear able to produce classifiers in ASL, they are unsure or unaware of their referential value. Alternatively, this pattern might be related to the second unexpected finding in the present study. In assessing the distribution of referential expressions across discourse statuses, we found that L2 signers used a smaller proportion of maintained reference, as well as a larger proportion of reintroduced reference, as compared to the native signers. Finding a difference in reintroduction frequency suggests to us that L2 signers are struggling with either the ASL specific way of approaching story telling. This might be due to the learners focusing on one clause at a time, and only keeping one referent active at a time. This would necessitate more frequent reintroductions compared to native signers who might utilize the affordances of the visual modality to a greater degree. By relying on having multiple articulators, native signers might be able to simultaneously keep multiple referents active in signing space, and thus in the discourse. To corroborate our speculations that this might be driving the difference in the proportions of reintroduced reference, we re-examined the narratives. We computed the ratio of signs held across clause boundaries to the number of clauses in the narrative, and compared the mean ratio of the L2 signers with that of the native signers. We found that, on average, native signers utilize this possibility more than twice as often as L2 learners do.

We propose that these results are in fact in line with the research on spoken L2 reference tracking, because they reflect the same underlying process. It has been proposed that the over-explicitness characterizing L2 reference tracking is a result of learners’ attempt at reducing cognitive load (Carroll & Lambert, 2003). This can be done by focusing on sentence level planning and eliminating the need for discourse level planning. In spoken languages this results in overspecificity of referential expressions. The low proportion of cross-clause holds in L2 ASL may well reflect the same principle of focusing on sentence level planning. If so, it suggests that the different affordances of the two modalities lead to bimodal second language learners facing different challenges than unimodal learners in the production of cohesive discourse. One caveat to our conclusions is that our study has only scratched the surface. The findings reported here should be confirmed through detailed statistical analyses, with more complex stimulus stories and with more homogenous learner groups. However, we think that studying reference tracking in an L2 in a different modality is an exciting avenue for viewing the problems of learner acquisition of cohesion from a new angle. Moreover, our findings point to the necessity of further studies of native reference tracking in the visual modality, so that we can gain a more complete understanding of the devices driving cohesion in signed discourse.
References


