Bilingualism: Language and Cognition 15 (2), 2012, 402–412 © Cambridge University Press 2012 doi:10.1017/S1366728911000587

Can experience with co-speech gesture influence the prosody of a sign language? Sign language prosodic cues in bimodal bilinguals*

DIANE BRENTARI
Department of Linguistics, University of Chicago
MARIE A. NADOLSKE
Department of Linguistics, Montclair State University
GEORGE WOLFORD
Department of Speech, Language and Hearing Sciences, Purdue University

(Received: September 9, 2011; final revision received: July 19, 2011; accepted: December 16, 2010; first published online January 4, 2012)

In this paper the prosodic structure of American Sign Language (ASL) narratives is analyzed in deaf native signers (L1-D), hearing native signers (L1-H), and highly proficient hearing second language signers (L2-H). The results of this study show that the prosodic patterns used by these groups are associated both with their ASL language experience (L1 or L2) and with their hearing status (deaf or hearing), suggesting that experience using co-speech gesture (i.e. gesturing while speaking) may have some effect on the prosodic cues used by hearing signers, similar to the effects of the prosodic structure of an L1 on an L2.

Keywords: sign language, gesture, language contact, prosody, second language acquisition, constructed action, character viewpoint, narrative, discourse

1. Introduction

Bilingualism takes many forms within a deaf community. The type of bilingualism we are addressing in this paper primarily concerns hearing people who use a spoken language and also use a sign language; these individuals are known as bimodal bilinguals. Some hearing people are native in both a signed and a spoken language because they are children of signing deaf parents, often referred to as CODAs (Children of Deaf Adults). Other hearing people learn sign language as a second language. In this paper we are interested in the potential effect that co-speech gestures (i.e., the gestures made spontaneously while speaking) may have on the prosodic structure of American Sign Language (ASL) in bimodal bilinguals. Previous studies have shown that there are influences of English on ASL in "contact" varieties of signing in matters of word order and lexical choice (Lucas & Valli, 1992), and of ASL on English in the co-speech gestures of English bimodal bilinguals (Casey & Emmorey, 2008; Emmorey & McCullough, 2008; Pyers & Emmorey, 2008).

There is evidence that, like the segmental patterns of a first language (L1), the prosodic structure of a spoken L1 has an effect on the acquisition of a spoken second language (L2) (McGory, 1997; Queen 2001; Trouvain & Gut, 2007). McGory (1997) reported on the different prosodic influences of Chinese vs. Korean when learning English as an L2. Queen (2001) found that bilingual speakers of Turkish and German produce a pattern of intonation that includes properties of both the L1 and L2 but is also distinct from both languages, which she called "language fusion". We expect that for a spoken L1 and a signed L2, the prosodic influence of the spoken L1 will be primarily from co-speech gesture, and to a lesser extent, from common properties that exist in both modalities, such as durational properties of words, signs, or pauses.

Two hypotheses will be tested with respect to possible effects of experience with co-speech gesture on sign language prosody. The prosodic pattern considered the target for the L2 signers is that of the native deaf group (called L1-D). Two hearing groups of signers will then be compared to this baseline: native hearing signers (CODAs) and second language learners of ASL (L1-H and L2-H, respectively). One hypothesis (Hypothesis 1) is that there is no influence on ASL from English within the prosodic domain. Since the two languages are in different communication modalities, it could be that the prosodic structures of the two systems are simply too different or incompatible for interaction effects to be possible. If this is the case, the baseline pattern of the L1-D group will be the same one exhibited by both hearing groups, although we might expect that the hearing groups would

* This research was supported by NSF grant BCS 0547554 to Diane Brentari and NIH DC5241 Ronnie B. Wilbur. We also thank two anonymous reviewers for helpful comments on earlier versions of this manuscript.

Address for correspondence:
Diane Brentari, Department of Linguistics, University of Chicago, 1010 East 59th Street, Chicago, IL 60637-1512, USA
dbrentari@uchicago.edu

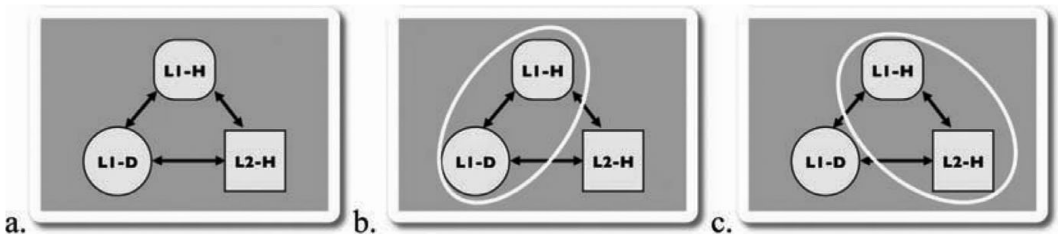

Figure 1. Schema of the hypotheses considered in this study: (a) no effect of American English co-speech gesture on ASL; (b) an effect of language experience (L1 vs. L2) on the prosodic patterns found; (c) an effect of hearing status (deaf vs. hearing) on the prosodic patterns found.

be less accurate; that is, we would not expect the hearing signers to add “innovations” to the system. Alternatively, one might hypothesize (Hypothesis 2) that gesture (as a part of English) can have an effect on bimodal bilingual signers’ use of ASL prosody, since some of the same cues are used in both gesture and sign. This effect might be the result of L2 acquisition or the result of gestural experience.

The possible effects of gesture on sign language prosody are exemplified in Figure 1. Figure 1a represents the case where no significant relationship among the three groups studied is found, supporting Hypothesis 1. Figures 1b–c represent cases that support Hypothesis 2. Figure 1b represents an effect of language experience (L1 vs. L2) on prosodic structure; the two L1 groups (deaf and hearing) pattern together. The explanation for such a pattern could be an interference effect of gesture on the L2-H signers, but alternatively, L2-H signers might simply not yet have acquired the L1 pattern, despite their fluent use of ASL. Evidence for Hypothesis 2 would be exemplified more clearly if our data were to show the pattern in Figure 1c, where the two hearing groups (L1 and L2) of signers pattern together, indicating that their gestural experience as hearing people has an effect on their production of sign language prosody, even in the case of L1-H signers.

1.1 Gestural cues in signed and spoken languages

The general idea that properties of gestural prosody and sign language prosody may be related is not completely new; Wilcox (2007) and Wilcox, Rossini and Pizzuto (2010) proposed a historical link between the two types of languages, but their analyses were based on grammaticalization of special lexical items and morphological properties; i.e., word-level effects. The current study, instead, is at the phrasal level, and to our knowledge there is no previous work of the sort being proposed here. This investigation is a way to introduce phrasal level prosodic comparisons as a field of investigation, focusing on the analysis of gestural properties and their phrasal prosodic domains in spoken and signed languages. The focus of this paper is on the Intonational Phrase (I-Phrase or IP) and the Utterance (U).

Within the prosodic hierarchy, the I-Phrase is the prosodic unit most closely aligned with the clause, while the U is the next largest prosodic unit and is most closely aligned with a complete thought; Us may contain one I-Phrase or more than one (Nespor & Vogel, 1986). Examples of I-Phrases and Us are given in (1) and (2), and while the prosodic cues of ASL and English that mark I-Phrases and Utterances are different (auditory/vocal vs. visual/gestural), the I-Phrase and Utterance constituency would be the same in (1) and (2). The string in (1) contains one U with two I-Phrases, while the one in (2) contains two Us with two I-Phrases (the subscripted letter ‘a’ refers to spatial locations that are co-referenced in the utterance).

(1) *One utterance, two I-Phrases*
  a. ASL: [MARY$_a$ POSS$_a$ TRAIN ARRIVE TIME-6:00] $_{\text{IP}}$ [INDEX-1SG PICK-UP$_a$] $_{\text{IP/U}}$
  b. English: [Mary’s train arrives at 6:00]$_{\text{IP}}$ [I’ll go pick her up] $_{\text{IP/U}}$

(2) *Two utterances, two I-Phrases*
  a. ASL: [MARY$_a$ POSS$_a$ TRAIN ARRIVE TIME-6:00] $_{\text{IP/U}}$ [DINNER EAT WHAT] $_{\text{IP/U}}$
  b. English: [Mary’s train arrives at 6:00] $_{\text{IP/U}}$ [What’s for dinner] $_{\text{IP/U}}$

Research on sign language prosody has made great strides in the last twenty years, particularly at the level of the I-Phrase and U, which is why we focus our attention on these units (Boyes Braem, 1999; Brentari, 1998; Brentari & Crossley, 2002; Brentari, González, Seidl & Wilbur, 2011; Nespor & Sandler, 1999; Nicodemus, 2010; Pfau & Quer, 2010; Sandler & Lillo-Martin, 2006; Tang, Brentari, González & Sze, 2010; Wilbur, 1994; Wilbur & Patschke, 1998). Almost all of this work has used deaf native (or near-native) signers to establish a foundation of the prosodic cues used in sign languages and their distribution. With a few notable exceptions, there is very little research on the prosodic cues in deaf individuals acquiring a sign language later in life (Boyes Braem, 1999), non-signers (Brentari, González, Seidl & Wilbur, 2011; Fenlon, Denmark, Campbell & Woll, 2007), or second language users of a sign language (Nicodemus, 2010), and there is no work that we are aware of on the prosody of hearing native signers.

The analysis of gesture in English has a long history (see Gullberg, 2006; Kendon, 2004, for excellent overviews), and work in this area has become more widely known in recent decades thanks to the work of Adam Kendon, David McNeill and colleagues (Kendon 2004; McNeill, 1992, 2000). Here only the aspects of this work that are most relevant to the prosodic analysis of this paper are described. Gestural patterns vary according to culture and language family (Kita, 2009), and this variation is also evident in balanced bilinguals (Stam, 2006). Moreover, Gullberg (2010) has shown that the study of gesture in L2 learners can provide insights into the process of L2 spoken language acquisition. In this paper, we are addressing the specific relationship between English and ASL.

Co-speech gestures are produced by the hands, eyebrows, face, and body. Iconic, metaphoric, and emblematic gestures contribute meaning to a message, and are called REPRESENTATIONAL gestures. BEAT gestures have no meaning and have often been considered most relevant for prosodic analyses (Duncan, 1996; Gullberg, 2006; McNeill, 1992, 2000); however, more recent work on the specific timing of gestures and speech suggests that ALL gestures have prosodically important timing (Krahmer & Swerts, 2005, 2007, 2009; Loehr, 2004, 2007; McClave, 1998). Loehr (2007) proposes that the difference between representational and beat gestures is that representational gestures have meaningful content, while beats are empty of meaningful content. In our analysis, representational gestures play the greater role because these are the ones that are most similar to the prosodic gestures of sign languages.

Nine prosodic cues that have been identified as cues marking I-Phrases and Us in ASL will be the target of our investigation. Many of them appear in co-speech gesture, as described in the next section, and we have also included duration, since it is a bi-modal property that can appear in speech, gesture, or sign language.

2. Methods

2.1 *Participants*

All of the participants were highly proficient ASL signers to ensure, to the greatest extent possible, that differences among the groups would not be due to a lack of general language proficiency. Because of the specific characteristics of our participants and the depth of our analysis, our participant group is small. Nine adult participants were distributed equally among the three participant groups: three deaf native ASL signers (L1-D group), whose mean age was 31 years and who had attended residential schools for the deaf; three native hearing signers who learned ASL from their deaf parents (L1-H group), whose mean age was 51 years; and three highly proficient second-language learners of ASL (L2-H group), whose mean age was 34 years. The L1-H and the L2-H signers were ASL–English interpreters and were certified either by the state of Indiana or by both the state and the national Registry of Interpreters for the Deaf (RID). These signers were selected on the basis of recommendations of a deaf interpreter consultant who was familiar with the interpreter population in the greater Indianapolis area and was aware of the importance of ASL proficiency for this task (instead of Signed English signing skills). Table 1 shows how the groups chosen isolate language experience from hearing status (see boxed elements).

Table 1. *The factors manipulated in this experimental design and how they are isolated by including the three groups of signers.*

	L1-D	L1-H	L2-H
ASL status	L1	L1	$\boxed{\text{L2}}$
ASL skill	Native	Native	Advanced/native-like
Parents	Deaf	Deaf	Hearing
Hearing status	$\boxed{\text{Deaf}}$	Hearing	Hearing
Interpreter	No	Yes	Yes

2.2 *Data collection*

Each signer sat across from the researcher, a proficient hearing L2 ASL signer who began learning ASL at age 8 years and taught ASL at Purdue University, with a laptop computer on a table or chair off to the side facing the signer. One camera filmed the signer and a second camera filmed the researcher. Each session began with a short ASL conversation to ease the signer into being filmed. The signer was asked to view a series of video clips, which were part of a larger study (Nadolske, 2009), and describe the clip to the researcher after watching each clip. The target stimulus for this study was presented second in the series. Signers were able to view each clip as many times as they wanted before responding. The researcher limited her responses to the signer's narration to nodding in comprehension, copying emotive facial expressions of the signer (e.g., smile or surprise), and occasionally signing OH-I-SEE. The target stimulus was a video clip selected from the fourth season of the cartoon television series *The Simpsons* from the episode *Homer the Heretic*. The stimulus lasted 31 seconds and contained no audio. This clip focused on two cars involved in a chase.

2.3 *Analysis of cues*

I-Phrases and Us were determined on the basis of three independent transcriptions by three highly proficient L2

https://doi.org/10.1017/S1366728911000587 Published online by Cambridge University Press

Table 2. *Cues transcribed in the ASL narratives and their definitions.*

Sign language cue	Sign language cue definition
Length/sign duration	Duration of the sign's articulation measured from the complete formation of the initial handshape of the sign to the time when the final handshape begins to deteriorate.
Hold	Period of time when the hand is kept in its particular shape and position.
Transition	The duration of the movement that starts at the end of the lexical movement or hold of one sign and ends at the beginning of the next sign.
Pause	The hold at the end of the sign plus the transitional movement between one sign and the next.
Blink	Inhibited, periodic eye blinks.
Drop hand	Deviation from the direct trajectory between the end of the preceding sign and the beginning of the next sign during the transitional movement: hands drop to the lap or to neutral position, or the wrists become lax.
Brow position	The position of the eyebrows – up, down, or back to neutral position.
Head position	A change in head position independent from changes in the torso – forward, back, sideways, or neutral.
Torso position	A change in position of the torso – forward, back, sideways, or neutral.

signers who were not participants in the study (two learned ASL at 7 and 8 years of age, respectively, and the third was a full-time certified interpreter). Judges were instructed to break the narratives into the largest prosodic units first, which were labeled Us, and the second largest units, which were labeled I-Phrases, without attending to specific prosodic cues or semantic content. Only those units on which there was agreement for two out of three judges on constituent type (U or I-Phrase) and placement of the boundary were analyzed further.

All of the prosodic cues transcribed for this study (provided in Table 2) have been discussed in the literature as robust cues of I-Phrases in ASL as well as other sign languages, although none of the cues is obligatory.

The transcriptions were done in EUDICO Linguistic Annotator (ELAN), a tool developed at the Max Planck Institute for Psycholinguistics, Nijmegen, for the analysis of language, sign language, and gesture. ELAN is a tool that allows the location and duration of each annotation to be indicated using time-aligned tier structures. The cues can be roughly divided into those that concern duration, those that occur at a boundary, and those that spread across an entire constituent domain and whose values change at a boundary.

The first four cues – SIGN DURATION, HOLD DURATION, TRANSITION DURATION and PAUSE DURATION – are similar in timing and function to positional lengthening effects in spoken English I-Phrases (Grosjean & Dechamps, 1975; Klatt, 1976). All of the durational ASL cues as a group create the positional effects of boundary strengthening, but any one of them could be responsible for this effect in ASL so each was measured independently. Sign duration was measured from the complete formation of the initial handshape of the sign to the time when the final handshape begins to deteriorate (Brentari & Crossley, 2002; Brentari et al., 2011, Tang et al., 2010; Tyrone, Nam, Saltzman, Mathur & Goldstein, 2010). Duration of signs in I-Phrase-initial, I-Phrase-final, and U-final positions was measured. Hold duration is the period of time when the hand is kept in its particular shape and position; the prosodic use of this cue typically occurs at the end of signs (Brentari, 1998; Liddell & Johnson, 1989; Perlmutter, 1992). Hold duration was measured in I-Phrase-final and U-final position. Transition duration is the duration of the transitional movement that starts at the end of the hold of one sign and ends at the beginning of the next sign (Brentari, 1998). Pause duration is a composite cue, consisting of the hold plus transition duration (Grosjean & Lane, 1977; Liddell, 1978; Tyrone et al., 2010). The composite cue pause is equivalent in function to pauses in spoken languages, but the two component parts – hold and transition – do not have a similar prosodic function in spoken languages (Brentari, 1998).

BLINK and DROP HANDS typically appear at boundaries and are conceptually similar to boundary tones in spoken languages. Blinks are inhibited, periodic eye blinks (Wilbur, 1994); these same periodic eye blinks do not appear at a prosodic boundary in spoken languages, but their likely counterparts are the breaths taken by speakers at the end of I-Phrases and Us (Baker & Padden, 1978; Tang et al., 2010). Drop hands is a deviation from the direct trajectory between the end of the preceding sign and the beginning of the next sign during the transitional movement: hands drop to the lap or to neutral position, or the wrists become lax (Nicodemus, 2010). We included this as a possible I-Phrase cue because "rest" position, which is a drop-hands position, is potentially a point of re-calibration in co-speech gesture (Kipp, Neff, Kipp & Albrecht, 2007).

In ASL, TORSO POSITION, BROW POSITION, and HEAD POSITION function as domain cues, conceptually similar to nasal harmony or vowel harmony in spoken languages, because these cues extend over an entire constituent,

https://doi.org/10.1017/S1366728911000587 Published online by Cambridge University Press

encompassing a sequence of signs (Pfau & Quer, 2010). The type, duration, beginning point, and ending point of each of these cues were transcribed.

In sign languages torso position has been reported to have three uses. The first is prosodic: Boyes Braem (1999) found that the points where the direction changed corresponded with the phrasal boundaries, noted in Swiss German Sign Language, illustrated in (3a) below.[1] The second use is pragmatic, seen in (3b), signaling direct discourse (Engberg-Pedersen, 1995, for Danish Sign Language; Gee & Kegl, 1993, for ASL; Morgan, 2000, for British Sign Language). The third, in (3c), is related to the second, used in CONSTRUCTED ACTION (Mazzoni, 2008, for Italian Sign Language; Quinto-Pozos & Mehta, 2010, for ASL), during which the signer assumes the role of a character in the narrative even if no actual discourse is uttered. In co-speech gesture TORSO SHIFT is associated with character viewpoint gestures (McNeill, 1992), and is also used when the speaker's gestures assume the role of a character in the narrative, see (3d). Typically, character viewpoint gestures express entire events or clauses (Parrill, 2010). The third use in sign languages and the use in co-speech are therefore related in both function and form. In this study, the annotations used for torso position (also referred to as torso shift) were "forward", "back", "sideways", or "neutral"; the three uses of torso position were also coded ("prosodic", "pragmatic", and constructed action), as illustrated in (3a–c).

(3) *Uses of torso position in sign languages and character viewpoint co-speech gesture*

a. Prosodic use (from Boyes Braem, 1999)
torso left____________________
ME FATHER BUY NEW SKIS
torso right_________________ torso center
AND SHOES-FITTED-TO-SKIS
"I and my father bought new skis and shoes fitted to skis."

b. Pragmatic use (direct address) (from Boyes Braem, 1999)
torso left torso right_____ torso left
I SAY STORE I GO-TO NEVER I
"I said I will not go to the store. I would never go!"

c. Constructed action (from Quinto-Pozos & Mehta, 2010)
torso shift_________
DON LOOK [surprised] . . .
"Don just looked at Buddy in surprise . . ."

d. English character viewpoint, co-speech gesture (from Parrill, 2010)
coded as use of torso
"Cat leans against door, sighs with relief."

[1] I-Phrases and phonological phrases were not differentiated in Boyes Braem (1999).

Brow position also has three main uses in ASL. Two are syntactic. A brow raise co-occurs with a topic structure, as in (4a), and constitutes an independent I-Phrase that occurs sentence-initially (Aarons, 1994; Liddell, 1978; Wilbur, 1994). A brow furrow signals a "*wh*-question" (Aarons, 1994; Sandler & Lillo-Martin, 2006), as in (4b). A third use of brow movements is affective and occurs during constructed action, as described above for the cue torso position, see (4c). In co-speech gesture BROW MOVEMENTS can occur as beats, and in such cases they are timed with pitch-accented syllables (Krahmer & Swerts, 2007). They also convey emotional affect, such as uncertainty, disbelief, surprise, and exasperation (Cohn, Reed, Ambadar, Xiao & Moriyama, 2004; Poggi, Pelachaud & De Rosis, 2000), especially during character viewpoint gestures, as described above and shown in (4d). Possible values for brow position were "up" (raise), "down" (furrow), or "neutral", as well as the three uses given in (3) above: "*wh*-question", "topic structure", and constructed action.

(4) *Uses of brow movement in ASL and English character viewpoint co-speech gesture*

a. ASL syntactic (topic) (from Aarons, 1994)
brow raise__
JOHN TEACHER REQUIRE [ø] LIPREAD MOTHER
"The teacher required John to lipread his mother."

b. ASL syntactic (*wh*-question) (from Sandler & Lillo-Martin, 2006)
______brow furrow
JOHN BUY WHAT
"What did John buy?"

c. ASL constructed action (affective) (from Quinto Pozos & Mehta, 2010)
brow raise_________
DON LOOK [surprised] . . .
Don just looked at Buddy in surprise . . .

d. English character viewpoint, co-speech gesture (from Parrill, 2010)
coded as use of affect
"Rabbit looks alarmed."

Head nods have been reported as an I-Phrase marker in Israeli Sign Language (Nespor & Sandler, 1999; Sandler & Lillo-Martin, 2006) and studied in co-speech gesture as well. House (2007) found that avatars are more human-like if rhythmic head nods are included, and McClave, Kim, Tamer and Mileff (2007) have found a variety of subtle, meaningful differences of head movements across several linguistic communities. In this study, the use of head nods as a prosodic boundary marker was our primary concern.

The ASL cues transcribed were normalized in order to analyze the cues across subjects and groups. For cues that could be counted (blink, drop hands, brow,

https://doi.org/10.1017/S1366728911000587 Published online by Cambridge University Press

head, and torso position), normalization was achieved by calculating by subject, and then by group for the proportion of these cues whose values appeared (in the case of blink and drop hands) or changed (in the case of brow, head, and torso position) at an I-Phrase boundary. For cues of duration (sign duration, and pause duration) normalization was achieved by calculating a ratio as follows. For sign duration, the mean durations of IP-phrase-internal, I-Phrase-initial, I-Phrase-final, and U-final signs were transcribed in milliseconds (ms), and then ratios were calculated among them, first by signer and then by group. For example, the ratio between I-Phrase-final and I-Phrase-internal means will indicate whether, for a given signer, there is lengthening of a sign in I-Phrase-final position. A ratio of 1 would indicate no difference between the two means. A ratio of <1 (e.g., 0.8) would indicate shortening I-Phrase-finally, whereas a ratio of >1 would indicate lengthening to some degree (e.g., 1.5). Likewise, normalized pause duration for each signer was achieved by calculating the mean pause duration between signs in IP-phrase-internal, I-Phrase-final, and U-final position, and then by calculating the ratios among them. For example, the ratio between the means for U-final and I-Phrase-final pauses would indicate, for a given signer, whether the pauses in U-final position are longer than those in I-Phrase-final position. A ratio of 1 would indicate no difference, whereas a ratio greater or lesser than 1 would indicate a difference – <1 (e.g., 0.8) would indicate shorter pauses in U-final position and >1 (e.g., 1.5) would indicate longer pauses in U-final position.

3. Results and discussion

The results address two dimensions of the data – overall consistency and differences among the groups. A cue was considered to be "consistent" if it was used at least 50% of the time at an I-Phrase boundary by any single signer in any of the three groups. The cues that fell below that criterion were hold duration, head position, and drop hands. The other cues exceeded that criterion – blink, brow position, torso position, sign duration, and pause duration.

Based on our experimental design, the cues that appeared consistently in the data revealed two patterns. Recall (from Figure 1 above) that if the two groups of native signers pattern together, differences will be attributed to language experience (L1 vs. L2), but if the two groups of hearing signers pattern together, the difference will be attributed to hearing status (deaf, hearing). Blink, brow, and pause demonstrated a mixed pattern, in which the influence of language experience and hearing status could not be determined with certainty. Sign duration and torso position demonstrated a different pattern, in which the difference between deaf and hearing groups could be attributed to hearing status. Table 3 shows the average, normalized use of each cue per group as described in the "Methods" section above, as well as the statistical analyses. A Mann-Whitney-U test of ranked comparisons was used to determine differences among groups because of the small sample size.

3.1 *Type 1: Cues used inconsistently*

While drop hands, head position, and holds have been reported in the literature as marking I-Phrases in several sign languages, they were not used frequently by any of the signers in our study.

IP-phrase-final holds were expected to be prevalent in the data based on previous literature (Liddell & Johnson, 1989; Perlmutter, 1992), but they were not used consistently by of the participants in this study (range: 5%–33%). This cue contributes to the composite cue of pause, but on its own it was not used as a reliable marker of I-Phrase boundaries. More recent work on this cue confirms this result (Brentari et al., 2011; Tang et al., 2010; Tyrone et al., 2010).

As noted in the introduction, head nods and other head position cues are used prosodically in co-speech gesture (House, 2007) and, in several sign languages, as an I-Phrase marker (see Tang et al., 2010, for Japanese Sign Language, Sandler & Lillo-Martin, 2006, for Israeli Sign Language). It was not used by any of the signers in these data as an I-Phrase marker (range: 0%–26%).

Drop hands occurs as an I-Phrase marker in Infant Directed Signing (Brentari et al., 2011), in the signing of interpreters while working (Nicodemus, 2010), and also in co-speech gesture as a "rest" between gestures, but it not it was not prevalent in the current data set (range: 6%–23%).[2]

Figure 2 shows the mean proportion by group used for the cues drop hands, head nods vs. holds vs. blinks, torso, and brows, which were used more consistently. No further analyses were performed on the three inconsistently used cues.

3.2 *Cues used consistently*

Cues without clear effects of language experience or hearing status

Blink, brow movements, and pause showed either marginal, equal, or mixed effects among groups and so the differences cannot be attributed to language experience or hearing status with any certainly. Blink was used prevalently in all groups of signers, but it showed only marginal differences both between L1 (n = 6) and L2 (n = 3) signers and between hearing (n = 6) and deaf (n = 3) signers ($p = .048$; range: 38%–82%). Recall that blinks are used cross-linguistically for this purpose by

[2] One of the reviewers suggested that interpreters may drop their hands while information is being received. This is certainly one possible use of this cue in the interpreting context.

https://doi.org/10.1017/S1366728911000587 Published online by Cambridge University Press

Table 3. *Summary of results showing the prosodic cues divided into three groups: those that did not meet the criterion of 50% for any signer are above the double line (hold, head, and drop hands). Those that met the criterion of 50% for any signer are below the double line: (blink, brow, and pause) had mixed results across groups effects of language experience or hearing status could not be clearly determined. Length and torso position showed significant differences between hearing and deaf signers (bottom two cues). Descriptive statistics are shown on the left, and Mann-Whitney-*U *results on the right (*U *value,* p *value; * indicates confidence levels of significance greater than .05).*

Cue	L1-D (n = 3)	L1-H (n = 3)	L2-H (n = 3)	L1(n = 6)/ L2(n = 3) *U*	L1(n = 6)/ L2(n = 3) *p =*	H(n = 6)/ D(n = 3) *U*	H(n = 6)/ D(n = 3) *p =*
Hold: total # _]IP	10	9	27				
% at _]IP	.16	.16	.25				
Head: total # _]IP	11	12	28				
% at _]IP	.16	.22	.26				
Drop hands: total #	8	8	10				
% at _]IP	.11	.10	.09				
Blinks: total # _]IP	36	61	77				
% at _]IP	.57	.81	.73	2	*.048	2	*.048
Brow: total #	49	74	47				
% at _]IP	.82	.49	.25	0	*.012	0	*.012
Pause: total # _]IP	63	60	77				
ms at _]IP	215	208	332	1	*.021	3	*.021
Pause: ratio U/IP	.98	1.25	1.93				
ms at _]U	210	260	642	0	*.012	1	*.021
Torso: total # _]IP	36	55	106				
% at _]IP	.30	.58	.40	3	.083	1	*.021
Length: ms at _]IP	604	796	965				
ratio_]IP/Wd	1.60	2.26	1.77	3	.083	1	*.021
ms at IP[_	459	320	581				
ratio IP[_/Wd	1.21	1.01	1.06	3	.083	1	*.021
ms at _]U	1662	1112	1181				
ratio_]IP/U	1.86	1.46	1.23	3	.083	2	*.048

many sign languages to mark I-Phrase boundaries, but not by hearing people during speech (Baker & Padden, 1978; Tang et al., 2010; Wilbur, 1994).

Brow movements showed a significant difference of the same magnitude between L1 and L2 ($p = .012$) and hearing and deaf ($p = .012$) signers (range: 20%–92%); in addition, this was the only cue for which there was a significant difference, pairwise, among all three groups: between L1-D (n = 3) and L1-H (n = 3) signers ($U = 0$, $p = .05$; range 48%–92%); between L1-H (n = 3) and L2-H (n = 3) signers ($U = 0$, $p = .05$; range 20%–52%), and between L1-D and L2-H signers ($U = 0$, $p = .05$, range: 20%–90%). One use of brow movements in ASL is syntactic, to signal topics or *wh*-questions, and another is affective, in constructed action. To further investigate the possible explanation that the syntactic or affective uses were responsible for the difference in the use of brow, the proportion of each use of brows was calculated. The results were inconclusive, so the source of this difference could be either increased affective use of brows in constructed action by hearing signers, which WOULD show a possible influence from gesture, or increased use of topics by L1 signers, which WOULD NOT show an influence of gesture.

Mann-Whitney-*U* comparisons on the cue of pause (both I-Phrase and U) showed mixed effects of language experience and hearing status. Equally significant differences between L1 and L2 signers and between hearing and

https://doi.org/10.1017/S1366728911000587 Published online by Cambridge University Press

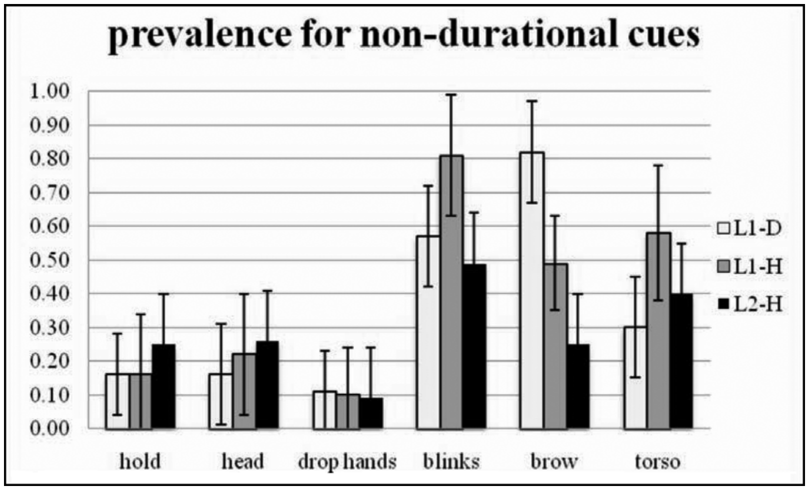

Figure 2. Average proportion for each group (L1-D, L1-H, and L2-H) for each of the cues indicated (total number of occurrences ÷ the number of occurrences at an I-Phrase boundary). Those on the left (hold, head, drop hands) had fewer than 50% consistency at I-Phrase boundaries in all signers.

deaf signers in I-Phrase final position ($U = 1$, $p = .021$, range: 208–332 ms) were found, but in U-final position there was a larger effect of language experience than for hearing status (L1 vs. L2: $U = 0$, $p = .012$, range: 210–642 ms; hearing vs. deaf: $U = 1$, $p = .021$, range: 210–642 ms). In U-final position, the L2 group produced longer pauses in U-final position than the L1 groups. A univariate ANOVA confirmed this finding using group (L1-D, L2-D, L2-H) and position (I-Phrase-final, U-final) as independent variables and duration in milliseconds as a dependent variable. There was a significant main effect for group, specifically in U-final position between L1 and L2 signers – $F(17,210) = 2.184$, SD = 265.437, $p = .007$). This cue was measured in two ways: as a composite cue as described in Table 2 and in the "Methods" section above, and as transition duration. Both tests were run on pause and transition duration and the results were identical, showing that pauses are grounded more firmly in transitions than in holds (recall the hold was used only inconsistently).

Pause has been shown to be important for L1 and L2 patterns in spoken languages at I-Phrase and U boundaries (Grosjean & Lane, 1977), and longer pause durations have also been reported in L2 speakers at U boundaries (Riazantseva, 2002). In L2 speakers, longer U-final boundaries have been attributed to processing demands, and we concur with that explanation in the case of ASL as well. In addition, pause has been shown to be the most salient I-Phrase cue in a perception experiment requiring signing and non-signing participants to make judgments about the presence of I-Phrase boundaries in ASL (Brentari et al., 2011). The fact that it is the most salient cue for both signers and sign-naïve speakers across modality suggests that the presence of a pause at I-Phrase boundaries is a possible language universal that is reliable in both spoken and signed modalities.

Cues showing clear effects of hearing status

Sign duration and changes in torso position were cues that showed a significant difference between hearing (n = 6) and deaf (n = 3) signers ($p = .021$; range: 16%–68%), but NOT between L1 and L2 signers ($p = .083$).

Sign Duration was measured in four positions: I-Phrase-initial, I-Phrase-medial, I-Phrase-final, and U-final positions. In all positions there was a significant difference between hearing and deaf signers (I-Phrase-initial: $U = 1$, $p = .021$, range (ratio): 10.01–1.21; I-Phrase-final: $U = 1$, $p = .021$, range (ratio): 1.60–2.26); U-final: $U = 2$, $p = .048$, range (ratio): 1.23–1.86). Based on our experimental design, the overall pattern can be attributed to hearing status, but the exact nature of the influence from English is not clear. Hearing signers had a stronger effect of lengthening in I-Phrase-final position, and deaf signers had a stronger effect in I-Phrase-initial and U-final positions. A univariate ANOVA was also performed to confirm this finding using group (L1-D, L2-D, L2-H) and position (I-Phrase-initial, I-Phrase-final, U-final) as independent variables and duration in milliseconds as a dependent variable. There was a significant main effect for group – $F(2,751) = 7.260$, SD = 477.641, $p = .027$ – and a significant interaction for group × position (I-Phrase-initial, I-Phrase-medial, I-Phrase-final, and U-final) – $F(8,751) = 2.184$, SD = 424.929, $p = .023$). Moreover, because the positional effects of LENGTH from English would be a case of transfer from speech prosody (i.e., the lengths of spoken words) rather than from co-speech gesture, more work will be needed to interpret this result.[3]

[3] Recall that there are positional effects for gesture length, since they occur most often as singletons, with a neutral rest position between them (Kipp et al., 2007)

https://doi.org/10.1017/S1366728911000587 Published online by Cambridge University Press

Mann-Whitney-*U* tests of ranked comparisons showed that changes in torso position were more consistently timed with I-Phrase boundaries in the two hearing groups of signers than in the deaf group (hearing (n = 6), deaf (n = 3); $U = 1$, $p = .021$, range: 16%–68%). This establishes that the domain of the I-Phrase is more relevant for hearing than deaf signers. In order to interpret this result, consider that the clause (prosodically an I-Phrase) is the relevant constituent for both character viewpoint co-speech gestures in English and in constructed action structures in ASL (Parrill, 2010; Quinto-Pozos & Mehta, 2010). Quinto-Pozos and Mehta (2010) have also argued that this is incorporation of gesture into sign. To determine if hearing signers use constructed action more frequently in their narratives overall, the proportion of constructed action in each narratives was calculated. All groups used constructed actions for the same proportion of the duration of their total narrative (L1-D, 10%; L1-H, 12%; L2-H, 10%). The specific issue here, however, is whether these structures were associated with a torso shift. It was found that both groups of hearing signers made more frequent use of torso shift than deaf signers to accompany their constructed actions (54% for L1-D vs. for 95% for L1-H, and 83% for L2-H signers). Quinto-Pozos and Mehta (2010) also note that torso shift is used less by two deaf native signers in non-formal settings among adults, which was the discourse context here.

Does this pattern come from co-speech character viewpoint gestures? There was no existing study that specifically coded the frequency or degree of torso shift, so to address this gap in the literature, the character viewpoint gestures in a set of three English co-speech narratives collected for a separate project on monolingual English adults were identified and coded using the system used to transcribe the ASL data in this study.[4] We found that 50% of character viewpoint gestures co-occurred with a torso shift. In (5) and in Figure 3, there are three examples of constructed action (one from each of groups in the current experiment, see (5a–c)) and one example, (5d), of a character viewpoint co-speech gesture from the English co-speech narratives.

(5) Examples of presence or absence of *torso shift*
  a. ASL (L1-D) without *torso shift*
     H-O-M-E-R ANGRY
     "Homer is angry."
  b. ASL (L1-H) with *torso shift*
     torso shift__
     H-O-M-E-R ANGRY.
     "Homer is angry."

Figure 3. Examples of constructed action in ASL and "character viewpoint" gesture in English co-speech gesture: (top left) L1-D; (top right) L1-H; (bottom left) L2-H; (bottom right) co-speech gesture. All produce a torso shift except the L1-D signer (top left).

  c. ASL (L2-H) with *torso shift*
     torso shift___________
     H-O-M-E-R ACCELERATES-GAS.
     "Homer accelerates angrily."
  d. English (co-speech gesture): with *torso shift*
     "Sylvester stands on the ledge
     torso shift________________________
     and looks at Tweety swinging his finger."

More work is needed on the use of torso shift in co-speech gesture, but these data demonstrate that: (i) the prosodic domain is the same in co-speech gesture and ASL prosody (i.e., the I-Phrase); (ii) there is a link between their uses (i.e., to assume the role of a character in the narrative); and (iii) in these data, there is more prevalent use of torso shift in hearing signers in this type of informal, narrative context in ASL. We therefore conclude that the pattern of use of torso position in hearing signers (both L1 and L2) is likely to be derived from its use in character viewpoint co-speech gesture.

4. Conclusion

The results of this study have several implications for the gestural prosody of signed and spoken languages, even

[4] These narratives were of "Canary Row" (a cartoon featuring the Tweety and Sylvester characters).

though this analysis was performed on a small number of participants. Some cues that were expected to be prevalent in the data were not – hold, head, and drop hands. This finding contributes to our general knowledge of sign language prosody. In addition, based on our experimental design, some cues showed mixed effects when comparing L1 vs. L2 signers and hearing vs. deaf signers – brow, blink, and pause – while others demonstrated differences solely between hearing and deaf signers, indicating a possible effect of English co-speech gesture on ASL for both L1 and L2 groups of hearing signers.

We have argued that the cue torso position demonstrated the clearest effect of English co-speech gesture on ASL. Our findings suggest that the similarity in form, function, and prosodic domain of character viewpoint co-speech gestures and constructed action ASL structures makes this transfer possible.

Despite the fact that these two languages are in different modalities, these results show that experience with producing gestures that accompany speech can affect a signer's use of prosodic cues. This result has confirmed our second hypothesis (Hypothesis 2), namely that experience with co-speech gesture influences the use of prosody in hearing signers to some degree. These results suggest that the gestural behaviors from a spoken L1 can influence a signed L2 in much the same way as prosodic features of a spoken L1 influence a spoken L2. Moreover, this work supports the view that the gestural elements of prosody in a spoken language are indeed a part of that language's multi-channel phonological structure.

References

Aarons, D. (1994). Aspects of syntax of American Sign Language. Ph.D. dissertation, Boston University.

Baker, C., & Padden, C. (1978). Focusing on the non-manual components of American Sign Language. In P. Siple (ed.), *Understanding language through sign language research*, pp. 27–57. New York: Academic Press.

Boyes Braem, B. (1999). Rhythmic temporal patterns in the signing of early and late learners of German Swiss Sign Language. *Language and Speech, 42,* 177–208.

Brentari, D. (1998). *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.

Brentari, D. (ed.) (2010). *Sign languages: A Cambridge language survey*. Cambridge: Cambridge University Press.

Brentari, D., & Crossley, L. (2002). Prosody on the hands and face: Evidence from American Sign Language. *Sign Language and Linguistics, 5,* 105–130.

Brentari, D., González, C., Seidl, A., & Wilbur, R. (2011). Sensitivity to visual prosodic cues in signers and nonsigners. *Language and Speech, 54 (1),* 49–72.

Brentari, D., & Poizner, H. (1994) A phonological analysis of a deaf Parkinsonian signer. *Language and Cognitive Processes, 9 (1),* 69–100.

Casey, S., & Emmorey, K. (2008). Co-speech gesture in bimodal bilinguals. *Language and Cognitive Process, 24,* 290–312.

Cohn, J., Reed, L. I., Ambadar, Z., Xiao, J., & Moriyama, T. (2004). Automatic analysis and recognition of brow actions and head motion in spontaneous facial behavior. *IEEE Internal Conference on Systems, Man, and Cybernetics*, pp. 210–616. http://ieeexplore.ieee.org (retrieved October 31, 2011).

Duncan, S. (1996). *Grammatical form and 'thinking-for-speaking' in Mandarin Chinese and English: An analysis based on speech-accompanying gestures*. Ph.D. dissertation, University of Chicago.

Emmorey, K., & McCullough, S. (2008). The bimodal bilingual brain: Effects of sign language experience. *Brain and Language, 109,* 124–132.

Engberg-Pedersen, E. (1995). Point of view expressed through shifters. In K. Emmorey & J. Reilly (eds.), *Language, gesture and space*, pp. 133–170. Mahwah, NJ: Lawrence Erlbaum.

Fenlon, J., Denmark, T., Campbell, R., & Woll, B. (2007). Seeing sentence boundaries. *Sign Language & Linguistics, 10,* 177–200.

Gee, J., & Kegl, J. (1993). Narrative/story structure, pausing, and American Sign Language. *Discourse Processes, 6,* 243–258.

Grosjean, F., & Dechamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français: Vitesse de parole et variables composantes, phénomènes. *Phonetica, 31,* 144–184.

Grosjean, F., & Lane, H. (1977). Pauses and syntax in American Sign Language. *Cognition, 5,* 101–117.

Gullberg, M. (2010). Methodological reflections on gesture analysis in second language acquisition and bilingualism research. *Second Language Research, 26,* 75–102.

Gullberg, M. (2006). Some reasons for studying gesture and second language acquisition (Hommage à Adam Kendon). *International Review of Applied Linguistics, 44,* 103–124.

House, D. (2007). Integrating audio and visual cues for speaker friendliness in multimodal speech synthesis. *Proceedings from the 8th Annual Interspeech Conference*, pp. 1250–1253. Antwerp: International Speech Communication Association.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kipp, M., Neff, M., Kipp, K., & Albrecht, I. (2007). Towards natural gesture synthesis: Evaluating gesture units in a data-driven approach to gesture synthesis. *Proceedings of the 7th International Conference on Intelligent Virtual Agents*, pp. 15–28. Berlin: Springer Verlag.

Kita, S. (2009). Cross-cultural variation in speech-accompanying gesture. *Language and Cognitive Processes, 24,* 145–167.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59,* 1208–1221.

Krahmer, E., & Swerts, M. (2005). How children and adults produce and perceive uncertainty in audiovisual speech, *Language and Speech, 48,* 29–53.

Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language, 57,* 396–414.

Krahmer, E., & Swerts, M. (2009). Audiovisual prosody – introduction to the special issue. *Language and Speech, 52 (2–3),* 129–133.

Liddell, S. (1978). Non-manual signals and relative clauses in American Sign Language. In P. Siple (ed.), *Understanding sign language through sign language research*, pp. 59–90. New York: Academic Press.

Liddell, S., & Johnson, R. (1989). American Sign Language: The phonological base. *Sign Language Studies, 64,* 197–277.

Loehr, D. (2004). Gesture and intonation. Ph.D. dissertation, Georgetown University.

Loehr, D. (2007). Aspects of rhythm in gesture and speech. *Gesture, 7 (2),* 179–214.

Lucas, C., & Valli, C. (1992). *Language contact in the American Deaf community*. New York: Academic Press.

Mazzoni, L. (2008). *Classificatori e impersonamento nella Lingua dei Segni Italiana*. Pisa: Plus.

McClave, E. (1998). Pitch and manual gestures. *Journal of Psycholinguistic Research, 27,* 69–90.

McClave, E., Kim, H., Tamer, R., & Mileff, M. (2007). Head movements in the context of speech in Arabic, Bulgarian, Korean, and African-American Vernacular English. *Gesture, 7 (3),* 343–390.

McGory, J. (1997). The acquisition of intonation patterns in English by native speakers of Korean and Mandarin. Ph.D. dissertation, The Ohio State University.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

McNeill, D. (2000). *Language and gesture*. Cambridge: Cambridge University Press.

Morgan, G. (2000). Discourse cohesion in sign and speech. *International Journal of Bilingualism, 4 (3),* 279–300.

Nadolske, M. (2009). Mastering the pressures of variation: A cognitive linguistic examination of advanced hearing ASL L2 signers. Ph.D. dissertation, Purdue University.

Nespor, M., & Sandler, W. (1999). Prosody in Israeli Sign Language. *Language and Speech, 42,* 143–176.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.

Nicodemus, B. (2010). *Prosodic markers and utterance boundaries in American Sign Language interpretation*. Washington, DC: Gallaudet University Press.

Parrill, F. (2010). Viewpoint in speech–gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes, 25,* 650–668.

Perlmutter, D. (1992). Sonority and syllable structure in American Sign Language. *Linguistic Inquiry, 23,* 407–442.

Pfau, R., & Quer, J. (2010). Nonmanuals: Their grammatical and prosodic roles. In Brentari (ed.), pp. 381–402.

Poggi, I., Pelachaud, C., & De Rosis, F. (2000). Eye communication in a conversational 3D synthetic agent. *AI Communications, 13,* 169–181.

Pyers, J., & Emmorey, K. (2008). The face of bimodal bilingualism: Grammatical markers in American Sign Language are produced when bilinguals speak to English monolinguals. *Psychological Science, 19,* 531–536.

Queen, R. (2001). Bilingual intonation patterns: Evidence of language change in Turkish–German bilingual children. *Language in Society, 30,* 55–80.

Quinto-Pozos, D., & Mehta, S. (2010). Register variation in mimetic gestural complements to signed language. *Journal of Pragmatics, 42,* 557–584.

Riazantseva, A. (2002). Second language proficiency and pausing: A study of Russian speakers. *Studies in Second Language Acquisition, 23,* 487–526.

Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge: Cambridge University Press.

Stam, G. (2006). Thinking for speaking about motion: L1 and L2 speech and gesture. *International Review Applied Linguistics, 44,* 145–171.

Tang, G., Brentari, D., González, C., & Sze, F. (2010). Crosslinguistic variation in the use of prosodic cues: The case of blinks. In Brentari (ed.), pp. 519–541.

Trouvain, J., & Gut, U. (2007). *Non-native prosody: Phonetic description and teaching practice*. Berlin: Mouton de Gruyter.

Tyrone, M., Nam, H., Saltzman, E., Mathur, G., & Goldstein, L. (2010). Prosody and movement in American Sign Language: A task-dynamics approach. Presented at the Prosody 2010 Conference, Chicago.

Wilbur, R. B. (1994). Eyeblinks & ASL phrase structure. *Sign Language Studies, 84,* 221–240.

Wilbur, R. B., & Patschke, C. G. (1998). Body leans and the marking of contrast in American Sign Language. *Journal of Pragmatics, 30,* 275–303.

Wilcox, S. (2007). Routes from gesture to language. In E. Pizzuto, P. Pietrandrea & R. Simone (eds.), *Verbal and signed languages: Comparing structures, constructs and methodologies*, pp. 107–131. Berlin: Mouton de Gruyter.

Wilcox, S., Rossini, P., & Pizzuto, E. A. (2010). Grammaticalization in sign languages. In Brentari (ed.), pp. 332–356.