

Prosody in a communication system developed without a language model

Lauren Applebaum^a, Marie Coppola^b, and Susan Goldin-Meadow^a

^aUniversity of Chicago / ^bUniversity of Connecticut

Prosody, the “music” of language, is an important aspect of all natural languages, spoken and signed. We ask here whether prosody is also robust across learning conditions. If a child were not exposed to a conventional language and had to construct his own communication system, would that system contain prosodic structure? We address this question by observing a deaf child who received no sign language input and whose hearing loss prevented him from acquiring spoken language. Despite his lack of a conventional language model, this child developed his own gestural system. In this system, features known to mark phrase and utterance boundaries in established sign languages were used to consistently mark the ends of utterances, but not to mark phrase or utterance internal boundaries. A single child can thus develop the seeds of a prosodic system, but full elaboration may require more time, more users, or even more generations to blossom.

Keywords: homesign, prosody, gesture, communication, sign language, intonational phrase

1. Introduction

Prosody, the intonation, rhythm, or “music” of language, is an important aspect of all natural languages. Prosody can convey structural information that, at times, affects the meaning we take from a sentence. For example, in English, we respond to a sentence ending in a high tone as if it is a question even if it is not syntactically a question (“you’re coming[↑]”; Bolinger 1983; Ladd 1992).

Prosodic structures, like all other language properties, are produced not only by the vocal cords of spoken language users, but also by the hands, faces, heads, and bodies of sign language users (e.g., Nespors & Sandler 1999; Dachkovsky &

Sandler 2009; Wilbur 2000; van der Kooij, Crasborn & Emmerik 2006). Although the physical realization of prosodic features differs in signed and spoken languages, their function is the same. In both modalities, prosody is used to mark structure at syntactic, semantic, and discourse levels of analysis (e.g., Benitez-Quiroz et al. 2014; Herrmann & Steinbach 2011; Nespor & Sandler 1999; Wilbur & Patschke 1998; van der Kooij, Crasborn & Emmerik 2006; Selkirk 1984).

We ask here whether prosody is robust not only across modality, but also across learning conditions. If a child were not exposed to conventional language and had to construct a communication system out of whatever materials are available, would that system contain prosodic structure? This question is difficult to address in typical language-learning environments as children are routinely exposed from birth to a model of a conventional language. However, some children cannot make use of the language model to which they are exposed — deaf children whose hearing losses prevent them from acquiring the spoken language that surrounds them. Moreover, most of these profoundly deaf children are born to hearing parents, who often choose not to expose them to a conventional sign language.

Despite their lack of a conventional model for language, deaf children in these circumstances develop their own gestural systems, called *homesigns*, which they use to communicate with the hearing people in their worlds. Homesign has been shown to have many of the properties found in natural languages (Goldin-Meadow 2003). We ask here whether homesign also has prosodic structure.

To begin our exploration of prosodic structure in homesign, we examine the literature on prosody in both spoken and signed languages. We begin with a description of prosody in spoken language, and then use these well-established findings as a foundation for our descriptions of prosodic structure in sign language. Finally, we use the combined literatures as the basis for our analyses of prosodic structure in homesign.

1.1 Prosody in spoken language

Prosody in spoken language manifests itself as “pitch, tempo, loudness, and pause” (Cutler, Dahan & Donselaar 1997). Prosody is a valuable component of language because it signals linguistic information suprasegmental to the words (Brentari & Crossley 2002), providing information that can disambiguate the semantics and syntax of a given utterance. A change in the intonation of a word in an utterance can alter the entire meaning of the message. For example, in the written phrase *cinnamon rolls and cookies*, we do not know whether the rolls and cookies are both made with cinnamon or whether the speaker is discussing cinnamon rolls and, separately, a batch of cookies. A pause or some other kind of rhythmic information — prosody — added to the utterance can help us make this distinction.

Prosodic phonology can be characterized by a set of hierarchical relationships, from the smallest level of analysis, the mora (most notably found in Japanese, Kubozono 1989), to the largest level of analysis, the phonological utterance (Nespor & Vogel 1986, 2007). Two structures are considered important at the phrase level: the phonological phrase and the intonational phrase. Each of these structures is marked by independent, unique features, which can differ by language (Nespor et al. 2008). Final lengthening and changes in pitch generally mark the phonological phrase; breaths, melodic contours, and changes in melody generally mark the intonational phrase.

Within the phonological hierarchy, the phonological phrase is subordinate to the intonational phrase, and one or more phonological phrases can often be found within a single intonational phrase. As each phonological phrase is fully contained within the intonational phrase, the first phonological phrase aligns with the beginning of the intonational phrase within which it is contained, and the last phonological phrase aligns with the end of the intonational phrase:

- (1) [[He was sad]_p [because his dog]_p [died]_p]_I

“He was sad because his dog died” constitutes an intonational phrase containing three phonological phrases; the beginning of the first phonological phrase, “he was sad,” aligns with the beginning of the intonational phrase, and the third phonological phrase, “died,” aligns with the end of the intonational phrase. In the event that the intonational phrase contains only one phonological phrase (e.g., if the entire utterance were “he was sad”), the beginning and end of the phonological phrase would coincide with the beginning and end of the intonational phrase; that is, they would be perfectly aligned. Thus, the features that mark phonological and intonational phrases coincide at intonational phrase boundaries. Because utterances in homesign tend to be short and therefore might contain only one phonological phrase, our initial focus is on the intonational phrase.

The intonational phrase has been related both to syntactic structures (like par-enthetical phrases; Nespor and Vogel 1986, 2007) and semantic structures (Selkirk 1984). Selkirk suggests that the intonational phrase is composed of “sense units,” or groups of words that both make sense together and can stand alone. This description of the intonational phrase as a grouping of semantically related information allows us to explore prosodic structure in a homesigner’s gestures. Homesigners string their gestures together in a semantically related way — their gesture strings express propositions (Goldin-Meadow & Mylander 1984; Goldin-Meadow 2003). These propositions are, in Selkirk’s (1984) terms, the sense units that comprise an intonational phrase. Although there is evidence that homesigns have rudimentary syntactic structures (e.g., Franklin, Giannakidou & Goldin-Meadow 2011), it is often difficult to distinguish syntactic from semantic structure in homesign (but see

Goldin-Meadow 1982). We therefore use the semantic proposition as the starting point for our prosodic analyses of homesign.

After having identified an appropriate unit of analysis, we need to establish which prosodic features to examine. We turn to the sign language literature to provide an inventory of manual (hands) and nonmanual (face, mouth, body) features that might play a prosodic role in homesign. There is widespread agreement that modern day conventional sign languages grew out of homesign systems (Newport & Supalla 2000; Supalla 2008; Coppola & Senghas 2010). By using the sign language literature as a guide to our analyses, we explore the features of an emerging nonverbal communication system, while at the same time exploring the roots of the very sign languages we use to guide our investigations.

1.2 Prosody in sign language

As mentioned earlier, even though prosody manifests itself differently in spoken and signed languages, the features in each modality serve similar functions. For example, the *breath*, which can mark intonational phrases in spoken languages, is comparable to the *blink* in sign languages (Wilbur 1994): both are necessary biological events; both can be only temporarily postponed; and both play similar roles by marking intonational phrases. Even beyond natural biological constraints, features in spoken language have their equivalent in sign language. For example, facial expressions are considered the melodies of sign languages (Dachkovsky & Sandler 2009); pitch accents and boundary tones are the melodies of spoken languages (Beckman & Pierrehumbert 1986). Research has also shown that prosodic markers used in sign languages, such as eye gaze, body leans, and eyebrow movements, are used and coordinated with speech in hearing individuals, a pattern that begins to develop as early as the second year of life (Balog & Brentari 2008).

As in the literature on spoken languages, research on established sign languages has identified markers for both phonological phrases and intonational phrases. Phonological phrase markers can appear on an individual sign. For example, in sign lengthening, the final movement of the sign is extended; in sign repetition, a sign is repeated more times than its citation form requires; for utterance internal pauses used as a prosodic feature, two signs are separated by a longer than usual period of time and the handshape becomes relaxed. Intonational phrases are marked by a different set of prosodic features: head tilts, body leans (Nespor & Sandler 1999; Wilbur & Patschke 1998), changes in facial expression (Dachkovsky & Sandler 2009), and eye blinks (Wilbur 1994; Tang et al. 2010). Because a signer can produce nonmanual markers along with manual markers, these features, which are not produced by the hands, can co-occur with the manual features that mark the phonological phrase (Brentari & Crossley 2002).

Although study of the world's sign languages is ongoing, research conducted thus far suggests that the features discussed above are near-universal. Their role in a specific sign language may differ, but the presence of both manual and nonmanual markers does not vary extensively across sign languages. For instance, Tang et al. (2010) compared the use of blinks to mark phrase boundaries across four unrelated sign languages: American Sign Language (ASL), Swiss-German Sign Language (DSGS), Hong Kong Sign Language (HKSL), and Japanese Sign Language (JSL). All four sign languages use blinks to mark intonational phrase boundaries, although the details of blink use differ between HKSL and the other three sign languages. This study, along with a number of other studies examining prosodic features cross-linguistically (blinks: Tang et al. (2010) for ASL, DSGS, HKSL, JSL; body leans: van der Kooij, Crasborn & Emmerik (2006) for Sign Language of the Netherlands (NGT), Wilbur & Patschke (1998) for ASL, Nespor & Sandler (1999) for Israeli Sign Language (ISL), Dachkovsky, Healy & Sandler (2013) for ISL and ASL), suggest that these features may be universal in sign languages.

Features such as head movements, final lengthening, repeated movements, etc. not only appear to be good candidates for sign language universals (Sandler & Lillo-Martin 2006), but their location and function are also very similar across sign languages. For example, Nespor and Sandler (1999) find brow movements, holds, and pauses, to be prominent at the ends of intonational phrases in Israeli Sign Language, and similar findings have been reported for ASL (Wilbur 1999; Malaia & Wilbur 2011) and NGT (Van der Kooij & Crasborn 2008). Many of the same features have thus been found to mark phonological or intonational phrase boundaries across sign languages.

Prosody clearly has its place in established languages, both signed and spoken. However, it is not clear that a child who does not have access to conventional language input will display prosodic structure in the absence of a community of signers. In the next section, we provide background on homesign and the homesigner who allows us to explore the robustness of prosody.

1.3 Background on homesign

Most children learn language from their parents, but homesigners do not. As mentioned earlier, homesigners are not able to learn the spoken language that surrounds them, and their hearing parents have not exposed them to a sign language. To communicate with others, homesigners typically turn to gesture. A series of studies by Goldin-Meadow and colleagues have shown that, even without conventional language input, homesigners introduce many language-like properties into their individual gesture systems. For example, Goldin-Meadow, Mylander, and Franklin (2007) and Goldin-Meadow, Mylander, and Butcher (1995) found that

young homesigners produce gestures that contain handshape and motion components that recombine to create new gestures, displaying a rudimentary morphological system. In addition, Goldin-Meadow, Butcher, Mylander, and Dodge (1994) found that the homesigner abbreviates gestures serving noun functions, and inflects gestures serving verb functions, thus drawing a distinction between two fundamental grammatical categories. As a final example, homesigners display consistent production probability patterns (e.g., they are more likely to produce a gesture for the entity playing a patient role than for the entity playing the agent role) and ordering patterns (e.g., they place gestures that represent entities playing a patient role before gestures that represent the action) in their gesture strings (Goldin-Meadow & Feldman 1977; Feldman, Goldin-Meadow & Gleitman 1978; Goldin-Meadow & Mylander 1984). Interestingly, these sentence- and word-level patterns are found in homesigners across different cultures (American, Chinese), despite the fact that the patterns are not seen in the gestures produced by the homesigners' hearing parents (Goldin-Meadow & Mylander 1983, 1998; Goldin-Meadow et al. 1994, 1995, 2007). This finding provides further evidence that homesigners are not learning a language, or even a gestural system, from their parents.

Given the importance of prosody in natural languages, and the fact that homesigners are able to incorporate many language-like features (e.g., recursion: Goldin-Meadow 1982; hierarchical structure: Hunsicker & Goldin-Meadow 2011; displacement: Butcher, Mylander & Goldin-Meadow 1991; Morford & Goldin-Meadow 1997) into their gesture systems, it is possible that prosody may also be a characteristic of homesign. We explore this possibility by examining manual and nonmanual prosodic features in the gestures produced by a young American homesigner whom we call "David."

1.4 Research questions and hypotheses

In this study, we explore two aspects of David's prosodic system. First, we ask whether David uses the prosodic elements found in established sign languages — the manual (e.g. holds, repetitions, and emphatic movements) and nonmanual (e.g., head tilts and head nods) movements that have been attested in other sign languages.

After searching for features found in established sign languages, our second goal is to look for evidence of prominence in David's prosodic system, prominence at the utterance level and prominence internal to the utterance (e.g., at intonational phrase boundaries). There is, in fact, mixed evidence for a *single* prosodic feature that reliably marks phrase, clause, or utterance boundaries in sign languages. For example, Tang et al. (2010) find that eye blinks reliably mark intonational phrase

boundaries in three sign languages, but Fenlon et al. (2007) find that prosodic markers are produced probabilistically at clause boundaries, and do not occur at very high rates even in native signers. In a study most comparable to our own in that it relies on naturalistic data, Hansen and Heßmann (2007) examined the relationship between formal features (manual signs, head nods, blinks, and pauses) and propositional content in a spontaneously signed conversation in Deaf signers of German Sign Language (DGS). They found that each of the features tends to appear at proposition-based sentence boundaries (i.e., sentence boundaries determined by meaning, which is how we define sentence boundaries). Importantly, however, none of the features functions exclusively as a boundary marker (i.e., in an all-or-none fashion). Taken together, the sign language literature suggests that if David does show evidence of a developing prosodic system, we should not expect to see extremely strong, all-or-nothing patterns.

In sum, we ask how a young homesigner uses features that have been found to mark prosodic structure in established sign languages at boundaries (both utterance boundaries and clause boundaries). The strength of our study is two-fold. First, unlike many sign language studies that examine prosodic structure in monologues (but see Hansen & Heßmann 2007), our data come from unscripted interactions in a home environment. Second, our study of prosody builds on analyses showing that the particular homesign system we examine has structure at a number of linguistic levels (Goldin-Meadow et al. 1994; Goldin-Meadow et al. 1995; Goldin-Meadow 1982; Goldin-Meadow & Mylander 1983; Goldin-Meadow & Mylander 1990). This rich, detailed context allows us to relate prosodic structure to structure at a number of other linguistic levels. Our study can thus shed light on the development of prosody in a homemade gesture system, a system that gives us a glimpse into the evolution of sign language.

2. Methods

2.1 Participant and procedure

David was born profoundly deaf (>90dB bilateral hearing loss) with no known cognitive deficits. His parents and siblings were all hearing, did not know sign language, and had no contact with signing Deaf individuals.¹ David attended a

1. We use the convention of referring to individuals who have a hearing loss and who use sign language as their primary form of communication with a capital “Deaf,” and to individuals who have a hearing loss but who use speech as their primary form of communication with lower case “deaf.”

preschool that used an oral method of deaf education. Oral methods advocate intense training in sound sensitivity, lipreading (or speechreading) and speech production, and discourage using conventional sign language and gesture with the child.

In general, a child with a severe hearing loss is unable to hear even shouted conversation and cannot learn speech by conventional means. A child with a profound loss such as David's hears only occasional loud sounds and these sounds may be perceived as vibrations rather than sound patterns. Amplification serves to increase awareness of sound but often does not increase the clarity of sound patterns (Mindel & Vernon 1971; Moores 1982). Although David wore hearing aids, they were largely ineffective in providing him with enough auditory information to understand human speech. Moreover, the visual information one gets from observing a speaker's lips is rarely sufficient to allow severely and profoundly deaf children to learn spoken language (Conrad 1979; Farwell 1976; Summerfield 1983). Visual cues are generally ambiguous with respect to speech; the mapping from visual cues to words is one-to-many. In order to constrain the range of plausible lexical interpretations, other higher-order classes of information (e.g., the phonological, lexical, syntactic, semantic, and pragmatic regularities of a language) must come into play during speechreading. The most proficient speechreaders are those who can use their knowledge of the language to interpret an inadequate visual signal (Conrad 1977), and post-lingually deafened individuals (people who had knowledge of a language before losing their hearing) are generally more proficient speechreaders than individuals who have been deaf from birth (Summerfield 1983). Since speechreading appears to require knowledge of a language to succeed, it is difficult for a deaf child like David to *learn* language solely through speechreading. Indeed, we found that David had only a few spoken words in his repertoire, all of which were recognizable in very specific contexts and were never combined with each other (see Goldin-Meadow & Mylander 1984).

In addition, at the time of our observations, David had no contact with signers (adults or children) at school or at home, and thus had no exposure to a conventional sign system, either American Sign Language or Signed English. A native ASL signer was asked to review David's videotapes to determine whether his homesigns had been influenced by ASL. The signer reported that David did not produce ASL signs for even the simplest concepts, thus confirming that he had not yet come into contact with a conventional sign language. Even at age 9, the handshapes David used in his gestures were structured differently from the handshapes found in ASL (Singleton, Morford & Goldin-Meadow 1993).

David did, however, see the gestures that his hearing parents produced as they talked to him. Hearing speakers routinely gesture when they speak (McNeill 1992) and David's parents were no exception. Although it is possible that the gestures

David's parents produced provided a model for the homesign system David developed, our previous work provides no evidence for this hypothesis. The structure that characterizes child homesign at both word- and sentence-levels cannot be traced back to the gestures that the children's hearing parents produced (Goldin-Meadow & Mylander 1983, 1984; Goldin-Meadow et al. 1994, 1995, 2007; Hunsicker & Goldin-Meadow 2012). Even the gestures that David's hearing sister produced when she was prompted to gesture in an experimental situation looked qualitatively different from the gestures David produced under the same conditions (Singleton et al. 1993).

In terms of gestural input to prosody, although the vocal cords are the primary vehicle for prosodic information in spoken language users, a speaker's hands and body can provide cues to this information (McNeill 1979; Alibali, Kita & Young 2000; Graf Cosatto, Strom & Huang 2002). Some aspects of prosody might then have been accessible to David through gesture, at least in principle. However, it is important to note that the relationship that these gestural cues hold to speech is not available to a deaf child, which would prevent David from learning about the role that the cues play in sentences, in other words, from learning about prosodic structure.

The goal of the present study is to determine whether prosody is a property of language that can be developed despite severely degraded learning conditions — David had no conventional sign language input and his hearing losses acted as a massive filter on reception of speech, preventing spoken language data from reaching him in an undistorted form (cf. Swisher 1989). Thus, the properties of language that appear in David's gestures have developed under radically atypical language-learning conditions; i.e., they are so over-determined that they arise even under acquisition conditions that are significantly degraded.

As part of a longitudinal study, David was videotaped at home every few months (Goldin-Meadow 1979; Goldin-Meadow & Mylander 1984). During a session, David interacted naturally with experimenters, his parents, or two older hearing siblings. A large toy bag, puzzles, and books were brought on each occasion to initiate play and interaction. Recording sessions lasted as long as David cooperated, generally about 2 hours. We coded and analyzed data from three sessions taken over a two year period when David was 3;05, 3;11, and 5;02 (years; months of age).

2.2 Identifying gestures and gesture utterances

The child's gestural communication was first segmented into individual gestures. To be included in the dataset, a gesture had to be communicative (i.e., directed toward a communication partner). In addition, a gesture had to be symbolic and not

a functional act. For example, attempting to twist off a jar lid would not constitute a gesture even though it does communicate information; in contrast, twisting the hand over (but not on) the jar, while looking at the communication partner, would constitute a gesture.

Gestures were described in terms of the formational parameters used to describe sign language: handshape, movement, place of articulation, and palm orientation. We coded the beginning of a gesture as the point at which the handshape was fully formed (i.e., not lax or moving into position). In cases where there was no handshape change to distinguish the end of one gesture and the beginning of another, we used the onset of movement. For example, when the homesigner produced two successive points, his handshape typically did not change, but his arm and hand moved from one location to another; this arm movement indicated the start of a new gesture. A gesture ended when the final handshape began to relax or when the child's arm began to drop (see Goldin-Meadow (1979) and Goldin-Meadow & Mylander (1984) for additional details).

Gestures could be produced in sequence. We borrowed a criterion often used in studies of sign language to determine the boundaries between gesture utterances. Relaxation of the hand accompanied by a pause or a drop of the hands after a gesture or series of gestures was taken to signal the end of a string; that is, to demarcate an utterance boundary (see Goldin-Meadow (2003, Chapter 7) for evidence validating this criterion). For example, if David pointed to a toy and then, without pausing or relaxing his hand, pointed to a table, the two pointing gestures were considered "within an utterance." The same two pointing gestures interrupted by a relaxation of the handshape and a pause or a drop of the hands would be classified as two isolated gestures. Note that we are not using the term "utterance" in the sense that it is often used in the linguistic literature, that is, to refer to a change in conversational turn. Rather, we adopt Schiffrin's (1994) view, which defines utterances as "units of language production [...] that are inherently contextualized"; in the context of our study, utterances are the spontaneous productions of the child in a naturalistic setting.

2.3 Attributing meaning to gestures and to gesture utterances

Gestures were classified into three categories: deictic gestures, characterizing gestures, and markers. Deictic gestures include points to objects, places, or people, as well as hold-ups (in which an object is held up in the partner's line of sight to draw attention to it). Homesigners use points to indicate entities and to stand for semantic arguments in utterances (Feldman et al. 1978), as do Deaf children learning sign language from their Deaf parents (cf. Hoffmeister 1978). In this sense, homesigners' deictic gestures function as early nouns or demonstratives do

in children acquiring a conventional language (see Hunsicker & Goldin-Meadow (2012), for evidence that David combined demonstrative points with nouns to form complex nominal constituents).

Characterizing gestures are iconic; their form captures an aspect of the intended referent, either action or attribute information. Depending on the context, characterizing gestures can function either as nouns, verbs, or adjectives (Goldin-Meadow et al. 1994). For example, the child moves his extended arms up and down to mimic the movement of a bird or butterfly, which would be glossed as FLY if used to refer to the object's actions, or BIRD if used to refer to the object itself. As another example, the child touches the index finger to the thumb forming a circle, which would be glossed as ROUND if used to refer to an attribute of an object, or as PENNY if used to refer to the object itself.

Markers are gestures that serve to modulate utterances. They typically are drawn from the gestures used by hearing speakers in the child's community; for example, a side-to-side headshake used to negate, a two-handed flip used to question or express emotion (see Franklin, Giannakidou & Goldin-Meadow 2011). A marker can take both manual (flip) and nonmanual (nod) forms.

As described in the preceding section, we used motoric criteria to determine the boundaries of a gesture utterance. We then determined how many different propositions were conveyed by the gestures within the bounds of that utterance (see Goldin-Meadow & Mylander 1984, for a more detailed description of the kinds of propositions young homesigners convey). We grouped gestures into propositions based on predicate frames, and assigned each gesture in the proposition a semantic role. For example, in propositions expressing transfer of an object to another location, homesigners produced gestures for the transferring act, the actor, the object being moved, and the location to which the object was moved. Homesigners rarely produced all of the semantic elements allowable in a predicate frame within a single utterance. For example, to indicate that the experimenter gave him a soldier toy, David explicitly produced gestures for the act (GIVE), the patient (*soldier toy*), and the recipient (*me*); the actor (*you/experimenter*) was not explicitly signed and was inferred from context. The probability with which each of the potential elements in a predicate frame was produced across sets of utterances provides evidence that the actor, although not explicitly gestured, was indeed part of the predicate frame underlying the utterance (see Goldin-Meadow 1985, 1987).

Utterances containing more than one proposition were considered complex. Example (2) displays a complex utterance containing two propositions. In this example, David is describing two pictures, one of a toy cowboy and another of a toy soldier. He describes the actions that each of these toys typically does (although neither toy is shown doing the action in the picture); each proposition contains the information relevant to the unique toy. David is thus conveying two propositions

(the cowboy sips a straw; the soldier beats a drum) within the bounds of a single utterance.²

(2) age 3;11

[SIP—cowboy—SIP—STRAW]₁—[soldier—BEAT]₂

Agreement between two independent coders ranged from 93% to 95% for identifying and assigning meanings to gestures and gesture strings (see Goldin-Meadow & Mylander (1984) for further discussion of the coding system).

2.4 Coding prosody in homesign

The first author, who did not do the form and meaning coding, conducted the prosodic coding. Internal reliability was assessed by having the coder recode a subset of the data two years after the initial coding. There was good intra-observer reliability on prosodic coding (kappa score = .82).

Our specific hypothesis was that we might find differences in the prosodic prominence of a gesture as a function of its position in the utterance. We included in the analyses only those gesture utterances containing more than one manual gesture simply because the gesture in a one-gesture utterance is both the beginning and the end of the utterance and thus not relevant to our hypothesis.

We used ELAN (Lausberg & Sloetjes 2009), a language annotation software package distributed by the Max Planck Institute for Psycholinguistics, Nijmegen, to code prosodic features.³ ELAN organizes annotations on multiple tiers, thus allowing a visual representation of simultaneously produced, overlapping prosodic features to be easily created. To determine whether David's use of prosodic features changed over developmental time, we examined prosodic features at age 3;05 (237 utterances), 3;11 (185 utterances) and 5;02 (194 utterances). Following the sign language literature (Baker & Padden 1978; Wilbur 2000; Nespor & Sandler 1999), we coded three manual prosodic features (holds, repetitions, emphatic movements) and two nonmanual features (head tilts, nods).

We recognize at the outset a technical limitation of our data. We were unable to code as many of the prosodic features (particularly the nonmanual features) that have been described in sign languages as we would have liked, largely due

2. In the utterance transcriptions, we gloss deictic gestures (points and hold-ups) with lower-case letters and characterizing/iconic gestures with capital letters. The English gloss for a single gesture may contain multiple words. Dashes separate the gestures within an utterance, and brackets divide the propositions. The numerical subscripts indicate the number of each proposition within the utterance.

3. <http://tla.mpi.nl/tools/tla-tools/elan/>

to the challenge of capturing the face of a young child during unconstrained play sessions. The prosodic features that we leave for future research include facial expression, eyegaze shifts, blinks, brow raises, and brow furrows. Given the central role that facial expression plays in defining intonational contours in established sign languages, our inability to capture David's face may limit the claims we can make about prosodic structure in his gestures. However, the nonmanual features that we were unable to code tend to mark larger phrase structures in sign languages (Nespor & Sandler 1999; Wilbur 1994; Wilbur & Patschke 1999), and these boundaries are also often marked by manual features. Thus, the five features that we were able to observe should provide us with basic information about David's ability to prosodically mark larger phrase structures.

2.4.1 *Manual prosodic features*

We coded three manual features — holds, repetitions, and emphatic movements — each of which has been shown to play a role in sign language prosody (Nespor & Sandler 1999; Fenlon et al. 2007; Wilbur & Martínez 2002). As described earlier, dropping the hand and/or pausing (i.e., a relax in handshape without subsequently dropping the hands) were used to identify the ends of utterances in our coding system (Goldin-Meadow & Mylander 1984). We therefore did not include these markers as prosodic features as, by definition, they occur at the ends of utterances. Manual prosodic features are superimposed on manual gestures. Because the features are punctate and can occur with only one gesture at a time, coding decisions were made using the individual gesture as the unit of analysis. In other words, each gesture in an utterance was annotated individually for the presence or absence of each of the three manual prosodic features. The three manual prosodic features that David used are described below. Note that a single gesture could, in principle, be marked by none, one, or more than one feature.

Holds. A hold was coded when the final articulation of the gesture remained in the same handshape and overall position for a minimum of 10 frames (there was



Figure 1. An example of a held sign

no frame maximum) at a camera speed of 30 frames/second (i.e., 300 ms duration). A hold ended when the handshape began to relax or the arm/elbow position began to change, whichever occurred first. An example of a hold can be seen in Figure 1. The time code at the top of each frame (hour:minute:second:frame; 30 frames per second) indicates how long David sustained the hold on this pointing gesture. The first two frames show him holding his finger in an extended position; the final frame shows the end of the hold (total hold time = 38 frames; approximately 1.25 seconds).

Repetitions. A repetition was coded when a gesture was reproduced in the same manner and in the same location immediately following its first production. Repetitions must be contiguous, i.e., the repetition immediately follows the first iteration of the gesture with no interruption by the interlocutor between repetitions. We excluded repetitions of iconic gestures with inherent aspect of iterative motion (e.g., a gesture indicating beating a drum) because it was difficult to determine whether the repetition was a prosodic feature or part of the lexical gesture.

Emphatics. An emphatic was coded when the gesture was produced with high velocity or force. High velocity or force typically involved the more distal movements of the shoulder or elbow. Emphatics were quick, forceful movements of the hand that contrasted with the more smoothly produced movements of the surrounding gestures.

2.4.2 *Nonmanual prosodic features*

The nonmanual features in our analyses were head tilts and nods, both of which have been shown to play a role in sign language prosody (Liddell 1978, 1980). Nonmanual prosodic features can co-occur simultaneously with manual gestures, but they can also occur on their own at the beginning or end of a string of gestures, or in between two gestures. When a nonmanual feature was produced on its own and occurred at the end of the utterance, it was assumed to mark the last gesture (see example (3) where the nod is coded as marking the sign MOVE at the end of the utterance).

- (3) age 5;02
 MOVE—toys—MOVE_{nod}

When a nonmanual feature was produced on its own between two gestures, we arbitrarily assumed that the feature marked the preceding gesture on the assumption that a nonmanual feature on its own is marking a semantic boundary and should therefore be “anchored” to the preceding gesture; see example (4) where the nod is coded as marking the gesture that precedes it, *light*. Theoretically, this coding decision applies to all nonmanual features; however, the only nonmanual

feature that David produced on its own between two gestures was the nod, which occurred between two gestures only once in our data.

- (4) age 5;02
light—nod—ROUND LIGHT—GO STRAIGHT

If a nonmanual feature was produced at the beginning of the utterance, it was not assigned to a gesture simply because no gesture preceded it. By not considering the first gesture marked, we remain consistent with our decision to attribute marking only to the gesture prior to a sequential nonmanual marker. Although this coding decision has the potential to bias our findings by reducing the number of utterance initial gestures that could potentially be marked, only a few nonmanual features were actually produced at the beginning of an utterance (4 utterances at 3;05, 2 at age 3;11, and 8 at age 5;02); moreover, if we remove these utterances from the database, the patterns described in subsequent sections do not change.

Nonmanual prosodic features can also co-occur with one or more gestures. Nonmanual features that continue over multiple gestures serve to group those gestures, and thus have two boundaries — one at the beginning of the group (the onset of the nonmanual) and one at the end (the offset of the nonmanual). As a result, when a group of two or more gestures that co-occurred with a nonmanual feature (e.g., a head tilt) was preceded by at least one other gesture (i.e., an utterance containing at least three gestures), we considered not only the offset of the nonmanual (i.e., the last gesture that co-occurred with the head tilt), but also the onset of the nonmanual (the gesture that came before the head tilt), to be marked. Coding the onset and offset of continuous nonmanual features allows us to capture prosodic *change* between gestures or between a series of gestures. For example, in (5) we consider both POUR (the gesture at the end of the nod) and the first *kitchen* (the deictic gesture preceding the nod) to be prosodically marked; in this way, we acknowledge the role that continuous nonmanual features can play in marking a group of gestures by separating those gestures off from the rest of the utterance.

- (5) age 5;02
_____ nod
 [kitchen—mother—kitchen]₁—[POUR (soda)]₂

When a nonmanual feature co-occurred with only one gesture (and thus did not serve to demarcate a group of gestures), we did not consider both its onset and offset to be marked and, in this way, did not artificially inflate the number of gestures coded as prosodically marked. Thus, in example (6), only *naughty* (the single gesture that co-occurred with the head tilt) was considered marked. The majority of nonmanual features in our dataset spanned a single gesture (25/44 at age 3;11; 47/63 at age 5;02; 43/64 at age 5;11).

(6) age 3;11

head tilt
[xylophone—BEAT]₁—[naughty]₂

There was one other exception to the practice of assigning two boundaries to non-manual features co-occurring with other gestures. If the nonmanual feature co-occurred with a gesture that began the utterance and the feature extended through to the final gesture, we did not consider the first gesture marked. Thus, in example (7), only the second GIVE at the end of the utterance, and not the first GIVE at the beginning of the utterance, was assumed to be marked with a head tilt. Assigning an end marking and not a beginning marking to utterances in which the non-manual feature extends over the entire utterance has the potential to artificially inflate end markings. However, the decision has the advantage of being consistent with our decision to count nonmanual features that extend over several gestures as a single marking on the final gesture (see example 5). There were, moreover, few instances of this type (4 at age 3:05, 8 at age 3;11, and 5 at age 5:02) and removing these utterances from the database does not alter the patterns described in the next section.

(7) age 3;11

tilt
GIVE—point to turtle toy—GIVE

As in our analysis of manual prosodic features, each gesture was evaluated for the presence or absence of a nonmanual prosodic feature. Each prosodic feature was coded separately; it was therefore possible for a gesture to be assigned more than one nonmanual feature. The two nonmanual prosodic features that David used are described below.

Head Tilts. A head tilt was coded when the head moved from a neutral position to one side; the tilt was considered completed when the head began its return



Figure 2. An example of a head tilt. David's forehead is tilted off of the vertical line of his body when he produces the second gesture (the point), but not the first gesture (hands behind back); as a result, the second gesture was considered marked by the head tilt (see text).

to a neutral position. In Figure 2, the child's head is in neutral space during the first gesture (hands behind his back), but is tilted during the second gesture (point at the toys behind the experimenter's back). Following our coding rule for non-manual features that co-occur with a single gesture, we considered only the gesture that co-occurred with the head tilt (i.e., the point) to be marked.

Nods. A nod was coded when the head moved down from its original position; the nod was considered completed when the head returned to its original position. Nods, like head tilts, can, in principle, co-occur with several gestures; however, all but two of the nods that David produced spanned only one gesture.

3. Results

3.1 The distribution of prosodic features in David's homesigns

If David's homesign patterns like established signed languages, we would expect him to produce more prosodic cues at the ends of his utterances than at other utterance locations. Ideally, we would compare his distribution to native child signers. Unfortunately, there are no analyses of these distributions in child signers (nor, for that matter, in deaf adults). Hansen and Heßman (2007) do provide some insight into which prosodic markers are produced by adults in spontaneous signing, and where in the utterance they appear. However, they do not compare how often prosodic features are produced at structural boundaries vs. non-boundaries. Although they conclude that prosodic features are produced at boundaries (see Table 2 in Hansen & Heßman 2007), they do not present production rates at non-boundary locations (i.e., a baseline rate). Our study is the first to establish a baseline by examining production rates at non-boundary locations, as well as boundary locations. We begin by describing the distribution of prosodic features in David's homesigns independent of position (Table 1); we then turn to an analysis of prosodic cues (manual and nonmanual) as a function of position (Tables 2 and 3).

Table 1 presents the frequency of each of the five prosodic features as a proportion of the total number of prosodic features that David produced. For example, at age 3;05, David produced a total of 363 prosodic features; of those 363 features, 115 (or 32%) were holds. Each of the manual and nonmanual features occurred at least once during the sessions analyzed, with some (emphatics) occurring as often as 130 times at a single developmental time point. Note that David used each of the three manual features more frequently than the two nonmanual features we were able to analyze, namely, head tilts and nods, although the proportion of each nonmanual did increase over development; the proportion of manual holds also

Table 1. The number of prosodic features of each type that David produced per session, expressed as a proportion of the total number of prosodic features for that session.

Age	Hold	Repetition	Emphatic	Head Tilt	Head Nod
3;05 Total = 363	0.32	0.18	0.36	0.07	0.07
3;11 Total = 368	0.45	0.19	0.18	0.16	0.03
5;02 Total = 283	0.41	0.17	0.11	0.19	0.12

increased slightly over time in conjunction with a decrease in the proportion of manual emphatics.

In sign languages, prosodic marking (or prominence) tends to be strongest at the end of an intonational phrase, which coincides with the end of the utterance (Nespor & Sandler 1999). To determine whether David's homesign system followed this pattern, we first divided the gestures in our analysis into those occurring at the End of an utterance, and those occurring at the Beginning or Middle of an utterance (recall that we excluded all one-gesture utterances from our database). We then calculated how many manual features (repetitions, emphatics, holds) and nonmanual features (nods, head tilts) were produced on Beginning, Middle, and End gestures, and divided that number by the number of gestures in that position. For example, in his two-gesture utterances, David produced 301 End gestures (and, of course, 301 Beginning gestures as well). He produced 156 manual prosodic features on those 301 End gestures, and 126, manual prosodic features on the 301 Beginning gestures, resulting in a manual prosodic marking score of .52 (156/301) for Ends and .42 (126/301) for Beginnings.

Table 2 presents the mean number of manual and nonmanual prosodic markers that David produced on gestures occurring at the End of the utterance, compared to gestures occurring at the Beginning or Middle of the utterance, collapsed over the three developmental time points. We performed the calculations separately for utterances containing only two gestures (with Beginning and End, but no Middle, gestures) and for utterances containing three or more gestures (with Beginning, Middle, and End gestures). By doing so, we allowed for the possibility that longer utterances displayed a different pattern from shorter utterances. However, the data in Table 2 indicate that David produced more prosodic features at the Ends of his utterances than at the Beginnings, regardless of the length of the utterance (2 gestures vs. 3 or more gestures).

Table 2 also highlights the fact that David used both manual and nonmanual prosodic features at the Ends of his utterances more often than at any other

Table 2. The mean number of manual and nonmanual prosodic features David produced on gestures occurring in utterances containing 2 gestures and utterances containing 3 or more gestures as a function of position (Beginning, Middle, End). Standard errors are in parentheses. N's are the number of gestures in a given utterance location.

Type of prosodic feature	Utterances containing 2 gestures		Utterances containing 3 or more gestures		
	Beginning	End	Beginning	Middle	End
	(n = 301)	(n = 301)	(n = 313)	(n = 630)	(n = 313)
Manual (repetitions, emphatics, holds)	.42 (.03)	.52 (.04)	.41 (.03)	.36 (.02)	.55 (.04)
Nonmanual (head tilts, nods)	.06 (.01)	.15 (.02)	.11 (.02)	.06 (.01)	.22 (.03)

position in the utterance. The proportions are higher for manual features than for nonmanual features because manual features were more common in David's gestures than nonmanuals, and because we coded 3 manual features and only 2 nonmanual ones. The important point, however, is that the pattern (Ends higher than Beginnings and Middles) was the same for both types of features.

To determine whether the End-marking pattern emerges over development, we calculated the average number of prosodic features David produced in utterances containing 2 gestures and utterances containing 3 or more gestures at each developmental time point. Because manual and nonmanual features followed similar patterns, we combined them in this and all subsequent analyses. Table 3 presents the data and shows that David produced more prosodic features at the Ends of his utterances than at any other position at each developmental session.

In order to statistically confirm this finding, we conducted a Poisson mixed effects regression with random effects for utterance ($sd=0.195$), fit using the

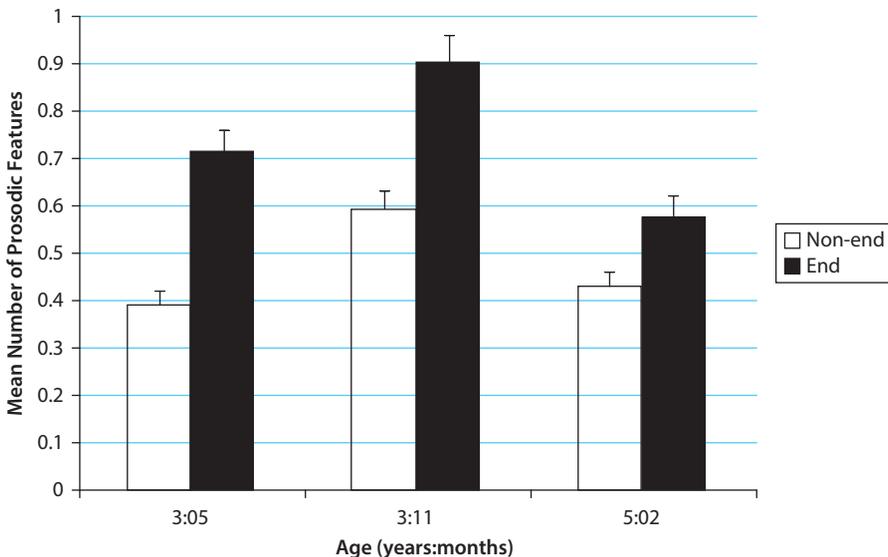
Table 3. The mean number of prosodic features David produced at each developmental session on gestures occurring in utterances containing 2 gestures and utterances containing 3 or more gestures as a function of position (Beginning, Middle, End). Standard errors are in parentheses. N's are the number of gestures in each position.

Age	Utterances containing 2 gestures		Utterances containing 3 or more gestures		
	Beginning	End	Beginning	Middle	End
3;05	.43 (.06) n = 121	.69 (.07) n = 121	.36 (.05) n = 115	.39 (.03) n = 267	.73 (.07) n = 115
3;11	.62 (.08) n = 81	.90 (.09) n = 81	.65 (.07) n = 103	.54 (.06) n = 156	.90 (.08) n = 103
5;02	.41 (.05) n = 99	.46 (.06) n = 99	.55 (.07) n = 95	.38 (.04) n = 207	.68 (.07) n = 95

Table 4. Results from the Mixed Effect Regression analysis comparing Non-Ends vs. Ends of Utterances.

	β	<i>SE</i>	<i>z</i>	<i>p</i>
(Intercept)	-0.90	0.06	-14.59	<.001***
Ends of Utterance	0.44	0.06	6.93	<.001***
Session 2	0.35	0.08	4.48	<.001***
Session 3	-0.03	0.08	-0.33	>.10

Laplace approximation. Because utterances containing 2 gestures and utterances containing 3 or more gestures showed similar patterns, we did not include utterance length as a factor in the statistical analysis. Table 4 provides estimated coefficients, standard errors, and significance levels of the fixed effects. In this model, we investigate the effects of both Session (Session 1 vs. 2 vs. 3) and Gesture Location (End of Utterance vs. Non-End; Beginnings and Middles were combined to form the Non-End category). Because we are comparing multiple categorical variables, one Session value (Session 1) and one Gesture Location value (Non-End) served as base categories and thus are not listed separately in Table 3 but are represented by the intercept value in the table. For this model, the positive and negative coefficients for Sessions 2 and 3 are thus interpreted in relation to Session 1. In other words, a positive coefficient suggests an increase in prosodic marking relative to Session 1; a negative coefficient suggests a decrease in prosodic marking relative to

**Figure 3.** Mean number of prosodic features on gestures occurring in Non-End and End positions in David's utterances containing 2 or more gestures at each age.

Session 1. Similarly, the coefficient for End of Utterance is relative to Non-End of Utterance, the base category for Gesture Location.

The significant positive coefficient for End of Utterance in Table 4 ($\beta = .44$, $p < .001$) suggests that David reliably marks Ends more frequently than Non-Ends. Note, however, that he also displays more prosodic marking in Session 2 relative to Session 1 ($\beta = .35$, $p < .001$), but not in Session 3 relative to Session 1 ($\beta = -0.03$, $p > .10$). In addition, we found that the decrease in prosodic marking between Sessions 2 and 3 is significant, $\beta = -.38$, $p < .001$. This pattern can be seen in Figure 3, which displays the development of David's prosodic feature marking over time. The important point, however, is that when we explored the interaction between Session and Gesture Location, we found that the added interaction term did not significantly improve the fit of the model, $\chi(2) = 4.15$, $p > .1$, indicating that David produced more prosodic features on the Ends of utterances (compared to Non-Ends) at all three sessions.

3.2 Is David marking the ends of sentences or the ends of propositions?

Our data suggest that David uses prosodic marking more often at the ends of his utterances than at the beginning or middle. The end of an utterance coincides with the end of a proposition in David's data. Perhaps David is, in fact, prosodically marking the ends of propositions, which also happen to occur at the ends of utterances. To determine whether David marks the ends of propositions or the ends of utterances, we need to examine utterances that contain two or more propositions; that is, complex utterances (cf. Goldin-Meadow 1982). We excluded the few utterances where it was difficult to unambiguously determine proposition boundaries within a complex utterance (16 at age 3;05; 19 at age 3;11; 11 at age 5;02). David thus produced 65 codable complex utterances at Session 1, 31 at Session 2, and 67 at Session 3. We divided gestures occurring in the complex utterances that David produced at these three ages into three categories: (1) End of Utterance gestures, which occurred at the end of an utterance (and at the end of a proposition); (2) End of Proposition gestures, which occurred at the end of a proposition but not at the end of an utterance; (3) Non-End gestures, which occurred in any other position. Example (8) illustrates how this coding system was implemented using the complex utterances presented earlier.

- (8) [SIP — cowboy — SIP — STRAW]₁ — [soldier — BEAT]₂
 Non-End Non-End Non-End End of Proposition Non-End End of Utterance

The end of an utterance necessarily co-occurs with the end of a proposition. Thus, David might have used his prosodic features to mark Ends of Propositions rather than Ends of Utterances. However, the data in Table 5 suggest that he did not

Table 5. The mean number of manual and nonmanual prosodic features David produced on gestures occurring in complex utterances as a function of position (End of Utterance; End of Proposition; Non-End). Standard errors are in parentheses; N's are the gestures on which each prosodic score is based.

Type of prosodic Feature	Non-End	End of Proposition	End of Utterance
Manual (repetitions, emphatics, holds)	0.38 (.04) n = 208	0.33 (.03) n = 281	0.52 (.05) n = 163
Nonmanual (head tilts, nods)	0.12 (.02) n = 208	0.07 (.02) n = 281	0.25 (.04) n = 163

— when we look at utterances containing two or more propositions, we find that David prosodically marked Ends of Propositions preferentially only when they co-occurred with the End of Utterances, and did so for both manual and nonmanual prosodic features.

Moreover, we find the same pattern at each developmental session. David produced more prosodic features on the Ends of Propositions only when they co-occurred with the Ends of Utterances (see Table 6).

Table 6. The mean number of prosodic features David produced at each developmental session on gestures occurring in complex utterances as a function of position (End of Utterance; End of Proposition; Non-End). Standard errors are in parentheses; N's are the gestures on which each prosodic score is based.

Age	Non-End	End of Proposition	End of Utterance
3;05	.42 (.07) n = 69	.45 (.05) n = 127	.77 (.10) n = 65
3;11	.59 (.10) n = 44	.42 (.09) n = 40	1.03 (.17) n = 31
5;02	.51 (.07) n = 95	.35 (.05) n = 114	.66 (.09) n = 67

The pattern seen in Tables 5 and 6 is supported by results from a Poisson mixed effects regression with random effects for utterance ($sd = 0.37$), fit using the Laplace approximation. Table 7 provides estimated coefficients, standard errors, and significance of the fixed effects. Coefficients for Sessions are compared to Session 1, the reference level for the model. Coefficients for End of Proposition and End of Utterance are compared to Non-End.

Once again, we see a significant positive coefficient for End of Utterance ($\beta = .44, p = .002$), suggesting that David preferentially marks the ends of his utterances. However, the marginally significant *negative* coefficient for End of Proposition ($\beta = -.20, p = .16$) indicates that he does not preferentially mark the end

Table 7. Results from the Mixed Effects Regression analysis for complex utterances.

	β	<i>SE</i>	<i>z</i>	<i>P</i>
(Intercept)	-.78	.14	-5.64	<.001***
End of Utterance	.44	.14	3.16	.002**
End of Proposition	-.20	.14	-1.40	.16
Session 2	.23	.17	1.36	>.1
Session 3	-.07	.14	-0.48	>.1

of a proposition unless that proposition is also the end of an utterance. Note that, although David displays more prosodic marking in his complex utterances during Session 2 relative to Session 1 ($\beta = .23$), the effect is not significant. Moreover, by Session 3, he is producing fewer prosodic markings relative to Session 1 ($\beta = -.07$), although this value, too, is not significant. We also found that the decrease in prosodic marking from Session 2 to Session 3 was marginally significant ($p = .08$). However, importantly, when we explored the interaction between Session and Gesture Location in David's complex utterances, we found that the interaction coefficients did not significantly improve the fit of the model, $\chi^2(4) = 3.14$, $p > .1$. The findings thus suggest that David produced more prosodic features on the gestures found at the ends of utterances, but not at the ends of propositions, and that this pattern does not change over time.

4. Discussion

4.1 Using prosodic cues to mark the ends of sentences

We have found that a young American homesigner, who did not have access to a conventional sign language model, nevertheless incorporated prosodic features that are found in many sign languages into his homesign system — manual features (holds, repetitions, and emphatics) and nonmanual features (head nods and head tilts) were found at every session studied. More importantly, David used these features more often at the ends of his gesture utterances than at any other utterance position. This pattern resembles many established sign languages, including American Sign Language and Israeli Sign Language (Nespor & Sandler 1999).

Note that David did not use his prosodic features only at the ends of utterances — he also used them at utterance-internal positions (i.e. intonational phrase boundaries), though not as often. There are several possible explanations for this finding. First, David's intonational phrases might have contained more than one phonological phrase, and David could have been using his prosodic features to

mark the boundaries of those smaller phrases. To explore this possibility, we would need to develop an independent criterion to identify the ends of phonological phrases in David's gesture system. Second, manual and nonmanual features play multiple roles in established sign languages (Sandler & Lillo-Martin 2006; Sandler 2010; Wilbur 2009), and it is possible that David used these features on gestures appearing in non-end positions to express linguistic information that is not prosodic. Finally, although we know that the prosodic features we have investigated do occur at the ends of sign utterances in established sign languages (e.g., Nespor & Sandler 1999), we do not know whether these prosodic features also occur at non-end positions in these languages (and if they do, how often they occur in these positions). Indeed, Wolford (2012) found that although prosodic features occur more frequently at the ends of utterances (particularly in native signing adults and native signing older children), they also occur at other positions within the utterance (see also Hansen & Heßmann 2007; Fenlon et al. 2007). Thus, prosodic markers are found in non-end positions even in established sign languages.

4.2 Gestural input to homesign

David did not have access to ASL (or, for that matter, to Signed English), and thus could not have copied the prosodic patterns found in his gestures from a model of an established signed language. However, he did see the manual and nonmanual gestures produced by his family members and school teachers (who were not signers), and he might have used those gestures as a model for his prosodic features, particularly since the hands and body have been shown to provide cues to prosodic structure in speakers (e.g., McNeill 1979; Alibali et al. 2000; Graf et al. 2002).

Although hearing speakers routinely use many of the nonmanual features found in sign languages, they use them differently from signers (e.g., Pyers & Emmorey 2008). Moreover, gesture and speech form an integrated system in hearing speakers (McNeill 1992; Balog & Brentari 2008), but David had access to only one part of that system — the gestures. Thus, even if the hearing speakers who surrounded him marked the ends of their spoken utterances with particular manual and nonmanual movements, David would have access only to the movements themselves — not to the fact that they occurred in particular positions in their spoken utterances. David may have borrowed his prosodic features from the hearing individuals in his world, but the pattern that he imposed on these features vis à vis his gestural system is not likely to have come from his hearing parents, particularly in light of the fact that David's mother rarely combined her gestures into strings and thus had no gestural combinatorial system to speak of (Goldin-Meadow & Mylander 1983, 1984).

4.3 Production and perceptual pressures that might lead to prosodic marking

Why might a homesigner produce prosodic cues more often at the ends of his gesture sentences than at other positions? One possibility is that there is a motoric basis to the pattern, that is, the head may naturally tilt or nod, and the hands may be held, repeat their movement, or be used for emphasis, at the end of a thought. If so, we might expect to see this pattern in hearing speakers as well (although, as mentioned earlier, it would be hard for a homesigner who does not have access to speech to detect the pattern). Future work will be needed to explore this possibility.

Another possibility is that the pattern serves a function for the listener. In fact, Fenlon et al. (2007) found that native signers and non-signers were able to identify sentence boundaries in a sign language with which they were unfamiliar. Using video clips of narratives in an unfamiliar sign language, Fenlon and colleagues asked participants to push a button when they thought they saw a sentence boundary. Because the participants were unfamiliar with the meaning of the signs in the language, the only cues they could use to make boundary decisions were the prosodic cues, which they were able to utilize reasonably well. Brentari et al. (2011) found similar results for hearing non-signers. Thus, there may be pressures on both the perception and production sides that would lead a homesigner to treat the ends of his sentences differently from other positions.

4.4 The impact of short sentences on prosodic marking

David's prosodic system is much simpler than systems found in conventional languages. Unlike conventional languages (both signed and spoken), which mark not only the ends of utterances but also units within the utterance, David used his prosodic marking only to mark the utterance-final boundary. However, we need to view this result with caution for several reasons. First, we coded only a subset of the prosodic features found in sign languages; the quality and nature of our videotapes made it impossible for us to code features marked on the face, including eye blinks. It is therefore possible that David did mark within-utterance boundaries using devices that we were unable to measure. Second, and related to this point, the devices that mark within-utterance boundaries might be different from those that mark utterance-final boundaries, as is found in established sign languages (Nespor & Sandler 1999). If we had looked at a second set of prosodic features, we might have found that David did mark utterance-internal boundaries (although it is worth noting that manual and nonmanual features patterned in the same way at both within-utterance and utterance-final boundaries in David's gestures). Finally, the relatively small number of complex utterances that David produced may have made it difficult to discover within-utterance boundary markings in these data.

Assuming for the moment that David's gesture system really does lack utterance-internal prosody, one way to account for this lack is to appeal to the length of David's utterances. David's utterances are generally very short (excluding one-gesture utterances, the mean number of gestures David produced in his utterances we observed was 3.02). Short sentences do not require parsing in the same way that long sentences do. Perhaps if, over time, David's utterances were to increase in length, we would find within-utterance boundary markers in the longer utterances (importantly, the mean length of David's gesture utterances did not increase over the three time points we observed: 3.13 at 3;05; 2.83 at 3;11; and 3.07 at 5:02).

4.5 Insights into language creation from homesign

The length of David's utterances may have obscured our finding more complex prosodic structure in those utterances. However, it is also possible that a child developing a homesign system without a communication partner to share the system may be unable to go further than David has gone in developing prosodic structure. This possibility is supported by the fact that the mean number of prosodic features David produced on the last gesture in his utterances did *not* steadily increase over the two-year period during which he was observed; if anything, the mean number declined at age 5;02. If a homesigner were allowed to use his or her gesture system into adulthood (e.g., adult homesigners in Nicaragua; see Coppola & Newport 2005; Brentari, Coppola, Mazzoni & Goldin-Meadow 2012), perhaps such a system would develop prosodic structure that is linguistically richer than the child homesigner's system.

Increases in the alignment of prosodic features have, in fact, been observed over longer periods of development and in larger communities. Consider, for example, the sign language that is currently evolving in a Bedouin tribe in Israel where many of the members carry a gene for deafness. Al-Sayyid Bedouin Sign Language (ABSL), as the language is called (Sandler et al. 2005; Aronoff et al. 2008), has been used for five generations, but many language properties do not yet seem to be fully developed or have only recently gained the consistency one would expect from a mature language. One example is prosody. Sandler and colleagues (2011) investigated the use of prosodic markers in two generations of ABSL users, and found an increase in the alignment of features (i.e., a "pile-up" of features at boundaries) over historical time; that is, the first-generation signers displayed less alignment than the later-generation signers. The fact that David did not display a steady increase in the alignment of prosodic features over ontogenetic time suggests that groups, and perhaps generations, of signers may be necessary for a communication system to take the next step toward a fully developed prosodic system.

We see what may be a similar step-wise pattern in the use of handshapes in classifier predicates. Brentari, Coppola, Mazzoni & Goldin-Meadow (2012) asked adult Nicaraguan homesigners to describe two types of vignettes, one in which a hand was shown putting an object on a surface (agentive), and another in which the object was shown sitting on the surface (non-agentive). The researchers then compared the homesigners' gestures to signs produced by Deaf signers of ASL or Italian Sign Language (LIS) asked to describe the same events, and to gestures produced by hearing adults asked to describe the two types of vignettes using only gesture (i.e., without speech). Signers of both ASL and LIS used greater finger complexity (i.e., handshapes involving more selected fingers) in the handshapes they used to represent *objects* than in the handshapes they used to represent the *handling* of objects. Hearing gesturers showed the opposite pattern: less finger complexity in *object* handshapes than in *handling* handshapes. The adult homesigners resembled Deaf signers with respect to finger complexity in *object* handshapes, suggesting that the seeds of this structural opposition are present in homesign. However, the homesigners differed from Deaf signers with respect to finger complexity in *handling* handshapes — they displayed more finger complexity than did the signers (although less complexity than did the silent gesturers). Thus, as in the development of prosodic structure, having a group, and perhaps generations, of signers may be necessary for homesigners to fully exploit handshape complexity as a marker for the contrast between agentive and non-agentive events.

5. Conclusion

We have shown that, even without benefit of a conventional language model, a child is able to incorporate prosodic marking into his homemade gestural system. Interestingly, however, the child uses his prosodic features only to mark the ends of utterances; that is, to distinguish between one utterance and the next — and not to mark within-utterance boundaries. Nor does the child display an increase with age in the alignment of prosodic features at utterance boundaries, an increase that has been observed over generations in a recently emerging sign language (Sandler et al. 2011). Although a single child is able to develop the seeds of a prosodic system, a fully elaborated system may require more time, perhaps more users, and even more generations to blossom.

Acknowledgements

Supported by R01DC00491 from the National Institute of Deafness and Other Communicative Disorders to SGM and P30DC010751 to Marie Coppola and Diane Lillo-Martin. We thank Carolyn Mylander for her assistance in data coding and management. We also thank Amy Franklin, Dea Hunsicker, Natasha Abner, Diane Brentari, and Anastasia Giannakidou for their helpful comments and discussions, and Matthew Carlson for his statistical assistance.

References

- Alibali, Martha W., Sotaro Kita & Amanda J. Young. 2000. Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes* 15(6). 593–613. DOI: 10.1080/016909600750040571
- Aronoff, Mark, Irit Meir, Carol Padden & Wendy Sandler. 2008. The roots of linguistic organization in a new language. *Interaction Studies* 9(1). 133–153. DOI: 10.1075/is.9.1.10aro
- Baker, Charlotte & Carol Padden. 1978. Focusing on the non-manual components of American Sign Language. In Patricia Siple (ed.), *Understanding language through sign language research*, 27–58. New York, NY: Academic Press.
- Balog, Heather L. & Diane Brentari. 2008. The relationship between early gestures and intonation. *First Language* 28(2). 141–163. DOI: 10.1177/0142723708088722
- Beckman, Mary. E. & Janet B. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255–309. DOI: 10.1017/S095267570000066X
- Benitez-Quiroz, C. Fabian, Kadir Gökgöz, Ronnie B. Wilbur & Aleix M. Martinez. 2014. Discriminant features and temporal structure of nonmanuals in American Sign Language. *PLoS ONE* 9(2): E86268. DOI:10.1371/journal.pone.0086268.
- Bolinger, Dwight. 1983. Intonation and gesture. *American Speech* 58(2). 156–174. DOI: 10.2307/455326
- Brentari, Diane, Marie Coppola, Laura Mazzoni & Susan Goldin-Meadow. 2012. When does a system become phonological? Handshape production in gesturers, signers, and homesigners. *Natural Language and Linguistic Theory* 30(1). 1–31. DOI: 10.1007/s11049-011-9145-1
- Brentari, Diane & Laurinda Crossley. 2002. Prosody on the hands and face: Evidence from American Sign Language. *Sign Language & Linguistics* 5(2). 105–130. DOI: 10.1075/sll.5.2.03bre
- Brentari, Diane, Carolina González, Amanda Seidl & Ronnie Wilbur. 2011. Sensitivity to visual prosodic cues in signers and nonsigners. *Language and Speech* 54(1). 49–72. DOI: 10.1177/0023830910388011
- Butcher, Cynthia, Carolyn Mylander & Susan Goldin-Meadow. 1991. Displaced communication in a self styled gesture system: Pointing at the nonpresent. *Cognitive Development* 6(3). 315–342. DOI: 10.1016/0885-2014(91)90042-C
- Conrad, Reuben. 1977. Lip-reading by deaf and hearing children. *British Journal of Educational Psychology* 47. 60–65. DOI: 10.1111/j.2044-8279.1977.tb03001.x
- Conrad, Reuben. 1979. *The deaf child*. London: Harper & Row.
- Coppola, Marie & Elissa L. Newport. 2005. Grammatical Subjects in home sign: Abstract linguistic structure in adult primary gesture systems without linguistic input. *Proceedings of*

- the National Academy of Sciences of the United States of America* 102(52), 19249–19253. DOI: 10.1073/pnas.0509306102
- Coppola, Marie & Ann Senghas. 2010. Deixis in an emerging sign language. In D. Brentari (ed.), *Sign languages: A Cambridge language survey*, 543–569. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511712203.025
- Cutler, Anne, Delphine Dahan & Wilma van Donselaar. 1997. Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40(2). 141–201.
- Dachkovsky, Svetlana, Christina Healy & Wendy Sandler. 2013. Visual intonation in two sign languages. *Phonology* 30(2). 211–252. DOI: 10.1017/S0952675713000122
- Dachkovsky, Svetlana & Wendy Sandler. 2009. Visual intonation in the prosody of a sign language. *Language and Speech* 52(2-3). 287–314. DOI: 10.1177/0023830909103175
- Farwell, Roberta M. 1976. Speech reading: A research review. *American Annals of the Deaf* 121. 19–30.
- Feldman, Heidi, Susan Goldin-Meadow & Lila Gleitman. 1978. Beyond Herodotus: The creation of language by linguistically deprived deaf children. In Andrew Lock (ed.), *Action, gesture and symbol: The emergence of language*, 351–414. New York, NY: Academic Press.
- Fenlon, Jordan, Tanya Denmark, Ruth Campbell & Bencie Woll. 2007. Seeing sentence boundaries. *Sign Language & Linguistics* 10(2). 177–200. DOI: 10.1075/sll.10.2.06fen
- Franklin, Amy, Anastasia Giannakidou & Susan Goldin-Meadow. 2011. Negation, questions, and structure building in a homesign system. *Cognition* 118. 398–416. DOI: 10.1016/j.cognition.2010.08.017
- Graf, Hans Peter, Eric Cosatto, Volker Strom & Fu Jie Huang. 2002. Visual prosody: Facial movement accompanying speech. *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*. 381–386.
- Goldin-Meadow, Susan. 1979. Structure in a manual communication system developed without a conventional language model: Language without a helping hand. In Haiganoosh Whitaker & Harry Whitaker (eds.), *Studies in neurolinguistics*, Vol. 4, 125–209. New York: Academic Press. DOI: 10.1016/B978-0-12-746304-9.50010-0
- Goldin-Meadow, Susan. 1982. The resilience of recursion: A study of a communication system developed without a conventional language model. In Eric Wanner & Lila R. Gleitman (eds.), *Language acquisition: The state of the art*, 51–77. New York: Cambridge University Press.
- Goldin-Meadow, Susan. 1985. Language development under atypical learning conditions: Replication and implications of a study of deaf children of hearing parents. In Keith E. Nelson (ed.), *Children's language*, Vol. 5, 197–245. Hillsdale, N.J.: Lawrence Erlbaum & Associates.
- Goldin-Meadow, Susan. 1987. Underlying redundancy and its reduction in a language developed without a language model: Constraints imposed by conventional linguistic input. In Barbara Lust (ed.), *Studies in the acquisition of anaphora: Vol. II, applying the constraints*, 105–133. Boston, MA: D. Reidel Publishing Company. DOI: 10.1007/978-94-009-3387-3_6
- Goldin-Meadow, Susan. 2003. *The resilience of language*. New York, NY: Psychology Press.
- Goldin-Meadow, Susan & Heidi Feldman. 1977. The development of language-like communication without a language model. *Science* 197(4301). 401–403. DOI: 10.1126/science.877567
- Goldin-Meadow, Susan & Carolyn Mylander. 1983. Gestural communications in deaf children: The non effect of parental input on language development. *Science* 221. 372–374. DOI: 10.1126/science.6867713

- Goldin-Meadow, Susan & Carolyn Mylander. 1984. Gestural communication in deaf children: The effects and non effects of parental input in early language development. *Monographs of the Society for Research in Child Development* 49(3). No. 207. DOI: 10.2307/1165838
- Goldin-Meadow, Susan & Carolyn Mylander. 1990. Beyond the input given: The child's role in the acquisition of language. *Language* 66. 323–355. DOI: 10.2307/414890
- Goldin-Meadow, Susan & Carolyn Mylander. 1998. Spontaneous sign systems created by deaf children in two cultures. *Nature* 391. 279–281. DOI: 10.1038/34646
- Goldin-Meadow, Susan, Cynthia Butcher, Carolyn Mylander & Mark Dodge. 1994. Nouns and verbs in a self styled gesture system: What's in a name? *Cognitive Psychology* 27. 259–259. DOI: 10.1006/cogp.1994.1018
- Goldin-Meadow, Susan, Carolyn Mylander & Cynthia Butcher. 1995. The resilience of combinatorial structure at the word level: Morphology in self-styled gesture systems. *Cognition* 56(3). 195–262. DOI: 10.1016/0010-0277(95)00662-1
- Goldin-Meadow, Susan, Carolyn Mylander & Amy Franklin. 2007. How children make language out of gesture: Morphological structure in gesture systems developed by American and Chinese deaf children. *Cognitive Psychology* 55(2). 87–135. DOI: 10.1016/j.cogpsych.2006.08.001
- Hansen, Martje & Jens Heßmann. 2007. Matching propositional content and formal markers. *Sign Language & Linguistics* 10(2). 145–175. DOI: 10.1075/sll.10.2.05han
- Herrmann, Annika & Markus Steinbach. 2011. Nonmanuals in sign languages. *Sign Language & Linguistics* 14(1). 3–8. DOI: 10.1075/sll.14.1.02her
- Hoffmeister, Robert. 1978. *The development of demonstrative pronouns, locatives and personal pronouns in the acquisition of American Sign Language by deaf children of deaf parents*. PhD dissertation, University of Minnesota.
- Hunsicker, Dea & Susan Goldin-Meadow. 2012. Hierarchical structure in a self-created communication system: Building nominal constituents in homesign. *Language* 88(4). 732–763. DOI: 10.1353/lan.2012.0092
- Kubozono, Haruo. 1989. The mora and syllable structure in Japanese: Evidence from speech errors. *Language and Speech* 32(3). 249–278.
- Ladd, D. Robert. 1992. An introduction to intonational phonology. In Gerard J. Docherty & Robert Ladd (eds.), *Papers in laboratory phonology II: Gesture, segment, prosody*, 321–334. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511519918.014
- Lausberg, Hedda & Han Sloetjes. 2009. Coding gestural behavior with the NEUROGES ELAN system. *Behavior Research Methods, Instruments, & Computers* 41(3). 841–849. DOI: 10.3758/BRM.41.3.841
- Liddell, Scott. 1978. Nonmanual signals and relative clauses in American Sign Language. In Patricia Siple (ed.), *Understanding language through sign language research*, 59–90. New York: Academic Press.
- Liddell, Scott. 1980. *American Sign Language syntax*. The Hague: Mouton.
- Malaia, Evie & Ronnie B. Wilbur. 2011. Kinematic signatures of telic and atelic events in ASL predicates. *Language and Speech* 55(3). 407–421. DOI: 10.1177/0023830911422201
- McNeill, David. 1979. *The conceptual basis of language*. Hillsdale, NJ: Erlbaum.
- McNeill, David. 1992. *Hand and mind*. Chicago, IL: University of Chicago Press.
- Mindel, Eugene D. & McCay Vernon. 1971. *They grow in silence: The deaf child and his family*. Silver Springs, MD: National Association of the Deaf.
- Moore, Donald F. 1982. *Educating the deaf: psychology, principles and practices* (2nd ed.). Boston: Houghton Mifflin.

- Morford, Jill P. & Susan Goldin-Meadow. 1997. From here and now to there and then: The development of displaced reference in homesign and English. *Child Development* 68(3). 420–435. DOI: 10.2307/1131669
- Nespor, Marina & Wendy Sandler. 1999. Prosody in Israeli Sign Language. *Language and Speech* 42(2-3). 143–176. DOI: 10.1177/00238309990420020201
- Nespor, Marina & Irene Vogel. 2007. *Prosodic phonology*. Berlin: Walter de Gruyter [originally published in 1986, Dordrecht: Foris]. DOI: 10.1515/9783110977790
- Nespor, Marina, Mohinish Shukla, Ruben van de Vijver, Cinzia Avesani, Hanna Schraudolf & Caterina Donati. 2008. Different phrasal prominence realizations in VO and OV languages. *Lingue e Linguaggio* 2. 1–29.
- Newport, Elissa L. & Ted Supalla. 2000. Sign language research at the millennium. In Karen Emmorey & Harlan L. Lane (eds.), *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima*, 103–114. Mahwah, NJ: Lawrence Erlbaum.
- Pyers, Jennie E. & Karen Emmorey. 2008. The face of bimodal bilingualism: Grammatical markers in American Sign Language are produced when bilinguals speak to English monolinguals. *Psychological Science* 19(6). 531–536. DOI: 10.1111/j.1467-9280.2008.02119.x
- Sandler, Wendy. 2010. Prosody and syntax in sign languages. *Transactions of the Philological Society* 108(3). 298–328. DOI: 10.1111/j.1467-968X.2010.01242.x
- Sandler, Wendy & Diane Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139163910
- Sandler, Wendy, Irit Meir, Svetlana Dachkovsky, Carol Padden & Mark Aronoff. 2011. The emergence of complexity in prosody and syntax. *Lingua* 121(13). 2014–2033. DOI: 10.1016/j.lingua.2011.05.007
- Sandler, Wendy, Irit Meir, Carol Padden & Mark Aronoff. 2005. The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences* 102(7). 2661–2665. DOI: 10.1073/pnas.0405448102
- Schiffrin, Deborah. 1994. *Approaches to discourse*. Oxford: Blackwell.
- Selkirk, Elizabeth. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Singleton, Jenny L., Jill P. Morford & Susan Goldin-Meadow. 1993. Once is not enough: Standards of well-formedness in manual communication created over three different timespans. *Language* 69. 683–715. DOI: 10.2307/416883
- Summerfield, A. Quentin. 1983. Audio-visual speech perception, lipreading and artificial stimulation. In M. E. Lutman & M. P. Haggard (eds.), *Hearing science and hearing disorders*, 132–179. New York, NY: Academic Press.
- Supalla, Ted. 2008. Sign language archeology: Integrating historical linguistics with fieldwork on young sign languages. In Ronice de Quadros (ed.), *Sign languages: Spinning and unraveling the past, present and future*, 574–583. Petrópolis, Brazil: Editora Arara.
- Swisher, M. Virginia. 1989. The language-learning situation of deaf students. *Teachers of English to Speakers Other Languages Quarterly* 23. 239–257.
- Tang, Gladys, Diane Brentari, Carolina González & Felix Sze. 2010. Crosslinguistic variation in prosodic cues. In Diane Brentari (ed.), *Sign languages*, 519–542. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511712203.024
- van der Kooij, Els & Onno Crasborn. 2008. Syllables and the word-prosodic system in Sign Language of the Netherlands. *Lingua* 118. 1307–1327. DOI: 10.1016/j.lingua.2007.09.013

- van der Kooij, Els, Onno Crasborn & Wim Emmerik. 2006. Explaining prosodic body leans in Sign Language of the Netherlands: Pragmatics required. *Journal of Pragmatics* 38(10). 1598–1614. DOI: 10.1016/j.pragma.2005.07.006
- Wilbur, Ronnie B. 1994. Eyeblinks and ASL phrase structure. *Sign Language Studies* 84. 221–240. DOI: 10.1353/sls.1994.0019
- Wilbur, Ronnie B. 1999. Stress in ASL: Empirical evidence and linguistic issues. *Language and Speech* 2-3. 229–250. DOI: 10.1177/00238309990420020501
- Wilbur, Ronnie B. 2000. Phonological and prosodic layering of nonmanuals in American Sign Language. In Karen Emmorey & Harlan L. Lane (eds.), *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima*, 213–244. Mahwah, NJ: Lawrence Erlbaum.
- Wilbur, Ronnie B. 2009. Productive reduplication in a fundamentally monosyllabic language. *Language Sciences* 31. 325–343. DOI: 10.1016/j.langsci.2008.12.017
- Wilbur, Ronnie B. & Aleix Martínez. 2002. Physical correlates of prosodic structure in American Sign Language. *Chicago Linguistic Society* 38. 693–704.
- Wilbur, Ronnie B. & Cynthia G. Patschke. 1998. Body leans and the marking of contrast in ASL. *Journal of Pragmatics* 30. 275–303. DOI: 10.1016/S0378-2166(98)00003-4
- Wilbur, Ronnie B. & Cynthia G. Patschke. 1999. Syntactic correlates of brow raise in ASL. *Sign Language & Linguistics* 2. 3–40. DOI: 10.1075/sll.2.1.03wil
- Wolford, George. 2012. *Acquisition of prosody in American Sign Language*. Master's thesis, Purdue University. West LaFayette, IN.

Corresponding author's address

Lauren Applebaum
Department of Psychology
University of Chicago
5848 South University Ave
Chicago, IL 60637-1554
USA

lapplebaum@uchicago.edu

Copyright of Sign Language & Linguistics is the property of John Benjamins Publishing Co. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.