Study on Segmentation Markers of Intonational Phrases in Chinese Sign Language

5. Study on Segmentation Markers of Intonational Phrases in Chinese Sign Language

Zhu Xiao
Ludong University, PRC

Abstract

In this chapter, I divide Shanghai sign utterances into intonational phrases based on the principle of "pause", study the duration of "pause", analyze the prosodic markers at the end of intonational phrases, such as manual markers include holding, amplitude and repetition of two hands and so on, and non-manual markers including head nod, eye blinks and so on. This chapter selects natural corpus and uses ELAN to label it in several layers and choose markers that appeared frequently to analyze and summarize them. According to the analysis, the "pause" time between intonational phrases varies according to gender, age, corpus property and environment, but most of them are distributed between 120ms and 300ms. At the end of an intonational phrase, there is one or more of such markers generally present, such as hand holding, large amplitude, repetition, nodding, blinking, and the like.

5.1. Introduction

American scholar William Stokoe (1960) explained the theory of sign language linguistics, and for the first time put forward that sign language and spoken language have the same functions. Nespor and Vogel (1986; 2007) proposed prosodic stratification from low to high: pitch, syllable, foot, prosodic word, phonological phrases, intonational phrases, and utterances.

Shanghai sign language is a dialect of Chinese sign language (CSL). The chapter takes Shanghai Sign Language discourse as a corpus in order to study the type and frequency of segmentation markers of intonational phrases (IP) in CSL. Gong (2009) pointed out that CSL is the language that is used by native signers belonging to the Chinese deaf community. The data we discuss comes from four signers who have been exposed to sign language from a very young age. Table 5-1 shows the corpus situation. Signer A provided a notice about "Severe Disability Subsidies". Signer B gave a story about "Ruban Learn Skills" through several pictures without Chinese words. Signer B reported a story named "Duck, chicken and wolf story" through several pictures without Chinese words. Signer D gave news about "House Demolition".

The four videos are input to ELAN (Eudico Linguistic Annotator) for hierarchical labeling and for creating a hierarchy of different markers. The complete time of the four

videos is about 8 minutes 40 seconds which are segmented into 445 IPs and we get statistics from them.

Table 5-1 Corpus situation

SIGNER	GENDER	AGE	DEAF AGE	PARENTS	PROPERTY	DURATION	IP
A	Male	48	1	Deaf	Notice	3min24s	175
B	Female	47	1	Hearing	Story	1min40s	80
C	Female	60	3	Hearing	Story	2min04s	100
D	Male	48	7	Hearing	News	1min38s	90

This chapter aims to:

Provide segmentation: segment IPs of the collected corpus according to "pause", the "pause" time between IPs was learned by doing statistics and observing the difference of "pause" time between different signers and the shortest and longest "pause" time.

Observe segmentation markers: after segmenting an utterance into several IPs, observe segmentation markers appearing at the boundary of IPs and learn the frequency at the end of IPs and the frequency of markers at the end in all IPs by doing statistics on them.

The chapter is structured as follows. Section 5.2 summarizes the "pause" time between two adjacent IPs after segmenting IPs by "obvious pause". Section 5.3 provides a list of cues that appeared in CSL including six frequent and three infrequent markers. Section 5.4 is the core of the chapter, in which I explore the importance of the cues that are listed above. Special emphasis is laid on the statistics of the frequency of those cues.

5.2. Pause between IPs

Pause here is not defined by no signing at all, but pause is the transitional part of signs. Most of the pauses are weak pauses, which means that the hands are still raised but relaxed instead of being dropped to the signer's lap. The pause between IPs is more obvious than that between phonological phrases or prosodic words. From Table 5-2, we can see that the pause duration is from 70ms to 480ms, and the average duration is about 150ms.

Table 5-2 Pause situation in four signers

SIGNER	NUMBER	DURATION (MS)	AVERAGE TIME (MS)
A	131	70–480	149
B	69	100–480	170
C	100	100–250	138
D	86	100–390	164

5.3. Prosodic Markers in CSL

The following display, Figure 5-1, shows the seven prosodic markers that most frequently appeared in CSL. Five manual markers contained hold (written as h), repeat (written as +++), amplitude (written as Amp.) for dominant hand and relax for non-dominant hand.

Three non-manual markers contained head nod (written as hn), head down/back/up (written as hd/hb/hu), eye blinks (written as eb). As the non-dominant hand is always relaxed between two IPs, compared to the non-dominant hand marker between phonological phrases, we just analyze the other six markers here.

This study refers to the intonational phrase markers found by Nakatani and Dukes 1977, Grosjean 1977; 1979, Baker and Padden 1978, Green 1984, Selkirk 1986, Wilbur 1994; 2009, Nespor and Sandler 1999, Sandler and Lillo-Martin 2006, Infantino *et al*. 2007, Fenlon 2010, Fenlon *et al*. 2007, Hoza 2008, Nicodemus 2008; 2009, Dachkovsky and Sandler 2009, Herrmann 2010, Sandler 2010, Sandler *et al*. 2011, Brentari *et al*. 2011, Borowsky *et al*. 2012, Crasborn *et al*. 2012, Khan *et al*. 2012, which research American sign language, Irish sign language, German sign language, Israeli sign language etc., I combine them with Shanghai Sign Language corpus to summarize the types of markers that appear in Shanghai Sign Language (as shown in Figure 5-1).
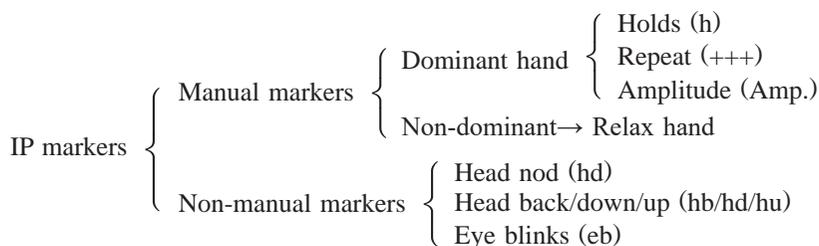

Figure 5-1   Seven markers that often appeared in four signers

5.3.1. Manual Markers
5.3.1.1 Hold

At the last prosodic word of IPs, one handshape would be held for a longer duration and the whole sign would be influenced to be longer. Once one sign has the feature of hold, the duration of the sign is longer than the IP–internal sign with the same meaning.

(1) [NAME LISTEN STRANGE—]$_{IP}$ [NAME LISTEN—]$_{IP}$
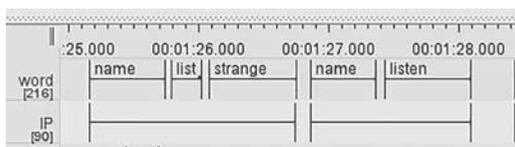   The name sounds strange.

Figure 5–2   Elan screenshot of annotation in sentence (1)

Figure 5–3   Elan screenshot of word duration in sentence (1)

The sign *listen* at the end of IP (Figure 5-3, line 136) is longer than that of IP-internal (Figure 5-3, line133), and the length of the two signs (Figure 5-3, line 134, 136) at the end of the two IPs is longer than others internally, which can be seen visually in Figure 5-2.

5.3.1.2 Repeat

At the end of IPs, one syllable would be repeated two or more times for prosodic emphasis. Take *SAME* (Figure 5-4) for example, the sign is repeated six times at the end of the IP in (2).

(2) [THERE LIVE SAME+++]IP
      People live in the same house there.
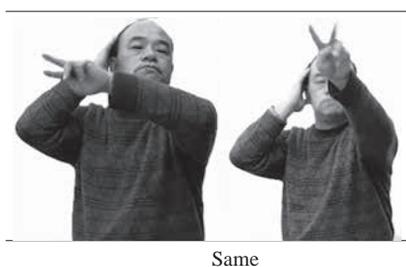
Same

Figure 5-4 Screenshot of *same* syllable

5.3.1.3 Amplitude

The same as hold and repeat, amplitude is also a marker for the end of an IP. At the end of IPs, the sign movement space is often larger than IP-internal signs. The height amplitude of the last sign *CROSS* (Figure 5-5) in (3) is larger than the other three signs internal.

(3) [BOAT SIT BOAT CROSS(Amp.)]IP
      Roban cross the river by boat.

| Boat | Sit | Boat | Cross |

Figure 5-5 Screenshot of prosodic words in sentence (3)

5.3.2. Non-manual Markers
5.3.2.1 Head Nod

At the end of IPs, there are often head nod features in the four signers' talking. This is illustrated in Figure 5-6. There is a head nod feature at the end of the two IPs in (4).
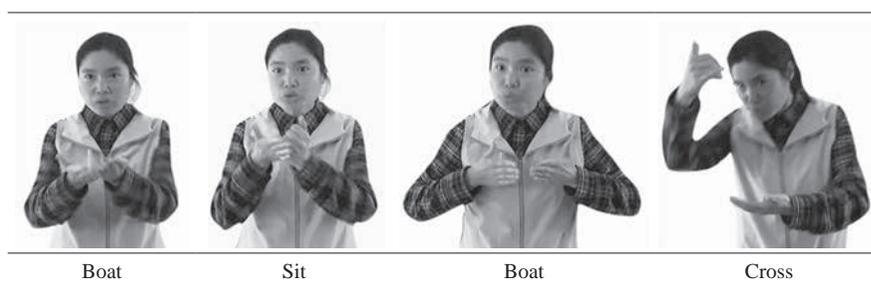
$$\underline{\hspace{1.5cm}\text{hn}\hspace{1.5cm}}\qquad\qquad\underline{\hspace{1.5cm}\text{hn}\hspace{1.5cm}}$$

(4) [LIVE HOTEL−1]$_{IP}$ [2 MARKET BUSINESS]$_{IP}$

First, there will build hotel and second, build market.

LIVE	HOTEL−1

2	MARKET	BUSINESE
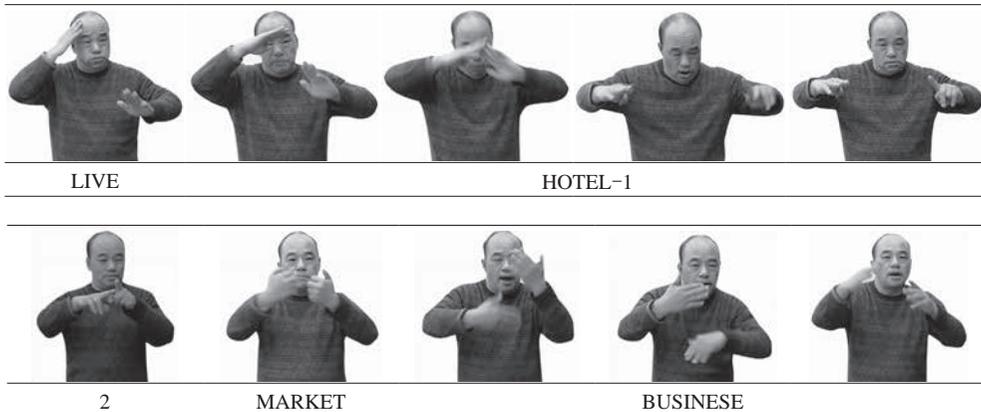
Figure 5-6   Screenshot of prosodic words in sentence (4)

5.3.2.2 Head Back/Up/Down

There is often a head back or head down marker at the beginning and a head up marker at the end of an IP. Head back means that the signer's head moves backward (Figure 5-7), which often happens at the beginning of IPs. Head down means that the signer's head moves downward (Figure 5-8), which also often happens at the beginning of IPs. Head back or head down occurs in different situations. There are head up markers at the end of IPs in both types.

$$\underline{\hspace{1.2cm}\text{hb}\hspace{1.2cm}}\qquad\qquad\underline{\hspace{1.2cm}\text{hu}\hspace{1.2cm}}$$

(5) [LIVE BUY HOUSE LIVE]$_{IP}$

People have to buy house to live.

LIVE	BUY	HOUSE	LIVE
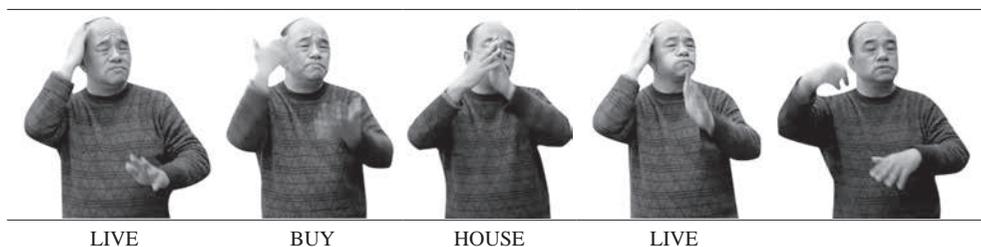
Figure 5-7   Screenshot of prosodic word in sentence (5)

<u>hd</u>   <u>hu</u>
(6) [SON THINGS PREPARE ON BACK]$_{IP}$
    The son prepares things and puts them on his back.

SON	THING	PREPARE	CARRY

Figure 5-8 Screenshot of prosodic words in sentence (6)

5.3.2.3 Eye Blinks

At the end of IPs, there are often eye blinks in the four signers' talking. As shown in Figure 5-9, there are eye blinks at the end of the two IPs in (7).

<u>eb</u>   <u>eb</u>
(7) [SON THINGS ON BACK]$_{IP}$ [CLIMBING]$_{IP}$
    The son prepares things and carries them on his back to climb the mountain.

SON	THING	PREPARE	CARRY

CLIMBING

Figure 5-9 Screenshot of prosodic word in sentence (7)

5.3.2.4 Body Shift

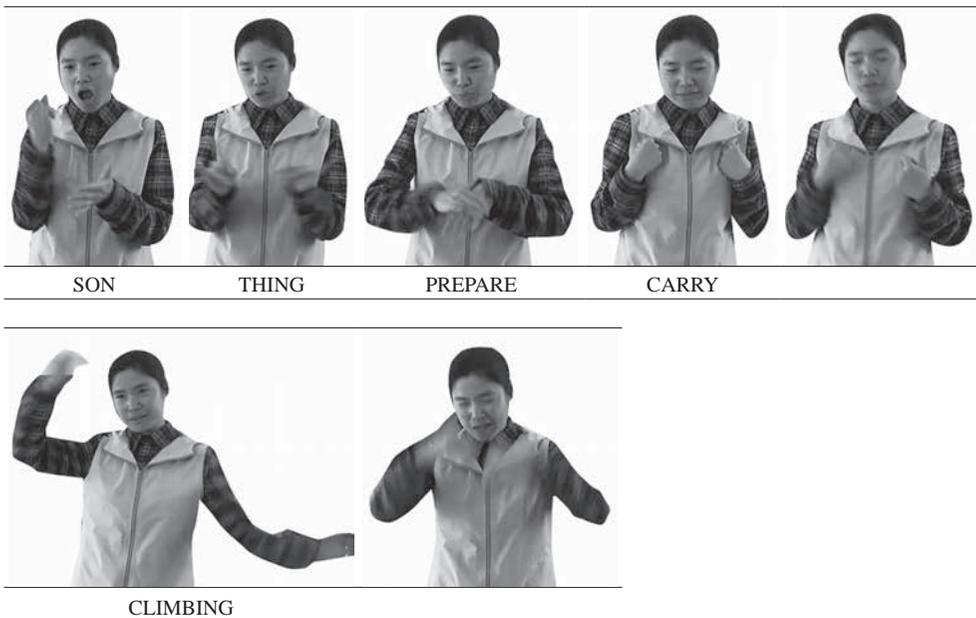Body shift refers to the rotation of a signer's body from one side to the other side and often occurs together with head shift. Body shift happens in CSL occasionally. There are often body shifts at IP boundaries in signer C's talking, while it's not obvious in the other three signers.

$$\text{bs}$$
(8)　[LOCK FINISHED]$_{\text{IP}}$ [DOG WATCH WALK LOOK]$_{\text{IP}}$
　　After the fox locked the door, the dog watched the door.
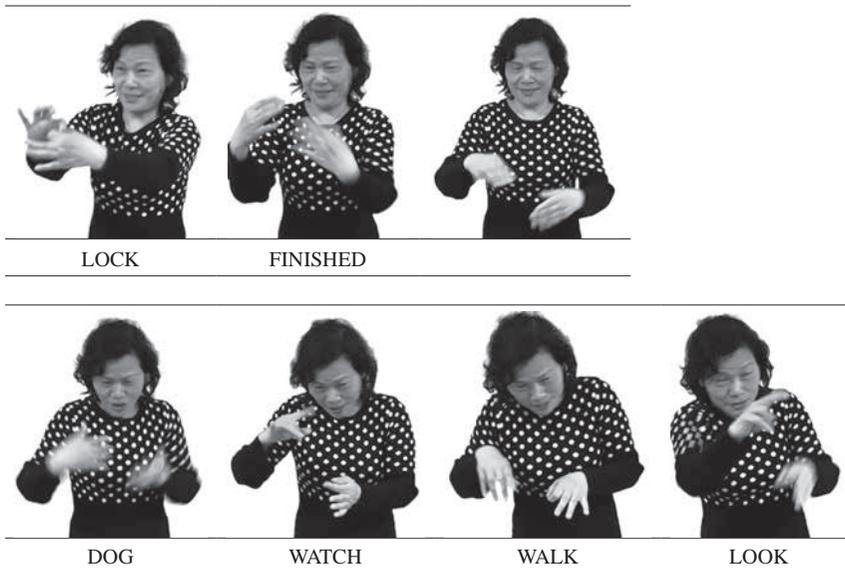
Figure 5-10　Screenshot of prosodic words in sentence (8)

Signer C's body is toward the right direction in the first IP and changes to the left direction in the second IP (Figure 5-10). As the video C corpus nature is "story" which talks about the story between the wolf and the rabbit. Most of the body shift acts to identity conversion or turn-taking. Although it also has the role of helping segmenting IPs, but the main role of body shift marks turn-shift.

5.3.2.5 Pointing

(9)　[P WHAT UNKNOW P]$_{\text{IP}}$
　　You know nothing.

| P | WHAT | UNKNOWN | P |
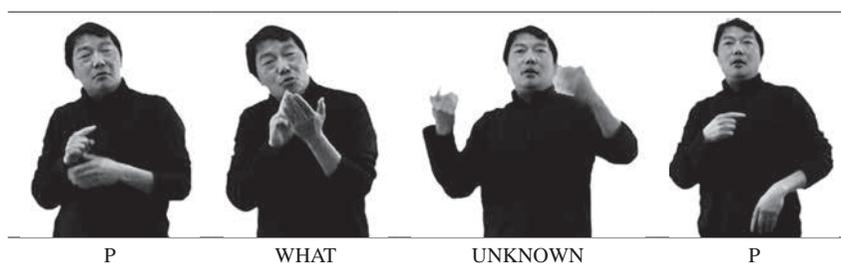
Figure 5-11   Screenshot of prosodic words in sentence (9)

The same as body shift, a pointing marker also occurred in signer A's talking, while it's not obvious in the other three signers' talking. As shown in figure 5-11, there is a pointing marker at the beginning and ending of the IP in (9).

5.3.2.6 Brow Raised or Furrowed

A brow raised or brow furrowed feature often expresses certain emotions, but sometimes occurs at IP boundaries. As in (10), the two IPs have different brow features and change at IP boundaries (Figure 5-12). It's not as reliable as eye blinks or head nod, and sometimes helps to segment IPs.

        <u>br</u>   <u>bf</u>

(10) [ALL MONEY INVEST MONEY]$_{IP}$ [万 DA 万]$_{IP}$

    All money is invested by Wanda.

| ALL | MONEY | INVEST | MONEY |

| 万 | DA | 万 |

Figure 5-12   Screenshot of prosodic words in sentence (10)

5.3.2.7 Dominant Hand Change

There is a unique phenomenon in signer A's talking, in which the dominant hand changed from one hand to another hand. It often happened at an IP boundary. For example, in (11) the dominant hand is the right hand in the first IP, and the dominant hand changed to the left hand in the second IP (Figure 5-13).

<div>
        <u>   rh   </u>   <u> lh </u>
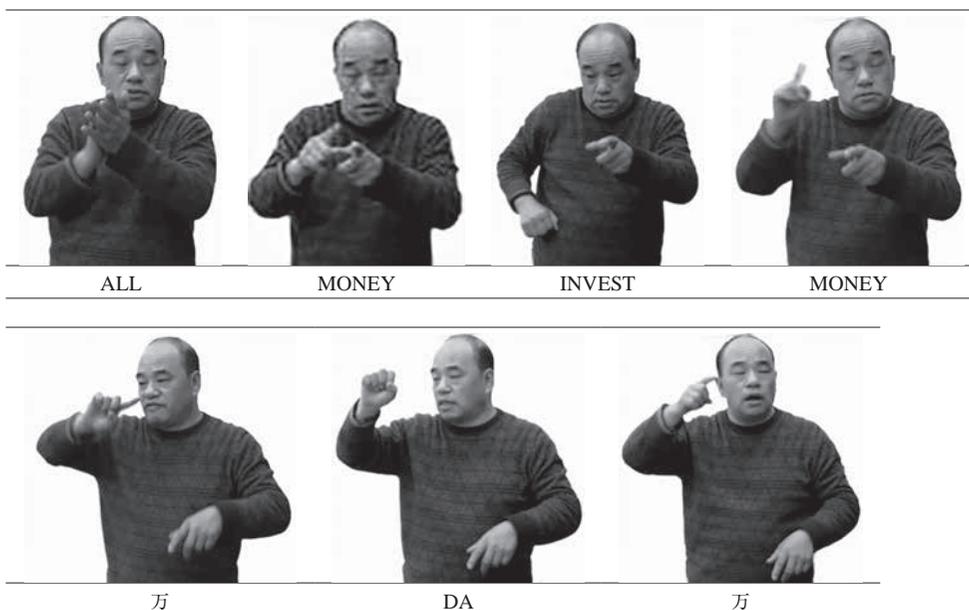</div>

(11) [CAN MUST CAN CATCH 300]$_{IP}$ [CAN]$_{IP}$

      You can get 300 yuan.

| CAN | MUST | CAN | CATCH | 300 | CAN |
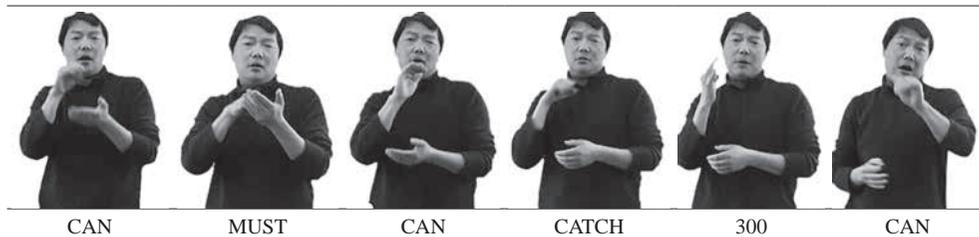
Figure 5-13   Screenshot of prosodic words in sentence (11)

5.4. Frequency of Certain Markers in CSL

In Section 5.3, we analyzed the markers that help to segment IPs in the four signers. This section tries to study the importance of different markers by doing statistics on certain markers, at the beginning, internal, and end of IPs, and the ratio of IPs with certain markers to all IPs in each signer's talking.

5.4.1 Hold

Table 5-3   Hold markers in four signers

Signer	IP Beginning	IP Internal	IP Ending	All Holds	IP Ending Frequency	All IPs	Ending Hold/ All IPs
A	4	4	28	36	78%	175	20%
B	6	1	17	24	71%	80	21%
C	6	2	25	33	76%	100	25%
D	5	2	18	25	72%	90	20%

As can be seen from Table 5-3, there are 36 holds in signer A's talking, four at the beginning of IPs, four are internal and the other 28 are at the end of IPs, so most of the hold cues appeared at the end of IPs in each of the four signers (all above 70%). In the four signers' talking, all above 20% IPs have hold markers, so hold feature is an important marker for IP segmentation.

5.4.2 Repeat

Table 5-4   Repeat in four signers' talking

Signer	IP Ending	IP Internal	IP Ending	All Reps	IP Ending Frequency	Ending/All IPs
A	6	1	23	30	77%	17%
B	2	5	11	18	61%	14%
C	8	3	15	26	58%	15%
D	3	2	20	25	80%	22%

Table 5-4 represents the number of repeat markers appearing at different position of IPs. Most of the repeat markers occurred at the end of IPs (above 58%). Frequency analysis reveals that about 15% of IP endings have repeat markers in the four signers.

5.4.3 Amplitude

Table 5-5   Amplitude marker in four signers' talking

Signer	IP Beginning	IP Middle	IP Ending	All Amps	Ending Frequency	All IPs	Ending Amps./ All IPs
A	4	1	29	33	88%	175	21%
B	8	1	10	19	53%	80	13%
C	3	0	21	24	88%	100	21%
D	3	0	24	27	89%	90	27%

Table 5-5 shows us the distribution of amplitude markers in different positions of IPs. This table shows us that most amplitude markers happen at the end of IPs (all above 53%, signers A, C, and D reached about 88%). In all IPs, above 13% IPs ending prosodic words have amplitude markers.

5.4.4 Head Nod

Table 5-6   Head nod situation in four signers' talking

Signer	IP Ending	All Nods	IP Ending Frequency	Ending/All IPs
A	20	27	74%	19%
B	14	15	93%	18%
C	19	23	83%	19%
D	14	19	74%	16%

Table 5-6 shows us that most head nod markers happen at the end of IPs (all above 74%, signer B even reached 93%). A similar proportion (16%–19%) of IPs have the head nod marker at the last prosodic word.

5.4.5 Head Up

Table 5-7   Head up situation in four signers' talking

SIGNER	TIMES	IP BEGINNING	IP ENDING	IP ENDING FREQUENCY	ENDING/ ALL IPS
A	39	4	30	77%	22%
B	31	5	25	81%	32%
C	32	3	27	84%	27%
D	15	0	15	100%	17%

Head back or head down at the beginning and head up at the end of IPs are important markers in CSL. Whether there are head back or head down markers at the beginning, there is always a head up at the end, so we give statistics of head up at the ending of IP instead of the two beginning markers. Table 5-7 statistics of the head up marker at the ending shoe all were above 77%, signer D even reached 100%. 17% of IPs ending have the boundary markers of head up.

5.4.6 Eye Blinks

Table 5-8   Eye blinks situation in four signers' talking

SIGNER	IP ENDING	ALL BLINKS	FREQUENCY	ENDING/ALL IPS
A	54	81	67%	39%
B	55	90	61%	71%
C	61	71	86%	61%
D	39	48	81%	44%

From Table 5-8, we can see that above 60% eye blinks appear at the end of IPs in the four signers' talking. Some are even above 80% (signer C: 86%; signer D: 81%). The frequency analysis reveals that eye blink markers are more frequent than other boundary markers, as the proportion (above 39%) of IP endings are marked with eye blinks by the four signers (signer A: 39%; signer B: 71%; signer C: 61%; signer D: 44%). So eye blinks are the most important cue to mark IP endings.

5.5. Conclusion

Table 5-9 below shows the proportion of the IPs with a certain marker in all IPs. As can be seen from Table 5-9, although the gender, age, corpus nature, environment is different through different signers, the rate of marked IPs is not much different over all. The statistics tell us that eye blinks are the most important prosodic markers to segment intonational phrases and hand hold, hand amplitude, hand repeat, head up, and head nod are also important prosodic markers for segmentation.

Table 5-9   Marker types and frequency in four signers' talking

MARKERS	FREQUENCY (IP ENDING/IP ALL)			
	SIGNER A	SIGNER B	SIGNER C	SIGNER D
Hold	28/175	17/80	25/100	18/90
Amplitude	29/175	10/80	21/100	24/90
Repeat	23/175	11/80	15/100	20/90
Head up	65/175	25/80	27/100	15/90
Head nod	21/175	14/80	19/100	14/90
Eye blinks	86/175	55/80	61/100	39/90

There is a clear "pause" between adjacent IPs of CSL. As the speed difference of different signers, pause interval is also different, but mostly between 0.1–0.3s (Table 5-2).

According to the analysis of the four selected videos, there are various manual markers at the end of IPs, such as long duration, large amplitude, and repetition of dominant hand and relaxation of non-dominant hand and some non-manual markers, such as eye blinks, head nod and head up at the end, and head down or head back at the beginning of IPs. Besides, there may be body shift, head shift, pointing, brow change and other markers for some signers. Although different signers have their own sign characteristics, such as body shift, head shift, pointing, brow change, etc., but dominant hand hold, amplitude, repetition, head nod, eye blinks and non-dominant hand relaxation appear frequency in CSL, and they are the main markers for IP segmentation from the statistics of the four signers' talking.

References

Baker, C. and C. Padden
    1978   Focusing on the Non-manual Components of American Sign Languages. In P. Siple (ed.) *Understanding Language through Sign Language Research*, pp. 27–57. New York: Academic Press.
Borowsky, T., S. Kawahara, T. Shinaya, and M. Sugahara (eds.)
    2012   *Prosody Matters: Essays in Honor of Elisabeth Selkirk*. Sheffield: Equinox Publishing.
Brentari, D., C. González, A. Seidl, and R. B. Wilbur
    2011   Sensitivity to Visual Prosodic Cues in Signers and Nonsigners. *Language and Speech* 54(1): 49–72.
Crasborn, O. A., Els. Van Der Kooij, and J. Ros
    2012   On the Weight of Phrase-final Prosodic Words in a Sign Language. *Sign Language & Linguistics* 15(1): 11–38.
Dachkovsky, S. and W. Sandler
    2009   Visual Intonation in the Prosody of a Sign Language. *Language and Speech* 52(2/3): 287–314.
Fenlon, J.
    2010   *Seeing Sentence Boundaries: The Production and Perception of Visual Markers Signalling Boundaries in Signed Languages.* London: University College London.

Fenlon, J., T. Denmark, R. Campbell, and B. Woll
    2007   Seeing Sentence Boundaries. *Sign Language & Linguistics* 10(2): 177−200.

Gong, Q.
    2009   A Linguistic Analysis of Sign Language and Chinese Problems in Deaf Education. *Chinese Journal of Special Education* 105: 63−67.

Green, K.
    1984   Sign Boundaries in American Sign Language. *Sign Language Studies* 42: 65−91.

Grosjean, F.
    1977   The Perception of Rate in Spoken and Sign Languages. *Perception & Psychophysics* 22 (4): 408−413.
    1979   A Study of Timing in a Manual and a Spoken Language: American Sign Language and English. *Journal of Psycholinguistic Research* 8(4): 379−405.

Herrmann, A.
    2010   The Interaction of Eye Blinks and Other Prosodic Cues in German Sign Language. *Sign Language & Linguistics* 13(1): 3−39.

Hoza, J.
    2008   Five Nonmanual Modifiers that Mitigate Requests and Rejections in American Sign Language. *Sign Language Studies* 8(3): 264−288.

Infantino, I., R. Rizzo, and S. Gaglio
    2007   A Framework for Sign Language Sentence Recognition by Commonsense Context. *Institute of Electrical and Electronics Engineers* 37(5): 1034−1039.

Khan, S., D. Bailey, and G. S. Gupta
    2012   Detecting Pauses in Continuous Sign Language. *Mechatronics & Machine Vision in Practice* 7196(4): 11−15.

Nakatani, L. H. and K. D. Dukes
    1977   Locus of Segmental Cues for Word Juncture. *The Journal of the Acoustical Society of America* 62(3): 714−719.

Nespor, M. and W. Sandler
    1999   Prosody in Israeli Sign Language. *Language and Speech* 42(2−3): 143−176.

Nespor, M. and I. Vogel
    1986   *Prosodic Phonology*. Dordrech-Hol land/Riverton: Foris Publications.
    2007   *Prosodic Phonology: With a New Foreword*. Berlin: Mouton de Gruyter.

Nicodemus, B.
    2008   The Use of Prosodic Markers to Indicate Utterance Boundaries in American Sign Language Interpretation. *Sign Language & Linguistics* 11(1): 113−122.
    2009   *Prosodic Markers and Utterance Boundaries in American Sign Language Interpretation*. Washington DC: Gallaudet University Press.

Sandler, W.
    2010   Prosody and Syntax in Sign Languages. *Transactions of the Philological Society* 108(3): 298−328.

Sandler, W. and D. Lillo-Martin
    2006   *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.

Sandler, W., I. Meir, S. Dachkovsky, C. Padden, and M. Aronoff
    2011   The Emergence of Complexity in Prosody and Syntax. *Lingua* 121(13): 2014−2033.
Selkirk, E.
    1986   On Derived Domains in Sentence Phonology. *Phonology Yearbook* 3: 371−405.
Stokoe, W.
    1960   Sign Language Structure: An Outline of the Visual Communication Systems of the
            American Deaf. Buffalo: University of Buffalo Press.
Wilbur, R. B.
    1994   Eye Blinks & ASL Phrase Structure. *Sign Language Studies* 84: 221−240.
    2009   Effects of Varying Rate of Signing on ASL Manual Signs and Non-manual Markers.
            Language and Speech 52(2/3): 245−285.