Representation of verbal event structure in sign languages

Evie Malaia
Indiana University
Bloomington, IN, USA
emalaia@indiana.edu

Ronnie B. Wilbur
Purdue University
West Lafayette, IN, USA
wilbur@purdue.edu

Abstract

Sign languages recruit physical properties of visual motion to convey linguistic information. The present experiment investigated the effect of sign position and grammatical aspect on kinematic parameters of predicates in American Sign Language (ASL) and Croatian Sign Language (Hrvatski Znakovni Jezik, HZJ) using motion capture data. Kinematic features of signs recorded on the dominant hand were affected by both grammatical aspect of the predicate, and its position within the sentence. The study demonstrates independent, but interactive effects of grammar and prosody on kinematic parameters of signs, and provides cross-linguistic confirmation that physical properties of articulator motion are recruited in sign languages to express linguistic features.

1 Introduction

Humans perceive and conceptualize reality in terms of discrete events, and use linguistic labels – verbs, or predicates – to denote these events. Event boundaries represented by the predicate have long been of interest to linguistic theory as possible semantic primitives (Dowty, 1979; Jackendoff, 1991; Pustejovsky, 1991; Ramchand, 2008; van Hout, 2001; Van Valin, 2007; Vendler, 1967; Verkuyl, 1972). Predicates denoting events with an inherent boundary representing a change of state (break, appear) are considered semantically telic, as opposed to predicates describing homogenous – atelic – events, such as swim or sew. These predicate properties are also known as event structure template, or aktionsart. Predicate telicity, or linguistic representation of event boundary, has been shown to affect syntactic structure of the sentence in spoken languages (Ramchand, 2008; Tenny, 1994), and thematic role assignment in online sentence processing (Malaia, Wilbur, & Weber-Fox, 2009).

Sign languages recruit physical properties of visual space and motion to convey linguistic information. Prior research has demonstrated that kinematic (motion-related) parameters are utilized for expression of linguistic features in a regular manner, both within linguistic modules and at their interfaces (Brentari, Gonzalez, Seidl, & Wilbur, in press). A growing body of research in psychology also indicates that perceptual segmentation of reality into discrete events is determined by kinematic properties of the scene, namely speed and acceleration in the motion of actors (Zacks, Kumar, Abrams, & Mehta, in press; Zacks, Swallow, Vettel, & McAvoy, 2006).

Interestingly, despite their mutual unintelligibility, sign languages (SLs) appear to be more similar to each other than spoken languages are (Newport & Supalla 2000). Sign components, especially for predicate signs, are grammaticalized from universally available physics of motion and geometry of space, which are therefore fundamentals on which more advanced meanings can be constructed (Wilbur 2003, 2005, 2008.) Cross-linguistic research on SLs can provide an explanation for their apparent visual similarity. At the same time, SL signs are grammaticalized units of meaning, which have to be learned as part of the linguistic system, and that distinguishes them from conventional gestures or pantomime.

Wilbur (2003) made the linguistic observation that ASL lexical verbs could be analyzed as telic (denoting a change of state, such as throw, fall) or atelic (denoting homogenous activities, such as swim, walk) based on their kinematic parameters: telic verbs appeared to have a sharper ending movement to a stop, reflecting the semantic end-state of the affected argument. The observation that semantic verb classes are characterized by certain movement profiles was

formulated as the Event Visibility Hypothesis (EVH).

Crosslinguistic quantitative research into event structure expression in SLs then became necessary to provide an insight into the interface between (possibly) language-independent perceptual cues of event structure used in SLs, and linguistic systems of different unrelated SLs.

HZJ presented an especially interesting case for investigation of event structure expression in sign kinematics, by virtue of being unrelated to ASL (on the basis of which EVH had been formulated), and having a member of Slavic language family as its spoken substrate. Slavic languages are characterized by a conflation of internal (event structure) and external (viewpoint) aspect within their lexicon, leading to fusion of temporal and aspectual domains in verbal predicates, equivalent to overt morphological specification of verbal event structure (Bertinetto, 2001; Borik, 2006; Filip, 1999). Hence, two unrelated sign languages (ASL and HZJ) were chosen in order to investigate kinematic parameters of both lexical (ASL) and grammatical (HZJ) expression of event structure.

2 Grammatical and prosodic markers in sign languages

In comparison on spoken languages, SLs are more likely to use simultaneous means of expressing grammatical markers ("layering"; Wilbur 2000). This strategy compensates for the longer time needed to articulate a sign compared to a spoken word. For example, adverbs can be made using lower face configurations while a verb is being signed on the hands.

Brentari (1998) demonstrated that the Prosodic Hierarchy, which is based on increasing breaks in rhythmic structure, is valid for sign languages: syllables contribute to prosodic words which combine into prosodic phrases which combine into intonational phrases. There is increasingly obvious Phrase Final Lengthening at these domains. Furthermore, the sign and pause durations are affected by signing rate (Wilbur 2009a).

In contrast to such rhythmic marking, components that are held in position from the beginning to the end of a domain generally mark the scope of syntactic and semantic operators. In ASL. Two such markers are lowered brow for wh-questions, and headshake for negation.

Beyond measuring sign and pause durations, previous investigations of sign kinematics have been lacking in quantitative measures, as motion capture equipment has only recently become more available.

3 Data collection and analysis

Various tests have been used in the literature to demonstrate that telicity is a relevant linguistic notion reflected in the grammatical system. The most widely used tests for spoken languages include the temporal adverbial modification test (Dowty, 1979; Verkuyl, 1972), and the conjunction test. Additionally, 'almost' modification has been used as a test in sign language research to identify telic predicates (Smith, 2007).

For the purposes of our study, a group of 50 ASL signs were tested in an interview with a native ASL signer/ linguistic consultant. The native signer's intuitions were elicited in the adverbial modification test, the conjunction test, the 'almost' modification test, and STOP/FINISH combinability test. Telicity of the predicates was established based on results of elicitation. For the adverbial modification test, ASL predicate signs were considered telic if they combined with 'IT TOOK AN HOUR'[1], and atelic if they combined with 'FOR AN HOUR'. Additionally, if the predicate combined with the adverbial meaning 'ALMOST (implemented as an adverbial, or as a modification of the formation of the sign's movement) yielding the meaning of "one did not complete doing X", we interpreted this as presence of end-point (which was not reached) in the event structure of the predicate. If the predicate combined with 'ALMOST' meant *only* "one did not start doing X", the predicate was considered atelic; as expected, some of the telic predicates allowed both interpretations.

For the conjunction test, we tested the predicates' meaning in the sentence 'she did V(erb) on Sunday and on Monday'. If the sentence was interpreted as denoting two discrete events, the predicate was considered telic; if the sentence referred to one long event, the predicate was considered atelic.

Finally, the predicates were examined for combinability with the signs FINISH and STOP. In cases where the predicate combined with FINISH with the 'completive' meaning (Fischer & Gough, 1999), it was interpreted as

[1] ASL expression best transcribed as 'IT TOOK AN HOUR' is equivalent to "in an hour"-type adverbials in spoken English, the temporal modifiers specifying the time elapsed to a referenced time-point.

having an inherent end-point (i.e. telic); if the predicate did not combine with FINISH meaning 'completed', but only with FINISH meaning 'already, in the past', and/or instead could only be combined with STOP, it was considered an atelic predicate. Results of these linguistic tests were then combined in order to classify the predicate as either telic or atelic. When telicity interpretations differed between the four tests for one predicate, signaling possibility of frame structure alternation (Levin, 1993) the predicate was eliminated from the final set of 40 stimuli, which included 24 telic and 16 atelic signs.

The following ASL predicates, which were identified as belonging to telic or atelic classes based on the results of all four linguistic tests, were selected for investigation:
Telic predicates (N=24): STING, THROW, HIT, PLUG-IN, APPEAR, CATCH-UP, OPEN-DOOR, RUIN, EAT-UP, CHECK, TAKE-FROM, ZIP, CLOSE-DOOR, SEIZE, DISAPPEAR, ARREST, BECOME, LOOK-AT, ARRIVE, DIE, RELAX, STEAL, SUGGEST, SHUT-DOWN-COMPUTER
Atelic predicates (N=16): TRAVEL, RIDE-IN, COLLECT, LIVE, PROCEED, SHAVE, FOLLOW, WRITE, STAY, INTERRUPT, DRAW, SEW-WITH-MACHINE, SEND, HAVE, INVESTIGATE, SWIM.

For the study of predicate production in HZJ, 120 imperfective-atelic Croatian verbs and 120 of their perfective counterparts were translated into HZJ in order to identify the mechanisms of temporal-aspectual category expression (Milkovič & Malaia, 2010). 3 major groups of temporal-aspectual verb pairs were identified. The largest group (104 signs) formed temporal-aspectual verb pairs based on the properties of sign kinematics: telic (perfective) signs in this group were formed by using shorter, sharper movement, as compared to atelic-imperfective roots. The second group did not allow formation of telic (perfective) signs from atelic-imperfective roots; the third group allowed formation of telic-perfective signs by suppletive means, including quantification of the internal argument, and use of verbal complements. A subset of 30 temporal-aspectual sign pairs from the first group was selected for further investigation using motion capture recording (see Table 1).

For the motion capture study of ASL predicates, 24 telic and 16 atelic signs were randomized, and elicited from 6 participants in the following linguistic conditions: in isolation, in the carrier phrase 'SIGN X AGAIN', sentence-medially 'SHE X TODAY', and in sentence-final position 'TODAY SHE X'. The conditions were the same for all participants: after completing a practice trial, they saw the stimuli in the same order, and signed to the camera while standing. One production per condition was collected for each signer (thus, we recorded 160 productions per signer for six signers). For motion capture study of HZJ, one participant followed the same protocol on 5 separate days of recording. A simultaneous video recording at 30fps rate was made with a NTSC video camera on a tripod outside the motion capture recording field. The positional data from the marker on the right wrist, tracking the movement of the dominant signing hand, was used for the analysis. Both the video and the 3-D positional data were imported into ELAN annotation software, and aligned using the audio marker and T-pose (the signer standing with hands extended to the sides at shoulder level) at the beginning and end of each recording. The video was annotated in ELAN by a native ASL signer, who marked the beginning and end of each target sign following procedures established by (Green, 1984), assuming the first frame of recognition of the sign-initial handshape as the beginning of each predicate, and either the point of contact, or maximal distance traveled by the hand, as the end of the sign. Thus, the onset and the ending of each sign were defined linguistically based solely on the video cues, without access to kinematic variables. The time points for the beginning and end of each sign were extracted from ELAN annotation of the video data, and processed in MATLAB to extract speed and acceleration profiles for each predicate from the recorded kinematic files.

The kinematic metrics for analysis were selected based on previous investigations in linguistics and psychology. Prior research in event perception has suggested that movement speed and acceleration/deceleration are the markers which enable humans to segment meaningful event from continuous reality (Zacks, Kumar, Abrams, & Mehta, in press). Event Visibility Hypothesis (Wilbur, 2003) proposed that sign languages denote event structure by the slope of deceleration from peak velocity to the end of the sign, which leads to concomitant changes in other kinematic properties of the sign – namely, sign duration, peak velocity, and timing of peak velocity within the predicate.

Table 1. Croatian verbs used as stimuli, and their English translations.

Imperfective form	English Translation	Perfective form	English Translation
buditi	to be waking up	probuditi	to wake up
putovati	to be travelling	otputovati	to take off
putovati	to be travelling	doputovati	to arrive
gledati	to be looking at	ugledati	to spot, to notice
gurati	to be pushing	gurnuti	to give a push
brisati	to be wiping	obrisati	to wipe off
crtati	to be drawing	nacrtati	to draw up
češljati	to be combing	počešljati	to comb through
čistiti	to be cleaning	očistiti	to clean up
čitati	to be reading	pročitati	to read through
dijeliti	to be dividing	podijeliti	to split
brijati	to be shaving	obrijati	to shave
bježati	to be fleeing	pobjeći	to run away
disati	to be breathing	udahnuti	to breathe in
dizati	to be lifting	dignuti	to pick up
dolaziti	to be coming	doći	to show up
donositi	to be carrying	donijeti	to bring
dopuštati	to tolerate	dopustiti	to permit (once)
dovoditi	to be bringing (someone)	dovesti	to bring (to someplace)
dovoziti	to be driving	dovesti	to drive up
govoriti	to be speaking	reći	to tell
gristi	to be biting	ugristi	to bite (someone)
gubiti	to be losing	izgubiti	to have lost
iskorištati	to be exploiting	iskoristiti	to take advantage of
oblačiti-se	to be dressing	obući-se	to put clothes on
odgovarati	to be responding	odgovoriti	to answer
prodavati	to be selling	prodati	to sell
propadati	to be decaying	propasti	to fail
birati	to be choosing	izabrati	to pick
grmjeti	to be thundering	zagrmjeti	to thunder

Based on these proposals, the following metrics were calculated for each verb sign:

a) the duration of the sign in milliseconds (duration);
b) peak instantaneous speed achieved within each sign (maxV);
c) the percent of sign movement elapsed to the moment where peak speed occurred (% elapsed), which is also the point at which deceleration starts,
d) minimum instantaneous negative acceleration (i.e. maximal deceleration) within each sign (minA);
e) the slope of deceleration, calculated as the difference between maxV and the following local minimum, divided by the number of milliseconds over which it occurred. The slope measured the overall steepness of the deceleration from maxV to the following minimum velocity, whereas minA measured the maximum instantaneous negative acceleration (deceleration).

Multivariate analysis of variance (MANOVA GLM) was conducted to determine the effect of each independent factor (Predicate, Position) and their interaction (Predicate x Position) on each of the dependent kinematic variables; the results for ASL and HZJ are presented in Tables 1 and 2, respectively.

4 Results

Kinematic features of verb signs were affected both by the verb type, and by its position within the sentence in a regular manner. Statistical analysis demonstrated regular kinematic distinctions between verb classes.

Table 2. Significant effects of Predicate Type and Position on ASL signs

Kinematic variable	Predicate Type			Position			Predicate Type x Position		
	F (1,916)	*p*<	η_p^2	*F* (1,916)	*p*<	η_p^2	*F* (1,916)	*p*<	η_p^2
duration	11.036	.001	.012	29.573	.001	.031			
maxV	78.301	.001	.079	13.092	.001	.014			
% elapsed	4.393	.036	.005	4.323	.038	.005	4.099	.043	.004
slope	29.645	.001	.031						
minA	52.614	.001	.054						

Table 3. Significant effects of Predicate Type and Position on HZJ signs

Kinematic variable	Predicate Type			Position			Predicate Type x Position		
	F (1,1170)	*p*<	η_p^2	*F* (1,1170)	*p*<	η_p^2	*F* (1,1170)	*p*<	η_p^2
duration	68.375	.001	.055	31.292	.001	.026			
maxV	641.448	.001	.354						
% elapsed	28.925	.001	.024	22.288	.001	.019			
minA	356.863	.001	.234	6.522	.011	.006			
slope	306.2	.001	.207	8.886	.003	.008	4.58	.033	.004

In ASL, measures of deceleration (slope, minA), and in HZJ - peak velocity, were robust to the prosodic effect of Phrase Final Lengthening. The findings showed that Event Visibility in kinematic parameters, demonstrated at the lexical level in ASL verbs, can be grammaticalized in sign languages, such as HZJ. The latter allows formation of temporal-aspectual verb classes from the same sign root, such that rapid deceleration following peak velocity constitutes a morphemic affix similar to those observed for various aspectual purposes, e.g. different types of reduplication (Wilbur, 2005, 2009b).

5 Conclusion

The motion capture data on sign production in two unrelated sign languages demonstrates that the final part of syllables in predicate signs denoting bounded (telic) events is marked by a rapid deceleration at the end of the sign, made even more prominent by higher peak velocity, as compared to verb signs denoting unbounded (atelic) events.

The two experiments show independent and interactive effects of grammar and prosody on kinematic parameters of verb sign, providing cross-linguistic confirmation that physical properties of articulator motion are recruited in sign languages to express linguistic features.

From the standpoint of linguistic theory, the significance of the finding that kinematics of sign production map onto event structure representation has implications for modeling the syntax-semantics interface in both signed and spoken languages. From the standpoint of computational linguistics, the evidence that minimal semantic feature (such as telicity) can affect multiple parameters of the sign's kinematic pattern, which merge the semantic and syntactic levels of a sign with its phonological level, can be utilized for machine translation of signed languages (cf. Malaia, Borneman & Wilbur, 2008).

Reference

Bertinetto, P. M. (2001). On a Frequent Misunderstanding in the Temporal-Aspectual Domain: the 'Perfective = Telic' Confusion. In C. Cecchetto, G. Chierchia & M. T. Guasti (Eds.), Semantic Interfaces [Reference, Anaphora and Aspect] (pp. 9-42). Stanford: CSLI Publications.

Borik, O. (2006). Aspect and Reference Time: Oxford University Press.

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

Brentari, D., Gonzalez, C., Seidl, A., & Wilbur, R. B. (in press). Sensitivity to visual prosodic cues in signers and nonsigners. Language and Speech.

Dowty, D. (1979). Word Meaning and Montague Grammar. Dordrecht: D. Reidel.

Filip, H. (1999). Aspect, Situation Types, and Noun Phrase Semantics. New York/London: Garland Publishing, Inc.

Green, K. (1984). Sign boundaries in American Sign Language. Sign Language Studies, 42, 65-91.

Jackendoff, R. (1991). Parts and boundaries. Cognition, 41(1-3), 9-45.

Malaia, E., Borneman, J., & Wilbur, R. B. (2008). Analysis of ASL Motion Capture Data towards Identification of Verb Type. In J. Bos & R. Delmonte (Eds.), Semantics in Text Processing. STEP 2008 Conference Proceedings (pp. 155-164): College Publications.

Malaia, E., Wilbur, R. B., & Weber-Fox, C. (2009). ERP evidence for telicity effects on syntactic processing in garden-path sentences. Brain and Language, 108(3), 145-158.

Milkovic, M., Malaia, E. (2010) Event visibility in Croatian Sign Language: separating Aspect and Aktionsart. *Theoretical Issues in Sign Language Research-10*, Purdue University, IN, USA

Newport, E. & T. Supalla (2000). Sign language research at the millennium. In Lane, H. and Emmorey, K. (eds.), The signs of language revisited: Festschrift for Ursula Bellugi and Edward Klima, pp. 103-114. Hillsdale, NJ: Lawrence Erlbaum.

Pustejovsky, J. (1991). The Syntax of Event Structure. Cognition, 41(1-3), 47-81.

Ramchand, G. (2008). Verb Meaning and the Lexicon - The First-Phase Syntax: Cambridge University Press.

Tenny, C. (1994). Aspectual roles and the syntax-semantics interface. Dordrecht: Kluwer Academic Press.

van Hout, A. (2001). Event Semantics in the Lexicon-Syntax Interface. In C. Tenny, Pustejovsky, J. (Ed.), Events as Grammatical Objects (pp. 239-282): CSLI Publications.

Van Valin, R. (2007). Some universals of verb semantics In R. Mairal & J. Gil (Eds.), Linguistic Universals): Cambridge University Press.

Vendler, Z. (1967). Linguistics in Philosophy. Ithaca, NY: Cornell University Press.

Verkuyl, H. (1972). On the Compositional Nature of Aspects. Dordrecht: Reidel.

Wilbur, R. B. (2000). Phonological and prosodic layering of non-manuals in American Sign Language. In Lane, H. and Emmorey, K. (eds.), The signs of language revisited: Festschrift for Ursula Bellugi and Edward Klima, pp. 213-241. Hillsdale, NJ: Lawrence Erlbaum.

Wilbur, R. B. (2003). Representation of telicity in ASL. In Chicago Linguistic Society 39 (pp. 354-368). Chicago, IL.

Wilbur, R. B. (2005). A reanalysis of reduplication in American Sign Language. In B. Hurch (Ed.), Studies in Reduplication (pp. 593-620). Berlin/New York: de Gruyter.

Wilbur, R. B. (2009a). Effects of varying rate of signing on ASL manual signs and nonmanual markers. Language and Speech 52(2/3): 245-285.

Wilbur, R. B. (2009b). Productive reduplication in ASL, a fundamentally monosyllabic language. Language Sciences, 31, 325-342.

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (in press). Using movement and intentions to understand human activity. Cognition.

Zacks, J. M., Swallow, K. M., Vettel, J. M., & McAvoy, M. P. (2006). Visual motion and the neural correlates of event perception. Brain Res, 1076(1), 150-162.

Acknowledgements

This work was supported by National Institutes of Health grant DC005241 and NSF BCS-0345314 grant to Ronnie Wilbur.