Analysis of ASL motion capture data towards identification of verb type

Evguenia Malaia
John Borneman
Ronnie B. Wilbur
SLHS Department, Purdue University
emalaya@purdue.edu

The Problem

- Imagine discourse with no “text”
- Relevance of word frequency as opposed to concept frequency?
- Ideal goal – to feed into programs for analysis and translation like any other language
- Barrier: extracting information from signal

This Paper

- Reviews analysis of ASL predicate 3D motion signatures;
- Considers further application of such data for computational processing of ASL video streams, and automatic recognition of predicate type based on 2D motion signatures.
- Particular attention paid to contribution of *slope of deceleration* at the end of signs, and to the values of the *maximum velocity and minimum acceleration* achieved during the sign motion.
- The focus of the study on predicates is determined by the fact that each sentence or clause in natural languages is built around a predicate.

Recent Work

- Preliminary data indicate that there is a significant difference in the motion signatures of lexical predicate signs that denote telic and atelic events.

- These results are empirical evidence for mapping between sign language (ASL) phonology/kinematics, and semantic decomposition of predicates.

Event Visibility Hypothesis (EVH)

- Event structure is visible in form of ASL predicate signs.

Model of Event Structure 1

- Events vs. States
  - An event is something that happens.
  - A state is condition or property that lasts over time (no change).
- Vendler (1967) verb classes
  - States
  - Activities
  - Achievements
  - Accomplishments

Vendler (1967)

- Vendler (1967) proposed a system of four basic syntactically relevant semantic types of predicates:
  - atelic States and Activities, and
  - telic Achievements and Accomplishments.
- Atelic events are homogenous - they may be divided into identical intervals
  - each of which is an instance of the event itself, i.e. ‘walking’ as an instance of ‘walking’.
- Telic events are composed of at least two subevents - one of which is the final state - hence heterogeneous
  - cannot be divided into identical intervals.

Model of Event Structure 2

- States, Processes, and Transitions

Sub-event type **< Pustejovsky 2001**	**Subevent**
States	S
Processes	P
Transitions: Achievements	S → S
Transitions: Accomplishments	P → S

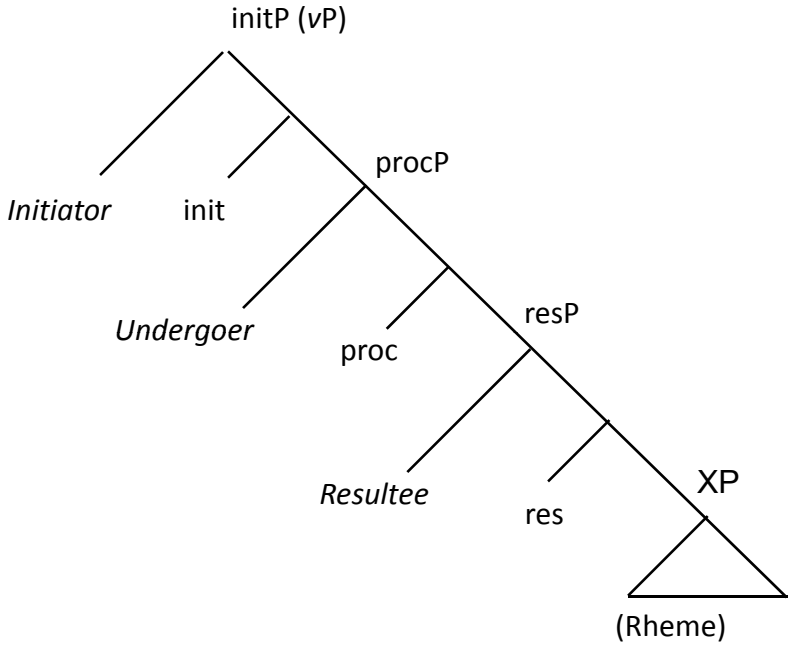
Atelic

Telic transition to end state

Model of Event Structure 3

- Ramchand (2008) - Divide events into
  - Initiation phrase InitP
  - Process phrase ProcP
  - Result phrase ResP
- Participant involved:
  - Initiation phrase: Initiator
  - Process phrase: Undergoer
  - Result phrase: Resultee

Ramchand Event Structure Model 3 Tree



Cross-linguistic investigations (e.g. Van Valin 2007) have further demonstrated that verbal morphology of individual languages can be demonstrated to represent individuated elements of event structure, allowing use of a single verbal root in multiple event structures, yielding distinctly telic or atelic meanings.

Class	Syntactic status	Semantic components	Examples
	[init, proc]		
I	Transitive	Initiator, Undergoer	Drive, push, paint
	Transitive	Initiator, Path	Eat, read, paint
II	Intransitive	$Initiator_i$, $Undergoer_i$	run
	[init, proc, res]		
III	Transitive	Initiator, $Undergoer_i$, $Resultee_i$	Throw, defuse
	Transitive	$Initiator_i$, $Undergoer_i$, Result-Rheme	enter
IV	Intransitive	$Initiator_i$, $Undergoer_i$, $Resultee_i$	Arrive, jump
V	Ditransitive	Initiator, Undergoer, Resultee	Give, throw
	[proc]		
VI	Intransitive	Undergoer	Melt, roll, freeze
	[proc, res]		
VII	Intransitive	$Undergoer_i$, $Resultee_i$	Break, tear
	[init, proc, N]		
VIII	N-conflation	$Initiator_i$, $Undergoer_i$	Dance, sleep
	[init, proc, A]		
IX	A-conflation	Undergoer	Dry, clear

Model of Event Structure 4

- Formal semantic model by Van Lambalgen and Hamm (2004)
- Provides model for representing meaning of concepts related to event structure, rather than just predicate/verb categories
- Not yet attempted to implement it

Testing with ‘almost’

- In English, ambiguous with telic predicates:
  - I almost sat down.
    - EITHER I never started
    - OR I never finished.
- But in ASL not ambiguous:
  - Unrealized Inceptive is different form than Incompletive
    - Almost sit down (but don’t because everyone else stays standing)
    - Almost sit down (but realize the seat is wet and stop myself before making contact with the seat).
- Not a modality issue, but a morphology issue.

Event Structure and SLs

- We already know that telicity plays a role in sign morphology:
  - [delayed completive] aspect only applies to telic stems (Brentari 1998).
  - Durative and continuative aspects cannot apply to telic predicates (Wilbur 2005, in press).
  - Some mouth NMs are distributed according to predicate telicity type (Schalber 2004, 2006; Schalber & Grose 2006) in Austrian SL and ASL.

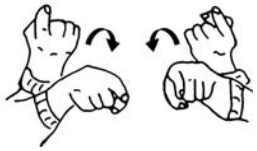
Event Visibility Hypothesis

- The Event Visibility Hypothesis (EVH) argues that the semantics of event structure (subevents) are visible in predicate sign formation.
- Can see if an end state is intended by way sign movement comes to a stop – rapidly for end states, regularly for no end state.

Examples of EVH from ASL

<u>Examples of movements in signs denoting Telic events:</u>

a. change of aperture handshape change
SEND



b. orientation change
HAPPEN

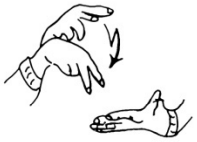
c. setting change proximal/distal
POSTPONE

d. change of location with contact
HIT

<u>Examples of movements in signs denoting Atelic events:</u>

(a) RUN [tracing: straight]

(b) PLAY (tracing + TM)



(c) READ (tracing + TM)

Motion Analysis

- ASL signs representing telic events appear to contain ‘perceptually significant rapid deceleration to a stop’.

- Hypothesis:

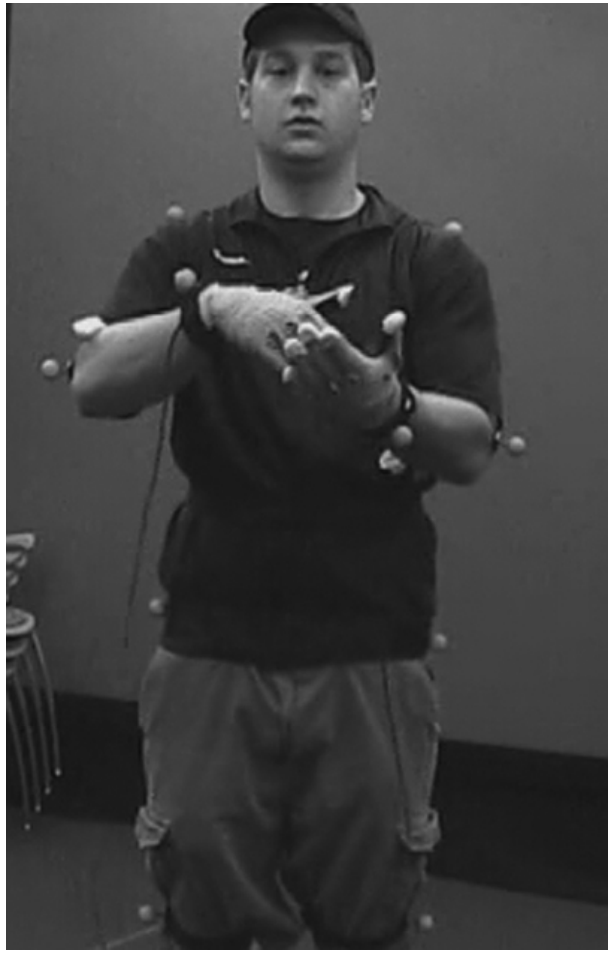
  Signs representing telic events will have steeper deceleration slopes than those representing atelic events, because it will provide an end-marking to indicate the final State.

Methodology

- Stimuli:
  - 29 telic and 21 atelic signs chosen, randomized and presented to native bilingual ASL signer.
- Conditions
  - Signs in isolation (done twice)
  - Signs in carrier phrase SIGN X AGAIN
  - Signs in medial sentence position IX_3 X TODAY
  - Signs in final position TODAY IX_3 X (NB: data analysis not yet finished).
  - Each sign produced 4 times by signer

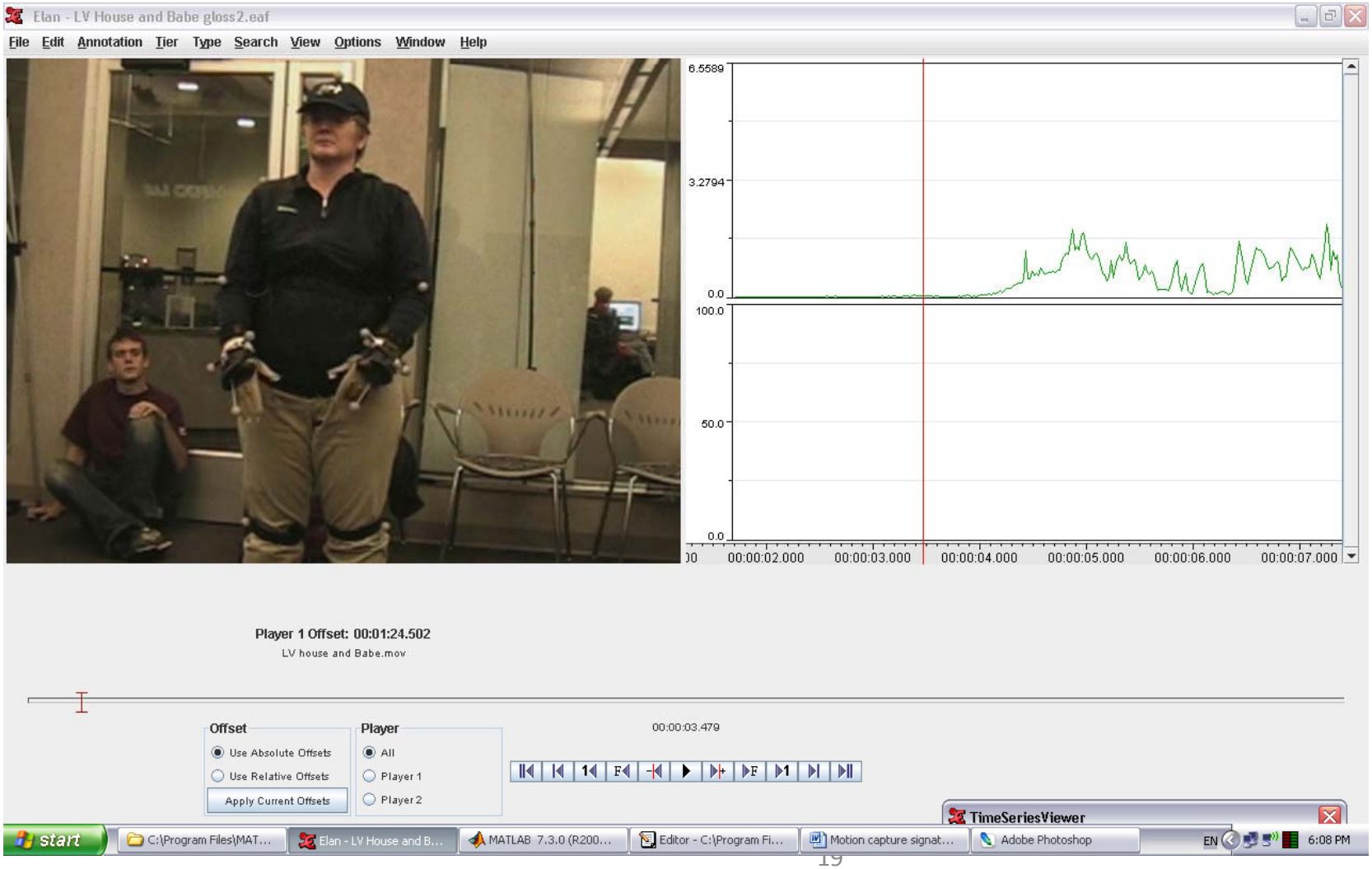
    x 29 for telic = 116 cases for telics

    x 21 for atelic = 84 cases for atelics.
- Equipment
  - Gypsy 3.0 wired motion capture suit
  - Pair of 18-sensor Metamotion Cybergloves.
  - Six special motion capture ceiling mounted cameras.

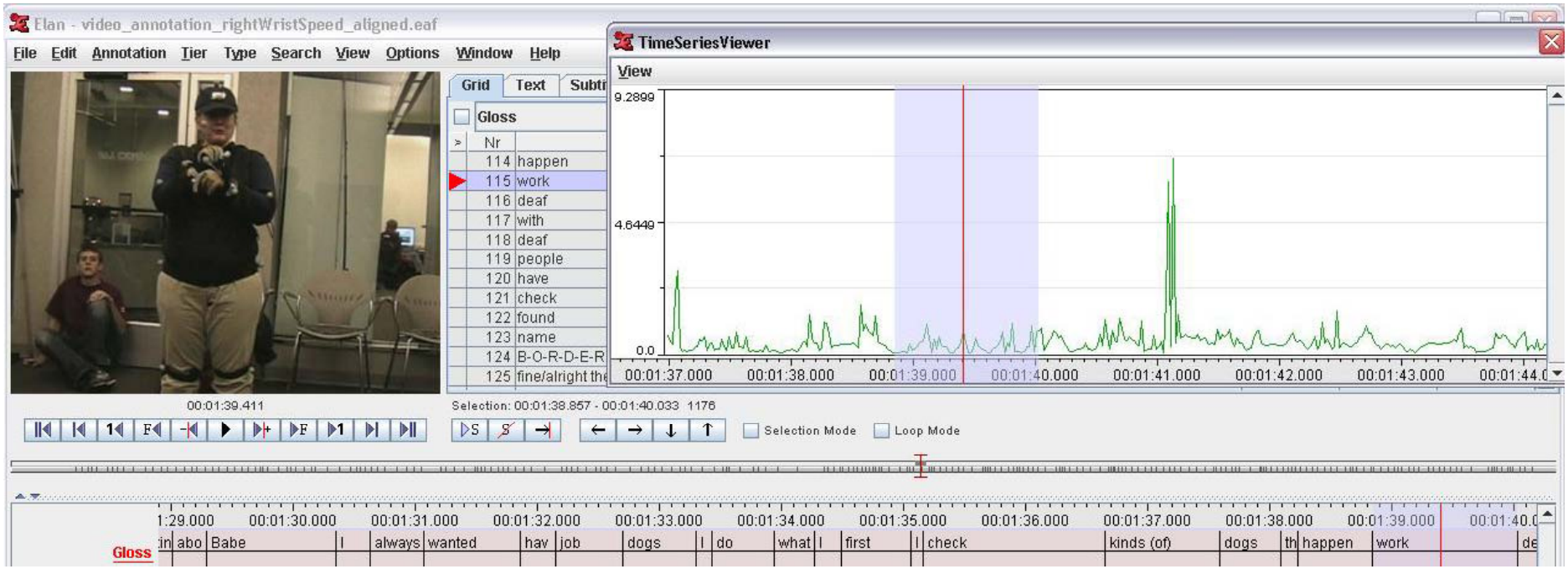


Data Processing

Video data and motion capture are aligned using audio and visual cues:



The Magic of ELAN < MPI



Kinematic study: data analysis

The following metrics were calculated:

- the maximal velocity (maxV),
- the local minimum velocity following it (minV),
- the slope of the drop from maxV to minV (slope),
- the minimum acceleration (minA) following the maximum velocity,
- duration of the predicate (in frames),
- the frame location of max V, minV, and minA,
- percentage of the predicate elapsed to maxV occurrence.

These metrics were chosen to allow for maximal homogeneity in predicate comparison.

Deceleration Results

Metric: the Slope from the velocity peak to the next velocity minimum, reflecting the deceleration to(ward) a stop.

Slope	Atelic	Telic	Ratio Telic/Atelic
Isolation 1 *	-.09	-.14	1.46
Isolation 2 *	-.12	-.18	1.46
Carrier Phrase**	-.12	-.23	1.97
Sentence 1 *	-.14	-.23	1.62

e.g. Mean telic slope is 1.46 times steeper than mean atelic slope in isolation

Difference between atelic and telic slopes is significant at *p<.05 **p<.001

From the perspective of syntax-semantics interface modeling theory (Ramchand, 2008, Wilbur, 2003), higher decelerations in motion signatures of telic ASL predicates also mark additional semantic arguments of the event, what Ramchand refers to as the 'Resultee'.

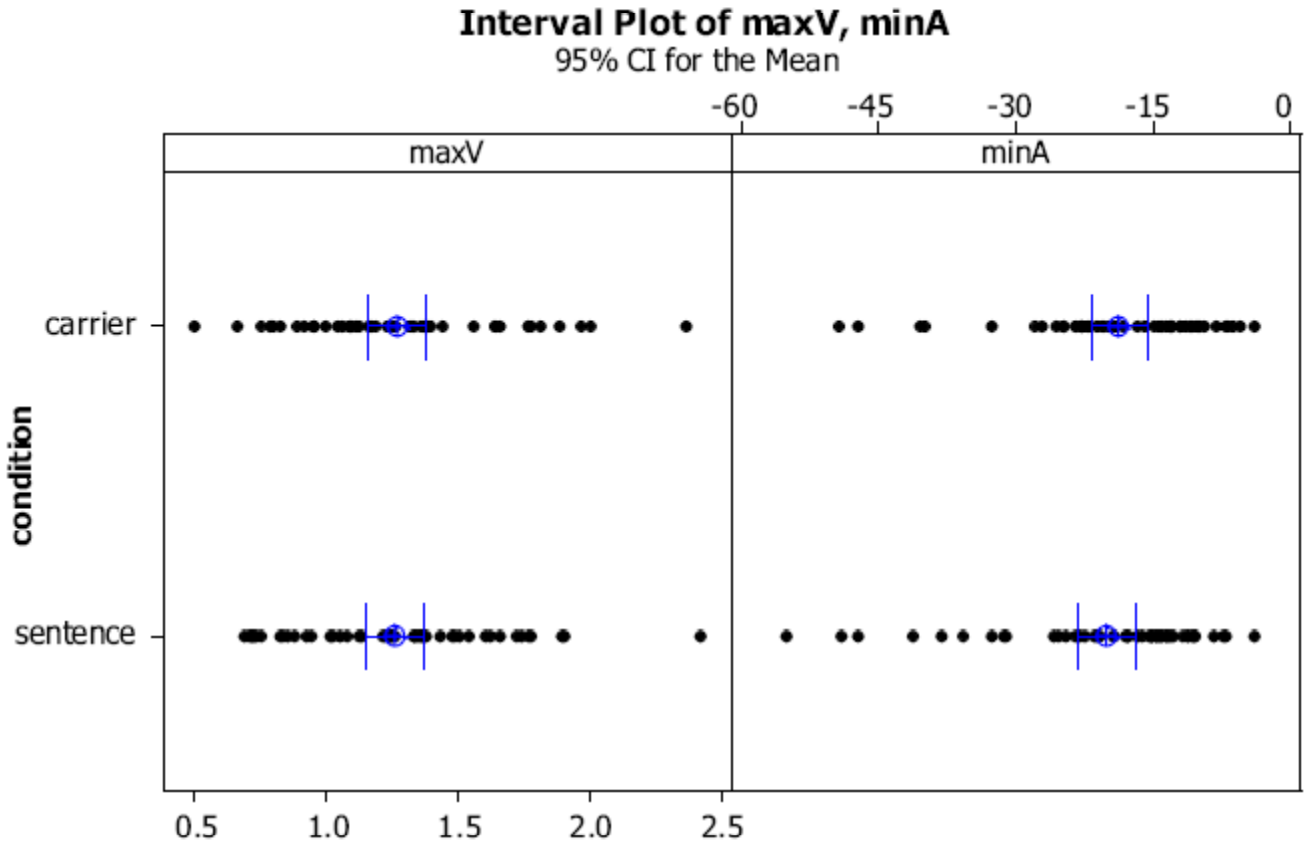
Implications

- For sign language research
  - Event structure theory is a good way to attack internal structure of predicate signs because it predicts a distinction that can be reliably measured
  - Sign movements have internal structure based on velocity changes
  - Signers are making production distinctions that people have not noticed before
  - These distinctions are categorical – either telic or atelic.

Further Processing

- Pooled carrier sentence and sentence data – overlapping distribution
- Eliminated isolation data from further processing at this point – two trials too distinct
- Performed regression to determine needed variables to make telic / atelic decision.

Pooled Carrier Phrase and Sentence Data based on Distribution



Further Analysis

- Binary logistic regression on pooled data
- Search for minimum number of variables that could be used to predict whether the signed predicate is telic or atelic.
- Logit function selected to calculate predicted probabilities.
- Applied to all measures mentioned
- Those above p=.05 discarded one by one

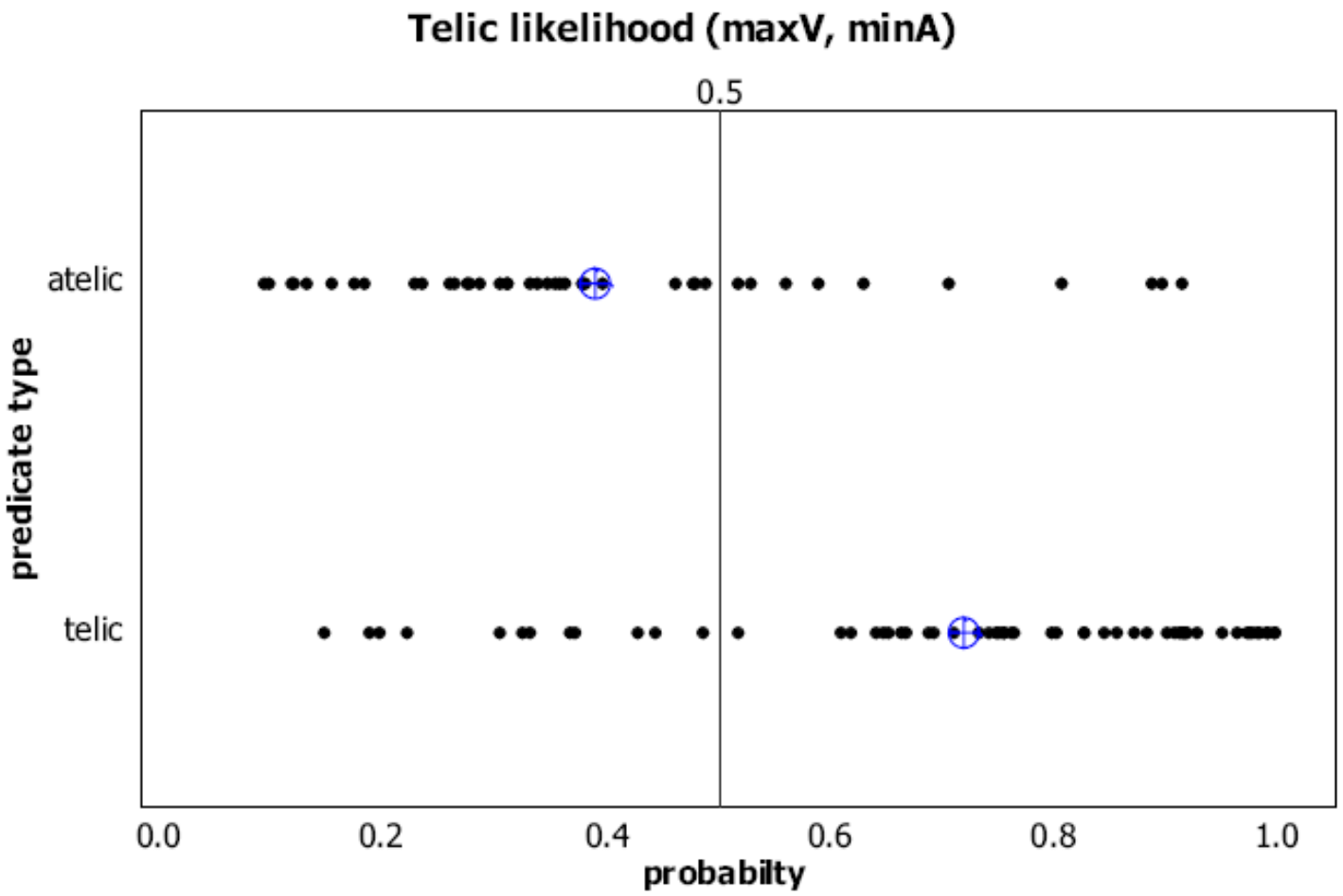
Model 1: maxV and minA

$$P_T = \frac{e^{\beta}}{1+e^{\beta}} \quad (1)$$

$$\beta = -4.46 + 2.63\ (\text{maxV}) - 0.097\ (\text{minA}) \quad (2)$$

P_T is probability of predicate being telic. With threshold $P_T > .05$ predicts telic and $P_T < .05$ predicts atelic, 46 out of 56 telic predicates (82.1%), and 32 out of 44 atelic predicates (72.7%) can be identified correctly using only maxV and minA measures.

Prediction of predicate type (1)

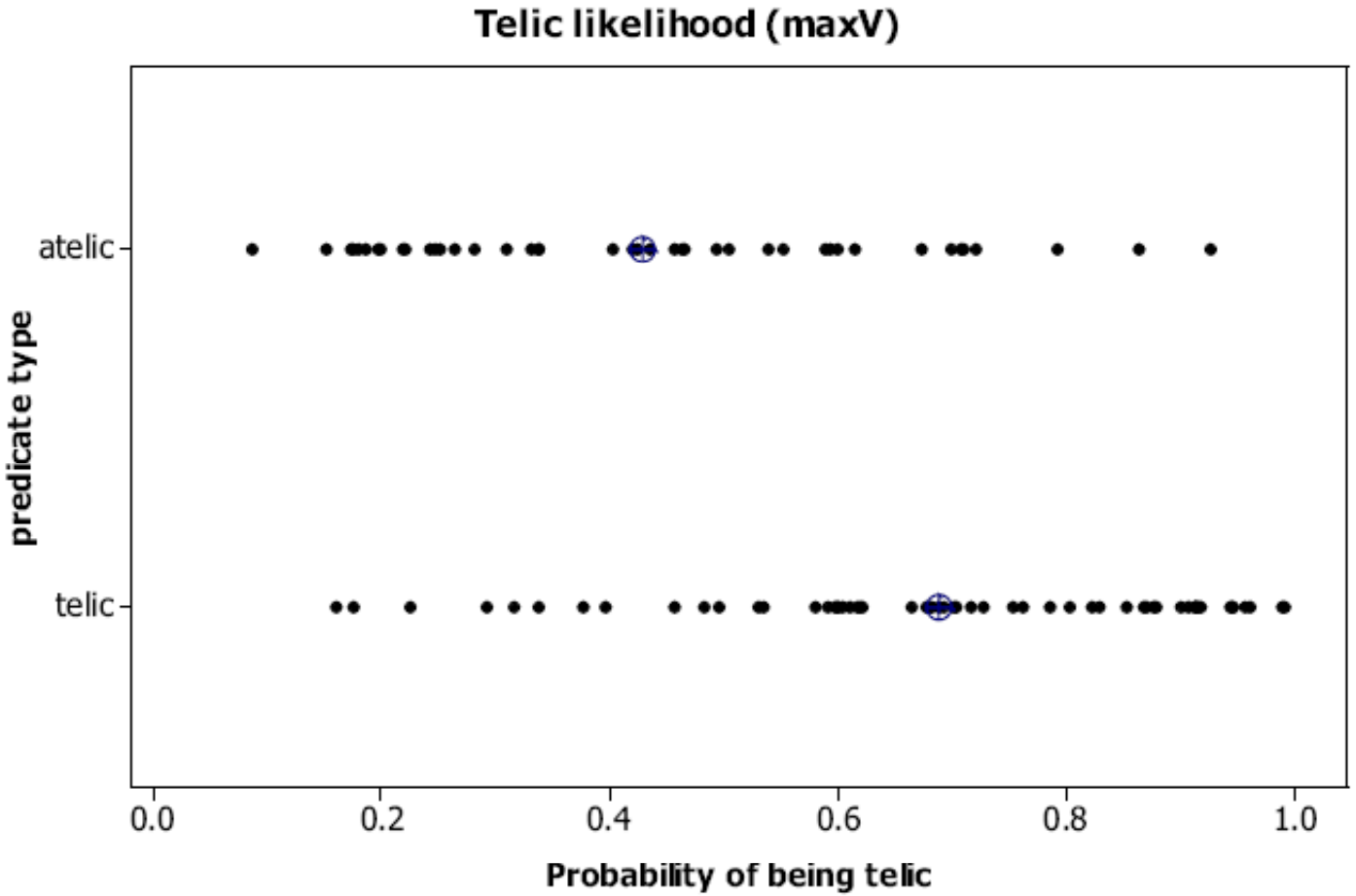


Model 2: maxV only

$$\beta = -4.19 + 3.71\ (\text{maxV}) \qquad (3)$$

Using this simplified equation on the pooled data ensures that 47 out of 56 telic predicates (83.9%) and 27 out of 44 atelic predicates (61.4%) can be identified correctly with a 50% probability threshold based only on maxV.

Prediction of predicate type (2)



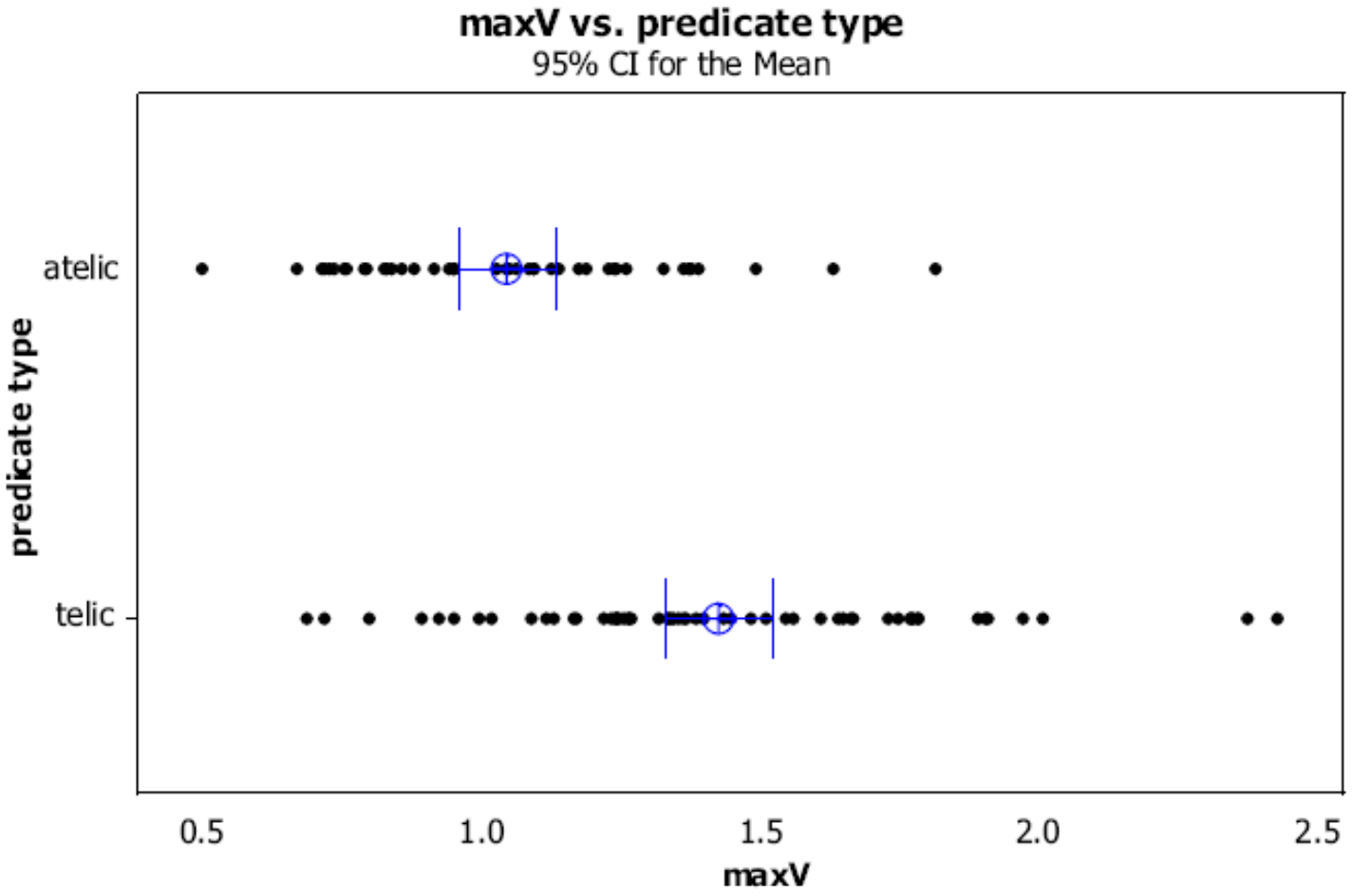
Model 2

- Data points represent telic and atelic predicates.
- Crosshairs represent the mean probability of correct predicate type identification for atelic (0.429) and telic (0.689) predicates.
- Direct analysis of the maximum velocity (maxV) data supports both the original model (eqn. 2) and the simplified model (eqn. 3).

Why?

- Telic predicates have a significantly higher maxV mean and distribution than atelic predicates.
- Supports hypothesized production difference between telic and atelic signs in ASL

Telic vs. atelic maxV



Remaining Problems

- Need higher data acquisition rates,
- Determine inter-signer variability in production differences between telic and atelic predicate signs,
- Determine reliability of maximal velocity and minimal acceleration metrics,
- Development of additional metrics, and
- Figure out how to find a predicate sign in a string.

Acknowledgments

This research is supported by NSF-RDE #0622900;

NIDCD #DC005241; NSF-Ling #0345314.

We thank:

- Nicoletta Adamo-Villani and personnel at the Purdue Envision Center for Data Perceptualization for assistance with motion capture;
- Robin Shay, Gabriel Masters and the Purdue and Indianapolis sign language communities for their ongoing support of Purdue Sign Language Linguistics research.

THANK YOU!!