

## **Synchronizing Modalities: A model for synchronization of gesture and speech as evidenced by American Sign Language**

A prominent account in American Sign language (Neidle, 1998 and Neidle, 2002) posits that wh-phrases landing in the right periphery is evidence that universal grammar must allow the option of rightward movement. This paper will show that extending a standard leftward account (Koopman, 1996 and Kayne, 1994) is sufficient for yielding an operative analysis of ASL. Consequently, and significantly, it also yields a model that properly and systematically parses multiple modalities as simultaneous, synchronized structures with coordinated domains, so that the patterns of one predict the patterns of the other. This new model applies to the multiple modalities of spoken languages as well. In the way that it parses manual and nonmanual expressions in signed languages as coordinated forms, it likewise produces parallel parsing of spoken and gestural speech in spoken languages.

The stipulations of the rightward analysis are biased by primary attentiveness to manual sign movement and subsequent interpretation of the nonmanual markings so that the latter conform to the structures of the former, rather than in coordination with them. For example, the construction proposed by Neidle (2002) moves the manual wh-phrase above TP first into the Spec of a higher projection, FP, then rightward into the Spec of CP. Then, as expected by principle [Ia], wh-marking co-occurs with the entire CP domain in sentence (3) because the wh-phrase in CP c-commands FP and TP. For this account, c-command suffices as grounds for spread in cases of wh-in situ, where spread is obligatory over all of CP, but it falls short in cases of [Ib] wh-movement as in sentence (4) where wh-marking may optionally co-occur only with the local wh-phrase. Furthermore, the rightward construction does not account for nonmanual boundaries on a phrasal level of interaction with manual phrases. It treats peaks of nonmanual intensity either as direct occurrences with a lexical item, such as WHO in (4)-(5), or as relations between two specified items, such as WHO and a +wh feature in (3). These conditions are unmotivated in light of a more comprehensive analysis.

The model extends Merge Grammar, a Chomskian-inspired minimal grammar framework, as described in Keenan-Stabler 1996 and Stabler 1997. It generates leftward building operations of movement, treats nonmanual nodes as separate lexical items for selection and transformation, and builds marking spreads on a phrasal level of operation. All its motivations are in principle with the actual behaviors of movement in ASL and its co-articulation of nonmanual gestures with manual signs.

The right edge of ASL questions is the edge most consistent with the location of the CP containing the wh-phrase—both before and after movement. So it is more intuitive that a large chunk of structure move *left* of this edge in order to land WHO at the right. The Koopman-Kayne analysis emerges. One could posit that the sentence (7) suffices for the criteria in [I] for wh-in situ. It would simply require a line extending its entire CP to denote spread, as governed by [Ia]. It is not enough, however, in an analysis that seeks to account for all the phenomena of synchrony across modalities. This parser includes nonmanual gestures as part of the lexicon; they are articulated simultaneously with the manual sign nodes in their c-command. Sentence (8) introduces the nonmanual node (WH), as one such lexical item, to denote wh-marking. This node occupies Spec-CXP, so that every manual sign node in its c-command (WHO in this case) co-occurs with nonmanual question marking.

In (9)-(10), the VP chunk, LOVE JOHN, moves out of the TP and leftward above C2P into C3P. The (WH) node begins in Spec of C2P. Sentence (9) demonstrates the first option in [Ib], where the spread is necessarily local and only C2P selects the nonmanual node (WH). In (10), the (WH) node moves from Spec-C2P to Spec-C4P, and so the spread grows to extend all across the wh-clause, C4P. Peak spreading correlates to the CXP that most immediately contains both the wh-lexical item as well as its +wh feature (the trace of WHO here). Therefore in (10) the peak is C2P, in accordance with principle [IIb], as shown in its X-Bar representation. Note that this peak naturally co-occurs with the wh-lexical item, WHO.

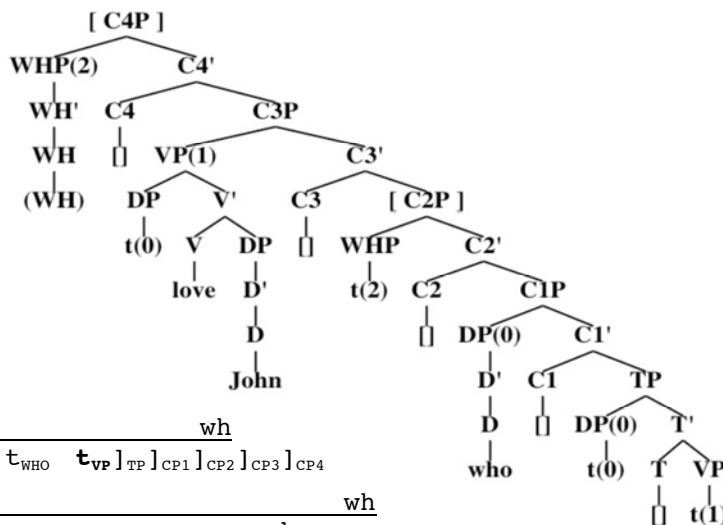
To achieve cases of wh-in situ, the chunk C2P now moves up and leftward to the higher projection, C5P, in sentence (11). The (WH) in Spec-C6P c-commands nonmanual spread over the entire wh-clause, by principle [Ia]. Again, the peak is the smallest phrase containing the perseveration of the lexical item, WHO, to its original +wh trace. So, as shown in (12), the peak is C5P. As the chunks progressively move leftward into higher projections, nonmanual markings grow leftward with them. This preserves the phrasal level transformation of nonmanual markings.

Ultimately, systematic building operations of a minimalist parser achieve parseable representations of synchronized, coordinated structures across multiple modalities. The leftward account remains intact and robust. This model holds potential application to many cross-linguistic platforms beyond ASL, such as the spoken and gestural modalities of natural spoken languages.

Nonmanual question markings are expressed in ASL by furrowed eyebrows and a slight, tilted headshake. Wh-movement is governed by these criteria: **[I]** Questions contain the wh-phrase either in situ or moved to the right periphery. **[Ia]** If wh is in situ, marking is articulated obligatorily with the entire wh-clause. So (3) is grammatical, but (2) is not. **[Ib]** If movement occurs, marking necessarily spreads local to the phrase and then optionally over the entire clause. **[II]** Peaks of nonmanual intensity occur at linguistically postulated positions for +wh features for question marking. **[IIa]** When wh-phrases remain in situ, the duration of peak obligatorily begins at the wh-lexical item and lasts the entire clause, as shown in (3). **[IIb]** With movement, peak occurs at the clause-final position, over the clause-final wh-phrase as in (4)-(5).

- (1) [JOHN SEE TEACHER YESTERDAY]  $\frac{wk}{pk}$
- (2) \* [JOHN SEE  $\frac{wh}{pk}$  WHO YESTERDAY] [+wh]
- (3) [JOHN SEE WHO YESTERDAY] [+wh]  $\frac{wh}{pk}$   
'John saw who yesterday.'
- (4) [JOHN SEE YESTERDAY] WHO  $\frac{wh}{pk}$
- (5) [JOHN SEE YESTERDAY] WHO  $\frac{wk}{pk}$   
'Who did John see yesterday.'
- (6) [[MARY [LOVE JOHN]<sub>VP</sub>] TP ] CP
- (7) [WHO [t<sub>WHO</sub> [LOVE JOHN]<sub>VP</sub>] TP ] CP
- (8) [(WH) [WHO [t<sub>WHO</sub> [LOVE JOHN]<sub>VP</sub>] TP ] CP<sub>1</sub>] CP<sub>2</sub>  $\frac{wh}{pk}$

The X-Bar representation of (10)



- (9) [[ [LOVE JOHN]<sub>VP</sub> [(WH) [WHO [t<sub>WHO</sub> t<sub>VP</sub>] TP ] CP<sub>1</sub>] CP<sub>2</sub>] CP<sub>3</sub>] CP<sub>4</sub>  $\frac{wh}{pk}$
- (10) [(WH) [[LOVE JOHN]<sub>VP</sub> [t<sub>(WH)</sub> [WHO [t<sub>WHO</sub> t<sub>VP</sub>] TP ] CP<sub>1</sub>] CP<sub>2</sub>] CP<sub>3</sub>] CP<sub>4</sub>  $\frac{wh}{pk}$
- (11) [(WH) [[t<sub>(WH)</sub> [WHO [t<sub>WHO</sub> t<sub>VP</sub>] TP ] CP<sub>1</sub>] CP<sub>2</sub> [t<sub>WHO</sub> [[LOVE JOHN]<sub>VP</sub>] CP<sub>3</sub>] CP<sub>4</sub>] CP<sub>5</sub>] CP<sub>6</sub>  $\frac{wh}{pk}$
- (12) [(WH) [[t<sub>(WH)</sub> [WHO [t<sub>WHO</sub> t<sub>VP</sub>] TP ] CP<sub>1</sub>] CP<sub>2</sub> [t<sub>WHO</sub> [[LOVE JOHN]<sub>VP</sub>] CP<sub>3</sub>] CP<sub>4</sub>] CP<sub>5</sub>] CP<sub>6</sub>  $\frac{wh}{pk}$

## References:

- i. Cassel, J. et al. 1994. Animated Conversation: Rule-based Generation of Facial Expression, Gesture & Spoken Intonation for Multiple Conversation Agents.
- ii. Cassel, J., Sullivan, J., Prevost, S., and Churchill, E. eds. 2000. *Embodied Conversational Agents*.
- iii. Hayes, B. 1995. *Metrical Stress Theory*
- iv. Kayne, R. 1994. *The Antisymmetry of Syntax*.
- v. Keenan, E., Stabler, E. 1994. Abstract syntax.
- vi. Kendon, Adam. 1995. Gestures as illocutionary and discourse structure markers in So. Italian conversation.
- vii. Koopman, H. 1996. The Spec-Head Configuration.
- viii. McClave, E. 1999. Linguistic functions of head movements in the context of speech.
- ix. Neidle, C., Language Across Modalities; ASL focus and question constructions. 2002.
- x. Neidle, C., et al. 1998. The Rightward Analysis of Wh-movement in ASL: A Reply to Petronio and Lillo-Martin 1997.
- xi. Neidle, C., et al. 2000. *The Syntax of ASL*
- xii. Neidle, C., Sclaroff S., Athitsos, V. 2001. SignStream: A Tool for Linguistic and Computer Vision Research on Visual-Gestural Language Data.
- xiii. Pierrehumbert, Janet. 1981. Synthesizing intonation.
- xiv. Stabler, E. 1997. Derivational minimalism.
- xv. Stabler, E. 1998. Acquiring languages with movement.